

**Affilin[®] binding proteins selected against a class B
GPCR ectodomain**

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I – Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

von Bo Song

geboren am 04. November 1978 in Hubei, China

Gutachter: PD Dr. Hauke Lilie

Prof. Dr. Harald Kolmar

Dr. Marcus Fändrich

Verteidigungsdatum: 20. 03. 2012

Acknowledgements

I would like to thank all the following people for their contributions to this thesis.

First of all I owe my gratitude to my former supervisor Prof. Dr. Rainer Rudolph, who gave me the excellent opportunity to perform my thesis. His constant support and scientific guidance were essential for my research. Although he has passed away, I will forever benefit from his broad-mindedness, patience and encouragement.

I am indebted to my PD Dr. Hauke Lilie for supervising and correcting my thesis, as well as much encouragement to advance. My sincere thanks are due to our group leader Dr. Sven Pfeifer for the intellectual stimulation, encouragement, support and personal guidance throughout my study. I greatly appreciate your selfless help for me and my family, from the beginning you picked us up when we had arrived in Halle for the first time on a late evening till now.

I am extremely grateful to all my colleagues of group "Artificial Binding Proteins" for nice working atmosphere and friendly help in a variety of ways in my study. I would like to specially thank Andreas Hoffmann for many wonderful discussions and assistances, Dr. Sladjana Tomić for her help in SPR experiments, Hannes Grabner for his efforts in Hit-ELISA experiment, Holger Herrmann for handling Hit-ELISA data, Dominik Schneider for DSF experiment and Qi Zhang for ITC experiment. I have to acknowledge Andreas Hoffmann for establishing the method of nPAC1-Rs production. I would thank all members of the Institute of Biotechnology for friendly assistance.

I would like to acknowledge Scil Proteins and all members of the Affilin[®] group so much for their great cooperation. I wish to give my special thanks to Dr. Ulrike Fiedler for her support and Dr. Arnd Steuernagel for his foresighted guidance. I would like to gratefully thank Dr. Anja Kunert for professional guidance, Dr. Madlen Zwarg for her essential assistance and correcting my thesis, Dr. Jörg Nerkamp for his always warm-hearted assistance. I greatly appreciate Dr. Markus Fiedler, Dr. Erik Fiedler and Dr. Thomas Göttler for their valuable discussions and suggestions, as well as Dr. Ilka Hänssgen for her assistance in DLS measurements.

The financial support from the BMBF InnoProfile program within the Unternehmen Region Innovation Initiative for the Neue Länder (Grant Number 03IP606) and the Excellence Initiative of the State of Sachsen-Anhalt of the Federal Republic of Germany is gratefully acknowledged.

Finally, I would like to express my deepest gratitude to my parents and sisters, my wife and son for their endless love, concern and support, which always encourage me to keep moving on.

Bo Song

Abstract

In the past twenty years, engineered antibodies and their fragments have been invaluable tools for the biotechnological and pharmaceutical applications. The hypervariable loop region of antibodies, supported by a structurally rigid framework, provides the vast repertoire of antigen-binding sites in the immune system. In the last decade, the concept of a universal binding site derived from antibody structure has been successfully transferred to alternative protein frameworks – the so-called “scaffolds”, which can tolerate a high degree of insertions, deletions or substitutions. Therefore a scaffold library containing large amount of variants can be created by randomizing the surface-exposed residues and selected against predefined molecular targets, generating target-specific artificial binding proteins.

The pituitary adenylate cyclase-activating polypeptide type I receptor (PAC1-R) belongs to the class B G-protein coupled receptor (GPCR) family, which includes receptors for a variety of therapeutically relevant peptide hormones. PAC1-R distributes abundantly in both central and peripheral nervous systems and has been proposed as a potential target for the treatment of epilepsy, neuro-degeneration and cognition disorders.

This work therefore aims to generate specific Affilin[®] binding proteins against the N-terminal extracellular domain of human PAC1-R (nPAC1-Rs) for structure determination of nPAC1-Rs and full length PAC1-R by co-crystallization, since membrane proteins could be stabilized by their binding proteins. The thesis describes the efficient selection and characterization of nPAC1-Rs-specific Affilin[®] binding proteins. The Affilin[®] library created by randomizing 15 surface-exposed residues on a dimeric ubiquitin-based scaffold molecule was selected against nPAC1-Rs, utilizing Tat-mediated phage display technology. Enrichment of binders was observed after four rounds of selection. After further screening with Hit-ELISA more than 700 variants were sequenced, obtaining 25 candidates exhibiting interesting binding properties and sequences. Characterization of binding affinity and specificity by ELISA and SPR experiments, as well as determination of expression yield, solubility and thermostability for promising variants revealed that several selected Affilin[®] binding proteins have the potential for further application in structure determination. The most promising candidates, showing low nanomolar to picomolar binding affinity to nPAC1-Rs, can now be utilized for co-crystallization with nPAC1-Rs and the full length PAC1-R.

Keywords: dimeric ubiquitin-based Affilin[®] library, Tat-mediated phage display, nPAC1-Rs, selection, characterization, Affilin[®] binding proteins

Table of contents

1. INTRODUCTION	1
1.1. Affinity proteins and their applications.....	1
1.1.1. Introduction of affinity proteins	1
1.1.2. Examples for the application of affinity proteins	1
1.2. Antibody.....	3
1.2.1. Introduction of antibody molecules	3
1.2.2. Application and limitation of antibody-based therapeutics.....	6
1.3. Scaffold–alternative artificial binding protein	7
1.3.1. Scaffold concept	7
1.3.2. Structural classes of promising scaffold proteins	10
1.3.3. Applications for alternative scaffold proteins.....	13
1.4. Affilin[®] – ubiquitin-based scaffold technology.....	15
1.4.1. General knowledge of ubiquitin	15
1.4.2. Affilin [®] concept	16
1.4.3. Dimeric ubiquitin-based Affilin [®] technology.....	17
1.5. Selection of artificial binding proteins by Tat-phage display	18
1.6. The N-terminal extracellular domain of PAC1-R as a target protein	20
1.6.1. G-protein coupled receptors	20
1.6.2. The N-terminal extracellular domain of PAC1-R.....	21
1.7. Project aims.....	25
2. MATERIALS AND METHODS.....	26
2.1. Materials	26
2.1.1. Chemicals	26
2.1.2. Bacterial strains	27
2.1.3. Helper phage	27
2.1.4. Plasmids.....	27
2.1.5. Oligonucleotides.....	28
2.1.6. Enzymes.....	28
2.1.7. Proteins and antibodies.....	29
2.1.8. Standards and kits	29
2.1.9. Buffers and solutions.....	29
2.1.10. Medium and antibiotics.....	30
2.1.11. Chromatography columns.....	31
2.1.12. Laboratory equipment.....	31

2.1.13.	Miscellaneous	32
2.2.	Methods	33
2.2.1.	General molecular biology methods	33
2.2.1.1.	Plasmid DNA preparation.....	33
2.2.1.2.	Agarose gel electrophoresis	33
2.2.1.3.	DNA recovery from agarose gel	33
2.2.1.4.	DNA concentration measurement	33
2.2.1.5.	Polymerase chain reaction (PCR)	34
2.2.1.6.	DNA digestion	35
2.2.1.7.	Dephosphorylation of vector DNA.....	35
2.2.1.8.	DNA ligation	35
2.2.1.9.	Transformation with chemically competent <i>E. coli</i> cells	35
2.2.1.10.	Transformation with electroporation competent <i>E. coli</i> cells	36
2.2.1.11.	DNA sequencing and analysis	37
2.2.2.	Production of the target protein nPAC1-Rs	37
2.2.2.1.	Expression of nPAC1-Rs.....	37
2.2.2.2.	Purification of nPAC1-Rs	37
2.2.3.	Selection of nPAC1-Rs-specific Affilin [®] binders by TDP technology	38
2.2.3.1.	Preparation of biotinylated nPAC1-Rs as a target	38
2.2.3.2.	Phage preparation for selections	38
2.2.3.3.	Selection against biotinylated nPAC1-Rs.....	39
2.2.3.4.	Screening by single phage ELISA	42
2.2.3.5.	Screening by high throughput Hit-ELISA	43
2.2.4.	Expression and purification of selected Affilin [®] binders	44
2.2.4.1.	Expression of Affilin [®] binders	44
2.2.4.2.	Purification of Affilin [®] binders.....	44
2.2.5.	Protein characterization	45
2.2.5.1.	Protein concentration measurement.....	45
2.2.5.2.	SDS-PAGE analysis	45
2.2.5.3.	Expression and solubility analysis	45
2.2.5.4.	Circular Dichroism (CD) Spectroscopy.....	46
2.2.5.5.	Isothermal Titration Calorimetry (ITC)	46
2.2.5.6.	Differential scanning fluorimetry (DSF).....	47
2.2.5.7.	Specificity test by ELISA.....	47
2.2.5.8.	Concentration-dependent ELSIA.....	47
2.2.5.9.	Surface plasmon resonance (SPR) experiment	48
3.	RESULTS.....	49
3.1.	Recombinant production of functional nPAC1-Rs	49
3.1.1.	Recombinant production of nPAC1-Rs.....	49
3.1.2.	Preparation of biotinylated nPAC1-Rs	51
3.1.3.	Biophysical characterization of nPAC1-Rs.....	54

3.2.	Selection of dimeric ubiquitin-based Affilin[®] binders against nPAC1-Rs.....	57
3.2.1.	Selection against nPAC1-Rs by Tat phage display	57
3.2.2.	Screening by single phage ELISA.....	59
3.2.3.	Screening by Hit-ELISA.....	61
3.2.4.	Data analysis of the screening by Hit-ELISA.....	64
3.2.5.	Sequence analysis of Affilin [®] variants	65
3.3.	Characterization of selected Affilin[®] binding proteins.....	68
3.3.1.	Analysis of expression and solubility	68
3.3.2.	Purification of Affilin [®] binding proteins	71
3.3.3.	Thermostability	73
3.3.4.	Affinity measurements by concentration-dependent ELISA.....	74
3.3.5.	Binding specificity of Affilin [®] binders.....	77
3.3.6.	Biosensor binding analysis.....	78
3.3.7.	Biosensor competitive assay	81
3.3.8.	Investigation of monomer ubiquitin variants	82
4.	DISCUSSION	86
4.1.	Recombinant production of target protein nPAC1-Rs	86
4.2.	Dimeric ubiquitin-based Affilin[®] library	88
4.3.	Selection of Affilin[®] binding proteins	89
4.4.	Screening of Affilin[®] binding proteins	91
4.5.	Characterization of selected Affilin[®] binding proteins.....	92
4.5.1.	Sequence analysis of Affilin [®] binding proteins.....	92
4.5.2.	Recombinant production of Affilin [®] binding proteins	94
4.5.3.	Thermostability of Affilin [®] binding proteins.....	94
4.5.4.	Specificity of Affilin [®] binding proteins.....	95
4.5.5.	Affinity of Affilin [®] binding proteins	95
4.5.6.	Monomer ubiquitin domains of Affilin [®] binding proteins.....	96
4.6.	Applications of selected Affilin[®] binding proteins.....	97
5.	SUMMARY.....	98
6.	REFERENCES	100
7.	SUPPLEMENTARY MATERIAL	116
7.1.	Construction of plasmid pTrS/nPAC1-Rs.....	116
7.2.	Construction of Affilin[®] library.....	116

8. ABBREVIATIONS	118
PUBLICATION LIST	120
CURRICULUM VITAE	121
ERKLÄRUNG	122

1. Introduction

1.1. Affinity proteins and their applications

1.1.1. Introduction of affinity proteins

Nearly every biological interaction in life processes is due to an initial binding event mediated by affinity proteins, for example the binding of antibodies and antigens, the interaction between receptors and hormones, as well as the binding of ribosomes to mRNAs. In general, such binding interactions can be characterized by binding affinity and specificity, classifying for a variety of biotechnological applications. The binding affinity describes the interaction strength between the participating biomolecules. The specificity is used as a term to measure the binding ability to target and other molecules. High specificity is essential to avoid any unwanted cross-reactivity to irrelevant molecules.

Various classes of affinity proteins are widely used for applications involving molecular recognition phenomena such as bioseparation, detection, diagnosis and therapy. The capability of an affinity protein to recognize and bind a target molecule can be utilized in many ways. Three typical formats involve the use of affinity proteins (i) as solid support-immobilized ligands for recovery of interacting target molecules (for example affinity chromatography), (ii) for in vitro target detection (such as enzyme-linked immunosorbent assay, immunohistochemistry and western blot), (iii) for in vivo targeting and/or functional blocking of cellular receptors in clinical applications.

Depending on the application, the affinity proteins must first of all have sufficient affinity. A typical affinity for the most frequently used affinity proteins is in the range of 10^{-5} - 10^{-10} M (dissociation constant, K_D). For example, a value of 10^{-5} M and below can be considered as a quantifiable binding affinity. A value of 10^{-6} - 10^{-12} M is preferred for chromatographic applications and 10^{-9} - 10^{-12} M for diagnostic and therapeutic applications.

1.1.2. Examples for the application of affinity proteins

Avidin – a biotin binding protein

Of all the affinity proteins, avidin and its derivatives probably are most employed for affinity capture applications. Streptavidin and Neutravidin bind the small molecule biotin with extraordinary high affinity of $K_D = 4 \times 10^{-14}$ M and 6×10^{-16} M, respectively (Green, 1975). Avidin has also been found to be extremely stable. The tertiary

structure and ability to bind biotin remain during exposure to high concentration of sodium dodecyl sulphate (SDS) at 70 °C, leading to numerous applications ranging from research to diagnostics to medical devices and pharmaceuticals (Wilchek and Bayer, 1990).

Some modified forms of avidin extend the applications successfully. Neutraavidin, a deglycosylated avidin with modified arginines, exhibits a more neutral isoelectric point (pI) and is available as an alternative to native avidin, wherein the problems of non-specific binding arise. Furthermore, due to the strong binding of avidin to biotin, the non-reversible property of the avidin-biotin complex limits the application of avidin in affinity chromatography. An avidin with reversible binding characteristics has been created by nitration or iodination of the binding site tyrosine. This modified avidin exhibits strong binding characteristics at pH 4 and releases biotin at pH 10 or higher pH (Morag *et al.*, 1996). A monomeric form of avidin with a reduced affinity for biotin is also applied in many commercial affinity resins. The monomer avidin is created by treatment of immobilized native avidin with urea or guanidine HCl (6-8 M), giving it a lower K_D of 10^{-7} M. This characteristic allows elution from the avidin matrix to occur under milder and non-denaturing conditions, using low concentration of biotin or lower pH conditions (Kohanski and Lane, 1990).

Bacterial receptins protein A, G and L

Some pathogenic bacterial species express cell surface displayed proteins capable of interacting with mammalian host proteins, for example albumin, fibronectin and immunoglobulins. Like many other affinity proteins, their biological functions haven't been fully understood, but one proposed explanation might be that the bacteria use the host molecule interaction as means for camouflaging itself to become more host-like and can escape the host immune defence (Achari *et al.*, 1992; Sauer-Eriksson *et al.*, 1995). The immunoglobulin binding receptins bind their targets involving other regions than complementarity determining regions (CDRs), like Fc (Fragment, crystallizable) and constant heavy chain region 1 (CH1), making them interesting as universal antibody binding tools in immunotechnology applications. Three most widely used immunoglobulin binding proteins are *Staphylococcus aureus* Protein A (SPA), *Streptococcal* Protein G (SPG) and *Peptostreptococcus magnus* Protein L (PPL).

The non-immune binding characteristics of these receptins provide important tools for diverse applications involved in purification, recognition and removal of immunoglobulins from different species and different isotypes, the use as amplification reagents in immunoassays such as enzyme-like immunosorbent assay (ELISA), and the use as fusion partners to facilitate purification and immobilization of recombinant proteins.

Antibody reagents

Antibodies are the most well-known and extensively used affinity proteins in

biotechnological and medical applications for over a hundred years. Antibodies are essential for the adaptive immune system of higher vertebrates. As a natural response to any immunogen like bacterium, virus, pollen or cancer cells, antibodies are secreted into body fluids by specialized plasma B cells. Antibodies have the capability of specific and tight binding to a given antigen, for example protein, nucleic acid, lipid, carbohydrate or haptan. The antibodies produced during humoral immunity defend the body through a variety of different ways, including opsonization, functional neutralization, antibody-dependent cellular cytotoxicity (ADCC) by nature killer (NK) cells, membrane attack complex (MAC) cytolysis, preventing bacterial adherence to host cells, agglutination of microorganisms as well as immobilization of bacteria and protozoans (Adams and Weiner, 2005). Antibodies can be isolated from the blood of immunized animals or supernatants from hybridoma cell cultures. With the invention of recombinant DNA technology (Mullis and Faloona, 1987), antibodies can be generated in vitro and produced in microbial hosts.

Due to the abilities and functions of antibodies, nature and engineered antibodies are widely used in many applications. In research field, antibodies are frequently used to identify, locate and isolate intracellular and extracellular proteins in various ways, including Affinity purification, immunofluorescence, ELISA, western blot, flow cytometry and immunohistochemistry. In medical applications, antibodies are commonly employed in serology test, infectious diseases diagnosis as well as the imaging and therapy of cancer.

1.2. Antibody

1.2.1. Introduction of antibody molecules

In general, most antibody molecules have a Y-shaped structure comprising a pair of identical shorter light chain and longer heavy chain linked by disulfide bonds (Fig. 1.1). The light chain consists of one variable domain VL and one constant domain CL. The heavy chain consists of one variable domain VH and three constant domains CH1, CH2 and CH3. Functionally, antibody is divided into two antigen-binding fragments (Fabs) consisting of one VH domain and one VL domain each, as well as Fc fragment comprising the hinge, CH2 and CH3 domains.

The antigen binding site is located on the tips of the Y-shaped antibody molecules. Both VH and VL regions contribute with three variable loops from each, called complementarity determining regions (CDRs) and arranged non-consecutively (Fig. 1.2). The CDRs form a continuous and highly plastic binding surface (Skerra, 2003). The marvelous variability and plasticity of the antigen binding site engender the ability to bind a variety of different antigens with high affinities and specificity. These adaptive binding characteristics might be donated by a high sequence diversity of the six variable loops, several possible CDR loop conformations or canonical structures,

fine-tuning of shape complementarity by single somatic mutations in the CDR framework or Verniet region, allowing for structural adaptation to different antigens (Foote and Winter, 1992; Dubreuil *et al.*, 2005).

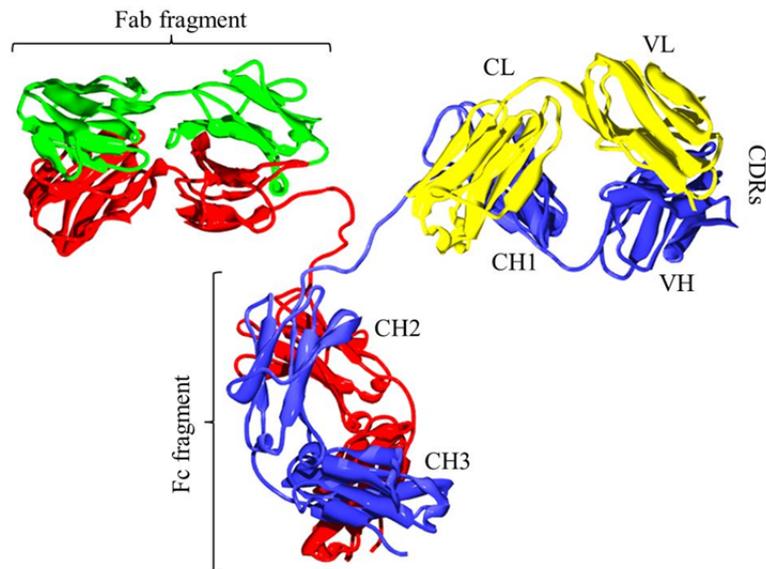


Figure 1.1 Structure model of an IgG2 antibody (PDB code 1IGT). The two light chains are in green and yellow, and two heavy chains are in red and blue. Figure modified from Tim Vickers.

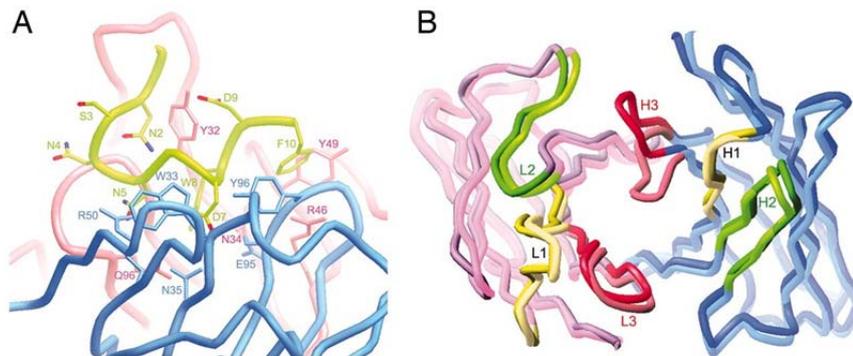


Figure 1.2 Crystal structure of the HzKR127 Fab-preS1 complex to represent CDRs. (A) Conformation of the preS1peptide bound to HzKR127 Fab. The light, heavy and peptide chains are shown in red, blue and green, respectively. (B) The backbone structure of the variable domains in the Fab is drawn as thick tubes, and the light and heavy chains are colored pink and blue, respectively. The free and bound forms are shown in lighter and darker colors, respectively. The CDR1, CDR2, and CDR3 loops are colored yellow, green and red, respectively. Figure modified from Chi *et al.*, 2007.

The large size and complicated structure of full length antibodies (~150 kDa) lead to certain limitations in many applications, related to the high cost production and undesired properties. The Fc-mediated immunological effector functions are only

desired for some certain applications. An inappropriate activation of Fc receptor-expression cells, like neutrophils, natural killer (NK) cells and macrophages can lead to unwanted side effects (Chatenoud, 2004). The bulky frame of antibodies also limits tissue penetration which might complicate some medical applications (Beckman *et al.*, 2007).

Advances in recombinant DNA and protein engineering technologies allow the development of smaller antibody fragments like scFv, Fab, F(ab')₂ fragments (25, 55 and 110 kDa, respectively) and other derivatives, including minibody (80 kDa) and diabody (55 kDa) as shown in Fig. 1.3 (Holliger and Hudson, 2005; Cater, 2006). The most frequently used unit to create novel antibody formats is the single-chain variable domain antibody fragment (scFv), which comprises V_H and V_L domains joined by a peptide linker. Both the scFv and Fab fragments are versatile for protein engineering. They can be tailored and produced at high yields in relatively cheap microbial hosts, and combined to a variety of effector functions in biospecific or tri-specific formats (Lonberg, 2005). By contrast to the traditional generation of antibodies by animal immunization and hybridoma technology, an alternative route has been established following the progress in the fields of antibody fragment cloning, engineering and production as well as powerful selection platforms.

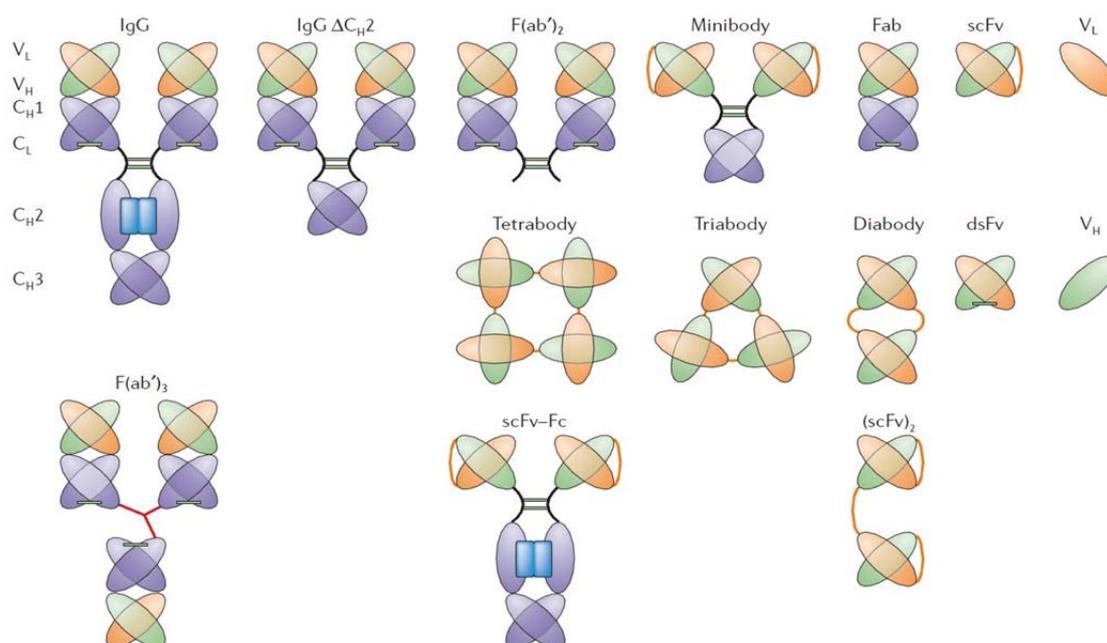


Figure 1.3 Representative antibody formats. The modular domain architecture of immunoglobulins has been exploited to create alternative antibody formats that spans a molecular weight range of 12–150 kDa and a valency (n) range from monomeric ($n=1$), dimeric ($n=2$) and trimeric ($n=3$) to tetrameric ($n=4$) and possibly higher. V_L, V_H and all constant domains are shown in orange, green and blue, respectively. The disulfide bonds, glycosylation, peptide and chemical linkers are depicted as green bars, dark blue blocks, orange and red lines, respectively. The dsFv is disulfide-stabilized scFv. Figure modified from Cater, 2006.

1.2.2. Application and limitation of antibody-based therapeutics

Polyclonal antibody preparations derived from serum are used for the treatment of infectious diseases and snake bites. Depending on how they are purified, such preparations may be of very uncertain composition and may cause severe side reactions in patients (Walsh, 2007). Typically several different animal species have been used for the generation of polyclonal antibodies. The choosing of animals for example rabbit, mouse, rat, hamster, goat, guinea pig and lately hen, depends on several parameters, such as the amount of antibody required and the phylogenetic distance between the donor of antigen and the recipient (Hanly *et al.*, 1995). First clinical use of polyclonal antibody approved by Food and Drug Administration (FDA) was ATG-Fresenius S in 1977.

Monoclonal antibodies (mAbs) are available as a scientific and pharmaceutical resource since the establishment of hybridoma technology (Kohler and Milstein, 1975). The first therapeutic monoclonal antibody approved by FDA was a murine IgG2a CD3-specific transplant rejection drug in 1986. More than 150 such drugs are in clinical development (Leader *et al.*, 2008) and >30 antibody-based therapeutics have been approved for the treatment of various diseases by FDA until 2011. Furthermore, hundreds of monoclonal antibodies have been widely used in research, diagnostic applications and chromatography.

The first developed hybridoma technology uses the mouse hybridomas generated from the stable fusion of immortalized myeloma cells with B cells from immunized mice, and makes great success in research. By contrast, a paltry success rate of 3% (Reichert, 2005) is approved in drug development mainly caused by the high immunogenicity of these exogenous proteins in humans, and inefficient effector functions of ADCC and complement dependent cytotoxicity (CDC) due to the weak interactions between mouse antibody Fc region and human IgG receptors (Fc γ Rs) and complement component 1q (C1q), respectively. In addition, mouse antibodies do not bind the human salvage receptor FcRn (Ober *et al.*, 2001), leading to a terminal half-life less than 20 hours (Carter, 2001; Presta, 2002).

A detailed understanding of the structures of V domains and the CDRs of antibodies provides the accessibility of antibody chimerization and humanization, which have been largely used and overcome the shortcoming of mouse-derived antibody in therapy (Kipriyanov and Le Gall, 2004; Gonzales *et al.*, 2005). Chimerization involves joining the VH and VL domains of mouse antibody to the constant domains of a human antibody (Boulianne *et al.*, 1984; Morrison *et al.*, 1984). The simplest of many approved humanization strategies involves grafting of CDRs from a mouse antibody to a human IgG, predominantly to the IgG1 isotype (Jones *et al.*, 1986; Verhoeyen *et al.*, 1988 and Riechmann *et al.*, 1988). In general, additional transfer of one or more framework-region residues from the parent mouse antibody is required for the generation of humanized antibodies with a high antigen-binding activity

(Gonzales, 2005). It should be pointed out that chimeric antibody and humanized antibody still contain approximately 33% and 10% rodent amino acid residues, respectively.

For human immunotherapy applications, monoclonal antibodies of human origin would be ideal reagents to avoid human anti-mouse-antibodies (HAMA) immune responses upon administration. Several routes have been used for the generation of human antibodies, including human hybridomas from patients (Brandlein and Vollmers, 2004; Illert *et al.*, 2005), antibody-cDNA cloning from single lymphocytes selected on antigen (Lagerkvist *et al.*, 1995; Babcook *et al.*, 1996), selection from recombinant antibody libraries (Hoogenboom, 2005) and transgenic mice that express human immunoglobulin genes (Green, 1999; Lonberg, 2005). Between 1997 and 2008, a total of 131 human mAbs entered clinical study, of which 88 were in active clinical development, with 30 in Phase I, 51 in Phase II and 7 in Phase III studies. A total of 7 were approved for marketing by the FDA, 3 are undergoing review by the FDA. Most of the growing numbers of antibodies entering clinical trials are completely human antibodies derived from phage display technology or transgenic mice.

In recent years, great progress has been made for the various antibodies and related formats towards the aim of using antibody as therapeutics. There is no doubt that the profound knowledge and accumulated experience will navigate novel antibody technologies and accelerate further developments. However, with increasing applications in research, biotechnology and medical therapy, some fundamental limitations of antibodies have emerged. For example, antibodies are rather larger molecules with composition of four individual protein chains, resulting inefficient tissue penetration and weak ADCC/CDC for the treatment of solid tumours. The stability of antibodies and related fragments relies on disulfide bonds, which do not form in reducing intracellular environments. Some antibody fragments tend to aggregate, especially when fused to additional domains, for instance added to achieve therapeutic efficacy. The constant Fc region mediates immunological effector functions that are only crucial for few pharmaceutical applications and often lead to undesired interactions. The production of glycosylated antibodies in mammalian expression technologies is time and cost-consuming, and facing limited capacities.

1.3. Scaffold–alternative artificial binding protein

1.3.1. Scaffold concept

Scaffold concept firstly emerged in the 1990s. The protein engineering and powerful library selection technologies developed for antibody libraries provided the possibility to recreate the function of the immune system *in vitro*. These technologies were extended to non-antibody proteins and generated the first examples of protease

inhibitors with matured affinity or changed specificity, such as neutrophil elastase inhibitor (Robert *et al.*, 1992) and ecotin (Wang *et al.*, 1995).

During the same period, with increasing knowledge about antibody, the limitations of antibody inspired scientists to transfer the concept of a universal binding site derived from the antibody structure to alternative protein frameworks – the so-called “scaffolds”, to generate non-antibody binding proteins. A simpler example is that peptides with known affinity to a certain target can be inserted into a rigid scaffold protein to combine its binding ability with the desired favorable characteristics of the scaffold, analogous to grafting of CDRs to antibodies (Ali *et al.*, 1999). More universal way is selection from a scaffold library created by genetic variation of a set of surface exposed residues against a certain target to obtain target-specific binding proteins.

Throughout the past decade, over 50 different scaffolds have been investigated and proved for the generation of alternative binding reagents. These scaffolds have different topologies and folds, and different structural elements mediate the target interactions. Many of them are approaching the pharmaceutical and industrial applications as shown in Tab. 1.1 (Nygren and Skerra, 2004; Binz and Pluckthun, 2005; Hey *et al.*, 2005).

In principle, a scaffold protein should ideally meet some general requirements (Skerra, 2009; Gebauer and Skerra, 2009). Usually, such a scaffold is derived from a robust and small soluble monomeric protein or from a stably folded extramembrane domain of a cell surface receptor, with a high tolerance for modifications for example multiple insertions, deletions or substitutions. Compared with antibodies or their recombinant fragments, these protein scaffolds often provide practical advantages including elevated stability and high production yield in microbial expression systems, together with an independent intellectual property situation.

In antibody variable domains, the vast repertoire of antigen-binding sites is provided by variation of length and sequence in three hypervariable loops region supported by structurally rigid framework, possessing a β -sandwich topology. Many different attempts have been utilized to provide such cavities or clefts for tight binding to different targets. As demonstrated in Fig. 1.4, they can be classified into three groups, (a) a single contiguous peptide loop on a scaffold, (b) more complex arrangements of several loops forming a continuous surface and (c) large non-contiguous regions based on secondary structural elements such as β -sheets or α -helices (Nygren and Skerra, 2004).

Table 1.1 Scaffold proteins for the generation of novel binding proteins for pharmaceutical and industrial applications (table adapted from Hey *et al.* 2005).

Scaffolds in use (underlying protein framework)	Application indication (target; status) ^a	Company
Nanobody: single-domain antibodies from the camelid family. Bacterial and yeast production	Therapy: rheumatoid arthritis (preclinical); inflammatory bowel disease (preclinical); thrombosis (arterial stenosis) (preclinical); psoriasis, solid tumours, Alzheimer's disease	Ablynx (Ghent, Belgium); http://www.ablynx.com
Affibody: three-helix bundle from Z-domain of Protein A from <i>S. aureus</i> . Bacterial production, chemical synthesis	Therapy: cancer (HER2/neu, preclinical); apheresis: Alzheimer's disease Chromatography: industrial-scale separation	Affibody (Bromma, Sweden); http://www.affibody.com
Maxibody: Scaffold not disclosed	Therapy: autoimmunity, inflammation, cancer	Avidia (Mountain View, CA, USA); http://www.avidia.com
Trans-body: human transferrin. Yeast production	Therapy: cancer	BioRexis (King of Prussia, PA, USA); http://www.biorexis.com
Tetranectin: monomeric or trimeric human C-type lectin domain. Bacterial production	Therapy: rheumatoid arthritis, Crohn's disease, HIV	Borean Pharma (Aarhus, Denmark); http://www.boreanpharma.com
iMabs: scaffold not disclosed	Chromatography: industrial-scale separation	Catchmabs (Wageningen, the Netherlands); http://www.catchmabs.com
AdNectin: human tenth fibronectin type III domain. Bacterial production	Therapy: anti-angiogenesis in cancer (VEGF, VEGF-R); rheumatoid arthritis, psoriasis, Crohn's disease (TNF α)	Compound Therapeutics (Waltham, MA, USA); http://www.compoundtherapeutics.com
Domain antibody: variable domain of human light or heavy chain. Bacterial and yeast production	Therapy: rheumatoid arthritis, Crohn's disease, (TNF α ; <i>in vivo</i> efficacy); rheumatoid arthritis (cytokine receptor; <i>in vivo</i> efficacy); autoimmune diseases (CD40L; <i>in vivo</i> efficacy); asthma (IL3, IL4), others	Domantis (Cambridge, UK); http://www.domantis.com
Kunitz-type domain of human and bovine trypsin inhibitor. Bacterial and yeast production	Therapy: hereditary angioedema (kallikrein; clinical phase II); open heart surgery (kallikrein; clinical Phase II); cystic fibrosis (neutrophil elastase; clinical Phase II)	Dyax (Cambridge, MA, USA); http://www.dyax.com
Evibody: human cytotoxic-associated Antigen (CTLA-4). Bacterial production	Therapy: cancer: tumour cell surface proteins	Evogenix (Sydney, Australia); http://www.evogenix.com
Ankyrin repeat protein. Bacterial production	Therapy: diagnostic	Molecular Partners (Zürich, Switzerland); http://www.molecularpartners.com
Anticalin: human lipocalins/lipocalins from <i>P. brassicae</i> butterfly. Bacterial and yeast production	Therapy: cancer, cardiovascular diseases	Pieris Proteolab (Freising, Germany); http://www.pieris.org
Affilin molecule: human γ -crystallin/human ubiquitin. Bacterial and yeast production	Therapy: cancer, ophthalmology, inflammation Chromatography: industrial-scale separation	Scil Proteins (Halle, Germany); http://www.scilproteins.com
Microbody: trypsin inhibitor II from <i>E. elaterium</i> . Bacterial and yeast production	Therapy: infectious diseases (<i>Pseudomonas aeruginosa</i> antigen); thrombosis, multiple sclerosis (α 4 β 1 integrin); rheumatoid arthritis (TNF α)	Selecure (Göttingen, Germany); http://www.selecure.com

^aAs disclosed on the corresponding company homepage in January 2005.

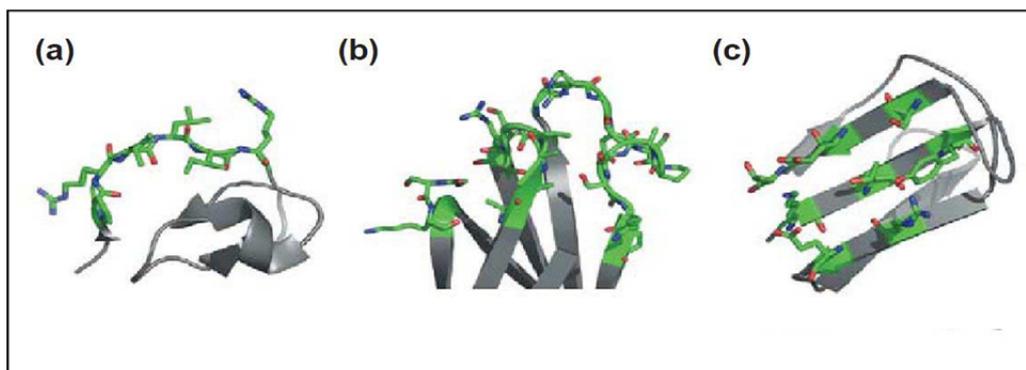


Figure 1.4 Structural bases of artificial binding sites on scaffolds. Residues involving artificial binding sites are in green and frameworks are in grey. Graphical representations of a, b and c were prepared using the software Pymol and coordinate sets 1H9H, 1FNA and 1AMM as deposited in the Protein Data Bank, respectively. Figure adapted from Hey *et al.*, 2005.

1.3.2. Structural classes of promising scaffold proteins

The protein scaffolds that have been proved with ability of generation artificial binding proteins can be classified into following four groups as shown in Fig. 1.5.

Antibody-related scaffolds

Some stable antibody derivatives have been engineered with favorable properties such as high affinities, high solubility, smaller in size and simpler structure as well as robust expression (Fig. 1.5a). Domain antibodies (Domantis) and Nanobodies (Ablynx) are using the variable domain of human light or heavy chain (Holt *et al.*, 2003; Jespers *et al.*, 2004) or single domain antibodies from camelidae family (Muyldermans, 2001; Huang *et al.*, 2010) for therapy and diagnosis, allowing the isolation of artificial binding proteins with affinities in a low nanomolar range. The increasing interest in antibody-related scaffolds as therapeutics is exemplified by that 7 of the 15 antibodies in Phase III clinical trials are antibody fragments (Reichert *et al.*, 2005).

Single loop on rigid scaffolds

The presentation of a single loop peptide on a scaffold (Fig. 1.5b) is probably the earliest attempt of modifying known binding activities, particularly in the context of protease inhibitors. The binding patch of this strategy is created by a loop either with a defined binding property, or that is hypervariable in length and substitutions of its residues (Roberts *et al.*, 1992; Dennis *et al.*, 1995). For example, Kunitz domain inhibitors are stable protein with ~60 residues, possessing three disulfide bonds, which act as slow but tight binding, reversible inhibitors of serine proteases. Two candidates have entered the phase II clinical trials for the treatment of cystic fibrosis (Wark, 2002) and hereditary angioedema (Williams and Baird, 2003). The Squash-type protease inhibitors, one of the smallest stable protein domains with only 30 residues are successfully used by modifying the original inhibitor loop (Christmann *et al.*, 1999; Baggio *et al.*, 2002; Hilpert *et al.*, 2003). Trans-bodies are scaffolds based on human serum transferrin comprising 679 residues. One or two surface exposed loops are suitable for the insertion of pre-selected or naïve polypeptides. The favorable properties of trans-bodies such as low toxic side effects and long circulatory half-life have been proved in clinical studies (Weave and Laske, 2003).

Multiple loops on rigid scaffolds

The closest way to imitate binding sites of antibody is variation of multiple hypervariable loops on a structurally rigid scaffold. In general, the artificial binding patch of these scaffold proteins is designed based on the loops which already present natural binding properties. Three these proteins are represented in Fig. 1.5c. For example, the human cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) ideally imitates the antibody concept. The extracellular domain (ECD) exhibits a V-like Ig fold. By randomization of nine residues of CDR-3-like loop, several CTLA-4-based

variants were selected out by phage display technology, with specific binding ability to human $\alpha\beta 3$ integrin (Hufton *et al.*, 2000). A similar library was used to select lysozyme-specific binding variants by ribosome display (Irving *et al.*, 2001).

The fibronectin type III domain is a small monomeric natural β -sandwich protein, which possesses seven β -strands with three connecting loops. The 10th domain of 15 repeating human FN3 was used as a scaffold by randomization of three N-terminal loops (Koide *et al.*, 2001). TNF α -specific variants with high affinities were successfully selected in vitro by mRNA-display (Xu *et al.*, 2002). A fibronectin-based scaffold (AdNectin) has been used to develop novel cancer therapeutics.

The lipocalins are a biologically widespread family, including ~60 structurally conserved members. They share a central β -barrel comprising eight antiparallel strands, which support a set of four loops that form the ligand-bind site. These four loops can be reshaped to generate artificial binding proteins (Anticalins) with high affinities against various haptens, peptides and protein targets (Beste *et al.*, 1999; Schlehuber *et al.*, 2000; Skerra, 2001; Schlehuber and Skerra, 2005). Remarkably, Anticalins with high affinities and specificities, down to the picomolar range, were successfully selected by phage display against several medically important targets, for example CTLA-4 (Schlehuber and Skerra, 2005) and *vascular endothelial growth factor* (Hohlbaum and Skerra, 2007).

Scaffolds providing interface on secondary structural elements

Compare to the loop-mediated binding mechanism described above, another type of molecular recognition can be achieved via amino acid residues that are positioned partially or completely within rigid secondary structure elements, for example on the surface of an α -helix bundle or β -sheets. The rigid binding domains seem to bind their target via a key-to-lock mechanism. It is speculated that such a rigid-body interaction may be advantageous both for affinity (low entropic costs upon binding) and specificity (conformational restriction) (Binz *et al.*, 2004).

One of the first protein scaffold investigated in this context is Affibody, a designed Z-domain variant of the IgG binding protein A (Nord *et al.*, 1995; Nord *et al.*, 1997). Affibody libraries are generated by randomization of up to 13 solvent exposed residues in the three- α -helical bundle (Wahlberg *et al.*, 2003). The libraries have typically been displayed on phages, followed by biopanning against desired targets. Using this strategy, Affibody molecules showing specific binding to a variety of different proteins (e.g., insulin, fibrinogen, transferrin, tumor necrosis factor- α , IL-8, gp120, CD28, human serum albumin, IgA, IgE, IgM, HER2 and EGFR) have been generated, demonstrating affinities in the micromolar to picomolar range (Nygren, 2008; Löfblom *et al.*, 2010). Importantly, due to their small size and rapid folding properties, Affibody molecules can be produced by chemical peptide synthesis (Engfeldt *et al.*, 2005). Affibody technology is towards bioseparation, imaging and therapeutic applications.

The designed ankyrin repeat proteins (DARPin)s represent another rigid scaffold protein with varying numbers of modular architectures. Each individual module comprises a β -turn and two antiparallel α -helices. The binding area is accomplished by randomization of six residues per module (Binz, 2003; Forrer *et al.*, 2004). The selected variants revealed affinities in the nanomolar to picomolar range via phage display (Steiner *et al.*, 2008) or ribosome display (Binz *et al.*, 2004). Company Molecular Partners is now exploiting the DARPin technology to generate artificial binding proteins for diagnostic and therapeutic applications.

The human γ B-crystallin is a small and remarkably stable protein with 176 amino acid residues. Eight residues at exposed positions in three neighboring β -strands were randomized to create a binding site (Fiedler and Rudolph, 2001; Ebersbach *et al.*, 2007). The γ B-crystallin variants with affinities up to low nanomolar range have been isolated against a variety of target proteins.

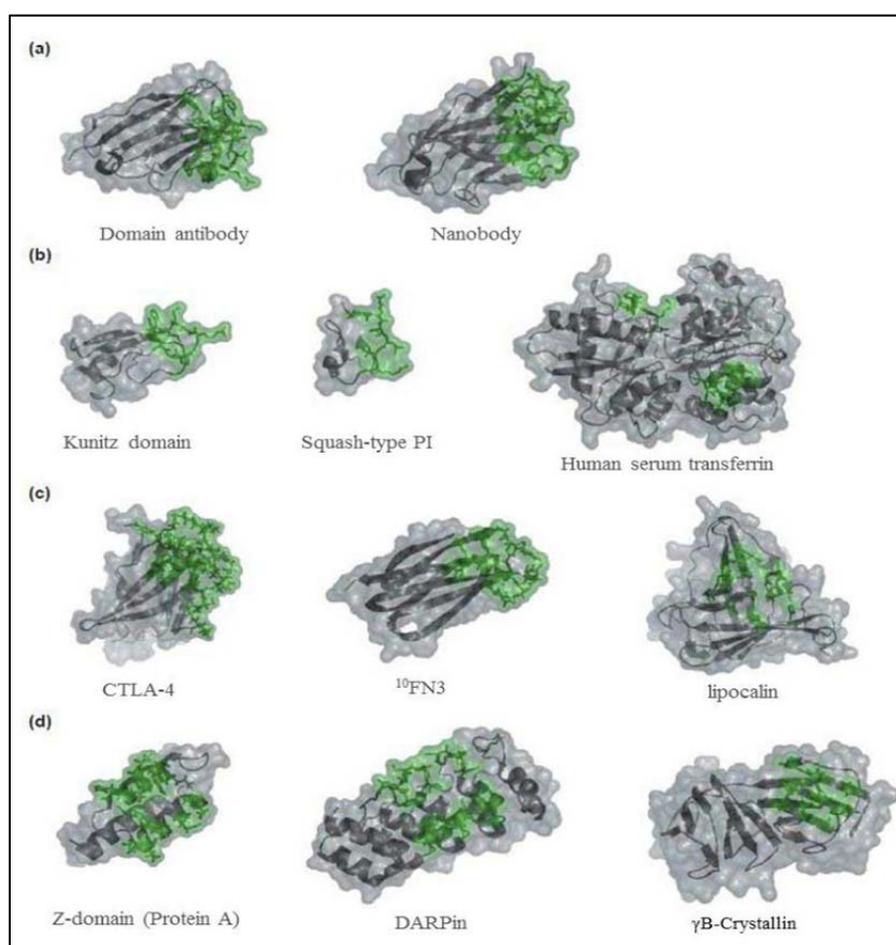


Figure 1.5 Four groups of scaffold proteins. The randomized surface regions are in green, whereas the conserved frameworks and surface are in grey. (a) Antibody derivatives: single-domain antibodies from human origin (left, chain A from PDB entry 1OHQ) and camelidea (right, chain E from 1KXQ); (b) Rigid protein frameworks for the insertion or randomization of single loop peptide, from left to right: Kunitz domain (chain A from 1AAP), Squash-type protease

inhibitor (chain I from 1H9H, residues 1 to 28) and human serum transferrin (1A8E); (c) Proteins imitating the antibody structure by presenting multiple loops for randomization on rigid frameworks, from left to right: human cytotoxic T lymphocyte-associated antigen 4 (1AH1), tenth fibronectin type III domain (1FNA) and lipocalin (chain A from 1BBP); (d) Proteins providing secondary structure elements for randomization, from left to right: Z-domain of Protein A (2SPZ), designed ankyrin repeat protein (1MJ0) and γ B-crystallin (1AMM). Graphical representations were prepared using the software Pymol. Figure modified from Hey *et al.*, 2005.

1.3.3. Applications for alternative scaffold proteins

In the past decade, the concept of “artificial binding protein” has been proved in various scaffold proteins. These novel scaffolds have already presented attractively intrinsic properties towards (i) therapeutics, (ii) diagnostics, (iii) intracellular applications, (iv) cocrystallization and (v) biotechnological applications. In addition, the criteria and technologies developed from antibodies are gradually accelerating the applications of alternative scaffold proteins.

Alternative scaffolds in therapeutic applications

A comparison of characteristics between scaffolds and antibodies is made in Tab. 1.2. Scaffolds do not have effector functions such as Fc region that can activate the complement system or cytotoxic cells. Their therapeutic potency seems to be limited to target neutralization. However, effector functions can be conferred to scaffolds by fusion to cytokines (Helguera *et al.*, 2002; Davis and Gillies, 2003), toxins (Pastan, 2003; Choudhary *et al.*, 2011) or Fc region (Rönnmark *et al.*, 2002). Fusion to Fc region might mediate dimerization or induce ADCC (Powers *et al.*, 2001).

Table 1.2 Characteristics of scaffolds and antibodies as therapeutics (table adapted from Hey *et al.*, 2005).

Property	Monoclonal antibodies (mABs)	Antibody fragments (Fab/scFv)	Scaffold-based proteins
Effector function	+	–	–
Neutralizing activity	+	+	+
Clearance	Slow	Fast	Fast
Tissue penetration	–	+	+
Ease of modification	+/-	+	+
Targeted delivery	+/-	+	+
Human origin	+/-	+/-	+
Intracellular activity	–	+/-	+/-
Oral delivery	–	–	+/-

Immunogenicity and safety profiles should be carefully assessed for all the scaffolds intended for therapy. Indeed, all protein drugs are potentially immunogenic, as the final molecule will be a nonhuman protein (Chirino *et al.*, 2004). Even fully human antibodies may elicit some degree of immune response (Koren *et al.*, 2002). However, different strategies are now emerging for rational reduction of protein immunogenicity, for example PEGylation (Chapman, 2002) and T-cell epitope engineering (Schirle *et al.*, 2001; Flower, 2003; Kim *et al.*, 2010). Nevertheless, the final effect of these strategies should be proved in a clinical trial.

Other characteristics must be also considered, including pharmacodynamic parameters, such as clearance rate, serum stability, tissue penetration and tissue-to-blood ratios. Due to the small size, most scaffolds exhibit improved tissue penetration but short serum half-lives. A common used strategy for size enlargement of biopharmaceuticals is the covalent attachment of polyethylenglycol (PEG) or fusion to plasma proteins such as serum albumin.

Currently more than 10 scaffold proteins are in clinical trials, of which 6 are in Phase II trials. Encouragingly none of them elicited severe adverse reactions or anti-drug antibody responses during Phase I clinical trials (Beck *et al.*, 2010).

Alternative scaffolds in imaging applications

As noted previously, most scaffolds have a small size and do not contain Fc region. They can show pharmacokinetic properties well suited for imaging applications, such as more efficient penetration of tumours and faster clearance from the body, providing improved resolution and low background (Tolmachev, 2008; Miao *et al.*, 2010).

Alternative scaffolds in intracellular applications

Many drug targets are located in the cytoplasm of cells. Proteins binding to these targets would be well suited for research on such drug targets and eventually for intracellular therapy. Various alternative scaffolds fold efficiently under reducing conditions, allowing cytoplasmic expression in soluble form. This could also enable the efficient binding to target proteins in cytoplasm (Koide *et al.*, 2002; Amstutz *et al.*, 2005) and offer an alternative to RNA interference or gene-knockout strategies (Couzin, 2004).

Alternative scaffolds for cocrystallization

High-resolution crystal structures of proteins not only contribute to the understanding of biological processes but also are important for rational drug design. Crystallization is difficult for many proteins due to intrinsic flexibility, for example the kinases, or because of the presence of only a small hydrophilic portion, such as membrane proteins (Iwata *et al.*, 1995). In both cases, specific binding proteins might support crystallization, either by locking the flexible protein in a specific conformation by providing a rigid surface for crystal contacts, or by increasing the hydrophilic portion. Two alternative scaffolds, Affibody and DARPIn have been successfully used for cocrystallization with different target proteins (Sennhauser *et al.*, 2007; Huber *et al.*, 2007).

Alternative scaffolds for biotechnological applications

Scaffold proteins with stable and compact fold are valuable affinity reagents for biotechnological applications, for example biosensor and affinity chromatography. These reagents should be stable enough to against a wide range of pH, temperature, chemicals and proteases, and maintain binding activities after tough regeneration conditions as generally used processes. The soluble expression in microbial hosts

might also offer the possibility of high yield and low-cost production. Affinity chromatography applications with scaffolds have been reported (Nord *et al.*, 2000; Reina *et al.*, 2002) for example affibodies, which work in a similar manner to their progenitor, protein A, a binder used in commercial immunoglobulin purification.

1.4. Affilin[®] – ubiquitin-based scaffold technology

1.4.1. General knowledge of ubiquitin

Ubiquitin is a small protein that ubiquitously exists in all known eukaryotic cells. In cytosol, it plays a crucial role in the regulation of the controlled degradation of cellular proteins, by attaching to the proteins destined for degradation and directing to the proteasome, which is a large protein complex in the cell that degrades and recycles unneeded proteins. Recent results have revealed that ubiquitin is involved in a diversity of cell functions. It participates in the G1 phase of the cell cycle, DNA repair, embryogenesis, immune defense, transcription, apoptosis and preventing self-pollination in plants (Marx, 2002; Ciechanover and Iwai, 2004).

Ubiquitin consists of a polypeptide with 76 amino acid residues folded in an extraordinarily compact α/β structure (Fig 1.6). Almost 87% of the polypeptide chain is involved in the formation of the secondary structural elements. 62% of the external solvent-accessible surface is covered with hydrophilic residues both charged and polar. The secondary structure of ubiquitin has three and a half α -helical turns, a short piece of 3_{10} helix (a helix with three residues per turn instead of 3.6 for α -helices), as well as a mixed β -sheet that contains five strands and seven reverse turns. Its core is organized in a $\beta(2)$ - α - $\beta(2)$ fashion known as the β -grasp fold (Vijay-Kumar *et al.*, 1987). Many other proteins share this kind of fold and are named for the protein family “ubiquitin-like proteins” (Murzin *et al.*, 1995).

Due to the favorable folding properties, ubiquitin is a highly stable and very soluble globular protein. Its folded structure is quite stable against thermo, chemicals and proteases. To unfold ubiquitin through heating in solution, the temperature needs to reach around 100 °C (Ibarra-Molero *et al.*, 1999). Ubiquitin can be produced by genetic engineering using microorganisms such as *Escherichia coli* in relatively large amounts either in cytosol or in periplasm. Because of its small size, ubiquitin can also be produced by chemical synthesis with native conformation (Ramage *et al.*, 1994).

The proteins of ubiquitin-like protein superfamily are highly conserved. Ubiquitin has an identical amino acid sequence in all mammals. The ubiquitin of human only has three amino acid residues different from yeast.

1.4.2. Affilin[®] concept

Due to the remarkable stability against chemical and physical denaturation, ubiquitin was chosen as a scaffold protein, to generate novel target-specific binding proteins by randomization of surface-exposed amino acid residues and selection against certain targets, including haptens, peptides, proteins, sugars, DNA etc. This is called Affilin[®] technology developed by Scil Proteins.

Surface-exposed amino acid residues are accessible to the surrounding solvent. Based on the X-ray crystallographic structure data of ubiquitin (1UBQ in Protein Data Bank), the residues with surface-exposed side chains that direct towards the solvent or a potential binding partner can be located by computerized analysis. If the accessibility of the amino acids is more than 8% compared to the accessibility of the amino acid in the model tripeptide Gly-X-Gly, the amino acids are defined surface-exposed (Shrake and Rupley, 1973; Connolly, 1983). Furthermore, those residues, whose random substitution presumably would have no or only a slightly negative effect on the stability of ubiquitin can be identified by computerized analysis. This information can provide a first indication as to the suitability of each single residue as an element of a binding site and would then require further experimental verification (Rudolph *et al.*, 2011).

The ubiquitin protein scaffold extraordinarily tolerates the extensive alterations performed in primary sequence without any negative effect on the folding of the polypeptide chain. Tolerance to substitutions within the beta sheet usually considered as rigid and inflexible and particularly of directly adjacent amino acids could not be a priori expected. Surprisingly, a modified ubiquitin variant could be obtained having 14 substitutions and a deletion while its original structure was maintained. Based on the total number of amino acids of ubiquitin, the number of modified residues corresponds to a percentage of about 20% (Rudolph *et al.*, 2011).

The amino acid residues at positions 2, 4, 6, 8, 62, 63, 64, 65, 66 in human ubiquitin are selected due to their surface-exposition and the tolerance of the overall structure to their randomization substitution. These positions are localized in spatial proximity to each other at the beginning of the first N-terminal β -sheet strand (positions 2, 4, and 6) or loop (position 8), as well as in the loop region (positions 62 and 63) or at the beginning of the C-terminal β -sheet strand (positions 64, 65 and 66). These positions form a contiguous region on the surface of ubiquitin. Thus, a novel binding property can be generated by modification of these residues in contiguous region, in which amino acid residues interact with their environment due to the charge, the spatial structure and the hydrophobicity/hydrophilicity of their side chains. The environment can be the solvent, generally water, or other molecules. This contiguous region can be named binding determining region (BDR) (Rudolph *et al.*, 2011).

1.4.3. Dimeric ubiquitin-based Affilin[®] technology

A strategy using dimeric ubiquitin-based scaffold is further developed to generate larger binding area and therefore select artificial binding proteins with higher affinities. With this strategy, a population of hetero-dimeric modified ubiquitin proteins, originating from monomeric modified ubiquitin proteins linked together in a head-to-tail arrangement by a linker, is created and selected against a certain target. The two binding determining regions created in the first and the second monomeric ubiquitin proteins respectively can bind a certain target together in a synergistic and combined manner. The advantage of this strategy is that while even more amino acid residues are modified, the protein chemical integrity is maintained without decreasing the overall stability of scaffold. Because less residues are modified in each monomeric ubiquitin molecule (Kunert *et al.*, 2011).

In this study, a dimeric ubiquitin-based Affilin[®] library was applied for selection. Two different DNA libraries encoding for monomeric modified ubiquitin proteins have been genetically fused to obtain DNA molecules encoding for the hetero-dimeric modified ubiquitin proteins. As shown in Fig. 1.6, the first library, SPW library contains 8 randomized residues in positions 2, 4, 6 and 62-66. The second library, SPF library has 7 randomized residues in positions 6', 8' and 62'-66'. The two modified monomeric ubiquitin proteins are connected with a linker "SGGSGGGIG" in head-to-tail format.

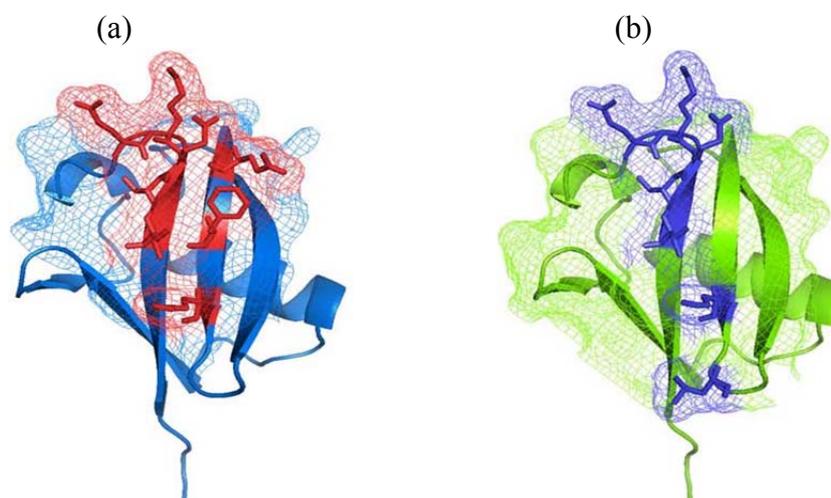


Figure 1.6 Structural views of the binding site of scaffold ubiquitin (PDB code 1UBQ). (a) Representation of ubiquitin molecule-SPW with 8 chosen residues (in red) for randomization. Framework is indicated in blue. (b) Representation of ubiquitin molecule-SPF with 7 chosen residues (in blue) for randomization. Framework is colored green. Figures were generated using software Pymol.

1.5. Selection of artificial binding proteins by Tat-phage display

In the past three decades, a variety of powerful selection technologies have been developed for selection of polypeptides with desired properties from a large collection of variants. The commonly used technologies are phage display, ribosome display, mRNA display and cell surface display (Smith, 1985; Mattheakis, 1994; Roberts and Szostak, 1997; Boder and Wittrup, 1997). All these selection technologies mimic three fundamental features of natural selection process in laboratory: (i) a diverse starting pool, (ii) a physical linkage between the phenotype and genotype, and (iii) a selection pressure that confers the protein variants with the desired traits a selection advantage, i.e. a functional selection of the fittest (Grimm, 2011).

The most commonly used selection technology is phage display technology based on the filamentous bacteriophage M13, with a monovalent approach using phagemid vector and helper phage (generally M13KO7) system. The filamentous bacteriophage M13 contains a single-stranded DNA (ssDNA) genome and is able to infect *E. coli* cells containing the F conjugative pilus. After infection, phage DNA is converted to and replicated as double stranded DNA. The phage coat is composed of five different coat proteins, of which protein III (pIII) is located at the tip of phage particle with 3-5 copies per phage particle and is the most commonly used coat protein as a fusion partner to display protein of interest (POI) (Russel, 1991; Karlsson *et al.*, 2003). Phagemid vectors carry the genes of pIII protein, POI and elements necessary for packing of the DNA into phage particles. The helper phages carry a complete set of phage proteins and a genetic mutation that reduces the efficiency of replication and packaging of the helper phage genome. Thus, the phage particles packed with phagemid genomes and statistically displaying monovalent POI are preferentially produced during phage propagation, and subsequently secreted from infected cells without cell lysis (Barbas, 2000).

A typical phage display selection experiment is illustrated in Fig. 1.7. A library gene pool encoding protein of interest (POI) is inserted into linearized phagemid DNA using recombinant DNA techniques and competent *E. coli* cells are transformed with successfully circularized vectors and then infected with helper phages, generating a library of phages displaying the POI. The phage library is exposed to immobilized target molecules, and the phages with appropriate specificity and affinity are captured, whereas the non-binding phages are washed out. Bound phages are eluted by conditions that disrupt the interaction between the displayed protein and the target molecule. Elute phages are then used to reinfect *E. coli* cells. The resulting amplified phage population is subsequently subjected for the following selection round. These steps are repeated, resulting in a phage population enriched in a limited number of variants with the desired binding affinity and specificity. After several rounds of bio-panning (generally 3-5), single phages may be selected and analyzed individually. The target-specific binding proteins can be identified by DNA sequencing analysis because a phenotype-genotype coupling occurs within each phage particle. These

clones may be subsequently subjected for characterization and other applications (Ruigrok *et al.*, 2011).

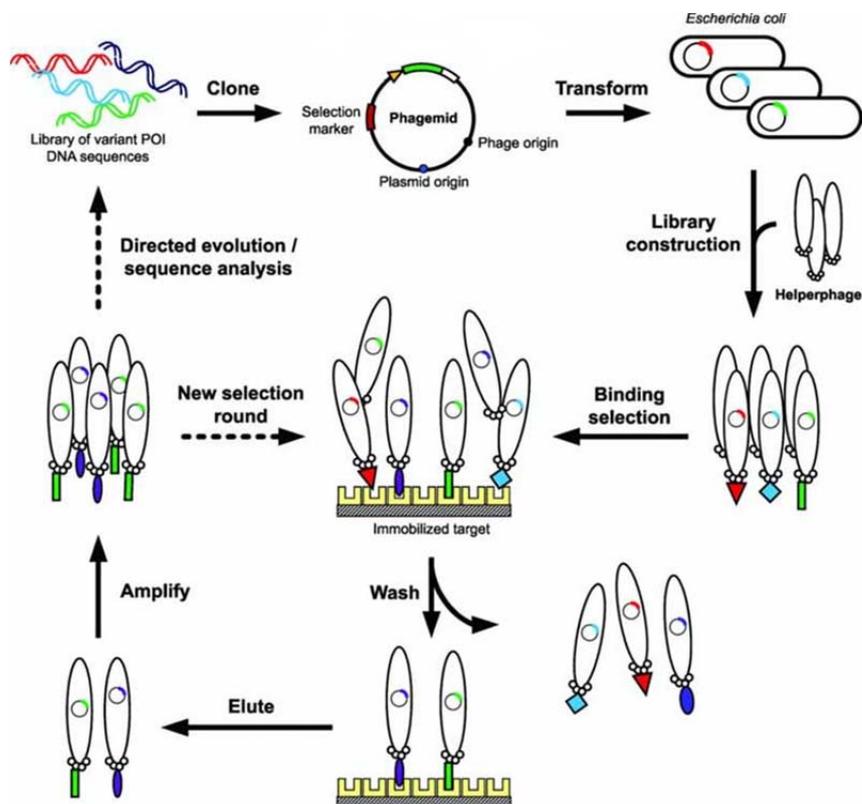


Figure 1.7 Schematic representation of monovalent phage display selection process. Figure adapted from Ruigrok *et al.*, 2011.

To ideally display protein of interest on phage particles, the protein of interest firstly need to be efficiently translocated to the bacterial envelop or periplasmic space. Three translocation pathways are therefore employed: the commonly used secretory (Sec) pathway translocates unfolded protein into periplasm. Considering that some fast-folding and stable proteins such as DARPins and thioredoxin that can't be efficiently translocated by traditional Sec pathway, alternative system is developed by using the signal recognition particle (SRP) pathway to direct the translocation of very stable and fast-folding proteins and scFv antibody fragments (Steiner *et al.*, 2006; Thie *et al.*, 2008). The third system utilizes the twin-arginine translocation (Tat)-mediated pathway that translocates proteins in their native and folded conformation. This Tat-based phage display technology is used for the display of proteins that fold in cytoplasm of *E. coli* cells (Fig. 1.8) (Paschke and Hohne, 2005; Paschke, 2006).

As ubiquitin is a fast-folding and stable protein, to efficiently display the dimeric ubiquitin-based Affilin[®] proteins, the Tat-mediated phage display technology is employed for selection in this study.

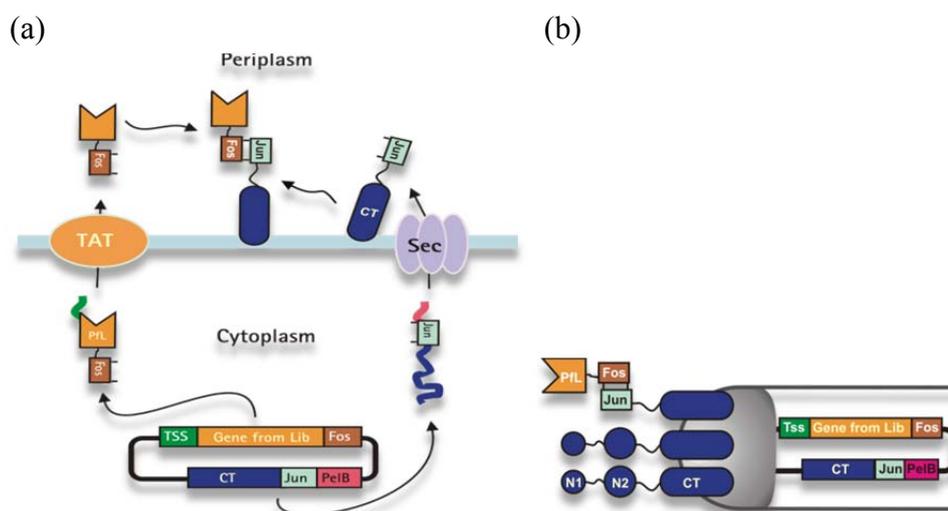


Figure 1.8 The principle of the Tat-mediated phage display (TDP) system. (a) Assembly of protein from library to M13 phage in TDP system. Two fusion proteins are produced. The first fusion protein consisting of folded protein from library (Pfl) with C-terminal Fos is translocated into periplasm by TSS (TorA Signal Sequence) via Tat pathway, while the second fusion protein consisting of truncated pIII protein (CT) with N-terminal Jun is directed by PelB leader peptide through Sec pathway. The two fusion proteins become covalently linked via the Jun/Fos interaction and the formation of two disulfide bridges in the periplasm. (b) A view of phage particle displaying Pfl. Figure adapted from Paschke and Hohne, 2005.

1.6. The N-terminal extracellular domain of PAC1-R as a target protein

1.6.1. G-protein coupled receptors

G-protein coupled receptors (GPCRs), known as seven-transmembrane domain receptors, comprise a protein superfamily of cell surface signaling proteins that transduce an enormous variety of extracellular signals into the intracellular effector pathways and ultimately activate cellular responses (Kilpatrick *et al.*, 1999; Pierce *et al.*, 1999; Chung *et al.*, 2008). The ligands that bind and activate GPCRs include light-sensitive compounds, ions, odors, pheromones, neurotransmitters and hormones, and vary in size from small molecules to peptides to large proteins (Bockaert and Pin, 1999).

GPCRs are integral membrane proteins that have a central common core consisting of seven transmembrane helices connected by three extracellular and three intracellular loops. The extracellular parts of the receptors can be glycosylated. These extracellular loops contain two highly conserved cysteine residues that form disulfide bonds to stabilize the receptor structure.

There are two principal signal transduction pathways involving the G-protein coupled receptors: the cAMP signal pathway and the Phosphatidylinositol signal pathway

(Gilman, 1987). When a ligand binds to a GPCR, it leads to a conformational change in the GPCR. The GPCR can sequentially activate an associated G protein by exchanging its bound GDP to a GTP. The α subunit of G protein, together with the bound GTP can then dissociate from the β and γ subunits to further affect intracellular signaling proteins or directly target functional proteins, depending on the type of α subunit (Wettschureck and Offermanns, 2005).

The superfamily of GPCRs includes about 900 members, which are only found in eukaryotes. GPCRs can be divided into at least five classes according to sequence homology and functional similarity (Attwood and Findlay, 1994; Foord *et al.*, 2005). Class A represents the largest group with about 700 members and is also termed the rhodopsin family. The GPCRs of class B have 15 members that structurally bind endogenous peptide hormones, including calcitonin, secretin, glucagon, the incretins glucagon-like peptide 1 (GLP-1), glucose-dependent insulinotropic polypeptide (GIP), parathyroid hormone (PTH), corticotropin-releasing factor (CRF), growth-hormone releasing factor (GRF), vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase activating polypeptide (PACAP). Structurally, Class B GPCRs have low homology with other families, but are highly conserved within the family. They all contain a relatively large N-terminal extracellular domain (ECD) that is critical for ligand binding. The N-terminal EC domains have some common features, typically comprising six conserved cysteine residues, two conserved tryptophan residues and an aspartate residue which has been proposed to be critical for ligand binding (Gaudin *et al.*, 1995; Laburthe *et al.*, 2002). Many peptide ligands of class B GPCRs are related in sequence and can bind to more than one receptor subtype (Vaudry *et al.*, 2000).

GPCRs are involved in many physiological functions and in multiple diseases, including the development of cancer and cancer metastasis. GPCRs are also the target of approximately 30% of all medicinal drugs (Overington *et al.*, 2006). Currently, the pharmacological manipulation of many GPCRs provides an excellent option to neutralize tumorigenic signals, promising the development of new GPCR-based therapeutics for cancer.

1.6.2. The N-terminal extracellular domain of PAC1-R

Pituitary adenylate cyclase-activating polypeptide (PACAP) is a 38-amino acid C-terminally α -amidated peptide that was first isolated from an ovine hypothalamic extract on the basis of its ability to stimulate cAMP formation in anterior pituitary cells (Miyata *et al.*, 1989). PACAP belongs to the vasoactive intestinal polypeptide (VIP)-secretin-growth hormone releasing hormone-glucagon superfamily. PACAP is widely distributed in the central and peripheral nervous system. Mature PACAP has two biologically active forms, PACAP-38 and PACAP-27. The N-terminal portion of PACAP is involved in interaction with the transmembrane domain and activating the receptor, whereas the remainder of the peptide is important for high affinity binding and receptor specificity (Laburthe and Couvineau, 2002).

The pituitary adenylate cyclase-activating polypeptide receptor (PAC-R) is a class B G-protein coupled receptor. Two subtypes of PAC-R have been characterized on the basis of their relative affinities for ligands PACAP and VIP. The subtype I receptor (PAC1-R) exhibits high affinity for PACAP ($K_D \approx 0.5$ nM) and much lower affinity for VIP ($K_D > 500$ nM) (Cauvin *et al.*, 1990; Gottschall *et al.*, 1990, 1991; Lam *et al.*, 1990; Suda *et al.*, 1992). By contrast, the subtype II receptor (e.g., VPAC1-R and VPAC2-R) possesses similar affinity for both PACAP and VIP ($K_D \approx 1$ nM) (Gottschall *et al.*, 1990; Lam *et al.*, 1990).

Three splice variants in the N-terminal ECD have been identified for PAC1-R with different binding affinity with ligands PACAP and VIP, as well as differential coupling to adenylate cyclase and phospholipase C pathways. The first full length PAC1-R normal (PAC1-Rn) encodes the entire N-terminus, whereas the second variant PAC1-R short (PAC1-Rs) is deleted by 21 amino acids (residues 89-109). The third variant, named PAC1-R very short (PAC1-Rvs) is deleted by 57 amino acids (residues 53-109) (Fig. 1.9).

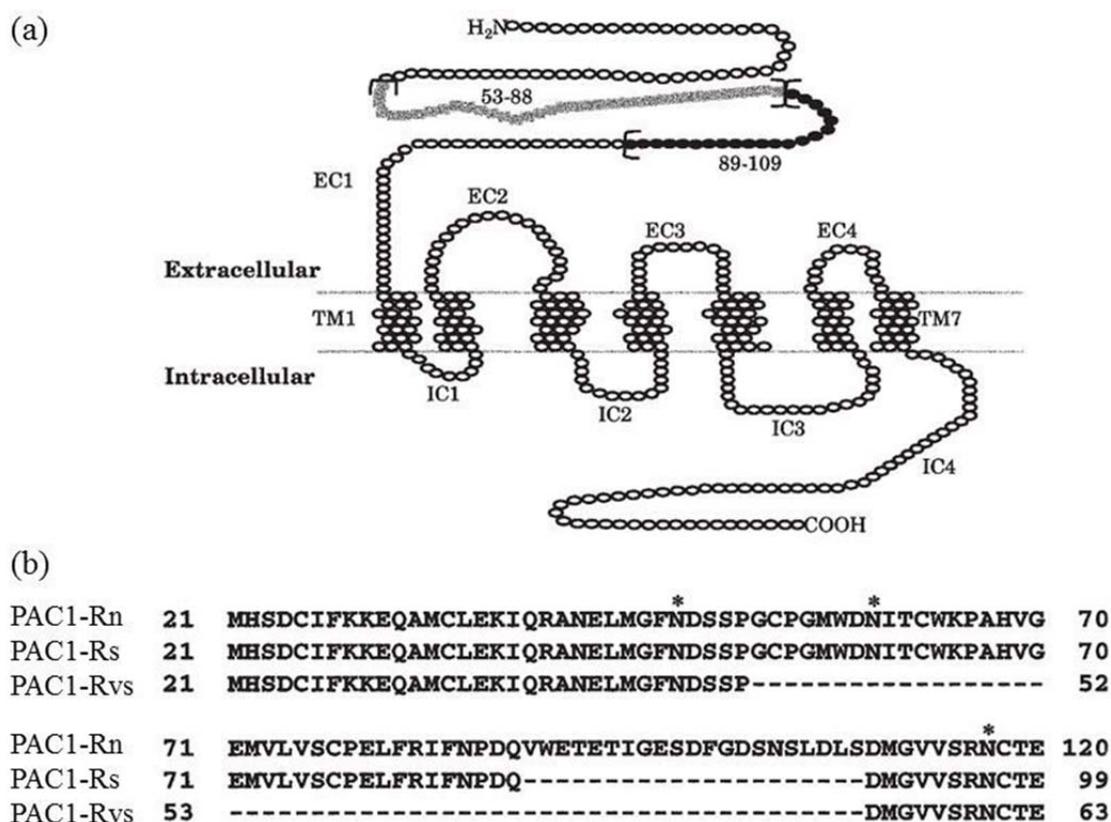


Figure 1.9 A two-dimensional model of PAC1-R (a) and sequence alignment of EC1 domain of PAC1-Rn, PAC1-Rs and PAC1-Rvs (b). (a) Residues 53-88, which are deleted only in PAC1-Rvs presented as squares. Residues 89-109, which are deleted both in PAC1-Rs and PAC1-Rvs, respectively, are shown as filled circles. (b) The deleted amino acids in PAC1-Rs and PAC1-Rvs are represented as dashes. N-glycosylation sites are indicated as asterisk. Figure modified from Dautzenberg *et al.*, 1999.

The binding and cAMP studies with transfected human embryonic kidney 293 (HEK293) revealed significant differences in the affinities and selectivities towards PACAP38, PACAP27 and VIP. PAC1-Rn binds to PACAP38 and PACAP27 with affinities in the low nanomolar range, and to VIP with up to 400-fold lower affinity. PAC1-Rvs binds to three ligands with respectively 100-fold lower affinities than PAC1-Rn. PAC1-Rs can unselectively bind all three ligands with high affinity. These data indicate that residues 53-88 within the N-terminal domain of PAC1-R are important for high affinity ligand binding, whereas residues 89-109 determine the receptor's ligand selectivity (Dautzenberg *et al.*, 1999).

Currently there is no experimentally determined structure of full length class B GPCRs available. Nevertheless, several N-terminal ECD structures of Class B GPCRs have recently been determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy towards a understanding for their structure and function. They are the murine type-2b CRF receptor (CRF-R2b) (Grace *et al.*, 2007), the human type-1 CRF receptor (CRF-R1) (Pioszak *et al.*, 2008), human GIP receptor (GIP-R) (Parthier *et al.*, 2007), human GLP-1 receptor (GLP-1R) (Runge *et al.*, 2008), human PTH type-1 receptor (PTH1-R) (Pioszak *et al.*, 2008) and human PACAP type-1 receptor (PAC1-Rs) (Sun *et al.*, 2007). These structures reveal a common form for the class B GPCR ECDs – an N-terminal helix and four β -strands forming two antiparallel sheets. These secondary structure elements are stabilized by three intermolecular disulfide bonds, forming a highly conserved fold termed “secretin family recognition fold” (Fig 1.10 a-b). Interestingly, the NMR structure of PAC1-Rs ECD exhibits a difference in loop 4 topology (Fig. 1.10c), which proceeds to β 4 “above” the terminal disulfide bond (in all other class B GPCR ECDs, this loop proceeds “below” this disulfide), resulting in an inverse direction of loop 4 compared to other ECDs (Parthier *et al.*, 2009).

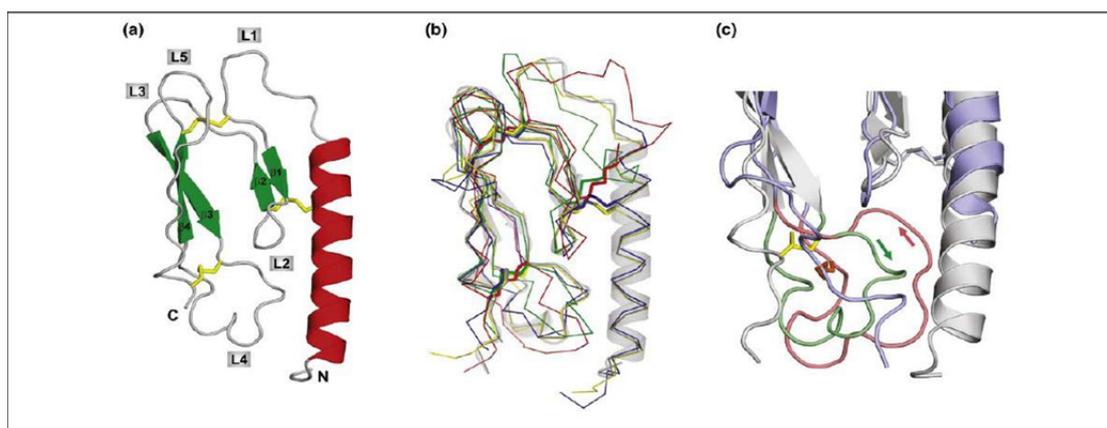


Figure 1.10 Structures of class B GPCR ECDs. Each ECD structure solved to date exhibits the “secretin family recognition fold”. (a) Common structural elements of class B GPCR ECDs as observed in the GIP-R ECD: an N-terminal α -helix (red), two β -sheets composed of strands β 1 to β 4 (green), with loop regions L1 to L5 (grey). The domains are stabilized by three conserved

disulfide bridges (yellow sticks). (b) Superposition of the polypeptide backbones of CRF-R2b ECD (red; PDB code: 2JND), CRF-R1 ECD (green; PDB code: 3EHU), GIP-R ECD (grey, in cartoon representation; PDB code: 2QKH), GLP-1R ECD (yellow; PDB code: 3C5T) and PTH1-R ECD (blue, PDB code: 3C4 M). (c) The aberrant topology of loop 4 in PAC1-Rs ECD (pink and lilac; PDB code: 2JOD) superimposed on the GIP-R ECD (light green and grey). The direction of the main chain of loop 4 (indicated by arrows) in PAC1-Rs ECD is opposite to that in GIP-R ECD and the other ECDs. Figure adapted from Parthier *et al.*, 2009.

More recently, a refined crystal structure of maltose binding protein (MBP)-fusion expressed PAC1-Rs ECD has been determined. This structure is very similar to all other class B GPCR ECDs, for example to the VIP2R structure (Fig. 1.11a). By contrast, the X-ray and previous NMR structures of PAC1-Rs ECD show differences in loop 4 region (Fig. 1.11b). In particular, the residue Pro78 is positioned very differently. Pro78 is a highly invariant residue among class B GPCRs. Pro78 and its corresponding residues in the family play important role in the structure stability of ECD (Karaplis *et al.*, 1998; Runge *et al.*, 2008). In the X-ray structure of PAC1-Rs ECD, Pro78 fills the hydrophobic cavity formed by Glu30, Ile61, Thr62, Leu80 and Phe81, showing the expected structural similarity to other class B GPCRs but quit different situation comparing to NMR structure of PAC1-Rs ECD (Kumar *et al.*, 2011).

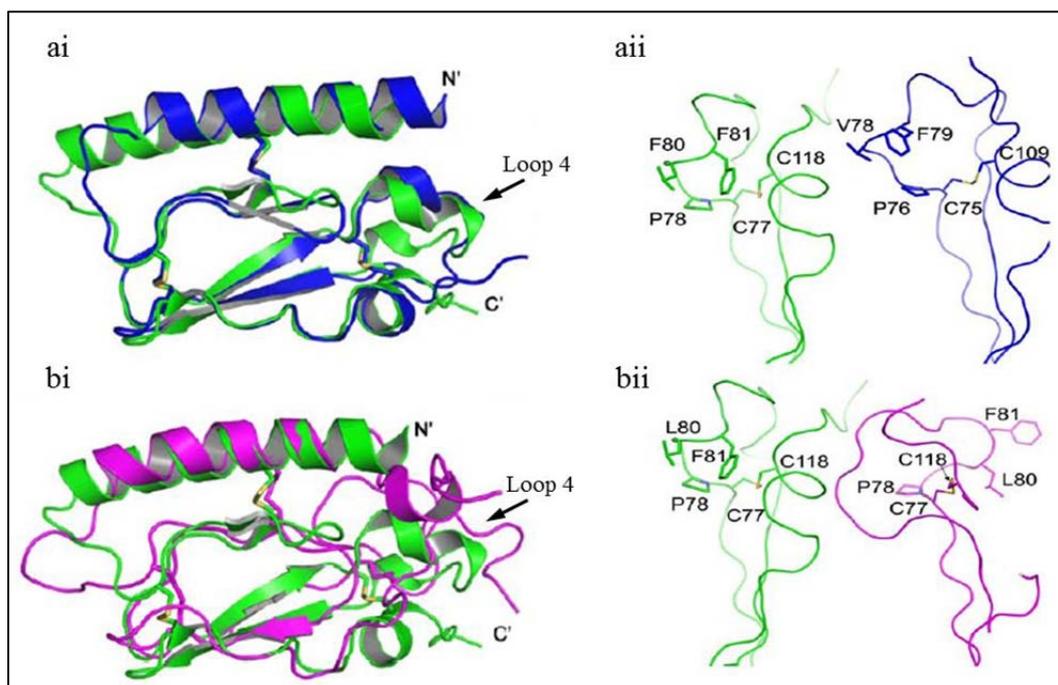


Figure 1.11 Structure superimposition of the PAC1-Rs ECD/VIP2-R ECD and X-ray/NMR structures of the PAC1-Rs ECD. (a) Structure superimposition of PAC1-Rs (green) and VIP2-R (blue) depicts the expected similarity in (i) the backbone and (ii) the position of the conserved residues. (b) (i) Structure superimposition of the X-ray (green) and NMR (magenta) structures of the PAC1-Rs ECD in the backbone. The two molecules were laterally separated (ii) the close up

of selected residues in the two structures. The loop 4 indicated by arrow is very different between the X-ray and NMR structures. Figure modified from Kumar *et al.*, 2011.

PACAP and its receptor PAC1-R have been implicated to play important roles in several cellular processes, including regulation of circadian rhythm, control of food intake, glucose metabolism, learning and memory, neuronal ontogenesis, apoptosis, immune system regulation and post-traumatic stress disorder (Ressler *et al.*, 2011). PAC1-R has therefore been proposed as a potential drug target for treatment of epilepsy, neurodegeneration and cognition disorders (e.g., Alzheimer's disease, schizophrenia and anxiety).

1.7. Project aims

The intrinsic properties of ubiquitin, for example the small size, remarkable stability and high solubility allow it to be a novel scaffold protein. Dimeric ubiquitin-based Affilin[®] technology has been successfully developed to generate target-specific artificial binding proteins with high affinity against various targets. We therefore aimed to validate efficient selection of Affilin[®] binding proteins against the target, N-terminal extracellular domain of human PAC1 receptor (nPAC1-Rs), and further characterize Affilin[®] binding proteins for the structure determination of PAC1 receptor by cocrystallization with nPAC1-Rs and full length PAC1 receptor. The detailed aims of this thesis include as following:

1. Production of functional and native nPAC1-Rs target protein for selection and characterization
2. Selection with dimeric ubiquitin-based Affilin[®] library against target protein nPAC1-Rs
3. Screening of nPAC1-Rs-specific Affilin[®] binding proteins by single phage ELISA and high throughput Hit-ELISA
4. Characterization of selected Affilin[®] binding proteins for structure determination of nPAC1-Rs and full length PAC1 receptor

2. Materials and methods

2.1. Materials

2.1.1. Chemicals

Chemical	Supplier
Acetone	Carl Roth GmbH & Co.
Agarose LE	Biozym
Agarose LE GP	Biozym
Agarose Sieve 3:1	Biozym
Agarose Sieve GP	Biozym
Bromophenol blue	AppliChem
Ethylene diamine tetraacetic acid (EDTA)	Carl Roth GmbH & Co.
Ethanol purest and technical	Carl Roth GmbH & Co.
Formaldehyde 37%	AppliChem
Glacial acetic acid	Carl Roth GmbH & Co.
D(+)-Glucose	Carl Roth GmbH & Co.
Glycerol 99.5%	AppliChem
Glycine	AppliChem
Hydrochloric acid	Carl Roth GmbH & Co.
Hydroxyethyl-piperazine-1-ethanesulfonic acid	AppliChem
Imidazole	E. Merck KGaA
Lactose monohydrate	Carl Roth GmbH & Co.
Magnesium chloride hexahydrate	AppliChem
Magnesium sulfate heptahydrate	AppliChem
β -Mercaptoethanol	AppliChem
2-(N-morpholino) ethanesulfonic acid (MES)	AppliChem
Phenylmethanesulfonylfluorid (PMSF)	AppliChem
Polyethylene glycol 6000 (PEG 6000)	Carl Roth GmbH & Co.
Potassium dihydrogen phosphate	AppliChem
Silver nitrate	AppliChem
Sodium carbonate	Carl Roth GmbH & Co.
Sodium chloride	AppliChem
Sodium dihydrogen phosphate monohydrate	Carl Roth GmbH & Co.
Sodium dodecyl sulfate (SDS)	AppliChem
di-Sodium hydrogen phosphate	AppliChem
Sodium hydroxide	Carl Roth GmbH & Co.
Sulfuric acid	Carl Roth GmbH & Co.
Trichloroacetic acid (TCA)	AppliChem

Chemical	Supplier
Tris	AppliChem
Tween-20	AppliChem
Urea	Carl Roth GmbH & Co.

2.1.2. Bacterial strains

<i>E. coli</i> strain	Genotype	Supplier
BL21(DE3)	<i>F⁻ ompT gal dcm lon hsdS_B(r_B⁻ m_B⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])</i>	Stratagene
DH5α	<i>F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K⁻ m_K⁺), λ-</i>	Department of Genetics (MLU)
ER2738	<i>F['] proA⁺ B⁺ lacI^q Δ(lacZ)M15 zzf::Tn10(Tet^R)/ fhuA2 glnV Δ(lac-proAB) thi-1 Δ(hsdS-mcrB)5</i>	NEB Lucigen
NovaBlue(DE3)	<i>endA1 hsdR17(rK12- mK12+) supE44 thi-1 recA1 gyrA96 relA1 lac (DE3)F'[proA+B+ lacI qZΔM15::Tn10] (TetR)</i>	Novagen
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F['] proAB lacIqZΔM15 Tn10 (Tetr)]</i>	Stratagene

2.1.3. Helper phage

Phage	Description	Supplier
M13KO7 Helper Phage	An M13 derivative which carries the mutation Met40Ile in gene II, with the origin of replication from P15A and the kanamycin resistance gene	Invitrogen

2.1.4. Plasmids

Plasmid	Description	Origin/Reference
pCD87SA	phagemid vector utilizing Tat translocation pathway	Dr. Matthias Paschke
pCD87SA-SPWF	pCD87SA phagemid containing inserted SPWF fragment encoding Affilin [®] variant	Scil Proteins GmbH
pET SUMO	N-terminal SUMO (small ubiquitin-like Modifier) fusion protein for increased expression and solubility of recombinant proteins and generation of native protein	Invitrogen
pTrS/nPAC1-Rs	pET SUMO derivative encoding Thioredoxin-SUMO-nPAC1-Rs fusion protein	NWG "Künstliche Bindeproteine"(KBP)

Plasmid	Description	Origin/Reference
pET23d(+)	Bacterial expression vector containing C-terminal 6× His tag	Novagen
pET23dk	pET23d(+) derivative with kanamycin resistance	NWG KBP
pET23dk-SPWF	pET23d vector containing inserted SPWF fragment encoding Affilin [®] variant	NWG KBP

2.1.5. Oligonucleotides

All the DNA oligonucleotides were synthesized by Thermo Scientific Biopolymers (Ulm, Germany).

Oligonucleotide	Sequence (5' to 3')	comment
Ptetop	ACCACTCCCTATCAGTGATAG	Sequencing primer for phagemid pCD87SA
T7_P	TAATACGACTCACTATAGGG	Sequencing primer for pET23dk vector
87SA_sub_fw	CACAACCTCGGCGGCTTAACC	Forward and reverse primers for subcloning of SPWF fragments from pCD87SA into pET23dk
SPF-rev-Xho	TGCAGCCATCTCGAGACCACCA CGTAAACGAAGAACTAAATGT	
SPW-23dk_sense	GTCTGCGTGGCGGTCTCGAGCA CCACC	Sense and antisense primers for construction of SPW monomer ubiquitin variant
SPW-23dk_antisense	GGTGGTGCTCGAGACCGCCAC GCAGAC	
SPF-23dk_sense	GAAGGAGATATACCATGGCCAT GCAAATTTTGTGACACG	Sense and antisense primers for construction of SPF monomer ubiquitin variant
SPF-23dk_antisense	CGTGTCACAAAAATTTGCATG GCCATGGTATATCTCCTTC	

2.1.6. Enzymes

Enzyme	Supplier
Alkaline phosphatase, Calf Intestinal (CIP)	NEB Lucigen
Benzonase nuclease	Novagen
FastDigest <i>NcoI</i> and <i>XhoI</i>	Fermentas
Lysozyme	AppliChem
Pfu DNA polymerase	Fermentas
Phusion [®] high-fidelity DNA polymerase	Finnzymes
SUMO protease	NWG KBP
T4 DNA ligase	Fermentas

2.1.7. Proteins and antibodies

Protein	Supplier
Bovine serum albumin (BSA)	Sigma-Aldrich
Human Serum	Sigma-Aldrich
N-terminal domain of parathyroid hormone type1 receptor (nPTH1-R)	NWG KBP
PACAP6-38	Bachem

Antibody	Detection protein	Source	Working concentration	Supplier
anti-M13/HRP	M13 phage major coat protein PVIII	Mouse Monoclonal	1:5000	GE Healthcare
anti-Ubi-Fab/POD	Monomer and dimer Ubiquitin	HuCAL [®]	1:6500	Morphosys

2.1.8. Standards and kits

Standard/Kit	Supplier
EZ-Link Sulfo-NHS-LC-Biotin	Pierce Protein
GeneRuler [™] 100 bp DNA ladder	Fermentas
GeneRuler [™] 1 kb DNA ladder	Fermentas
PageRuler [™] pre-stained protein ladder	Fermentas
PageRuler [™] unstained protein ladder	Fermentas
PageBlue [™] protein staining solution	Fermentas
PureYield [™] Plasmid Maxiprep system	Promega
SYPRO Orange protein gel stain	Sigma-Aldrich
TMB Plus Substrate Solution	Kem-En-Tec
Wizard [®] Plus SV Minipreps DNA Purification Kit	Promega
Wizard [®] SV Gel and PCR Clean-Up Kit	Promega
6× DNA Loading Dye	Fermentas
10× HBS-EP+	GE Healthcare
10× Sigma Blocking buffer	Sigma-Aldrich

2.1.9. Buffers and solutions

Buffer	Components
Buffer E	200 mM Glycine, 2.7 mM KCl, 8 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄ , adjust to pH 2.2 with HCl
Cell lysis buffer	500 mM NaCl, 50 mM NaH ₂ PO ₄ , 10 mM Imidazole, 2.5 mM MgSO ₄ , pH 8.0
Neutralizer	1 M Tris, adjust to pH 9.1 with HCl

Chapter 2: Materials and methods

Buffer	Components
PBS	137 mM NaCl, 2.7 mM KCl, 8 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄ , pH 7.4
PBS-T	PBS containing 0.1% (v/v) Tween-20
1× MES running buffer	50 mM MES, 50 mM Tris, 1 mM EDTA, 0.1% (w/v) SDS, pH 7.3
PEG/NaCl solution	20% (w/v) PEG 6000, 2.5 M NaCl
Resuspension buffer	8 M Urea, 100 mM Tris, 1 mM EDTA, pH 8.0
TAE buffer	20 mM Tris-HCl, 40 mM acetic acid, 2 mM EDTA, pH 8.5
5× SDS-PAGE sample buffer	250 mM Tris-HCl, 5% (w/v) SDS, 50% Glycerol, 0.05% (w/v) Bromophenol blue, 5% (v/v) β-Mercaptoethanol, pH 8.0

Chromatography buffer	Components
NPI-20	150 mM NaCl, 50 mM NaH ₂ PO ₄ , 20 mM Imidazole, pH 8.0
NPI-50	500 mM NaCl, 50 mM NaH ₂ PO ₄ , 50 mM Imidazole, pH 8.0
NPI-500	150 mM NaCl, 50 mM NaH ₂ PO ₄ , 500 mM Imidazole, pH 8.0
Desalting buffer	150 mM NaCl, 25 mM HEPES, 20 mM Imidazole, pH 8.0
Gel filtration buffer	150 mM NaCl, 50 mM NaH ₂ PO ₄ , pH 8.0
HEPES buffer	150 mM NaCl, 25 mM HEPES, 1 mM EDTA, pH 7.4

Silver staining solution	Components
Fixing solution	60 ml 50% Acetone, 1.5 ml 50% TCA, 25 μl Formaldehyde (37%)
Pretreatment I solution	60 ml 50 % Acetone
Pretreatment II solution	60 ml ddH ₂ O, 100 μl 10% Na ₂ S ₂ O ₃ (Sodium thiosulfate) solution
Impregnation solution	60 ml ddH ₂ O, 800 μl AgNO ₃ (20% solution), 600 μl Formaldehyde
Developing solution	60 ml ddH ₂ O, 1.2 g Na ₂ CO ₃ , 25 μl Formaldehyde, 25 μl 10 % Na ₂ S ₂ O ₃
Stopping solution	60 ml 1 % Glacial acetic acid

2.1.10. Medium and antibiotics

All components of medium for cultivation of *E. coli* were purchased from Becton Dickinson (Heidelberg, Germany). The antibiotics were supplied by Carl Roth GmbH & Co.

Chapter 2: Materials and methods

Medium	Components
Autoinduction medium ZYM-5052	1% N-Z-amine, 0.5% Yeast Extract, 25 mM Na ₂ HPO ₄ , 25 mM KH ₂ PO ₄ , 50 mM Na ₂ SO ₄ , 0.5% Glycerol, 0.05% Glucose, 0.2% Lactose, add distilled H ₂ O to a final volume, autoclave (Studier, 2005)
LB medium	10 g Tryptone, 5 g Yeast Extract, 5 g NaCl, add distilled H ₂ O to final 1L, autoclave
LB agar	LB medium, 15 g Agar, add distilled to final 1L, autoclave
SOBCG agar	20 g Tryptone, 5 g Yeast Extract, 0.5 g NaCl, 15 g Agar, add distilled to final 940 ml, autoclave, cool to 50-60 °C, add further components to final concentration of 10 mM MgCl ₂ , 2 % Glucose and 30 µg/ml Chloramphenicol
2YT medium	17 g Tryptone, 10 g Yeast Extract, 5 g NaCl, add distilled H ₂ O to final 1L, autoclave
2YT agar	2YT medium, 15 g Agar, add distilled H ₂ O to final 1L, autoclave

Antibiotic	Stock	Working concentration
Chloramphenicol	30 mg/ml in Ethanol	30 µg/ml
Kanamycin	50 mg/ml in H ₂ O	50 µg/ml
Tetracycline	5 mg/ml in Ethanol	0.1 µg/ml

2.1.11. Chromatography columns

Column	Supplier
HiLoad16/60 Superdex 75 prep grade column	GE Healthcare
HiPrep 26/10 desalting column	GE Healthcare
HisTrap HP 1 ml column	GE Healthcare
HisTrap HP 5 ml column	GE Healthcare

2.1.12. Laboratory equipment

Equipment	Supplier
ÄKTA explorer system	GE Healthcare
ÄKTA xpress system	GE Healthcare
Allegra X-15R centrifuge with rotors FX6100 and SX4750A	Beckman Coulter
Avanti J-26 XP centrifuge with rotors JA-25.50, JLA-16.250, JLA-8.1000 and JS-5.3	Beckman Coulter
Balance SI-234, SI-2002 and MXX-412	Denver Instrument
Biacore T100 instrument	GE Healthcare
BIOMEK 3000	Beckman Coulter
BIOMEK FX Laboratory Automation Workstation	Beckman Coulter

Equipment	Supplier
BioTek ELx405 Microplate Washer	Beckman Coulter
Bio-Vision-3000 Gel Documentation	Vilber Lourmat
DU-730 UV/Vis Spectrophotometer	Beckman Coulter
ECM 630 Electroporator	BTX
Forma-86 C ULT Freezer	Thermo Scientific
Galaxy MiniStar table centrifuge	VWR
HAAKE N3 Circulating Bath	HAAKE
Heraeus Fresco 21 and Pico 17 centrifuge	Thermo Scientific
Hera Safe KS 12 Safety Cabinet	Thermo Scientific
Innova 42, 44 and 4230 Incubator Shaker	New Brunswick Scientific
InoLab PH 720 pH meter	WTW
J-810 Circular Dichroism Spectropolarimeter	Jasco
LightCycler 480 II	Roche
Memmert Incubator INB 400	Memmert
MR Hei-Standard Magnetic Stirrer	Heidolph
MTP Incubator TH15	Edmund Bühler GmbH
Multimode Detector DTX880	Beckman Coulter
PARADIGM detection platform	Beckman Coulter
Peristaltic Pumps 505S	Watson Marlow
QPix 2 Colony Picker	Genetix
Scotsman AF80 Ice Flaker	Scotsman
SENSQUEST Lab Cycler	SENSQUEST
SRT9 Roller Mixer	Stuart
Systec V-75A Autoclave	Systec
Thermomixer Comfort	Eppendorf
Unimax 1010 Platform Shaker	Heidolph
Vibra Cell VC 750 Ultrasonic processor	Sonics
Vortex Genie 2	Scientific Industries
VP-ITC Isothermal Titration Calorimeter	MicroCal
X cell SureLock Electrophoresis Cell	Invitrogen

2.1.13. Miscellaneous

Miscellaneous	Supplier
Cellulose Nitrate Filter 0.45 μm	Sartorius Stedim
Dynabeads [®] MyOne [™] Streptavidin C1	Invitrogen
Dynabeads [®] MyOne [™] Streptavidin M-270	Invitrogen
Electroporation Cuvette 1 mm gap	BTX
Enzyme-linked immunosorbent assay (ELISA) Plate	Nunc
NuPAGE 4-12% Bis-Tris Gel 1.0 mm	Invitrogen
Reacti-Bind NeutrAvidin Coated strip Plate	Pierce protein
Q-Tray	Genetix
Sartorius Vivaflow 50	Sartorius Stedim

Miscellaneous	Supplier
Sensor Chip SA	GE Healthcare
Slide-A-Lyzer MWCO 3500 Dialysis Cassette	Thermo Scientific
Spectra/Por 7 Tubing MWCO 3000	Spectrum Laboratories
Syringe Filter 0.22 and 0.45 μm	TPP
96-well Microtiter Plate, PP	Greiner bio-one
96-well Microtiter Plate, PS	Greiner bio-one
96-well Deep Microtiter Plate, PP	Greiner bio-one

2.2. Methods

2.2.1. General molecular biology methods

2.2.1.1. Plasmid DNA preparation

The plasmid DNA preparations were performed either with Promega Wizard[®] Plus SV Minipreps DNA Purification Kit for 5 ml of transformed *E. coli* bacterial cell cultures or PureYield[™] Plasmid Maxiprep kit for 100 ml of cell cultures according to the manufacturer's protocols.

2.2.1.2. Agarose gel electrophoresis

The agarose gel electrophoresis was used for analysis and separation of DNA fragments. The DNA samples mixed with DNA loading dye were loaded on agarose gel prepared from four kinds of agarose respectively, depending on different DNA length and purpose. The 2% (w/v) Sieve 3:1 agarose and Sieve GP agarose were used for analysis and recovery of DNA fragments less than 1 kb, respectively. The 1% (w/v) LE agarose and LE GP agarose were utilized for analysis and recovery of the fragments longer than 1 kb, respectively. The running was conducted at 8 V/cm with 1 \times TAE buffer. The gel was stained for 20 min in 2 $\mu\text{g/ml}$ ethidium bromide solution. BIO-Vision-3000 was equipped for imaging and documentation.

2.2.1.3. DNA recovery from agarose gel

The DNA fragments were recovered from gel slices or PCR products with Promega Wizard[®] SV Gel and PCR Clean-Up Kit according to the manufacturer's protocol.

2.2.1.4. DNA concentration measurement

The DNA concentration and purity was determined by measuring the absorbance at 260 nm and 280 nm by DU-730 Life Science UV/Vis Spectrophotometer. The DNA

concentration was calculated using following equation:

$$\text{Concentration } (\mu\text{g/ml}) = A_{260} \times \text{Absorbance factor} \times \text{dilution factor} \quad \text{Equation 1}$$

A_{260} : absorbance at 260 nm, Absorbance factor of dsDNA: 50 $\mu\text{g/ml}$

The DNA purity can be estimated by performing ratio absorbance measurement at A_{260}/A_{280} . As a general rule any preparation with an A_{260}/A_{280} between 1.7 and 2.0 is looked as pure.

2.2.1.5. Polymerase chain reaction (PCR)

For the subcloning of SPWF fragments from pCD87SA phagemid vector into expression vector pET23dk, the PCR was prepared as following protocol:

Component	volume	final concentration
5× HF buffer	10 μl	1×
dNTPs, 10 mM each	1 μl	0.2 mM each
DNA template,	variable	100 ng
87SA_sub_fw	1 μl	0.2 μM
SPF_rev_XhoI	1 μl	0.2 μM
DMSO	1.5 μl	3%
Phusion DNA polymerase	0.5 μl	1 unit
Nuclear-free water	to 50 μl	

PCR was performed in a volume of 50 μl using Phusion High-fidelity DNA polymerase, including a denaturation step (180 sec at 98 °C) and 25 cycles (30 sec at 98 °C, 30 sec at 55 °C, and 15 sec at 72 °C) followed by a final elongation step (300 sec at 72 °C). The PCR products were purified by using Wizard[®] SV Gel and PCR Clean-Up Kit (2.2.1.3), and subsequently applied for *NcoI/XhoI* double digestion.

In addition, PCR was employed to construct the plasmid containing single domain ubiquitin variants (either SPW-domain variants or SPF-domain variants) with QuikChange mutagenesis strategy. The protocol was shown as blow:

Component	volume	final concentration
10× <i>Pfu</i> buffer with Mg^{2+}	5 μl	1×
dNTPs, 10 mM each	1 μl	0.2 mM each
DNA template,	variable	50 ng
Sense primer	1 μl	0.2 μM
Antisense primer	1 μl	0.2 μM
<i>Pfu</i> DNA polymerase	1 μl	2.5 units
Nuclear-free water	to 50 μl	

The PCR program included a denaturation step (30 sec at 95 °C) and 18 repeated cycles (denaturation for 30 sec at 95 °C, primer annealing for 60 sec at 53 °C, and

DNA synthesis for 600 sec at 72 °C) followed by a final elongation step (900 sec at 72 °C). The PCR products were used for *DpnI* digestion.

2.2.1.6. DNA digestion

The DNA digestion with restriction endonucleases was performed under the conditions recommended by manufacturer. The preparative digestion was carried out at 37 °C for 60 min using 0.4 µg of amplified SPWF fragments or 2.5 µg plasmid DNA and 2 FDU (FastDigest[®] Unit, Fermentas) *NcoI* as well as 2 FDU *XhoI* in 50 µl reaction volume. FastDigest enzymes were deactivated by followed thermal inactivation step at 80 °C for 10 min. The digested DNA fragments were separated by agarose gel electrophoresis (2.2.1.2, page 33) and purified by Wizard[®] SV Gel and PCR Clean-Up Kit (2.2.1.3, page 33).

DpnI was used to digest the methylated template DNA in QuikChange PCR reaction. 0.5 µl *DpnI* (10 U/µl, Fermentas) was directly added into 50 µl PCR reactions, incubated at 37 °C for 60 min, and deactivated at 80 °C for 20 min. The DNA samples were purified by Wizard[®] SV Gel and PCR Clean-Up Kit (2.2.1.3, page 33) and used for transformation.

2.2.1.7. Dephosphorylation of vector DNA

Alkaline phosphatase, Calf Intestinal (CIP) catalyzes the removal of 5' phosphate groups from DNA and RNA. Since CIP-treated fragments lack 5' phosphoryl termini required by ligases, they can't self-ligate. Therefore dephosphorylation can be used to decrease the vector' self-ligation background in cloning strategies. 5 µg of *NcoI/XhoI* double digested vector DNA pET23dk were dephosphorylated by incubating at 37 °C for 60 min with 5 U CIP (NEB) in 50 µl reaction. CIP was deactivated incompletely at 65°C for 20 min, therefore the vector DNA was further purified by Wizard[®] SV Gel and PCR Clean-Up Kit (2.2.1.3, page 33).

2.2.1.8. DNA ligation

A 3:1 molar ratio of insert: vector was used for all ligation reactions. 25 ng of linear vector DNA pET23dk and 10 ng digested SPWF fragment were incubated overnight at 16 °C in 10 µl reaction, in presence of 0.5 mM ATP and 2.5 U T4 DNA ligase (5U/µl, Fermentas). T4 DNA ligase was heat inactivated at 65 °C for 10 min. The ligation products were purified by Wizard[®] SV Gel and PCR Clean-Up Kit (2.2.1.3, page 33) and used for transformation.

2.2.1.9. Transformation with chemically competent *E. coli* cells

For preparation of chemically competent *E. coli* cells, 5 ml of LB medium was

inoculated with a single colony picked from a fresh LB-agar plate and shaken at 37 °C overnight. The overnight culture was inoculated into 300 ml LB medium at a ratio of 1:100 and grown at 37 °C with shaking at 200 rpm until OD₆₀₀ reached 0.4-0.5. The cells were transferred into 6 Falcon tubes (50 ml each) and centrifuged at 2400 x g for 10 min at 4 °C after placing on ice for 30 min. All cell pellets were resuspended gently with 300 ml of ice-cold sterile 0.1 M CaCl₂. After centrifugation at 2400 x g for 10 min at 4 °C, the cell pellets were resuspended with 10 ml of ice-cold sterile 0.1 M CaCl₂ solution and placed on ice for 30 min. The competent cells were finally frozen at -80 °C in 100 µl aliquots containing 20% glycerol.

Prior transformation the competent cells aliquot was thawed on ice for 30 min. After incubation with 50 ng of plasmid DNA or 2 µl of ligation reaction for 15 min, the cells were heat-shocked by incubating at 42 °C for 90 sec. Following the incubation on ice for 2 min, the transformed cells were added with 900 µl of SOC medium immediately, and shaken at 37 °C and 150 rpm for 60 min. Finally 100 µl of transformed cells were plated on LB-agar plates containing appropriate antibiotics and incubated overnight at 37°C.

2.2.1.10. Transformation with electroporation competent *E. coli* cells

To achieve high transformation efficiency, electroporation was used for library construction and pool DNA transformation. For preparation of electroporation competent *E. coli* cells, 5 ml of 2YT medium was inoculated with a single colony picked from a fresh 2YT-agar plate and shaken overnight at 37 °C. The overnight culture was inoculated into 500 ml of LB medium at a ratio of 1:100 and grown at 37 °C with shaking at 200 rpm until OD₆₀₀ reached 0.5-0.6. After placing on ice for 30 min, the cells were transferred into a 1 liter centrifuge bottle and centrifuged at 1600 x g for 16 min at 4 °C to sediment the cells. The cell pellet was resuspended gently in 500 ml of pre-cold sterile Millipore water. After centrifugation at 1800 x g for 18 min at 4 °C, the cell pellet was resuspended with 250 ml of ice-cold sterile Millipore water, transferred into 250 ml centrifuge bottle and centrifuged at 4°C and 2000 × g for 20 min. The cell pellet was subsequently resuspended in 50 ml of pre-cold sterile 10% glycerol solution (dissolved with Millipore water and filtered through 0.45 µm filters), transferred to a 50 ml sterile Facon tube and centrifuged at 4°C and 1600 × g for 14 min. Finally the cell pellet was resuspended in 2.5 ml of pre-cold sterile 10% glycerol solution, aliquoted to 80 µl and stored at -80 °C.

Prior transformation the competent cells aliquot was thawed on ice for 30 min, mixed with 0.5 µl of ligation products or 2 µl of deionized DNA sample and transferred into pre-chilled 1 mm gap electroporation cuvette. The electroporation was performed under voltage of 1800 V, resistance of 200 Ω and capacitance of 25 µF. The transformed cells were added with 900 µl of pre-warmed SOC medium, shaken at 37 °C and 150 rpm for 60 min and plated on 2YT-agar plate containing appropriate antibiotics. The plate was incubated overnight at 37 °C.

2.2.1.11. DNA sequencing and analysis

The sequencing of all single-plasmid DNA samples was performed by QIAGEN (Hilden, Germany). For the sequencing of samples in 96-well plate format, overnight culture was inoculated into LB-agar wells containing appropriate antibiotic and sent to GATC Biotech (Konstanz, Germany). The sequences were analyzed and aligned by using software Clone Manager, BioEdit and Microsoft Excel.

2.2.2. Production of the target protein nPAC1-Rs

2.2.2.1. Expression of nPAC1-Rs

The overnight culture was prepared by inoculating 20 ml of 2YT medium containing 100 µg/ml of kanamycin with a single colony of freshly transformed *E. coli* BL21(DE3) cells harboring plasmid pTrS/nPAC1-Rs and shaking overnight at 37 °C and 220 rpm. Totally 2 liter of autoinduction medium ZYM-5052 (Kana: 100µg/ml) was inoculated with overnight culture at a ratio of 1:500 and shaken at 220 rpm and 30 °C for 24 hours for expression. The cells were harvested by centrifugation at 4000 × g for 10 min and stored at -80 °C.

2.2.2.2. Purification of nPAC1-Rs

The nPAC1-Rs protein was purified in five steps consisting of 1st ion mental affinity chromatography (IMAC), desalting, SUMO protease digestion, 2nd IMAC and size exclusion chromatography (SEC). All purification steps were controlled by ÄKTA explorer system and carried out at 4 °C.

The cell pellets were thawed on ice for 30 min and resuspended in 150 ml of cell lysis buffer containing 30 mg lysozyme and 15 µl of Benzonase nuclease (25U/µl, Novagen). After incubation at RT for 60 min on a roller shaker, the cells were disrupted by ultrasonication for 5 × 20 sec with amplitude of 40% and pause of 30 sec between pulses. Following an additional incubation at RT for 30 min on a roller shaker, the cells were centrifuged at 4 °C and 50,000 × g for 30 min. The soluble fraction was filtered through 0.45 µm filters and applied for purification.

The soluble fraction was loaded onto 5 ml HisTrap HP column pre-equilibrated by buffer NPI-20 for 1st IMAC purification. Following the washing step with 8 column volume (CV) of buffer NPI-50, the bound proteins were eluted by buffer NPI-500. The elution fraction was loaded onto HiPrep 26/10 desalting column, buffer-exchanged to desalting buffer and subsequently digested by SUMO protease (1:200 diluted) overnight at 4 °C on a roller shaker. The digested protein reaction was concentrated with Vivaflow 50 (MWCO 3000) and applied to 2nd IMAC purification

by loading onto pre-equilibrated 5 ml HisTrap HP column and collecting the flow through fraction. Thereafter, the flow through fraction was further purified by size exclusion chromatography with HiLoad16/60 Superdex 75 pg column. The nPAC1-Rs protein was eluted from the column in HEPES buffer and stored at 4 °C until use.

2.2.3. Selection of nPAC1-Rs-specific Affilin[®] binders by TDP technology

2.2.3.1. Preparation of biotinylated nPAC1-Rs as a target

The purified nPAC1-Rs protein was dialyzed against 1 × PBS buffer for 3 times at 4°C using Spectra/Por 7 tubing (MWCO 3000). The Biotinylation was carried out by mixing nPAC1-Rs with EZ-Link Sulfo-NHS-LC-Biotin in appropriate molar ratio and incubating overnight at 4°C on a roller shaker. Excess non-reacted biotin and reaction byproducts were removed by another three times dialysis against HEPES buffer. The biotinylated nPAC1-Rs protein was applied for further analysis and selection.

A pull-down experiment was employed to analyze the biotinylation efficiency. In this experiment, 1 µg of biotinylated nPAC1-Rs protein was mixed and incubated at RT for 60 min with 30 µl of Dynabeads Streptavidin C1, which was pre-blocked by 1 ml of 10× Sigma blocking buffer and washed with 1 ml PBS-T (PBS containing 0.1% Tween 20) for three times. The supernatant and beads were separated by focusing reaction tube on magnetic rack for 2 min. The beads were washed three times with 1 ml PBS-T, which was subsequently pooled as wash fraction. The residual protein in wash fraction was precipitated by trichloroacetic acid and resuspended in 10 µl PBS, while the biotinylated nPAC1-Rs was bound to Streptavidin beads and eluted by SDS sample buffer as elute fraction. Finally the supernatant fraction, wash fraction and elute fraction were mixed with 5 µl of 5 × SDS sample buffer respectively and heated at 80 °C for 10 min. all these three fractions were loaded into NuPAGE 4-12% Bis-Tris Gel for analysis.

2.2.3.2. Phage preparation for selections

The phage particles displaying Affilin[®] variants were rescued by helper phage M13KO7 from *E. coli* ER2738 cells harboring phagemid pCD87SA-SPWF (2.1.4, page 33) and applied for selection.

For the phage propagation in 1st round of selection, the freshly transformed *E. coli* ER2738 cells were inoculated in 1 liter of 2YT medium supplemented with chloramphenicol (30 µg/ml) to get a starting OD₆₀₀ of 0.1. The cells were incubated at 37 °C and 220 rpm until OD₆₀₀ reached 0.4. Following an ice bath to cool down to 25 °C, the cells were superinfected with 1 ml of M13KO7 helper phage (2.1.3, multiplicity of infection (MOI) = 1) and incubated for 30 min at 26 °C and 50 rpm. Kanamycin was subsequently added to a final concentration of 50 µg/ml. After

incubation for 30 min at 26 °C and 220 rpm, the cells were centrifuged for 10 min at 4 °C and 3000 × g, and then resuspended in pre-warmed 1 liter of 2YT medium containing 30 µg/ml chloramphenicol, 50 µg/ml kanamycin and 0.1 µg/ml tetracycline. The culture was incubated overnight at 26 °C and 220 rpm.

For phage harvest, the overnight culture was centrifuged at 4 °C and 10,000 × g for 20 min. The supernatant was filtered through 0.45 µm syringe filters. The phage particles in supernatant were precipitated by adding 1/4 volume of PEG/NaCl solution (2.1.9, page 30) and incubating for 1 hour on ice. Following centrifugation for 20 min at 4 °C and 12,000 × g, the phage pellets were resuspended in total 50 ml of sterile cold PBS and solubilized by incubating on ice for 30 min. The insoluble particles were removed by centrifugation for 5 min at 4 °C and 20,000 × g. The phages were further purified by adding 1/4 volume of PEG/NaCl solution, incubating on ice for 30 min, centrifuging for 30 min at 4 °C and 17,000 × g and resuspending in 2 ml of sterile cold PBS. The insoluble particles were removed by centrifugation for 5 min at 4 °C and 17,000 × g. The concentrated phages were applied for 1st round of selection against biotinylated nPAC1-Rs.

To calculate the titer of input phage, 10 µl of concentrated phage sample was separated and diluted as 10-fold series dilution. 10 µl of phage sample from each series dilution was used to infect 190 µl of exponentially growing *E. coli* ER2738 cells. After incubating for 30 min at 37 °C, the infected cells were plated on SOB-CG plates and incubated overnight at 32 °C.

For the subsequent three selection rounds, phages were rescued in 100 ml culture and purified using the protocol described above in a smaller volume.

2.2.3.3. Selection against biotinylated nPAC1-Rs

1st round of selection on NeutrAvidin-coated strip

For the 1st selection round, approximate 8.3×10^{12} phage particles of the Affilin[®] phage library (2.2.3.2, page 38) were added to 8 NeutrAvidin-coated wells, which had been blocked with 200 µl/well of 10 × Sigma blocking buffer overnight at 4 °C, washed 2 times with 200 µl/well PBS and incubated for 30 min at RT with 2.5 µg/well of biotinylated nPAC1-Rs in a volume of 100 µl/well PBS followed by 3 times washing with 200 µl/well PBS. After incubating on thermomixer at RT for 2 hours and rinsing the strip 5 times with 200 µl/well PBS as well as 5 times with 200 µl/well PBS-T respectively, the phage particles were eluted 2 times with 100 µl/well of buffer E for 5 min. Eluates were pooled and neutralized with 240 µl neutralizer, followed by infecting 10 ml of exponentially growing *E. coli* ER2738 cells. After incubating for 30 min at 37 °C, 10 µl of infected cells were separated, diluted as 10-fold series dilution and plated on SOB-CG petri dishes, while the majority cells were centrifuged for 10 min at 4 °C and 3500 × g, resuspended in 2 ml of 2YT medium and plated on a SOB-CG Q-Tray. All plates were incubated overnight at 32 °C. The petri dishes were

used to calculate output phage titer, while the cells in Q-Tray plate were scraped with 12 ml of 2YT-chloramphenicol medium, subjected to inoculate 100 ml of 2YT medium containing 30 µg/ml chloramphenicol, and grown at 37 °C for 2 hours and 220 rpm. This pre-culture was used to make glycerol stocks and inoculate 100 ml of 2YT-chloramphenicol medium for next round of phage rescue as described in section 2.2.3.2, page 38.

Following rounds of selection with Streptavidin beads

For the subsequent three rounds of selection, approximate 10^{12} of amplified phage particles were used and incubated with decreasing amount of target protein coated M270 streptavidin Dynabeads from 100 µl to 10 µl. In the meanwhile, the amount of biotinylated nPAC1-Rs protein decreased from 2.5 µg to 0.25 µg and washing times were increased from 10 times to 20 times (detailed selection strategy presented in Table 2.1).

Two different selection strategies were used in following three selection rounds. In rounds 2 and 3, selection was carried out with both Selection-In-Solution (SIS) and Selection-On-Immobilized-Target (SOIT) strategies, while the selection round 4 was only performed with SIS. For SIS strategy, phage particles were firstly incubated with biotinylated nPAC1-Rs protein for 2 hours at RT, and then the complex of phage-nPAC1-Rs was captured by M-270 streptavidin Dynabeads, which had been blocked by incubating overnight with 1 ml of $10 \times$ Sigma blocking buffer at 4 °C and washed twice with 1 ml PBS. For the SOIT strategy, pre-blocked beads were firstly incubated with biotinylated nPAC1-Rs protein at RT for 30 min on a roller shaker, followed by 3 times washing with 1 ml PBS. These target immobilized beads were subsequently incubated with phage particles for 2.5 hours at RT.

Two different elution methods were used in last three selection rounds. In round 2, the bound phages were simply eluted by acid elution. In round 3 and 4, the bound phages were eluted by both competitive elution and acid elution (detailed elution strategies presented in Tab. 2.1). For the competitive elution method, the complex of M-270 beads and bound phages were resuspended with non-biotinylated nPAC1-Rs solution, which was 1000 molarity excess compared to the immobilized biotinylated nPAC1-Rs. After incubation for 2 hours at RT on a mixer, eluate was used to infect 10 ml of exponentially growing *E. coli* ER2738 cells. For the acid elution method, bound phage particles were eluted 2 times from M-270 beads with 200 µl/well of buffer E for 5 min. Eluates were sequentially pooled and neutralized with 60 µl neutralizer, followed by infecting 10 ml of exponentially growing *E. coli* ER2738 cells. Phage titration (for the titer of output phage) and reamplification were carried out as the process described in 1st round of selection.

Table 2.1 Selection strategy of Affilin® SPWF library against the target biotinylated nPAC1-Rs. Two parallel selections (A and B) were performed in both selection rounds 3 and 4.

Round	appendix	Matrix	Matrix-Preparation	Phage	Target incubation	Washing	Elution
1		8 Wells NA-Strip	Blocked wells + 1.25 µg/well biotinylated nPAC1-Rs	from SPWF library	215 pmol/well coated biotinylated nPAC1-Rs, 2 h at RT	5 x PBST and 5 x PBS	2 x 100ul buffer E
2		2 x 50 µl SA-Beads	a) Blocked beads b) Blocked beads +1.25 µg biotinylated nPAC1-Rs	from cycle 1	a)SIS (432 µl): 250 nM biotinylated nPAC1-Rs for 2 h at RT, then captured by blocked beads b)SOIT: 108 pmol immobilized biotinylated nPAC1-Rs	pool beads, wash 5 x PBST and 5 x PBS	2 x 200ul buffer E
3	A	2 x 25 µl SA-Beads	a) Blocked beads b) Blocked beads +0.63 µg biotinylated nPAC1-Rs	from cycle 2	a)SIS (432 µl): 125 nM biotinylated nPAC1-Rs for 2 h at RT, then captured by blocked beads b) SOIT: 54.3 pmol immobilized biotinylated nPAC1-Rs	pool beads, wash 13 x PBST and 3 x PBS	a) compete with 125 µM nPAC1-Rs in 432 µl PBS for 2 h at RT b) 2 x 200ul buffer E
	B	2 x 25 µl SA-Beads	a) Blocked beads b) Blocked beads +0.63 µg biotinylated nPAC1-Rs	from cycle 2	a)SIS (432 µl): 125 nM biotinylated nPAC1-Rs for 2 h at RT, then captured by blocked beads b)) SOIT: 54.3 pmol immobilized biotinylated nPAC1-Rs	pool beads, wash 13 x PBST and 3 x PBS	b) 2 x 200ul buffer E
4	A	10 µl SA-Beads	Blocked beads	from cycle 3A-a	SIS (360 µl): 60 nM biotinylated nPAC1-Rs for 2 h at RT, then captured by blocked beads	wash 17 x PBST and 3 x PBS	a) compete with 60 µM nPAC1-Rs in 360 µl PBS for 2 h at RT b) 2 x 200ul buffer E
	B	10 µl SA-Beads	Blocked beads	from cycle 3A-b and 3B-b	SIS (360 µl): 60 nM biotinylated nPAC1-Rs for 2 h at RT, then captured by blocked beads	wash 17 x PBST and 3 x PBS	b) 2 x 200ul buffer E

The enrichment factor (E) can be calculated by using the equation:

$$E = (R_n \text{ output/input}) / (R_1 \text{ output/input}), \quad \text{Equation 2}$$

where R_1 output/input is the ratio of output/input in round 1, and R_n output/input is the ratio of output/input in a certain selection round (Mutuberría *et al.*, 1999).

To compare the difference between competitive elution and acid elution, two selections were performed in parallel for both round 3 and 4. In selection round 3A, the bound phages were eluted by competitive elution (3A-a) followed by acid elution (3A-b), while in round 3B, the bound phages were directly eluted by acid elution (3B-b). The competitive eluate (3A-a) was used for selection round 4A, which was subsequently eluted by competitive elution (4A-a) followed by acid elution (4A-b). Eluates 3A-b and 3B-b were combined and applied for selection 4B, which was simply eluted by acid elution (4B-b) (Fig. 2.1).

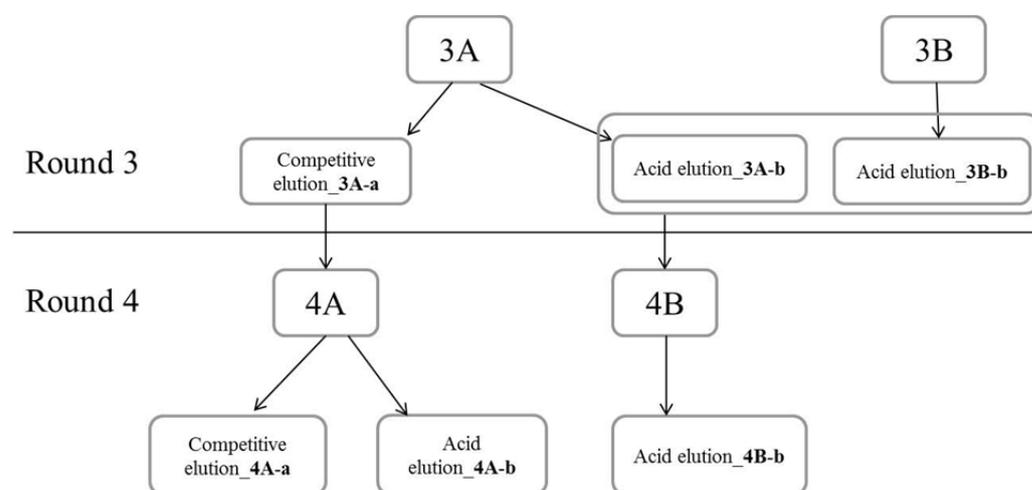


Figure 2.1 Schematic representations of two parallel selections in both round 3 and 4.

2.2.3.4. Screening by single phage ELISA

To preliminarily detect the Affilin[®] binders against target protein nPAC1-Rs after selection, single phage ELISA experiment was utilized with single clone picked from round 3 and 4. The phages displaying Affilin[®] variants were propagated in Deep-Well-Plate (DWP) and used in ELISA experiment.

Phage propagation

The single clone (corresponding phagemid in *E. coli* ER2738) was inoculated into 600 μ l/well of 2YT medium containing 30 μ g/ml chloramphenicol and incubated by shaking at 30 °C and 250 rpm for 14 hours. In next day 600 μ l/well of fresh 2YT medium containing 30 μ g/ml chloramphenicol in a new DWP was inoculated with 15 μ l of overnight culture and shaken at 37 °C and 250 rpm for 1.5 hours. Following an ice bath to cool down to 26 °C, the cells were superinfected with 10 μ l/well of M13KO7 helper phage (1×10^{11} pfu/ml) and incubated for 30 min at 26 °C without

shaking. Subsequently 10 μ l/well of diluted kanamycin (3 mg/ml) was added and incubated for 15 min at 26 °C and 250 rpm. After centrifugation for 10 min at 4 °C and 4000 \times g, the cells were resuspended with 600 μ l/well of fresh 2TY containing 30 μ g/ml chloramphenicol, 50 μ g/ml kanamycin and 0.1 μ g/ml tetracycline and shaken overnight at 26 °C and 250 rpm. The overnight culture was centrifuged for 15 min at 4 °C and 4500 \times g. 540 μ l/well of supernatant was taken into a new DWP and mixed with 60 μ l/well of 10 \times Sigma blocking buffer for single phage ELISA experiment.

Single phage ELISA

For the single phage ELISA, 80 μ l of above prepared phage samples were applied to each well (Nunc, MediSorp plates) immobilized with 200 ng nPAC1-Rs, 200 ng anti-C-myc antibody as well as negative controls 500 ng BSA, 300 ng lysozyme and 300 ng nPTH1-R respectively, followed by incubating at RT for 2 hours. 200 ng of biotinylated nPAC1-Rs protein was coated on NeutrAvidin plate. 1:50 diluted human serum was also coated as a negative control. After washing the wells three times with 300 μ l PBS-T, wells were incubated at RT for 1 hour with 50 μ l/well of anti-M13 antibody horseradish peroxidase conjugate (GE Healthcare, USA, 1:5000 diluted in PBS-TB). Following three times washing with PBS-T and three times washing with PBS, the ELISA was developed using 50 μ l/well of TMB Plus substrate solution. The chromogenic reaction was stopped by adding 50 μ l/well of 0.5 M H₂SO₄ and absorbance was measured at 450 nm (Abs 450 nm) and 620 nm by Multimode Detector DTX880.

2.2.3.5. Screening by high throughput Hit-ELISA

From the five selected phage pools in round 3 and 4, the DNA fragments encoding the Affilin[®] inserts were amplified (2.2.1.5, page 34) and subcloned into the expression vector pET23dk via *NcoI/XhoI* (2.2.1.6-2.2.1.8, page 35).

Single selected colonies were picked by colony picker and expressed using *E. coli* NovaBlue(DE3) cells in 96-well plates. In brief, 150 μ l of 2YT medium containing 50 μ l/ml kanamycin and 1% glucose was inoculated with a randomly picked colony of *E. coli* NovaBlue(DE3) harboring pET23dk-SPWF plasmid encoding one of the selected Affilin[®] variants and incubated overnight at 37 °C with shaking at 750 rpm in MTP incubator TH15. 150 μ l of fresh 2YT medium containing 50 μ l/ml kanamycin was inoculated with overnight culture by 2 transfers using high capacity long pin. After incubation at 37 °C with shaking at 750 rpm for 5 hours (OD₆₀₀ \approx 0.4), 200 μ l of fresh autoinduction medium ZYM-5052 containing 50 μ l/ml kanamycin was inoculated with the pre-culture (OD₆₀₀ \approx 0.4) by 1 transfer using high capacity long pin and shaken for 20 hours at 37 °C and 750 rpm. Cells were harvested by centrifugation at 4000 \times g and 4 °C for 20 min and stored at -80 °C.

The cell pellets were resuspended in 100 μ l of cell lysis buffer containing 0.3 mM PMSF, 0.1 mg/ml lysozyme and 5 U/ml of Benzonase nuclease by vigorous shaking

for 30 min at RT. After 5 cycles of freeze and thaw, cell debris were removed from cell lysate by centrifugation at $6800 \times g$ and $4\text{ }^{\circ}\text{C}$ for 30 min.

The cell lysates were 10-fold diluted with PBS by BIOMEK 3000. The following high throughput Hit-ELISA was performed based on BIOMEK FX Laboratory Automation Workstation. 50 μl of diluted cell lysate was applied to each well (Nunc, MediSorp plates) which had been coated overnight with target protein (200 ng nPAC1-Rs per well) or negative control (300 ng lysozyme per well) respectively and blocked by $1 \times$ sigma blocking buffer for 1.5 hours at RT. After incubation at RT for 1.5 hours and washing the wells three times with 300 μl PBS-T, wells were incubated with anti-Ubi-Fab-POD antibody (1:10000 in PBS, 50 μl per well) for 1.5 hours at RT. Following three times washing with PBS, the bound Affilin[®] proteins were detected with 50 μl /well of substrate TMB plus 2 by incubating for 30 min at RT. The chromogenic reaction was stopped by adding 50 μl /well of 0.5 M H_2SO_4 solution. Absorbance was sequentially measured at 450 nm and 620 nm.

2.2.4. Expression and purification of selected Affilin[®] binders

2.2.4.1. Expression of Affilin[®] binders

The overnight culture was prepared by inoculating 5 ml of LB medium containing 50 $\mu\text{g}/\text{ml}$ of kanamycin with a single clone of freshly transformed *E. coli* NovaBlue(DE3) cells harboring plasmid pET23dk/SPWF and shaking overnight at $37\text{ }^{\circ}\text{C}$ and 220 rpm. For small scale and large scale expression, from 5ml to 1 L of fresh autoinduction medium ZYM-5052 (Kanamycin: 50 $\mu\text{g}/\text{ml}$) was inoculated with overnight culture at a ratio of 1:100 and shaken at 220 rpm and $30\text{ }^{\circ}\text{C}$ for 24 hours. The cells were harvested by centrifugation at $4000 \times g$ for 10 min and stored at $-80\text{ }^{\circ}\text{C}$.

2.2.4.2. Purification of Affilin[®] binders

All of the Affilin[®] binders were purified by IMAC combined with size exclusion chromatography. The purification was controlled by ÄKTA xpress system and carried out at RT.

Cell pellets were thawed on ice for 30 min and resuspended in cell lysis buffer containing 0.1 mg/ml lysozyme and 5 U/ml of Benzonase nuclease with 1/10 of culture volume. After incubation at RT for 60 min on a roller shaker, the cells were disrupted by ultrasonication for 5×20 sec with amplitude of 40% and pause of 30 sec between pulses. Following an additional incubation at RT for 30 min on a roller shaker, the cells were centrifuged at $4\text{ }^{\circ}\text{C}$ and $50,000 \times g$ for 30 min. The soluble fraction was filtered through 0.45 μm filters and applied for purification.

The soluble fraction was loaded onto 1 ml or 5 ml HisTrap HP column pre-equilibrated

by buffer NPI-20 for IMAC purification. Following washing step with 8 CV of buffer NPI-50, the bound proteins were eluted by buffer NPI-500. The elution fraction was subsequently purified by size exclusion chromatography with pre-equilibrated HiLoad16/60 Superdex 75 pg column. The Affilin[®] protein was eluted from the column in gel filtration buffer and fractionated.

2.2.5. Protein characterization

2.2.5.1. Protein concentration measurement

The concentration of purified or dialyzed proteins was determined by measuring the absorbance at 280 nm. The protein concentration was calculated according to the following equation which was derived from the Beer-Lambert law:

$$C = A \times MW / (E \times L) \quad \text{Equation 3}$$

C: concentration (mg/ml), A: absorbance at 280 nm, MW: molecular weight (g/mol), E: extinction coefficient ($M^{-1} \text{ cm}^{-1}$), and L: cell path length = 1 (cm).

All extinction coefficient values of different proteins were calculated by ProtParam tool at website <http://www.expasy.org>.

2.2.5.2. SDS-PAGE analysis

The SDS-PAGE was performed with NuPAGE 4-12% Bis-Tris precast Gels and 1 × MES running buffer. Protein samples were mixed with 5 × SDS sample buffer, heated at 95 °C for 5 min and loaded on the gel assembled in X cell SureLock Electrophoresis Cell. Electrophoresis was run at 200 V for 45 min in 1 × MES running buffer. The gels were either stained in PageBlue[™] protein staining solution or applied for following silver staining procedures.

For the silver staining procedures, the gel was firstly incubated with fixing solution for 5 min. Following the washing steps with water for 3 × 5 sec, 5 min and 3 × 5 sec respectively, the gel was incubated successively in pretreatment I solution for 5 min, pretreatment II solution for 1 min and water for 5 min. The gel was subsequently incubated with impregnation solution for 8 min, washed with water for 2 × 5 sec and developed by incubating with developing solution for 10-20 sec. The developing was stopped by soaking the gel in stopping solution for 1 min. After washing with water for 5 min, the gel was imaged by documentation system.

2.2.5.3. Expression and solubility analysis

To analyze the expression level and solubility of Affilin[®] binders, 1 ml of cell culture with $OD_{600}=1.0$ was divided and centrifuged at 4 °C as well as 3000 × g for 5 min. Cell pellet was resuspended completely in 100 µl PBS containing 20 µg lysozyme and 3 µl of 10 mM PMSF by incubating at RT and 600 rpm for 20 min using thermomixer.

The cell resuspension was frozen in liquid nitrogen and thawed subsequently by incubating at 20 °C and 750 rpm in thermomixer. After five times of freeze-and-thaw cycle, 0.3 µl of Benzonase nuclease and 0.2 µl of 1 M MgSO₄ were added. Following the incubation for 30 min at 4 °C, 6 µl of disrupted cells were taken out as the whole cell lysate sample. The residual cells were centrifuged at 21,000 × g and 4 °C for 30 min and 6 µl supernatant was transferred to a new reaction tube as the soluble protein fraction. Cell pellet was washed by 150 µl PBS, centrifuged at 21,000 × g and 4 °C for 30 min and resuspended in 94 µl of resuspension buffer. 6 µl of pellet resuspension was taken into a new reaction tube as insoluble protein fraction.

All three samples were mixed with 3 µl of 5 × SDS sample buffer respectively and heated at 94 °C for 5 min. Thereafter, samples were loaded on the gel directly for SDS-PAGE analysis.

2.2.5.4. Circular Dichroism (CD) Spectroscopy

Far-UV CD spectra were measured on a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD) at 20 °C in a 1 mm quartz cuvette. The scanning range was 260-185 nm at a rate of 5 nm / min, the bandwidth and data pitch were 1 nm, and the spectra were accumulated 25 times with a response time of 4 sec.

2.2.5.5. Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry experiment was carried out using a MicroCal ITC titration calorimeter (MicroCal, USA). PACAP6-38 and nPAC1-Rs protein were dialyzed overnight respectively in 100 mM NaH₂PO₄ buffer at pH 7.0. A typical experiment used 1.8 ml of nPAC1-Rs (15 µM) in sample cell. The ligand PACAP6-38 (150 µM) was titrated into nPAC1-Rs solution at 25 °C in 24 injections of 10.5 µl, 300 sec apart. An initial injection of 2 µl was made to clear the syringe of any nPAC1-Rs from the sample cell which might have mixed with PACAP6-38 in the syringe during equilibration.

The data were analyzed with Origin software (MicroCal Software, MA, USA) based on the amount of heat liberated upon association (Wiseman *et al.*, 1998). The change in Gibbs free energy upon association, ΔG° , was calculated using the equation,

$$\Delta G^\circ = -RT \ln K_a \quad \text{Equation 4}$$

R is the gas constant and T is the absolute temperature. The entropy change of association, ΔS° , was calculated using,

$$\Delta G^\circ = \Delta H^\circ - T \Delta S^\circ \quad \text{Equation 5}$$

ΔH° is the measured enthalpy of association of the reaction at temperature T . The equivalence number, n , K_a , and ΔH° could be determined for the interaction.

2.2.5.6. Differential scanning fluorimetry (DSF)

The thermostability of selected Affilin[®] binders was measured by differential scanning fluorimetry (DSF) experiment. The SYPRO orange dye was supplied in DMSO at 5000 times the working concentration and diluted to 100 times with gel filtration buffer (2.1.9) prior to being added into protein solutions, to obtain a final 2 times the working concentration in assay samples. Optical 96-well reaction plate was used with 50 μ l of solution per well, containing 10 μ g protein sample. Fluorescence measurements were performed using LightCycler 480 II instrument. The fluorescence emission was collected at 510 nm with a fixed excitation wavelength at 465 nm. During the DSF experiment the temperature was increased from 20 to 90 °C with a heating rate of 1 °C/min. The data were analyzed by using a DSF analysis program established by Dominik Schneider. The apparent melting temperature T_m is defined as the midpoint transition temperature between folded and unfolded states.

2.2.5.7. Specificity test by ELISA

The specificity of selected Affilin[®] binders was tested by ELISA experiment. Target protein nPAC1-Rs and negative control BSA, lysozyme, nPTH1-R were coated overnight at 4 °C on Nunc MediSorp plate with amount of 200 ng, 500 ng, 300 ng and 300 ng per well respectively, while 1:50 diluted human serum was coated as negative control under the same conditions. The wells were washed once by PBS and blocked by 1 \times Sigma blocking buffer for 2 hours at RT. After three times washing by PBS-T, the wells were incubated with 80 μ l/well of purified Affilin[®] variant for 2 hours at RT. Following three times washing by PBS-T, 80 μ l/well of anti-Ubi-Fab antibody POD conjugated (1:6500 diluted in PBS-TB) was added and incubated for 1 hour at RT. Wells were washed three times by PBS-T and three times by PBS. Thereafter, the bound Affilin[®] was detected by incubating with 80 μ l/well of substrate TMB plus solution for 30 min at RT. The reaction was stopped by adding 40 μ l/well of 0.5 M H₂SO₄ solution, and absorbance was measured at 450 nm and 620 nm.

2.2.5.8. Concentration-dependent ELSIA

The concentration-dependent ELISA was employed for extensive determination of the binding affinities of selected Affilin[®] binders to the target nPAC1-Rs. In this experiment, purified Affilin[®] binders were 3-fold serially diluted and applied to the wells which had been coated with 200 ng/well nPAC1-Rs and blocked with 300 μ l/well of 1 \times sigma blocking buffer. The following procedures are the same as described in section 2.2.5.7, page 47.

To determine the apparent dissociation constant (K_D), the values of Abs 450 nm were plotted against protein concentration utilizing software SigmaPlot 11.0. The K_D values were calculated according to the “ligand binding and one site saturation” equation

derived from the law of mass action (Raghava and Agrewala, 1994; Voss and Skerra, 1997):

$$y = B_{max} \times x / (K_D + x) \quad \text{Equation 6}$$

y : specific binding data, B_{max} : maximum number of binding sites, x : concentration of free ligand, K_D : concentration of ligand to reach half maximal binding.

2.2.5.9. Surface plasmon resonance (SPR) experiment

The interactions of Affilin[®] binders or the ligand PACAP6-38 with the target protein nPAC1-Rs were analyzed by real-time biospecific interaction analysis using a Biacore T100 instrument (GE Healthcare). The biotinylated nPAC1-Rs protein was diluted in 10 mM sodium acetate solution at pH 4.0 to a final concentration of 0.5 µg/ml and immobilized as ligand in one cell of SA sensor chip at a flow rate of 5 µl/min and for a contact time of 90 sec. One blank flow cell was used as reference surface to correct the non-specific binding and bulk refractive-index change. 1 × HBS-EP (10 mM HEPES, 150 mM NaCl, 3 mM EDTA and 0.005% surfactant P20 at pH 7.4) was used as running buffer and for analyte dilutions. Purified and filtered Affilin[®] binders were diluted to different concentrations with running buffer and flowed through blank cell and ligand immobilized cell in parallel. Regeneration between cycles was performed using 10 mM Glycine-HCl solution at pH 2.0 and run at 30 µl/min with a contact time of 45 sec. A mid concentration for each analyte was taken and repeated following the highest concentration for duplicate analysis. Before data evaluation, the response from the reference surface was subtracted from the ligand surface response. Biacore T100 evaluation software v.1.1 (GE Healthcare) was used for data evaluation.

Determination of kinetic rate constants

Injections of Affilin[®] binders diluted from 2560 nM to 2.5 nM respectively were performed using SA sensor chip containing biotinylated nPAC1-Rs protein. The assay was run at a rate of 50 µl/min with a sample contact time of 120 sec and a dissociation time of 360 sec. A sensorgram from a running buffer HBS-EP injection was subtracted from the sample injections before data analysis. The kinetic rate constants k_{on} and k_{off} , as well as the equilibrium dissociation constant K_D , were calculated using 1:1 binding model.

Ligand competition assay

Affilin[®] binders were mixed with 1600 nM PACAP6-38 respectively, incubated for 30 min at RT, and sequentially injected at 50 µl/min over a SA sensor chip containing the immobilized biotinylated nPAC1-Rs protein. The final concentrations of mixed Affilin[®] binders were 80 nM for P1A02, P2A08 and P2A12, and 640 nM for 2H07-1, respectively.

3. Results

3.1. Recombinant production of functional nPAC1-Rs

Highly pure and functional target protein is important for efficient selection of specific binding molecules by phage display technology. The expression, purification and biophysical characterization of the target protein used in this study, the N-terminal extracellular domain of the human PAC1 receptor (nPAC1-Rs) are described in the following sections.

3.1.1. Recombinant production of nPAC1-Rs

The expression of nPAC1-Rs was accomplished by using the fusion protein thioredoxin-SUMO-nPAC1-Rs. The corresponding plasmid pTrS/nPAC1-Rs (see Supplementary material 7.1, page 116) was transformed into *E. coli* BL21(DE3) cells and expression was performed using autoinduction medium ZYM-5052. After shaking at 30 °C for 24 hours, the cell culture reached an OD₆₀₀ of 9-10. Approximately 40 g of cell mass was obtained from 2 liters of cell culture by centrifugation.

The soluble cell extract was firstly purified by the 1st Immobilized metal-ion affinity chromatography (IMAC). The eluate from the 1st IMAC containing components of Trx-SUMO-nPAC1-Rs and Trx-SUMO fusion proteins (Fig. 3.2a) was applied to a desalting column, allowing a buffer-switch to SUMO protease digestion buffer. Cleavage of Trx-SUMO-nPAC1-Rs fusion protein by SUMO protease resulted in production of native nPAC1-Rs with no extra amino acids added between the cleavage site and the start amino acid residue of nPAC1-Rs protein. After digestion, the Trx-SUMO and incompletely digested Trx-SUMO-nPAC1-Rs fusion proteins as well as the SUMO protease were removed in the 2nd IMAC due to these proteins contained a 6×His-tag, whereas the tag-free nPAC1-Rs protein directly flowed through the IMAC column. The flow through fraction containing nPAC1-Rs protein was further purified by size-exclusion chromatography (SEC).

In size-exclusion chromatography with a HiLoad16/60 Superdex 75 pg column, three elution peaks were observed in the retention volume of 47.49 ml, 65.77 ml and 77 ml, presumably corresponding to the aggregates, dimer and monomer nPAC1-Rs protein, respectively (Fig. 3.1). The majority of the nPAC1-Rs protein subjected to SEC was in monomer form.

The various fractions of nPAC1-Rs during purification were loaded onto SDS-PAGE gel under reducing conditions (Fig. 3.2b). The size of Trx-SUMO-nPAC1-Rs and Trx-SUMO fusion proteins is approximately 37.3 kDa and 26 kDa respectively. The

molecular weight of nPAC1-Rs protein is approximately 11.6 kDa.

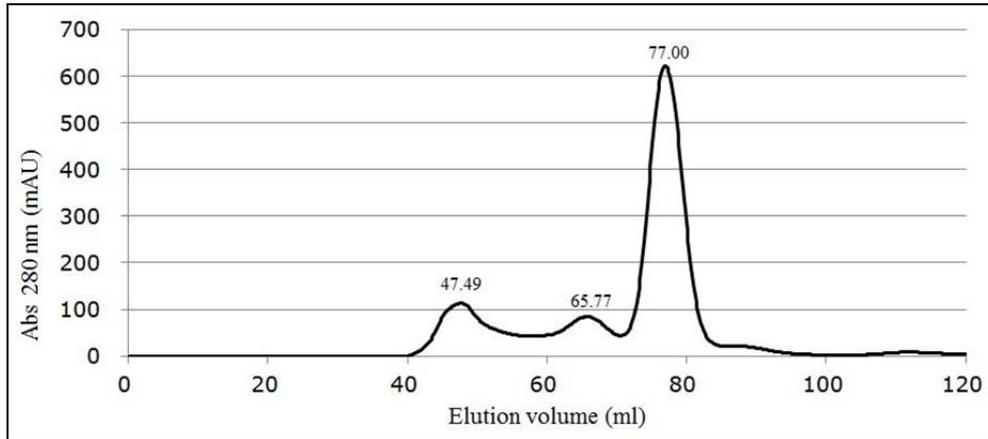


Figure 3.1 Elution profile of nPAC1-Rs protein from a HiLoad16/60 Superdex 75 pg size-exclusion chromatography column. The flow through fraction of the 2nd IMAC containing native nPAC1-Rs protein was applied to the gel filtration column and subsequently eluted with HEPES buffer.

(a)



(b)

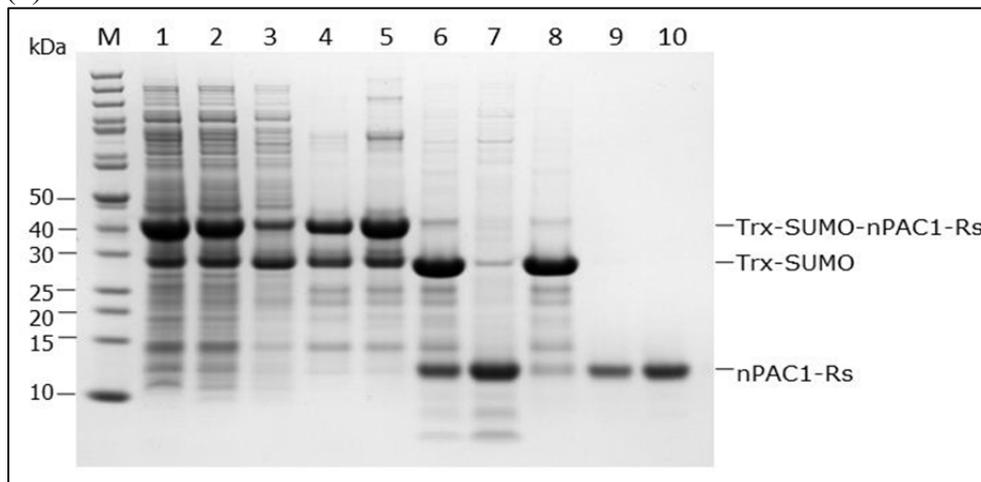


Figure 3.2 Purification of nPAC1-Rs protein. (a) Fusion proteins presented in purification. (b) Various fractions during the purification of nPAC1-Rs protein analyzed by SDS-PAGE. M, PageRuler™ unstained protein ladder; Expression: lane 1, whole cell lysate; lane 2, soluble fraction; The 1st IMAC: lane 3, flow through; lane 4, elute; Desalting: lane 5, elute from desalting column; The 2nd IMAC: lane 6, sample after SUMO protease digestion; lane 7, flow through; lane 8, elute; Gel filtration: lane 9, 1 µg of nPAC1-Rs protein; lane 10, 2 µg of nPAC1-Rs protein.

By comparing the protein amount between lane 1 and 2 (Fig. 3.2b), we know that the fusion protein Trx-SUMO-nPAC1-Rs was successfully expressed with high yield and solubility in the cytoplasm of *E. coli* cells. The bands of Trx-SUMO-nPAC1-Rs and Trx-SUMO fusion proteins presented in lane 3, indicating that the HisTrap HP column was overloaded. From lane 9 and 10, the purity of produced nPAC1-Rs protein was estimated to be greater than 95%. Finally, approximately 10 mg of purified nPAC1-Rs protein was obtained from 2 liters of cell culture.

3.1.2. Preparation of biotinylated nPAC1-Rs

It was reported that conventional immobilization of proteins by direct passive adsorption to plastic surfaces results in partial or complete denaturation of the proteins (Schwab *et al.*, 1986 and Suter *et al.*, 1986). To avoid partial protein denaturation, an approach of the streptavidin/NeutrAvidin-biotin specific interaction was used for immobilization of the target protein nPAC1-Rs during selection.

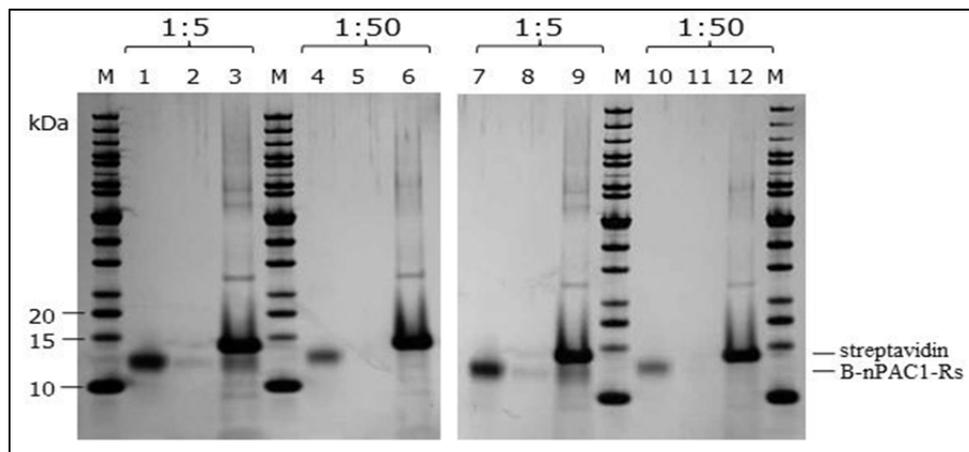
The biotinylation of nPAC1-Rs target protein was carried out using EZ-Link Sulfo-NHS-LC-Biotin kit (Pierce protein). To get high biotinylation efficiency, different pH conditions and molar ratios between nPAC1-Rs protein and biotin were tested. The biotinylation efficiency was evaluated by pull-down experiments (in section 2.2.3.1, page 38), in which each fraction was loaded on SDS-PAGE gel as shown in Fig. 3.3. Biotinylation of nPAC1-Rs protein was firstly tested in PBS buffer at pH 7.4 and pH 8.6 with different molar ratio of nPAC1-Rs and biotin at 1:5 and 1:50, respectively. The distribution of nPAC1-Rs protein in each fraction from pull-down experiments revealed the biotinylation efficiency in different conditions.

As presented in Fig. 3.3a, the molecular weight of streptavidin and biotinylated nPAC1-Rs is approximately 14 and 12 kDa respectively. There was almost no nPAC1-Rs band detected in all supernatant fractions (in lane 2, 5, 8 and 11). In contrast, the band of biotinylated nPAC1-Rs appeared clearly in elution fractions (in lane 3 and 9), showing that the nPAC1-Rs protein was successfully biotinylated and captured by streptavidin beads. The biotinylation performed at pH 7.4 showed slightly higher efficiency than at pH 8.6 by comparing the amount of biotinylated nPAC1-Rs in lane 3 and 9.

Interestingly, the band of biotinylated nPAC1-Rs protein was presented neither in supernatant fraction nor in elution fraction when the biotinylation was performed with a molar ratio of 1:50 between nPAC1-Rs and biotin as shown in lane 5, 6, 11 and 12 of Fig. 3.3a. A similar situation was observed in lane 10, 11 and 12 of Fig. 3.3b when the molar ratio of nPAC1-Rs and biotin increased to 1:10. Only much less biotinylated nPAC1-Rs was found in elution fraction (lane 12) comparing to control (lane 9), and no band was detected in both supernatant and washing fractions (lane 10 and 11). It was assumed that nPAC1-Rs protein was excessively labeled by biotin under a high molar ratio of biotin and nPAC1-Rs, allowing tight binding of biotinylated nPAC1-Rs

to streptavidin, which may not be denatured completely during sample treatment of SDS-PAGE. Furthermore, it is not desirable to get excessively biotin-labeled target protein, which may negatively affect the interaction of target protein and binders during selection. Therefore, lower molar ratios from 1:2 to 1:10 were used for further biotinylation tests.

(a)



(b)

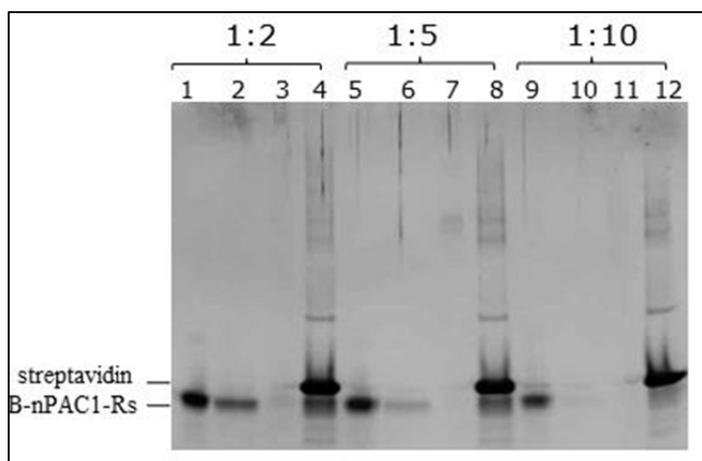


Figure 3.3 Analyses of biotinylation efficiency of nPAC1-Rs protein by pull-down experiments. In this experiment, 1 μg of biotinylated nPAC1-Rs protein (B-nPAC1-Rs) was mixed and incubated with 30 μl of Dynabeads Streptavidin C1. The supernatant and beads were separated by focusing reaction tube on magnetic rack. The beads were washed three times with 1 ml PBS-T, which was subsequently pooled and precipitated by TCA as wash fraction. The biotinylated nPAC1-Rs bound to Streptavidin beads was eluted by SDS sample buffer as elute fraction. The supernatant fraction, wash fraction and elution fraction of each pull-down experiment were analyzed by SDS-PAGE. (a) Biotinylation of nPAC1-Rs was performed in PBS buffer at pH 7.4 (left part, lane 1-6) and at pH 8.6 (right part, lane 7-12). The molar ratio of nPAC1-Rs and biotin used for biotinylation was 1:5 and 1:50 respectively as labeled above the lanes. M, PageRulerTM unstained protein ladder; lane 1, 4, 7 and 10, 1 μg of biotinylated

nPAC1-Rs as control; lane 2, 5, 8 and 11, supernatant fraction after pull-down; lane 3, 6, 9 and 12, elution fraction from streptavidin beads. (b) Biotinylation of nPAC1-Rs was performed in PBS buffer at pH 7.4 with different molar ratio of nPAC1-Rs and biotin at 1:2, 1:5 and 1:10 respectively, as labeled above the lanes. lane 1, 5 and 9, 1 μg of biotinylated nPAC1-Rs as control; lane 2, 6 and 10, supernatant fraction after pull-down; lane 3, 7 and 11, wash fraction; lane 4, 8 and 12, elution fraction from streptavidin beads.

As illustrated in Fig. 3.3b, further biotinylation tests were carried out at pH 7.4 with lower molar ratio of nPAC1-Rs and biotin varied from 1:2 to 1:10. The amount of biotinylated nPAC1-Rs in the supernatant fraction (lane 2) was slightly less than the elution fraction (lane 4). Therefore the biotinylation efficiency using a molar ratio of 1:2 was estimated to be about 60%. When molar ratio increased to 1:5, the amount of biotinylated nPAC1-Rs in supernatant fraction (lane 6) decreased comparing to lane 2. The same situation happened to elution fraction (lane 8). Finally, considering both biotinylation efficiency and the level of incorporation, biotinylation of nPAC1-Rs protein was finally performed in PBS buffer at pH 7.4 with a molar ratio of 1:2 between nPAC1-Rs and biotin.

To avoid underloading of streptavidin beads during selection, the binding capacity of M-270 beads to biotinylated nPAC1-Rs was also determined by pull-down experiment. A series of different amounts of biotinylated nPAC1-Rs were incubated with 20 μl of M-270 streptavidin beads respectively. The supernatant and elution fractions were analyzed by SDS-PAGE (Fig. 3.4).

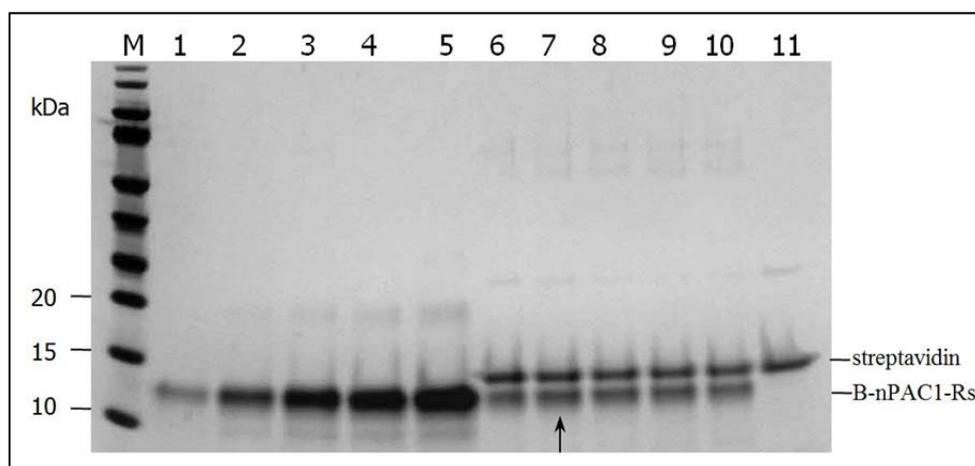


Figure 3.4 Binding capacities of M-270 streptavidin beads to biotinylated nPAC1-Rs protein tested by pull-down experiment. A series of different amounts of biotinylated nPAC1-Rs (B-nPAC1-Rs) were incubated with 20 μl of M-270 streptavidin beads respectively. The supernatant and elution fractions were analyzed by SDS-PAGE. M, PageRuler™ unstained protein ladder; lane 1-5, supernatant fraction after pull-down, the corresponding initial amount of biotinylated nPAC1-Rs protein was 0.25, 0.5, 0.75, 1.0 and 1.25 μg respectively; lane 6-10, elution fraction from M-270 streptavidin beads, the corresponding initial amount of biotinylated nPAC1-Rs was 0.25, 0.5, 0.75, 1.0 and 1.25 μg respectively.

As shown in Fig 3.4, increased amount of biotinylated nPAC1-Rs was observed in elution fraction lane 7 indicated by back arrow when compared to lane 6. The corresponding protein amount used in lane 7 was 0.5 μg . An almost equal amount of biotinylated nPAC1-Rs protein presented from lane 7 to 10, showing that the binding ability of 20 μl of M-270 streptavidin beads was saturated. Therefore the binding capacity of 20 μl of M-270 streptavidin beads was estimated up to 0.5 μg of biotinylated nPAC1-Rs protein.

3.1.3. Biophysical characterization of nPAC1-Rs

Circular Dichroism (CD) Spectroscopy

To obtain information about the secondary structure of purified nPAC1-Rs protein, far-UV-CD spectra was measured. The concentration of nPAC1-Rs protein was 10.6 μM in 10 mM KH_2PO_4 pH 7.0. Data were measured at 1 nm of band width over the range of 260-185 nm with a response time of 4 sec/data point. The spectrum was accumulated 25 times. The spectrum was characteristic of a folded protein with secondary structure elements dominated by α -helix, β -sheet and turns (Fig. 3.5). It showed a negative band with a minimum at 208 nm and a mean residue ellipticity of approximately $-9000 \text{ deg cm}^2 \text{ dmol}^{-1}$ at this wavelength.

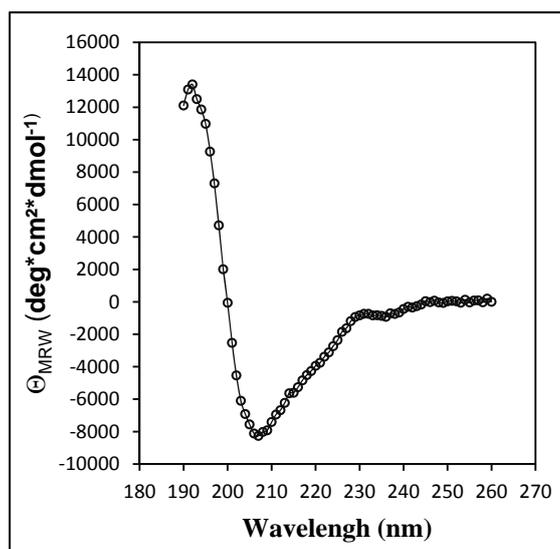


Figure 3.5 Circular Dichroism spectroscopy of purified nPAC1-Rs protein. The concentration of nPAC1-Rs protein was 10.6 μM in 10 mM KH_2PO_4 pH 7.0. Data were measured at 1 nm of band width over the range of 260-185 nm with a response time of 4 sec/data point. The spectrum was accumulated 25 times, corrected for buffer contributions, and converted to mean residue ellipticity according to Schmid (Schmid, 1997).

Ligand binding experiments

To analyze the functionality of produced nPAC1-Rs protein, binding characteristics of nPAC1-Rs to its ligand PACAP6-38 were firstly determined with surface plasmon

resonance (SPR) on a Biacore T100 instrument. Biotinylated nPAC1-Rs protein was immobilized onto a SA sensor chip by biotin-streptavidin interaction to approximately 314 response units (RU). Serial PACAP 6-38 concentrations ranging from 400 nM to 6.4 μ M were applied to the chip as displayed in Fig. 3.6a. The binding of ligand PACAP6-38 to immobilized biotinylated nPAC1-Rs protein showed very fast association and dissociation rates. From the concentration dependent steady state signals, a K_D value of 605.3 nM was calculated (Fig. 3.6b).

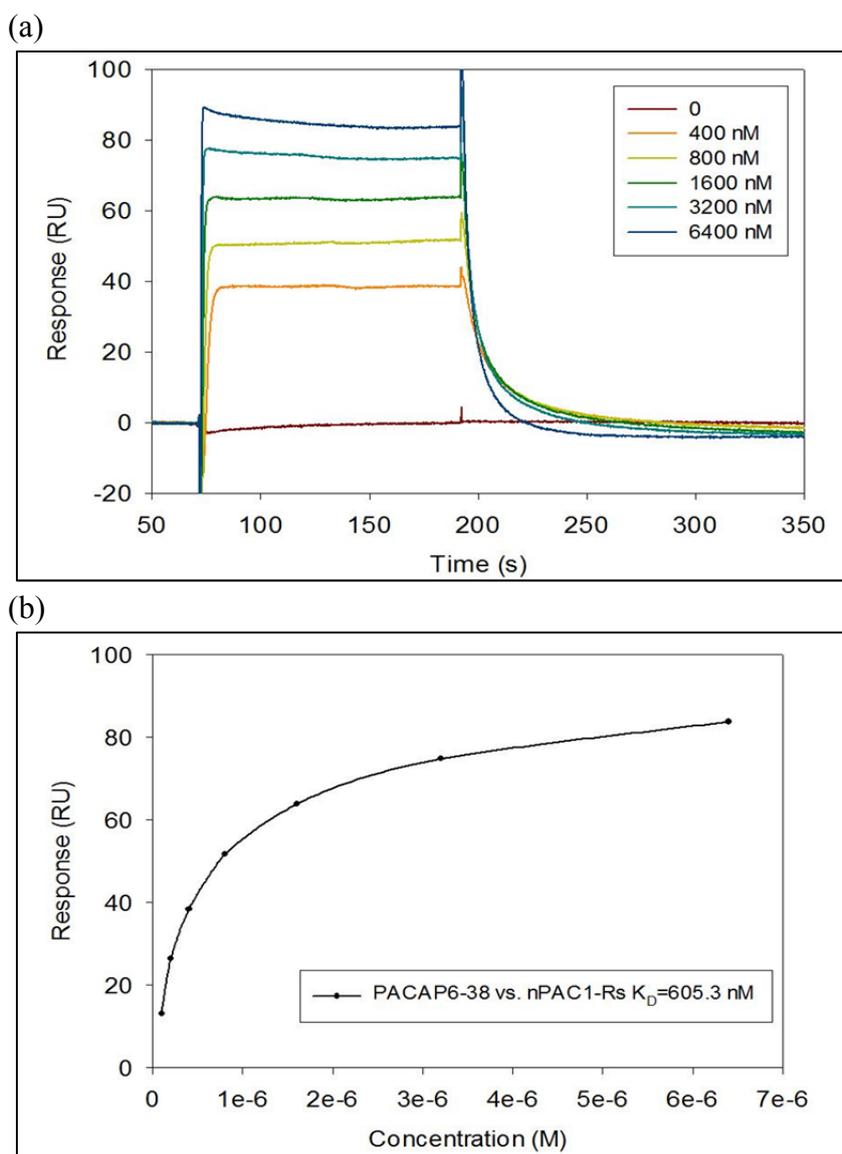


Figure 3.6 Surface plasmon resonance (SPR) analysis of the binding of ligand PACAP6-38 to nPAC1-Rs. (a) The sensorgrams represent the ligand PACAP6-38 with concentrations varied from 400 nM to 6400 nM, respectively, injected over a SA sensor chip immobilized with biotinylated nPAC1-Rs protein. The biotinylated nPAC1-Rs protein was immobilized with response units (RU) of 314. (b) The binding affinity of the ligand PACAP6-38 to nPAC1-Rs is 605.3 nM calculated from the concentration dependent steady state signals by using Biacore T100 Evaluation Software.

Isothermal titration Calorimetry (ITC) was used as a second method to quantitatively determine the binding of the ligand PACAP6-38 to nPAC1-Rs protein. Titration of PACAP6-38 with nPAC1-Rs protein led to the release of heat upon binding (Fig. 3.7). The integrated areas were averaged and subtracted from the raw data using software Origin 7, MicroCal LLC ITC. Analysis of data revealed an apparent ΔH of $-6860 (\pm 77.35) \text{ cal mol}^{-1}$, and $T\Delta S$ of $4.63 \text{ cal mol}^{-1} \text{ deg}^{-1}$. The calculated K_D value is $909 (\pm 52) \text{ nM}$.

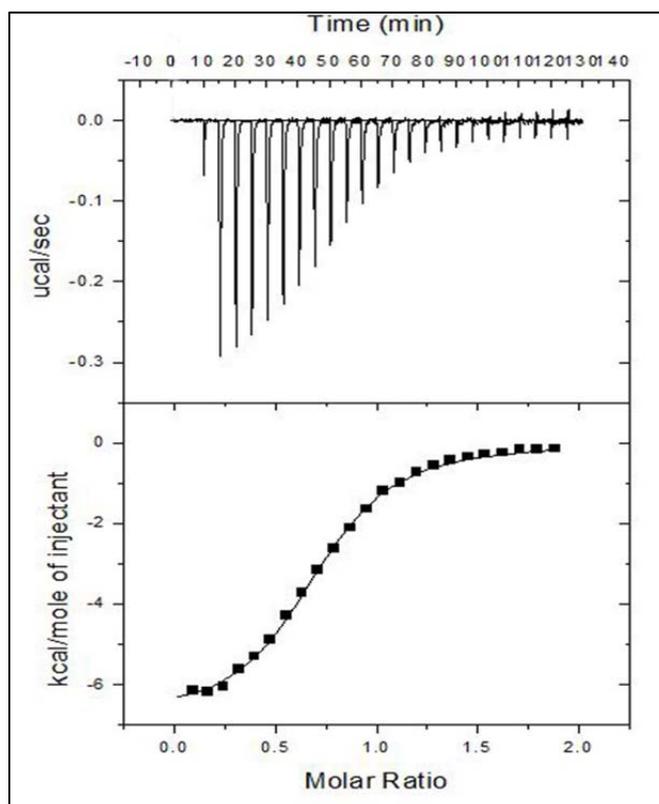


Figure 3.7 Isothermal titration Calorimetry results of the interaction between PACAP6-38 and nPAC1-Rs. 1.8 ml of nPAC1-Rs (15 μM) was used in sample cell. The ligand PACAP6-38 (150 μM) was titrated into nPAC1-Rs solution at 25 $^{\circ}\text{C}$ in 24 injections of 10.5 μl , 300 sec apart. An initial injection of 2 μl was made to clear the syringe of any nPAC1-Rs from the sample cell which might have mixed with PACAP6-38 in the syringe during equilibration. Both protein samples were dialyzed overnight in 100 mM NaH_2PO_4 buffer at pH 7.0, respectively. Top, the baseline subtracted raw data; bottom, the peak-integrated and concentration-normalized enthalpy change vs. PACAP 6-38/nPAC1-Rs ratio.

In conclusion, nPAC1-Rs protein was expressed as a soluble fusion protein in recombinant *E. coli* cells. The native nPAC1-Rs protein was obtained after purification, with a purity of $\geq 95\%$. The biological activity of produced nPAC1-Rs protein was demonstrated by its binding ability to the ligand PACAP6-38 determined via both SPR and ITC experiments.

3.2. Selection of dimeric ubiquitin-based Affilin[®] binders against nPAC1-Rs

According to accumulated knowledge and experiences on monomeric ubiquitin-based library, a dimeric ubiquitin-based Affilin[®] library was designed by Scil Proteins. In this library, two ubiquitin domains are connected by a (SG₄)₂ linker. Totally 15 surface exposed residues are randomized by 19 natural amino acids without cysteine, 8 in the first domain-SPW domain (in position of 2, 4, 6, 62, 63, 64, 65 and 66) and 7 in the second domain-SPF domain (in position of 6', 8', 62', 63', 64', 65' and 66'). Thereby a contiguous binding patch is created, with a size of approximately $2 \times 400\text{\AA}^2 + X$, depending on the length and orientation of linker (Fig. 3.8). Theoretically, this Affilin[®] library (SPWF library) synthesized using Slonomics[®] technology contains up to 1.5×10^{19} individual variants. In this study, the Affilin[®] library is constructed utilizing Tat-mediated phage display technology, possessing a diversity of 7×10^8 (see Supplementary material 7.2, page 116).

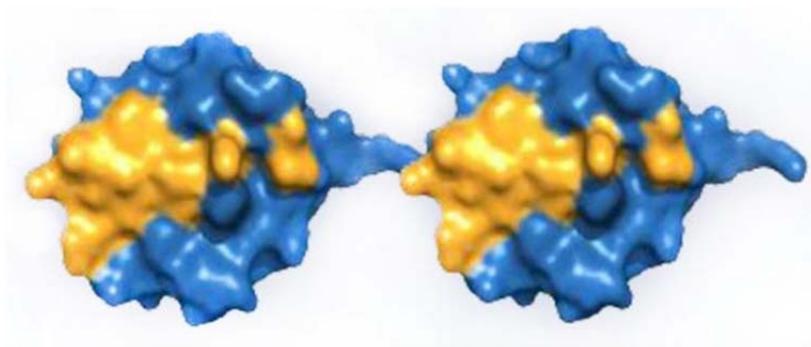


Figure 3.8 Model of dimeric ubiquitin-based Affilin[®] molecule. The randomized residues are shown in yellow. The framework is shown in blue. Figure adapted from Scil Proteins.

3.2.1. Selection against nPAC1-Rs by Tat phage display

To enrich specific binders from the Affilin[®] library for the target protein nPAC1-Rs, four rounds of selection were performed using the strategies as shown in Tab. 2.1 (page 41). In order to decrease the background and favor the Affilin[®] binders with slower off-rates, stringency was increased by decreasing the amount of target protein and increasing washing times between rounds.

In particular, the first selection round was carried out against biotinylated nPAC1-Rs immobilized on a matrix of NeutrAvidin strip. The following three rounds were performed using the M-270 streptavidin beads. Practically, in each selection cycle, phage particles displaying Affilin[®] variants were rescued and purified according to the protocol described in the Material and Methods section 2.2.3.2, page 38. These phages were incubated with target protein biotinylated nPAC1-Rs. The unbound phages were removed during washing step, while the other phages specifically binding to nPAC1-Rs protein were remained and eluted either by competitive elution method (method a), or by acid elution method (method b) as shown in Fig. 2.1, page

42. These bound phages were propagated by re-infecting into exponentially growing *E. coli* cells and subjected to next selection cycle.

To get efficient selection against the nPAC1-Rs protein, two approaches were used to capture the phage particles displaying Affilin[®] binders. One was named “Selection On-Immobilized-Target” (SOIT). In this method, biotinylated nPAC1-Rs protein was firstly immobilized on NeutrAvidin coated wells or streptavidin coated beads by biotin-streptavidin/NeutrAvidin interaction. The phage particles displaying Affilin[®] binders could be further captured due to the interaction between Affilin[®] binder and target nPAC1-Rs molecules. The other was so-called “Selection-In-Solution” (SIS). In this method, the phage particles displaying Affilin[®] binders were firstly bound to biotinylated nPAC1-Rs. These complexes could be sequentially captured by NeutrAvidin or streptavidin coated wells or beads via biotin-streptavidin/NeutrAvidin interaction. Considering that the biotinylation efficiency of nPAC1-Rs protein is about 60%. Using the method of “Selection-In-Solution” may lead to losing of 40% binders, which bind to non-biotinylated nPAC1-Rs and could not be captured by NeutrAvidin and streptavidin. Therefore, both selection methods were used simultaneously in round 2 and 3 before the efficient enrichment of Affilin[®] binders.

In the first two rounds of selection, bound phages were simply eluted by acid elution method. To compare the difference between competitive elution and acid elution, two selections were performed in parallel for both selection rounds 3 and 4 as shown in Fig. 2.1, page 42. All the titration results of input phage and output phage in each selection round were presented in Tab.3.1.

Table 3.1 Phage titers of input and output in all 4 rounds of selection. Two parallel selections (A and B) were performed in both selection rounds 3 and 4, in which elute a and elute b stand for competitive elution method and acid elution method, respectively.

Round		Input (CFU)	Output (CFU)	output/input
1		8.3×10^{12}	5.0×10^5	6.0×10^{-8}
2		3.5×10^{12}	2.4×10^6	6.7×10^{-7}
3	A	5.8×10^{12}	Elute a: 1.2×10^7	4.8×10^{-6}
			Elute b: 1.5×10^7	
	B	4.8×10^{12}	2.0×10^7	4.3×10^{-6}
4	A	1.8×10^{12}	Elute a: 1.2×10^7	9.7×10^{-6}
			Elute b: 6.0×10^6	
	B	1.6×10^{12}	2.8×10^7	1.8×10^{-5}

As shown in Tab. 3.1, for the first selection round, about 10^{13} CFU (colony forming units) of phages rescued from 1 liter culture was used as input. For the subsequent three rounds of selection, more than 10^{12} phages (CFU) amplified from 100 ml culture

were used as input. Output phages increased from 5.0×10^5 in round 1 to about 10^7 in round 4. By comparing the ratio of output/input between round 4 and round 1 using equation 2 (2.2.3.3, page 42), an enrichment factor of approximately 300 times was obtained in the 4th selection round.

3.2.2. Screening by single phage ELISA

To initially identify positive Affilin[®] binders after selection, single colonies randomly picked from five elution pools of selection rounds 3 and 4 were applied to phage rescue in 96-well microtiter plates as described in section 2.2.3.4, page 42.

In order to compare the difference between directly immobilized non-biotinylated nPAC1-Rs and functionally immobilized biotinylated nPAC1-Rs in ELISA experiment, both proteins were coated as target protein on MediSorp plates and NeutrAvidin plate, respectively. To monitor if Affilin[®] variants were successfully displayed on phage particles, anti C-myc antibody was immobilized on MediSorp plates, because the C-terminus of Affilin[®] variants was followed by a fused C-myc tag, which could be recognized by anti C-myc antibody. Several controls were also coated on Nunc MediSorp plates to test the specificity of Affilin[®] variants, including BSA, lysozyme, human serum and the N-terminal domain of parathyroid hormone type 1 receptor (nPTH1-R), where the nPTH1-R protein has approximately 30% similarity to nPAC1-Rs and also belongs to the class B GPCRs family. Phage samples were incubated with these coated proteins. The bound phages were detected by anti-M13 antibody HRP conjugate in a chromogenic reaction according the protocol in section 2.2.3.4, page 42.

As demonstrated in Fig. 3.9, most Affilin[®] variants can specifically bind nPAC1-Rs. The binding to non-biotinylated nPAC1-Rs and biotinylated nPAC1-Rs were highly consistent except variant C9 (Fig. 3.9c), which only binds biotinylated nPAC1-Rs protein, showing that this variant may bind to NeutrAvidin. All variants binding to nPAC1-Rs showed nice binding to anti C-myc antibody, indicating that these Affilin[®] variants had been nicely displayed on M13 phage particles. No binding was detected to negative controls BSA, lysozyme, human serum and nPTH1-R.

Among 285 analyzed clones, 182 clones (~64%) were positive for binding to target protein nPAC1-Rs, as shown in Fig. 3.10. Two pools from round 3 (pools 3A-a and 3A-b+3B-b) presented a high percentage of positive clone (~60%), indicating that the positive binders had efficiently enriched after three rounds of selection. Therefore, a slightly increased positive percentage was observed in all three pools of round 4 (~60-70%). The comparison of positive clone percentage between different pools, which were eluted by competitive elution or acid elution both in rounds 3 and 4, revealed that the pools eluted by competitive elution have a higher positive clone percentage and faster enrichment.

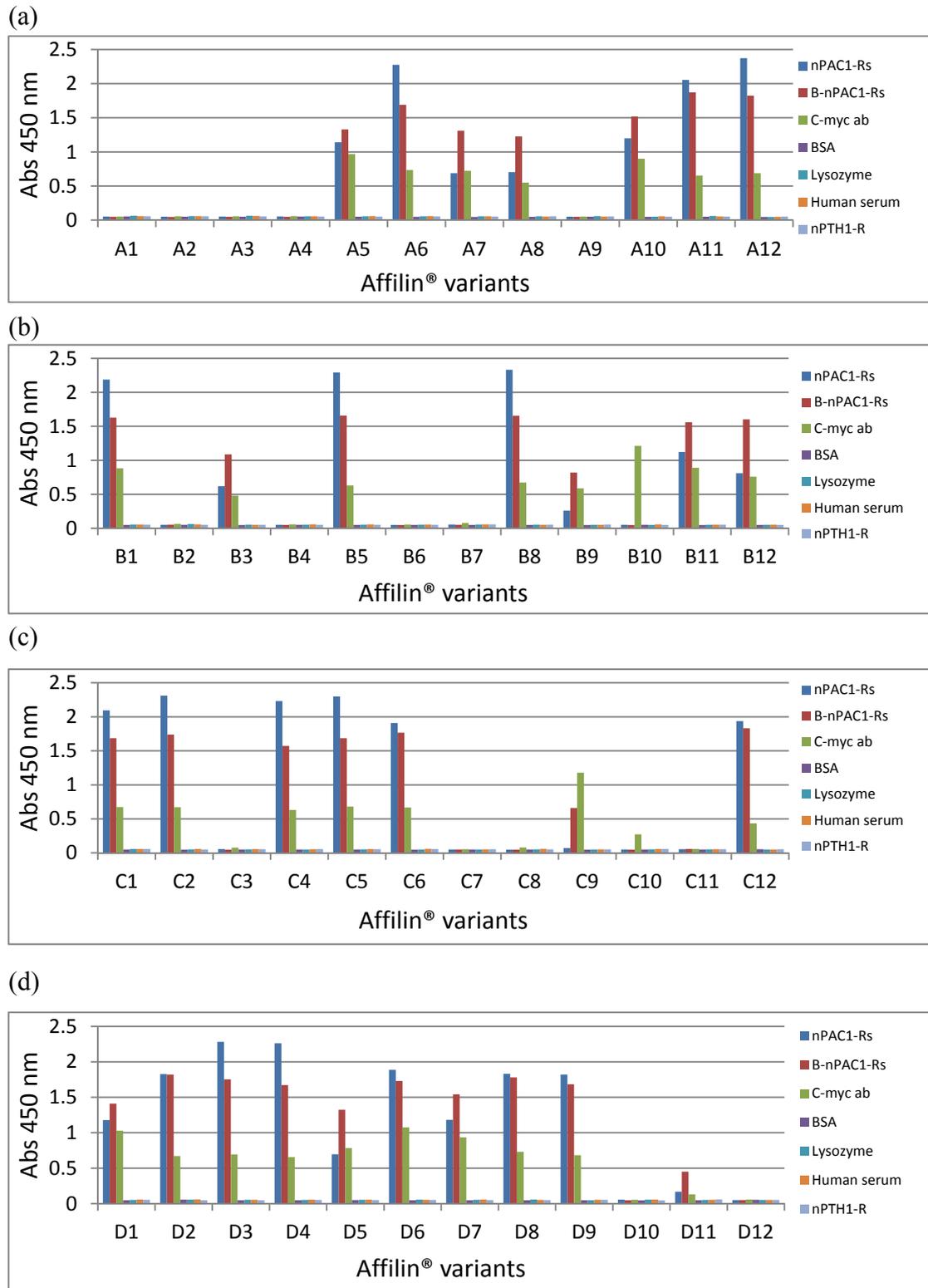


Figure 3.9 Single phase ELISA screening to identify Affilin® binders with binding affinity to nPAC1-Rs. Phage samples prepared from randomly picked colonies (in the elution pools of selection rounds 3 and 4) were analyzed for binding to immobilized nPAC1-Rs, biotinylated nPAC1-Rs (B-nPAC1-Rs) and controls (C-myc antibody, BSA, lysozyme, human serum and nPTH1-R) respectively. (a)-(d) typically exemplify the single phase ELISA results for elution pools.

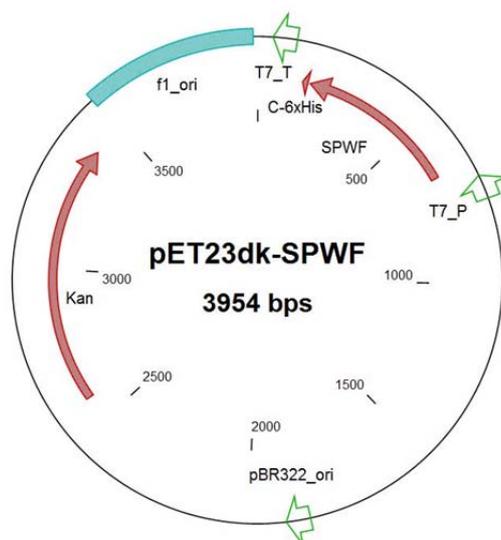


Figure 3.11 Schematic representation of the plasmid pET23dk-SPWF. T7_P, T7 promoter; SPWF, DNA fragment encoding dimeric ubiquitin-based Affilin[®] variant; T7_T, T7 transcription terminator; f1_ori, f1 origin of replication; Kan, kanamycin resistance gene; pBR322_ori, origin of replication from pBR322 plasmid.

To investigate the cell density of the culture in each well of 96-well microtiter plates, OD₆₀₀ was measured on PARADIGM detection platform after expression. In addition, a blank well only containing ZYM-5052 medium with 50 µg/ml kanamycin was setup in each plate to monitor if there was any contamination during inoculation and cultivation. As shown in Fig. 3.12, the OD₆₀₀ results of first plate presented a typical result for all 16 plates. The absorbance for all sample wells was around 2.0, while for the blank well was approximately 0.1, showing nicely cultivated cells without contamination.

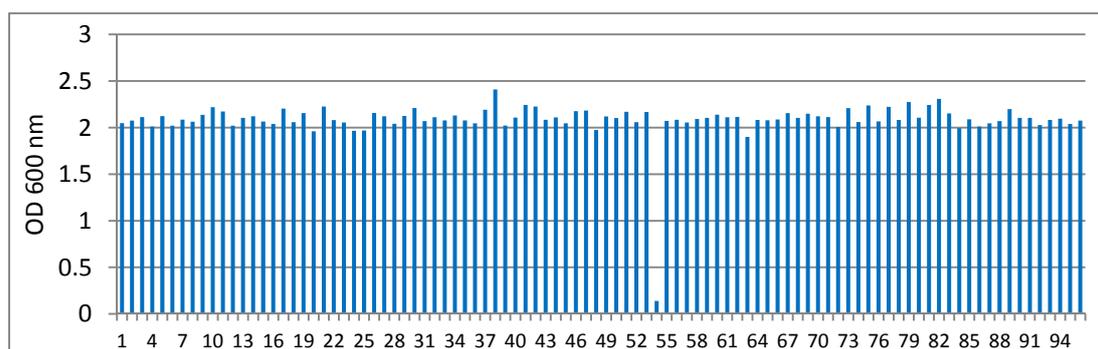


Figure 3.12 OD₆₀₀ measurement result of one plate typically exemplified the cultivation in all 16 plates. The well of No. 54 was filled only with ZYM-5052 medium with 50 µg/ml kanamycin to monitor any possible contamination.

Chapter 3: Results

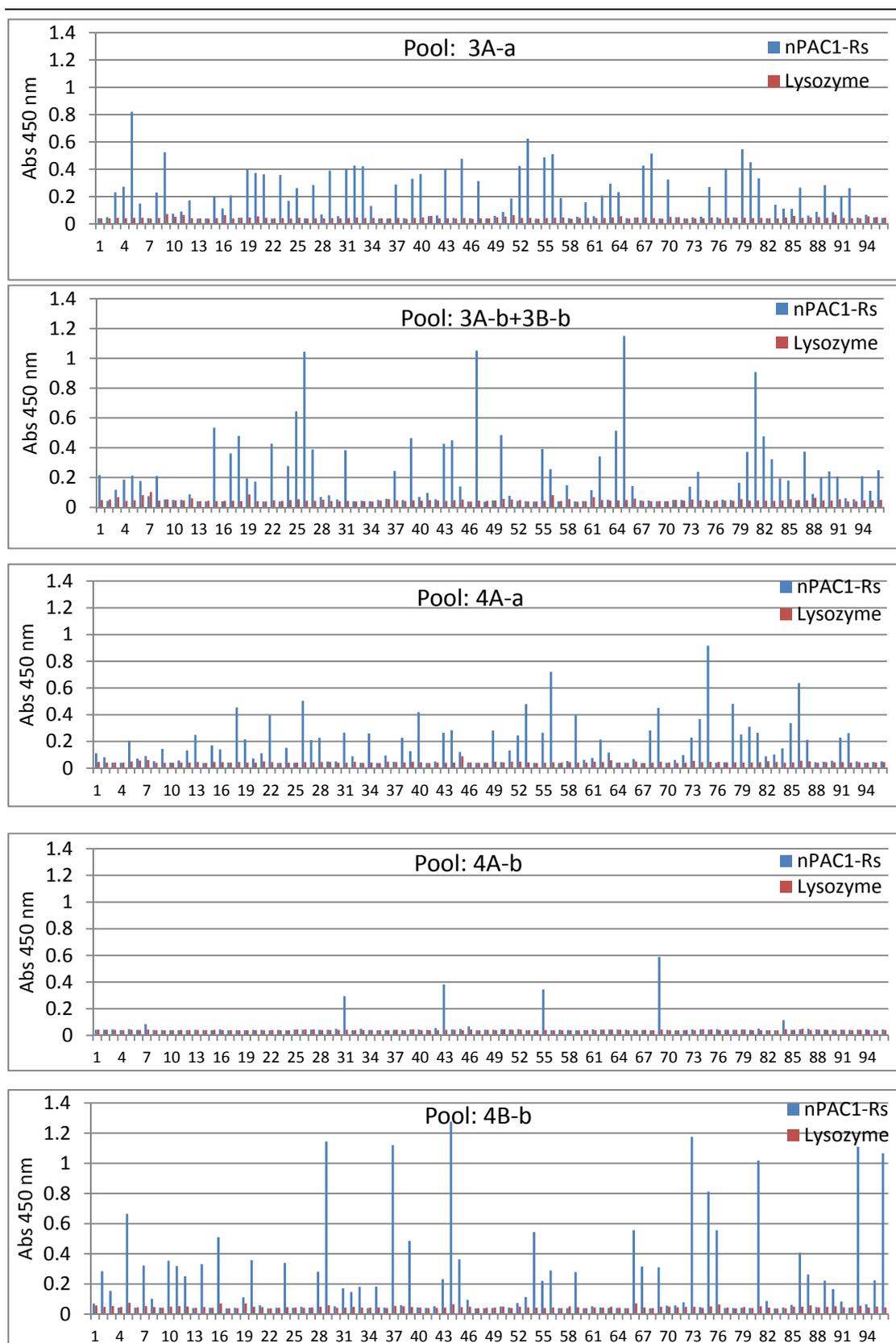


Figure 3.13 Typical Hit-ELISA results for each selected pool in round 3 and 4. Blue bars show the binding signals of Affilin[®] variants to target protein nPAC1-Rs, while the red bars present the binding signals to control protein lysozyme.

Considering that positive binders were efficiently enriched in selection rounds 3 according to single phage ELISA results, the majority of analyzed clones were originated from two elution pools of round 3. In short, 552 clones were picked from elution pool 3A-a, 368 from elution pool 3A-b+3B-b, 276 from elution pool 4A-a, 184 from elution pool 4A-b and 92 from elution pool 4B-b. After expression, the cell lysate was incubated with target protein nPAC1-Rs and the control lysozyme both immobilized on MediSorp plates. The bound Affilin[®] variants were detected by anti-Ubi-Fab-POD antibody as described in section 2.2.3.5, page 43. For each elution pool, one plate of Hit-ELISA result was exemplified in Fig. 3.13 to typically present the binding signals to target and control. The two pools (3A-a and 4A-a) eluted by competitive elution method (a) showed a higher hit percentage and lower binding signal intensity when comparing to the other three pools eluted by acid elution method (b).

3.2.4. Data analysis of the screening by Hit-ELISA

As charted in Fig. 3.14, to simply analyze all Hit-ELISA data, the binding signal intensity was initially evaluated by absorbance data ($A_{450\text{nm}}$) of target nPAC1-Rs subtracted by the control lysozyme as X axis, while the binding specificity was evaluated by using the ratio of absorbance data ($A_{450\text{nm}}$) between target nPAC1-Rs and control lysozyme as Y axis. The red points, defined as hits, gave significant signals with $A_{450\text{nm}} \geq 0.3$ and specific binding to target nPAC1-Rs with a ratio ≥ 8 .

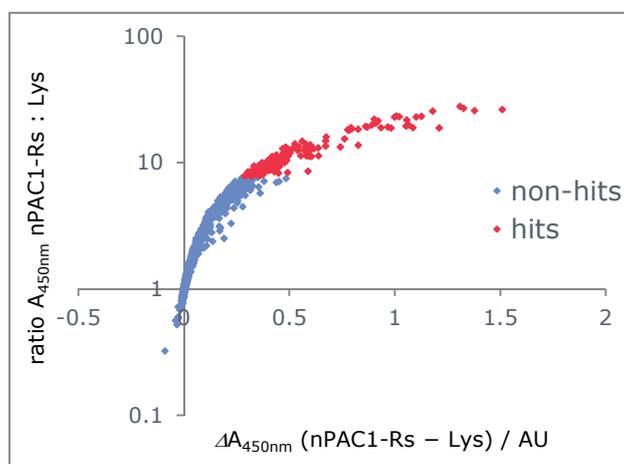


Figure 3.14 Specificity vs. signal intensity in high throughput Hit-ELISA results. High throughput Hit-ELISA for Affilin[®] variants randomly picked from 5 elution pools of selection rounds 3 and 4 was performed against target protein nPAC1-Rs and negative control lysozyme.

Previously data analysis provided initial evaluation of the signal intensity and specificity for Hit-ELISA results, but ignored the deviations between different samples and different plates. A more precise processing and ranking might be necessary for identification of most promising Affilin[®] binding proteins. Therefore, the Hit-ELISA data were further processed in two approaches: (i) standard hit analysis and (ii) statistically normalized hit analysis.

(i) In standard hit analysis, the equation for calculating the value of standard ratio T:C was as follows:

$$\text{Std.ratioT:C} = \text{TargetA}_{450\text{nm}} / \text{ControlA}_{450\text{nm}} \quad \text{Equation 7}$$

where the $\text{TargetA}_{450\text{nm}}$ and $\text{ControlA}_{450\text{nm}}$ are the absorbance values of a certain Affilin[®] variant against target nPAC1-Rs and control lysozyme, respectively.

In order to correct the signal deviation between all samples, a value of standard Z-scoreT:C was calculated using equation:

$$\text{Std.Z-scoreT:C} = [(\text{Std.ratioT:C}) - \text{Mean}(\text{Std.ratioT:C})] / \text{SD}(\text{Std.ratioT:C}) \quad \text{Equation 8}$$

where $\text{Mean}(\text{Std.ratioT:C})$ is a mean value of Std.ratioT:C, $\text{SD}(\text{Std.ratioT:C})$ is a standard deviation value of Std.ratioT:C.

A ranking list was sequentially created according to the Std.Z-scoreT:C value. The standard rank of the Affilin[®] variant with highest Std.Z-scoreT:C value was defined to be No. 1. Its corresponding score was regarded as 1000. Therefore, the score of standard hit ratio (HR) was converted from the rank by using equation:

$$\text{Score Std.HR} = 1000 - [\text{INT}(\text{Std.rank} / 1.472)] \quad \text{Equation 9}$$

(INT means taking the integer portion of a number; 1.472 comes from 1472 divided by 1000; totally 1472 clones are analyzed).

(ii) In normalized hit analysis, the value of normalized ratio $T_N:C_N$ was calculated using equation to correct the signal deviation between target plate and control plate:

$$\text{Norm.ratioT}_N:C_N = [\text{TargetA}_{450\text{nm}} / \text{Mean}(\text{TargetA}_{450\text{nm}})] / [\text{ControlA}_{450\text{nm}} / \text{Mean}(\text{ControlA}_{450\text{nm}})] \quad \text{Equation 10}$$

To correct the signal deviation between all samples, the Normalized Z-score was calculated using equation:

$$\text{Norm.Z-scoreT}_N:C_N = [(\text{Norm.ratioT}_N:C_N) - \text{Mean}(\text{Norm.ratioT}_N:C_N)] / \text{SD}(\text{Norm.ratioT}_N:C_N) \quad \text{Equation 11}$$

A ranking list was created according to the Norm.Z-score $T_N:C_N$ value. The normalized rank of the Affilin[®] variant with highest Norm.Z-score $T_N:C_N$ value was defined to be No. 1. Its corresponding score was also regarded as 1000. Therefore, the score of normalized hit ratio (HR) was converted from the rank by using equation:

$$\text{Score Norm.HR} = 1000 - [\text{INT}(\text{Norm.rank} / 1.472)] \quad \text{Equation 12}$$

(functions such as Mean, SD and INT refer to (i) standard hit analysis)

The final rank score was calculated using following equation:

$$\text{Rank score} = (\text{Score Std.HR} + \text{Score Norm.HR}) / 2 \quad \text{Equation 13}$$

According to the final rank score, the best 174 clones were picked out for sequencing and further analysis.

3.2.5. Sequence analysis of Affilin[®] variants

The best 174 hits screened by Hit-ELISA experiment were sequenced and aligned as described in section 2.2.1.11, page 37. DNA sequence alignment of all 174 clones resulted in 59 different Affilin[®] variants, showing a very high diversity after selection. The common motifs among all sequences were revealed by a nearest neighbor

clustering method as presented in Fig. 3.15, 3.16 and 3.17. All these sequences were clearly different from the wild type human ubiquitin molecule without any amino acid residue remaining conserved among all 15 randomized positions.

In Fig 3.15, amino acid sequences in randomized positions of 25 Affilin[®] hits were given, whereas the other 34 shown slight differences in sequence and lower scores in Hit-ELISA were not subjected to further analysis. These sequences of 25 Affilin[®] hits can be divided into four clusters according to the position of randomized amino acid residues. The identical sequences were labeled with same color within each cluster. In the first cluster, including randomized positions 2, 4 and 6, three predominant consensus sequences were detected as “TNI”, “KWF” and “DTI”. Five sequences “VSHPN”, “PDVER”, “RRANV”, “WPHDV” and “HRNKN” were presented dominantly in the second cluster (positions 62, 63, 64, 65 and 66). The preponderant sequences were “DD”, “DN”, “DA” and “SR” in the third cluster (positions 6’ and 8’), as well as “KPPPF”, “RPPGW”, “RHPDW” and “GEWNF” in the fourth cluster (positions 62’, 63’, 64’, 65’ and 66’).

One Affilin[®] variant, donated the hit No. 6-P1F11, dominated by occurring in 92 different clones from the best 174 sequenced hits and being present in all five elution pools. The variants No.15-P2A08, No.16-P2C05 and No. 14-P2A12 appeared 6, 6 and 7 times respectively in the best 174 clones (Fig 3.15). Enrichment of these Affilin[®] binders as well as the high hits percentage presented in both single phage ELISA and Hit-ELISA screenings indicated that the selection of Affilin[®] library against the nPAC1-Rs target protein was efficient.

High proline content was observed in many sequenced Affilin[®] variants (Fig 3.15). Proline was found in randomized positions 62-65 and 62’-65’, particularly in three positions 63’, 64’ and 65’ of SPF domain with high frequency. Generally proline is regarded as a structural disruptor in the middle of regular secondary structure elements such as α -helices and β -sheets, due to the exceptional conformational rigidity of side chain. The randomized positions 64’ and 65’ are localized at the beginning of the C-terminal beta sheet strand. Therefore it might be possible that this high proline content may negatively affect the folding of the second domain-SPF domain.

Compared to the wild type dimeric human ubiquitin, the selected Affilin[®] binders presented several hydrophobic amino acid substitutions with high frequency. The randomization positions of wild type dimeric human ubiquitin molecule contain two hydrophobic amino acid residues, Phe (F) in position 2 of SPW domain and Leu (L) in position 8’ of SPF domain, according to the hydropathy index value of amino acid residues (Kyte *et al.*, 1982). As shown in Fig. 3.16, the hydrophobic amino acid residues (labeled in yellow) frequently occurred in positions 6, 62, 64 and 66 of SPW domain as well as in positions 6’, 8’, 63’ and 66’ of SPF domain. These hydrophobic substitutions could mediate the binding to target protein via hydrophobic interaction.

Chapter 3: Results

Randomized position	SPW domain					Linker	SPF domain										
	Cluster 1			Cluster 2			Cluster 3		Cluster 4								
	2	4	6	62	63		64	65	66	6'	8'	62'	63'	64'	65'	66'	
Wild type ubiquitin	Q	F	K	Q	K	E	S	T			K	L	Q	K	E	S	T
1 2H03-1	T	N	I	P	D	V	E	R			D	D	P	R	G	T	A
2 2H03-2	T	N	I	P	D	V	E	R			D	D	T	D	P	P	Y
3 2H07-1	T	N	I	R	R	A	N	V			D	D	G	E	W	N	F
4 P1A02	T	N	I	R	R	A	N	V			D	D	R	P	P	G	W
5 2H07-2	T	N	I	R	R	A	N	V			D	D	R	V	P	P	W
6 P1F11	T	N	I	V	S	H	P	N			S	R	K	P	P	P	F
7 P1D10	T	N	I	V	S	H	P	N			D	N	K	P	P	P	F
8 P1G07	T	N	I	W	P	H	D	V			D	N	K	P	P	P	F
9 16H09-1	T	N	I	W	P	H	D	V			D	A	R	P	P	G	W
10 P8B07	T	N	I	A	S	N	N	W			D	A	R	P	P	G	W
11 P1H11	T	N	I	A	P	N	T	K			I	K	R	H	P	D	W
12 1G07-2	T	N	I	G	W	K	T	D			G	W	R	H	P	D	W
13 P2H09	T	N	I	H	R	N	K	N			S	R	K	P	P	P	F
14 P2A12	T	N	I	H	R	N	K	N			D	A	K	P	P	P	F
15 P2A08	T	N	I	H	R	N	K	N			D	A	R	P	P	G	W
16 P2C05	K	W	F	H	R	N	K	N			D	A	R	P	P	G	W
17 P16A12	K	W	F	H	R	N	K	N			T	K	R	H	P	D	W
18 P16H06	K	W	F	H	R	N	K	N			D	A	R	L	T	R	P
19 P16C08	K	W	F	H	R	N	K	N			D	T	R	L	T	R	P
20 P15C12	D	T	I	P	I	G	E	D			W	Q	A	D	V	P	W
21 2F02-1	D	T	I	E	Q	R	D	T			I	N	R	M	P	P	W
22 P12A10	Q	W	V	A	H	V	R	K			N	S	R	N	P	N	W
23 P16H03	Y	W	Y	A	P	N	T	K			I	K	R	H	P	D	W
24 P12C08	I	T	V	W	P	H	D	V			L	S	G	E	W	N	F
25 P12C04	K	H	V	G	N	Q	R	W			M	M	G	I	H	K	K

Figure 3.15 Amino acid sequences in randomized positions aligned and clustered for 25 Affilin® hits and wild type human ubiquitin. Four clusters are compartmentalized according to the position of randomized amino acid residues, leading to the first cluster (positions 2, 4 and 6), the second cluster (positions 62, 63, 64, 65 and 66), the third cluster (positions 6' and 8') and the fourth cluster (positions 62', 63', 64', 65' and 66'). The identical sequences are labeled in same color within each cluster.

Randomized position	SPW domain					Linker	SPF domain										
	Cluster 1			Cluster 2			Cluster 3		Cluster 4								
	2	4	6	62	63		64	65	66	6'	8'	62'	63'	64'	65'	66'	
Wild type Ubiquitin	Q	F	K	Q	K	E	S	T			K	L	Q	K	E	S	T
1 2H03-1	T	N	I	P	D	V	E	R			D	D	P	R	G	T	A
2 2H03-2	T	N	I	P	D	V	E	R			D	D	T	D	P	P	Y
3 2H07-1	T	N	I	R	R	A	N	V			D	D	G	E	W	N	F
4 P1A02	T	N	I	R	R	A	N	V			D	D	R	P	P	G	W
5 2H07-2	T	N	I	R	R	A	N	V			D	D	R	V	P	P	W
6 P1F11	T	N	I	V	S	H	P	N			S	R	K	P	P	P	F
7 P1D10	T	N	I	V	S	H	P	N			D	N	K	P	P	P	F
8 P1G07	T	N	I	W	P	H	D	V			D	N	K	P	P	P	F
9 16H09-1	T	N	I	W	P	H	D	V			D	A	R	P	P	G	W
10 P8B07	T	N	I	A	S	N	N	W			D	A	R	P	P	G	W
11 P1H11	T	N	I	A	P	N	T	K			I	K	R	H	P	D	W
12 1G07-2	T	N	I	G	W	K	T	D			G	W	R	H	P	D	W
13 P2H09	T	N	I	H	R	N	K	N			S	R	K	P	P	P	F
14 P2A12	T	N	I	H	R	N	K	N			D	A	K	P	P	P	F
15 P2A08	T	N	I	H	R	N	K	N			D	A	R	P	P	G	W
16 P2C05	K	W	F	H	R	N	K	N			D	A	R	P	P	G	W
17 P16A12	K	W	F	H	R	N	K	N			T	K	R	H	P	D	W
18 P16H06	K	W	F	H	R	N	K	N			D	A	R	L	T	R	P
19 P16C08	K	W	F	H	R	N	K	N			D	T	R	L	T	R	P
20 P15C12	D	T	I	P	I	G	E	D			W	Q	A	D	V	P	W
21 2F02-1	D	T	I	E	Q	R	D	T			I	N	R	M	P	P	W
22 P12A10	Q	W	V	A	H	V	R	K			N	S	R	N	P	N	W
23 P16H03	Y	W	Y	A	P	N	T	K			I	K	R	H	P	D	W
24 P12C08	I	T	V	W	P	H	D	V			L	S	G	E	W	N	F
25 P12C04	K	H	V	G	N	Q	R	W			M	M	G	I	H	K	K

Figure 3.16 Amino acid properties in randomized positions for wild type human ubiquitin and 25 Affilin® hits. Hydrophobic, basic and acidic amino acid residues are highlighted in yellow, blue and red respectively.

Additionally, to collect sequence information for all five elution pools, including the sequences of both hits and flops, one 96-well plate of samples for each pool was sequenced respectively. The amino acid sequences of randomized positions of 30 Affilin[®] flops were demonstrated in Fig. 3.17. These flops showed equally low binding signals to both target nPAC1-Rs and negative control lysozyme with a value of approximately 0.05 (Abs 450 nm). As color labeled in Fig. 3.17, one or more consensus sequences were presented in the same clusters as Affilin[®] hits shown in Fig. 3.15. For instance, two flops No.19-P12D10 and No.20-P8B04 presented highest consensus sequences among all the Affilin[®] flops when comparing with Affilin[®] hit No.6-P1F11. They had only three different residues in the first cluster among all 15 randomized positions. The possible reasons for this phenomenon will be discussed in section 4.5.1, page 93.

		SPW domain						Linker	SPF domain							
		Cluster 1			Cluster 2				Cluster 3		Cluster 4					
Randomized position		2	4	6	62	63	64	65	66	6'	8'	62'	63'	64'	65'	66'
	Wild type ubiquitin	Q	F	K	Q	K	E	S	T							
1	P16G07	T	N	I	V	S	H	P	N	S	R	Q	K	S	F	M
2	P8D03	T	N	I	V	S	H	P	N	P	Q	V	D	P	P	W
3	P12E04	T	N	I	V	S	H	P	N	D	N	P	Y	W	W	Q
4	P12A05	T	N	I	P	H	P	R	K	R	W	K	P	P	P	F
5	P12C03	T	N	I	N	L	G	A	P	I	P	P	K	L	I	F
6	P8C09	T	N	I	H	S	M	Q	T	R	T	P	P	V	W	K
7	P12F07	T	N	I	A	P	N	T	K	I	K	R	I	L	L	V
8	P3A05	T	N	I	A	D	Y	D	I	D	N	P	Y	W	W	Q
9	P3H09	T	N	I	W	P	H	D	V	K	T	K	P	P	P	F
10	P3A02	T	N	I	W	P	H	D	V	L	S	G	K	P	I	T
11	P12B05	T	S	T	W	P	H	D	V	L	S	A	P	T	R	S
12	P3D09	T	S	I	E	H	S	I	A	Q	V	Y	S	H	E	D
13	P8A02	T	W	F	A	P	W	P	H	F	T	P	E	A	H	H
14	P3A03	W	H	W	W	E	Q	D	T	N	S	S	V	T	Q	P
15	P12G02	V	W	F	D	N	L	P	I	D	E	R	Y	P	P	F
16	P3E05	L	T	F	W	P	W	D	V	L	D	R	D	P	H	W
17	P12B08	I	T	V	W	P	H	D	V	L	S	Q	V	D	T	D
18	P8D10	S	Y	S	V	W	W	D	V	M	M	K	P	P	P	F
19	P12D10	I	T	V	V	S	H	P	N	S	R	K	P	P	P	F
20	P8B04	G	V	L	V	S	H	P	N	S	R	K	P	P	P	F
21	P16H04	K	W	F	H	L	L	T	V	D	A	P	F	E	L	T
22	P3C06	I	K	F	G	P	R	E	H	Q	S	H	Q	W	P	P
23	P16E02	N	W	F	G	D	G	S	P	S	R	A	D	V	P	W
24	P12C09	G	I	A	E	Q	R	D	T	I	N	K	P	P	P	F
25	P8E11	E	M	W	E	P	D	W	H	S	E	R	G	W	S	G
26	P12A01	D	T	I	R	R	A	N	V	S	R	A	G	W	F	H
27	P15A08	D	T	I	V	S	H	P	N	W	Q	A	D	V	P	W
28	P16F03	D	G	G	V	M	G	P	M	P	S	R	H	P	D	W
29	P12C06	A	T	K	G	P	R	E	H	Q	S	H	Q	W	P	P
30	P3B12	A	E	S	W	P	H	D	V	L	S	G	E	W	N	F

Figure 3.17 Amino acid sequences in randomized positions aligned and clustered for wild type human ubiquitin and 30 Affilin[®] flops. Clusters are made as the same rules as in Fig. 3.15, page 67. The identical sequences are labeled in same color within each cluster.

3.3. Characterization of selected Affilin[®] binding proteins

3.3.1. Analysis of expression and solubility

To analyze the expression level and solubility of Affilin[®] hits after selection, some

Affilin[®] binders presenting favorable binding abilities in Hit-ELISA with interesting sequences were expressed and analyzed by SDS-PAGE as described in the section 2.2.5.3, page 45. Two different expression temperatures were compared as shown in Fig. 3.18. The molecular weight of Affilin[®] with C-terminal 6×His-tag is approximately 19 kDa. For most variants, the expression level at 30 °C and 37 °C was almost equal. Therefore, expression at 30 °C was used for following expression.

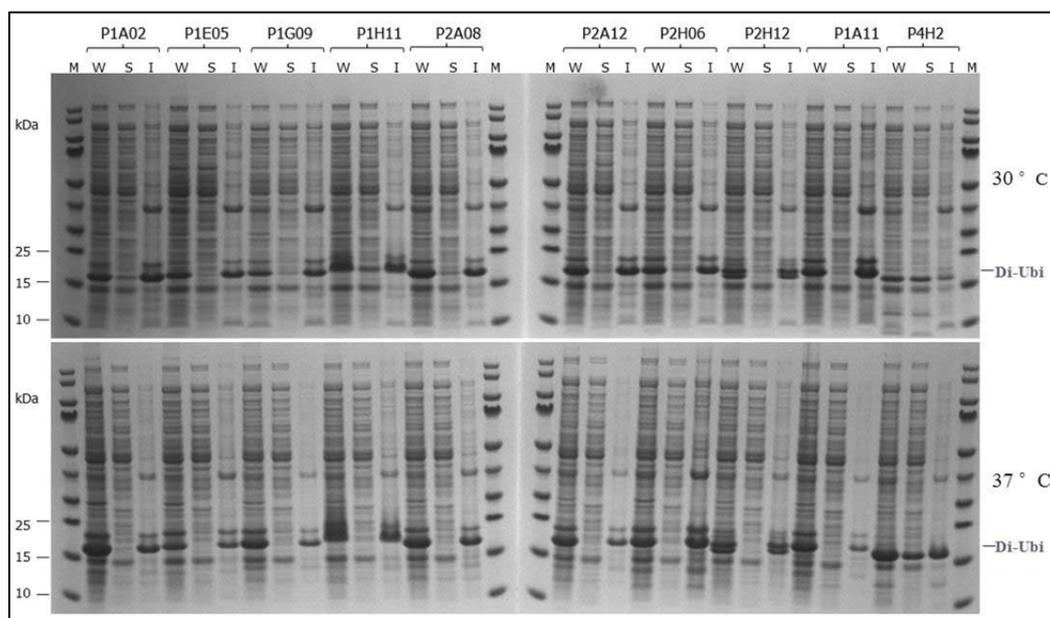
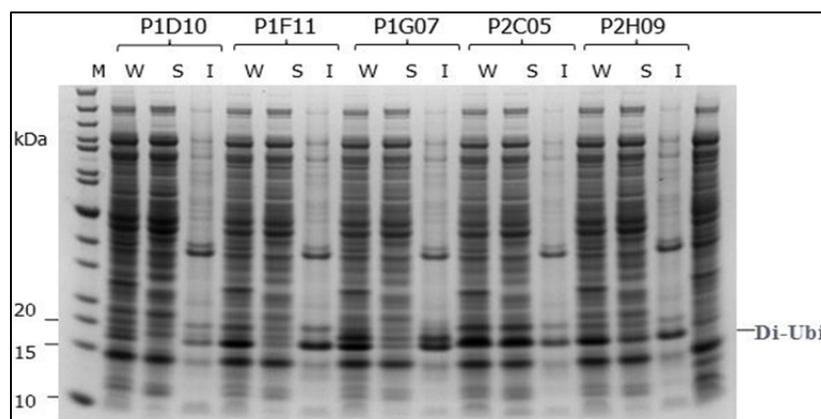


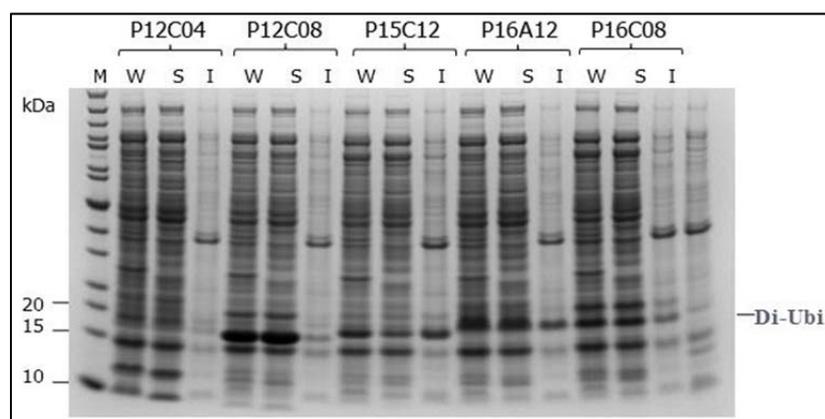
Figure 3.18 Expression of selected Affilin[®] binders at 30 °C and 37 °C. Affilin[®] binders were expressed in autoinduction medium ZYM-5052 at 30 °C (upper part) and 37 °C (lower part), respectively. M, PageRuler[™] prestained protein ladder; W, whole cells extraction; S, soluble fraction; I, insoluble fraction.

As illustrated in Fig. 3.19, to extensively investigate the expression level and solubility of selected Affilin[®] binders, semiquantitative analysis was introduced. For example, for variant P12C08 (Fig. 3.19b), the expression level was defined to be 10 as the highest, while the solubility was considered as 100%. In contrast, for the variant P2C05 (Fig. 3.19a), the expression level and solubility were considered as 8 and 70%, respectively.

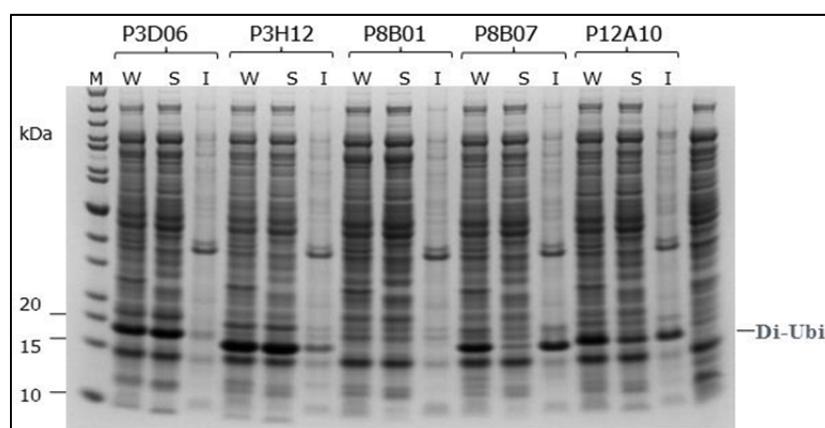
(a)



(b)



(c)



(d)

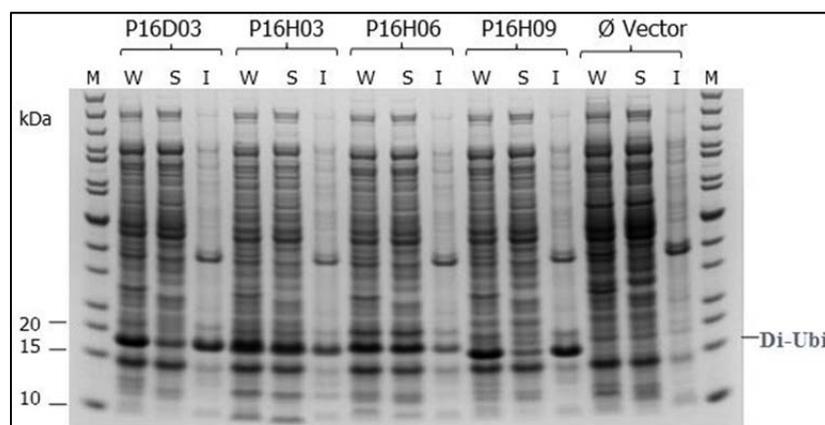


Figure 3.19 Expression and solubility analysis of selected Affilin[®] binders by SDS-PAGE. M, PageRuler[™] unstained protein ladder; W, whole cells extraction; S, soluble fraction; I, insoluble fraction. (a)-(d) typically exemplify the expression and solubility analysis results of selected Affilin[®] binders.

According to the analysis data and definition above, the expression level and solubility of 25 selected Affilin[®] binders as well as a dimeric wild type human ubiquitin were summarized in Tab. 3.2.

Table 3.2 Expression level and solubility of 25 selected Affilin[®] binders as well as a dimeric wild type human ubiquitin. N/A means not available.

Randomized position	2	4	6	62	63	64	65	66	6'	8'	62'	63'	64'	65'	66'	Solubility	Expression level	
Wild type ubiquitin	Q	F	K	Q	K	E	S	T	K	L	Q	K	E	S	T	100%	10	
1	2H03-1	T	N	I	P	D	V	E	R	D	D	P	R	G	T	A	N/A	N/A
2	2H03-2	T	N	I	P	D	V	E	R	D	D	T	D	P	P	Y	N/A	N/A
3	2H07-1	T	N	I	R	R	A	N	V	D	D	G	E	W	N	F	N/A	N/A
4	P1A02	T	N	I	R	R	A	N	V	D	D	R	P	P	G	W	5%	8
5	2H07-2	T	N	I	R	R	A	N	V	D	D	R	V	P	P	W	N/A	N/A
6	P1F11	T	N	I	V	S	H	P	N	S	R	K	P	P	P	F	5%	8
7	P1D10	T	N	I	V	S	H	P	N	D	N	K	P	P	P	F	5%	5
8	P1G07	T	N	I	W	P	H	D	V	D	N	K	P	P	P	F	5%	8
9	16H09-1	T	N	I	W	P	H	D	V	D	A	R	P	P	G	W	N/A	N/A
10	P8B07	T	N	I	A	S	N	N	W	D	A	R	P	P	G	W	10%	8
11	P1H11	T	N	I	A	P	N	T	K	I	K	R	H	P	D	W	20%	8
12	1G07-2	T	N	I	G	W	K	T	D	G	W	R	H	P	D	W	N/A	N/A
13	P2H09	T	N	I	H	R	N	K	N	S	R	K	P	P	P	F	30%	8
14	P2A12	T	N	I	H	R	N	K	N	D	A	K	P	P	P	F	5%	10
15	P2A08	T	N	I	H	R	N	K	N	D	A	R	P	P	G	W	5%	10
16	P2C05	K	W	F	H	R	N	K	N	D	A	R	P	P	G	W	70%	10
17	P16A12	K	W	F	H	R	N	K	N	T	K	R	H	P	D	W	60%	10
18	P16H06	K	W	F	H	R	N	K	N	D	A	R	L	T	R	P	80%	10
19	P16C08	K	W	F	H	R	N	K	N	D	T	R	L	T	R	P	60%	8
20	P15C12	D	T	I	P	I	G	E	D	W	Q	A	D	V	P	W	30%	8
21	2F02-1	D	T	I	E	Q	R	D	T	I	N	R	M	P	P	W	50%	4
22	P12A10	Q	W	V	A	H	V	R	K	N	S	R	N	P	N	W	30%	8
23	P16H03	Y	W	Y	A	P	N	T	K	I	K	R	H	P	D	W	70%	10
24	P12C08	I	T	V	W	P	H	D	V	L	S	G	E	W	N	F	100%	10
25	P12C04	K	H	V	G	N	Q	R	W	M	M	G	I	H	K	K	50%	5

Comprehensively considering the sequence, binding ability, expression level and solubility, 25 Affilin[®] binders shown in Fig. 3.15 (page 67) were finally subjected to purification and further characterization.

3.3.2. Purification of Affilin[®] binding proteins

To obtain proteins with high purity and quantity, the expression of Affilin[®] binders was scaled up to 1 liter culture in baffled flasks and performed in *E. coli* NovaBlue(DE3) cells as described in section 2.2.4.1. The proteins were purified from soluble fraction by IMAC and further by size exclusion chromatography (section 2.2.4.2, page 44). 19 Affilin[®] binders were successfully purified with protein yields varied from 0.2 to 135 mg per liter culture, depending on the cell mass, expression level and solubility (Tab. 3.3). The protein yields of 6 Affilin[®] binders was defined to be “0” because they are almost completely insoluble and failed to be purified.

The chromatograms and SDS-PAGE of four Affilin[®] binders were illustrated in Fig. 3.20 to exemplify the typical purification of Affilin[®] binders with high solubility. These four variants showed high solubility of 60%-80% and high expression level in previous analysis. The peaks in the retention volume of around -20 ml (x-axis) indicated the protein elution from the IMAC purification. The peak 3 (P3) indicated monomer Affilin[®] proteins as expected, while peak 1 (P1), peak 2 (P2) and peak 4 (P4) indicated presumably aggregates, dimer Affilin[®] protein and digested Affilin[®] by cytosolic protease, respectively. Various fractions from different elution peaks were

analyzed by SDS-PAGE as shown below. Comparison of the protein amounts in each elution peak clearly revealed that the monomer Affilin® is the main species in the soluble fractions. The protein yields are from 21 mg to 62.6 mg per liter culture after purification.

Table 3.3 Protein yields per liter culture of 25 selected Affilin® binders.

Affilin variant			Protein Yield(mg/L)			Affilin variant			Protein Yield(mg/L)		
1	2H03-1	12.6	14	P2A12	3.5						
2	2H03-2	8.7	15	P2A08	0.26						
3	2H07-1	2.6	16	P2C05	39						
4	P1A02	0.96	17	P16A12	62.6						
5	2H07-2	0	18	P16H06	44.3						
6	P1F11	0.45	19	P16C08	29.4						
7	P1D10	0	20	P15C12	4.6						
8	P1G07	2.4	21	2F02-1	0						
9	16H09-1	1.48	22	P12A10	23.7						
10	P8B07	0	23	P16H03	21						
11	P1H11	0	24	P12C08	135						
12	1G07-2	0.2	25	P12C04	17.6						
13	P2H09	0									

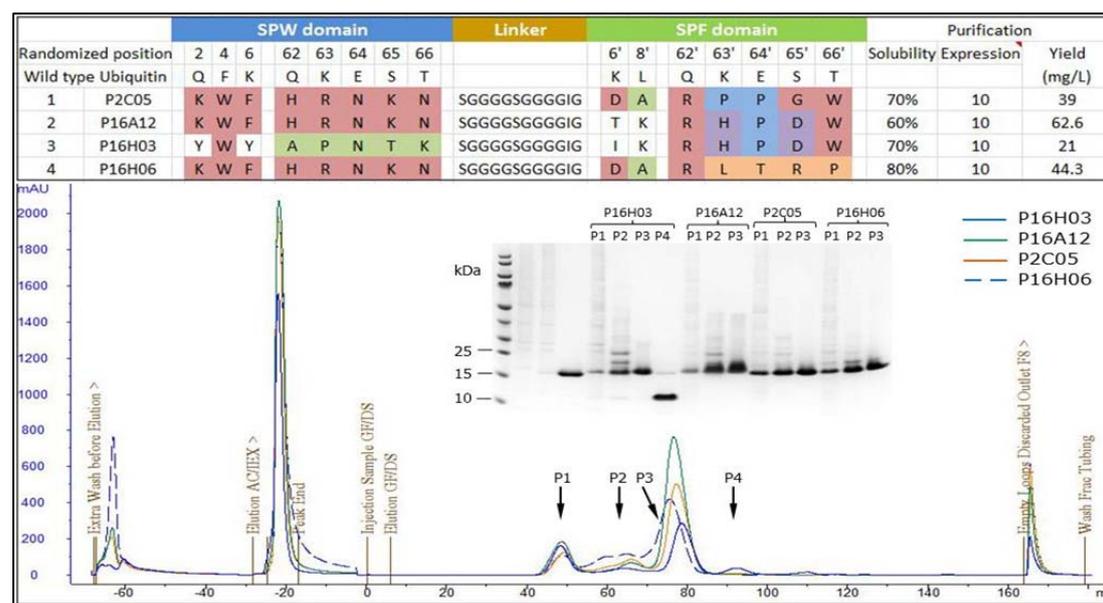


Figure 3.20 Chromatograms of four Affilin® binders illustrating the purification of highly soluble Affilin® proteins via IMAC and subsequent size exclusion chromatography. Basic information (upper part) and chromatograms (lower part) of Affilin® binders P16H03, P16A12, P2C05 as well as P16H06 are shown. In chromatograms, plotted on the x-axis is the retention volume in unit of ml (retention volume -70-0 ml represents the IMAC chromatogram and 0-160 ml represents the size exclusion chromatogram. Peak at -20 ml represents the elution from IMAC.) and the y-axis is the absorbance at 280 nm in unit of mAU. Black arrows (P1, P2, P3 and P4) indicate the elution peaks from HiLoad16/60 Superdex 75 pg column. SDS-PAGE demonstrates different elution fractions of all four Affilin® binders.

However, the following chromatograms and SDS-PAGE of another four Affilin[®] binders were illustrated in Fig. 3.21 to present the typical purification of Affilin[®] binders with low solubility. These four variants also showed high expression level but with low solubility of 5%-30% in previous analysis. The peaks in the retention volume of around -20 ml (x-axis) indicated the protein elution from the IMAC purification. The P3 indicated monomer Affilin[®] proteins as expected. The P1 and P2 presumably indicated aggregates and dimer Affilin[®] proteins, respectively. Various fractions from different elution peaks were analyzed by SDS-PAGE as shown below. In this case, the majority of soluble fraction turned into aggregates during purification, leading to the poor protein yields from 0 to 0.96 mg per liter cell culture.

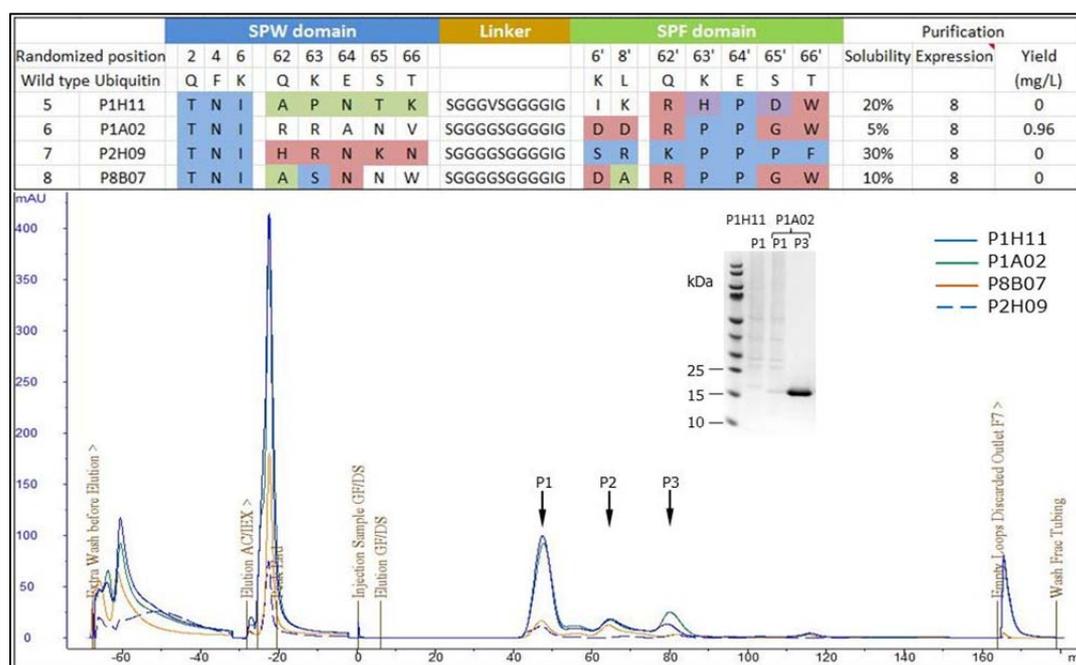


Figure 3.21 Chromatograms of four Affilin[®] binders illustrating the purification of Affilin[®] proteins with low solubility via IMAC and subsequent size exclusion chromatography. The basic information (upper part) and chromatograms (lower part) of Affilin[®] binders P1H11, P1A02, P8B07 as well as P2H09 are shown. In chromatograms, plotted on the x-axis is the retention volume in unit of ml (the retention volume -70-0 ml represents the IMAC chromatogram and 0-160 ml represents size exclusion chromatogram. Peak at -20 ml represents the elution from IMAC.) and the y-axis is the absorbance at 280 nm in unit of mAU. Black arrows (P1, P2 and P3) indicate the elution peaks from HiLoad16/60 Superdex 75 pg column. SDS-PAGE presents the different elution fractions of Affilin[®] binders.

3.3.3. Thermostability

The thermostability of selected Affilin[®] binders was measured by DSF experiments. Due to the high throughput capability and fast screening, DSF is usually employed as a high-throughput screening assay for initial assessment. The typical DSF profile generally consists of a sharp sigmoid-like increase to the maximum level, followed by

a decrease in fluorescence intensity. The thermograms of four Affilin[®] binders were exemplified in Fig. 3.22, as well as the thermogram of wild type dimeric ubiquitin molecule. Interestingly, the wild type dimeric ubiquitin melt with two transitions, probably indicating the partially and fully unfolded status. The midpoint temperature of thermal unfolding (apparent melting temperature) for all Affilin[®] binders was presented in Fig. 3.22c, varying from 50 to 65 °C. Compared to the apparent T_m value (over 90 °C) of wild type dimeric ubiquitin determined by CD spectroscopy, the T_m values of all Affilin[®] binders measured by DSF experiment were much lower and unrealistic. The possible reason will be discussed in section 4.5.3, page 94.

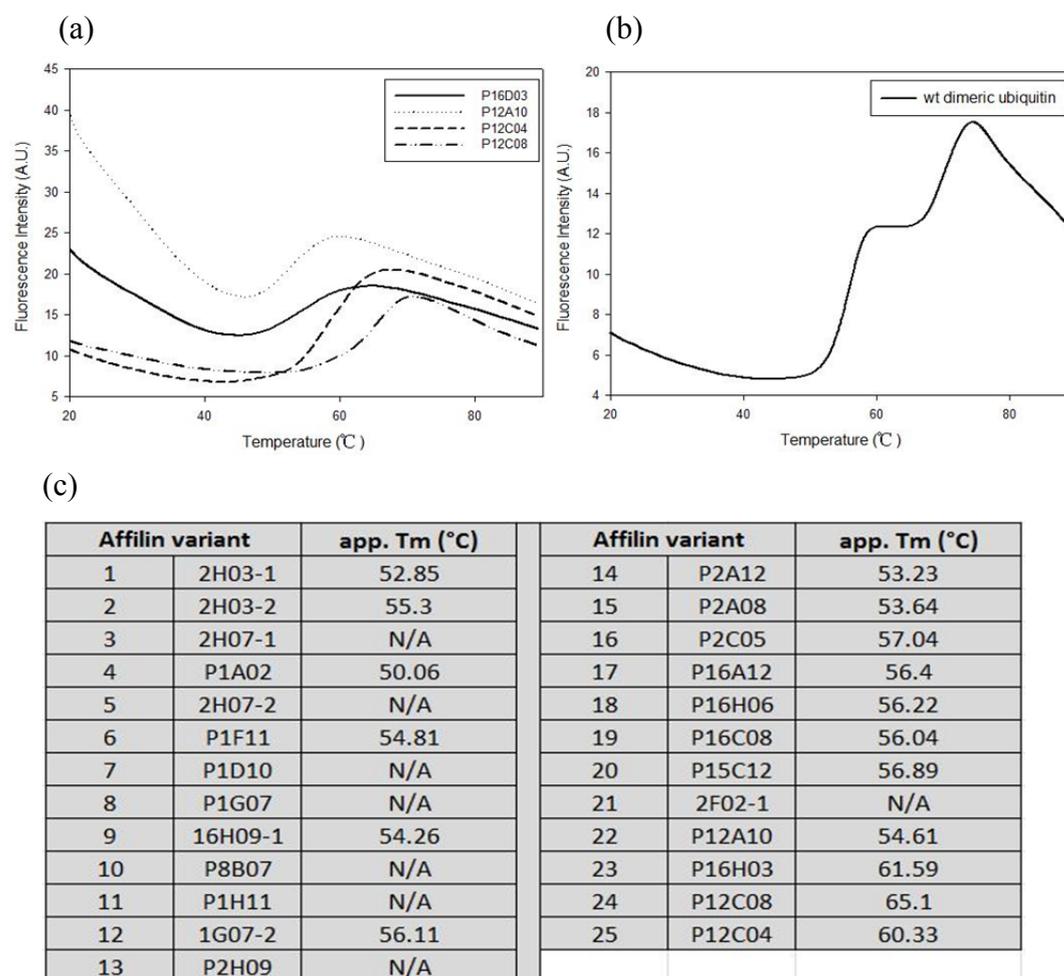


Figure 3.22 Thermostability of Affilin[®] binders measured by DSF experiment. (a), thermograms of four Affilin[®] binders exemplified; (b), thermogram of wild type dimeric ubiquitin; (c), apparent melting temperatures for all Affilin[®] binders. N/A means not available.

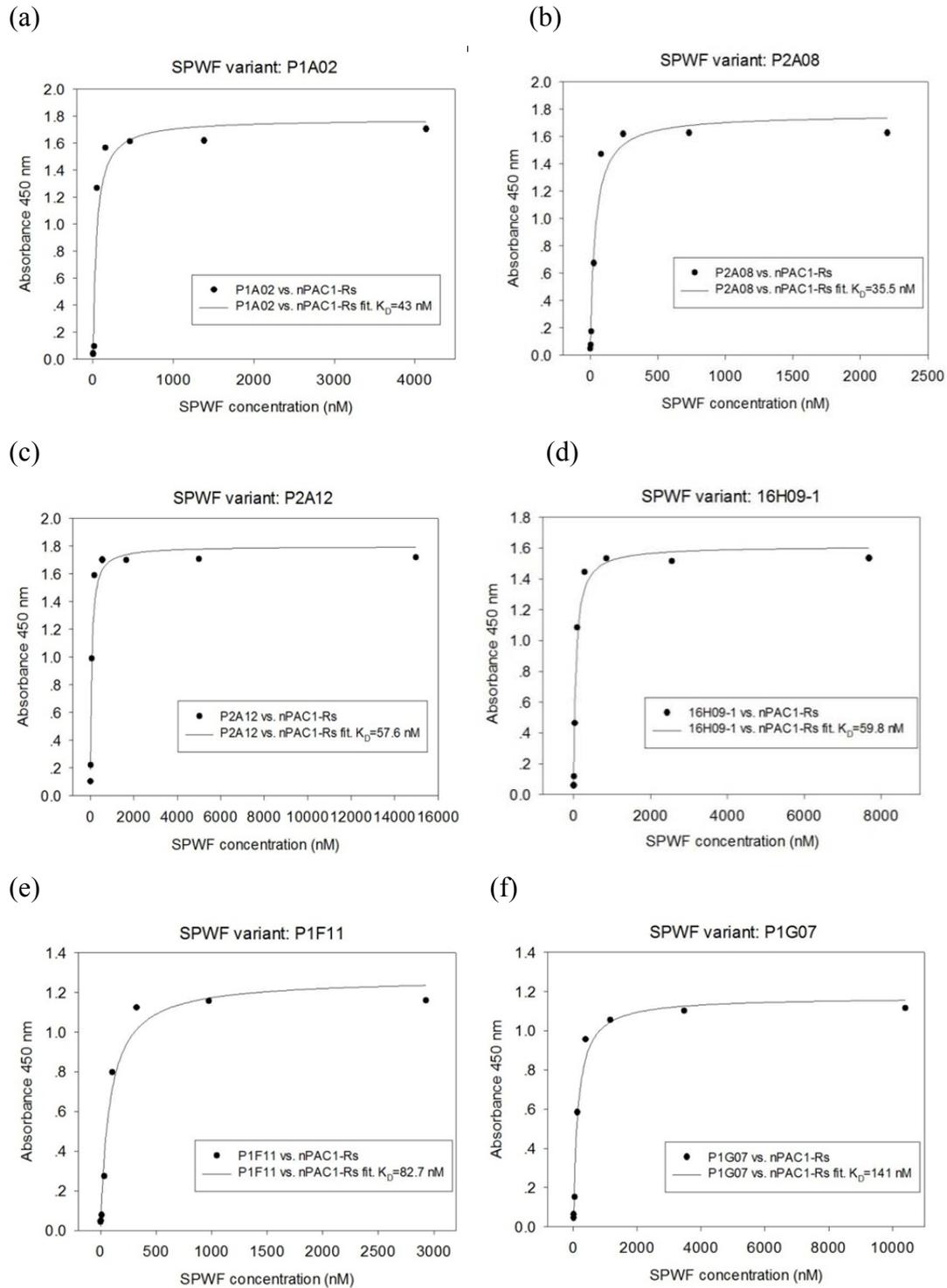
3.3.4. Affinity measurements by concentration-dependent ELISA

To extensively evaluate the binding affinities of selected Affilin[®] binders, the concentration-dependent ELSIA experiment was performed with purified 19 Affilin[®] binders as described in section 2.2.5.8. The initial concentration varied from 2 μ M up

Chapter 3: Results

to 50 μM , depending on the achieved concentration of purified protein samples. The apparent dissociation constants (K_D) were determined by nonlinear regression analysis of the binding data according to equation 6 (see section 2.2.5.8, page 48).

As shown in Fig. 3.23, most Affilin[®] binders presented binding affinities to the target nPAC1-Rs in a nanomolar range. In particular, the variants P1A02, P1F11, P2A08, P2A12 and 16H09-1 presented binding affinities in a low nanomolar range.



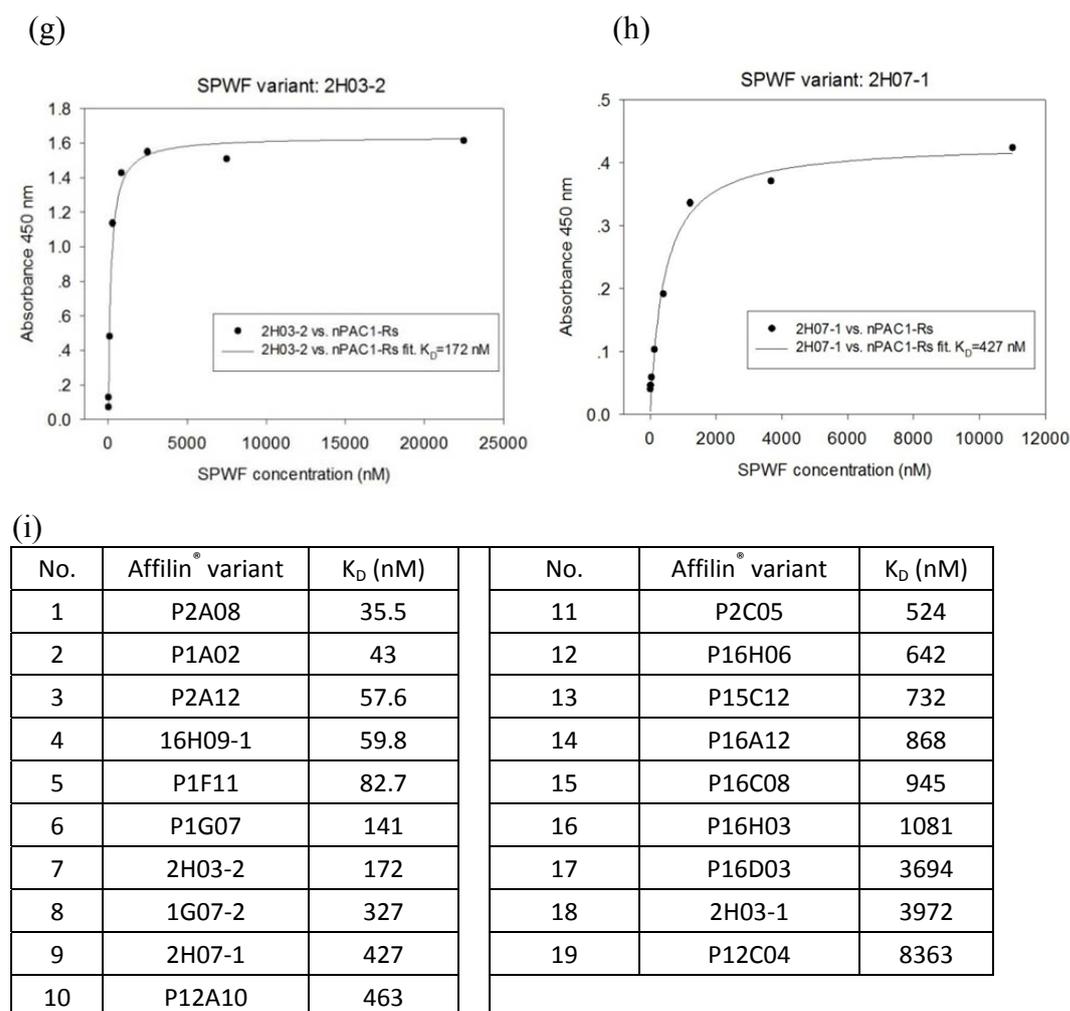


Figure 3.23 Binding affinities of 19 selected Affilin[®] binders to nPAC1-Rs determined by concentration-dependent ELISA. Target protein nPAC1-Rs was immobilized on the wells of MediSorp microtiter plates and incubated with a series of diluted Affilin[®] binders, respectively. The bound Affilin[®] protein was detected by POD conjugated anti-Ubi-Fab antibody in a chromogenic reaction. The curves represented the nonlinear fit of the ELISA data to equation 6 (see section 2.2.5.8, page 48). Figures (a)-(h) illustrate the ELISA data and nonlinear fitted curves of 8 Affilin[®] binders as examples. Figure (i) presents the apparent dissociation constant (K_D) of all 19 Affilin[®] binders.

As mentioned in section 3.2.5 (page 66), the sequence analysis revealed that proline was frequently found in randomized positions, particularly in positions 63', 64' and 65' of SPF domain. Generally proline is structural unfavorable due to the rigidity of side chain. Interestingly, these Affilin[®] binders represented attractive binding affinities in concentration-dependent ELISA experiment. Comprehensively considering the sequence, affinity and solubility, we therefore chose four Affilin[®] binders (Fig. 3.24) with different proline contents for further characterization. The proline-free variant 2H07-1 exhibits affinity in a mid nanomolar range. Both variants P1A02 and P2A08 contain two proline residues in positions 63' and 64', showing binding affinity in low nanomolar range. The variant P2A12 contains three proline residues in positions 63',

64' and 65', respectively and has high affinity in low nanomolar range. Furthermore, variants P1A02 and 2H07-1, as well as variants P2A08 and P2A2 have identical sequence in the first three clusters of randomized positions, respectively.

Randomized position	SPW domain						Linker	SPF domain						Hit-ELISA		Purification			Thermal Stability	Affinity				
	2	4	6	62	63	64	65	66	6'	8'	62'	63'	64'	65'	66'	Ranking	Total Score	Solubility	Expression	Yield (mg/L)	app.Tm [°C]	K _D (nM)		
Wild type Ubiquitin	Q	F	K	Q	K	E	S	T			K	L	Q	K	E	S	T							
P1A02	T	N	I	R	R	A	N	V			D	D	R	P	P	G	W	2	999	5%	8	0.96	50.06	43
P2A08	T	N	I	H	R	N	K	N			D	A	R	P	P	G	W	1	999.5	5%	10	0.26	53.64	35.5
P2A12	T	N	I	H	R	N	K	N			D	A	K	P	P	P	F	5	997	5%	10	3.5	53.23	57.6
2H07-1	T	N	I	R	R	A	N	V			D	D	G	E	W	N	F	N/A	N/A	N/A	N/A	2.6	N/A	427

Figure 3.24 Sequence characteristics of four selected Affilin[®] binders for further analysis.

3.3.5. Binding specificity of Affilin[®] binders

In previous single phase ELISA experiment, the binding specificity of selected Affilin[®] binders was preliminarily tested, showing no binding to the negative controls BSA, lysozyme, human serum and nPTH1-R. To confirm the specificity for the selected 19 Affilin[®] binders, the purified protein samples were applied in specificity ELISA experiment using the protocol described in section 2.2.5.7. Binding was only seen against the target protein nPAC1-Rs, and not against BSA, lysozyme and human serum (Fig. 3.25).

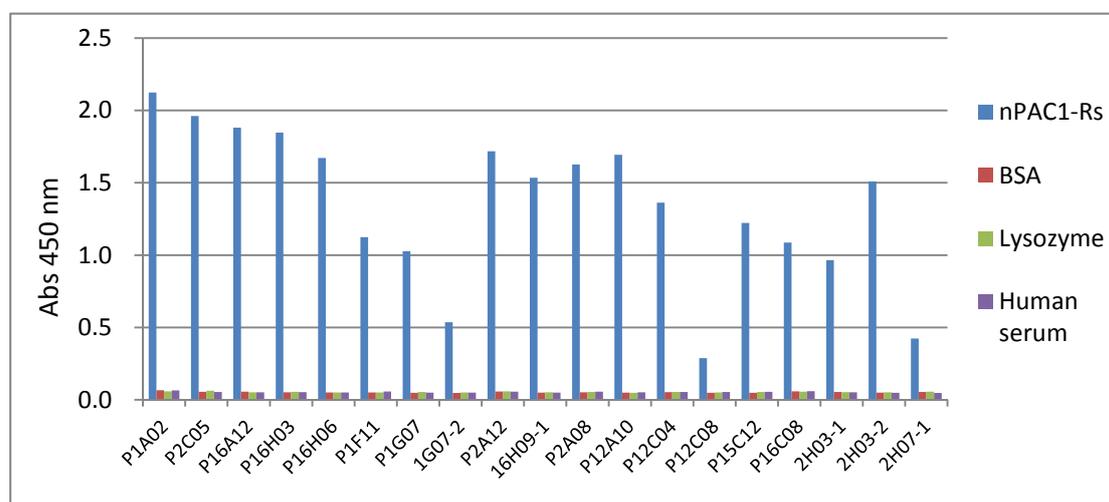


Figure 3.25 Specificity ELISA tests of 19 Affilin[®] binders against target protein nPAC1-Rs and controls BSA, lysozyme and human serum. These samples were coated to ELISA plate and incubated with Affilin[®] binders. The concentration of Affilin[®] binders varied from 2 μ M up to 50 μ M, depending on the achieved concentration of purified protein samples. The bound Affilin[®] protein was detected by POD conjugated anti-Ubi-Fab antibody in a chromogenic reaction.

A further specificity ELISA was performed with four selected Affilin[®] binders as shown in Fig. 3.26. The nPTH1-R protein was included as a control protein. Both

variants P1A02 and P2A12 showed unspecific binding to nPTH1-R and human serum weakly in absence of detergent tween-20 (Fig. 3.26a), while only P2A12 exhibited unspecific binding to nPTH1-R in presence of 0.5% tween-20.

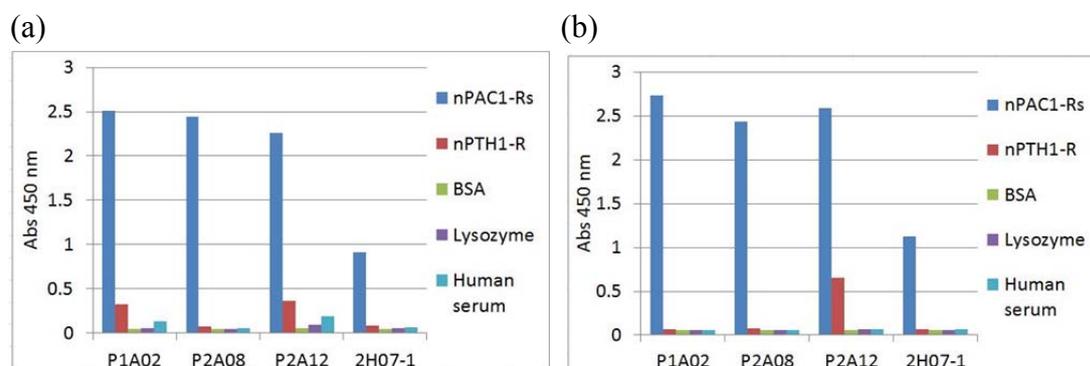


Figure 3.26 Specificity ELISA tests of 4 Affilin® binders in absence and presence of tween-20. Affilin® binders were incubated with immobilized target protein nPAC1-Rs and controls nPTH1-R, BSA, lysozyme and human serum in absence (a) and presence (b) of 0.5% tween-20, respectively. The bound Affilin® protein was detected by POD conjugated anti-Ubi-Fab antibody in a chromogenic reaction.

3.3.6. Biosensor binding analysis

The binding to the target protein nPAC1-Rs was confirmed by SPR measurements for four chosen Affilin® binders P1A02, P2A08, P2A12 and 2H07-1 using a Biacore T100 instrument. Biotinylated nPAC1-Rs protein was immobilized on a SA sensor chip at low concentrations, resulting in the immobilization level of 314.5 RU. The Affilin® binders filtered by 0.25 μ M filters were applied to blank flow cell and target immobilized flow cell simultaneously at increasing concentrations as shown in Fig. 3.27. All variants were found to bind to the immobilized target protein in a concentration-dependent manner. The repeated runs under mid concentration (20 nM for variants P1A02 and P2A08, 80 nM for P2A12 and 2H07-1 respectively) were highly consistent, indicating that the analyte was completely removed and binding ability of immobilized biotinylated nPAC1-Rs protein was remained after the regeneration cycles of the chip.

The kinetic binding curves for four Affilin® binders were fitted to a Langmuir 1:1 binding model using Biacore evaluation software. The kinetic binding constants against nPAC1-Rs were presented in Fig. 3.28. The first three variants P1A02, P2A08 and P2A12 exhibited much slower off-rate than variant 2H07-1. The K_D against the target nPAC1-Rs was determined to be 518 pM, 4.4 nM, 2.7 nM, and 53.6 nM for Affilin® binders P1A02, P2A08, P2A12 and 2H07-1, respectively.

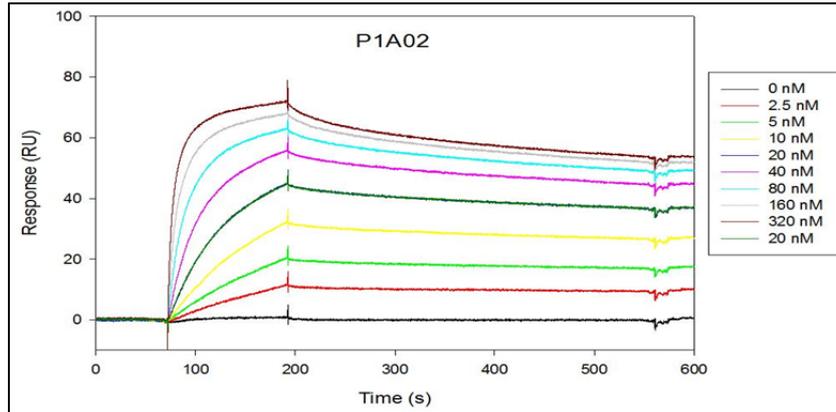
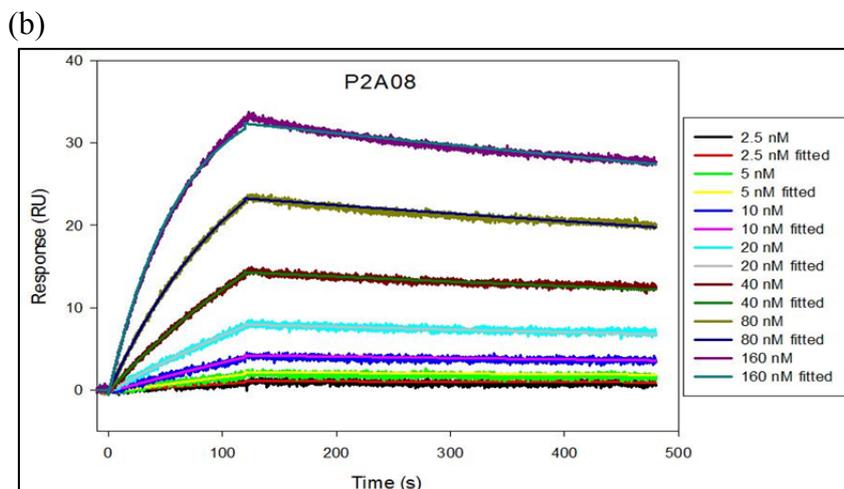
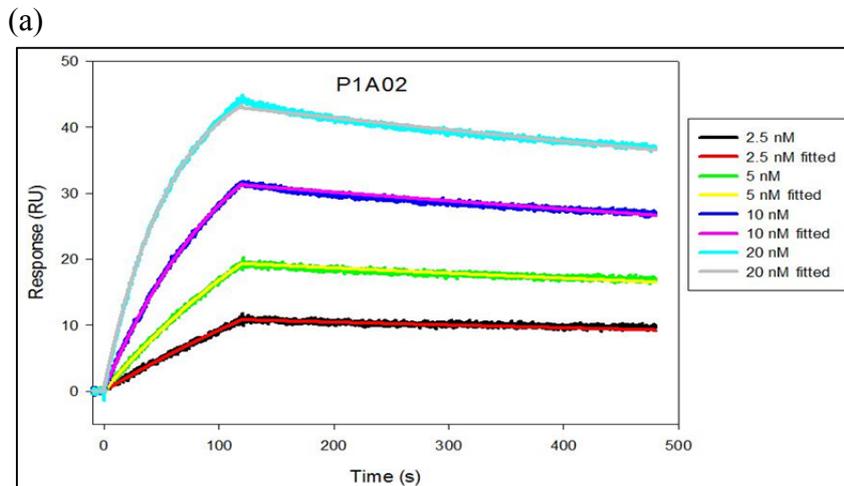


Figure 3.27 Affilin® binder P1A02 exemplified the Affinity determination using SPR. The kinetics of the binding of Affilin® variant P1A02 to nPAC1-Rs was monitored using Biacore T100. Biotinylated nPAC1-Rs protein was immobilized on a flow cell of SA sensor chip. Binding of the Affilin® variant to immobilized target protein was compared to binding to an empty flow cell at increasing concentrations of Affilin® variant. Figure illustrates the kinetics of variant P1A02 with increasing concentrations (2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, 80 nM, 160 nM and 320 nM). A duplicate cycle at the concentration of 20 nM was run following the highest concentration.



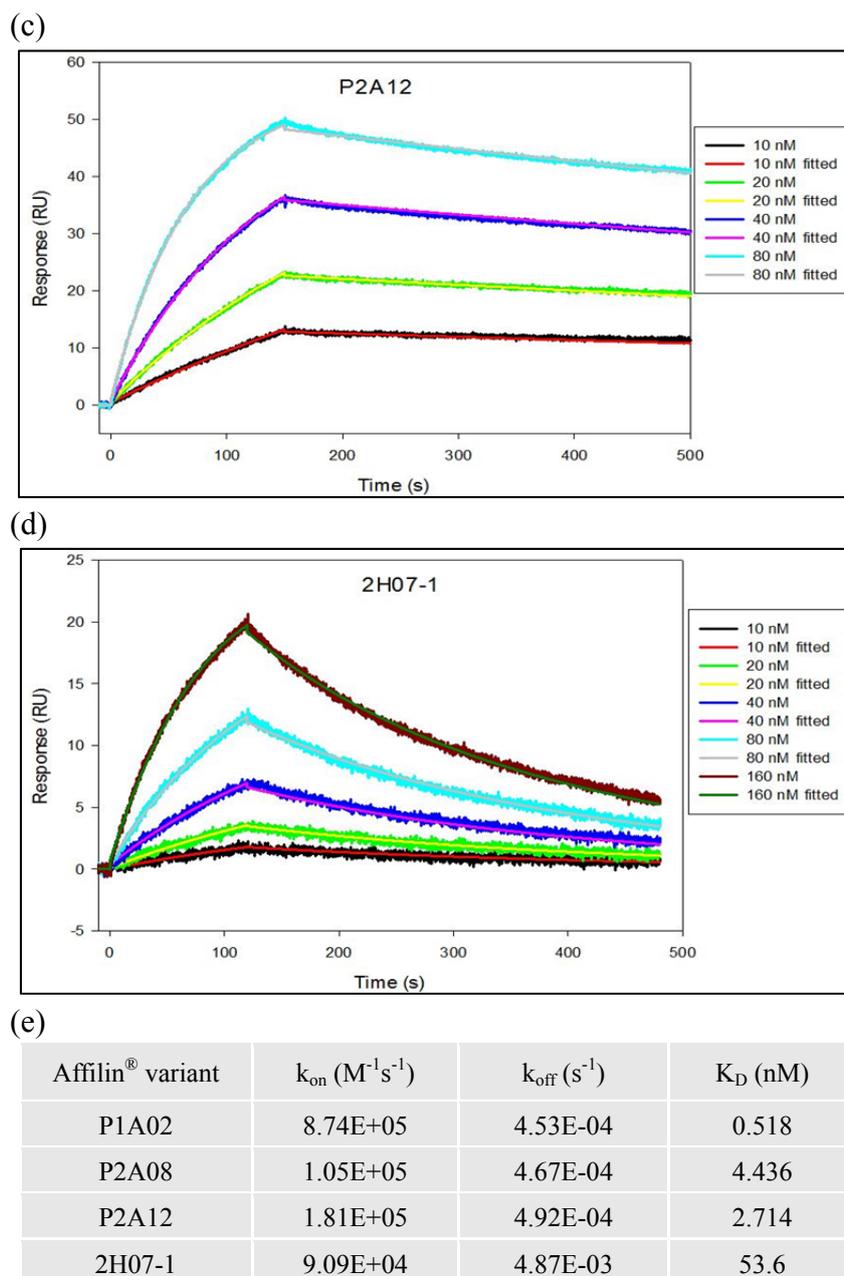


Figure 3.28 Kinetic analyses for four selected Affilin[®] binders using SPR. Experimental data from SPR measurements were presented (thick line) and fitted to a Langmuir 1:1 binding model (slim line). (a), kinetic binding curves and fitted curves of Affilin[®] variant P1A02 at concentrations of 2.5 nM, 5 nM, 10 nM and 20 nM; (b), kinetic binding curves and fitted curves of variant P2A08 at concentrations of 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, 80 nM and 160 nM; (c), kinetic binding curves and fitted curves of variant P2A12 at concentrations of 10 nM, 20 nM, 40 nM and 80 nM; (d), kinetic binding curves and fitted curves of variant 2H07-1 at concentrations of 10 nM, 20 nM, 40 nM, 80 nM and 160 nM; (e), kinetic parameters determined from SPR experiments.

3.3.7. Biosensor competitive assay

Previous SPR and ITC experiments revealed that the ligand PACAP6-38 could bind nPAC1-Rs protein with an affinity of 600-900 nM. To detect if there is any competitive binding of the ligand PACAP6-38 and selected Affilin[®] binders to target protein nPAC1-Rs, four Affilin[®] binders P1A02, P2A08, P2A12 and 2H07-1 were further analyzed by biosensor competitive assay. The concentration of variants P1A02, P2A08 and P2A12 was 80 nM, while a higher concentration was used at 640 nM for variant 2H07-1 due to the lower response. The concentration of ligand PACAP6-38 was 1600 nM. As shown in Fig. 3.29a-c, the sensorgram for the mixture of Affilin[®] variant and PACAP6-38 was very close to the accumulated sum of individual Affilin[®] variant and PACAP6-38 at a Affilin[®]/ PACAP6-38 molar ratio of 1:20, indicating that the binding of these Affilin[®] binders and the ligand PACAP6-38 to nPAC1-Rs was not competitive. A similar situation happened to variant 2H07-1 (Fig. 3.29d).

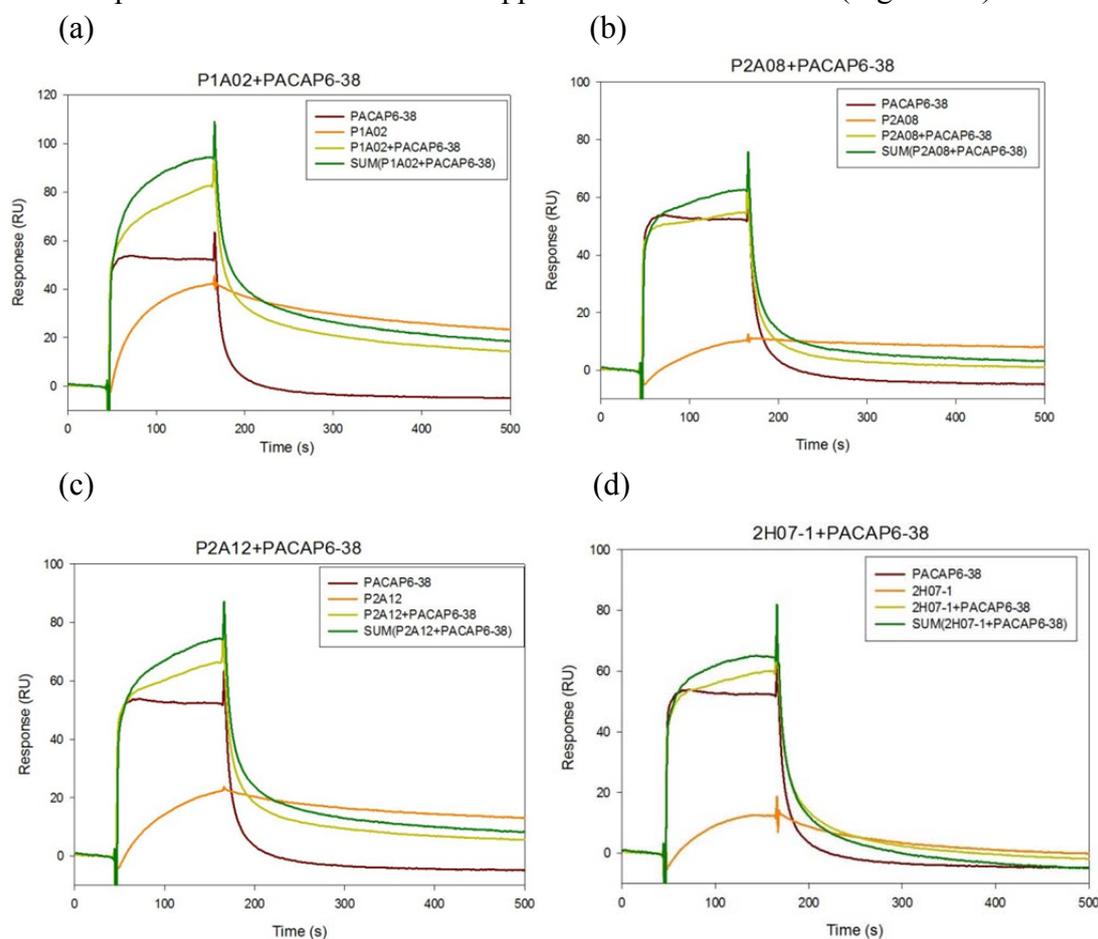


Figure 3.29 Biosensor competitive assay using four selected Affilin[®] binders. Affilin[®] variant, ligand PACAP6-38 as well as the mixture of Affilin[®] variant and PACAP6-38 were respectively injected over a SA sensor chip containing the immobilized biotinylated nPAC1-Rs. The responses were compared to each other. Sensorgrams represent Affilin[®] variant in orange, PACAP6-38 in dark red, mixture of Affilin[®] variant and PACAP6-38 in laurel green, as well as the accumulated sum of individual Affilin[®] variant and PACAP6-38 in green. Figures (a)-(d) demonstrate the competitive assay for Affilin[®] binders P1A02, P2A08, P2A12 and 2H07-1, respectively.

3.3.8. Investigation of monomer ubiquitin variants

In previous solubility analyses and affinity measurements, less solubility and lower protein yield were observed for the majority of selected Affilin[®] binders with high affinity. Furthermore, the high proline content in the second domain (SPF domain) of some Affilin[®] binders raised an interesting question on solubility of these SPF-domain variants. We therefore chose three Affilin[®] binders P1A02, P2A08 and P2A12, and carried out further analyses with their monomer ubiquitin domains.

The amino acid residues on randomized positions of these three Affilin[®] binders were shown in Fig. 3.30a. The variants P2A08 and P2A12 have the same sequence in the SPW domain, therefore two SPW-domain variants (P1A02-SPW and P2A08-SPW) and three SPF-domain variants (P1A02-SPF, P2A08-SPF and P2A12-SPF) were constructed using QuikChange mutagenesis strategy as described in section 2.2.15. The calculated molecular weight of monomer ubiquitin variants is approximately 10 kDa. As shown in Fig. 3.30b, all five variants have very high expression levels. The solubility of the variants P1A02-SPW and P2A08-SPW is 5% and 20% respectively. Surprisingly, all three SPF-domain variants presented the high solubility of 90%-100%, implying that the low solubility of their parental Affilin[®] binders might be caused by the worse solubility of their first SPW domains.

(a)

	SPW domain					Linker	SPF domain									
Randomized position	2	4	6	62	63	64	65	66		6'	8'	62'	63'	64'	65'	66'
Wild type Ubiquitin	Q	F	K	Q	K	E	S	T		K	L	Q	K	E	S	T
P1A02	T	N	I	R	R	A	N	V	SGGGGSGGGGIG	D	D	R	P	P	G	W
P2A08	T	N	I	H	R	N	K	N	SGGGGSGGGGIG	D	A	R	P	P	G	W
P2A12	T	N	I	H	R	N	K	N	SGGGGSGGGGIG	D	A	K	P	P	P	F

(b)

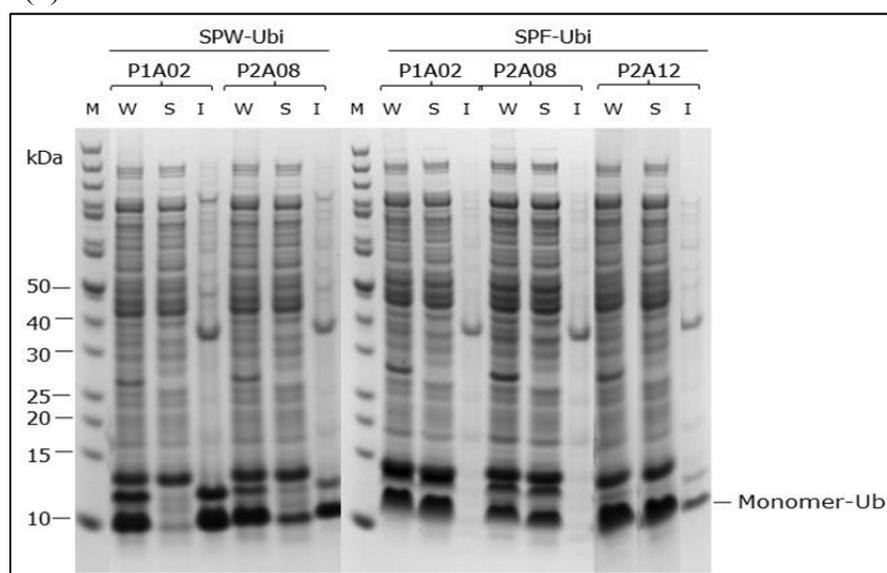


Figure 3.30 Expression level and solubility analyses of five monomer ubiquitin variants by

SDS-PAGE. (a) Sequences of three selected Affilin[®] binders; (b) SDS-PAGE of five monomer ubiquitin variants. M, PageRuler[™] unstained protein ladder; W, whole cells extraction; S, soluble fraction; I, insoluble fraction.

The expression and purification of monomer ubiquitin variants were performed as the standard procedures of Affilin[®] variants. As shown in Fig. 3.31, the variant P1A02-SPW was failed to be purified, the main components in the soluble fraction turned into aggregates during purification. While for the variant P2A08-SPW, except the aggregates presented in the retention volume of 50 ml, the peak appeared in the retention volume of 80 ml presumably is its dimer or different protein species. In contrast, all three SPF-domain variants show no tendency to aggregation and are completely monomeric.

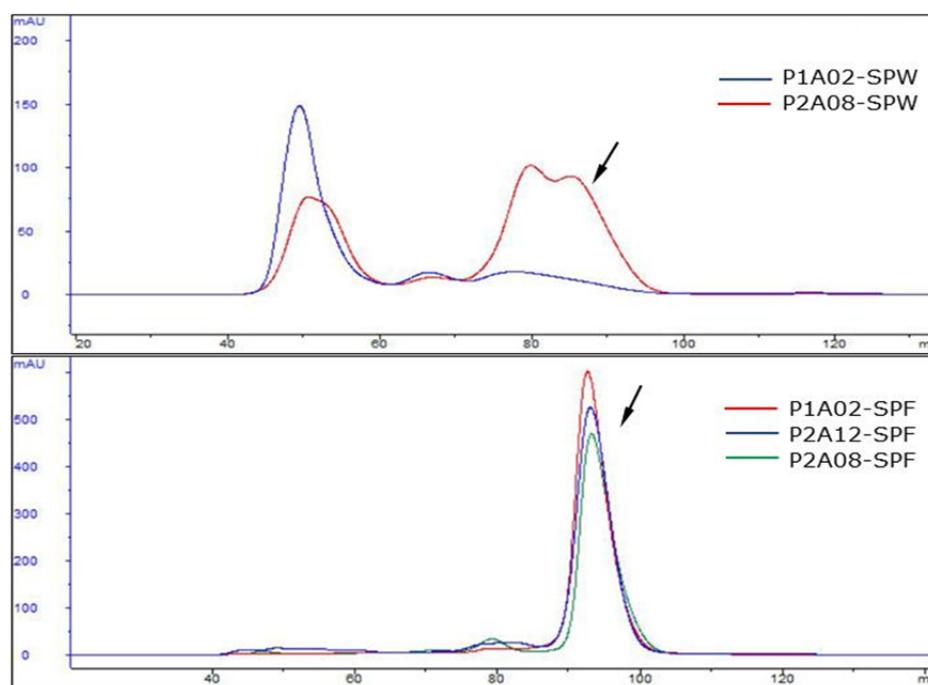


Figure 3.31 Size exclusion chromatography of five monomer ubiquitin variants. The chromatograms of two SPW-domain variants (top) and three SPF-domain variants (bottom) are shown. Arrows indicate the expected monomer proteins.

The secondary structure of three purified SPF-domain variants was measured by far-UV-CD spectroscopy on a Jasco J-810 spectropolarimeter. The measurements were performed with 5 μ M of protein samples in 10 mM KH_2PO_4 pH 7.4. The spectra were characteristic of a folded protein with secondary structure elements dominated by α -helix and β -sheet (Fig. 3.32). All three SPF-domain variants have similar spectra to the wild type ubiquitin, indicating that these SPF-domain variants are structured and their secondary structure composition corresponds to natural ubiquitin protein. The other two SPW-domain variants can't be applied in this experiment due to the low protein concentrations.

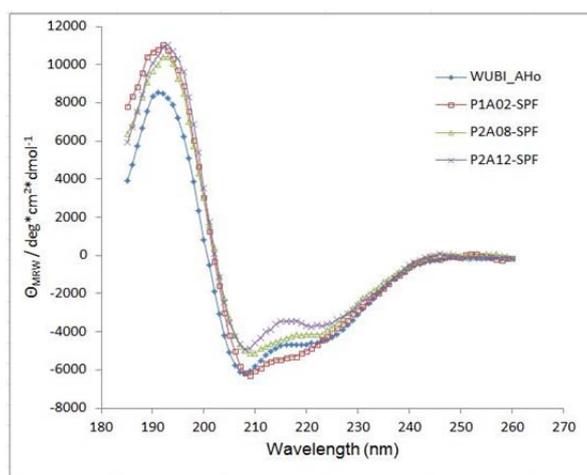
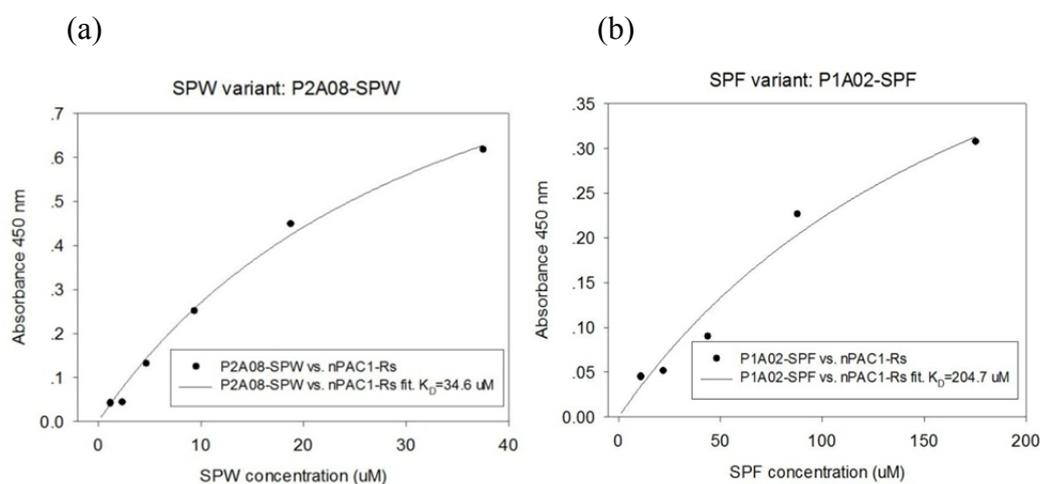


Figure 3.32 Circular Dichroism spectra of three SPF-domain variants. The spectra of P1A02-SPF, P2A08-SPF and P2A12-SPF are compared to the spectrum of the wild type human ubiquitin WUBI (F45W).

As shown in Fig. 3.33, one SPW-domain variant P2A08-SPW and three SPF-domain variants (P1A02-SPF, P2A08-SPF and P2A12-SPF) were subjected to the concentration-dependent ELISA experiment. All these four monomer ubiquitin variants presented binding ability to the target protein nPAC1-Rs in a concentration-dependent manner, with affinities from 28 μM to 204.7 μM .

The affinities of these four monomer ubiquitin variants were compared to their parental Affilin[®] binders as shown in Fig. 3.33e. The affinity of the variant P1A02-SPW couldn't be measured because it was almost completely insoluble and failed to be purified. Interestingly, the multiplication value of the K_D of variants P2A08-SPW and P2A08-SPF is almost equal with the K_D of their parental Affilin[®] variant P2A08. The same situation also happens to the Affilin[®] variant P2A12. These results suggest that two binding patches of SPW and SPF domains have formed a new binding patch as they are connected by a linker, ideally matching the initial design of hetero-dimeric ubiquitin based scaffold protein.



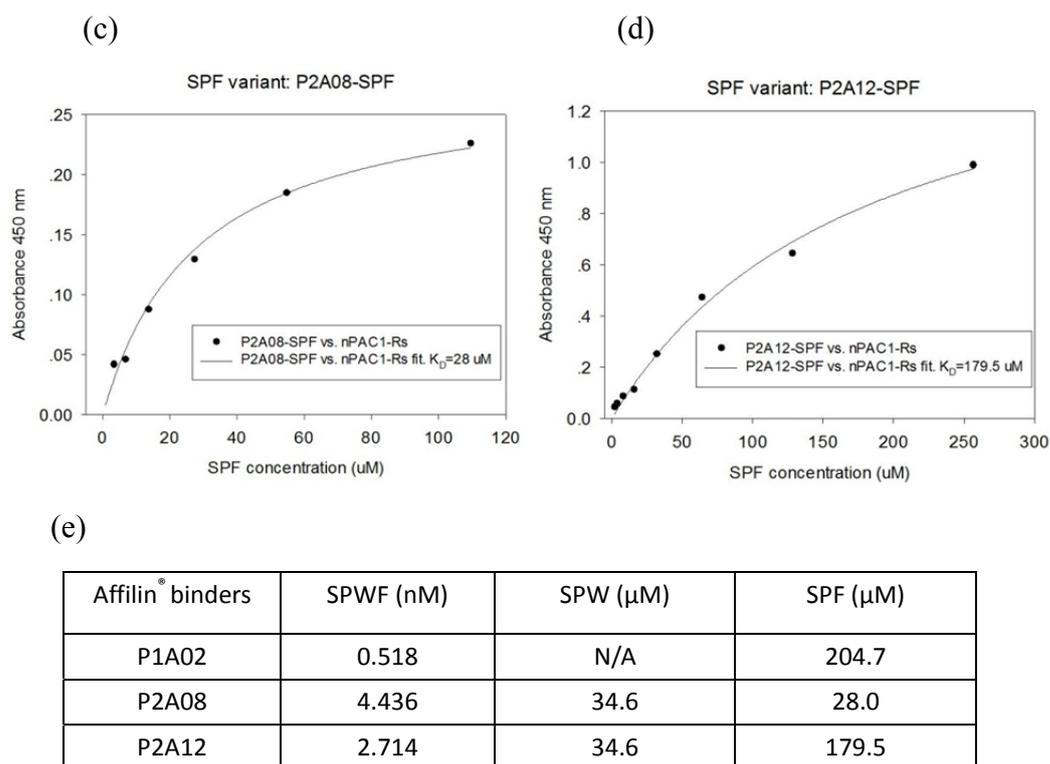


Figure 3.33 Binding affinities of 4 monomer ubiquitin variants to nPAC1-Rs. Target protein nPAC1-Rs was immobilized on wells of MediSorp ELISA plates and incubated with a series of diluted monomer ubiquitin variants respectively. The bound ubiquitin variants were detected by POD conjugated anti-Ubi-Fab antibody in a chromogenic reaction. The curves represented the nonlinear fit of the ELISA data to equation 6 (see section 2.2.5.8, page 48). Figures (a)-(d) illustrate the ELISA data and nonlinear fitted curves of P2A08-SPW, P1A02-SPF, P2A08-SPF and P2A12-SPF. Figure (f) presents the apparent dissociation constant (K_D) of all monomer ubiquitin variants (determined by concentration-dependent ELISA) and their parental Affilin[®] binders (determined by SPR). N/A means not available.

4. Discussion

In the past decade, a rapidly increasing number of studies have clearly illustrated the concept of non-antibody scaffold proteins. On the one hand, more and more proteins with stable architecture have been recruited as scaffold proteins to generate novel artificial binding proteins. This relies (i) on the better understanding of protein structures and protein interactions, (ii) on the rational design, especially for the spatial arrangement of binding sites and the choice of suitable amino acid positions to be randomized, as well as (iii) on the powerful selection technologies. On the other hand, the applications of scaffold protein need to be further explored. Currently the principal applicability of scaffold proteins has been demonstrated in various biotechnological and research areas exclusively restricted to antibodies. But as the alternative therapeutics to antibodies, the efficacy of scaffold drugs utilizing different approaches to implement effector functions is still unknown. The real potential of the various scaffold proteins in medical areas will only become clear once a number of phase I/II trials are completed in the near future (Binz *et al.*, 2005; Gebauer and Skerra, 2009).

The aims of the current investigation are to select and further characterize dimeric ubiquitin-based Affilin[®] binding proteins against a class B GPCR ectodomain (human nPAC1-Rs). Some Affilin[®] binding proteins might be applicable in addressing and affinity purification of the receptor, in cocrystallization, in diagnosis or even in therapy.

4.1. Recombinant production of target protein nPAC1-Rs

Like the other members of class B GPCRs family, human nPAC1-Rs has six conserved cysteine residues, forming three disulfide bonds to stabilize the structure of N-terminal extracellular (EC) domain. It was reported that the expression system using thioredoxin as the fusion partner is useful for high level production of soluble fusion proteins in the *E. coli* cytoplasm, particularly for the proteins containing disulfide bonds, due to the fact that thioredoxin can catalyze disulfide bond formation in a reducing condition of *E. coli* cytoplasm (LaVallie *et al.*, 1993; Stewart *et al.*, 1998). SUMO fusions may also increase the expression of recombinant proteins, enhance the solubility of partially insoluble proteins and generate native protein with any desired N-terminal residue except proline (Saitoh *et al.*, 1997; Butt *et al.*, 2005). Therefore, to produce functional and native target protein for selection, the nPAC1-Rs protein was fused to the C-terminus of thioredoxin-SUMO fusion protein and expressed in *E. coli* BL21(DE3) cells. The thioredoxin-SUMO-nPAC1-Rs fusion protein was expressed in soluble form, with a high expression yield accounting for approximately 50% of the total *E. coli* proteins (Fig. 3.2, page 50). After purification,

the final protein yield was 5 mg per liter cell culture, with a purity of more than 95%. Previous ligand binding studies have revealed that the peptide hormone PACAP38 can bind the full length PAC1 receptor with an affinity of 0.5 nM. The N-terminal residues of PACAP38 are involved in the interaction with other parts of the receptor outside of the N-terminal EC domain, such as the extracellular loops or the transmembrane helices of the receptor (Vandermeers *et al.*, 1992). The peptide PACAP6-38 exhibits an affinity of 350 nM to the nPAC1-Rs protein as determined by a fluorescence polarization anisotropy (FPA) competition assay (Sun *et al.*, 2007). In this study, the biological activity of produced nPAC1-Rs was demonstrated by its ability to bind hormone peptide PACAP6-38. The binding affinity was determined by both SPR and ITC experiments, showing the K_D of approximately 605 nM and 909 nM, respectively. These values are lower than earlier reported affinity but still in the consistent and reasonable range. The peptide PACAP binds the nPAC1-Rs with about 1000-fold lower affinity when compared with the full length PAC1 receptor. This might be due to the influence of the membrane environment, where the peptide could preform into a helical structure and diffuse in two dimensions to interact with the receptor (Inooka *et al.*, 2001). Another possibility is that the full length PAC1 receptor has three glycosylation sites in the N-terminal EC domain, which may lead to a conformational difference when compared to the nPAC1-Rs expressed in *E. coli* cells.

The structure of human nPAC1-Rs protein consists of an N-terminal helix and four β -strands forming two antiparallel sheets. These secondary structure elements are locked by three disulfide bridges between C34 and C63, C54 and C97, as well as C77 and C113. The recombinant production of nPAC1-Rs protein in *E. coli* cells has been attempted previously by using either thioredoxin-nPAC1-Rs fusion protein (Sun *et al.*, 2007) or MBP-nPAC1-Rs fusion protein (Kumar *et al.*, 2011). Both methods yield correctly folded nPAC1-Rs protein, indicating that the three disulfide bridges of nPAC1-Rs intrinsically tend to correctly form. In this study, the production of nPAC1-Rs has shown highly repeatable and stable. The nPAC1-Rs protein produced in this project presents identical ligand binding affinity and Far-UV spectrum to previously produced nPAC1-Rs, which was subjected to disulfide pattern analysis by *offline* nano-HPLC/MALDI-TOF/TOF mass spectrometry and shown correct disulfide bridges as published (data from Andreas Hoffmann). Moreover, the proved ligand binding activity of the nPAC1-Rs in this thesis is an indirect evidence for correct disulfide bridges, as long as S-S bonds are essential for receptor function.

In addition, the approach of fusion expression with thioredoxin-SUMO has been attempted to the other ectodomains of class B GPCRs, such as the calcitonin receptor (nCalcR), secretin receptor (nSecR), VIP receptor (nVIPR), type 1 and type 2 GLP receptor (nGLP1R and nGLP2R) and growth-hormone releasing hormone receptor (nGHRHR). All these fusion proteins express in completely or partially insoluble form and are non-active (Rahfeld, 2010). Therefore, the production of functional nPAC1-Rs by fusion expression with thioredoxin-SUMO is only a special case, which is mainly due to the intrinsic property of nPAC1-Rs protein.

4.2. Dimeric ubiquitin-based Affilin[®] library

Most commonly, successful selections from a combinatorial library against various target molecules can be strongly attributed to three factors: (i) the rational design of library, including the certain scaffold protein used and the diversification strategy employed; (ii) the high diversity of the combinatorial DNA library, as well as the high library quality at protein level; (iii) the efficient display of protein of interest by employed display technology. These factors together determine the functional diversity of the library, which finally affects the selection of artificial binding proteins (Steiner *et al.*, 2008).

Ubiquitin is a highly stable and soluble protein with small size and compact fold. It can extraordinarily tolerate extensive substitutions in primary sequence, retaining its fold well. These attractive properties allow ubiquitin to be an ideal scaffold protein. A monomeric ubiquitin-based Affilin[®] library has been successfully selected against several target proteins, such as nGLP1-R, the Fc portion of human immunoglobulin M (Fc IgM), hydrocortisone (HC), tumor necrosis factor-alpha (TNF α), and vascular endothelial growth factor (VEGF), generating specific binding proteins with affinities up to low nanomolar range (Rudolph *et al.*, 2011). A further developed dimeric ubiquitin-based Affilin[®] library utilizes hetero-dimeric modified ubiquitin proteins connected by a linker, providing a contiguous binding patch with a size larger than 800 Å². Using this strategy, Affilin[®] binding proteins showing specific binding to different targets, for example nerve growth factor (NGF), IgM, TNF α , and the extra domain B of fibronectin (EDB) have been generated, demonstrating affinities up to sub-nanomolar range (Kunert *et al.*, 2011).

Due to the limitations of display technologies, the finally used libraries can physically include 10⁶ to 10¹⁴ members. In spite of this, a relatively large number of residues (10 to 24) are usually involved in randomized positions (Nygren and Skerra, 2004). The dimeric ubiquitin-based Affilin[®] library possesses 15 randomized positions distributed in two ubiquitin domains connected by a flexible linker, providing a vast repertoire (1.5×10^{19}) of binding sites as well as flexible spatial conformation adapted for different targets. Theoretically, the dimeric ubiquitin-based Affilin[®] library should have higher possibility to generate first generation binding proteins against more target molecules when compared to monomeric ubiquitin-based Affilin[®] library with the same diversity.

There are many ways to introduce mutagenesis into a library. From the practical point of view, cassette mutagenesis is the most popularly used method to introduce codons for any set of desired amino acids at specific chosen positions and regions with no codon bias, for example the trinucleotide phosphoramidites (Virnekäs *et al.*, 1994) or Slonomics[®] technology. In this study, the dimeric ubiquitin-based Affilin[®] library was entirely synthesized by Slonomics[®] technology. 19 codons for 19 amino acids (without cysteine) were equally introduced into each randomized position, allowing

covering the whole library with DNA diversity as small as possible. The synthesized Affilin[®] library was nearly completely functional at DNA level and protein level.

Tat-mediated phage display (TPD) system was employed to display fast folding and stable proteins such as Affilin[®] proteins, which are refractory to conventional phage display that utilizes Sec pathway to translocate the protein of interest (POI) into the periplasm. Affilin[®] binding proteins with high affinities have been selected against a broad range of target proteins using the TPD system. In this project, by only three rounds of selection, enrichment factor of approximately 200-fold was achieved. In single phage ELISA experiments, both Affilin[®] proteins and C-terminally followed C-myc tag could be nicely detected by their corresponding antibodies. All these evidences indicate that Affilin[®] proteins can be efficiently display on phage surface by Tat-mediated phage display technology.

4.3. Selection of Affilin[®] binding proteins

Several strategies were used to ensure an efficient selection of Affilin[®] library against the target protein nPAC1-Rs. First, only biotinylated nPAC1-Rs protein was used as target protein, either in solution or immobilized by NeutrAvidin strips or streptavidin beads, providing completely functional target as well as better accessibility for the interaction with Affilin[®] binding proteins. This strategy might also contribute to the selection of binding proteins with high affinity and a variety of diversity, because of the sufficient exposure of target protein. Second, the selection was performed with increased stringency, such as decreasing target amount and increasing washing times between rounds. These strategies provided high pressure for selection and favored the enrichment of binding proteins. Finally, the selection might benefit from some measures, for example, all phages subjected to selection were precipitated twice to get pure phage sample, as well as fresh reaction tubes were always used for each washing step. These experimental procedures can accelerate the removal of unbound Affilin[®] variants and decrease the background binding.

Several innovative selection strategies can be used for phage display technology to facilitate the selection of binding proteins with optimized binding activity and favorable properties (Holliger and Hudson, 2005). These strategies are based on the fact that phage particles are extraordinarily stable against a variety of conditions such as extremes in pH, treatment with denaturants, nucleases or proteolytic enzymes. Therefore, direct selection for favorable biophysical properties can be achieved using chemical denaturants, high temperature, and reducing agents or proteases (Hoogenboom, 2005). As shown in Tab. 4.1, several selected Affilin[®] binding proteins such as No. 7-P1D10 and No. 10-P8B07 have very poor protein yields, resulting from their insoluble expression. These Affilin[®] binders with unfavorable properties may be excluded using the selection strategy that transiently heats phage display Affilin[®] library.

Table 4.1 Overview of 25 selected Affilin[®] binding proteins.

		SPW domain					Linker	SPF domain						Hit-ELISA		Purification			Thermostability	Binding affinity				
Randomized position		2	4	6	62	63	64	65	66		6'	8'	62'	63'	64'	65'	66'	Ranking	Total Score	Solubility	Expression	Yield(mg/L)	app.Tm [°C]	K _D (nM)
Wild type Ubiquitin		Q	F	K	Q	K	E	S	T		K	L	Q	K	E	S	T						56.0	
1	2H03-1	T	N	I	P	D	V	E	R		D	D	P	R	G	T	A					12.6	52.9	3972
2	2H03-2	T	N	I	P	D	V	E	R		D	D	T	D	P	P	Y					8.7	55.3	172
3	2H07-1	T	N	I	R	R	A	N	V		D	D	G	E	W	N	F					2.6	N/A	427
4	P1A02	T	N	I	R	R	A	N	V		D	D	R	P	P	G	W	2	999	5%	8	1.0	50.1	43
5	2H07-2	T	N	I	R	R	A	N	V		D	D	R	V	P	P	W					0.0	N/A	N/A
6	P1F11	T	N	I	V	S	H	P	N		S	R	K	P	P	P	F	33	978.5	5%	8	0.5	54.8	82.7
7	P1D10	T	N	I	V	S	H	P	N		D	N	K	P	P	P	F	26	984	5%	5	0.0	N/A	N/A
8	P1G07	T	N	I	W	P	H	D	V		D	N	K	P	P	P	F	11	993	5%	8	2.4	N/A	141
9	16H09-1	T	N	I	W	P	H	D	V		D	A	R	P	P	G	W					1.5	54.3	59.8
10	P8B07	T	N	I	A	S	N	N	W		D	A	R	P	P	G	W			10%	8	0.0	N/A	N/A
11	P1H11	T	N	I	A	P	N	T	K		I	K	R	H	P	D	W	6	996.5	20%	8	0.0	N/A	N/A
12	1G07-2	T	N	I	G	W	K	T	D		G	W	R	H	P	D	W					0.2	56.1	327
13	P2H09	T	N	I	H	R	N	K	N		S	R	K	P	P	P	F	19	987	30%	8	0.0	N/A	N/A
14	P2A12	T	N	I	H	R	N	K	N		D	A	K	P	P	P	F	5	997	5%	10	3.5	53.2	57.6
15	P2A08	T	N	I	H	R	N	K	N		D	A	R	P	P	G	W	1	999.5	5%	10	0.3	53.6	35.5
16	P2C05	K	W	F	H	R	N	K	N		D	A	R	P	P	G	W	87	942	70%	10	39.0	57.0	524
17	P16A12	K	W	F	H	R	N	K	N		T	K	R	H	P	D	W		734	60%	10	62.6	56.4	868
18	P16H06	K	W	F	H	R	N	K	N		D	A	R	L	T	R	P		656	80%	10	44.3	56.2	642
19	P16C08	K	W	F	H	R	N	K	N		D	T	R	L	T	R	P		552.5	60%	8	29.4	56.0	945
20	P15C12	D	T	I	P	I	G	E	D		W	Q	A	D	V	P	W		700.5	30%	8	4.6	56.9	732
21	2F02-1	D	T	I	E	Q	R	D	T		I	N	R	M	P	P	W	34	977.5	50%	4	0.0	N/A	N/A
22	P12A10	Q	W	V	A	H	V	R	K		N	S	R	N	P	N	W		608	30%	8	23.7	54.6	463
23	P16H03	Y	W	Y	A	P	N	T	K		I	K	R	H	P	D	W		714	70%	10	21.0	61.6	1081
24	P12C08	I	T	V	W	P	H	D	V		L	S	G	E	W	N	F		676	100%	10	135.0	65.1	N/A
25	P12C04	K	H	V	G	N	Q	R	W		M	M	G	I	H	K	K		626	50%	5	17.6	60.3	8363

A standard procedure to identify binding proteins from a naïve library commonly comprises selection, screening and characterization of the selected proteins. The last two processes are quite time and labor consuming. Furthermore, the first generation binding proteins usually are not suited for applications. Maturation steps are necessary for producing the second generation binding proteins with improved properties. The dimeric ubiquitin-based Affilin[®] library has a diversity of approximately 10^{19} , which destines that only a very small part of the library can be covered by a phage display library (typically 10^{9-11}) or ribosome display library (typically 10^{12-14}). Compared to traditional maturation route, a built-in maturation process may be performed, aiming at a rapid selection of favorable Affilin[®] binding proteins. In general, if a selection experiment is successfully carried out obvious enrichment could be observed after 2-3 rounds of selection. This enriched pool after three selection rounds can be directly used for built-in maturation without further characterization. In the case of dimeric ubiquitin-based Affilin[®] library, either N-terminal or C-terminal ubiquitin domain can be retained, whereas the other ubiquitin domain is substituted by its corresponding naïve library after three to four rounds of selection against a certain target. Thus an entire monomeric ubiquitin library is reintroduced and diversity is increased by the reintroduced library. Following another 2-4 rounds of selection, characterization can eventually be carried out. This route exploits the advantages of dimeric ubiquitin-based Affilin[®] library that can be divided into two monomeric libraries, as well as the diversity of each monomeric domain library may be completely covered by popular display technologies such as phage display. By this way, Affilin[®] binding proteins with favorable properties can be rapidly selected from the entire dimeric Affilin[®] library.

4.4. Screening of Affilin[®] binding proteins

Among all clones randomly picked from the elution pools in rounds 3 and 4, about 60-70% clones showed specific binding to the target nPAC1-Rs, indicating that positive binders had been enriched efficiently. These results are compatible with the phage titration results, which presented approximately 200-fold and 300-fold enrichment factors in rounds 3 and 4, respectively. The other 30-40% clones are Affilin[®] variants with frameshift mutations and the background binding, which are difficult to be completely removed from the selection system (Azzazy and Highsmith, 2002).

High throughput Hit-ELISA was employed for initial and extensive identification of positive clones, the so-called hits. Combined with an automated operation platform and data post-processing, the most promising clones were accurately selected from the Hit-ELISA experiment. As demonstrated in Tab.4.1, the Affilin[®] binders with the highest scores in Hit-ELISA showed the lowest K_D values measured by concentration-dependent ELISA. For example, the variants No. 4-P1A02 and No. 15-P2A08 took the top 2 in Hit-ELISA ranking list; they also presented highest affinities in the low nanomolar range. The Hit-ELISA results and affinity data also

revealed that most selected Affilin[®] binding proteins with higher affinities come from the pools eluted by acid elution method, whereas the pools eluted by competitive elution method possess higher positive hit percentage.

4.5. Characterization of selected Affilin[®] binding proteins

4.5.1. Sequence analysis of Affilin[®] binding proteins

The sequence analysis of Affilin[®] variants revealed a very high diversity for the output pools of selection rounds 3 and 4. In total, 59 different Affilin[®] variants were identified from 174 clones. 25 amino acid sequences at randomized positions were aligned in Tab. 4.1. According to the randomized positions, these sequences were divided into four clusters as described in Fig. 3.15, page 67. Interestingly, a high sequence identity was observed within each cluster or each domain. This phenomenon looks like an intermolecular recombination had even occurred during selection, but actually it could not happen, either at DNA level or at protein level. One possible explanation is that this might be caused by the bias of Slonomics[®] technology during the initial library DNA synthesis.

Three Affilin[®] binding proteins are exemplified to represent the hydrophathy profiles of substitutions (Tab. 4.2) and electrostatic surface potential (Fig. 4.1). The hydrophobic amino acids, which could mediate the binding to target protein via hydrophobic interactions, were frequently found within randomized positions. The electrostatic surface potential was shown in monomeric ubiquitin format based on human ubiquitin (PDB code 1UBQ). Compared to wild type ubiquitin, distinct differences are found in the region comprising positions 62-66, that was positively charged for all domains. Interestingly, all three SPF domains strongly present characteristics with negative charge in position 6.

Table 4.2 Properties of amino acids located in randomized positions for wild type ubiquitin and three exemplified Affilin[®] binders. Hydrophobic, basic and acidic amino acid residues are highlighted in yellow, blue and red, respectively. The dissociation constant (K_D) values were determined by SPR experiment.

	SPW domain						Linker	SPF domain						Binding affinity			
Randomized position	2	4	6	62	63	64	65	66		6'	8'	62'	63'	64'	65'	66'	K_D (nM)
Wild type Ubiquitin	Q	F	K	Q	K	E	S	T		K	L	Q	K	E	S	T	
1 P1A02	T	N	I	R	R	A	N	V		D	D	R	P	P	G	W	0.518
2 P2A08	T	N	I	H	R	N	K	N		D	A	R	P	P	G	W	4.436
3 P2A12	T	N	I	H	R	N	K	N		D	A	K	P	P	P	F	2.714

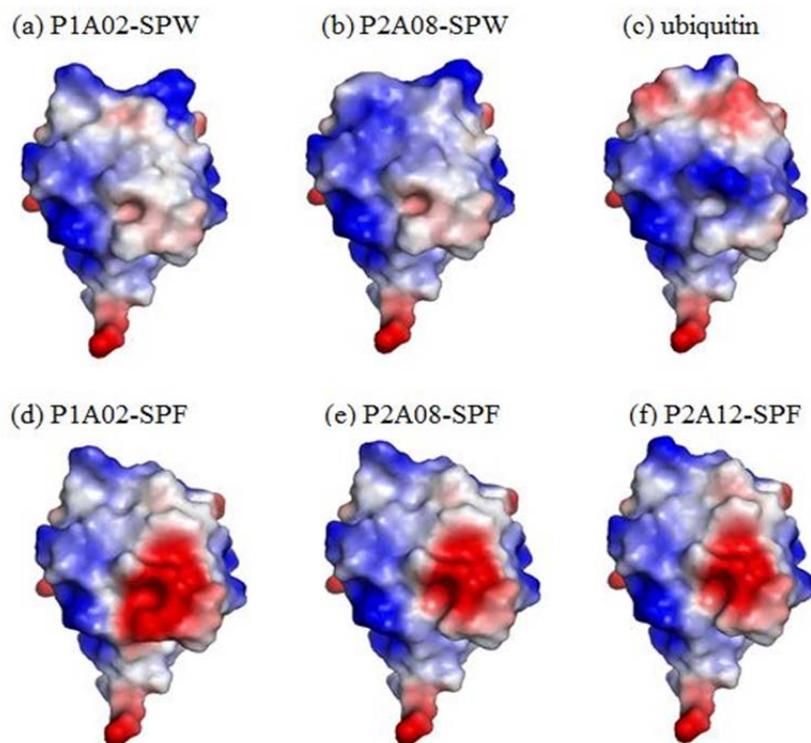


Figure 4.1 Electrostatic surface potential of wild type ubiquitin (c), modeled structures of N-terminal ubiquitin domains (a and b) as well as C-terminal ubiquitin domains (d, e and f) for three Affilin[®] binders. The surface is colored blue for positively charged residues (K, R or H) and red for negatively charged residues (D or E). The other residues are indicated in grey. The saturation of the color is proportional to the degree of electrostatic charge. Models based on coordinates and structure factors of human ubiquitin (PDB code 1UBQ) were prepared for N- and C-terminal domains of Affilin[®] binders P1A02, P2A08 and P2A12 using software Pymol.

Although the properties of selected Affilin[®] binding protein are collectively determined by its entire amino acid sequence, sequence consensus within randomized positions might provide partial but valuable information for characterization. As listed in Tab. 4.1, Affilin[®] binders containing “TNI” motif in the first cluster tended to represent higher affinity as well as lower stability, solubility and protein yields. A further comparison between variants No. 15-P2A08 and No. 16-P2C05, which only have sequence difference in the first cluster, revealed that the “TNI” motif might contribute to the high affinity and low stability, as well as low solubility.

For some Affilin[®] flops, they have a high sequence consensus when compared to the Affilin[®] hits. For example, two Affilin[®] flops No.19-P12D10 and No.20-P8B04 (Fig. 3.17, page 68) have only three different residues in the first cluster among all 15 randomized positions compared with the Affilin[®] hit No.6-P1F11. It might suggest that the first domain (SPW) of these flops is destructured by the substitutions in the first cluster. Although their second domain (SPF) is probably stable and can bind target protein with low affinity in high micromolar range, the weak binding and low stability of entire proteins can't ensure them to be identified as hits.

Among all sequences of selected Affilin[®] binding proteins, proline was found in randomized positions 62-65 and 62'-65', which locate in a loop region and at the beginning of the C-terminal β -sheet strand. In particular, proline was frequently observed in three positions 63', 64' and 65' of SPF domain. Moreover, proline residue also appeared in positions 6, 62, 63 and 65 in several other Affilin[®] binding proteins selected against targets nGLP1-R and Fc IgM (Rudolph *et al.*, 2011). The common trait of these positions is that they either locate in a loop region, or address the terminus of a β -strand adjacent to a loop region. As the proline residue in peptides is restricted by the N-C α torsion angle, proline preferentially generates the turn conformation according to empirical prediction (Rose *et al.*, 1985). We therefore speculate that proline residues in these positions may allow them to escape from the well-defined structure of β -strands, leading to partial or complete destruction of the β -strand.

4.5.2. Recombinant production of Affilin[®] binding proteins

The expression analysis of selected Affilin[®] binders revealed a high expression level for most variants. The expression of some Affilin[®] binders was associated with the formation of inclusion bodies. Depending on the solubility and expression level, the yields of soluble protein of 19 Affilin[®] binders varied from 0.2 to 135 mg per liter cell culture. By contrast, the expression of wild type dimeric ubiquitin is highly soluble in recombinant *E. coli* cells with protein yield of 140 mg per liter cell culture. It suggests that the amino acid substitutions in the artificial binding site have influenced the protein folding and stability, depending on the intrinsic property of scaffold protein. Currently the scaffold protein with highest stability is DARPin. It was reported that randomly picked library members were expressed at a level of 10-30% of total *E. coli* protein with up to a protein yield of 200 mg per liter cell culture (Kohl *et al.*, 2003).

4.5.3. Thermostability of Affilin[®] binding proteins

The thermostability of selected Affilin[®] binding proteins was determined by DSF experiment with fluorescent dye SYPRO orange. The question firstly arose from the thermostability data of a control, the wild type dimeric ubiquitin protein, which comprises two ubiquitin domains connected in head-to-tail format by a linker. The apparent T_m value of wild type dimeric ubiquitin is over 90 °C (determined by Scil Proteins, data not published). By contrast, in DSF experiment, the apparent T_m value was 56 °C where the first transition occurred (Fig. 3.22b, page 74). A hypothesis could be that the transitions observed in DSF experiment are not obtained from the unfolding of the ubiquitin domains. They might be caused by the disruption of interdomain interaction or by the disruption of oligomers. In the diplom thesis of Marcus Böhme (2011), it was described that dimeric ubiquitin protein might form higher oligomers in high concentration. Therefore, all the measured apparent T_m values in DSF experiment are not reliable. The domains of selected Affilin[®] binding

proteins are still intact after heating. Other methodologies such as CD spectroscopy are preferentially utilized for thermostability measurements of dimeric ubiquitin-based Affilin[®] proteins.

4.5.4. Specificity of Affilin[®] binding proteins

The prerequisite for selected Affilin[®] binding proteins was that they could bind target protein nPAC1-Rs with high affinity. On the other hand, nPAC1-Rs-specific binding is also an important requirement for any application. The Affilin[®] binding proteins selected in this study showed highly specific binding to human nPAC1-Rs, without any cross-reactivity against BSA, lysozyme and human serum, either in single phage ELISA or specificity ELISA. The binding ability of selected Affilin[®] binding proteins was examined against another control protein nPTH1-R, which belongs to class B GPCRs family and has ~ 30% similarity with nPAC1-Rs. Although nPTH1-R is a very sticky protein, the selected Affilin[®] binding proteins didn't exhibit any binding ability except one variant P2A12, which presented weakly unspecific binding to nPTH1-R (Fig 3.26, page 78).

4.5.5. Affinity of Affilin[®] binding proteins

Utilizing concentration-dependent ELISA, the binding affinities of selected Affilin[®] binding proteins were determined against target protein nPAC1-Rs, varying from 35.5 nM to 8.4 μ M. Compared to these results, four further analyzed Affilin[®] variants (P1A02, P2A08, P2A12 and 2H07-1) showed over 10-fold higher affinities in the following surface plasmon resonance (SPR) experiments. These differences were probably caused by deviation in protein concentration measurements and crude operation in concentration-dependent ELISA, as where 19 variants were simultaneously analyzed and hundreds of protein samples were diluted. This speculation was subsequently confirmed by a repeated concentration-dependent ELISA for the Affilin[®] variant P2A08, which presented a K_D value (data from Scil Proteins) equal with the SPR data. Although the affinities measured by concentration-dependent ELISA are not reliable, there is a significant correlation between these affinities and their scores in Hit-ELISA as well as the affinities determined by SPR experiments. Therefore, the method of concentration-dependent ELISA is still capable for initial and extensive investigation of affinities. These affinities are valuable for choosing the most promising Affilin[®] binders for further characterization.

In the SPR experiment, remarkable off-rate was exhibited by three Affilin[®] binding proteins (P1A02, P2A08 and P2A12), in the range of 4.5×10^{-4} to 4.9×10^{-4} s⁻¹. A variety of applications could benefit from such slow dissociation rates, for example the purification of Affilin[®] binder/nPAC1-Rs complex.

It is widely accepted that the affinities of selected antibodies mostly correlate with the size of the library. Binders with affinities up to 10 nM could be selected from phage antibody fragment libraries containing 10^7 to 10^9 members. By contrast, binders with sub-nanomolar affinities have been selected from libraries with over 5×10^9 members (Ling, 2003; Hoogenboom, 2005). In this study, after four rounds of primary selection, Affilin[®] binding proteins with sub-nanomolar affinities were selected from the dimeric ubiquitin-based Affilin[®] library containing 7×10^8 individual members. This might be attributed to the rational design and high functionality of Affilin[®] library, as well as the efficient display by Tat-mediated phage display platform. Moreover, various phage display libraries with over 10^{10} members have been constructed using traditional transformation route. It is worth the effort to build an Affilin[®] library with around 10^{10} members. As a consequence, Affilin[®] binders with improved properties may be selected from such a large library.

The binding affinities of selected Affilin[®] binders in this thesis are comparable to the other first-generation Affilin[®] binders selected against targets such as IgM, TNF α and EDB. These Affilin[®] binders have affinities to their targets in the sub-nanomolar to micromolar range (Kunert *et al.*, 2011). Compared to the other first-generation scaffold proteins, only DARPin primarily selected from a very large phage library containing more than 10^{10} members are reported to bind a variety of targets with such high affinities (Steiner *et al.*, 2008). Most binding proteins based on alternative scaffolds usually have affinities in a nanomolar to micromolar range (Hosse *et al.*, 2006), which are at least equal with the affinities of antibodies typically isolated from naïve or synthetic libraries (Yau *et al.*, 2005; Wark and Hudson, 2006).

4.5.6. Monomer ubiquitin domains of Affilin[®] binding proteins

Three SPF monomer ubiquitin domains (P1A02-SPF, P2A08-SPF and P2A12-SPF) with high proline content have been isolated from their parental Affilin[®] binding proteins. Interestingly, the solubility analysis by SDS-PAGE revealed that they are completely soluble. Their elution profiles on size exclusion chromatography also represented highly homogeneous protein species. Moreover, they all showed similar spectra to a control protein human ubiquitin (F45W), which is a correctly folded ubiquitin protein. All these evidences clearly indicate that these monomer ubiquitin domains are still structured and relatively stable. This result implies that proline substitutions occurred in positions 63-65 only lead to a slight conformation change. The compact fold of these ubiquitin domains is maintained.

The binding affinity was determined by concentration-dependent ELISA for one SPW monomer ubiquitin domain (P2A08-SPW) and three SPF monomer ubiquitin domains (P1A02-SPF, P2A08-SPF and P2A12-SPF). They showed relatively low affinities against nPAC1-Rs, varying from 28 μ M to 204.7 μ M, which were over 10^4 - 10^5 folds lower than their parental Affilin[®] binding proteins. The considerable affinity difference could suggest that the SPW and SPF domains bind to different epitopes that address in

target protein nPAC1-Rs, because a simple dimerization of two binding domains which bind the same epitope could not lead to such a large affinity difference. Compared to individual SPW domain and SPF domain, the combination of SPW and SPF domains generates a new binding area in their parental Affilin[®] binding protein, with much higher binding affinity. This also demonstrates the advantage of dimeric ubiquitin-based Affilin[®] library.

By the characterization of individual SPW and SPF domains as well as the sequence alignment, we could learn that the poor protein yields of some Affilin[®] variants (such as variants P1A02, P2A08 and P2A12) are largely caused by low stability of the SPW domain, dominated by “TNI” motif in the first cluster. Therefore, maturation could be carried out by retaining several sequences of SPF domain variants and reintroducing the entire SPW domain library or alternatively reintroducing mutagenesis only within the first cluster of SPW domain. Affilin[®] binding proteins with improved stability may be selected from such a second generation library. The favorable properties such as robust expression, high stability and solubility are often selected together (Holliger and Hudson, 2005).

4.6. Applications of selected Affilin[®] binding proteins

For cocrystallization purpose, several milligrams of Affilin[®] binding protein are required, with relatively high concentration and affinity to the target protein. In this study, the Affilin[®] candidates represented proper protein yields and binding affinities to nPAC1-Rs can now be subjected to cocrystallization or NMR experiment for structure determination. For example, the Affilin[®] binders No. 2-2H03-2, No. 8-P1G07, No. 16-P2C05, and No. 22-P12A10 (Tab. 4.1) showed protein yields between 2.4 and 39 mg per liter cell culture, as well as affinities ranging from low nanomolar to middle nanomolar.

The selected Affilin[®] binding proteins also have the potential for other applications. For example, they might be used for addressing the expression of PAC1 receptor or for diagnosis. These Affilin[®] binders have no competition with ligand binding, but they probably have the potential for therapy, because the binding of these Affilin[®] proteins might affect the signal transduction of PAC1 receptor. Moreover, the Affilin[®] binding proteins with affinities in the range of middle nanomolar to micromolar might be used for the affinity purification of full length PAC1 receptor. These Affilin[®] binders can provide sufficient binding ability and allow elution to occur under mild and non-denaturing conditions.

5. Summary

As alternatives to antibodies, artificial binding proteins based on scaffold concept have been strongly promoted for medicinal and industrial applications. Artificial binding proteins are also powerful tools in research field, for example for structure determination of membrane protein by cocrystallization. It is suggested that artificial binding protein can associate the crystallization of membrane protein by locking membrane protein in a stable conformation and increasing the hydrophilic portion of membrane protein. In addition, artificial binding proteins provide favorable characteristics such as robustness and high soluble expression in microbial hosts, which are more competitive than antibodies and their fragments.

In this thesis, dimeric ubiquitin-based Affilin[®] binding proteins were selected against the target protein N-terminal extracellular domain of human PAC1-R (nPAC1-Rs) and further characterized, aiming at the structure determination of nPAC1-Rs and full length PAC1 receptor by cocrystallization.

The nPAC1-Rs target protein was fused to C-terminus of thioredoxin-SUMO fusion protein and expressed highly soluble in recombinant *E. coli* cells. Native nPAC1-Rs protein was obtained after purification, with protein yield of approximate 5 mg per liter cell culture and a purity of higher than 95%. The functionality of produced nPAC1-Rs protein was demonstrated by its binding ability to the ligand PACAP6-38 determined via both SPR and ITC experiments.

The dimeric ubiquitin-based Affilin[®] library was selected against biotinylated nPAC1-Rs target protein with increasing stringency utilizing Tat-mediate phage display technology. After four rounds of selection, a 300-fold enrichment factor was achieved by comparing the ratio of output/input phage titers between round 4 and round 1. Following initial screening by single phage ELISA experiment revealed that nPAC1-Rs-specific Affilin[®] binding proteins had been efficiently enriched in the elution pools of round 3 and round 4. For subsequent high throughput screening, the DNA fragments encoding Affilin[®] inserts were subcloned from the selected phage pools of both round 3 and round 4 into expression vector pET23dk. 1472 colonies were screened by Hit-ELISA on a robotic workstation. 25 Affilin[®] binding proteins were chosen out of 700 sequenced variants, exhibiting interesting binding properties and sequences.

The properties of 25 selected Affilin[®] binding proteins were extensively investigated. Expression level and solubility of these variants were analyzed by SDS-PAGE. 19 /25 Affilin[®] binding proteins were successfully purified by IMAC and size exclusion chromatography with protein yields varied from 0.2 mg to 135 mg per liter cell culture, depending on the cell mass, expression level and solubility. In a specificity

Summary

ELISA experiment, nearly all of the Affilin[®] binding proteins exhibited high binding specificity to the target nPAC1-Rs. No binding was detected against controls BSA, lysozyme, nPTH1-R and human serum. Further characterization of binding affinity was performed by ELISA and SPR experiments. The most promising Affilin[®] binding proteins showed binding affinity to nPAC1-Rs in low nanomolar to picomolar range. Therefore, the Affilin[®] binding proteins with high affinities in low nanomolar range and proper protein yields in range of milligram per liter cell culture can be utilized for co-crystallization with nPAC1-Rs and the full length PAC1-R respectively, for example the Affilin[®] binding proteins P1G07, 2H03-2 and 16H09-1.

Lower solubility and protein yield were observed for the majority of selected Affilin[®] binding proteins with high affinity. Moreover, the high proline content in the C-terminal domain (SPF domain) of these proteins also raised an interesting question on their solubility. To study these questions, two N-terminal domain (SPW domain) variants and three C-terminal domain (SPF domain) variants were isolated from their parental Affilin[®] binding proteins P1A02, P2A08 and P2A12. Characterization of these five monomeric ubiquitin variants revealed that two SPW domain variants represented low stability, whereas the other three SPF domain variants with high proline content were highly soluble and stable. This result suggested that proline substitutions occurred in positions 63-65 only led to a slight conformation change. The compact fold of these monomeric ubiquitin domains was remained. Further sequence alignment of the SPW domain variants demonstrated that low stability of these variants was mainly caused by the “TNI” substitution motif in the randomized positions 2, 4 and 6. It implied that maturation might focus on these three positions of SPW domain or on the entire SPW domain of these Affilin[®] binding proteins, to get the second generation Affilin[®] binding proteins with improved solubility, stability, protein yield as well as affinity.

6. References

- Achari, A., Hale, S.P., Howard, A.J., Clore, G.M., Gronenborn, A.M., Hardman, K.D., Whitlow, M. (1992). 1.67-Å X-ray structure of the B2 immunoglobulin-binding domain of streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry*. 31(43):10449-57.
- Adams, G.P., Weiner, L.M. (2005). Monoclonal antibody therapy of cancer. *Nat. Biotechnol.* 23:1147-57.
- Ali, S.A., Joao, H.C., Hammerschmid, F., Eder, J., Steinkasserer, A. (1999). Transferrin trojan horses as a rational approach for the biological delivery of therapeutic peptide domains. *J. Biol. Chem.* 274:24066-73.
- Attwood, T.K., Findlay, J.B. (1994). Fingerprinting G-protein-coupled receptors. *Protein Eng.* 7 (2):195-203.
- Azzazy, H.M., Highsmith, W.E. Jr. (2002). Phage display technology: clinical applications and recent innovations. *Clin. Biochem.* 35(6): 425-45.
- Babcock, J.S., Leslie, K.B., Olsen, O.A., Salmon, R.A., Schrader, J.W. (1996). A novel strategy for generating monoclonal antibodies from single, isolated lymphocytes producing antibodies of defined specificities. *Proc. Natl. Acad. Sci. U. S. A.* 93:7843-8.
- Baggio, R., Burgstaller, P., Hale, S.P., Putney, A.R., Lane, M., Lipovsek, D., Wright, M.C., Roberts, R.W., Liu, R., Szostak, J.W., Wagner, R.W. (2002). Identification of epitope-like consensus motif using mRNA display. *J. Mol. Recognit.* 15:126-34.
- Barbas, C.F. (2000). Phage display: a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Bazarsuren, A., Grauschopf, U., Wozny, M., Reusch, D., Hoffmann, E., Schaefer, W., Panzner, S., Rudolph, R. (2002). *In vitro* folding, functional characterization, and disulfide pattern of the extracellular domain of human GLP-1 receptor. *Biophys. Chem.* 96:305-18.
- Beck, A., Wurch, T., Bailly, C., Corvaia, N. (2010). Strategies and challenges for the next generation of therapeutic antibodies. *Nat. Rev. Immunol.* 10(5):345-52.
- Beckman, R.A., Weiner, L.M., Davis, H.M. (2007). Antibody constructs in cancer therapy: protein engineering strategies to improve exposure in solid tumors. *Cancer*.

References

109: 170-9.

Beste, G., Schmidt, F.S., Stibora, T., Skerra, A. (1999). Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold. *Proc. Natl. Acad. Sci. U. S. A.* 96: 1898-903.

Binz, H.K. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J. Mol. Biol.* 332:489-503.

Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grütter, M.G., Plückthun, A. (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* 22:575-82.

Bockaert, J., Pin, J.P. (1999). Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.* 18(7):1723-9.

Boder, E.T., Wittrup, K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* 15:553-7.

Boulianne, G.L., Hozumi, N., Shulman, M.J. (1984). Production of functional chimaeric mouse/human antibody. *Nature.* 312: 643-6.

Brandlein, S., Vollmers, H.P. (2004). Natural IgM antibodies, the ignored weapons in tumour immunity. *Histol. Histopathol.* 19:897-905.

Butt, T.R., Edavettal, S.C., Hall, J.P., Mattern, M.R. (2005). SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* 43(1):1-9.

Byla, P., Andersen, M.H., Holtet, T.L., Jacobsen, H., Munch, M., Gad, H.H., Thøgersen, H.C., Hartmann, R. (2010). Selection of a novel and highly specific tumor necrosis factor alpha (TNFalpha) antagonist: insight from the crystal structure of the antagonist-TNFalpha complex. *J. Biol. Chem.* 285(16):12096-100.

Carter, P. (2001). Improving the efficacy of antibody-based cancer therapies. *Nature Rev. Cancer.* 1:118-29.

Carter, P.J. (2006). Potent antibody therapeutics by design. *Nat. Rev. Immunol.* 6(5):343-57.

Cauvin, A., Buscail, L., Gourlet, P., De Neef, P., Gossen, D., Arimura, A., Miyata, A., Coy, D.H., Robberecht, P., Christophe, J. (1990). The novel VIP-like hypothalamic polypeptide PACAP interacts with high affinity receptors in the human neuroblastoma cell line NB-OK. *Peptides.* 11:773-7.

References

- Chapman, A.P. (2002). PEGylated antibodies and antibody fragments for improved therapy: a review. *Adv. Drug Deliv. Rev.* 54:531-45.
- Chatenoud, L. (2004). Anti-CD3 antibodies: towards clinical antigen-specific immunomodulation. *Curr. Opin. Pharmacol.* 4:403-7.
- Chi, S.W., Maeng, C.Y., Kim, S.J., Oh, M.S., Ryu, C.J., Kim, S.J., Han, K.H., Hong, H.J., Ryu, S.E. (2007). Broadly neutralizing anti-hepatitis B virus antibody reveals a complementarity determining region H3 lid-opening mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 104(22):9230-5.
- Chirino, A.J., Ary, M.L., Marshall, S.A. (2004). Minimizing the immunogenicity of protein therapeutics. *Drug Discov. Today.* 9:82-90.
- Choudhary, S., Mathew, M., Verma, R.S. (2011). Therapeutic potential of anticancer immunotoxins. *Drug Discov. Today.* 16(11-12): 495-503.
- Christmann, A., Walter, K., Wentzel, A., Krätzner, R., Kolmar, H. (1999). The cystein knot of a squash-type protease inhibitor as a structural scaffold for Escherichia coli cell surface display of conformationally constrained peptides. *Protein Eng.* 12:797-806.
- Chung, S., Funakoshi, T., Civelli, O. (2008). Orphan GPCR research. *Br. J. Pharmacol.* 153:339-46.
- Ciechanover, A., Iwai, K. (2004). The ubiquitin system: from basic mechanisms to the patient bed. *IUBMB Life.* 56(4):193-201.
- Connolly, M.L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science.* 221(4612):709-13.
- Dautzenberg, F.M., Mevenkamp, G., Wille, S., Hauger, R.L. (1999). N-terminal splice variants of the type I PACAP receptor: isolation, characterization and ligand binding/selectivity determinants. *J. Neuroendocrinol.* 11(12):941-9.
- Davis, C.B., Gillies, S.D. (2003). Immunocytokines: amplification of anti-cancer immunity. *Cancer Immunol. Immunother.* 52:297-308.
- DeLisa, M.P., Lee, P., Palmer, T., Georgiou, G. (2004). Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway. *J. Bacteriol.* 186:366-373.
- DeLisa, M.P., Tullman, D., Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc. Natl. Acad. Sci.*

References

U. S. A. 100:6115- 20.

Dennis, M.S., Herzka, A., Lazarus, R.A. (1995). Potent and selective Kunitz domain inhibitors of plasma kallikrein designed by phage display. *J. Biol. Chem.* 270:25411-7.

Dubreuil, O., Bossus, M., Graille, M., Bilous, M., Savatier, A., Jolivet, M., Ménez, A., Stura, E., Ducancel, F. (2005). Fine tuning of the specificity of an anti-progesterone antibody by first and second sphere residue engineering. *J. Biol. Chem.* 280(26):24880-7.

Ebersbach, H., Fiedler, E., Scheuermann, T., Fiedler, M., Stubbs, M.T., Reimann, C., Proetzl, G., Rudolph, R., Fiedler, U. (2007). Affilin-novel binding molecules based on human gamma-B-crystallin, an all beta-sheet protein. *J. Mol. Biol.* 372(1):172-85.

Engfeldt, T., Renberg, B., Brumer, H., Nygren, P.A., Karlström, A.E. (2005). Chemical synthesis of triple-labelled three-helix bundle binding proteins for specific fluorescent detection of unlabelled protein. *Chembiochem.* 6(6):1043-50.

Etzerodt, M., Holtet, T.L., Graversen, N.J.H., Thogersen, H.C. (2002). Combinatorial libraries of proteins having the scaffold structure of C-type lectin-like domains. *International Patent Application.* WO 02/48189.

Fiedler, U., Rudolph, R. (2001). Fabrication of beta-pleated sheet proteins with specific binding properties. *International Patent Application.* WO 01/04144.

Flower, D.R. (2003). Towards *in silico* prediction of immunogenic epitopes. *Trends. Immunol.* 24:667-74.

Foord, S.M., Bonner, T.I., Neubig, R.R., Rosser, E.M., Pin, J.P., Davenport, A.P., Spedding, M., Harmar, A.J. (2005). International Union of Pharmacology. XLVI. G protein-coupled receptor list. *Pharmacol. Rev.* 57 (2):279-88.

Foote, J., Winter, G. (1992). Antibody framework residues affecting the conformation of the hypervariable loops. *J. Mol. Biol.* 224:487-99.

Forrer, P., Binz, H.K., Stumpp, M.T., Plückthun, A. (2004). Consensus design of repeat proteins. *Chem. Bio. Chem.* 5:183-9.

Gaudin, P., Couvineau, A., Maoret, J.J., Rouyer-Fessard, C., Laburthe, M. (1995). Mutational analysis of cysteine residues within the extracellular domains of the human vasoactive intestinal peptide (VIP) 1 receptor identifies seven mutants that are defective in VIP binding. *Biochem. Biophys. Res. Commun.* 211(3):901-8.

References

Gebauer, M., Skerra, A. (2009). Engineered protein scaffolds as next-generation antibody therapeutics. *Curr. Opin. Chem. Biol.* 13:245-55.

Gilman, A.G. (1987). G Proteins: Transducers of Receptor-Generated Signals. *Annu. Rev. Biochem.* 56:615-49.

Gonzales, N.R., De Pascalis, R., Schlom, J., Kashmiri, S.V. (2005). Minimizing the immunogenicity of antibodies for clinical application. *Tumour. Biol.* 26:31-43.

Gottschall, P.E., Tatsuno, I., Arimura, A. (1991). Hypothalamic binding sites for pituitary adenylate cyclase-activating polypeptide: Characterization and molecular identification. *FASEB J.* 5:194-9.

Gottschall, P.E., Tatsuno, I., Miyata, A., Arimura, A. (1990). Characterization and distribution of binding sites for the hypothalamic peptide, pituitary adenylate cyclase-activating polypeptide. *Endocrinology.* 127:272-7.

Grace, C.R., Perrin, M.H., Gulyas, J., Digruccio, M.R., Cattle, J.P., Rivier, J.E., Vale, W.W., Riek, R. (2007). Structure of the N-terminal domain of a type B1 G protein-coupled receptor in complex with a peptide ligand. *Proc. Natl. Acad. Sci. U. S. A.* 104:4858-63.

Grauschopf, U., Lilie, H., Honold, K., Wozny, M., Reusch, D., Esswein, A., Schafer, W., Rucknagel, K.P., Rudolph, R. (2000). The N-terminal fragment of human parathyroid hormone receptor 1 constitutes a hormone binding domain and reveals a distinct disulfide pattern. *Biochemistry.* 39:8878-87.

Green, L.L. (1999) Antibody engineering via genetic engineering of the mouse: XenoMouse strains are a vehicle for the facile generation of therapeutic human monoclonal antibodies. *J. Immunol. Methods.* 231:11.

Green, N.M. (1975). Avidin. *Adv. Protein. Chem.* 29:85-133.

Grimm, S. (2011). Ribosome display for selection and evolution of affibody molecules. *Ph.D Dissertation.*

Hanly, W.C., Artwohl, J.E., Bennett, B.T. (1995). Review of Polyclonal Antibody Production Procedures in Mammals and Poultry. *ILAR J.* 37(3):93-118.

Helguera, G., Morrison, S.L., Penichet, M.L. (2002). Antibody-cytokine fusion proteins: harnessing the combined power of cytokines and antibodies for cancer therapy. *Clin. Immunol.* 105:233-46.

Hilpert, K., Wessner, H., Schneider-Mergener, J., Welfle, K., Misselwitz, R., Welfle,

References

- H., Hocke, A.C., Hippenstiel, S., Höhne, W. (2003). Design and characterisation of a hybrid miniprotein that specifically inhibits porcine pancreatic elastase. *J. Biol. Chem.* 278:24986-93.
- Hohlbaum, A.M., Skerra, A. (2007). Anticalins: the lipocalin family as a novel protein scaffold for the development of next-generation immunotherapies. *Expert. Rev. Clin. Immunol.* 3:491-501.
- Holliger, P., Hudson, P.J. (2005). Engineered antibody fragments and the rise of single domains. *Nature Biotechnol.* 23:1126-36.
- Holt, L.J., Herring, C., Jespers, L.S., Woolven, B.P., Tomlinson, I.M. (2003). Domain antibodies: proteins for therapy. *Trends Biotechnol.* 21:484-90.
- Hoogenboom, H.R. (2005). Selecting and screening recombinant antibody libraries. *Nature Biotechnol.* 23:1105-16.
- Hosse, R.J., Rothe, A., Power, B.E. (2006). A new generation of protein display scaffolds for molecular recognition. *Protein Sci.* 15(1):14-27.
- Huang, L., Muyldermans, S., Saerens, D. (2010). Nanobodies[®]: proficient tools in diagnostics. *Expert. Rev. Mol. Diagn.* 10(6):777-85.
- Huber, T., Steiner, D., Roethlisberger, D., Pluckthun, A. (2007). In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: the Na⁺ - citrate symporter CAS as an example. *J. Struct. Biol.* 159:206-21.
- Hufton, S.E., van Neer, N., van den Beuken, T., Desmet, J., Sablon, E., Hoogenboom, H.R. (2000). Development and application of cytotoxic T lymphocyte-associated antigen 4 as a protein scaffold for the generation of novel binding ligands. *FEBS Lett.* 475:225-31.
- Ibarra-Molero, B., Loladze, V.V., Makhatadze, G.I., Sanchez-Ruiz, J.M. (1999). Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry.* 38(25):8138-49.
- Illert, B., Otto, C., Vollmers, H.P., Hensel, F., Thiede, A., Timmermann, W. (2005). Human antibody SC-1 reduces disseminated tumor cells in nude mice with human gastric cancer. *Oncol. Rep.* 13:765-70.
- Inooka, H., Ohtaki, T., Kitahara, O., Ikegami, T., Endo, S., Kitada, C., Ogi, K., Onda, H., Fujino, M., Shirakawa, M. (2001). Conformation of a peptide ligand bound to its

References

- G-protein coupled receptor. *Nat. Struct. Biol.* 8(2):161-5.
- Irving, R.A., Coia, G., Roberts, A., Nuttall, S.D., Hudson, P.J. (2001). Ribosome display and affinity maturation: from antibodies to single V-domains and steps towards cancer therapeutics. *J. Immunol. Methods.* 248:31-45.
- Iwata, S., Ostermeier, C., Ludwig, B., Michel, H. (1995). Structure at 2.8 Å resolution of cytochrome c oxidase from *Paracoccus denitrificans*. *Nature.* 376:660-9.
- Jespers, L., Schon, O., Famm, K., Winter, G. (2004). Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat. Biotechnol.* 22:1161-5.
- Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S., Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature.* 321:522-5.
- Karaplis, A.C., He, B., Nguyen, M.T., Young, I.D., Semeraro, D., Ozawa, H., Amizuka, N. (1998). Inactivating mutation in the human parathyroid hormone receptor type 1 gene in Blomstrand chondrodysplasia. *Endocrinology.* 139: 5255-8.
- Karlsson, F., Borrebaeck, C.A., Nilsson, N., Malmberg-Hager, A.C. (2003). The mechanism of bacterial infection by filamentous phages involves molecular interactions between TolA and phage protein 3 domains. *J. Bacteriol.* 185(8):2628-34.
- Kastrup, J.S., Nielsen, B.B., Rasmussen, H., Holtet, T.L., Graversen, J.H., Etzerodt, M., Thøgersen, H.C., Larsen, I.K. (1998). Structure of the C-type lectin carbohydrate recognition domain of human tetranectin. *Acta. Crystallogr. D Biol. Crystallogr.* 54:757-66.
- Kilpatrick, G.J., Dautzenberg, F.M., Martin, G.R., Eglen, R.M. (1999). 7TM receptors: the splicing on the cake. *Trends. Pharmacol. Sci.* 20(7):294-301.
- Kim, Y., Sette, A., Peters, B. (2010). Applications for T-cell epitope queries and tools in the Immune Epitope Database and Analysis Resource. *J. Immunol. Methods.* Epub ahead of print.
- Kipriyanov, S.M., Le Gall, F. (2004). Generation and production of engineered antibodies. *Mol. Biotechnol.* 26:39-60.
- Kohl, A., Binz, H.K., Forrer, P., Stumpp, M.T., Plückthun, A., Grütter, M.G. (2003). Designed to be stable: crystal structure of a consensus ankyrin repeat protein. *Proc. Natl. Acad. Sci. U. S. A.* 100(4):1700-5.
- Kohanski, R.A., Lane, M.D. (1990). Monovalent avidin affinity columns. *Methods*

References

Enzymol. 184:194-200.

Koide, A., Bailey, C.W., Huang, X., Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. *J. Mol. Biol.* 284:1141-51.

Koren, E., Zuckerman, L.A., Mire-Sluis, A.R. (2002). Immune responses to therapeutic proteins in humans-clinical significance, assessment and prediction. *Curr. Pharm. Biotechnol.* 3:349-60.

Kumar, S., Pioszak, A., Zhang, C., Swaminathan, K., Xu, H.E. (2011). Crystal structure of the PAC1R extracellular domain unifies a consensus fold for hormone recognition by class B G-protein coupled receptors. *PLoS One.* 6(5):e19682.

Kunert, A., Nerkamp, J., Steuernagel, A., Fiedler, M., Fiedler, E., Goettler, T. (2011). A method for identifying hetero-multimeric modified ubiquitin proteins with binding capability to ligands. *International Patent Application.* WO 2011/073214.

Kyte, J., Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157 (1):105-32.

Laburthe, M., Couvineau, A. (2002). Molecular pharmacology and structure of VPAC Receptors for VIP and PACAP. *Regul. Pept.* 108(2-3):165-73.

Laburthe, M., Couvineau, A., Marie, J.C. (2002). VPAC receptors for VIP and PACAP. *Receptors Channels.* 8(3-4):137-53.

Lagerkvist, A.C., Furebring, C., Borrebaeck, C.A. (1995). Single, antigen-specific B cells used to generate Fab fragments using CD40-mediated amplification or direct PCR cloning. *Biotechniques.* 18:862-9.

Lam, H.C., Takahashi, K., Ghatei, M.A., Kanze, S.M., Polak, J.M., Bloom, S.R. (1990). Binding sites of a novel neuropeptide pituitary-adenylate-cyclase-activating polypeptide in the rat brain and lung. *Eur. J. Biochem.* 193:725-9.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Biotechnology (NY).* 11(2):187-93.

Leader, B., Baca, Q.J., Golan, D.E. (2008). Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug. Discov.* 7:21-39.

Ling, M.M. (2003). Large antibody display libraries for isolation of high-affinity antibodies. *Comb. Chem. High Throughput Screen.* 6:421-32.

References

- Lonberg, N. (2005). Human antibodies from transgenic animals. *Nature Biotechnol.* 23:1117-25.
- Löfblom, J., Feldwisch, J., Tolmachev, V., Carlsson, J., Ståhl, S., Frejd, F.Y. (2010). Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. *FEBS Lett.* 584(12):2670-80.
- Marx, J. (2002). Cell biology. Ubiquitin lives up to its name. *Science.* 297(5588):1792-4.
- Mattheakis, L.C. (1994). An in vitro Polysome Display System for Identifying Ligands from very Large Peptide Libraries. *Proc. Natl. Acad. Sci. U.S.A.* 91:9022-6.
- Miao, Z., Levi, J., Cheng, Z. (2010). Protein scaffold-based molecular probes for cancer molecular imaging. *Amino Acids.* Epub ahead of print.
- Miyata, A., Arimura, A., Dahl, R.R., Minamino, N., Uehara, A., Jiang, L., Culler, M.D., Coy, D.H. (1989). Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. *Biochem. Biophys. Res. Commun.* 164:567-74.
- Morag, E., Bayer, E.A., Wilchek, M. (1996). Reversibility of biotin-binding by selective modification of tyrosine in avidin. *Biochem. J.* 316 (Pt 1):193-9.
- Morag, E., Bayer, E.A., Wilchek, M. (1996). Immobilized nitro-avidin and nitro-streptavidin as reusable affinity matrices for application in avidin-biotin technology. *Anal. Biochem.* 243(2):257-63.
- Morrison, S.L., Johnson, M.J., Herzenberg, L.A., Oi, V.T. (1984). Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc. Natl. Acad. Sci. U. S. A.* 81:6851-5.
- Mullis, K.B., Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335-50.
- Mutuberria, R., Hoogenboom, H.R., van der Linden, E., de Bruïne, A.P., Roovers, R.C. (1999). Model systems to study the parameters determining the success of phage antibody selections on complex antigens. *J. Immunol. Methods.* 231(1-2):65-81.
- Muyldermans, S. (2001). Single domain camel antibodies: current status. *J. Biotechnol.* 74:277-302.
- Nielsen, B.B., Kastrop, J.S., Rasmussen, H., Holtet, T.L., Graversen, J.H., Etzerodt, M., Thøgersen, H.C., Larsen, I.K. (1997). Crystal structure of tetranectin, a trimeric

References

- plasminogen-binding protein with an alpha-helical coiled coil. *FEBS Lett.* 412:388-96.
- Nord, K., Nilsson, J., Nilsson, B., Uhlén, M., Nygren, P.A. (1995). A combinatorial library of an alpha-helical bacterial receptor domain. *Protein Eng.* 8:601-8.
- Nord, K., Nilsson, J., Nilsson, B., Uhlén, M., Nygren, P.A. (1997). Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat. Biotechnol.* 15:772-7.
- Nord, K., Gunneriusson, E., Uhlén, M., Nygren, P.A. (2000). Ligands selected from combinatorial libraries of protein A for use in affinity capture of apolipoprotein A-1M and *Taq* DNA polymerase. *J. Biotechnol.* 80:45-54.
- Nygren, P.A. (2008). Alternative binding proteins: affibody binding proteins developed from a small three-helix bundle scaffold. *FEBS J.* 275:2668-76.
- Nygren, P.A., Skerra, A. (2004). Binding proteins from alternative scaffolds. *J. Immunol. Methods.* 290(1-2):3-28.
- Ober, R.J., Radu, C.G., Ghetie, V., Ward, E.S. (2001). Differences in promiscuity for antibody-FcRn interactions across species: implications for therapeutic antibodies. *Int. Immunol.* 13:1551-9.
- Overington, J.P., Al-Lazikani, B., Hopkins, A.L. (2006). How many drug targets are there? *Nat. Rev. Drug Discov.* 5(12):993-6.
- Parthier, C., Kleinschmidt, M., Neumann, P., Rudolph, R., Manhart, S., Schlenzig, D., Fanghänel, J., Rahfeld, J.U., Demuth, H.U., Stubbs, M.T. (2007). Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *Proc. Natl. Acad. Sci. U. S. A.* 104:13942-7.
- Parthier, C., Reedtz-Runge, S., Rudolph, R., Stubbs, M.T. (2009). Passing the baton in class B GPCRs: peptide hormone activation via helix induction? *Trends. Biochem. Sci.* 34(6):303-10.
- Paschke, M. (2006). Phage display systems and their applications. *Appl. Microbiol. Biotechnol.* 70(1):2-11.
- Paschke, M., Hohne, W. (2005). A twin-arginine translocation (Tat)-mediated phage display system. *Gene.* 350:79-88.
- Pastan, I. (2003). Immunotoxins containing *Pseudomonas* exotoxin A: a short history. *Cancer Immunol. Immunother.* 52:338-41.

References

- Perrin, M.H., DiGrucchio, M.R., Koerber, S.C., Rivier, J.E., Kunitake, K.S., Bain, D.L., Fischer, W.H., Vale, W.W. (2003). A soluble form of the first extracellular domain of mouse type 2beta corticotropin-releasing factor receptor reveals differential ligand specificity. *J. Biol. Chem.* 278:15595-600.
- Pierce, K.L., Premont, R.T., Lefkowitz, R.J. (2002). Seven-transmembrane receptors. *Nature Rev. Mol. Cell Biol.* 3:639-50.
- Pioszak, A.A., Parker, N.R., Suino-Powell, K., Xu, H.E. (2008). Molecular recognition of corticotropin-releasing factor by its G-protein-coupled receptor CRFR1. *J. Biol. Chem.* 283:32900-12.
- Pioszak, A.A., Xu, H.E. (2008). Molecular recognition of parathyroid hormone by its G protein-coupled receptor. *Proc. Natl. Acad. Sci. U. S. A.* 105:5034-9.
- Powers, D.B., Amersdorfer, P., Poul, M., Nielsen, U.B., Shalaby, M.R., Adams, G.P., Weiner, L.M., Marks, J.D. (2001). Expression of single-chain Fv-Fc fusions in *Pichia pastoris*. *J. Immunol. Methods.* 251(1-2):123-35.
- Presta, L.G. (2002). Engineering antibodies for therapy. *Curr. Pharm. Biotechnol.* 3:237-56.
- Rahfeld M. (2010). Lösliche Herstellung von N-terminalen G-protein-gekoppelten Rezeptor-domänen der Klasse B in *E. coli*. Diplom thesis.
- Raghava, G.P., Agrewala, J.N. (1994). Method for determining the affinity of monoclonal antibody using non-competitive ELISA: a computer program. *J. Immunoassay.* 15(2):115-28.
- Ramage, R., Green, J., Muir, T.W., Ogunjobi, O.M., Love, S., Shaw, K. (1994). Synthetic, structural and biological studies of the ubiquitin system: the total chemical synthesis of ubiquitin. *Biochem. J.* 299 (Pt 1):151-8.
- Reichert, J.M., Rosensweig, C.J., Faden, L.B., Dewitz, M.C. (2005). Monoclonal antibody successes in the clinic. *Nature Biotechnol.* 23:1073-8.
- Reina, J., Lacroix, E., Hobson, S.D., Fernandez-Ballester, G., Rybin, V., Schwab, M.S., Serrano, L., Gonzalez, C. (2002). Computer-aided design of a PDZ domain to recognize new target sequences. *Nat. Struct. Biol.* 9:621-7.
- Ressler, K.J., Mercer, K.B., Bradley, B., Jovanovic, T., Mahan, A., Kerley, K., Norrholm, S.D., Kilaru, V., Smith, A.K., Myers, A.J., Ramirez, M., Engel, A., Hammack, S.E., Toufexis, D., Braas, K.M., Binder, E.B., May, V. (2011). Posttraumatic stress disorder is associated with PACAP and the PAC1 receptor.

References

Nature. 470:492-7.

Riechmann, L., Clark, M., Waldmann, H., Winter, G. (1988). Reshaping human antibodies for therapy. *Nature*. 332:323-7.

Roberts, B.L., Markland, W., Ley, A.C., Kent, R.B., White, D.W., Guterman, S.K., Ladner, R.C. (1992). Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. *Proc. Natl. Acad. Sci. U. S. A.* 89(6):2429-33.

Roberts, R.W., Szostak, J.W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U. S. A.* 94:12297-302.

Rose, G.D., Gierasch, L.M., Smith, J.A. (1985). Turns in peptides and proteins. *Adv. Protein Chem.* 37:1-109.

Rönmark, J., Hansson, M., Nguyen, T., Uhlén, M., Robert, A., Ståhl, S., Nygren, P.A. (2002). Construction and characterization of affibody-Fc chimeras produced in *Escherichia coli*. *J. Immunol. Methods*. 261:199-211.

Rudolph, R., Fiedler, U., Fiedler, M. (2011). Generation of artificial binding proteins based on ubiquitin proteins. *International Patent Application*. WO 04/106368.

Ruigrok, V.J., Levisson, M., Eppink, M.H., Smidt, H., van der Oost, J. (2011). Alternative affinity tools: more attractive than antibodies? *Biochem. J.* 436(1):1-13.

Runge, S., Schimmer, S., Oschmann, J., Schiødt, C.B., Knudsen, S.M., Jeppesen, C.B., Madsen, K., Lau, J., Thøgersen, H., Rudolph, R. (2007). Differential structural properties of GLP-1 and exendin-4 determine their relative affinity for the GLP-1 receptor N-terminal extracellular domain. *Biochemistry*. 46(19):5830-40.

Runge, S., Thøgersen, H., Madsen, K., Lau, J., Rudolph, R. (2008). Crystal structure of the ligand-bound glucagonlike peptide-1 receptor extracellular domain. *J. Biol. Chem.* 283:11340-7.

Russel, M. (1991). Filamentous phage assembly. *Mol. Microbiol.* 5(7):1607-13.

Saitoh, H., Pu, R.T., Dasso, M. (1997). SUMO-1: Wrestling with a New Ubiquitin-Related Modifier. *Trends Biochem. Sci.* 22:374-6.

Sauer-Eriksson, A.E., Kleywegt, G.J., Uhlén, M., Jones, T.A. (1995). Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. *Structure*. 3:265.

References

- Schirle, M., Weinschenk, T., Stevanovic, S. (2001). Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T-cell epitopes from defined antigens. *J. Immunol. Methods.* 257:1-16.
- Schlehuber, S., Beste, G., Skerra, A. (2000). A novel type of receptor protein, based on the lipocalin scaffold, with specificity for Digoxigenin. *J. Mol. Biol.* 297:1105-20.
- Schlehuber, S., Skerra, A. (2005). Lipocalins in drug discovery: from natural ligand-binding proteins to “anticalins”. *Drug Discov. Today.* 10:23-33.
- Schmid, F.X. (1997). Optical spectroscopy to characterize protein conformation and conformational changes, in: T.E. Greighton (Ed.), *Protein Structure: A Practical Approach*, Oxford University Press, Oxford, UK, pp. 261-97.
- Schwab, C., Bosshard, H.R. (1992). Caveats for the use of surface-adsorbed protein antigen to test the specificity of antibodies. *J. Immunol. Methods.* 147(1):125-34.
- Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O., Grütter, M.G. (2007). Drug export pathway of multidrug exporter AcrB revealed by DARPIn inhibitors. *PLoS Biol.* 5(1):e7.
- Shrake, A., Rupley, J.A. (1973). Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* 79(2):351-71.
- Skerra, A. (2000). Engineered scaffolds for molecular recognition. *J. Mol. Recognit.* 13:167-87.
- Skerra, A. (2001). ‘Anticalins’: a new class of engineered ligand-binding proteins with antibody-like properties. *J. Biotechnol.* 74:257-75.
- Skerra, A. (2003). Imitating the humoral immune response. *Curr. Opin. Chem. Biol.* 7:683-93.
- Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science.* 228(4705):1315-7.
- Steiner, D., Forrer, P., Plückthun, A. (2008). Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *J. Mol. Biol.* 382(5):1211-27.
- Steiner, D., Forrer, P., Stumpp, M.T., Pluckthun, A. (2006). Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat. Biotechnol.* 24:823-31.
- Stewart, E.J., Aslund, F., Beckwith, J. (1998). Disulfide bond formation in the

References

- Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. *MBO J.* 17(19):5543-50.
- Suda, K., Smith, D.M., Ghatei, M.A., Bloom, S.R. (1992). Investigation of the interaction of VIP binding sites with VIP and PACAP in human brain. *Neurosci. Lett.* 137:19-23.
- Sun, C., Song, D., Davis-Taber, R.A., Barrett, L.W., Scott, V.E., Richardson, P.L., Pereda-Lopez, A., Uchic, M.E., Solomon, L.R., Lake, M.R., Walter, K.A., Hajduk, P.J., Olejniczak, E.T. (2007). Solution structure and mutational analysis of pituitary adenylate cyclase-activating polypeptide binding to the extracellular domain of PAC1-RS. *Proc. Natl. Acad. Sci. U. S. A.* 104(19):7875-80.
- Suter, M., Butler, J.E. (1986). The immunochemistry of sandwich ELISAs. II. A novel system prevents the denaturation of capture antibodies. *Immunol. Lett.* 13(6):313-6.
- Thie, H., Schirrmann, T., Paschke, M., Dubel, S., Hust, M. (2008). SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli. *N. Biotechnol.* 25:49-54.
- Tolmachev, V. (2008). Imaging of HER-2 overexpression in tumors for guiding therapy. *Curr. Pharm. Des.* 14(28):2999-3019.
- Uchiyama, F., Tanaka, Y., Minari, Y., Tokui, N. (2005). Designing scaffolds of peptides for phage display libraries. *J. Biosci. Bioeng.* 99(5):448-56.
- Vandermeers, A., Vandenborre, S., Hou, X., de Neef, P., Robberecht, P., Vandermeers-Piret, M.C., Christophe, J. (1992). Antagonistic properties are shifted back to agonistic properties by further N-terminal shortening of pituitary adenylate-cyclase-activating peptides in human neuroblastoma NB-OK-1 cell membranes. *Eur. J. Biochem.* 208(3):815-9.
- Vaneycken, I., D'huyvetter, M., Hernot, S., De Vos, J., Xavier, C., Devoogdt, N., Caveliers, V., Lahoutte, T. (2011). Immuno-imaging using nanobodies. *Curr. Opin. Biotechnol.* Epub ahead of print.
- Vaudry, D., Gonzalez, B.J., Basille, M., Yon, L., Fournier, A., Vaudry, H. (2000). Pituitary adenylate cyclase-activating polypeptide and its receptors: from structure to functions. *Pharmacol. Rev.* 52(2):269-324.
- Verhoeyen, M., Milstein, C., Winter, G. (1988). Reshaping human antibodies: grafting an antilysozyme activity. *Science.* 239:1534-6.
- Vijay-Kumar, S., Bugg, C.E., Cook, W.J. (1987). Structure of ubiquitin refined at 1.8

References

- A resolution. *J. Mol. Biol.* 194(3):531-44.
- Virnekäs, B., Ge, L., Plückthun, A., Schneider, K.C., Wellnhofer, G., Moroney, S.E. (1994). Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Res.* 22(25):5600-7.
- Voss, S., Skerra, A. (1997). Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein Eng.* 10(8):975-82.
- Wahlberg, E., Lendel, C., Helgstrand, M., Allard, P., Dinibas-Renqvist, V., Hedqvist, A., Berglund, H., Nygren, P.A., Härd, T. (2003). An affibody in complex with a target protein: structure and coupled folding. *Proc. Natl. Acad. Sci. U. S. A.* 100:3185-90.
- Walsh, G. (2007). *Pharmaceutical Biotechnology: Concepts and Applications*, Wiley.
- Wang, C.I., Yang, Q., Craik, C.S. (1995). Isolation of a high affinity inhibitor of urokinase-type plasminogen activator by phage display of ecotin. *J. Biol. Chem.* 270(20):12250-6.
- Wark, K.L., Hudson, P.J. (2006). Latest technologies for the enhancement of antibody affinity. *Adv Drug Deliv Rev.* 58(5-6):657-70.
- Wark, P.A. (2002). DX-890 (Dyax). *IDrugs.* 5:586-9.
- Weaver, M., Laske, D.W. (2003). Transferrin receptor ligandtargeted toxin conjugate (Tf-CRM107) for therapy of malignant gliomas. *J. Neurooncol.* 65:3-13.
- Wettschureck, N., Offermanns, S. (2005). Mammalian G proteins and their cell type specific functions. *Physiol. Rev.* 85 (4):1159-204.
- Wilchek, M., Bayer, E.A. (1990). Avidin-biotin technology. *Methods Enzymol.* 184:5-13.
- Williams, A., Baird, L.G. (2003). DX-88 and HAE: A developmental perspective. *Transfus. Apheresis Sci.* 29:255-8.
- Wiseman, T., Williston, S., Brandts, J.F., Lin, L.N. (1998). *Anal. Biochem.* 179:131-7.
- Wong, J.Y. (2006). Basic immunology of antibody targeted radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 66(2 Suppl):S8-14.
- Xu, L., Aha, P., Gu, K., Kuimelis, R.G., Kurz, M., Lam, T., Lim, A.C., Liu, H., Lohse, P.A., Sun, L., Weng, S., Wagner, R.W., Lipovsek, D. (2002). Directed evolution of

References

high-affinity antibody mimics using mRNA display. *Chem. Biol.* 9:933-42.

Yau, K.Y., Dubuc, G., Li, S., Hiram, T., Mackenzie, C.R., Jermutus, L., Hall, J.C., Tanha, J. (2005). Affinity maturation of a V(H)H by mutational hotspot randomization. *J Immunol Methods.* 297(1-2):213-24.

7. Supplementary material

7.1. Construction of plasmid pTrS/nPAC1-Rs

The gene encoding nPAC1-Rs (Fig. 7.1) was synthesized by GeneArt and inserted into pET-Thioredoxin-SUMO expression vector (pTrS vector, Fig. 7.2) by Andreas Hoffmann. The codon usage of the DNA sequence was optimized for expression in *E. coli* cells.

1	M H S D G I F K K E Q A M C L E K I Q R A N E L
1	ATGCATAGCGAT GGCATCTTCAAA AAAGAACAGGCC ATGTGCCTGGAA AAAATTCAGCGT GCGAACGAACTG
73	M G F N D S S P G C P G M W D N I T C W K P A H
73	ATGGGCTTTAAC GATAGCTCTCCG GGGTGTCCGGGC ATGTGGGATAAC ATTACCTGCTGG AAACCGGCGCAT
145	V G E M V L V S C P E L F R I F N P D Q D M G V
145	GTGGGCGAAATG GTGCTGGTGAGC TGCCCGGAACTG TTTCTATTTTC AATCCGGATCAG GATATGGGCGTG
217	V S R N C T E D G W S E P F P H Y F D A C G F D
217	GTGAGCCGTAAC TGCACCGAAGAT GGTGGAGCGAA CCGTTCCGCAT TATTTGATGCG TCGGGCTTCGAT
289	E Y E S E T * *
289	GAATATGAAAGC GAAACCTAATAA

Figure 7.1 DNA sequence of nPAC1-Rs optimized for expression in *E. coli* cells.

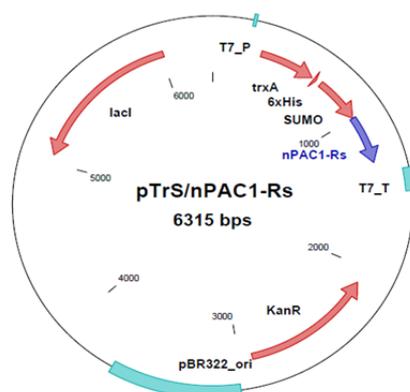


Figure 7.2 Schematic representation of the plasmid pTrS/nPAC1-Rs. T7_P, T7 promoter containing a *lac* operator; *trxA*, gene encoding thioredoxin; 6×His, hexahistidine; SUMO, small ubiquitin-like modifier; nPAC1-Rs, gene encoding target protein nPAC1-Rs; T7_T, T7 transcription terminator; KanR, kanamycin resistance gene; pBR322_ori, origin of replication from pBR322 plasmid; *lacI*, gene encoding the *lac* repressor.

7.2. Construction of Affilin[®] library

The DNA fragments of dimeric ubiquitin-based Affilin[®] library (SPWF) were synthesized using Slonomics[®] technology and inserted into Tat-phage display based

phagemid vector pCD87SA via *sfiI/sfiI* as shown in Fig. 7.3. The use of the Tat pathway ensures the efficient display of fast-folding Affilin[®] variants on phage particles. Different codon usage for each amino acid residue was combined to generate dissimilar DNA sequences in framework region for SPW and SPF genes.

The purified ligation products of pCD87SA-SPWF were transformed into *E. coli* ER2738 cells, yielding an original Affilin[®] SPWF library with a complexity of 7×10^8 . The plasmid DNA extracted from this Affilin[®] SPWF library was aliquoted and frozen at $-80\text{ }^{\circ}\text{C}$ for stock. Before selection, the library DNA was transformed into commercial ER2738 electrocompetent cells. The freshly transformed *E. coli* ER2738 cells with an overrepresentation of 37 folds were sequentially applied for selection. This part of work was carried out by Scil Proteins.

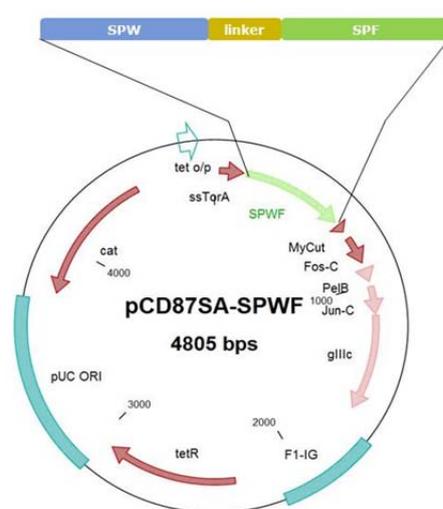


Figure 7.3 Schematic representation of plasmid pCD87SA-SPWF. The DNA fragments (SPWF) of dimeric ubiquitin-based Affilin[®] library were inserted into phagemid vector pCD87SA via *sfiI/sfiI*. Tet o/p, Tet operator/promoter sequence; ssTorA, Tor A leader sequence; SPWF, DNA fragment of dimeric ubiquitin-based Affilin[®] library; MyCut, combined myc epitope and TEV protease cleavage site; Fos-C, cysteine modified Fos leucine zipper domain; PelB, PelB leader sequence; Jun-C, cysteine modified Jun leucine zipper domain; gIIIc, C-terminal domain of M13 phage minor coat protein III (250-406 aa); F1-IG, filamentous F1 phage intergenic region; tetR, Tet o/p repressor; pUC ORI, ColE1 (PUC) origin of replication; cat, chloramphenicol acetyltransferase gene.

8. Abbreviations

A, C, G, T	Adenine, Cytosine, Guanine, Thymine
Abs 450 nm	Absorbance at 450 nm
ADCC	Antibody dependent cellular cytotoxicity
ATP	Adenosine triphosphate
bp	base pair
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CD	Circular Dichroism
CDC	Complement dependent cytotoxicity
CDRs	Complementarity determining regions
CFU	Colony forming units
CV	Column volume
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides triphosphate
DSF	Differential scanning fluorimetry
ECD	Extracellular domain
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>exempli gratia</i> (for example)
EDTA	Ethylene diamine tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
<i>et al.</i>	<i>et alii</i> (and others)
Fab	Fragment antigen binding
GPCRs	G protein-coupled receptors
HBS	HEPES buffered saline
HEPES	2-{(4-(hydroxyethyl)-1-piperazin)}ethanesulfonic acid
HRP	Horseradish peroxidase
IgG	Immunoglobulin G
IMAC	Immobilized metal-ion affinity chromatography
ITC	Isothermal titration calorimetry
K_D	dissociation constant
kDa	kilodalton
LB	Luria-Bertani
MES	2-(N-morpholino) ethanesulfonic acid
min	minute
MOI	Multiplicity of infection
mRNA	messenger ribonucleic acid
MWCO	Molecular weight cut off
nPAC1-Rs	N-terminal extracellular domain of human PAC1 receptor splice variant

Abbreviations

nPTH1-R	N-terminal extracellular domain of parathyroid hormone type 1 receptor
OD ₆₀₀	Optical density at 600 nm
PAC1-R	PACAP type 1 receptor
PACAP6-38	Pituitary adenylate cyclase activating polypeptide (residues 6-38)
PBS	Phosphate buffered saline
PBS-T	PBS buffer containing 0.1% (v/v) Tween-20
PBS-TB	PBS-T buffer containing 1× Sigma blocking buffer
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
pH	Percentage of hydrogen
pI	Isoelectric point
PMSF	Phenylmethanesulfonylfluorid
POI	Protein of interest
POD	Peroxidase
rpm	Revolutions per minute
RT	Room temperature
RU	Response unit
scFv	single-chain variable domain antibody fragment
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	second
SEC	Size exclusion chromatography
SPR	Surface plasmon resonance
SUMO	small ubiquitin-like modifier
TAE	Tris-acetate-EDTA buffer
Tat	Twin arginine translocation
TCA	Trichloroacetic acid
TMB	3,3',5,5'-tetramethylbenzidine
Tris	N-{tris-(hydroxymethyl-)}aminomethane
Trx	Thioredoxin
V	Voltage
v/v	Volume per volume
w/v	Weight per volume
2YT	Two times yeast tryptone
β-ME	β-Mercaptoethanol

Publication list

Publication:

This is a cooperation project between Nachwuchsgruppe “Künstliche Bindeproteine” and Scil Proteins, GmbH. According to the cooperation agreements, parts of this thesis will be allowed to publish after the application of concerned patent, which is now in progress.

Poster:

Song, B., Hoffmann, A., Kunert, A., Zwarg, M., Rudolph, R., Pfeifer, S. (2011). Efficient selection of dimeric Ubiquitin-based Affilin[®] binding proteins against a class B GPCR ectodomain. *25th Anniversary Symposium of the Protein Society 2011*, Boston.

Curriculum Vitae

Name: Bo Song
Date of birth: 04.11.1978
Place of birth: Hubei, China
Nationality: Chinese
Address: Brandbergweg 35A-407, D-06120 Halle (available until 11/2011)
(Postal address: Bo Song, c/o Frau Qi Zhang, Richard-Paulick Str.
13/Zimmer 606, D-06124 Halle)
Email: symbolsb@hotmail.com

Educational and research experience:

- 09/2007 – now PhD student in the InnoProfile NWG “Artificial binding proteins” led by Dr. Sven Pfeifer, Institut für Biochemie und Biotechnologie, Martin-Luther-Universität Halle-Wittenberg, Supervisor: Prof. Dr. Rainer Rudolph (09/2007-12/2009) and PD Dr. habil. Hauke Lilie (since 12/2009).
- 06/2004 – 02/2007 Research assistant, Acon Biotech Co., Ltd., Hangzhou, China.
- 09/2001 – 06/2004 Master student in the Institute of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, China, Thesis: Site-directed mutation of the β -subunit of phycocyanin and phycoerythrocyanin and the study of reconstitution *in vitro*, Supervisor: Prof. Kai-Hong Zhao, Degree: Master of Science in Biochemistry.
- 09/1997 – 07/2001 Bachelor student in the College of Life Science, Huazhong Normal University, Wuhan, China, Major: Biochemistry, Degree: Bachelor of Science in Biochemistry.
- 09/1994 – 07/1997 High school student, The first high school of Zhijiang, Hubei, China.

Erklärung

Hermit erkläre ich an Eides statt, dass ich mich bisher mit dieser Arbeit weder an der Martin-Luther-Universität Halle-Wittenberg, noch an einer anderen Einrichtung um die Erlangung eines akademischen Grades beworben habe. Ich versichere weiterhin, dass die vorliegende Arbeit selbstständig und nur unter Benutzung der angegeben Quellen und Hilfsmittel erstellt wurde. Den benutzten Werken wörtlich oder inhaltlich entnommene Stellen sind als solche gekennzeichnet.

Halle (Saale), 02. 11. 2011

Bo Song