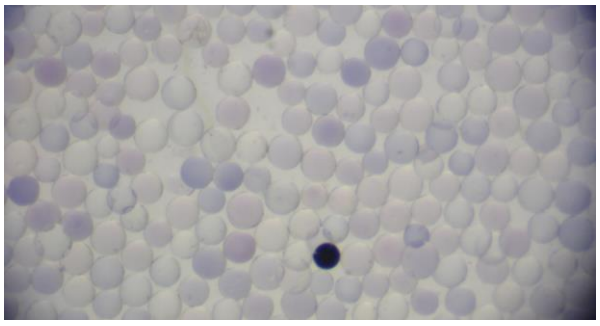
Chapter 2

Evolution of the OBOC Peptide Library

Screening Protocol



2.1 Introduction

2.1.1 Iterative In Situ Click Chemistry for Protein-Catalyzed Capture (PCC) Agent Development

Previous work in the Heath lab demonstrated that the flexibility of chemically synthesized One-Bead, One-Compound (OBOC) peptide libraries could be combined with the selective power of the *in situ* click process to develop multi-peptide ligand capture agents that can serve as drop-in antibody replacements in assays[1]. These peptide ligands can be made in large quantities entirely by robots, making the scale-up cheap and robust. They are also highly stable agents that can be used in a variety of assays, removing the need for the gold-standard antibodies in a variety of protein detection techniques[2,3].

The iterative *in situ* click screen to develop a capture agent starts with the discovery of a peptide ligand that binds to a protein target through the use of OBOC library screening. Once a peptide has been discovered, labeled the "anchor peptide," it is appended with a click handle and screened again against the protein in the presence of a new OBOC library that contains the opposing click handle, as seen in **Figure 2-1**. When a library member binds to the surface of the protein in close proximity to the anchor ligand and is held in place through a high-affinity for the protein target, a click reaction between the anchor and library-bound ligand can occur without the use of the Cu(I) catalyst. The addition of this new ligand, the secondary ligand, forms a "biligand" in complex with the original anchor. This selection technique allows the protein target itself to catalyze the formation of the peptide ligands that bind to it with the highest affinity and selectivity. This iterative process can be performed as many times as necessary to produce a ligand with the desired affinity and specificity for the target, and serves as the basis for the iterative *in situ* click chemistry technique for protein-catalyzed capture (PCC) agent production. After a PCC agent has been discovered using this technique, the Cu(I) catalyst can be brought back in order to scale-up the final click triazole-containing product in high quantities.

The technology as presented by Agnew, *et al*[1] provided a solid foundation for the construction of these PCC agents, but the methods, discussed in section 2.3.1, were time-consuming and labor-intensive, making rapid ligand discovery very difficult. After this BCAii proof-of-concept PCC agent was completed, the next stage of technology development required an optimization of the techniques involved in order to increase the robustness and output of the overall process. This chapter describes the transformation of the OBOC iterative *in situ* click technology into an efficient and robust technique.

2.1.2 Prostate Specific Antigen (PSA)

Prostate Specific Antigen (PSA) is a serum protease produced by the prostate. The accurate detection of PSA levels in the blood can be a strong indicator of the presence of prostate cancer, but this result is confounded by the elevated PSA levels also seen in Benign Prostate Hyperplasia (BPH), a non-cancerous condition[4]. In serum, PSA is partially in complex with α_1-antichymotrypsin (ACT), with 60-95% generally found as a PSA-ACT complex while the rest of the PSA remains free. It has also been discovered that the PSA-ACT fraction is larger in prostate cancer, whereas BPH has more PSA free in serum[4]. It was hypothesized, therefore, that a better PSA detection test could be designed to measure this through the use of PCC agents, and much of the screening strategies developed in this chapter were focused on the design of this agent.

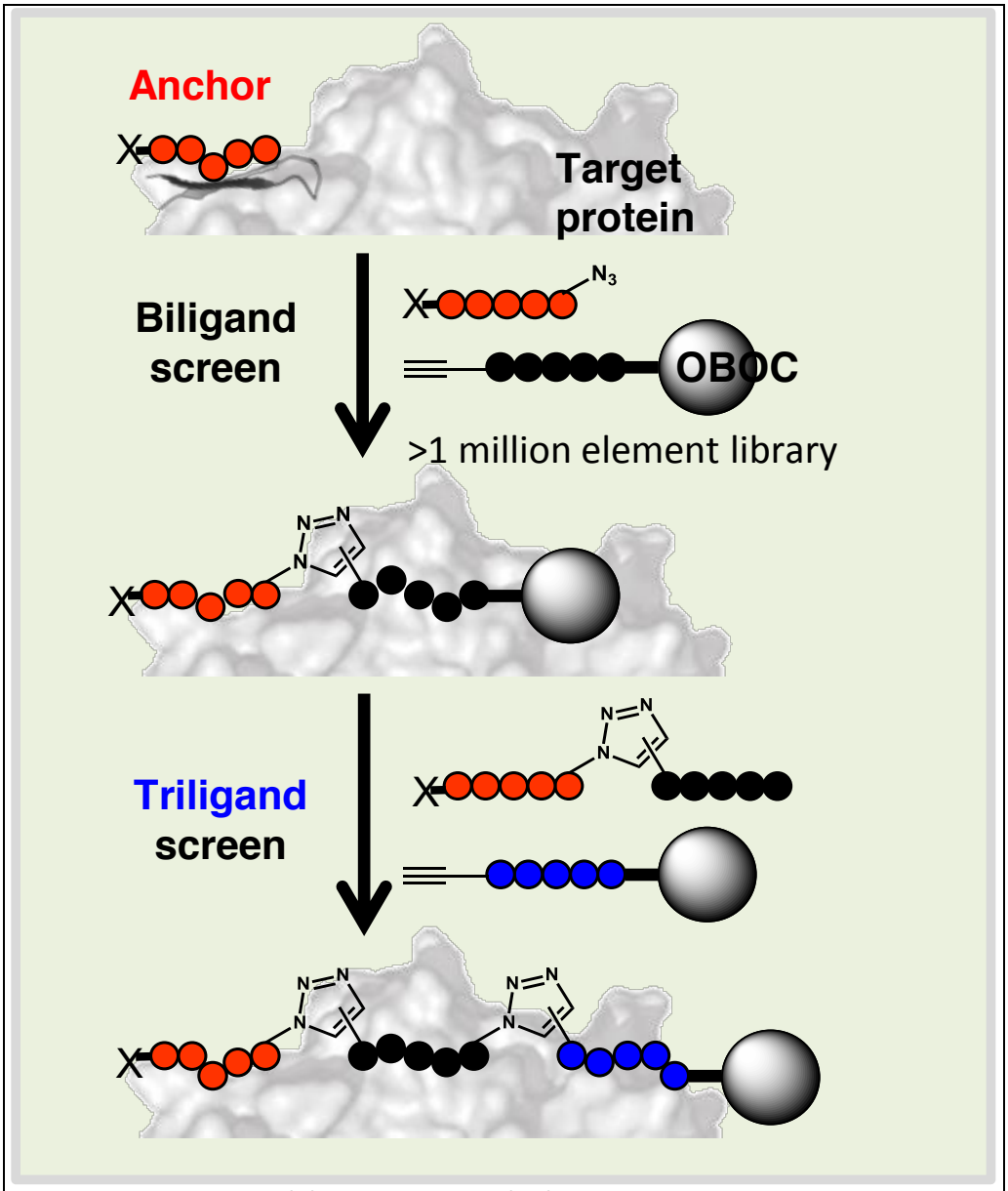


Figure 2-1: Iterative In Situ Click Screening Core Technology.
An anchor ligand that binds to the protein target can be appended with a click handle. In the presence of the protein and a OBOC library appended with the opposite click handle, the anchor can click onto the library to form a biligand. The click only occurs when the anchor and library bead are held long enough on the protein surface, so the protein selects ligands with high affinities and selectivities. This process can be repeated as many times as necessary.

2.2 Materials and Methods

2.2.1 *Standard Materials*

All amino acids were purchased from Aapptec as the FMOC carboxylic acid with the standard TFA side-chain protecting groups. HATU (2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) and PEG_5 (Fmoc-NH-PEG_5-CH_2CH_2COOH, Fmoc-18-amino-4,7,10,13,16-pentaoxaoctadecanoic acid) were purchased from ChemPep. DIEA (diisoproylethylamine), TES (triethylsilane), and TFA (trifluoroacetic acid) were purchased from Sigma. TentaGel beads were purchased as 90µm S-NH2 beads, 0.29mmol/g, $2.86x10^6$ beads/g from Rapp Polymere (Germany), and Rink Amide resin was purchased from Anaspec.

2.2.2 *Peptide Library Construction*

Peptides and peptide libraries were synthesized by hand until the summer of 2009, when they were then synthesized on a Titan 357 split-and-mix automated peptide synthesizer (Aapptec) via standard FMOC SPPS coupling chemistry[5] using 90µm TentaGel S-NH_2 beads. Libraries contain 18 D-stereoisomers of the natural amino acids, minus cysteine and methionine (unless otherwise stated), at each of five randomized positions and an azide or alkyne *in situ* click handle. At least a five-fold excess of beads is used when synthesizing libraries to ensure efficient oversampling of each sequence. Amino acid side-chains are protected by TFA labile protecting groups that are removed all at once following library synthesis.

Figure 2-2: OBOC Peptide Library *constructed on TentaGel Resin. Where X is comprised of all of the naturally occurring D – amino acids except Cys and Met.*

2.2.3 *Bulk Peptide Synthesis*

Bulk synthesis of peptide sequences was performed using standard FMOC SPPS peptide chemistry on either the Titan 357 automated peptide synthesizer (AAPPTEC) or a Liberty 1 microwave peptide synthesizer (CEM Corporation). The typical scale was 300mg on Rink Amide Resin, unless otherwise noted. Peptides were cleaved from the beads with side-chains deprotected using a 95:5:5 ratio of TFA: H_2O: TES. The peptides were purified on a prep-scale Dionex U3000 HPLC with a reverse-phase C18 column (Phenomenex).

2.2.4 *Typical Screening Protocol for Fluorescent Dye-labeled Protein Target Detection*

Hit beads in the initial OBOC screens were detected via a fluorescent probe attached to the protein target of interest. The target protein was labeled using an Alexa-Fluor 647 Microscale Protein Labeling Kit, following all manufacturer's instructions. The activity of the target enzymes was then tested before screening to ensure that the dye label did not disturb function or folding.

Screens were conducted using a OBOC library of 5-amino-acid-long peptides composed of the D - isomers of 19 naturally occurring amino acids (no Cys, for stability reasons). 100mg of dried library was weighed for screening (~280,000 unique sequences, ~42% sampling of sequence space) and swelled in 1xTBS buffer (25mM Tris, 150mM NaCl, 10mM $MgCl_2$, pH = 7.5) containing 0.05% NaN_3, 0.1% BSA, and 0.1% Tween-20 (TBSTBNaN_3). The library was then blocked for one hour in this buffer, then 50nM protein in 1.5mL TBSTBNaN_3 was added, the screen wrapped in foil to protect the light-sensitive dye label, and incubated overnight on a 180° shaking arm. In the morning, the buffer containing the protein was drained from the beads, which were then washed three times with TBSTBNaN_3, three times with TBS + 0.1% Tween-20 (TBST), then three times with 1xTBS. The beads were then dried on a vacuum and spread to a monolayer on approximately 10 clean microscope slides for about 10mgs of beads per slide. The slides were imaged on a GenePix

Pro 5.1 microarray scanner at 635nm to view beads containing bound fluorescent protein target. The dye saturated the color signal of the GenePix, and the hit "beads" that were considered appeared white in a sea of red, due to the background auto fluorescence of the TentaGel library (**Figure 2-4**). These hit beads were then removed from the microscope slides using a needle, stripped of protein with 7.5M pH = 2.0 Guanadine-HCl buffer, rinsed in water, and sequenced via Edman degradation on an Applied Biosystems Procise CLC 494 system.

2.2.5 *Typical Screening Protocol for Antibody Signal Amplification Target Only Screens*

100mg of library beads were prepared, washed and blocked for one hour as for the fluorescent detection screen. The library was then incubated with about 50nM, which differed slightly based on the exact screen, of protein overnight at room temperature. In the morning, the library was washed five times with 1xTBS + 0.1% BSA + 0.1% Tween-20 (1xTBSTB). The primary anti-protein target antibody was incubated with the library for 1 hour, washed five times with the 1xTBSTB buffer, then incubated with the secondary anti-mouse alkaline-phosphatase antibody for one hour. The library was then washed five times with the TBSTB buffer, three times five minutes each in high salt buffer (1xTBS + 600mM NaCl), and five times in 1xTBS. The screen was developed with a two part BCIP/NBT system: 10mL TBS + 26 µL BCIP + 13 µL NBT. This detection cocktail was mixed with the library beads, which were poured into a large polystyrene dish for visualization of the color change under an optical microscope. Hit library members appear as dark purple among the normally clear beads (**Figure 2-6**), and are removed using a pipet. They are washed, stripped, and sequenced as above.

2.2.6 *Typical Screening Protocol for an Anti-Screen*

The library beads (typically 250-500mg) swelled in 1xTBS were blocked 2 hours to overnight in 5% milk in 1xTBS, washed three times with 1x TBS, then incubated with an off-target protein in 0.5% milk in 1xTBS for one hour on the shaking arm at room temperature. The beads were washed three times with 1x TBS, then incubated with the anti-off-target protein - alkaline phosphatase conjugated antibody in 0.5% milk for one hour at room temperature. The antibody used here must be the same antibody used in the target screen in order to ensure that the library members that bind to this antibody are removed and not mistaken for hits. The library resin was then washed three times with high salt buffer and let shake for one hour in high salt at room temperature before being washed three times with BCIP buffer (100mM Tris-Cl, 150mM NaCl, 1mM $MgCl_2$, pH = 9.0) and developed by adding 15mL BCIP buffer plus 13µL BCIP and 26µL NBT. The beads that turned purple bound to both mutant and wildtype protein or to the detection antibodies, and were discarded. The beads that remained clear after this step were picked and washed with guanidine-HCl to remove any bound proteins.

The off-target protein can be a different version of the target, such as a wildtype protein when detecting for a mutation, or a protein lacking a certain domain or post-translational modification of interest, such as a phosphorylation site or glycosylation. Anti-screens can also be designed to clear against any number of interferents, such as whole human serum, to remove any generally sticky peptide sequences. For these anti-screens, the antibody used for detection is an anti-whole human serum antibody followed by a secondary alkaline-phosphatase conjugated antibody.

2.2.7 *Typical Target Screening Procedure During a Multi-Step Screen (Figure 2-3)*

The library beads were blocked in 5% milk in 1x TBS for two hours to overnight. They were then washed three times with 1x TBS. The target protein and anchor peptide or small molecule targeting agent[6] were pre-incubated in 3-5mL of 0.5% milk in an approximately a 10:1 ratio, ensuring the same concentration of anchor peptide used in the preclear. This solution was added to the blocked library beads and incubated for either 5 hours or overnight to allow an *in situ* click reaction to occur. In the morning, the beads were washed three times with 1x TBS, then incubated with the same dilution of an anti-target alkaline phosphatase conjugated antibody that was used in the anti-screen in 0.5% milk for one hour. The beads were then washed three times with a high salt TBS, then incubated on the shaking arm for one hour with the high salt buffer. They were then washed three times with BCIP buffer and developed as previously. Hit beads turned purple and were removed and washed in NMP for four hours to decolorize, then guanidine-HCl to denature and remove and remaining protein.

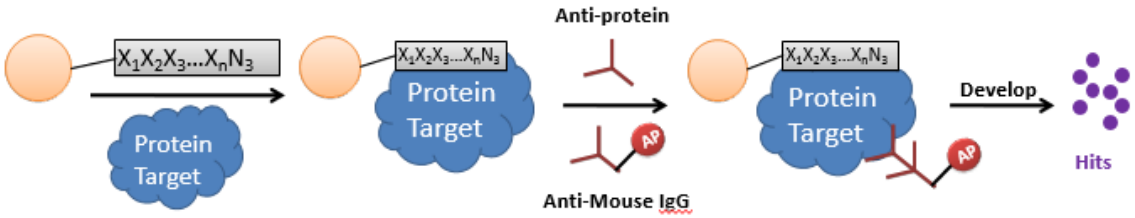


Figure 2-3: Typical Antibody-Detected Target Screen. *The library is incubated with the protein target, which is detected via antibodies conjugated to alkaline phosphatase. The screen is developed with BCIP/NBT, and hit beads turn purple.*

2.2.8 *Typical Screening Protocol for a Preclear*

Swelled library beads (250-500mg) were blocked overnight in 5% w/v dried non-fat milk in 1x TBS, then washed with 1x TBS three times. The beads were incubated with a µM solution of any anchor peptide or small molecule for one hour, then washed 3x with 1xTBS. Five milliliters of either a 1:10,000 dilution of streptavidin-alkaline phosphatase conjugate in 0.5% milk in TBS or an anti-biotin antibody were added to the beads and incubated with shaking at room temperature for one hour. If the anti-biotin antibody was used, a secondary antibody conjugated to alkaline

phosphatase was then incubated with the library for 1 hour after it was washed three times in 1xTBS. The beads were washed with a high-salt TBS buffer three times, then were left to shake in high salt buffer for one hour. The beads were then washed three times with BCIP and developed as for the anti-screen. After one hour, the purple beads were removed by pipette and discarded. The remaining beads were incubated in NMP 4 hours to remove trace purple precipitate from the BCIP/NBT reaction, then were washed five times with methanol, five times with water, five times with TBS and blocked overnight in 5% milk.

2.2.9 *Typical Screening Protocol for a Click Product Screen*

The beads that pass through the target and anti-screen were washed three times with 1x TBS. They were then incubated with a 1:10,000 dilution of either streptavidin – alkaline phosphatase conjugate or anti-biotin antibody (whichever was used in the preclear) in 0.5% milk for one hour. The beads were washed three times with high salt TBS then let shake for one hour with high salt buffer before being washed three times with BCIP buffer and developed as previously. The beads that turned purple contained the anchor peptide covalently bound to the bead and had formed a protein-catalyzed *in situ* click reaction. These beads were collected and stripped with guanidine-HCl for one hour, washed ten times with water, and sequenced via Edman degradation.

2.2.10 *Peptide Sequencing Strategies*

The OBOC peptide library sequencing method most commonly used by Caltech is Edman degradation. This process involves treating a peptide with a free amine terminus with phenylisothiocyanate, which reacts stoichiometricly with the N-terminus of the peptide to form a phenylthiocarbamyl (PTC)-peptide derivative. This PTC derivative is then treated with TFA to cleave it off from the rest of the peptide, leaving behind a new N-terminus to react during the

next cycle. Meanwhile, the PTC amino acid is then analyzed via HPLC, and the peak is compared to standards of all of the PTC-amino acids in order to determine the residue. One cycle per amino acid residue is performed and analyzed, providing the sequence of the peptide on the hit library bead[7]. This method is slow, but highly accurate and has been automated by Applied Biosystems into the Procise CLC 494 Automated Edman Degradation machine used by Caltech.

Hit peptide sequences can also be determined through MALDI-TOF/TOF MS. For this method, the library must be specially made. The peptide must be attached to the library through a methionine amino acid, and no other methionine can be present in the library. The isobaric amino acids, isoleucine and leucine, lysine and glutamine, are doped by anther amino acid in order to properly call the sequence by mass. Glutamine is doped with a 6% molar equivalent of glycine, and isoleucine is doped with a 7% molar equivalent of alanine. While reading the mass of these amino acids on the MALDI, any residue that has one of these amino acids can be distinguished by the presence or absence of the small satellite parent mass corresponding to the same sequence plus glycine or alanine[8].

In order to sequence the library hit by MALDI-TOF/TOF, the bead is first treated with cyanogen bromide in order to cleave the peptide from the bead at the methionine amino acid. It can then be dissolved in MALDI matrix and spotted onto the plate. The peptide parent peak is first discovered using MALDI-TOF, then is fragmented again in order to break it up into smaller amino acid ions. These ions can be analyzed using standard peptide MS techniques to determine the sequence[8].

2.3 Results and Discussion

2.3.1 *Screening via Fluorescent Dye-labeled Protein Target Detection*

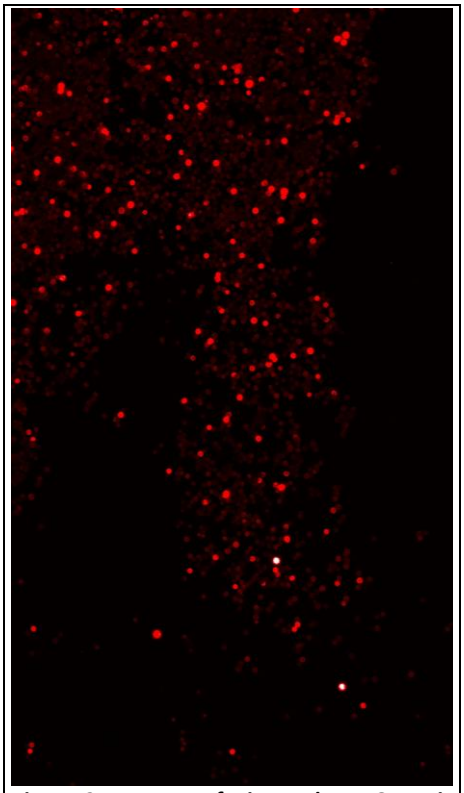


Figure 2-4: Image of Hit Beads on GenePix Microarray Scanner*. The bright white beads are saturating the fluorescence and are considered "hits" above the background TentaGel auto fluorescence.*

The initial OBOC peptide screening strategies developed by Heather Agnew[1] relied on a fluorescent dye-labeled protein in order to detect hit binding. The target protein of interest was labeled with a dye, and any library beads that bound to the target were detected on a GenePix microarray reader. As seen in **Figure 2-4**, the TentaGel library beads also auto-fluoresce, meaning that all screens conducted in this fashion were highly subjective, and the hit quantity depended entirely on the gain settings of the microarray. AlexaFluor-647 was also the only dye that was used, as the beads auto fluoresce the least in the range of this dye. These hits were mostly picked using a light microscope, meaning that the images from the microarray had to be used as a "map" to guide the bead picker to the correct clear bead on a slide of thousands. This process was highly inefficient, requiring up to an hour to pick each individual hit bead. These picked hits were always imaged again on the GenePix to ensure that each bead that had been selected was a highly fluorescent bead, indicating that the correct one had been chosen based on the map. It was possible to use a COPAS automatic bead sorter to separate out the hit beads, though one was not available at Caltech.

The sequences from a typical fluorescent target screen are shown in **Table 2-1**. The hits were generally dominated by the positively charged residues, arginine and lysine. This overwhelming charged signal is most likely due to the overall (-3) charge on the AlexaFluor 647

dye,[9] which is attracting the positively charged amino acid sequences and creating a significant level of noise in the final hits. Most screens had to be run many times in order to find enough quality hit sequences, meaning ones that did not contain almost exclusively arginine and lysine residues, because of this high background. Generally, a hit that contained 3 or more positively charged amino acids was considered background and removed from the pool. One screen rarely yielded more than a handful of hits that appeared to be binding to the surface of the protein and not just to the dye.

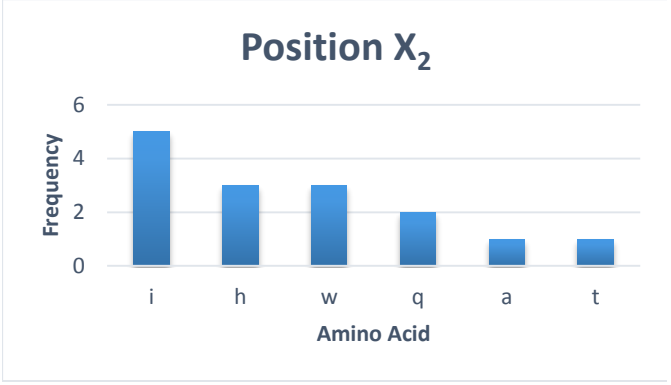


Figure 2-5: Histogram of Position X_2 in PSA Screens. *The hits from multiple PSA screens were pooled and analyzed. This sample chart shows the frequency of an amino acid at position X2 in the library, and is used to synthesize the focused library.*

Focused screens were also used in order to hone in on target-binding peptide sequences. The focused libraries used in these screens were designed based on histograms of the amino acids that were seen at each library position, meaning X_1 -> X_5 as seen in **Figure 2-2**, after the removal of the dye label background sequences. As can be seen in **Figure 2-5**, in this particular PSA screen, there were only six amino acids that were seen at position 2, so only these six amino acids were built into the focused library at position 2. This reduction in total amino acids present in each position allowed for the synthesis of a much smaller library that could be oversampled in each screen to permit a more thorough sampling of the sequence space. Only about 100mg of beads were usually screened, but 100mg could frequently oversample the sequence space of a focused library, compared to that of naïve libraries where less than half of the space was sampled. Due to this increase in sequence space sampling, focused libraries were generally extended by one or two amino acid positions in the hopes that a slightly longer peptide would have a higher affinity

and selectivity for the protein target. The screening was then repeated with the focused libraries, and the same process for analyzing hits was repeated until the peptide sequences converged in sequence homology and produced a peptide ligand that showed near μM affinity for the protein target.

This convergence frequently required the use of two to three separate focused libraries with accompanying screening and sequencing. The overall time required to determine one peptide ligand that bound to the target protein of interest could easily take more than six months. These ligands also regularly bound in the range of low μM affinities, which are generally considered to be fairly weak binders.

Table 2-1: Screens from Sample Target Screen Using Fluorescent Protein Detection. *This screen was performed against PSA protein labeled with AlexaFluor 647 dye. Note the high prevalence of "r" and "k" positively charged amino acids. See Figure 2-2 for a visualization of the X amino acid positions on bead.*

X_1	X_2	X_3	X_4	X_5
y	r	r	r	r
r	i	f	r	r
r	f	l	r	a
r	r	k	r	f
m	r	r	w	r
r	r	r	w	p
r	r	w	i	r
r	r	r	f	l
r	l	r	w	r
r	f	r	i	r
l	s	r	r	r
r	r	r	y	t
r	r	m	r	w
r	r	k	p	r
f	y	r	r	r
r	k	w	l	w
k	r	r	m	r

2.3.2 *Screening via Antibody Signal Amplification Target Only Screens*

Detecting hit peptides via fluorescence was a very time-consuming process in which the high noise from the overwhelming presence of positively charged amino acids meant that very little meaningful output was obtained. For this reason, a new method of screening was developed using a tag-less protein to switch the screening focus from the charged dye label back to the target. This technique relied on anti-target antibodies conjugated to alkaline phosphatase, which is an enzyme that can form a dark purple precipitate in the presence of its BCIP/NBT substrate. This meant that any "hit" now showed up as a very dark purple bead in a sea of clear. The label-less detection technique, therefore, provided the additional benefit of a colorimetric readout of a hit, allowing for the much easier separation of these beads from the rest of the library.

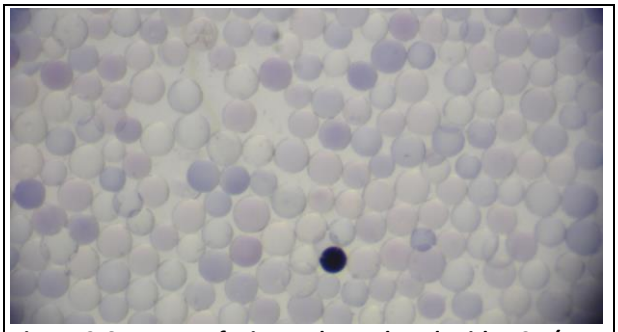


Figure 2-6: Image of Hit Bead Developed with BCIP/NBT. *High background lighter purple surrounding beads could be removed through later preclear and antiscreen steps.*

As can be seen in sample screen results in **Table 2-2**, the high prevalence of positively charged amino acids is gone. In fact, the comparison between **Table 2-1** and **Table 2-2** is startling, considering that the only difference between these two screens is the target detection method. This demonstrates that the dye label was having a dramatic effect on the quality of hit sequences and was responsible for much of the large time investment that was devoted to screening. This huge reduction in noise now meant an instant reduction in the number of screens that needed to be run and sequenced in order to see homology. The colorimetric hit visualization also permitted larger numbers of beads to be screened much faster, so the overall number of library sequences that were sampled went up even though fewer screens were run. One BCIP/NBT-developed screen could sample the same number of beads as up to five different fluorescent screens in less

time, as all of the hits could be picked in the time it used to take to pick one. With this increase in both sampled sequence space and in the overall signal to noise seen in the sequences, hit quality and screening speed improved dramatically in a much shorter overall time.

Table 2-2: Screens from Sample Antibody Amplification Screen Using BCIP/NBT Protein Detection*. Screen was performed against unlabeled (PSA), detected with PS2 mouse mAb anti-PSA antibody and anti-mouse-AP secondary antibody with BCIP/NBT readout.*

X_1	X_2	X_3	X_4	X_5
n	g	m	e	d
e	t	q	m	d
w	t	d	e	m
s	e	d	d	t
a	n	d	e	e
n	y	d	p	e
G	n	m	d	d
e	d	v	l	i
f	e	n	d	a
e	i	n	e	l
v	e	f	G	e
e	h	d	a	y
d	e	t	a	t
i	w	n	m	e
y	d	d	s	l
d	d	e	a	G
e	n	t	i	d

2.3.3 Introduction of an Anti-screen

The antibody development technique dramatically improved the quality of hit peptides by visual inspection (**Table 2-1** versus **Table 2-2**), but also introduced a hidden source of noise into the screens. The presence of several different antibodies and a new detection agent in the screen itself provided more "off-target" sources of library binding. This was conclusively demonstrated by Steve Millward while screening for an Akt capture agent. He developed a biligand using the standard *in situ* click chemistry technique with antibody development, and proceeded to test the affinity of this ligand via SPR. The SPR was set up to immobilize an anti-FLAG antibody (the same used in screening) to the flow cell in order to capture the much less stable Akt protein that might

not survive the required EDC/NHS coupling step. A blank flow cell of only Anti-FLAG antibody without Akt was used as a chip blank. The data from these SPRs is seen in **Figure 2-7**. The sensorgram on the right shows binding to the Akt, as to be expected, but the sensorgram from the blank flow cell on the left shows an identical signal. In conjunction with data (not shown) from the anchor ligand that has almost no binding to the anti-FLAG flow cell, we can conclude that the biligand is actually binding to the anti-FLAG antibody, present in both of those flow cells, and not to the desired Akt target protein. It is only logical that we would see "hits" of peptide sequences that bind to these antibodies, because the presence of the detection antibody bound to a library bead would show BCIP precipitation exactly like the presence of the target protein. A new screening step was needed that would remove the signal seen from the binding of these other proteins used in the screening process.

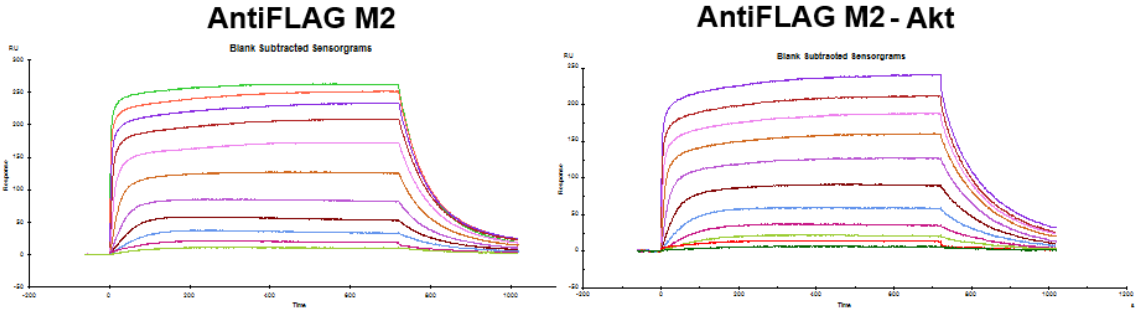


Figure 2-7: SPR Data from Akt Biligand. Data from Steve Millward's Akt Biligand Capture Agent. *An anti-FLAG antibody was immobilized onto an SPR chip via standard EDC/NHS coupling techniques. It was used to capture a FLAG-tagged Akt protein for testing. As seen from the figure on the left, the capture agent bound equally as well to the flow cell immobilized with only Anti-FLAG antibody, supposed to be the chip blank, as to the one on the right that also contained immobilized Akt indicating that the biligand's affinity actually stems from the Anti-FLAG antibody and not the Akt protein target.*

Around this time, there was interest in developing capture agents for proteins containing post-translational modifications, such as phosphorylations or glycosylations. It was hypothesized that hits specific for a post-translational modification could be discovered by screening against the protein target containing the modification, then *anti-screening* against the protein target with

the post-translational modification removed, since everything else in the screen would be identical (**Figure 2-8**). These screens entailed first "target screening," as per usual antibody detection screens, to find all of the hit beads that have an affinity for the target. These beads were then be stripped of their purple color and bound proteins and incubated with the off-target protein that had the post-translational modification removed. Any purple hits from the anti-screen were thrown out as not specific for the modification, since they demonstrated binding in a screen that did not contain the site of interest. This new screening step has the added benefit of removing all of the hits that also have an affinity for the antibodies or developing solution that was used in the screen. An anti-screen like this would have prevented the development of a biligand with an affinity for the anti-FLAG antibody, as these hits would have been detected in both the target and the anti-target screen, and would have been discarded.

The anti-screen is an important step that is now incorporated into each screen that is run in the lab, and is responsible for a significant reduction in background hits. For example, an anti-screen that was run for the PSA protein eliminated 91% of the hit beads from the target screen, indicating that approximately 91% of what was previously considered to be a target hit was just background. For visualization purposes (**Table 2-3**), this means that a screen run with 250mg of beads went from 167 hits down to 15 after this step. This cut down on not only sequencing and hit analysis/testing time, but also eliminated the time that was usually spent trying to tease out signal from noise. Focused screens were also no longer necessary, as that step was designed to help enrich for signal, eliminating a significant chunk of time necessary for developing a capture agent.

Current screening protocols have evolved significantly to include stringent anti-serum anti-screens in order to make capture agents that can function in the most complex mediums, such as out of blood and in cells. For these anti-screens, the decolorized target hit beads are

incubated with anywhere from 1% - 25% human serum to remove even the marginally sticky peptides from the pool of potential candidates.

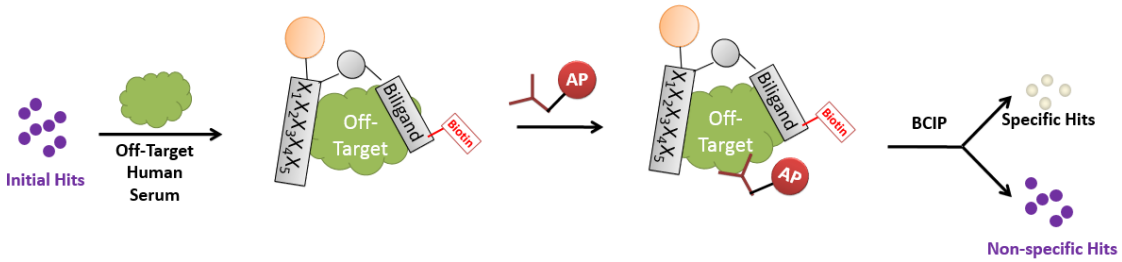


Figure 2-8: Sample Anti-Screen Step. *The target hits are stripped of target, decolorized, and incubated with an off-target or general interferent such as human serum. The off-target is detected with an off-target antibody (the target antibody should also be included if it is different), and these purple beads are non-specific binders, which are removed. The clear beads are specific for the target.*

2.3.4 Introduction of a Click Product Screen

The *in situ* screening process has an inherent screening advantage that had not yet been exploited. A covalently-linked product is formed on the surface of the bead during the screen that can be detected separately from target binding. This means that in addition to probing the library for beads that bind to the target, the library can be searched additionally for the presence of the *in situ* click product – a completely complementary screen.

Once an anchor ligand has been discovered, the next step in the *in situ* screening process (**Figure 2-1**) involves the clicking of a new peptide ligand onto this anchor ligand. In order to accomplish this, the anchor peptide is appended with a click handle and pre-incubated with the target protein, and then both are incubated with the OBOC library. This step searches for a library peptide that binds in close proximity to the anchor peptide on the surface of the protein target, and will "click" onto the anchor if held in position long enough. This click reaction covalently attaches the anchor peptide onto the library bead. By first appending the anchor peptide with a biotin tag, the presence of the anchor peptide on bead, or the ability of this library candidate to "click" to the anchor, can be probed independently of the presence of the target on bead. These screens involve harsh, denaturing wash steps that ensure that everything not covalently attached

to the library will be removed and not detected by either the streptavidin conjugated to alkaline phosphatase or an anti-biotin antibody. These detection agents will bind to the biotin label on the anchor that will only be present after a covalent reaction has occurred, and can therefore detect which library members have formed a click product (**Figure 2-9**).

Table 2-3: PSA Screening Statistics. *These hit bead statistics are taken from a screen against PSA. The percent column indicates the percent of beads that passed from one stage of the screen to the next.*

	Beads	**Percent**
Start	375,500	
Target Screen	167	0.04%
Anti-screen	15	9%
Product Screen	7	47%

Continuing the comparison with the PSA screens from above, only 7 of the 15 remaining beads after the anti-screen showed the presence of a click product. The other 8 beads could very easily have been hits that would be a different anchor ligand – a peptide ligand that is binding specifically to the target protein, but is not close enough to the original anchor for a click to form. The sequences from these hits, shown in **Table 2-4**, are very nearly identical peptides, and contrast sharply with the previously identified hits from the anti-screen in **Table 2-2**. This indicates that the sequences are more than likely all binding very strongly to the exact same location and in close proximity to the anchor ligand, allowing for the formation of the click product.

Table 2-4: PSA Hit Bead Sequences from Product Screen. *The product hits shown in Table 2-3 were sequenced. There is an enormous sequence homology, meaning that the same part of the target is being targeted. The end of the last sequence and the 7th hit were lost due to machine error.*

X_1	X_2	X_3	X_4	X_5
Y	G	w	r	e
Y	d	w	r	q
L	G	w	r	e
e	G	w	r	e
a	d	w	r	q
a	G	-	-	-

The product screen is an elegant step in the screening process that allows for the very specific narrowing of the sequence space. It has become such a huge part of the success of the OBOC capture agent development process that naïve anchor screens, which inherently cannot

include product screens, have been completely eliminated. This switch to all *in situ* click screens has greatly increased both the specificity and affinity of the original anchor ligands, dramatically improving the quality of the final PCC agent. Details of the rationale and results from these more targeted screens can be seen in Chapter 3.

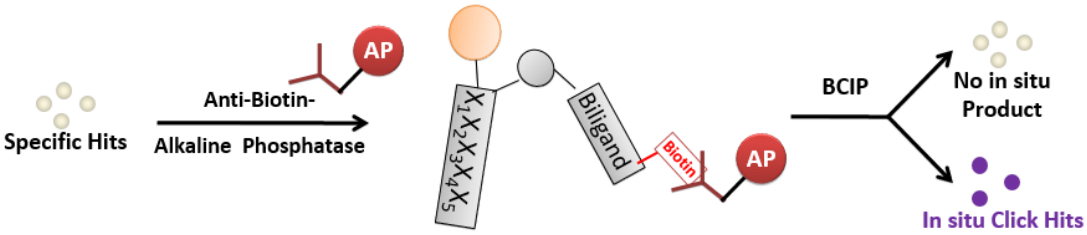


Figure 2-9: Sample Product Screening Step. *The specific hits that survive the anti-screen are stripped of all non-covalent binders and incubated with anti-biotin alkaline phosphatase (or streptavidin). The purple hits from this screen indicate those in which a click reaction has covalently attached the anchor ligand onto the library bead.*

2.3.5 Introduction of a Preclear

Three of the candidates from **Table 2-4** were scaled up. In order to do this, the secondary arm is clicked to the original anchor using Cu(I) to form a "biligand," and is tested for binding to the PSA protein. Unfortunately, none of the biligand candidates shown in **Table 2-4** demonstrated binding to the PSA protein in either ELISA assays or SPR, even though the anchor ligand by itself was still able to bind (indicating that all of the parts of the assays were working). The secondary ligands themselves also did not show any binding to the PSA protein, independent of the anchor ligand. These ligand sequences from the click screen, however, were very homologous, indicating that they were all binding in the same place, which was somewhere they could click onto the anchor peptide. It would be impossible to see that level of similarity in the hit sequences, otherwise. Unfortunately, during the screening process, the anchor ligand itself is present in ten times higher quantity than the protein target, and can also bind to the library beads. It was hypothesized, therefore, that the anchor ligand itself bound to those library sequences tightly enough to catalyze the click product that was detected in the final screen. This scenario would

explain why the biligands showed no binding to the protein – the anchor could no longer even bind to the target with another ligand, potentially blocking those binding sites. It also explains why the secondary ligands showed no affinity for PSA. They were not ligands that bound to the target, and wouldn't have an affinity for it.

To counter this effect, a new screening step was added at the beginning of the process to remove all of the library peptides that bound to the anchor ligand before the anchor ligand even saw the target protein (**Figure 2-10**). These screens still detect the biotin label on the anchor ligand, and the detection with streptavidin or anti-biotin in this "preclear" step eliminates the need to use these detection agents in the anti-screen. The preclear screens generally remove 1-10% of the library beads, depending on the library, and also reduce the percentage of beads that need to be removed in the anti-screen.

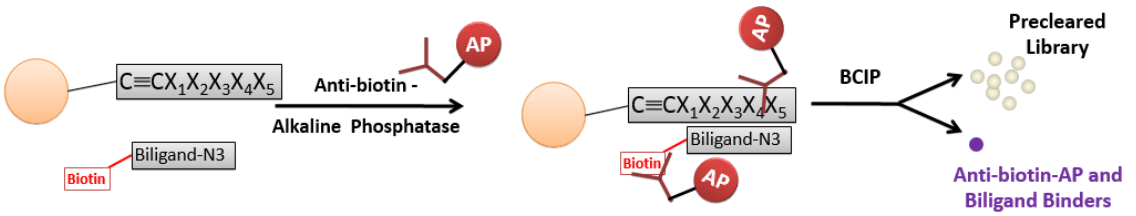


***Figure 2-10: Sample Preclear Screening Step**. The library is incubated with an anchor or biligand, which is then detected with an anti-biotin alkaline phosphatase conjugated antibody (or streptavidin), and developed with BCIP. The purple beads from this screen bind to either the initial ligand or the detection antibody.*

2.3.6 Use of Alkyne Versus Azide Libraries

Throughout the course of technology development, certain seemingly trivial details become important. For the OBOC screens, different libraries and slightly different conditions produced vastly different results. The first issue with the propargylglycine alkyne-containing amino acids surfaced initially after the addition of multiple stages to the screening process. After undergoing more than three rounds of screening, washing and denaturing, the libraries containing the alkyne were no longer able to be successfully sequenced via any method - Edman degradation or MALDI TOF/TOF. The Edman spectra were entirely blank, indicating that the amino acid

residues were probably not cleaving from the beads, and the MALDI TOF/TOF was unable to identify a parent peak that contained the fixed alkyne amino acid. The alkyne-containing amino acid was the N-terminal residue, the first residue that needed to cleave via Edman, and anything modifying this amino acid would affect the cleavage. It was hypothesized that the BCIP/NBT developing solution was modifying these amino acids, which was confirmed by the use of C-terminal alkyne libraries. Even after undergoing four screening steps, the libraries still sequenced correctly using Edman degradation up to the alkyne amino acid. These same library hits, though, were not able to be sequenced using MALDI-TOF/TOF. Because the TOF/TOF would be greatly affected by an unknown change to an amino acid, it was assumed that the alkyne was somehow being modified during these screening steps. For this reason, azide-containing libraries are now always used when undergoing more than three screening steps, unless a C-terminal alkyne library with Edman degradation sequencing is appropriate.

It was also noticed that the libraries that contained a propargylglycine seemed to have more difficult preclears, meaning more purple hits to remove, than the libraries that contained the Lys(N_3) azide amino acid. To test this, two libraries, identical except for their N-terminal azide or alkyne click handle, were blocked in 5% milk in TBS. The libraries were washed three times in TBS, then developed with the BCIP/NBT solution used in the methods section. After 45 minutes, about 5% of the beads in the alkyne library turned bright yellow, indicating binding of the NBT substrate. The azide library did not show this background substrate turnover/binding, and it was assumed that this was related to the sequencing issues with the alkyne libraries. If the NBT substrate is somehow changing or appending to the propargylglycine amino acid, it could explain why the sequences no longer appear as they should during screening.

2.3.7 Typical Flow of Screening

With a multi-stage screening process now in place, the some of the steps need to be conducted in a certain order to achieve the correct results. The first step is the preclear. This occurs before the anchor ligand sees the protein target, and has a chance to form legitimate clicked-hit peptides on bead. These screens look for anything that binds to streptavidin, alkaline-phosphatase, BCIP/NBT, and the anchor peptides. Usually, a screen begins with 300-500mg of library beads, and 1-10% are removed. Typically, any bead that has turned even the lightest shade of purple is removed in order to reduce the overall background as much as possible. This means that any bead that passes through this stage of the screening process has remained clear.

The next step is the target and click-catalyzed screen. The beads that remained clear in the preclear are incubated with the target of interest and the anchor ligand overnight for a click reaction to occur. These beads are then probed for the presence of target. Any bead bound to target will turn purple, and passes through to the next stage of screening. Even though the on-bead click has occurred during this screen, probing for the click product occurs at a later stage.

The hits from the target screen are then decolorized and incubated with an off-target protein or proteins. Any library bead that binds and turns purple in this screen demonstrates off-target interactions with other proteins, and is removed from the pool. At the end of this screen, only beads that remain entirely clear are kept. Even slight purple can indicate undesirable interactions and background binding, and are removed from the pool of hits.

The final screening stage probes for the presence of the clicked product on bead. After harsh denaturing and washing conditions, the beads are probed for the presence of biotin. These beads will turn purple only if biotin is linked to the bead, which is only possible if the *in situ* click reaction was successful. These purple hits have proven to have no affinity to the screening agents in the preclear, an affinity for the target but not off-target interactions in the target and anti-screens, and then have also shown involvement in the covalent click reaction. The clear-purple-

clear-purple pattern of hit detection also ensures that the beads are behaving properly at each stage in the process.

Screens following this pattern now have several produced high-affinity ligands that are very selective to their target of interest. This methodology has an incredibly high success rate that is only getting better as the process continues to grow and develop.

2.4 Conclusions

Over the past ten years in which the project has been in existence, protein-catalyzed capture (PCC) agents have proven to be highly effective detection agents that are incredibly stable and easy to synthesize[1,2,3,6,10]. These agents can be made almost entirely with robotics for ease of scale-up, and the capture agents are highly modular, so the addition of labeling tags is trivial. The exact chemical structures of each of these capture agents are known, eliminating the batch to batch variability that is common with antibodies and can cause multiplexed assays to be expensive and difficult to produce. A spin-off company, InDi Molecular, is in place for commercialization of these agents. PCC agents are also completely stable, demonstrating no degradation upon incubation with mouse liver enzymes, and full functionality after being stored at 65°C as a powder for weeks[2], demonstrating their excellence for use in anything from clinical work to detection of diseases in third world countries[3]. The technology development discussed in this chapter has revolutionized how screening for PCC agents occurs, and the robustness of these techniques has provided a solid foundation for the rapid discovery of a multitude of additional agents for a wide range of purposes [1,2,3,6,10].

2.5 Acknowledgements

The work described in this chapter was done in conjunction with Heather Agnew and Steve Millward. The initial OBOC screening techniques used were developed by Heather, and the colorimetric screening development was done by Steve. The rest of the work described herein was performed in conjunction with both Heather and Steve, as well as Arundhati Nag and Rosemary Rohde.

2.6 References

1. Agnew, H. D.; Rohde, R. D.; Millward, S. W.; Nag, A.; Yeo, W.-S.; Hein, J. E.; Pitram, S. M.; Tariq, A. A.; Burns, V. M.; Krom, R. J.; Fokin, V. V.; Sharpless, K. B.; Heath, J. R., Iterative In Situ Click Chemistry Creates Antibody-like Protein-Capture Agents. *Angewandte Chemie International Edition* **2009,** *48* (27), 4944-4948.
2. Farrow, B.; Hong, S. A.; Romero, E. C.; Lai, B.; Coppock, M. B.; Deyle, K. M.; Finch, A. S.; Stratis-Cullum, D. N.; Agnew, H. D.; Yang, S.; Heath, J. R., A Chemically Synthesized Capture Agent Enables the Selective, Sensitive, and Robust Electrochemical Detection of Anthrax Protective Antigen. *ACS Nano* **2013,** *7* (10), 9452-9460.
3. Pfeilsticker, J. A.; Umeda, A.; Farrow, B.; Hsueh, C. L.; Deyle, K. M.; Kim, J. T.; Lai, B. T.; Heath, J. R., A Cocktail of Thermally Stable, Chemically Synthesized Capture Agents for the Efficient Detection of Anti-Gp41 Antibodies from Human Sera. *PLoS ONE* **2013,** *8* (10), e76224.
4. Stenman, U.-H.; Leinonen, J.; Zhang, W.-M.; Finne, P., Prostate-specific antigen. *Seminars in Cancer Biology* **1999,** *9* (2), 83-93.
5. Coin, I.; Beyermann, M.; Bienert, M., Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nature protocols* **2007,** *2* (12), 3247-56.
6. Nag, A.; Das, S.; Yu, M. B.; Deyle, K. M.; Millward, S. W.; Heath, J. R., A Chemical Epitope-Targeting Strategy for Protein Capture Agents: The Serine 474 Epitope of the Kinase Akt2. *Angewandte Chemie International Edition* **2013,** *52* (52), 13975-13979.
7. Procise Protein Sequencing System. http://tools.lifetechnologies.com/content/sfs/manuals/cms_041125.pdf (accessed May 1).
8. Lee, S. S.; Lim, J.; Tan, S.; Cha, J.; Yeo, S. Y.; Agnew, H. D.; Heath, J. R., Accurate MALDI-TOF/TOF sequencing of one-bead-one-compound peptide libraries with application to the identification of multiligand protein affinity agents using in situ click chemistry screening. *Anal Chem* **2010,** *82* (2), 672-9.
9. Sobek, D. J.; Aquino, C.; Schlapbach, D. R., Analyzing the Properties of Fluorescent Dyes Used for Labeling DNA in Microarray Experiments. In *BioFiles*, Simga-Aldrich: Online, 2011; Vol. 6.3.
10. Millward, S. W.; Henning, R. K.; Kwong, G. A.; Pitram, S.; Agnew, H. D.; Deyle, K. M.; Nag, A.; Hein, J.; Lee, S. S.; Lim, J.; Pfeilsticker, J. A.; Sharpless, K. B.; Heath, J. R., Iterative in situ click chemistry assembles a branched capture agent and allosteric inhibitor for Akt1. *Journal of the American Chemical Society* **2011,** *133* (45), 18280-8.