Chapter 2: Development of branched capture agents against bovine carbonic anhydrase II

2.1 Introduction

Biologically[1,2] or chemically synthesized protein capture agents that can detect protein in diagnostics platforms and can be easily functionalized remain a very lucrative research area. Peptides discovered through phage display[3-5], antibody or protein[6] fragments[7] and nucleic acid[5,8,9] aptamers represent some classes of protein capture agents. Each class of artificial capture agents have unique advantages and disadvantages[6,10,11].In the past we have reported development of a new class of protein capture agents. An in situ protein-catalyzed multiligand capture agent was developed against bovine carbonic anhydrase II. The capture agent showed nanomolar binding affinity and could be used as an antibody substituent[12]. The PCC agent developed was linear, meaning that peptides, acting as modules, were joined either N terminal to N terminal or C terminal to N terminal. In the following chapter we extend the earlier work to develop branched capture agents against bCAII starting from the same monoligand peptide. Branching of peptides introduces geometric constraints to the PCC agent and reduces the conformational entropy of the molecule[13], which should result in higher affinity. Branching of peptide have also been reported to increase the protease resistance[14]. In the current work we explore two approaches to find the most appropriate branching point of an anchor peptide towards branched ligands via in situ click screen. In one approach a branching point is introduced in the biligand stage through screening of peptide libraries containing azido-modified artificial amino acids at random positions. In the second approach, an amino acid within a linear biligand is replaced with an artificial amino acid containing an azide or acetylene functional group. Hydrophobicity and structural similarity are used as criteria for the replacement. The branched biligands are further used in triligand screens to develop branched triligands. Some of the branched triligands are found to have ~3 fold higher affinity than the linear triligand developed earlier. We also develop a mimic of the 1,5 disubstituted 1,2,3 triazole linkage and determine if the 1,4 vs 1,5 orientation of the peptide arms in the developed multivalent ligands have any effect on the affinity for the protein.

2.2 Materials and Methods

2.2.1 Materials

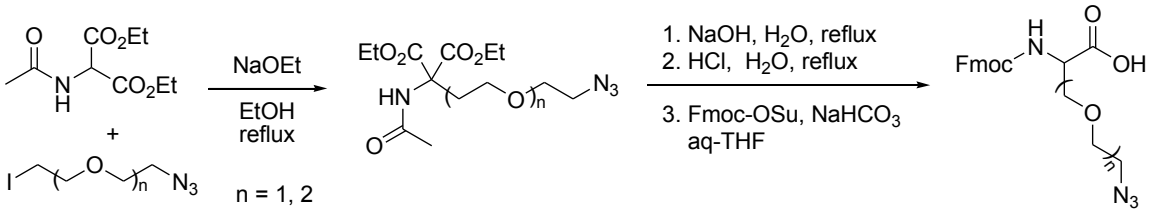
Fmoc D amino acids were purchased from Anaspec (San Jose, CA) and AAPPTec (Louisville, KY). TentaGel S-NH_2 resin (diameter 90 μm, capacity 0.28 mmol/g) used in OBOC library construction was bought from Anaspec (San Jose, CA). Bulk peptide synthesis was done either on Biotin NovaTag™ resin (EMD) or Rink Amide MBHA resin (Anaspec). Peptide synthesis reagents NMP (1-methyl-2 -pyrrolidinone), HATU (2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylammonium hexafluorophosphate) and DIEA (N,N′-diisopropylethylamine) were bought from EMD Chemicals, Inc. (Gibbstown, NJ), ChemPep (Miami, FL), and Sigma-Aldrich (St. Louis, MO) respectively. The target protein Bovine carbonic anhydrase II from bovine erythrocytes was bought from Sigma-Aldrich.

2.2.2 Methods

2.2.2.1 Synthesis of azide and alkyne containing artificial amino acids

Artificial amino acids 21- 35 (Figure 2.7) containing either the free azide functionality or alkyne functionality were synthesized. Amino acids **21**-**27**, **34**-**35** were synthesized following literature methods[15-17]. Amino acids **28**-**33** were developed by modification of literature methods[18,19]. The details of synthesis of amino acids **28**-**33** are as follows:

Figure 2.1: Synthesis of unnatural amino acids containing PEGylated azide side chains.



To a solution of pre-dried diethyl acetamidomalonate (7.35g, 3.3mmol), sodium ethoxide solution (0.82 g, 12 mmol in 40 ml of ethanol) was added at ambient temperature and stirred for 2 hours at 40°C. Azidoalkyliodide (4.0 mmol) was added and the reaction mixture was refluxed at 85 °C for 24 hours. The mixture was extracted with ethyl acetate, washed with water and brine, dried over $MgSO_4$ and then concentrated. Flash column chromatography (hexane:ethyl acetate = 2:1) was performed to give a purified product diethyl 2-acetamido-2-(3-azidopropyl) malonate (> 80%) as a clear oil. Diethyl 2-acetamido-2-(3-azidopropyl) malonate (2.6 mmol) was dissolved in 35 ml of 10% NaOH solution and the mixture was refluxed for 4 hours. The solution was then acidified to pH = 2.0 with concentrated HCl and concentrated under reduced pressure. The residue was dissolved in 70 ml of H_2O and refluxed for 4 hours. After cooling to ambient temperature, the mixture was acidified to pH 3.0 with 10% HCl solution and extracted with ethyl acetate. The organic layers were concentrated to give yellow oil, into which 54 ml of 2.0 N HCl solution was added. The mixture was refluxed for 2 hours and then concentrated in vacuum. The salt was then dissolved in a THF/H_2O (55:45) mixture and $NaHCO_3$ (10 eq.) was added with stirring at 0 °C. Fmoc-OSu (1.1 eq.) was added and the reaction mixture was stirred for 4 hours. THF was removed under reduced pressure and the resulting mixture was extracted with diethyl ether. The combined aqueous layers were acidified to pH 2.0 using concentrated HCl and extracted with ethyl acetate. The organic layers were washed with brine, dried over dried over $MgSO_4$ and then vacuum concentrated to obtain the final product which was further purified by column chromatography to yield the white product.

28 (n = 1):

^{1}H-NMR (400MHz, DMSO-d_6): δ 1.77-1.85 (m, 2H), 1.94-2.02 (m, 2H), 3.30-3.60 (m, 6H), 4.04-4.10 (m, 1H), 4.18-4.30 (m, 3H), 7.33 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.63 (d, 1H, *J* = 8.0 Hz), 7.72 (d, 2H, *J* = 7.4 Hz), 7.89 (d, 2H, *J* = 7.4 Hz), 12.85 (s, 1H).

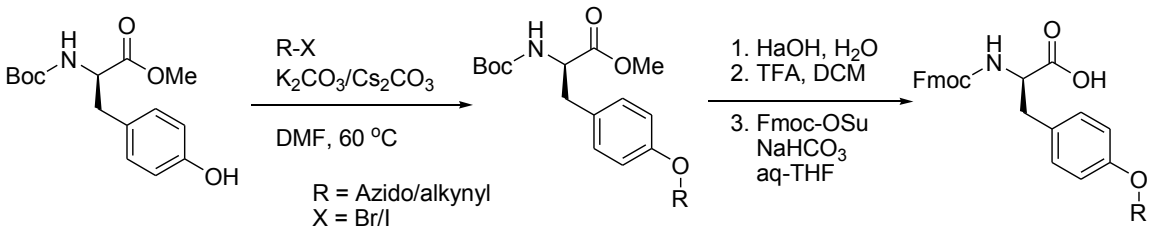
^{13}C-NMR (100MHz, DMSO-d_6): δ 31.4, 47.1, 50.4, 51.2, 66.0, 67.1, 69.5, 120.6, 125.7, 127.5, 128.1, 141.2, 144.3, 156.6, 174.4.

29 (n = 1):

^{1}H-NMR (400MHz, $CDCl_3$): δ 2.15-2.20 (m, 2H), 3.36-3.40 (m, 2H), 3.60-3.70 (m, 8H), 4.24 (t, 2H, *J* = 6.4 Hz), 4.40-4.44 (m, 2H), 4.45-4.52 (m, 2H), 6.04 (bd, 1H, *J* = 7.2 Hz), 7.32 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.61 (t, 2H, *J* = 7.4 Hz), 7.77 (d, 2H, *J* = 7.4 Hz).

^{13}C-NMR (100MHz, $CDCl_3$): δ 31.0, 31.4, 47.2, 50.7, 52.5, 67.0, 67.9, 69.8, 70.3, 120.0, 125.1, 127.1, 127.7, 141.3, 143.8, 156.2, 175.0.

Figure 2.2: Modification of D-phenyl alanine with an azido/acetylene group to synthesize amino acids 30 -33.



A bromide or iodide was added to Boc-protected D-phenyl alanine methyl ester in anhydrous DMF. Anhydrous K_2CO_3 (3 eq.) and Cs_2CO_3 (0.2 eq.) were added to the solution and heated at 60^{o}C overnight. Water was added to the mixture and the product was extracted with ethyl acetate. The ethyl acetate solution was washed with water three times. A crude mixture was obtained by removing the solvent under reduced pressure. The mixture was purified by column chromatography using hexane: ethyl acetate (1:5). After the product was dissolved in H_2O: MeOH: THF (1:1:3), NaOH (2 eq.) in water was added to the solution. After concentrating the mixture under reduced pressure, 2 M HCl solution was added to adjust the acidity to pH 3. The product was extracted with ethyl acetate two times. After removing the solvent, 50% TFA in DCM was added to the product and stirred overnight. After evaporating, the product was

dissolved in 40% aqueous THF. $NaHCO_3$ (10 eq.) and Fmoc-OSu (1.1 eq.) were added sequentially and stirred overnight. After removal of THF, diethyl ether was added to extract unreacted Fmoc-Osu. After discarding the organic layers, the pH of the aqueous solution was adjusted to 2.0 by adding 2M HCl solution. The product was extracted with ethyl acetate and then purified through column chromatography using DCM: MeOH (50:1).

30: ^{1}H-NMR (400MHz, $CDCl_3$): δ 2.03 (q, 2H, *J* = 6.4 Hz), 3.07 (dd, 1H, *J* = 14.4, 6.4 Hz), 3.16 (dd, 1H, *J* = 14.4, 5.2 Hz), 3.50 (t, 2H, *J* = 6.4 Hz), 3.98 (t, 2H, *J* = 6.0 Hz), 4.21 (t, 1H, *J* = 7.2 Hz), 3.36 (dd, 1H, *J* = 10.8, 6.8 Hz), 3.16 (dd, 1H, *J* = 10.8, 6.8 Hz), 4.65-4.70 (m, 1H), 5.23 (bd, 1H, *J* = 8.0 Hz), 6.82 (d, 2H, *J* = 8.4 Hz) , 7.17 (d, 2H, *J* = 8.4 Hz), 7.31 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.55 (t, 2H, *J* = 7.4 Hz), 7.78 (d, 2H, *J* = 7.4 Hz).

^{13}C-NMR (100MHz, $CDCl_3$): δ 28.8, 36.9, 47.1, 48.2, 54.7, 64.4, 67.1, 114.6, 120.0, 125.1, 127.1, 127.7, 127.8, 130.4, 141.3, 143.6, 143.7, 155.8, 157.9.

31: ^{1}H-NMR (400MHz, $CDCl_3$): δ 1.40-1.52 (m, 4H), 1.60-1.68 (m, 2H), 1.76-1.82 (m, 2H), 3.04-3.20 (m, 2H), 3.28 (t, 2H, *J* = 6.8 Hz), 3.92 (t, 2H, *J* = 6.4 Hz), 4.21 (t, 1H, *J* = 7.2 Hz), 4.34-4.40 (m, 1H), 4.43-4.49 (m, 1H), 4.65-4.71 (m, 1H), 5.21 (bd, 1H, *J* = 8.0 Hz), 6.82 (d, 2H, *J* = 8.4 Hz) , 7.17 (d, 2H, *J* = 8.4 Hz), 7.31 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.55 (t, 2H, *J* = 7.4 Hz), 7.78 (d, 2H, *J* = 7.4 Hz).

^{13}C-NMR (100MHz, $CDCl_3$): δ 23.0, 23.7, 25.7, 26.5, 28.9, 30.3, 47.1, 51.4, 67.1, 67.7, 114.6, 120.0, 125.1, 127.1, 127.8, 130.4, 141.3, 143.7, 158.3.

32:^{1}H-NMR (400MHz, $CDCl_3$): δ 3.04-3.18 (m, 2H), 3.41 (t, 2H, *J* = 4.8 Hz), 3.74 (t, 2H, *J* = 4.8 Hz), 3.85 (t, 2H, *J* = 4.8 Hz), 4.10 (t, 2H, *J* = 4.8 Hz), 4.37 (dd, 1H, *J* = 10.8, 6.8 Hz), 4.46 (dd, 1H, *J* = 10.8, 7.2 Hz), 4.62-4.70 (m, 1H), 5.27 (bd, 1H, *J* = 8.0 Hz), 6.85 (d, 2H, *J* = 8.8 Hz) , 7.06 (d, 2H, *J* = 8.8 Hz), 7.31 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.55 (t, 2H, *J* = 7.4 Hz), 7.79 (d, 2H, *J* = 7.4 Hz).

^{13}C-NMR (100MHz, $CDCl_3$): δ 36.9, 47.1, 50.7, 54.7, 67.1, 67.3, 69.8, 70.2, 114.8, 120.0, 125.1, 127.1, 127.8, 130.4, 141.3, 143.7, 155.8, 157.9, 175.8

33: ^{1}H-NMR (400MHz, $CDCl_3$): δ 3.03-3.18 (m, 2H), 3.38 (t, 2H, *J* = 4.8 Hz), 3.65-3.70 (m, 4H), 3.71-3.76 (m, 2H), 3.83-3.88 (m, 2H), 4.07-4.11 (m, 2H), 4.21 (t, 1H, *J* = 6.8 Hz), 4.36 (dd, 1H, *J* = 10.8, 6.8 Hz), 4.45 (dd, 1H, *J* = 10.8, 7.2 Hz), 5.33 (bd, 1H, *J* = 8.0 Hz), 6.84 (d, 2H, *J* = 8.8 Hz) , 7.05 (d, 2H, *J* = 8.8 Hz), 7.31 (t, 2H, *J* = 7.4 Hz), 7.41 (t, 2H, *J* = 7.4 Hz), 7.55 (t, 2H, *J* = 7.4 Hz), 7.79 (d, 2H, *J* = 7.4 Hz).

^{13}C-NMR (100MHz, $CDCl_3$): δ 36.9, 47.1, 50.6, 54.7, 67.0, 67.3, 69.8, 70.1, 70.7, 70.8, 114.7, 120.0, 125.1, 127.1, 127.8, 130.4, 141.3, 143.8, 155.8, 157.9, 175.4, 176.8.

2.2.2.2 Procedure for peptide synthesis

Peptides were synthesized using standard Solid Phase Fmoc chemistry (SPPS)[20] either manually or automatically. Peptides were synthesized on Rink Amide AM resins (Anaspec, loading = 0.31 mmol/g), Biotin Novatag Resin (EMD Biosciences, loading 0.5 from mmol/g), and Rink Sieber Amide resin (Anaspec, loading = 0.4 mmol/g) on a 200 mg scale. Peptides were manually synthesized in fritted polypropylene tubes using a 180° shaker (St. John Associates; Beltsville, MD). A 24-port SPE vacuum manifold system (Grace; Deerfield, IL) was used for draining the coupling and deprotection solutions. Automated synthesis was done using the Titan 357 (AAPPTec) automated peptide synthesizer. Before the peptide synthesis, the resin was swelled in NMP for 2 hours. For coupling amino acids to the resin, it was treated for 30 minutes with 2 equivalents of an Fmoc protected amino acid, (0.2 M in NMP), 2 equivalents of TBTU (0.2 M solution in NMP), and 5 equivalents of DIEA (0.2 M solution in NMP). For removal of the Fmoc group from the N terminal, the resin was treated with 20% pyridine in NMP for 15 minutes twice. In between the coupling and deprotection steps, the resin was washed five times for 1 minute with NMP. For N- terminal acylation, the peptide containing resin was swelled in DMF and treated with 5:6:10 solution of acetic acid, 2,6 lutidine and DMF for 10 minutes twice.

Following the treatment the excess reagent was removed by multiple washes with DMF. For peptide cleavage off the resin, the resin was treated with a trifluoroacetic acid cleavage cocktail (TFA/water/triisopropylsilane (TIS) (95/2.5/2.5, v/v/v)) mixture at room temperature for 2 hours. The cleavage solution was concentrated and then added drop-wise into ice-cold diethyl ether to precipitate the peptide. The resulting crude peptides were lyophilized. All the peptides were purified using a gradient of double distilled water and HPLC grade acetonitrile and 0.1% TFA on the RP-HPLC (Beckman Coulter System Gold 126 Solvent Module and 168 Detector) using a C18 reversed phase semi-preparative column (Phenomenex Luna 10 µm, 250 × 10 mm).

Protocol for on bead copper catalyzed azide alkyne cycloaddition (CuAAC) click reaction:

On bead Cu catalyzed click reactions were performed with the azide on bead and the alkyne in solution. The resin was treated with 2 equivalents of the relevant alkyne, 1.5 equivalents of CuI (Sigma) and 2.5 equivalents of ascorbic acid (Sigma), in a solution of 20% piperidine in DMF. The reaction was performed overnight at room temperature. The excess copper was removed from the resin by washing extensively with a Cu chelating solution (5% (w/v) sodium diethyl dithiocarbamate, 5% (v/v) DIEA in DMF).

Synthesis of biotin tagged peptides:

Biotin was incorporated at the C terminal for all the peptides, with a polyethylene glycol between the biotin tag and the rest of the peptide sequence. The peptides were synthesized either by directly using biotin novatag resin (which incorporates a biotin on the C terminal of the peptide) or by appending a biotin on the side chain of a lysine on Rink Amide MBHA resin. For the later method, Fmoc-lysine(Mtt)-OH (Mtt = 4-methyltrityl) was coupled to Rink Amide resin using standard SPPS chemistry. The Mtt group was selectively removed by three consecutive washes with trifluoroacetic acid (3%), triisopropylsilane (3%) and DCM (94%). 1.5 equivalents of Biotin-NHS and 5 equivalents of DIEA were added to the resin Fmoc-$(PEG)_2$-CO_2H (20

atoms, EMD Biosciences) was coupled to the resin to introduce a polyethylene glycol (PEG) spacer. The synthesized peptides were purified to > 98% purity.

2.2.2.3 Measurement of the Relative Hydrophobicity Coefficients of amino acids

To compare the relative hydrophobicities of the synthesized artificial amino acids with natural amino acids, a literature method was adapted[21]. Briefly, a series of Ac-X-GLF-OH peptides were synthesized (X is the natural or unnatural amino acid) and their retention times were measured using an autosampler equipped Waters Acquity UPLC system (Waters Acquity UPLC® BEH C18 column, flow-rate of 0.3 ml/minutes, temperature 25°C, eluent of 10% acetonitrile and 0.1% TFA in water). The hydrophobicity coefficients of the amino acid X in the synthesized peptides were calculated as the retention time of the X-substituted peptide minus the retention time of Ac-GGLF-OH peptide. The relative hydrophobicity coefficients are presented in Figure 2.8.

2.2.2.4 Construction of Peptide Libraries

Randomized OBOC libraries of penta or heptapeptides were synthesized using an automatic synthesizer Titan 357 (AAPPTec) via standard split-and-mix methods, on TentaGel S-NH2 beads. Unless otherwise specified, non-natural D-stereoisomers were used at every possible position in the peptide sequence. For all the coupling steps, standard SPPS method with Fmoc chemistry was used. The resin was swelled in NMP for 2 hours in a big Collective Vessel (CV) and divided into multiple equal-mass aliquots for the cycles of coupling and deprotection with different amino acids in the smaller Reaction Vessels (RV). For libraries used at IBN, Fmoc-methionine was coupled to the entire resin in the first coupling step to make the library MALDI-TOF compatible. The amino acid side-chain protective groups were then removed by treatment with TFA cleavage cocktail mixture for 2 hours. The on bead peptide library was washed with DCM, methanol and water, and stored. The peptide libraries used for various screens are summarized in Table 2.1.

Table 2.1: One Bead One Compound (OBOC) peptide libraries used in screens.

Library	Sequence	Amino acids used	Number of unique sequences
A	$X_1X_2X_3X_4X_5X_6$-TG	X_i= D-Trp, D-Lys, D-Ile, Gly, D-Leu, D-Val, D/L- azidolysine	117649
B	$X_1X_2X_3$-D/L-Az4-X_5X_6-D-Met-TG	X_i = D-Trp, D-Lys, D-Leu, D-Val, D-Ile, Gly	15552
C	$X_1X_2kX_4$-D/L-Az4-X_6w-D-Met-TG	X_1= all 18 D amino acids except D Met and D-Cys; X_2= D-Lys(Alloc)-OH, D-Trp, D-Val; X_4= D-Arg, D-Asn, D-Gln, D-Asp, D-Lys, D-Ser, D-Thr, D-His, D-Ala, Gly, D-Val, D-Trp; X_6 = D-Arg, D-Asn, D-Gln, D-Asp, D-Lys, D-Ser, D-Thr, D-His, D-Ala, Gly, D-Leu, D-Phe	15552
D	Acetyl-D-Pra-$X_1X_2X_3X_4X_5X_6$-L-Met-TG	D-Pra = D-Propargylglycine,; X_i = D-Ala, Gly, D-Leu, D-Ile, D-Val, D-Phe, D-Trp, D-Arg, D-His, D-Lys, D-Asp, D-Glu, D-Asn, D-Gln, D-Ser, D-Thr, D-Tyr, D-Pro	34012224
E	$X_1X_2X_3X_4X_5X_6$-L-Met-TG	X_1= D-Lys, D-Arg, D-Val, D-Leu, D-Phe, D-Ile, D-Tyr, D-Trp; X_2 = D-Ile, D-Arg, D-Phe, D-Val, D-Leu, D-Lys, D-Trp, D-Tyr; X_3= D-Phe, D-Arg, D-Ile, D-Val, D-Leu, D-Lys, D-Thr; X_4 = D-Tyr, D-Phe, D-Arg, D-Trp, D-Val, D-Lys, D-Ile, D-Leu X_5 = D-Arg, D-Ile, D-Lys, D-Phe, D-Val, D-Tyr, D - His, D-Asn; X_6 = D-Arg, D-Lys, D-Phe, D-Tyr, D-Val, D-Trp, D-His, D-Ile	262144
F	D/L-Az4-$X_1X_2X_3X_4X_5$-TG	Xi = all 18 D amino acids except D-Met and D-Cys	3,779,136

2.2.2.5 Screening procedures

For screens with the florescent labeled protein, bCAII was labeled with Alexa Fluor 647 (Microscale Protein Labeling Kit, Invitrogen) following the manufacturer's protocol. The different screens for biligand and triligands are summarized in Table 2.2. Details of each type of screen i.e., target, or product, and each type of sorting, i.e., automatic or manual is provide below.

Table 2.2: Summary of screens for the development of branched PCC agents for bovine carbonic anhydrase II.

Screen	Library	bCAII concentration	Peptide anchor	Peptide concentration	Type of Screen	Type of sorting	Ligand/ motif Isolated
BB1	A	10 nM	**37**	20 μM	target	manual	xxx-Az4-xx
BB2	B	10 nM, 1 nM and 500 pM	**37**	20 μM, 2 μM, 1 μM	target	manual	x-k-x-Az4-x-w
BB3	C	10 nM	**37**	40 μM	target	manual	***38, 39***
TRI1	D	10 nM	**42**	20 μM	target	automated	N/A
TRI2	E	10 nM	**42**	20 μM	target	automated	***45, 46***
TRI3	F	10 nM	**40**	20 μM	product	manual	***48***, ***49, 50, 51***

In situ click target screen:

Nanomolar solutions of Alexa Fluor 647-labeled bCAII (PBS, pH 7.4, 0.1% Tween-20, 0.1% bovine serum albumin (BSA), 0.05% NaN_3) was pre-incubated for 2 hours with the micromolar solutions of the corresponding anchor peptide (table 2.2) in binding buffer for 2 hours at 25^{o}C in a polypropelene vessel. 200 mg of the relevant OBOC peptide library was swelled in the binding buffer in a separate vessel for 1 hour at 25^{o}C. The pre-incubated solution of Alexa Fluor® 647-labeled bCAII and the biligand peptide anchor 40 were added to the equilibrated on bead peptide library. The library was incubated for ~16 hours at 25^{o}C. Non-specifically bound

proteins were eliminated by multiple washes, first with the binding buffer and then with 0.1%Tween 20 in PBS. The details of each individual screen are tabulated in table 2.2.

Automated sorting of fluorescent protein:

The assayed library was transferred into the sample vessel of a COPAS Plus (Union Biometrica) automated bead sorter, and diluted with 200 ml of PBS buffer (pH=7.4). A two-step sorting was applied. The first sorting was performed at high flow rate (> 80 beads per seconds) to sort out 1~2% of beads according to fluorescence intensity. The sorted bright beads were washed with DI water several times, transferred into the sample vessel, and diluted with 200 ml of DI water. The second sorting was carried out at low flow rate (5 to 10 beads per second). In the second sorting, positive beads were directly sorted into a 96 titer well plate with conical-shaped wells. Sequencing of these beads was performed with MALDI-TOF/TOF and the semi-automated algorithm.

Manual sorting of fluorescent protein:

Following the screen and washes, the library was transferred to glass microscope slides and imaged using the GenePix 4200 array scanner (λ_{ex} = 635 nm). The beads with the highest fluorescence intensities were considered hit beads and picked up manually using a glass micropipette. The picked beads were washed with 7.5 M guanidium hydrochloride (pH= 2.0) for 1 hour followed by ten washes with double distilled water. Each bead was loaded on a cartridge on the Edman Sequencer (494 Procise cLC Sequencer, Applied Biosciences).

In situ click product screen:

Bovine carbonic anhydrase II was incubated with ~2000x concentrated biotinylated anchor peptide overnight at 25 ^{o}C. The OBOC peptide library is swelled in the binding buffer overnight and then added to the preincubated protein solution. The screen was done for 7.5 hours at room temperature, followed by washes with the binding buffer, wash buffer 1 (PBS + 0.05% NaN3 + 0.1% Tween 20) and PBS. The beads were treated with 7.5 M guanidium hydrochloride

(pH = 2) for an hour and washed ten times with Millipore water. The beads were swelled back in binding buffer and incubated for an hour with a 1:10,000 diluted solution of anti-biotin mAb-alkaline phosphatase (Sigma Aldrich) in the binding buffer. The beads were washed multiple times with the wash buffer and PBS. The beads in AP buffer (100mM Tris-HCl [pH 9.0], 150mM NaCl, 1mM MgCl2) were transferred to plastic petri dishes. 10 ml of BCIP solution (33 μl BCIP/ 10 ml AP buffer) was added to each petri dish. After 30 minutes, beads that have anti-biotin mAb-alkaline phosphatase binding to them turn turquoise due to the reaction of BCIP with the alkaline phosphatase. These beads were picked up manually using a micropipette and decolorized by washing with dimethyl formamide. Following treatment with 7.5 M guanidium hydrochloride (pH= 2.0) for 1 hour and stringent water washes, the beads were equilibrated in binding buffer. The library was then treated with biotin saturated anti-biotin –alkaline peroxide antibody. After an hour, the beads were washed as in the previous screen and treated with BCIP solution. Due to ligand competition, the beads that bind to the biotin-binding site of the antibody were colorless in this screen while the nonspecific binders to the antibody turned turquoise. The colorless beads were picked up manually, washed with 7.5 M guanidium hydrochloride (pH= 2.0) for 1 hour followed by ten washes with Millipore water. Each bead was separately loaded on an Edman Sequencer cartridge and sequenced on the Procise sequencer.

Table 2.3: Hit sequences from the first biligand screen BB1.

The library A of form $X_1X_2X_3X_4X_5X_6$-TG (where X_i= D-Trp, D-Lys, D-Ile, Gly, D-Leu, D-Val, D/L- azidolysine TG = tentagel S NH2 resin, diversity 117649) is screened with anchor peptide **37** with the Alexa-647 labeled bCAII protein.

X_1	X_5	X_3	X_4	X_5	X_6
I	f	k	Az4	v	w
i	k	v	Az4	i	w
f	k	w	Az4	i	w
v	k	v	Az4	i	w
w	k	v	Az4	i	w
w	k	l	Az4	i	w
f	k	f	Az4	f	f
Az4	k	f	Az4	i	w
Az4	k	w	G	G	l

Table 2.4: Hit sequences from the second generation biligand screen BB2.

The library B of form $X_1X_2X_3$-D/L-Az4-X_5X_6-D-Met-TG (where X_i = D-Trp, D-Lys, D-Leu, D-Val, D-Ile, Gly, Met=L-Methionine, TG = tentagel S NH2 resin, 15552 unique sequences) is screened with anchor peptide **37** with the Alexa-647 labeled bCAII protein in the biligand screen.

Protein concentration	X_1	X_2	X_3	Az4	X_5	X_6
10 nM	w	k	v	Az4	l	w
10 nM	v	k	v	Az4	i	w
10 nM	w	k	l	Az4	i	w
10 nM	i	k	w	Az4	f	w
10 nM	w	k	f	Az4	i	w
10 nM	v	k	w	Az4	i	w
10 nM	w	k	i	Az4	v	w
10 nM	v	k	w	Az4	i	w
10 nM	i	k	f	Az4	i	w
10 nM	f	k	f	Az4	f	w
10 nM	v	k	w	Az4	f	w
1 nM	v	k	w	Az4	f	w
500 pM	w	k	v	Az4	l	w

Table 2.5: Hit sequences from the third generation biligand screen BB3.

The library C of form $X_1X_2kX_4$-Az4X_6w-D-Met-TG (where X_1= all 18 D amino acids except D Met and D-Cys; X_2= D-Lys(Alloc)-OH, D-Trp, D-Val; X_4= D-Arg, D-Asn, D-Gln, D-Asp, D-Lys, D-Ser, D-Thr, D-His, D-Ala, Gly, D-Val, D-Trp; X_6 = D-Arg, D-Asn, D-Gln, D-Asp, D-Lys, D-Ser, D-Thr, D-His, D-Ala, Gly, D-Leu, D-Phe) is screened with anchor peptide **37** and Alexa-647 labeled bCAII protein .

X_1	**X_2**	**k**	**X_4**	**Az4**	**X_6**	**w**
k	w	k	w	Az4	l	w
r	v	k	w	Az4	i	w
v	w	k	v	Az4	i	w
k	w	k	v	Az4	l	w
a	v	k	v	Az4	l	w
i	v	k	w	Az4	l	w
i	w	k	v	Az4	f	w

Table 2.6: Hit sequences from the first generation triligand screen TRI1

Library D of form Acetyl-D-Pra-$X_1X_2X_3X_4X_5X_6$-L-Met-TG (where D-Pra = D-Propargylglycine,; X_i = D-Ala, Gly, D-Leu, D-Ile, D-Val, D-Phe, D-Trp, D-Arg, D-His, D-Lys, D-Asp, D-Glu, D-Asn, D-Gln, D-Ser, D-Thr, D-Tyr, D-Pro) is screened against bCAII- biligand **42** complex.

1	vgvwfk	38	nyfkrr	77	hrfyrr
2	viifkr	39	fttlvk	78	rGfykr
3	wklwfr	40	lrfyli	79	yryrlr
4	hiifnk	41	aiflrr	80	klkyrr
5	alfiir	42	kyrylr	81	srryry
5	hdtfvr	43	rfrrfa	82	rlfyth
6	afffrl	44	ldnvvk	83	Gwyvrr
7	hiryrk	45	ylfflr	84	krwwnr
7	vwtygh	46	lvallk	85	frkrhr
8	iwvryd	47	lrlllr	86	tiiyrh
9	vflrnv	48	kwrirl	87	kvyfnr
10	nvarhp	49	ffrtlv	88	tllykk
11	lrlyfr	50	iviwkr	89	rrfrlr
12	rfrlwl	51	kwrhwr	90	yyrfrk
13	vkiyvr	52	iaiwlr	91	khyrrw
14	kiyvrf	53	llvvyr	92	kiifsr
15	vvtsvr	54	nwkryr	93	rvvwfr
16	vrlylk	55	llhvfr	94	vklflk
17	rilfhw	56	rfwklk	95	frvwfg
18	kwyfyr	57	sryrhr	96	ryffkw

19	yifflr	58	kfhrrk	97	ftvllr
20	lraylr	59	yafflr	98	riyvrv
21	wrrfrr	60	kfyyrv	99	lrkwlw
22	ytGlfk	61	kiryfr	100	rfvkvf
23	pypyl	62	nwkwrk	101	wpherd
24	nrGnhr	63	khwrrr	102	pfdlw
25	Iiiyrs	64	rkawlr	103	fyyrk
26	yylvkr	65	fitrkf	104	pwfwG
27	llhltk	66	kwvver	105	rfvkvf
28	wpvpvf	67	ywlvkr	106	fkrkir
29	frvysf	68	lffrwv	107	wriyir
30	nfyyri	69	fafyvr	108	lfirly
31	rwklrr	70	wirirk	109	vfvkkl
32	kwtrei	71	hifirk	110	llrlay
33	wirGfy	72	rifvfr	111	rlrfhk
34	vyrkyk	73	llfyrk	112	prfyky
35	iyifrk	74	hyrkkw	113	rvkwkk
36	yrwrkf	75	ywflkk		
37	rkywkr	76	fkkyyr		

Table 2.7: Hit sequences from the second generation triligand screen TRI2

Library E of form $X_1X_2X_3X_4X_5X_6$-L-Met-TG (X_1= D-Lys, D-Arg, D-Val, D-Leu, D-Phe, D-Ile, D-Tyr, D-Trp; X_2 = D-Ile, D-Arg, D-Phe, D-Val, D-Leu, D-Lys, D-Trp, D-Tyr; X_3= D-Phe, D-Arg, D-Ile, D-Val, D-Leu, D-Lys, D-Thr; X_4 = D-Tyr, D-Phe, D-Arg, D-Trp, D-Val, D-Lys, D-Ile, D- Leu X_5 = D-Arg, D-Ile, D-Lys, D-Phe, D-Val, D-Tyr, D - His, D-Asn; X_6 = D-Arg, D-Lys, D-Phe, D-Tyr, D-Val, D-Trp, D-His, D-Ile) is screened against Alexa-647 labeled bCAII in presence of biligand **42**.

1	**yrlfvr**	18	**lilfnr**	35	**yklffr**
2	kyfakf	19	**lklwfk**	36	**vklwfk**
3	**vrlffk**	20	**kyyfrf**	37	**kyffrf**
4	**vklwlk**	21	**lklwlk**	38	**kyffrf**
5	rfwfkw	22	**vklwfk**	39	yvtfvk
6	yirnrr	23	**ykifvr**	40	**kyffrf**
7	**kyffrf**	24	**kyffrf**	41	yrrfri
8	**kyffkf**	25	kfffrv	42	**vkifvr**
9	iitfnr	26	ifyrh	43	**lilfnr**
10	**vrlffk**	27	**vklfvr**	44	rfffrv
11	**rifvri**	28	kiyffr	45	ryflrf
12	kiyffr	29	kfffr	46	**irlflk**
13	kfffri	30	**iklwlr**	47	**vivfnr**
14	**ivlfnr**	31	**irlwlk**	48	**ivkfvr**
15	**lrlwlk**	32	**fkifvr**	49	kwyffr
16	**kffyrf**	33	**lrfflk**	50	**kyffrf**
17	rkrfrh	34	lrkfri		

Table 2.8: Hit sequences from the first generation screen TRI3 using the product screening method.

The library F of form D/L-Az4- $X_1X_2X_3X_4X_5$-TG (Xi = all 18 D amino acids except D-Met and D-Cys) is screened with biligand anchor peptide **40** with the Alexa-647 labeled bCAII protein in the triligand screen.

Az4	X_1	X_2	X_3	X_4	X_5
Az4	h	d	t	f	y
Az4	h	d	t	G	f
Az4	h	d	e	G	G
Az4	y	s	q	w	a

2.2.2.6 Dot blots

The affinity of the capture agents for bCAII was demonstrated through dot blot experiments in 5% nonfat dry milk in TBST [25 mM Tris, 150 mM NaCl, 2 mM KCl, 0.1% Tween 20 (pH 7.0)]. A bCAII stock solution (10 mg/mL) was prepared in PBS buffer (pH 7.4). A dilution series of bCAII solution was applied to a nitrocellulose membrane, typically ranging from 1 μg to 20 ng per spot. The membrane was blocked at room temperature for 1 hour in 5% nonfat milk/TBST, which was then washed three times with TBST for 10 minutes each. A biotinylated peptide ligand solution was prepared at 100 nM or 1 μM in 0.5% nonfat milk/TBST and incubated over the membrane for 1 hour at room temperature. After washing three times with TBST for 10 minutes, the membrane was incubated with 1:3000 streptavidin-HRP (Abcam) in 0.5% milk/ TBST for 1 hour followed by subsequent washing with TBST for 10 minutes. The

resulting membrane was detected using the chemiluminescent reagents (Amersham ECL plus Western blotting detection reagents, GE Healthcare).

2.2.2.7 ELISA with PCC agent as capture agent

Streptavidin coated 96 well plates were washed three times with 0.1% BSA/TBST. 100 μl of 200 nM solution of the biotinylated capture agent (made by dilution in 0.1% BSA/TBST of 1 mM DMSO stock) was added to each well and incubated for 2 hours with shaking at room temperature. The wells were then washed three times with 200 μl of 0.1% BSA/TBST. 100 μl of His_6 tagged recombinant carbonic anhydrase II protein solution, (from 10 μM to 1 nM in 0.1% BSA/TBST) was added to the wells. After overnight incubation of the plate at 4°C, each well was washed three times with 0.1% BSA/TBST. 100 μl of 1:1000 diluted anti-His_6 mouse monoclonal antibody (Abcam) was added to each well and incubated with shaking for an hour at 4°C. Following three washes with 0.1% BSA/TBST, 100 μl of a 1:10,000 diluted solution of anti – mouse IgG goat antibody conjugated with horseradish peroxidase enzyme (Abcam) was added and incubated for one hour at 4°C. After one hour the wells were washed five times with TBST and then 100 μl TMB (3,3',5, 5'-tetramethylbenzidine) substrate (KPL) was added per well. Upon color development, each well was quenched with 100 μl of 1 M sulfuric acid in the order the TMB substrate was added. The absorbance at wavelength 450 nm was measured using a Spectramax ELISA plate reader.

2.3 Results and discussion

The PCC agent can be visualized as a multivalent binding agent that links several non-competitive weak affinity peptide ligands to allow cooperative binding. We have developed in the past a linear triligand capture agent **44** against bovine carbonic anhydrase II (bCAII). Starting with the development of a weak peptide binder (with high micromolar dissociation constant) **36**, a specific and selective multiligand peptide binder with high nanomolar affinity had been

developed. The peptide ligands had been selected through screens such that they could be linked linearly from N terminal to C terminal. In the current work, we explore the different connectivities of the peptide ligands to create branched peptide multiligands for bCAII starting with the same monomeric peptide binder. Two different methods of developing a branched bivalent PCC agent are explored. In the first method, the reported peptide **36** is retained as the monoligand and a new 2° peptide arm is discovered thorough screening. In the second method, the earlier reported linear biligand **41** is modified in the 2° ligand arm. The newly developed branched biligands are used to screen for branched trivalent PCC agents. It is explored if the variation in connectivity of the peptide arms in the PCC agent has any positive effect on the affinity of the ligand for the bCAII protein.

2.3.1 Development of branched biligand using randomized azide library

In the previous work, the primary peptide binder **36** to the protein that had been obtained via two-generation screens, first against a comprehensive random library, and then against a smaller focused library. This ligand **36** had been modified with D-propargylglycine at the C terminal to form the anchor ligand, **37,** and used as the anchor peptide in the linear biligand screens. To develop a branched biligand containing the same primary ligand, we utilize the highest occurring amino acids from the linear library screens to develop a small focused library. To allow the protein to scan the entire conformational space the azidolysine is used as an amino acid component like all others. This results in the synthesis of an OBOC heptapeptidic library with a random number and random position of the azidolysine (library A). The library was screened against 10 nM Alexa 647-labelled bCAII pretreated 20 μM peptide **37.** After buffer washes, the beads with the highest fluorescent intensities were manually picked up. Out of the ten sequences containing azidolysine, nine contained one azidolysine, while one sequence had two azidolysines Amino acid D/L-azidolysine was conserved at X_4 (Table 2.3). This suggested that the protein actually prefers the in situ click occurring not end to end but from one end of the

anchor to the middle of the second peptide arm. X_2 and X_6 positions of the hit sequences were also found to be highly conserved. A second focused library was synthesized, where the position of the D/L-azidolysine was fixed at X_4. The library was screened against three different concentrations (10 nM, 1 nM and 500 pM) of Alexa 647-labeled bCAII in presence of peptide **37**. The beads with highest fluorescence intensity were picked. This yielded eleven beads from the 10 nM screen, and one bead each from the 1 nM and 500 pM screens (Table 2.4). The trends noticed in the first screen were even more noticeable in this screen, yielding two motifs, wkX-Az4-Xw and vkX-Az4-Xw. Biligands **38** and **39** were synthesized in which vkw-Az4-lw and wkv-Az4-lw were clicked to the peptide **37**.

To find the preference of the protein for further branching, the protein was screened with a third heptameric focused library. The aim of this library was to increase the length of the second peptide arm while simultaneously looking at the selectivity of the protein for branching at X_5 position by incorporation of D-Lys(Alloc)-OH. In situ click screens of two copies of the OBOC library with 10 nM bCAII pretreated with 40 μM anchor, after buffer washes, was imaged and the highest fluorescent intensity beads manually picked. The sequences of the beads are presented in Table 2.5. The amino acid at X_5 was always D-Trp or D-Val, indicating a strong preference of these two amino acids at this position. A hit sequence k(alloc)-wkv(Az4)lw obtained twice within the hit sequences was chosen as the candidate to further extend the second peptide arm. The alkyne functionality was incorporated on the side chain of the N terminal D-Lys by coupling Fmoc-D-Pra to the lysine side chain.

Figure 2.3: Molecular structures of the reported monoligand 36 and the monoligand anchor 37 used in the branched biligand screen.

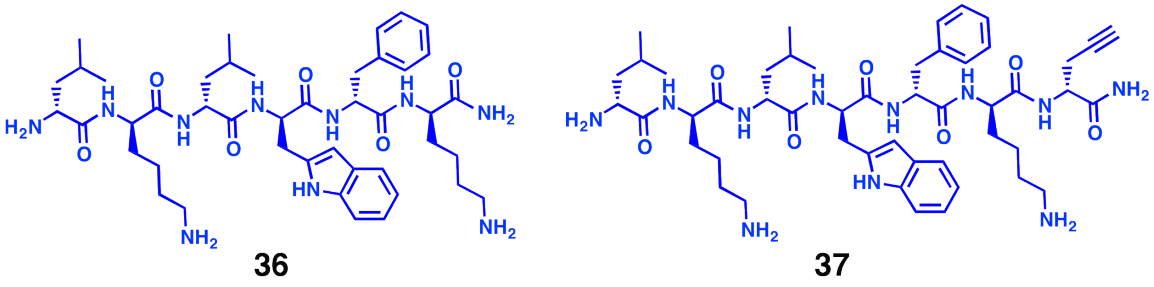


Figure 2.4: Schematic representation of target screen against bovine carbonic anhydrase II using OBOC library with multiple azides for developing a branched biligand.

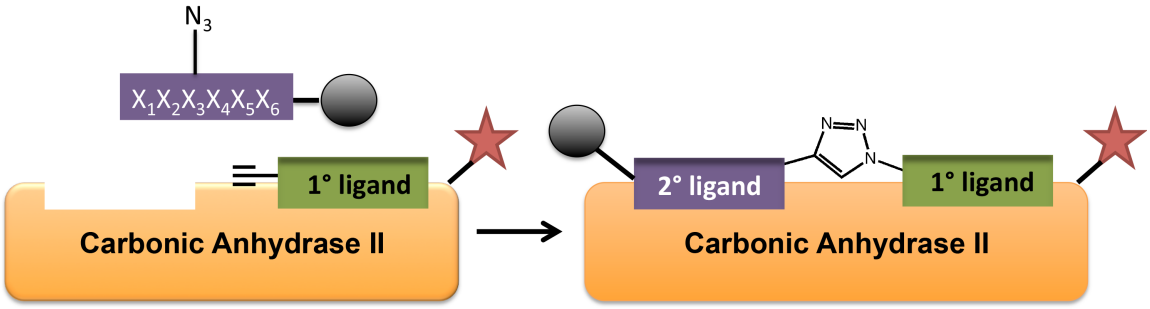


Figure 2.5: Motifs obtained by screening against bovine carbonic anhydrase II in presence of monoligand anchor peptide 37.

Screening against a library with variable number and position of azides, it is observed that the protein favors one azidolysine at X_4. Screens against focused library B yield motifs wkXAz4Xw and vkXAz4Xw. Screen against a further focused library C containing these two motifs yields the final biligand anchor arm k(alloc)-wkv-Az4-lw.

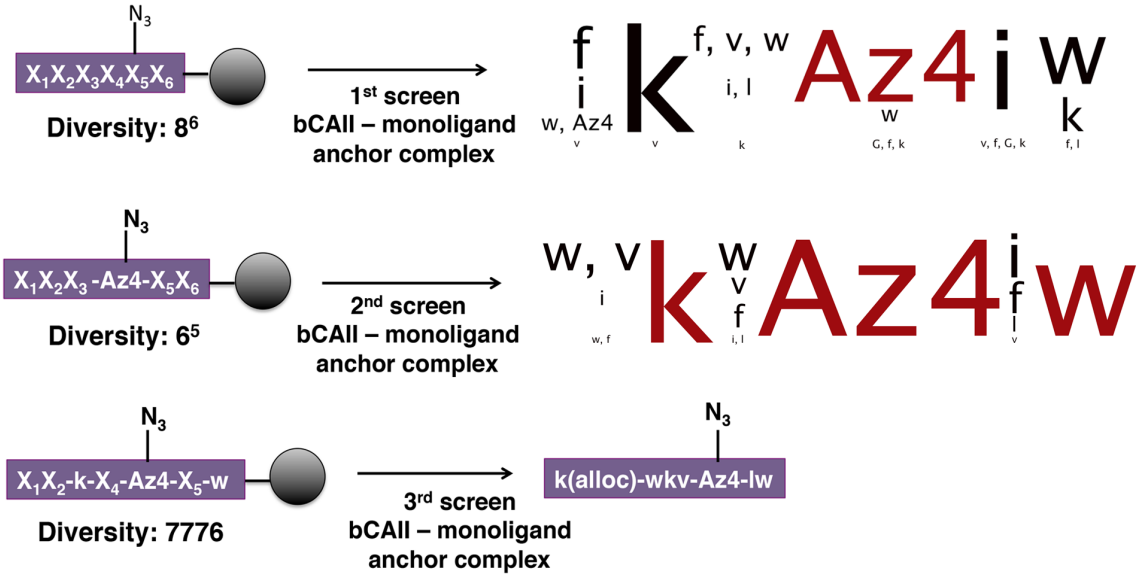
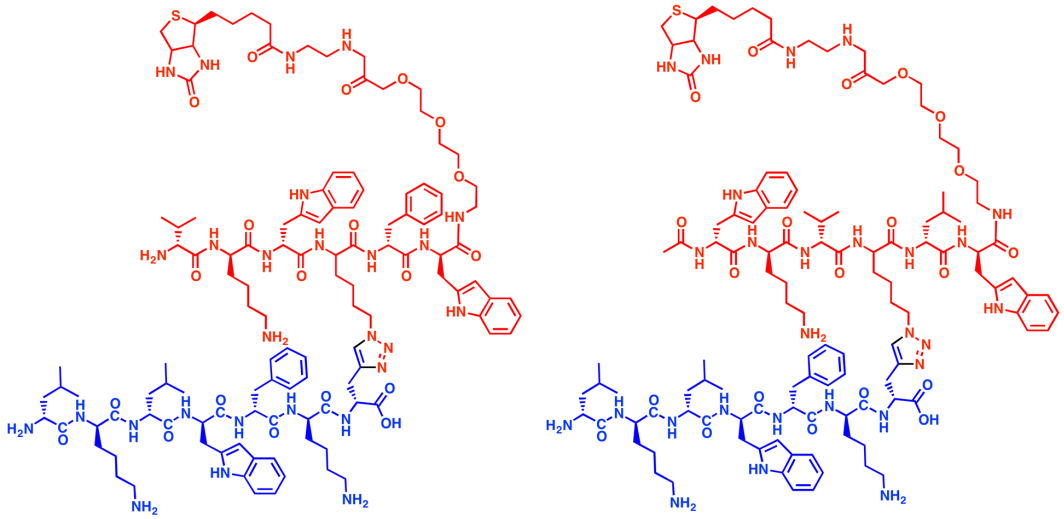
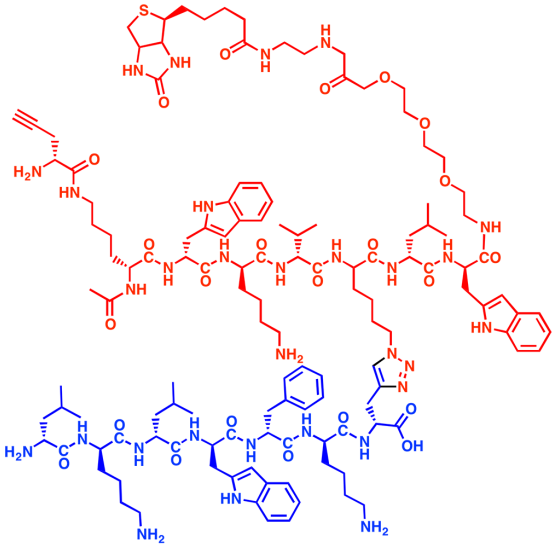


Figure 2.6: Molecular structures of branched biligands 38 and 39 and biligand anchor 40. In the branched biligands **38** and **39**, the 2° arm isolated through screens (in red) is linked by CuAAC chemistry to the monoligand anchor **37** (in blue). An alkyne is appended to the biligand **39** to synthesize biligand anchor **40**. The biotin tag spaced by a polyethylene glycol linker is utilized in assays and in the triligand screen.



38

39



40

2.3.2 Substitution of amino acid in linear ligand to developed branched biligand

An alternative approach to develop a branched biligand would be through amino acid substitution of the developed linear ligand. For this strategy to succeed, one would need to replace amino acids in the linear biligand with an artificial amino acid with similar properties, so as to not disrupt the binding affinity of the ligand. To incorporate a reactive click handle in the reported linear biligand peptide to exploit branched peptides, we synthesized various artificial amino acids containing azide and alkynes in the side chain (Figure 2.7). Hydrophobicity of an amino acid is an important property that can be used to guide replacement of a component amino acid, with minimal disruption in binding properties[22,23]. The Relative Hydrophobicity Coefficient (RHC) is a measure of the hydrophobicity. The RHC values of the naturally occurring as well as the synthesized artificial amino acids were determined by modifying a literature method[21]. In this method, the RHC of an amino acid X is proportional to the retention times of the peptide Ac-X-GLF by HPLC. The RHC values of six of the synthesized amino acids are comparable to those of the hydrophobic naturally occuring amino acids, Leu, Ile, Phe and Trp. Amino acid **21** has RHC value (2.21) lower than those of L-Tyr (**17**, RHC = 2.44) and L-Val (**18**, RHC = 2.61). So **21** can replace L-Tyr or L-Val within a peptide without a big change in hydrophobicity. Note that the same RHC values apply for replacing the D-amino acid with an artificial amino acid within a D-peptide.

For in situ click screens, this strategy was used to replace an amino acid of the D-anchor peptide with an artificial amino acid of similar hydrophobicity index (Figure 2.9). Starting with the earlier developed linear biligand **41** (Figure 2.9) we replaced D-Trp (RHC = 5.74) with the racemic artificial amino acid, Az5 (**25**, RHC = 6.50) to synthesize branched biligand **42**.

The other strategy of amino acid replacement commonly followed in literature is to replace an amino acid with an artificial derivative having similar molecular structure. Following this strategy, D-phenylalanine of the biligand **41** was replaced with 4-azido-D-phenylalanine to synthesize biligand 43**.** The binding affinities of biligands **42** and **43** for bCAII as determined by SPR were similar to the binding affinity of linear biligand **41** for bCAII. So both replacement strategies do not compromise the binding affinity of the modified biligand in SPR measurements (Figure 2.10). However, the yield of **43** was not high due to instability of phenyl azide under the harsh deprotection conditions during peptide synthesis.

Figure 2.7: Synthesized artificial amino acids with side chains containing azide or alkyne functional side chains.

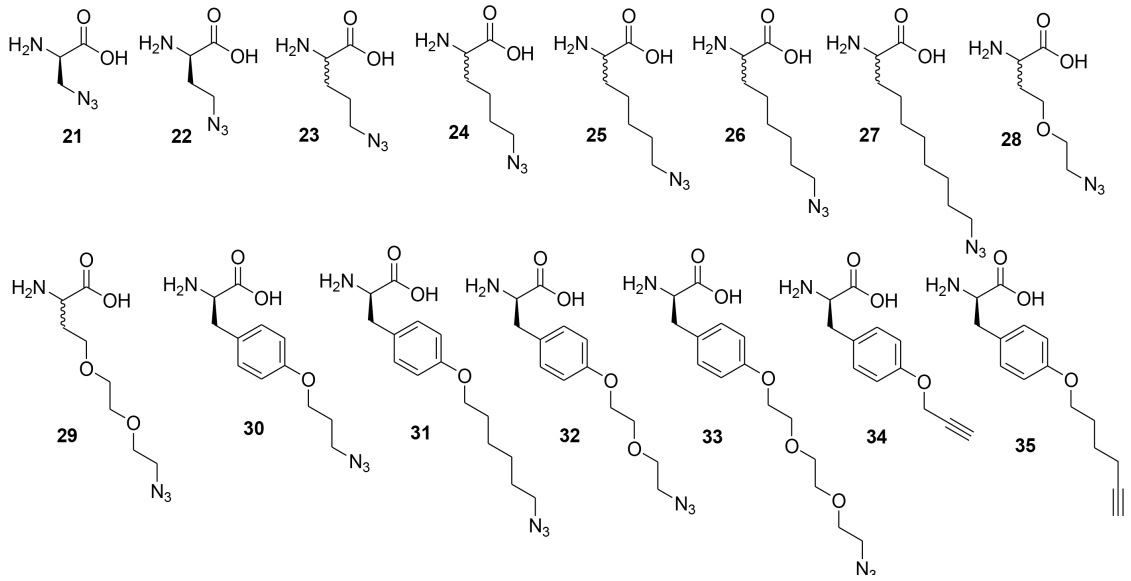


Figure 2.8: Comparison of Relative Hydrophobicity Coefficients (RHC) of natural amio acids with those of synthesized artificial amino acids containing azide or alkyne functional group in the side chain.

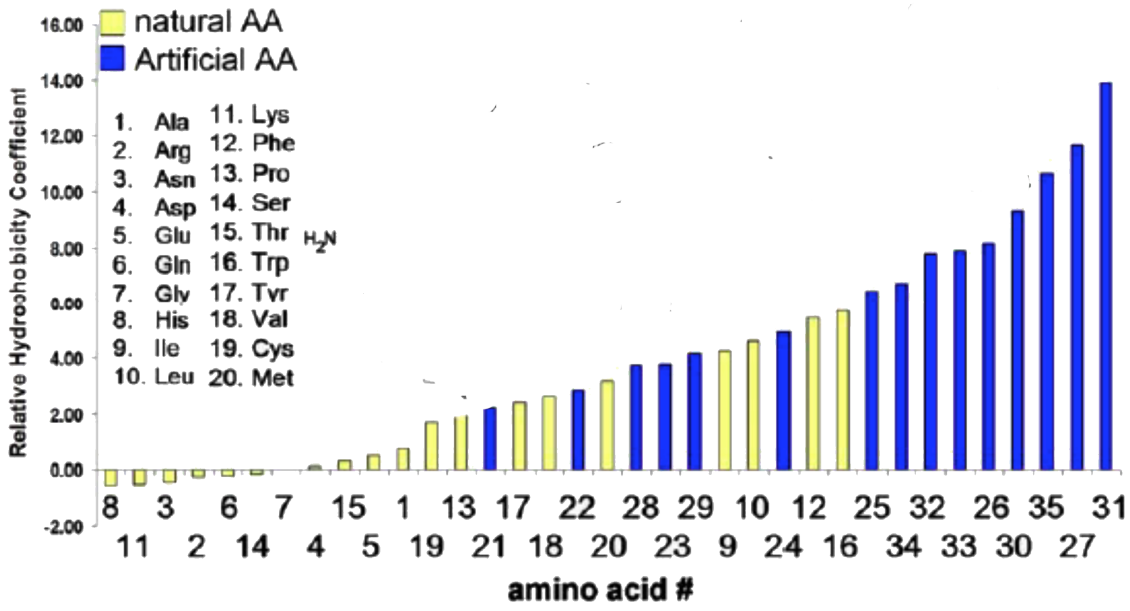


Figure 2.9: Amino acid substitution in the linear biligand to develop branched biligands. The linear biligand peptide **41** developed earlier is modified through amino acid substitutions to yield branched biligands **42** and **43**.

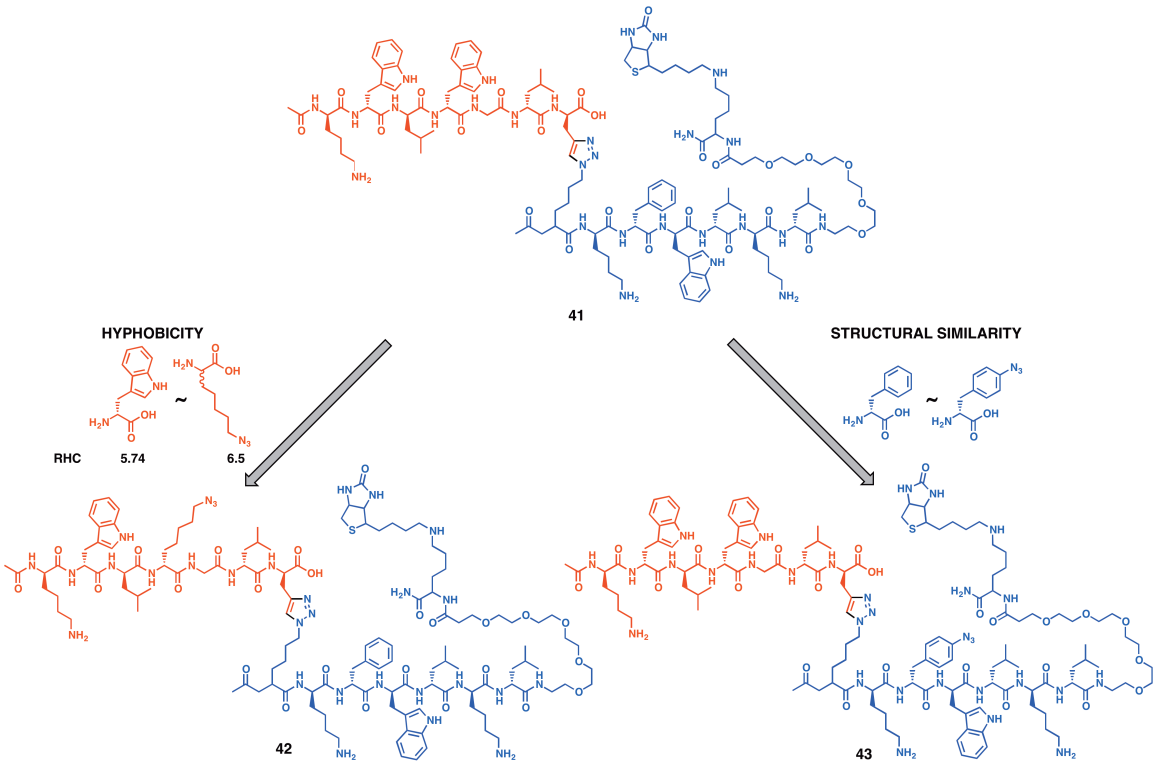
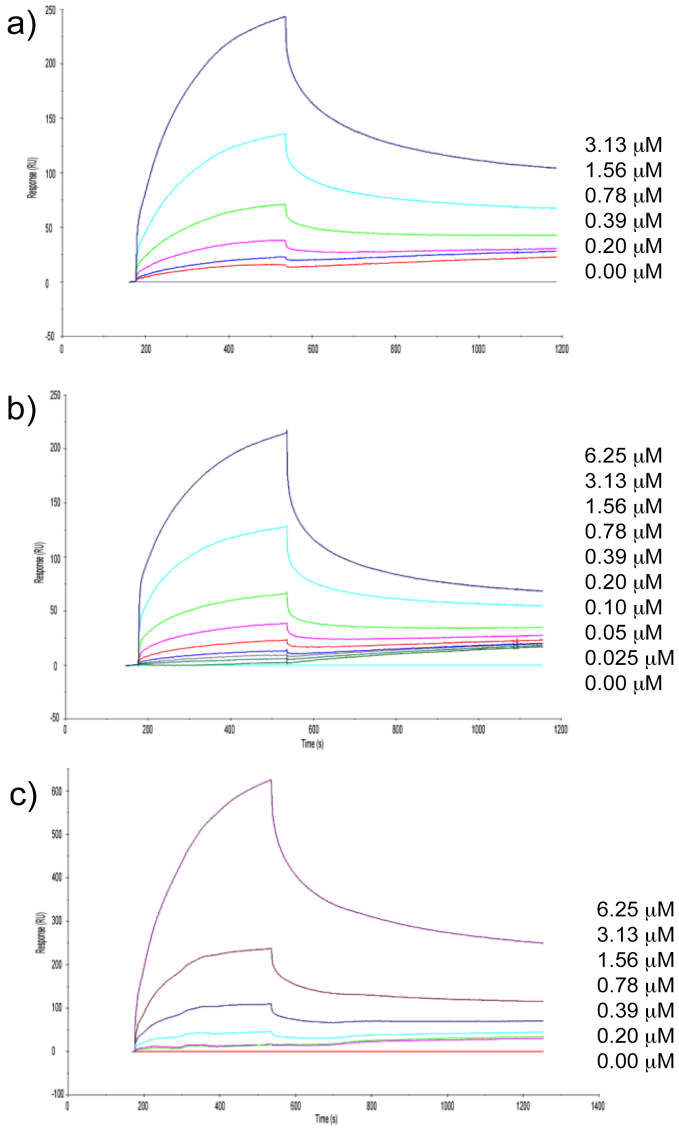


Figure 2.10: SPR response sensograms of the branched biligands developed by amino acid substitution of the linear biligand.

Sensograms of a) linear biligand **41** and branched biligands b) **42** and c) **43.** The immobilization level of bCAII is 4000 RU.



2.3.3 Branched triligand development with substituted biligand

The biligand **42** obtained by modifying the linear biligand **41** using the hydrophobicity approach was utilized as a new anchor peptide for a branched triligand peptide development. In the in situ click target screen, the Alexa Fluor 647 labeled bCAII pretreated with biligand **42** was screened against library D. The beads with high fluorescence intensity are automatically sorted using COPAS plus and sequenced by Maldi TOF/TOF[24] (Table 2.6). The 8 most frequently occurring amino acids were used to generate the focused library. Screening this focused library gave highly homologous peptide sequences (Table 2.7). The hit sequences are found to contain into three motifs (denoted by red, green and blue colors in the sequences). One peptide was chosen from each of the three motifs as the triligand arm. Triligand **45** was constructed with the sequence Ac-(D-Pra)fkifvr as the third peptide arm. Other two triligands **46** and **47** were synthesized with Ac-(D-Pra)irlflk and Ac-(D-Pra)kyffrf as the third peptide arm respectively.

Figure 2.11: Schematic representation of the in situ click target screens TRI1 and TRI2 for developing branched triligand.

The Alexa Fluor 647 (denoted by star) labeled bCAII protein is treated with biligand **42** and screened against on bead peptide library.

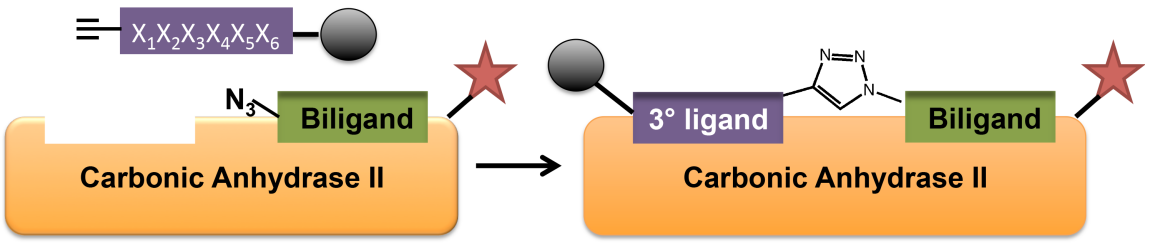


Figure 2.12: Branched triligand development through two generation target screen starting from a comprehensive randomized library.

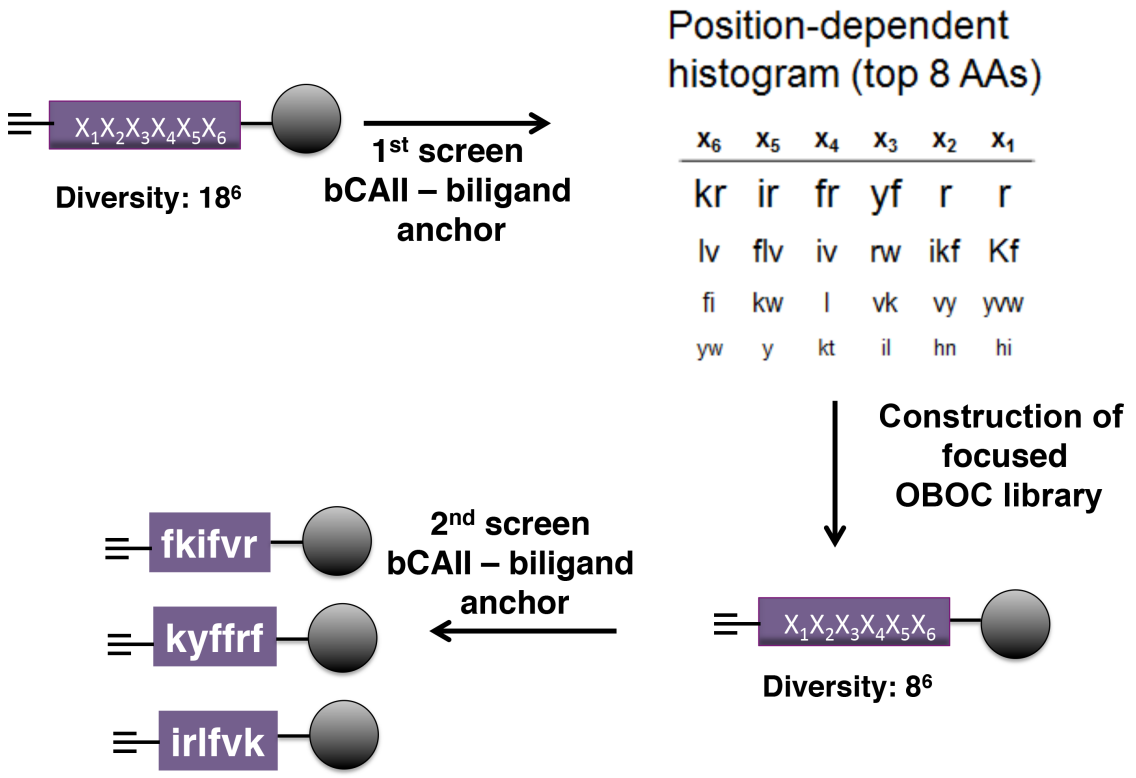
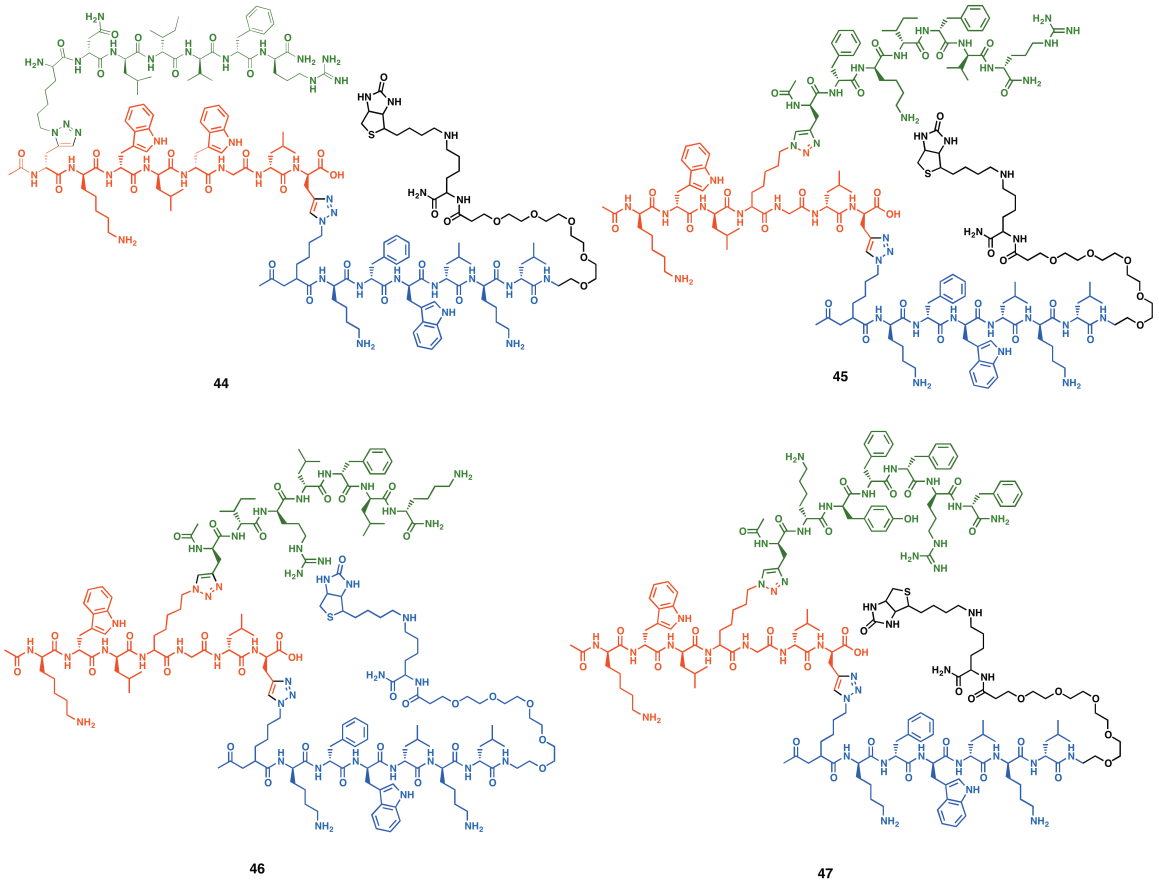


Figure 2.13: Molecular structures of the linear triligand 44 and the three branched triligands 45, 46 and 47 obtained through the target screening process using modified biligand 42 and bovine carbonic anhydrase II protein.



2.3.4 Development of branched triligands using biligand anchor isolated through screens

To select the best triligand binders for bCAII using biligand anchor **40** a product screen was performed. In the earlier target screens, beads with the high fluorescent intensity had been picked as hit beads. These beads contain peptide sequences that bind to the fluorescent protein. As the protein is pretreated with a high concentration of the anchor peptide, it is assumed that the hit sequences bind effectively to the protein in presence of a anchor peptide bound to the protein surface. An alternate method of screening is one that can directly detect the click product formation on the hit beads. This can be achieved by tagging the anchor peptide with the biotin tag and detecting the biotin rather than the protein. After the screen of the OBOC library with the mixture of the protein and the biotin labeled biligand anchor peptide **40**, the protein and peptide binding nonspecifically to the beads are removed by harsh washes with denaturants like guanidium hydrochloride and dimethyl formamide. Only the anchor peptide covalently bound to the on bead peptide via triazole formation remain and are detected by an anti biotin antibody. To eliminate beads binding to the detection antibody, the beads are rescreened. The screen is identical to the first but uses biotin saturated anti biotin antibody for detection. The biotin containing beads remain colorless as the biotin binding site on the antibody is already occupied. The colorless hit beads were picked and sequenced.

The four sequences obtained after the second screen are listed in Table 2.8. Note that the four sequences were obtained as the triligand arms starting from a randomized library containing ~2 million unique sequences, while the sequences for the biligand were obtained from much smaller focused libraries. This process of direct detection of the clicked product is much more stringent than the earlier screen formats and do not need development and screening of consecutively focused libraries. The triligands **48**, **49**, **50** and **51** were synthesized by CuAAC

reaction of the biligand anchor with triligand arm Az4-hdtgf, Az4-hdtfy, Az4-hdegg and Az4-ysqwa, respectively.

Figure 2.14: Scheme for in situ click product screen TRI3 for development of branched triligand.

The click product is directly detected on bead by detecting the biotin tag on the biligand anchor peptide.

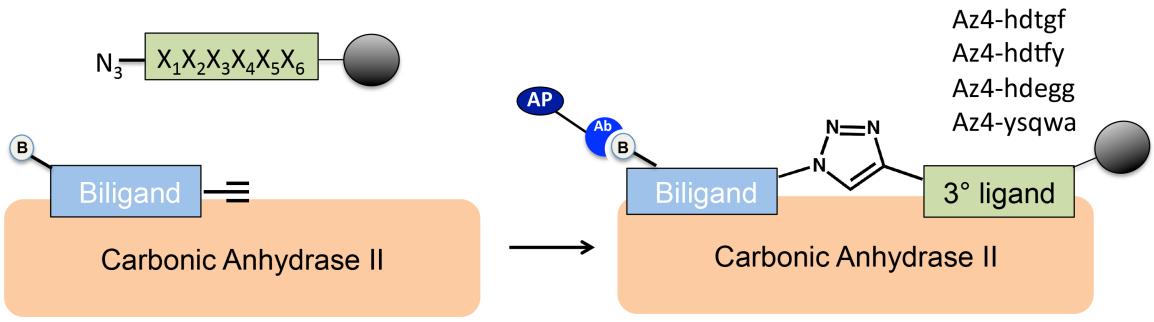
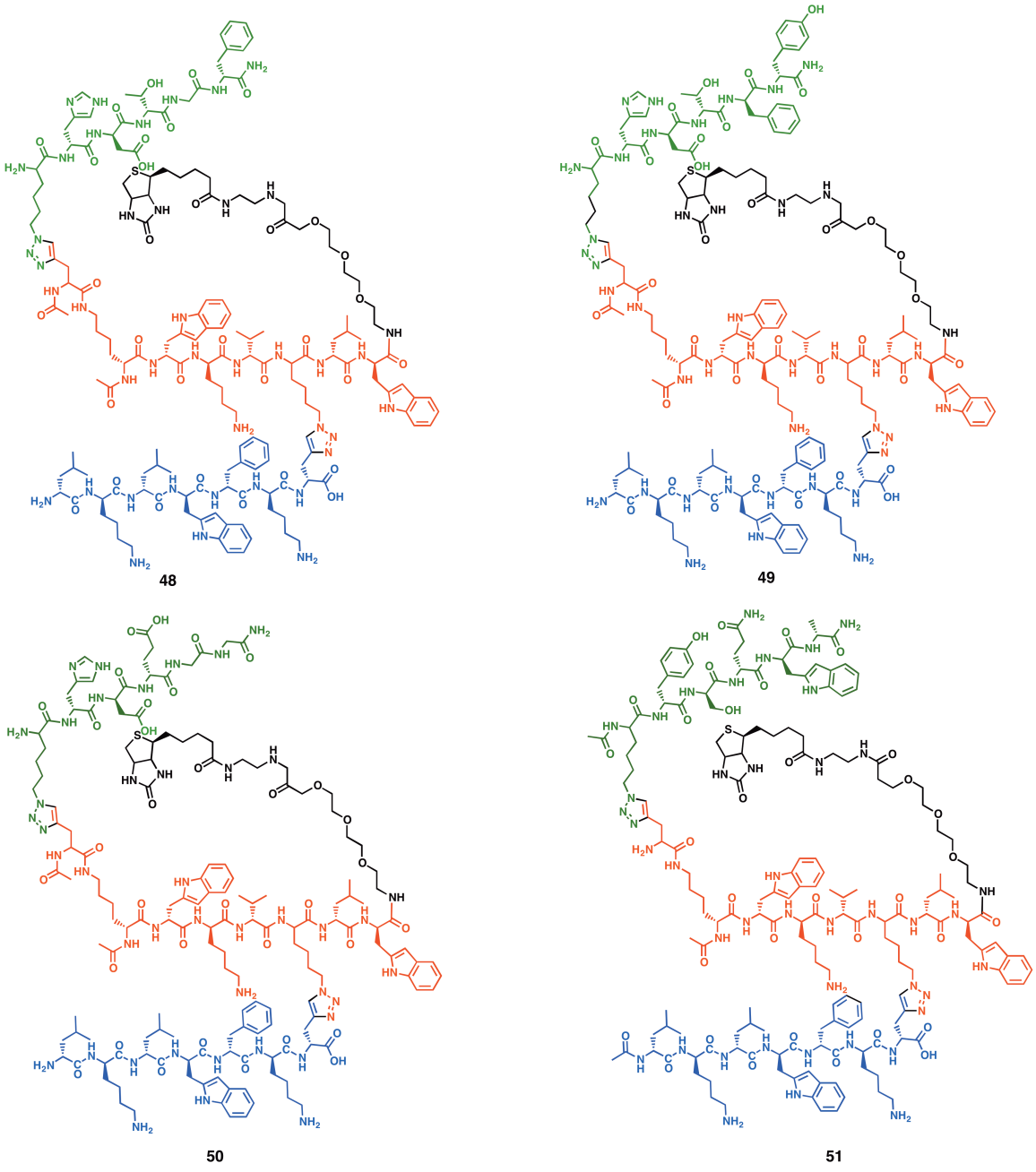


Figure 2.15: Molecular structures of branched triligands developed through product screen against the bovine carbonic anhydrase II.

The triligand arm (in green) is linked to the biligand anchor through CuAAC reaction.



2.3.5 Characteristics of developed capture agents

The branched biligands and triligands developed were tested for their binding affinities, their selectivity and specificty for bCAII. The bovine and human carbonic anhydrase II have high homology in the amino acid sequence and protein structure and it had been noted that the monoligand and linear triligand had similar affinity for both. The same trend was observed for the ligands developed here.

To demonstrate the efficiency of the ligands as detection agents they are used to detect bCAII spotted on a nitrocellulose membrane. The dot blot experiment is performed with a series of concentrations of the protein to determine the minimum amount of protein that can be detected (L.O.D). The L.O.D for the several triligands **45, 46** and **47** is 50 ng, as compared to 100 ng for the linear triligand **44**. The L.O.Ds for all the ligands are noted in the Table 2.9. By nature, the dot blot experiments are dependent on the concentration of the detecting reagent. To cross verify the L.O.D.s for the ligands, two different concentrations of the ligands, 1 μM or 10 nM, are used in different sets of experiments. Although the values are different for the two sets of experiments, the same trend is observed relative to each other.

The use of the developed ligands as capture agents is demonstrated in a series of sandwich ELISA assays. The biotinylated ligands were immobilized on a streptavidin plate and titrated with varying concentrations recombinant human carbonic anhydrase II. The absorbance at 450nm wavelength was proportional to the binding of the capture agent to the protein and was used to calculate the relative affinity curves of the various ligands.

The sandwich ELISA is a more sensitive assay than the dot blot and gives us better insight in the change in binding affinities. In the ELISA experiment we observe that when the D-Trp in the linear biligand **41** was replaced by D,L-azidolysine to create the substituted branched biligand **42**, there was a ~ 2 fold decrease in the binding affinity of the biligand for the protein, despite amino acid **25** and tryptophan having similar RHC values. Thus substitution of an amino

acid in a linear ligand to make it branched did not make it a better binder in the biligand stage. However, when the substituted biligand **42** was used in the next phase of multiligand development process, the resultant triligand candidates **46** (EC_{50} 50 nM) and **47** (EC_{50} 58 nM) have ~3 times higher affinities than the earlier reported linear triligand **44** (EC_{50} 144 nM).

Figure 2.16: Biligands and triligands used as detection agents for bCAII.

Dot blot using biligands for protein detection. The picture on the left specifies the amount of bCAII protein spotted on the membrane.

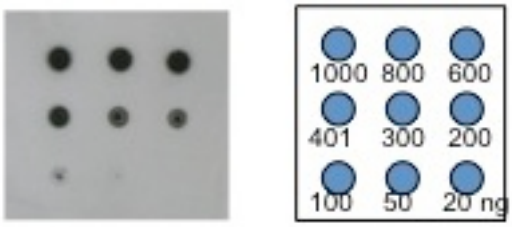
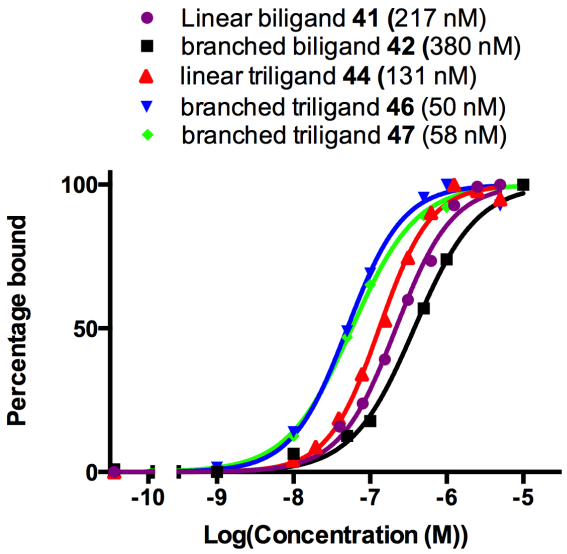


Table 2.9: Table summarizing the limits of detection (L.O.D) of the bovine carbonic anhydrase II protein using the developed multiligands as detection agents.

Ligand	**L.O.D. using 10 nM ligand solution for detection (ng)**	**L.O.D. using 1 uM ligand solution for detection (ng)**
Linear biligand **41**	200	Not determined
Branched biligand **42**	300	Not determined
Linear triligand **44**	100	Not determined
Branched triligand **45**	50	50 ng
Branched triligand **46**	100	50 ng
Branched triligand **47**	50	50 ng
Branched triligand **48**	200	50 ng
Branched triligand **49**	200	50 ng
Branched triligand **50**	200	50 ng
Branched triligand **51**	200	50 ng

Figure 2.17: ELISA demonstrating relative binding affinities of biligands and triligands. Binding of linear biligand **41** (purple), branched biligand **42** (black), linear triligand **44** (red), and branched triligands **46** (blue) and **47** (green) for the CAII protein. The biotinylated ligands are immobilized on a SA plate and treated with different concentrations of the protein. The individual absorbance values are denoted as points and the fits are denoted as solid lines.



2.3.6 Substitution of triazole linkage by triazole mimic

The orientation of the alkyne and azide during in situ protein catalyzed click chemistry dictates the nature of the triazole linkage between the two functionalities. If the protein scaffolds and orients the azide and alkyne is a syn orientation, a 1,5 - disubstituted triazole linkage is formed by the in situ click reaction. If, on the other hand, the protein scaffolds and orients the azide and alkyne is an anti orientation, a 1,4 - disubstituted triazole linkage is formed by the in situ click reaction. A protein, when catalyzing the click reaction, may yield only one triazole product exclusively. So it is important to check both the triazole products for binding affinity to the protein. While the 1,4 substituted triazole linked peptide biligand can be synthesized relatively simply by copper-catalyzed alkyne azide chemistry (CUAAC), it is synthetically

challenging to synthesize the 1,5 disubstituted triazole linked peptide biligand in bulk for different assays. To overcome this difficulty, a mimic of the 1,5 disubstituted triazole linker was synthesized. As proline has been successfully replaced by the 1,5 disubstituted triazole molecule in proteins[25], we reversed the concept to replace 1,5 disubstituted triazole with proline. To recreate the length of the 1,5-triazole link that is formed between L-azidolysine and D-propargyl glycine, as in the in situ protein click reaction, D-Pro and D-Asp was coupled to the acid side chain of L-diaminobutyric acid.

For evaluating the binding affinity of the 1,5-triazole containing ligands as compared to the 1,4-triazole containing ligands for the protein, two branched biligands **38** and **39** were synthesized containing the 1,5-triazole mimic linker. Each of the ligands were immobilized on an ELISA platform and titrated with the protein. For both the branched triligands **38** and **39,** the 1,5-triazole mimic linked ligands have higher affinities for the CAII than the corresponding 1,4 triazole linked ligands. The binding affinity of the 1,5 triazole mimic linked biligand was ~3.3 times the binding affinity of the 1,4 triaozle linked ligand **38**. The binding affinity of the 1,5-triazole mimic linked biligand for protein CAII was ~3 times that for 1,4 triazole linked branched biligand **39.** This indicates that at the biligand stage, the protein catalyzes the formation of the 1,5 triazole linked biligand, where the two peptide arms are turned towards each other rather than the 1,4 triazole linked biligand, where the two peptide arms are placed away from each other.

Figure 2.18: Molecular structure comparison of 1,5 disubstituted triazole linked biligand and its mimic.

A dipeptide D-Pro-D-Asp is coupled to the side chain of L-diaminobutyric acid to mimic the 1,5 triazole linker. The number of atoms in the linker and the orientation of the triazole are maintained. The two peptide arms are syn to each other in the 1,5 mimic.

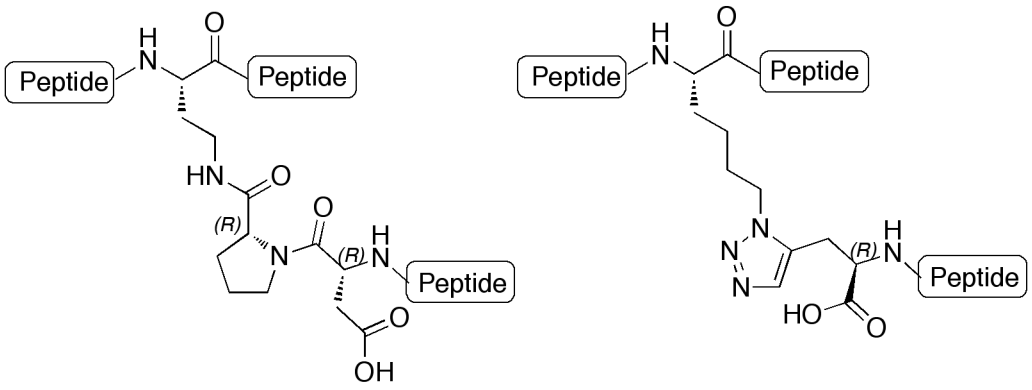
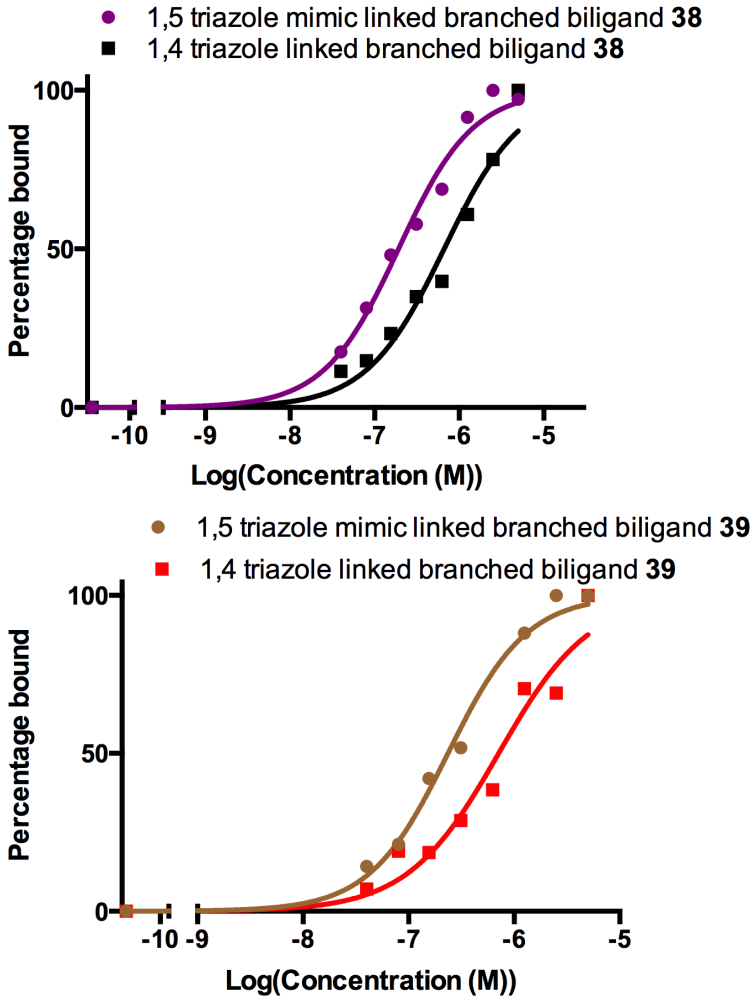


Figure 2.19: Comparison of the binding affinities of 1,4 and 1,5 disubstituted regioisomers of biligands to carbonic anhydrase II protein.



2.4 Conclusions

In this work we evaluate the effect of branching a linear peptide ligand starting with the same weak binding monoligand peptide arm. Two approaches are described. One involves replacing an amino acid in the isolated linear biligand with an amino acid of similar hydrophobicity, to make the biligand branched. The protein, in presence of this branched biligand searches for a third linear peptide arm. The resulting triligands developed are branched. The developed triligands ligands **45**, **46** and **47** have either similar or lower limits of detection than the linear triligand previously developed. The other strategy involves using an OBOC peptide library with random number and position of the branching amino acid azidolysine to find the natural preference of the protein binding in presence of the weak monoligand peptide. Using focused libraries, this pattern is further resolved as wX(Az4)xw or vX(Az4)Xw. The protein in presence of this branched biligand anchor is screened against a linear peptide library to find triligand binders. Of the four ligands isolated though the screens, two triligand arms are found to have a pattern of hdtXX. However despite the high homology of the triligand sequences, their limit of detection are higher than the other branched triligands using the semi quantitative dot blot method. The ligands are also used as capture agents in an ELISA platform. The ELISA assays show the same trend as the dot blots, with branched triligands **46** and **47** having ~3 times higher affinity than the linear triligand **44**.

While the protein might form the 1,4 or 1,5 triazole linkage between the two peptides, where the relative peptide orientations are very different, bulk synthesis and characterizations of ligands are performed with 1,4 triazole linked peptide ligands. Here we determine whether selecting a particular regioisomer for further screens can lead to a change in binding affinity. To circumvent the difficult synthesis of the 1,5 triazole linked peptide, a mimicking strategy is applied that conserves the linker length and triazole cis or trans orientation. All the biligands are

found to favor the triazole cis orientation, particularly the branched biligands developed through screening. This might partially explain why we do not attain the desired affinity of the PCC agent in the triligand stage, as we screened with the 1,4 diastereomer of the biligand with low affinity. In the triligand stage also the 1,5 regioisomers might have better affinities than the 1,4 regioisomers. In the development of future PCC agents, it would be beneficial to check the binding affinity of the different regioisomers of the multiligand from the biligand stage and choose the best binder before proceding to triligand development.

REFERENCES:

(1) Binz, H. K.; Amstutz, P.; Kohl, A.; Stumpp, M. T.; Briand, C.; Forrer, P.; Grutter, M. G.; Pluckthun, A. *Nature Biotechnology* **2004**, *22*, 575.

(2) Hey, T.; Fiedler, E.; Rudolph, R.; Fiedler, M. *Trends Biotechnol* **2005**, *23*, 514.

(3) Ladner, R. C.; Sato, A. K.; Gorzelany, J.; de Souza, M. *Drug Discovery Today* **2004**, *9*, 525.

(4) Smith, G. P.; Petrenko, V. A. *Chem Rev* **1997**, *97*, 391.

(5) Strehlitz, B.; Nikolaus, N.; Stoltenburg, R. *Sensors* **2008**, *8*, 4296.

(6) Holliger, P.; Hudson, P. J. *Nature Biotechnology* **2005**, *23*, 1126.

(7) Desmyter, A.; Decanniere, K.; Muyldermans, S.; Wyns, L. *J Biol Chem* **2001**, *276*, 26285.

(8) Cox, J. C.; Hayhurst, A.; Hesselberth, J.; Bayer, T. S.; Georgiou, G.; Ellington, A. D. *Nucleic Acids Res* **2002**, *30*.

(9) McCauley, T. G.; Hamaguchi, N.; Stanton, M. *Anal Biochem* **2003**, *319*, 244.

(10) Kehoe, J. W.; Kay, B. K. *Chemical Reviews* **2005**, *105*, 4056.

(11) Brody, E. N.; Gold, L. *J Biotechnol* **2000**, *74*.

(12) 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. *Angewandte Chemie-International Edition* **2009**, *48*, 4944.

(13) Doig, A. J.; Sternberg, M. J. *Protein Sci* **1995**, *4*, 2247.

(14) Falciani, C.; Lozzi, L.; Pini, A.; Corti, F.; Fabbrini, M.; Bernini, A.; Lelli, B.; Niccolai, N.; Bracci, L. *Chem Biol Drug Des* **2007**, *69*, 216.

(15) Bohrsch, V.; Mathew, T.; Zieringer, M.; Vallee, M. R.; Artner, L. M.; Dernedde, J.; Haag, R.; Hackenberger, C. P. *Org Biomol Chem* **2012**, *10*, 6211.

(16) Le Chevalier Isaad, A.; Papini, A. M.; Chorev, M.; Rovero, P. *J Pept Sci* **2009**, *15*, 451.

(17) Sun, D.; Jones, V.; Carson, E. I.; Lee, R. E.; Scherman, M. S.; McNeil, M. R.; Lee, R. E. *Bioorg Med Chem Lett* **2007**, *17*, 6899.

(18) Wilkinson, B. L.; Bornaghi, L. F.; Houston, T. A.; Innocenti, A.; Supuran, C. T.; Poulsen, S.-A. *J Med Chem* **2006**, *49*, 6539.

(19) Chenault, H. K., Dahmer, J., Whitesides, G.M. *J. Am. Chem. Soc.* **1989**, *111*, 6354.

(20) Fields, G. B.; Noble, R. L. *International Journal of Peptide and Protein Research* **1990**, *35*, 161.

(21) Kovacs, J. M.; Mant, C. T.; Hodges, R. S. *Biopolymers* **2006**, *84*, 283.

(22) Kiyota, T.; Lee, S.; Sugihara, G. *Biochemistry* **1996**, *35*, 13196.

(23) Wieprecht, T.; Dathe, M.; Beyermann, M.; Krause, E.; Maloy, W. L.; MacDonald, D. L.; Bienert, M. *Biochemistry* **1997**, *36*, 6124.

(24) Franz, A. H.; Liu, R.; Song, A.; Lam, K. S.; Lebrilla, C. B. *Journal of Combinatorial Chemistry* **2003**, *5*, 125.

(25) Tam, A.; Arnold, U.; Soellner, M. B.; Raines, R. T. *Journal of the American Chemical Society* **2007**, *129*, 12670.