

# Capture Compound Mass Spectrometry: A Technology for the Investigation of Small Molecule Protein Interactions

Hubert Köster,<sup>1</sup> Daniel P. Little,<sup>2</sup> Peng Luan,<sup>3</sup> Rolf Muller,<sup>4</sup> Suhaib M. Siddiqi,<sup>5</sup>  
Subramanian Marappan,<sup>5</sup> and Ping Yip<sup>6</sup>

**Abstract:** One of the major hurdles in the post-genomic era is to understand the function of genes and the interplay of many different cellular proteins. This is especially important for drug development. Capture compound mass spectrometry (CCMS) addresses this challenge by selectively reducing the complexity of the proteome. Capture compounds are trifunctional molecules: a selectivity function reversibly interacts via affinity with proteins; a reactivity function irreversibly forms a covalent bond outside the affinity binding site; and a sorting/pullout function allows the captured protein(s) to be isolated from cellular lysate for mass spectrometric analysis and characterization by database queries. In the present study, we demonstrate the use of a CCMS capture compound with a sulfonamide drug analog as its selectivity function, isolating an expected target protein from cell lysates containing a large excess of other “non-target” proteins. A future application of CCMS is to define or confirm drug target proteins and their mechanisms of drug action, or to discover off-target proteins that cause side effects, enabling subsequent drug structure optimization.

## Introduction

**A**NALYZING THE STRUCTURE AND FUNCTION of all cellular proteins (the proteome) with the goal of understanding the time-dependent interplay of the different biochemical and signaling pathways is a daunting task. Protein structural changes (*e.g.*, phosphorylation, glycosylation, or other posttranslational modifications) and differences in the level of protein expression are involved in the development of certain disease states. Associating such protein changes with disease by directly focusing on the expressed proteins bypasses the low correlation of protein abundance with the corresponding mRNA transcripts.<sup>1–3</sup>

There are various approaches to analyze complex protein mixtures.<sup>4,5</sup> Despite sophisticated mass spectrometry

(MS) and bioinformatics there is a need for intelligent chemistries to reduce the complexity of the proteome. Classical approaches include two-dimensional gel electrophoresis<sup>6</sup> and affinity chromatography<sup>7</sup> combined with MS. These important analytical technologies are hampered by the limitation to work only in an aqueous environment. This is especially problematic when analyzing membrane proteins. Therefore there is a need for protein analysis technologies, based on a different technological approach, that are complementary to the existing proven analytical tools. Recently chemical probes have been introduced.<sup>8,9</sup> “Activity-based probes” utilize active site-directed probes, which form a covalent enzyme-inactivating bond within the active site to determine the functional state of enzymes.<sup>9,10</sup> Activity-based probes are limited to the analysis of active enzymes.

<sup>1</sup>Caprotec GmbH, Berlin, Germany.

<sup>2</sup>Melrose, MA.

<sup>3</sup>Portola, Inc., South San Francisco, CA.

<sup>4</sup>Biomatrix, Inc., San Diego, CA.

<sup>5</sup>Helios Biosciences, Inc., Cambridge, MA.

<sup>6</sup>Correlogic Systems, Inc., Rockville, MD.

**ABBREVIATIONS:** CA, carbonic anhydrase; CCMS, capture compound mass spectrometry; HEK293, human embryonic kidney cell line 293; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; NMR, nuclear magnetic resonance; TOF, time-of-flight.

This paper introduces capture compounds, which represent simple, versatile, and broadly applicable chemical probes combined with MS (capture compound MS [CCMS]). The CCMS capture compounds are designed for the rapid and qualitative investigation of small molecule–protein interactions. Small molecule ligands serve as a “selectivity function” for affinity-based interactions, resulting in a biased proteome–small molecule screening approach.

Capture compounds are trifunctional probes, with (1) a reversibly interacting selectivity group, (2) a covalent bond-forming reactivity group, and (3) a sorting/pullout function. Matrix-assisted laser desorption ionization (MALDI)/time-of-flight (TOF) MS is employed to directly analyze the captured proteins. The CCMS technology does not require fluorescent, chemiluminescent, radioactive, or other common detection markers, since the molecular mass of the captured proteins serves as a very informative signal.

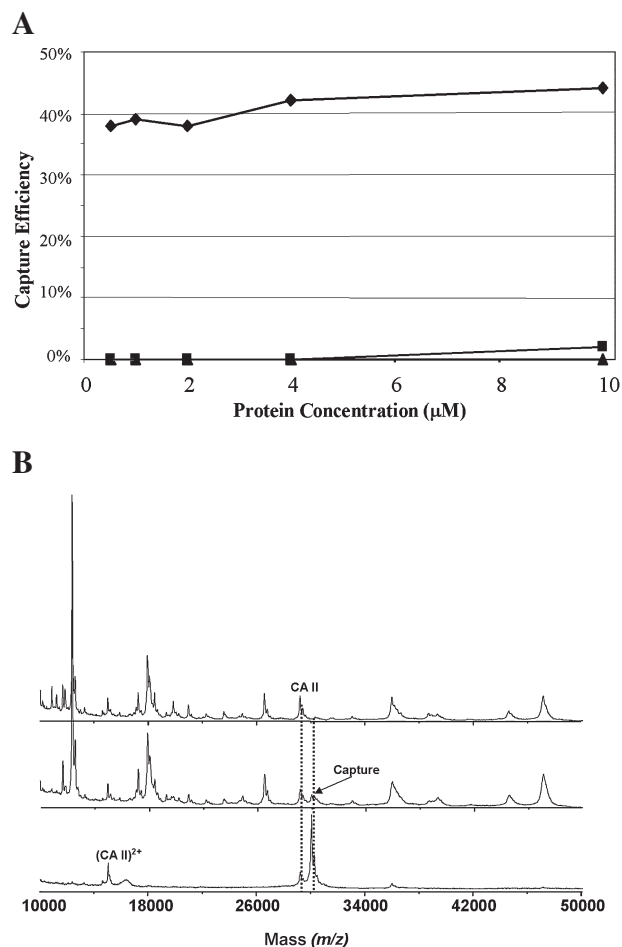
## Materials and Methods

### Materials

Lysozyme (chicken egg, 14,300 Da), alcohol dehydrogenase (baker’s yeast), and carbonic anhydrase (CA) I and II (human erythrocytes, 28,743 Da and 29,100 Da, respectively) were from Sigma (St. Louis, MO). The human embryonic kidney cell line 293 (HEK293) was obtained from American Tissue Culture Collection. Red blood cells were prepared from whole blood (obtained from a local blood bank). MALDI-TOF MS was performed on a Voyager DE-STR (Applied Biosystems, Foster City, CA) using  $\alpha$ -cyano-4-hydroxycinnamic acid or sinapinic acid (Sigma) as the matrix. MS-quality water was obtained from Sigma. Cells were grown according to the American Tissue Culture Collection protocol. The S100 fraction was used as the cytosolic fraction, which was further fractionated by fast protein liquid chromatography using a BioLogic™ DuoFlow™ system (Bio-Rad, Hercules, CA) over a cationic exchange column (UNO Q1) to obtain the flow-through fraction (Fig. 1B). SoftLink™ avidin resin (Promega, Madison, WI) was used for the pullout of biotinylated captured proteins. Pierce (Milwaukee, WI) spin columns were used to remove excess of capture compounds after the capturing reaction, where it was necessary.

### Solutions

Solvents and reagents of analytical grade were purchased from Sigma. Capture compounds were kept at 10 mM stock solution in dimethyl sulfoxide at  $-20^{\circ}\text{C}$  in the dark. CA II (1 mg of lyophilized powder) was dissolved in 200  $\mu\text{l}$  of 20 mM HEPES buffer, pH 7.2. A working

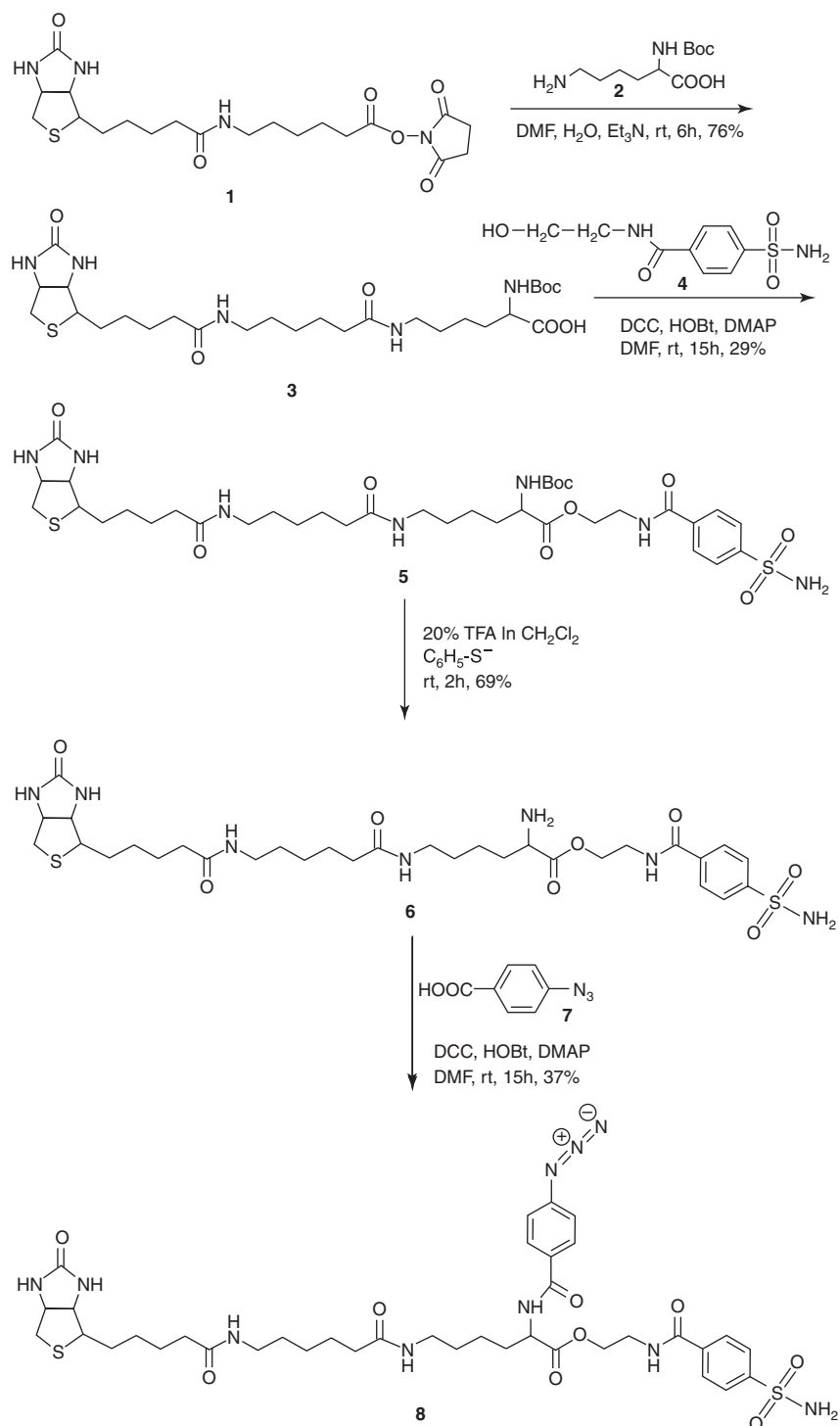


**FIG. 1.** (A) The capture compound **8** ( $5 \mu\text{M}$ ) after interaction with a mixture of CA II ( $\blacklozenge$ ), lysozyme ( $\blacksquare$ ), and alcohol dehydrogenase ( $\blacktriangle$ ). (B) Top spectrum, 200 nM CA II added to fast protein liquid chromatography “flow-through” fraction of cell line HEK293; middle spectrum, capture compound **8** added; bottom spectrum, spectrum after pullout.

stock solution of  $40 \mu\text{M}$  CAII in the same buffer was made subsequently and aliquoted and kept frozen at  $-20^{\circ}\text{C}$ . Washing buffer for the pullout was 10 mM HEPES (pH 7.2), 1 M NaCl, 0.1% Triton X-100, 2 mM EDTA, and 4 mM dithiothreitol.

### Chemical synthesis of capture compound **8** (Scheme 1)

2-tert-Butoxycarbonylamino-6-{6-[5-(2-oxo-hexahydro-thieno-[3,4-d]imidazol-4-yl)-pentanoylamino]-hexanoylamino}-hexanoic acid (**3**). *N*-Hydroxysuccinimide biotin analog **1** was prepared according to standard procedures or purchased from Sigma or Pierce. tert-Butoxycarbonyl-Lys-OH **2** (Sigma) (3.4 g, 14 mmol) dissolved in *N,N*-dimethylformamide:H<sub>2</sub>O:triethylamine (100:10:2 ml) mixture was added to compound **1** (6 g, 13.2 mmol), and the reaction mixture was stirred at room temperature for 6 h. The solvent was removed under vac-



**SCHEME 1.** The chemical synthesis of capture compound **8**. Boc, *tert*-butoxycarbonyl; DCC, dicyclohexyl carbodiimide; DMAP, 4-*N,N*-dimethylaminopyridine; DMF, *N,N*-dimethylformamide; Et<sub>3</sub>N, triethylamine; HOBT, hydroxybenzotriazol-1-ol; rt (r.t.), room temperature; TFA, trifluoroacetic acid.

uum, and the solid obtained was washed with tetrahydrofuran (100 ml) and dried, yielding biotin lysine coupling product **3** as a white solid with a yield of 5.84 g (76%). Analysis gave the following: <sup>1</sup>H nuclear magnetic

resonance (NMR) (300 MHz [Bruker, Billerica, MA], CD<sub>3</sub>OD), δ 4.49 (dd, 1 H), 4.31 (dd, 1 H), 4.04 (m, 1 H), 3.18 (m, 5 H), 2.92 (dd, 1 H), 2.70 (d, 1H), 2.18 (m, 4 H), 1.8–1.4 (m, 18 H), 1.43 (s, 9 H); MS (MALDI) for

$C_{27}H_{47}N_5O_7S$   $m/z$ , calculated 585.8, 624.3, found 608.3 ( $M + Na$ )<sup>+</sup>, 624.3 ( $M + K$ )<sup>+</sup>.

2-*tert*-Butoxycarbonylamino-6- $\{6-[5-(2-oxo-hexahydro-thieno[3,4-d]imidazol-4-yl)-pentanoylamino]-hexanoylamino\}$ -hexanoic acid 2-(4-sulfamoyl-benzoylamino)-ethyl ester (**5**). To a solution of 2-*tert*-butoxycarbonylamino-6- $\{6-[5-(2-oxo-hexahydro-thieno[3,4-d]imidazol-4-yl)-pentanoylamino]-hexanoylamino\}$ -hexanoic acid (**3**) (1 g, 1.7 mmol) dissolved in 20 ml of anhydrous *N,N*-dimethylformamide were added dicyclohexylcarbodiimide (400 mg, 1.94 mmol) and hydroxybenzotriaz-1-ol (270 mg, 2.0 mmol), and the reaction mixture was stirred at room temperature for 15 min. To this stirred solution sulfonamide **4** (450 mg, 1.84 mmol) and 4-*N,N*-dimethylaminopyridine (250 mg, 2.0 mmol) were added. This reaction mixture was stirred at room temperature overnight, the solvent was removed under high vacuum, and the solid residue obtained was purified by silica gel column chromatography using 20%  $CH_3OH-CH_2Cl_2$  mixture as a solvent. Evaporation of the solvent afforded compound **5** as a white crystalline solid with a yield of 400 mg (29%). Analysis gave the following: <sup>1</sup>H NMR (300 MHz,  $CD_3OD$ ),  $\delta$  7.98 (m, 4 H), 4.46 (m, 2 H), 4.36 (dd, 1 H), 3.72 (t, 2H), 3.52 (t, 2H), 3.16 (m, 5 H), 2.90 (dd, 1 H), 2.69 (d, 1H), 2.18 (t, 2H), 2.14 (t, 2 H), 2–1.2 (m, 18 H), 1.42 (s, 9 H); MS (MALDI) for  $C_{36}H_{57}N_7O_{10}S_2$   $m/z$  calculated 812.03, found 813.43 ( $M + H$ )<sup>+</sup>.

2-Amino-6- $\{6-[5-(2-oxo-hexahydro-thieno[3,4-d]imidazol-4-yl)-pentanoylamino]-hexanoylamino\}$ -hexanoic acid 2-(4-sulfamoyl-benzoylamino)-ethyl ester (**6**). To a suspension of biotin sulfonamide derivative **5** (300 mg, 0.36 mmol) in  $CH_2Cl_2$  (2 ml) were added thioanisole (0.1 ml) and trifluoroacetic acid (0.4 ml). The clear solution formed in 10 min was allowed to stir at room temperature for 2 h. The solvent was evaporated, and the residue obtained was purified by silica gel column chromatography using 50%  $CH_3OH-CH_2Cl_2$  as a solvent. Removal of the solvent gave compound **6** as a white solid with a yield of 180 mg (69%). Analysis gave the following: <sup>1</sup>H NMR (300 MHz,  $CD_3OD$ ),  $\delta$  7.98 (m, 4 H), 4.46 (m, 2 H), 4.36 (dd, 1 H), 3.72 (t, 2H), 3.52 (t, 2H), 3.16 (m, 5 H), 2.90 (dd, 1 H), 2.69 (d, 1H), 2.18 (t, 2H), 2.14 (t, 2 H), 2–1.2 (m, 18 H); MS (MALDI) for  $C_{31}H_{49}N_7O_8S_2$   $m/z$  calculated 711.91, found 712.36 ( $M + H$ )<sup>+</sup>.

2-(4-Azido-benzoylamino)-6- $\{6-[5-(2-oxo-hexahydro-thieno[3,4-d]imidazol-4-yl)-pentanoylamino]-hexanoylamino\}$ -hexanoic acid 2-(4-sulfamoyl-benzoylamino)-ethyl ester (**8**). To a solution of 4-azidobenzoic acid (**7**) (30 mg, 0.18 mmol) in anhydrous *N,N*-dimethylformamide (2 ml) dicyclohexyl carbodiimide (40 mg, 0.19 mmol) followed by benzotriaz-1-ol (27 mg, 0.2 mmol)

was added, and the reaction mixture was stirred at room temperature for 15 min. Compound **6** (90 mg, 0.12 mmol) dissolved in anhydrous *N,N*-dimethylformamide (3 ml) and 4-*N,N*-dimethylaminopyridine (24 mg, 0.2 mmol) was added to the solution, and the reaction mixture was stirred at room temperature overnight. Thereafter solvent was removed under vacuum, and the residue obtained was purified by silica gel column chromatography using 20%  $CH_3OH-CH_2Cl_2$  as a solvent. Evaporation of the solvent afforded desired compound **8**, as a white crystalline solid with a yield of 40 mg (37%). Analysis gave the following: <sup>1</sup>H NMR (300 MHz,  $CD_3OD$ ),  $\delta$  7.98 (m, 4 H), 7.90 (d, Hk, 2 H), 7.11 (d, Hl, 2 H), 4.46 (m, Ha & Hc, 2 H), 4.36 (dd, Hb, 1 H), 3.16 (m, Hd & He & Hj, 5 H), 2.90 (dd, Hf, 1 H), 2.69 (d, Hg, 1H), 2.18 (t, Hh, 2H), 2.14 (t, Hi, 2 H), 2–1.2 (m, 18 H); MS (MALDI) for  $C_{38}H_{53}N_{10}O_9S_2$   $m/z$ , calculated 825.97, 797.9 ( $M - N_2$ )<sup>+</sup>, found 826.4 ( $M + H$ )<sup>+</sup>, 798.2 ( $M - N_2$ )<sup>+</sup>.

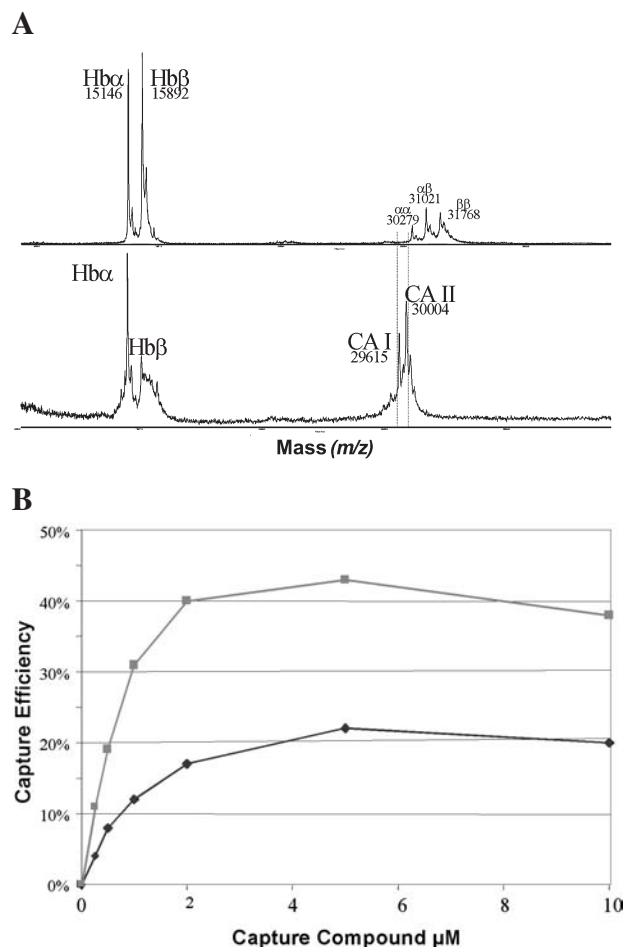
#### Capture of CA II from HEK293 flow-through fraction with capture compound **8**

As a representative example for capture and pullout experiments the procedure generating the results of Fig. 1B is given. SoftLink avidin resin (Promega) was washed and equilibrated three times with 20 mM HEPES (pH 7.2), and 5  $\mu$ l of resin slurry was used for one pullout.

To obtain the HEK293 flow-through fraction, cells were harvested and lysed in 20 mM HEPES (pH 7.2), 250 mM sucrose, 2 mM EDTA, and 1 mM dithiothreitol, supplemented with 1 $\times$  protease inhibitor cocktail (Promega). After centrifugation at 100,000  $g$  for 60 min the supernatant (S100) was collected as the total cytosolic protein extract from HEK293 cells. The S100 was then processed by high performance liquid chromatography using an UNO Q1 column (Bio-Rad). Flow-through fractions were pooled, concentrated using a Millipore (Bedford, MA) Centricon device, and buffer-exchanged into 20 mM HEPES, pH 7.2 (total protein concentration was 0.5 mg/ml).

CA II (200 nM final concentration) was added in 25  $\mu$ l of HEK293 flow-through. Capture compound **8** (1  $\mu$ M) was added, and the final volume was adjusted to 50  $\mu$ l using 20 mM HEPES buffer, pH 7.2. The reaction mixture was incubated in the dark at room temperature for 30 min. Photoreaction was carried out using a high-intensity broadband photography flash lamp (B1600 from Alien Bees, Nashville, TN) with 20 shots applied under cooling to remove excessive heat.

CAII pullout was performed by addition of 5  $\mu$ l of resin slurry to 20  $\mu$ l of the photoreaction mixture in a polymerase chain reaction tube and rotation for 30 min at room temperature. The resin was washed four times each with washing buffer and MS-quality water, respectively, and resuspended in 2  $\mu$ l of water. One microliter



**FIG. 2.** (A) Capture of endogenous CA I and CA II out of crude red blood cell lysate with capture compound **8**: top spectrum, red blood cell lysate from 16  $\mu\text{l}$  of blood (same MS spectra were obtained before and after the capture event due to the high excess of hemoglobin [Hb]); bottom spectrum, after capture with capture compound **8** and pullout. The molar ratio of Hb:CAI:CA II equals 1,000:10:1. (B) Determination of relative binding constants between capture compound **8** and CA I and CA II.

was transferred onto an MS sample plate, and 1  $\mu\text{l}$  of matrix solution was added, mixed, and let air-dry. MALDI-TOF MS was done under standard conditions.

#### Capture of endogenous CA I and CA II from red blood cells with capture compound **8** (Fig. 2A)

Erythrocytes from human blood were prepared using standard procedures. Red blood cells were washed three times with phosphate-buffered saline. Total red blood cell lysate was obtained by three cycles of freeze-thaw and centrifugation at 80,000  $g$  (Beckman Coulter, Fullerton, CA) for 30 min. The supernatant was used for the capture experiments and endogenous CAs (I and II) were isolated as described (see above). Capture compound **8** (5  $\mu\text{M}$ ) was used, and pullout was done with

5  $\mu\text{l}$  of capture reaction mixture without any further processing.

#### Dilution test for sensitivity of pullout of CAs from red blood cell lysate with the capture compound **8**

Red blood cell lysate following capture compound exposure and light activation was added into a very large volume of uncaptured cell lysate before pullout. As little as 2  $\mu\text{l}$  of red blood cell lysate after protein capture could be added to 6 ml of untreated red blood cell extract in one pullout experiment without significantly affecting the quality and quantity of the CA II signal in the mass spectrum. This procedure diluted the precaptured target protein by a factor of 3,000 while keeping the background protein concentration unchanged.

#### Determination of binding constants ( $K_D$ )

This approach is based on the very fast time scale of photolytic reactions (nanoseconds to milliseconds). Under ideal conditions the photolytic reaction can take a snapshot of an enzyme–substrate complex in equilibrium. The amount of covalently cross-linked enzyme–substrate complex is, however, dependent on various parameters (see below). The amount of the covalently cross-linked enzyme–substrate complex is measured by MALDI TOF MS.

#### Equilibrium analysis

The dissociation constant  $K_D$  is defined as

$$K_D = [S][E]/[ES]$$

where [S], [E], and [ES] are the concentrations of the free substrate, free enzyme, and substrate–enzyme complex, respectively. Introduction of measurable parameters,  $[S_0]$  = the initial concentration of substrate and  $[E_0]$  = the initial concentration of enzyme, makes this equation directly useful, leading to

$$K_D = ([S_0] - [ES])([E_0] - [ES])/[ES]$$

#### After photolysis

Not all of the bound enzyme substrate complex in equilibrium will be transformed into a covalently cross-linked enzyme–substrate complex. The photolytically created activated species of the bound enzyme–substrate complex can follow different reaction pathways. Part of it will form the covalently cross-linked enzyme–substrate complex, which will be measured by MS. The surrounding aqueous buffer may deactivate some part before it can form a covalent cross-link. This part, as well as the deactivated capture compound in solution, will be removed by the pullout and subsequent washing steps and remains

undetected during the analytical process. Some activated species may not find an appropriate bond in the protein-binding site chemically available for covalent bond formation. In order to account for quenching of light-activated bound enzyme substrate complex by surrounding buffer, the efficiency of covalent bond formation of the crosslink, and quantum yield, a factor  $\alpha$  is introduced. This factor  $\alpha$  represents the overall conversion efficiency of bound enzyme–substrate complex to covalently cross-linked enzyme–substrate complex, which is measured by MS. The concentration of the covalently cross-linked enzyme–substrate complex is thus  $\alpha[\text{ES}]$ .

During pullout the covalently cross-linked enzyme–substrate complex is isolated through the interaction of its biotin moiety with the avidin-coated beads. The pullout process is for practical reasons, not quantitative. With the pullout efficiency  $\beta$  the peak area,  $A$ , in a MALDI TOF mass spectrum gives a direct measurement for the concentration of the isolated pulled-out cross-linked enzyme–substrate complex:

$$A = \beta \times \alpha \times E_0 / (1 + K_D / [S_0])$$

For the determination of  $K_D$  there is no need to quantify  $\alpha$  and  $\beta$  as long as one obtains a calibration curve.

For the determination of absolute  $K_D$  values the equation is transformed to

$$\ln(A) = \ln(\beta) + \ln(\alpha) + \ln(E_0) - \ln(1 + K_D/[S_0])$$

Experiments will be performed preferentially at high substrate concentrations,  $K_D \ll [S_0]$ :

$$\ln(A) = \ln(\beta) + \ln(\alpha) + \ln(E_0) - K_D/[S_0]$$

Thus by plotting  $\ln(A)$  versus  $1/[S_0]$ ,  $K_D$  can be obtained from the slope of the linear fit.

An external standard might be needed to normalize the spectra taken from samples with different values of  $[S_0]$ .

Relative  $K_D$  values can be obtained when the use of an external standard is not available or not desirable, *e.g.*, two isoenzymes interacting with the same compound/substrate. With CCMS it is reasonable to assume that the photolytic and pullout efficiencies are very similar. With dissociation constants  $K_D^1$  and  $K_D^2$ , initial enzyme concentrations  $E_0^1$  and  $E_0^2$ , and MALDI peak areas  $A^1$  and  $A^2$ , respectively:

$$\ln(A^1/A^2) = \ln(E_0^1/E_0^2) - (K_D^1 - K_D^2)/[S_0]$$

allows by plotting the natural log of the relative MALDI MS peak areas against  $1/[S_0]$  the determination of the difference in dissociation constants ( $K_D^1 - K_D^2$ ) directly from the slope of the linear fit. The appealing feature of

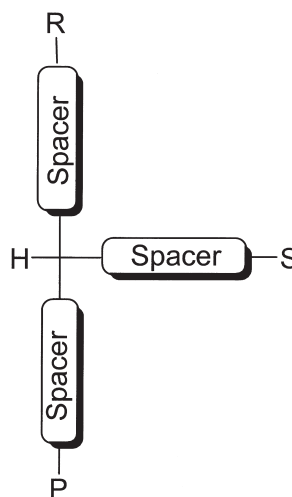
this analysis is that since relative areas are determined, there is no need to normalize the areas from different spectra.

## Results

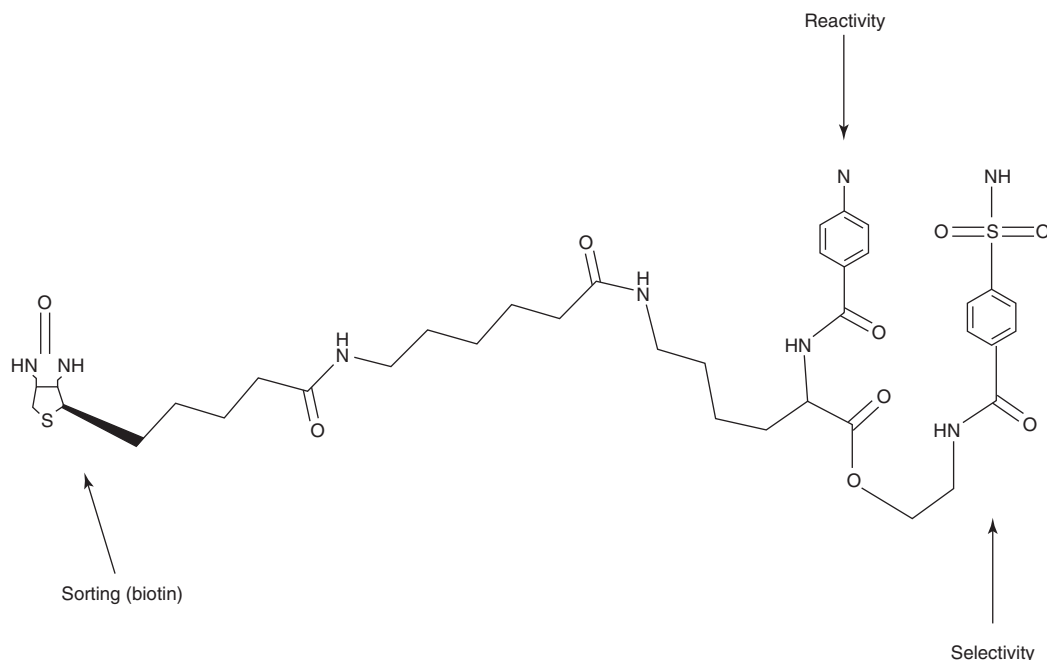
### Design of CCMS capture compounds

CCMS capture compounds are trifunctional small, “star-like” molecules containing a selectivity, reactivity, and sorting (pullout) function (Scheme 2).

The capture compounds are designed to qualitatively investigate small molecule–protein interactions (biased applications) without the need to use fluorescent, chemiluminescent, colorimetric, radioisotopic, or other detection markers. The selectivity function may be, for example, a receptor ligand, a substrate, a transition-state analogue, an enzyme inhibitor, a drug molecule, or a drug metabolite. The three functions are linked to a flexible and sterically unrestrained “star-like” core. Spacers of variable lengths could further modify accessibility and solubility of the capture compounds, *i.e.*, by introducing poly-/oligo-ethylene oxide spacers. The structure is optimized to maximize the independent accessibility of the selectivity and reactivity functions to the protein space in the biological sample and the sorting/pullout function to the avidin-coated beads. After the selectivity function is brought to equilibrium with its corresponding protein binding site(s), a photoactivatable function such as phenyl azide, trifluoromethyl diazirine, or benzophenone is activated by light to form a covalent bond<sup>11,12</sup> near but not within the binding site. Ideally the reactivity group



**SCHEME 2.** Generalized structure of the type of capture compound introduced. S, selectivity function; R, reactivity function; P, sorting or pullout function; H, hydrogen. The spacer can be lipophilic or hydrophilic to influence solubility of the capture compound.



**SCHEME 3.** Chemical structure of capture compound **8** to isolate CA. The selectivity function is the derivative of a prescription drug for the treatment of glaucoma.

is chemically stable and transformed only by irradiation with light into a highly active species with very low selectivity. This will influence the yield of covalent bond formation to create the cross-link. It should also be attached relatively close to the anchor point of the selectivity function so that the covalent cross-link can occur only and close to the affinity binding site of the selectivity group. Linking the selectivity function moiety to the capture compound scaffold in multiple orientations makes different functional and structural features accessible for protein interaction. The results of these interactions may be used to identify drug molecule features involved in target and off-target protein binding.

The experimental procedure is a simple, fast, one-pot reaction amenable to a microtiter plate format for automation and scale-up: a cell lysate from the biological sample (cell line, tissue sample, biopsy, etc.) is directly incubated with the appropriate capture compounds. For sample preparation, gel electrophoresis or chromatography may not be necessary. After affinity interaction and photochemical irradiation yielding covalent bond formation, avidin-coated beads are added. Proteins captured and bound to the beads are extensively washed and either directly or after elution analyzed by MS.

The capture compound **8** in Scheme 3 was designed to capture the enzyme CA. A phenyl sulfonamide group, used here as the selectivity function, is known to reversibly interact with the enzyme binding site of CA. A phenyl azide group serves as the reactivity function in forming an irreversible covalent bond with CA II. Biotin,

used here as the sorting (pullout) function, isolates captured proteins from the cell lysate material upon presentation to avidin-coated beads.

#### *Chemical synthesis of capture compound 8*

The synthesis (Scheme 1) uses a commercially available biotin derivative **1**, with an extended spacer arm attached through an amide bond to optimize efficient binding to the avidin-coated beads. The *N*-hydroxysuccinimide ester moiety enables nucleophilic substitution by the epsilon amino group of *N*-*tert*-butoxycarbonyl-protected lysine **2**. Three steps follow to furnish the capture compound **8**. Compound **4** is the commercially available derivative of a drug molecule prescribed for the treatment of glaucoma.

#### *Use of capture compound 8 for capturing CA*

CA has been extensively studied to design inhibitors for the treatment of different medical indications.<sup>13,14</sup> Derivatives of aromatic sulfonamides are used to treat glaucoma. The mechanism of action is well established through biochemical and X-ray crystallographic studies. Of the two most abundant cytosolic CA isoenzymes in mammalian tissues (CA I and CA II), CA II has a higher catalytic activity and binding affinity to unsubstituted sulfonamides. CA I is 10-fold more abundant in red blood cells, but has lower catalytic activity and medium affinity. In the capture compound of Scheme 3 the drug molecule was coupled to the scaffold, leaving the unsubsti-

tuted sulfonamide pharmacophore functionality available for protein interactions. The amino nitrogen atom is known to reversibly interact with the catalytically essential zinc atom in the active site, thereby inhibiting its catalytic action.<sup>13</sup> In control experiments the specificity and proper function of the selectivity group were tested by measuring the influence of the other major structural features of capture compound **8** on enzyme inhibition (Table 1). Neither biotin with hydrocarbon spacer nor the capture compound scaffold with the reactivity group in combination with the biotin moiety was found to inhibit enzyme activity due to the absence of the sulfonamide selectivity function. The sulfonamide ligand **4** has a  $K_D$  value of  $\sim 80$  nM. Biotin linked to the ligand **4** and the capture compound **8** inhibited the enzymatic activity with a  $K_D$  value of  $\sim 120$  nM and  $\sim 100$  nM, respectively. Only in the latter case was a capture event observed through the mass spectrometric detection of covalently attached capture compound **8** to CA. Finally, if the primary amino group was not available for interaction by transforming it into a photoactive azido group no inhibition and no covalent attachment to the enzyme were found. The modified ligand **4** had no affinity interaction within the active site and no ability to covalently cross-link to the enzyme.

Interaction of capture compound **8** with CA II (molecular mass  $\sim 30$  kDa) in the presence of lysozyme ( $\sim 14$  kDa) and alcohol dehydrogenase ( $\sim 65$  kDa) (Fig. 1A) revealed the highly selective interaction with CA II at 1  $\mu$ M concentration (maximum peak intensity in the MALDI-TOF MS spectrum). Alcohol dehydrogenase from yeast was not captured even at 10  $\mu$ M protein concentration. However, a very low unspecific interaction with the small lysozyme molecule could be observed. Capture efficiencies exceeding 40–46% were not obtained under the conditions used, probably due to quenching of the bound photoactivated species by excess of buffer, efficiency of bond formation of the cross-link, and/or quantum yield.

To test selective properties of capture in a more complex environment, we added CA II into a fast protein liquid chromatography “flow-through” fraction (Fig. 1B) and into the complete cytosole (data not shown) of HEK293 cells at 200 nM and 1  $\mu$ M concentration, respectively. Here CA II was present in a mixture of about 150 to several thousand different proteins, respectively. Under the capture conditions used again a capture of about 40% of CA II was observed by the integration of the two protein peaks (captured CA II identified by the distinct mass shift through addition of 830 Da of the scaffold to CA II with molecular mass of 29,100 Da). Analysis of the mass spectrum after pullout is straightforward. The results were in both cases very similar. The predominant peaks were generated by the parent and dependent ions from CA; Fig. 1B revealed no other peaks, whereas in the interaction of the capture compound **8** with

the HEK293 cytosol two proteins in very low abundance could be detected at 22.5 kDa and 42 kDa, respectively. CA was not only purified, but the signal significantly amplified as a result of the increased ionization efficiency after elimination of most other proteins.

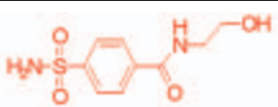
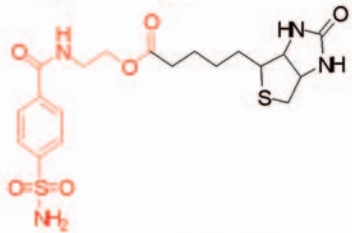
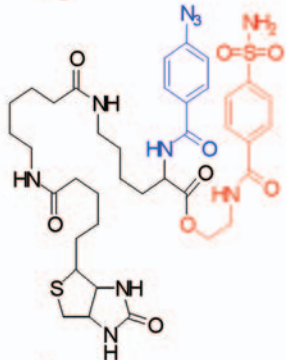
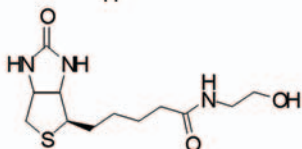
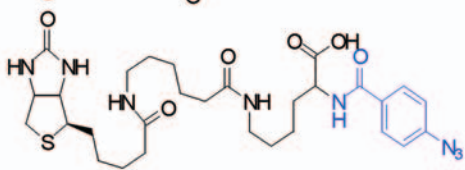
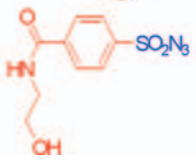
Through formation of a covalent bond, CCMS allows with its integrated pullout function the isolation of weak drug–protein interactions, in contrast to most other methods such as two-dimensional gel electrophoresis, affinity chromatography, and protein chips. By increasing the capture compound concentration from between 1 and 5  $\mu$ M for strong interactions to 100  $\mu$ M the equilibrium of weak binding proteins can be shifted to the right, allowing the identification and isolation of weakly interacting proteins (data not shown).

In Fig. 2A, capture compound **8** was added to a crude lysate of red blood cells to capture the endogenous CA I and CA II. After pullout CA I and CA II are detectable with significantly amplified signal intensities. In a control experiment (data not shown) it was demonstrated that a 3,000-fold dilution with untreated red blood cell lysate does not significantly reduce the detection efficiency of captured proteins. This reflects the extraordinary high affinity of the biotin-avidin system. The small amount of the two highly abundant hemoglobin chains is due to non-specific interaction with avidin-coated beads (no mass shift proves no capture). The lower peak intensity of the more abundant CA I relative to CA II was due to its lower affinity to ligand **4** (Fig. 2B). Employing an increasing excess of the capture compound over a constant protein concentration allowed the determination of binding constants using MALDI-TOF MS. Using capture compound **8** (Fig. 2B), equilibrium was reached at about fivefold excess at a different capture level for CA I and CAII, reflecting the different binding strengths of the two enzymes to the same ligand. Algorithms have been developed for the determination of absolute and relative binding constants from these curves (see Materials and Methods).

## Discussion

The CCMS capture compounds offer a complementary approach for proteome analysis due to the different technological approach (chemical probe vs. chromatography and electrophoresis). The capture compounds have several distinct advantages over existing technology. The CCMS capture compounds are easy-to-synthesize small molecules, whose structural identity and purity can be unambiguously proven by analytical methods. This is difficult with insoluble material as used in affinity chromatography and protein chips. In the latter two cases the accessibility of the polymer-bound ligand has to be established in using proper solvent conditions, which opti-

TABLE 1.  $K_D$  VALUES OF CAPTURE COMPOUND **8**, INTERMEDIATES, AND CONTROL STRUCTURES

Compound	Structure	Approximate $K_D$ ( $\mu\text{M}$ ) (activity assay)	MALDI
S		0.08	N/A
S+P		0.12	N/A
S+R+P		0.10	Capture
P		$> 10^3$	N/A
P+R		$> 10^3$	No Capture
D*+R		$> 10^3$	No Capture

N/A, not applicable, S (red), drug molecule; D\* (red), modified and activated drug molecule; P, biotin with spacer arm; R (blue), reactive moiety.

mally swell the polymeric network. The density of the cross-links of the polymeric network has a limiting influence on the size of the proteins that can travel through the network in reaching the ligand. With CCMS the critical affinity interaction takes place in solution, eliminating steric interference with solid surfaces and allowing for unencumbered interactions of the small capture compound molecule with large protein molecules. The solubility can be adjusted to be in line with the proper biologically functional environment. This can enable the

CCMS capture compounds to interact with lipophilic proteins such as membrane proteins.

The presence of all the three functions (selectivity, reactivity, and pullout) in the same capture compound enables a highly specific physical isolation of interacting proteins from complex biological samples. The integration of an analytical MS step allows for facile molecular weight-based identification and characterization.

Weak and strong interacting proteins can be separately isolated by varying the ratios of capture compound to the

proteins in the lysate. Since the selectivity group (the ligand) occupies the protein affinity-binding site, the covalent bond is typically formed adjacent to, but not within, the ligand-specific binding site. The use of a photoactivatable reactivity group may allow for the determination of absolute and relative binding constants by shifting the equilibrium employing different capture compound concentrations prior to covalent bond formation.

The strength of CCMS is its qualitative ability to directly isolate proteins of interest from any biological sample and subsequently to investigate and understand the molecular basis of the binding between the protein(s) and its respective ligand(s).

There are several important applications for drug development. By performing the drug protein capture experiments at different capture compound concentrations, interacting proteins can be identified. At the same time the respective binding constants are being determined. If a target protein is known, the drug structure could be optimized by using a focused combinatorial library of structural variants for protein capture; the strongest binder (or the weakest to reduce binding to off-target proteins) can be directly determined. Using CCMS for the determination of binding constants between small molecules and proteins offers a very simple and informative alternative to otherwise laborious and less reproducible methods. Capture compounds in which the selectivity group is the substrate or cofactor of an enzyme are useful for the isolation of that specific class of enzymes. As an example, nucleic acid and protein methyl transferases can be isolated from any biological sample when using *S*-adenosyl homocysteine, an effective analogue of the naturally occurring cofactor *S*-adenosyl methionine, as the selectivity function (H. Köster, C. Dalhoff, and E. Weinhold, unpublished data).

### Acknowledgments

The authors acknowledge gratefully the contributions of Wendy Cordier and Patricia d'Avis (molecular biology, cell biology), Matthew Grealish (chemical synthesis), and Julius Apuy (MS). We further gratefully acknowledge support of Ling Ma for preparing the figures. H.K., D.L., S.S., and P.Y. have financial interests in Caprotec GmbH.

### References

1. Gygi SP, Rochon Y, Franza BR, Aebersold R: Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 1999;19:1720–1730.
2. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, *et al.*: Integrated genomic and proteomic analyses of systematically perturbed metabolic network. *Science* 2001;292:929–934.
3. Griffin TJ, Gygi SP, Ideker T, Rist B, Eng J, Hood L, Aebersold R: Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 2002;1:323–333.
4. Aebersold R, Mann M: Mass spectrometry-based proteomics. *Nature* 2003;422:198–207.
5. Patterson SD, Aebersold R: Proteomics: the first decade and beyond. *Nat Genet* 2003;33:311–323.
6. Lilley KS, Razzaq A, Dupree P: Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. *Curr Opin Chem Biol* 2002;6:46–50.
7. Heftmann E, ed.: Affinity chromatography. *J Biochem Biophys Methods Special Issue* 2001;49:1–750.
8. Jeffery DA, Bogoy M: Chemical proteomics and its application to drug discovery. *Curr Opin Biotechnol* 2001;14: 87–95.
9. Adam GC, Sorensen EJ, Cravatt BF: Chemical strategies for functional proteomics. *Mol Cell Proteomics* 2002;1: 781–790.
10. Hagenstein MC, Musgnug JH, Lotte K, Plessow R, Brockhinke A, Kruse O, *et al.*: Affinity-based tagging of protein families with reversible inhibitors: a concept for functional proteomics. *Angew Chem Int Ed* 2003;42:5635–5638.
11. Nesnas N, Rando RR, Nakanishi K: Synthesis of biotinylated retinoids for cross-linking and isolation of retinal binding proteins. *Tetrahedron* 2002;58:6577–6584.
12. Dorman G, Prestwich GD: Using photolabile ligands in drug discovery and development. *Trends Biotechnol* 2000;18:64–77.
13. Supuran CT, Scozzafava A: Carbonic anhydrase inhibitors. *Curr Med Chem Immunol Endocrinol Metab Agents* 2001;1:61–97.
14. Tripp BC, Smith K, Ferry JG: Carbonic anhydrase: new insights for an ancient enzyme. *J Biol Chem* 2001;276: 48615–48618.

Address reprint requests to:

Dr. Hubert Köster

Caprotec GmbH

Max-Dohrn-Strasse 10

D-10589 Berlin, Germany

E-mail: hubert.koester@caprotec.com