Solid-Phase Synthesis of Double-Headed Epoxysuccinyl Activity-Based Probes for Selective Targeting of Papain Family Cysteine Proteases

Steven H. L. Verhelst and Matthew Bogyo*[a]

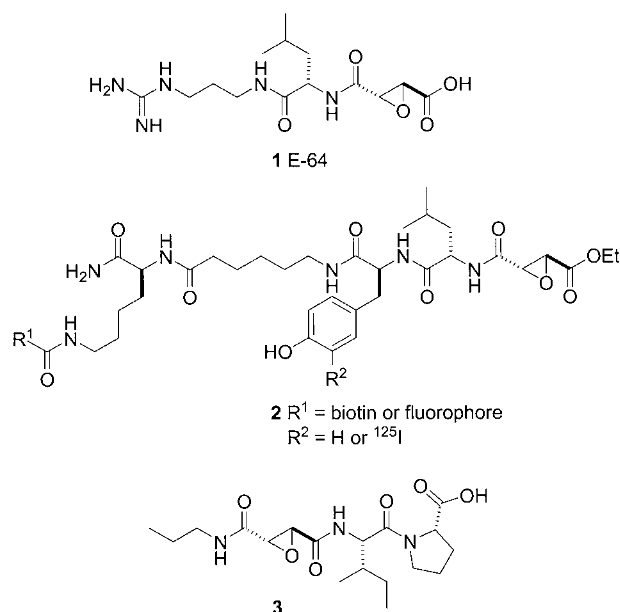
The completion of genome sequencing of several organisms, including man, has provided global information on gene regulation. Emphasis has now shifted to understanding protein function.[1] To facilitate functional studies, small molecules that label subsets of enzymatic proteins have been developed as a means to simplify complex proteomes and allow bulk profiling of enzyme activity.[2] These activity-based probes (ABPs) combine tags for visualization or purification with moieties that covalently attach to the active sites of enzymatic targets in an activity-dependant manner. By using this approach, a specific protein or protein family can be readily monitored in complex protein mixtures, in intact cells, and even in vivo.[3] Furthermore, enzyme-class-specific probes can be used to develop screens for small-molecule inhibitors that can be used in functional studies.[4,5] Therefore, methods that will facilitate the development of novel ABPs have great value for advancing the use of this technology.

Proteolysis controls a number of essential biological processes ranging from cell division to cell death. Amongst the cysteine proteases, the lysosomal cathepsins play important roles in a number of human diseases.[6] However, the members of this family display similar substrate preferences, thus making the development of tools to study individual cathepsins a challenging task.

The natural product E-64 (**1**; Scheme 1) inhibits cysteine proteases by covalent attachment to the active-site sulfhydryl nucleophile. This reagent contains a leucine residue that mimics the critical P2 residue of a substrate and therefore binds efficiently in the S2 specificity pocket of virtually all cysteine cathepsins. As a result this reagent is a broad-spectrum inhibitor and activity probe.

Recently, our laboratory has reported peptide-based ABPs based on the E-64 scaffold.[4,5,7] One of the first ABPs reported with the epoxide moiety, DCG-04 (**2**) has found widespread use for a number of functional studies of the cysteine cathepsins. However, this general probe makes use of a single peptide piece that makes contacts with only one side of the protease active site.

A recent crystal structure of a double-headed epoxide inhibitor showed that the entire structure binds along the active-site cleft, with contact made on both sides of the reactive cysteine nucleophile.[8] Thus, we reasoned that development of a method that would allow facile attachment of diverse peptide sequences on both sides of the epoxide moiety would facilitate the synthesis of probes with increased selectivity compared to the general DCG-04 probe. Here we report our preliminary results involving the solid-phase synthesis of cathepsin B-specific ABPs.



Scheme 1. *Epoxysuccinyl-based cysteine protease inhibitors E-64 (**1**), CA-074 (**3**), and activity-based probe DCG-04 (**2**).*

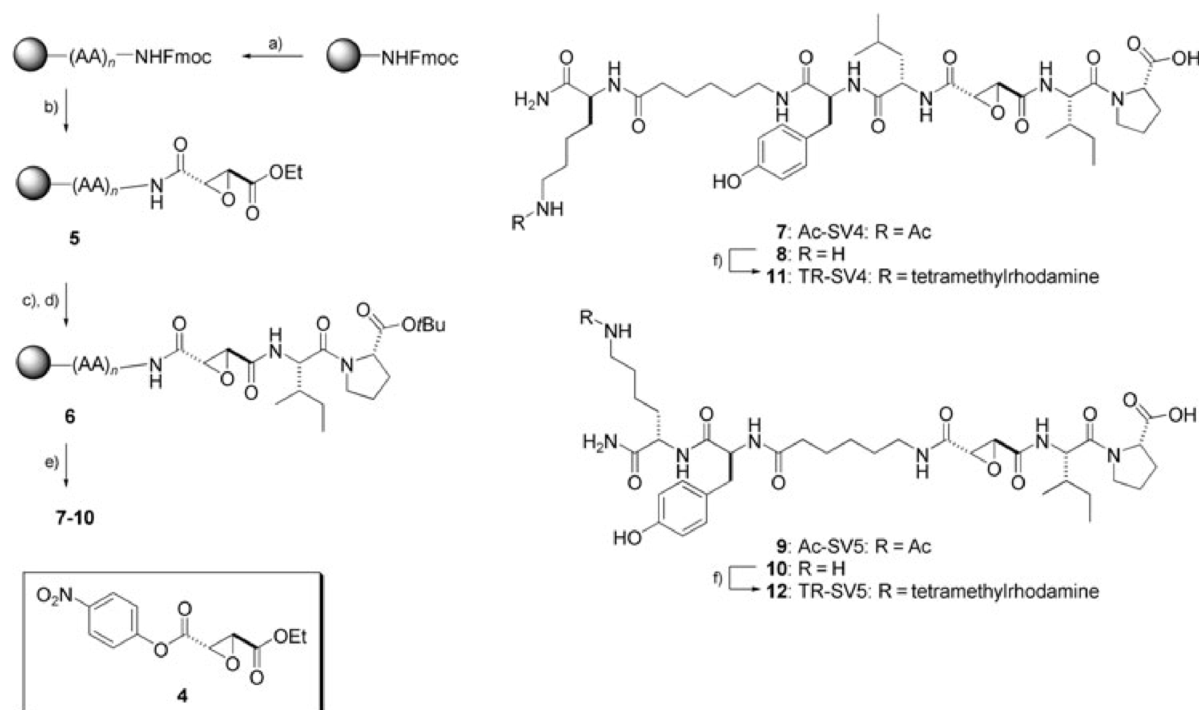
In the past, selective inhibitors for both cathepsins B and L have been reported. Most of these compounds, such as CA-074 (**3**),[9] have taken advantage of diversity elements on both sides of the epoxide. In fact, virtually all of the selectivity for cathepsin B has derived from specific interactions with residues on the occluding loop found only in this protease. This unique specific interaction has been used to generate small molecules that target cathepsin B. However, the previously reported inhibitors and probes were obtained by time-consuming solution-phase synthesis,[10] thus limiting the extent of structure–activity studies that could be performed.

In order to efficiently synthesize double-headed epoxysuccinyl probes, we needed to design a solid-phase synthesis protocol that would allow peptide synthesis to be carried out followed by attachment of the epoxide group and then further elongation of the probe. Our initial probe design was based on the broad-spectrum ABP DCG-04 (**2**) containing the dipeptide of CA-074 (**3**). Scheme 2 outlines the basic strategy which makes use of standard solid-phase peptide chemistry followed by capping of the N terminus with the epoxysuccinyl synthon by using the activated nitrophenyl ester **4**.[11] The epoxysuccinyl ester **5** can then be hydrolyzed on resin thereby allowing elongation of the probe on the other side of the epoxide moiety.

The use of a PEG-based Rink resin permitted the use of alcohols as a solvent system during this ester-bond cleavage. Thus, hydrolysis of ethyl ester **5** was achieved by treatment with a

[a] *Dr. S. H. L. Verhelst, Dr. M. Bogyo*
Departments of Pathology, and Microbiology and Immunology
Stanford University School of Medicine
300 Pasteur Drive, Stanford, CA 94305 (USA)
Fax: (+1) 650-725-7424
E-mail: mbogyo@stanford.edu

 DOI: 10.1002/cbic.200400377 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 2. Reagents and conditions*: a) Elongation with Fmoc-AA-OH (3 equiv), DIC (3 equiv), HOBt (3 equiv) in DMF, 1–2 h, followed by Fmoc deprotection (piperidine/DMF 1:4), 15 min. b)* ***4*** *(3 equiv), DMF, 1 h. c) KOH (0.25 M in EtOH), 20 min. d) H-Ile-Pro-OtBu (3 equiv), PyBOP (3 equiv), DIEA (6 equiv), DMF. e) 95 % TFA, 2.5 % TIS, 2.5 % H_2O. f) Tetramethylrhodamineoxysuccinimide ester, mixed isomers (1 equiv), DIEA (2 equiv), DMSO, 1 h.*

0.25 M solution of KOH in ethanol. LC-MS analysis after test cleavage showed that clean hydrolysis was complete in approximately 20 min.

Since elongation at the distal side of the epoxide takes place in the opposite direction to standard solid-phase peptide synthesis, it was necessary to minimize the couplings performed after ester hydrolysis. Therefore, protected Fmoc-Ile-Pro-O*t*Bu was synthesized in solution and Fmoc-deprotected prior to coupling to the free acid on a solid support. This coupling was achieved by using (benzotriazol-1-yl)oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) as a coupling reagent.

By using the above-described protocol, compounds **7**–**10** were obtained in 20–36 % yield after HPLC purification. Free amines **8** and **10** were conjugated to a tetramethylrhodamine (TR) fluorophore. Probes were synthesized that contained a leucine in the P-2 position (Ac-SV4 and TR-SV4) and that lacked the leucine residue (Ac-SV5 and TR-SV5).

The potency and selectivity of the probes was determined by labeling of cathepsin cysteine proteases in crude rat-liver homogenates. Labeling with ^{125}I-labeled versions of Ac-SV4 and AC-SV5 indicated that Ac-SV5 is highly specific for cathepsin B, whereas Ac-SV4, while somewhat selective for cathepsin B, targets all the primary active cathepsins (B, Z, H, J, and C) at high concentrations (Figure 1). Control samples that were preheated confirm that labeling is activity dependent. Correspondingly, the fluorophore-containing probes show similar potency and specificity patterns, however with a slightly higher nonspecific labeling of other proteins.

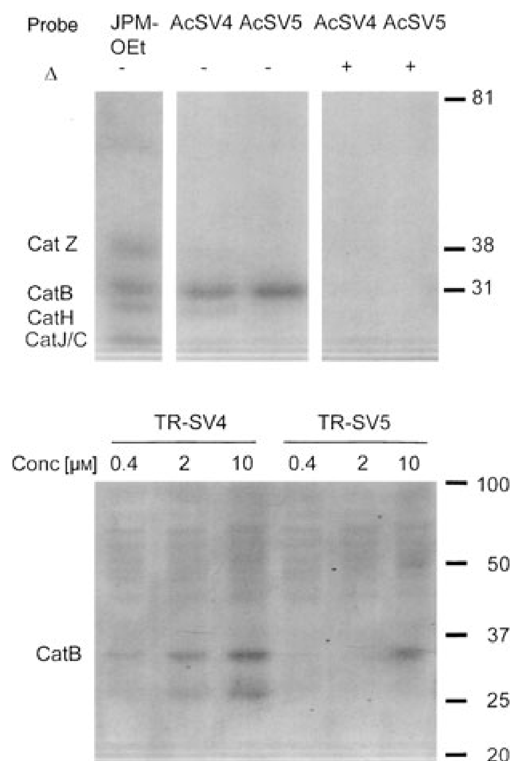


Figure 1. *Top: direct labeling of rat-liver homogenates with radiolabeled Ac-SV4 and Ac-SV5. Bottom: direct labeling with fluorophore-labeled TR-SV4 and TR-SV5.*

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

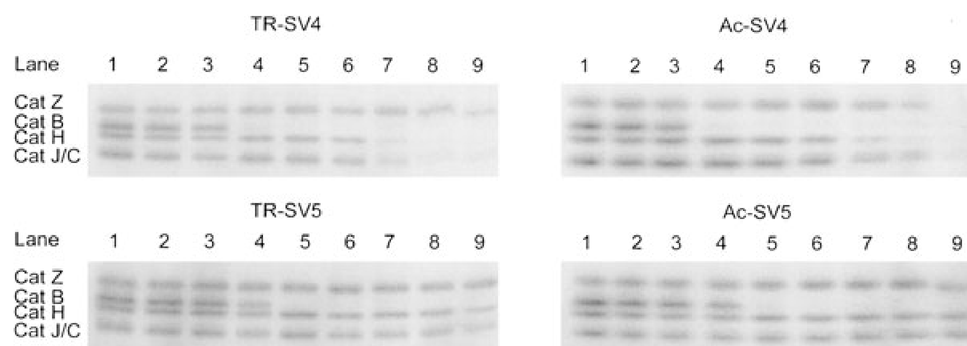


Figure 2. Competition experiments in rat-liver homogenates with increasing concentrations of probe (lanes 1 to 9: 0 nм, 0.6 nм, 3.2 nм, 16 nм, 80 nм, 400 nм, 2 μм, 10 μм and 50 μм, respectively). Radiolabeled general cysteine protease ABP JPM-OEt was added to monitor remaining cysteine protease activity.

To quantify the potency of the SV4 and SV5 probes, competition experiments were performed with the general cysteine protease ABP JPM-OEt.[12] After addition of the double-headed ABPs over a range of concentrations, radiolabeled JPM-OEt was added to monitor the remaining activity of the cathepsins in the rat-liver homogenates. Both the SV4 and SV5 compounds have a defined preference for cathepsin B.

The competition data from Figure 2 were quantified by using image-analysis software and used to determine the concentrations at which 50% of the enzyme activity was inhibited (apparent IC_{50} values; see Table 1). The SV4 probes have IC_{50} values for cathepsin B of ≈7–8 nм; this makes them slightly more potent than the SV5 probes, which showed IC_{50} values of 15–22 nм. However, the loss in potency is accompanied by a tenfold increase in selectivity for cathepsin B relative to the SV4 series of probes. Thus, the SV5 probes were more than 2300 times more reactive towards cathepsin B than the other predominant cathepsins.

Table 1. Apparent IC_{50} values of ABPs for different cathepsins in rat-liver homogenates.

ABP	IC_{50} [nм] Cat B	Cat Z	Cat H	Cat J/C	Cat B selectivity
Ac-SV4	8	15×10^3	1.2×10^3	4.0×10^3	1.5×10^2
TR-SV4	6.8	11×10^3	1.1×10^3	1.5×10^3	1.6×10^2
Ac-SV5	22	$>50\times10^3$	$>50\times10^3$	$>50\times10^3$	$>2.3\times10^3$
TR-SV5	15	$>50\times10^3$	$>50\times10^3$	$>50\times10^3$	$>3.3\times10^3$

The selectivity of the majority of the previously published cathepsin B-specific inhibitors is derived from distinct interactions with the unique occluding loop on cathepsin B. Crystal structures of enzyme–inhibitor complexes have shown that the isoleucine–proline dipeptide portion of compounds like CA-074 binds such that the free carboxylate projects into the so-called S′ region of the active site. For the dipeptidyl peptidase cathepsin B, a loop structure protrudes into this region and makes direct contacts with the free carboxylate through hydrogen bonds to two conserved histidine residues.[13] While these contacts are what drive the primary specificity for cathepsin B, the interaction of a leucine residue in the hydrophobic P2 pocket is optimal for virtually all of the papain-fold cysteine proteases. Thus, the higher selectivity of the SV5 probes is most likely due to the loss of the hydrophobic leucine residue, thereby eliminating the general high affinity interactions with the P2 pockets of other cathepsins. This hypothesis also explains the slight reduction in overall inhibitor potency observed for the SV5 probes. This general paradigm suggests that by optimization of both the P2 binding element and the P′ binding elements it should be possible to generate higher selectivity of probes for individual members of this protease family.

In summary, we describe here a solid-phase method for the synthesis of double-headed epoxide inhibitors of cysteine proteases. Key features of this method include the on-resin hydrolysis of the epoxysuccinyl ethyl ester and subsequent coupling of diversity elements.

Using this method, we show the efficient synthesis of two different classes of activity-based probes carrying either a radioactive or a fluorescent tag. Studies of these probes in complex proteomes show that removal of the hydrophobic P2 residue increases selectivity for cathepsin B, resulting in a novel, highly selective cathepsin B label. Furthermore, this method will facilitate the synthesis of additional probe families with a range of diversity elements on both sides of the reactive moiety that are likely to yield additional protease-specific reagents.

Experimental Section

Synthesis of probes: After Rink amide Novagel (NovaBiochem) had been loaded with the first Fmoc-protected amino acid, elongation took place by standard solid-phase peptide chemistry by using 20% piperidine in DMF to cleave Fmoc-protecting groups and a combination of 1,3-diisopropylcarbodiimide (DIC; 3 equiv) and HOBt (3 equiv) to condensate each amino acid (3 equiv). After final Fmoc-deprotection, nitrophenyl ester **4** (3 equiv) in DMF was added to the resin, and the mixture was allowed to react for 1 h to cap the terminal amine functionality. Next, the ethyl ester was saponified with KOH (0.25 м) in EtOH for approximately 20 min, and resin was subsequently washed with 1% AcOH in EtOH, then with EtOH and with dichloromethane. Finally, H-Ile-Pro-O*t*Bu (3 equiv) was coupled to the free carboxylic acid under the influence of PyBOP (3 equiv) and *N*,*N*-diisopropylethylamine (DIEA; 6 equiv) in

DMF. Deprotection of the probes and concomitant cleavage from the solid support was effected with TFA/H_2O/TIS (95:2.5:2.5; TIS = triisopropylsilane). Probes were precipitated with ether, collected by centrifugation, and purified by HPLC.

Evaluation in proteomes: Rat-liver homogenates (1 mg mL^{-1} total protein) were used in reaction buffer (50 μL) of pH 5.5 (50 mM sodium acetate, 2 mM dithiothreitol, 5 mM $MgCl_2$). Controls were preheated for 5 min at 90 °C. Samples were incubated for 0.5 h with either ^{125}I-radiolabeled (10^6 cpm) or fluorophore-conjugated probes, and analyzed by SDS-PAGE. For competition experiments, ABPs were added to rat-liver homogenates at the indicated concentrations and incubated at room temperature for 0.5 h. Subsequently, samples were treated with radiolabeled JPM-OEt (10^6 cpm) for an additional 30 min prior to subjection to gel electrophoresis. Data were quantified by using NIH ImageJ[14] and analyzed with GraphPad Prism (San Diego, CA).

Acknowledgements

This work was funded by a stipend from the Netherlands Organisation for Scientific Research (NWO; to S.V.) and a National Technology Centers for Networks and Pathways grant, NIH grant U54 RR020843 (to M.B.).

Keywords: activity-based probes · cysteine proteases · inhibitors · proteomics · solid-phase synthesis

[1] For example, see: a) M. Tyers, M. Mann, *Nature* **2003**, *422*, 193–197; b) G. L. G. Miklos, R. Maleszka, *Proteomics* **2001**, *1*, 169–178; c) S. D. Patterson, R. H. Aebersold, *Nat. Genet.* **2003**, *33*, 311–323.

[2] Recent reviews include: a) D. A. Campbell, A. K. Szardenings, *Curr. Opin. Chem. Biol.* **2003**, *7*, 296–303; b) J. W. Kozarich, *Curr. Opin. Chem. Biol.* **2003**, *7*, 78–83; c) D. A. Jeffery, M. Bogyo, *Curr. Opin. Biotechnol.* **2003**, *14*, 87–95; d) A. E. Speers, B. F. Cravatt, *ChemBioChem* **2004**, *5*, 41–47.

[3] For example, see: J. A. Joyce, A. Baruch, K. Chehade, N. Meyer-Morse, E. Giraudo, F.-Y. Tsai, D. C. Greenbaum, J. H. Hager, M. Bogyo, D. Hanahan, *Cancer Cell* **2004**, *5*, 443–453.

[4] D. C. Greenbaum, W. D. Arnold, F. Lu, L. Hayrapetian, A. Baruch, J. Krumrine, S. Toba, K. Chehade, D. Broemme, I. D. Kuntz, M. Bogyo, *Chem. Biol.* **2002**, *9*, 1085–1094.

[5] D. Greenbaum, A. Baruch, L. Hayrapetian, Z. Darula, A. Burlingame, K. F. Medzihradszky, M. Bogyo, *Mol. Cell. Proteomics* **2002**, *1*, 60–68.

[6] F. Lecaille, J. Kaleta, D. Bromme, *Chem. Rev.* **2002**, *102*, 4459–4488.

[7] D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569–581.

[8] I. Stern, N. Schaschke, L. Moroder, D. Turk, *Biochem. J.* **2004**, *381*, 511–517.

[9] a) M. Murata, S. Miyashita, C. Yokoo, M. Tamai, K. Hanada, K. Hatayama, T. Towatari, T. Nikawa, N. Katunuma, *FEBS Lett.* **1991**, *280*, 307–310; b) T. Towatari, T. Nikawa, M. Murata, C. Yokoo, M. Tamai, K. Hanada, N. Katunuma, *FEBS Lett.* **1991**, *280*, 311–315.

[10] For example, see: a) Z. Huang, E. B. McGowan, T. C. Detwiler, *J. Med. Chem.* **1992**, *35*, 2048–2054; b) B. J. Gour-Salin, P. Lachance, C. Plouffe, A. C. Storer, R. Ménard, *J. Med. Chem.* **1993**, *36*, 720–725; c) N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, D. Turk, L. Moroder, *Bioorg. Med. Chem.* **1997**, *5*, 1789–1797; d) N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, L. Moroder, *FEBS Lett.* **1998**, *421*, 80–82; e) N. Katunuma, E. Murata, H. Kakegawa, A. Matsui, H. Tsuzuki, H. Tsuge, D. Turk, V. Turk, M. Fukushima, Y. Tada, T. Asao, *FEBS Lett.* **1999**, *458*, 6–10; f) N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, T. Lassleben, C. P. Sommerhoff, L. Moroder, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 677–680; g) N. Schaschke, Assfalg-I. Machleidt, T. Lassleben, C. P. Sommerhoff, L. Moroder, W. Machleidt, *FEBS Lett.* **2000**, *482*, 91–96.

[11] a) M. Tamai, C. Yokoo, M. Murata, K. Oguma, K. Sota, E. Sato, Y. Kanaoka, *Chem. Pharm. Bull.* **1987**, *25*, 1098–1104.

[12] M. Bogyo, S. Verhelst, V. Bellingard-Dubouchaud, S. Toba, D. Greenbaum, *Chem. Biol.* **2000**, *7*, 27–38.

[13] D. Turk, M. Podobnik, T. Popovič, N. Katunuma, W. Bode, R. Huber, V. Turk, *Biochemistry* **1995**, *34*, 4791–4797.

[14] NIH ImageJ can be obtained as freeware at http://rsb.info.nih.gov

Received: October 22, 2004
Published online on ■■ ▮▮, 2005

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim