Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters
Tetrahedron Letters 46 (2005) 4053–4056

An affinity-based probe for the proteomic profiling of aspartic proteases

Souvik Chattopadhaya,[a] Elaine W. S. Chan[b] and Shao Q. Yao[a,b,*]

[a]*Department of Biological Sciences, National University of Singapore, 3 Science Drive 3, Singapore 117543, Singapore*
[b]*Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore 117543, Singapore*

Received 6 March 2005; revised 28 March 2005; accepted 5 April 2005
Available online 19 April 2005

Abstract—We have developed an affinity-based probe for the proteomic profiling of aspartic proteases. Our probe was shown to be selective towards aspartic proteases over other proteins. It was also shown that the strategy may be used to label selectively aspartic proteases in the presence of a large excess of other proteins, thus making it useful for future proteome profiling experiments.

Recent advances in genomics have opened up new prospects in the life sciences arena. However proteins, not genes, are ultimately responsible for the biological processes inside the cell. Given the extremely complex and diverse post-transcriptional and post-translational processes involved in the >30,000 human genes, it has been estimated that there may be between 100,000 and 1,000,000 functionally different proteins in the human proteome.[1] This presents enormous challenges to the field of proteomics, which aims to identify, characterize and assign biological functions of all proteins expressed by the genome. Technical limitations associated with two-dimensional gel electrophoresis (2D-GE), the traditional powerhouse for proteome profiling experiments, are numerous and remain to be adequately addressed.[2] One of the most pressing issues of 2D-GE is its inability to detect proteins on the basis of their biological activities. Recently, a novel chemical approach that complements the existing 2D-GE technologies has been reported.[3] This activity-based profiling approach uses active site-directed, small molecule probes that chemically react with certain classes of enzymes in a complex proteome, and can therefore report unique profiles of enzymes on the basis of their catalytic activities. The key component of an activity-based probe is its so-called reactive unit, usually made of a class-specific, mechanism-based inhibitor that reacts covalently and selectively with enzymes in accordance to their activity. Thus far, activity-based probes have been reported for enzyme classes that possess known mechanism-based suicide inhibitors, such as serine hydrolases, cysteine proteases and protein tyrosine phosphatases.[3,4] Alternative approaches have also been developed that target other classes of enzymes, also in an activity-dependent manner.[5]

The current activity-based profiling strategies are limited in that they are applicable only to a few classes of enzymes which possess covalent intermediates in their catalytic mechanism. In order to profile enzymes which do not possess covalent intermediates, affinity-based strategies have recently been reported.[6] Hagenstein et al. used benzophenone (BP)-tagged isoquinolinesulfonamides to profile kinases.[6a] We disclosed a similar strategy for the proteomic profiling of metalloproteases by taking advantage of small molecule probes containing a zinc-binding hydroxamate chelator and a photo-active diazirine moiety.[6b] Saghatalien et al. independently verified our approach by using BP-hydroxamates to profile matrix metalloproteases.[6c] In this report, we have extended our affinity-based strategy to the study of another subclass of enzymes, the aspartic proteases.

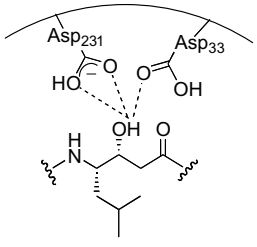
Aspartic proteases have been widely studied owing to their enormous ramifications in human diseases: renin is implicated in hypertension, cathepsin D in breast cancer metastasis, β-secretase in Alzheimer's disease, plasmepsin in malaria and HIV-1 protease in AIDS. Aspartic proteases are characterized by two catalytic aspartic acid residues located in their active sites and in most cases, by the conserved Asp-Thr-Gly (DTG)

Keywords: Affinity-based profiling; Activity-based profiling; Proteomics; Aspartic proteases; Enzymes; Photolabile probe.
* Corresponding author. Tel.: +65 68741683; fax: +65 67791691; e-mail: chmyaosq@nus.edu.sg

0040-4039/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2005.04.015

sequence in the primary structure.[7] Although the catalytic mechanism of aspartic proteases is poorly understood, it has been generally accepted that the aspartic residues in the enzyme active site bind a molecule of water through extensive H-bonding. The oxygen in water is activated and participates in the subsequent enzymatic reaction in which a tetrahedral intermediate is generated but is not covalently bound to the enzyme active site. In searching for an affinity-binding probe suitable for the proteomic profiling of aspartic proteases, we settled on pepstatin, a naturally occurring reversible inhibitor of pepsin (an aspartic protease), as well as a general inhibitor of many aspartic proteases. The hexapeptidyl pepstatin, Ival-Val-Val-Sta-Ala-Sta, comprises the central core unit, (3*S*,4*S*)-statine, which functions as an isosteric transition state analogue to replace the P_1–P_1' residues in the protease substrate (Scheme 1).[8] The hydroxyl group on statine binds tightly to the enzyme active site through H-bonding with the catalytic aspartic residues, thereby replacing the active site-bound water molecule.

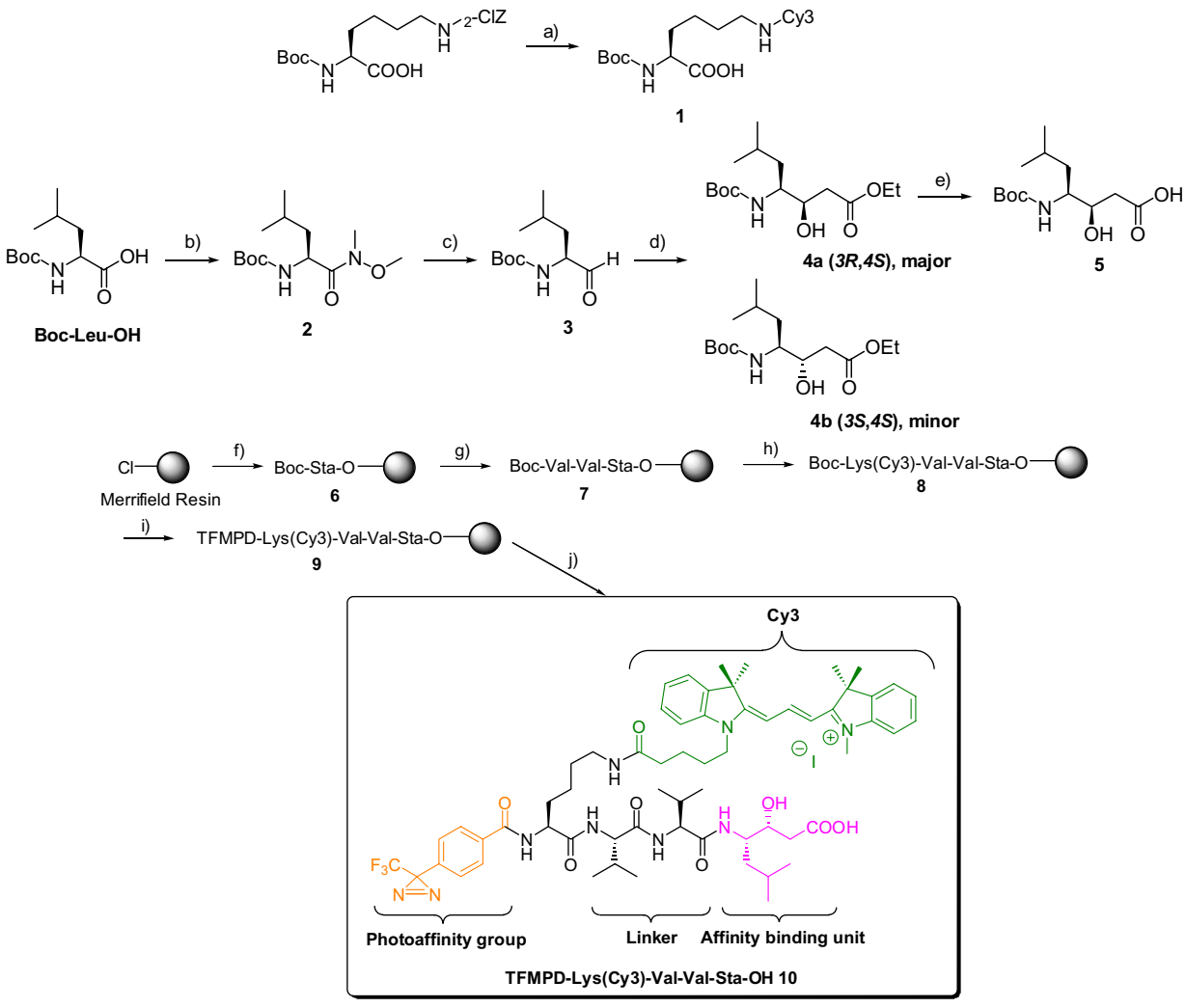
We designed our affinity-based probe (compound **10** in Scheme 2) to be broad-based in general by using a truncated analogue of pepstatin, Z-Val-Val-Sta, which was shown previously to retain reasonable activity against aspartic proteases (i.e., $K_i = 1.90 \times 10^{-7}$ M against pepsin).[9] Statine was retained in our probe to function as an affinity binding unit, while the two valine residues served as a linker and, to a lesser extent, retain some substrate recognition towards pepsin. (3*R*,4*S*)-Statine was selected in lieu of its naturally occurring diastereomer for easy chemical synthesis and also because it did not severely affect binding to pepsin.[9] A photolabile group, 3-trifluorophenylmethyl diazirine, was incorporated into the probe structure for covalent attachment to target enzymes. We previously showed that diazirine is ideal for affinity-based proteomic profiling experiments.[6b] Upon irradiation at 360 nm, the diazirine moiety undergoes homolytic C–N bond cleavage to generate a reactive carbene species that adds irreversibly across any C–H bonds in the enzyme active site. The inclusion of a fluorescent tag Cy3 in the probe through a lysine handle, resulted in the final trifunctional compound, **10**. To synthesize **10**, a solid-phase strategy was conceived, in anticipation that it may be applicable in future for the convenient synthesis of other statine-containing probes. Briefly (Scheme 2), Cy3 was first attached to the ϵ-amino group of Boc-lysine-OH, giving **1**. Boc-leucinal, **3**, was afforded from reduction of Boc-Leu-*N*,*O*-dimethylhydroxamate **2**. Boc-(3*R*,4*S*)-statine ethyl ester **4a** was subsequently synthesized, via an aldol reaction as previously reported,[10] and the major product isolated from its diastereomer by column chromatography. Hydrolysis of the ethyl ester with 20% potassium carbonate afforded Boc-protected statine, **5**. Subsequently, **5** was neutralized to its cesium salt with Cs_2CO_3 and loaded onto Merrifield resin. The following two valine residues were coupled to the statine-functionalized resin **6** using standard solid-phase peptide synthesis with Boc chemistry. The lysine handle containing the fluorophore, Boc-Lys(Cy3)-OH **1**, was subsequently coupled at the N-terminus of the tripeptide. Upon removal of the α-amino Boc protecting group with neat TFA, followed by coupling of the photolabile diazirine moiety, the final probe **10** was released from the resin by TFMSA-mediated cleavage.[11]



Scheme 1. Binding mode of statine to an aspartic protease.

Pepsin (EC 3.4.23.1) was chosen as the model aspartic protease for our enzyme labelling studies. It confers a high specificity towards large hydrophobic residues at the P_1 and P_1' positions (e.g., Phe and Leu). The active pH profile for pepsin ranges from pH 1–6, with pH optimum of 3.5.[12] Hence all labelling reactions were done at acidic pH to ensure optimal enzyme activity. Under reversible binding conditions (in the absence of UV cross-linking), the probe **10** was found to inhibit the activity of pepsin with an IC_{50} value of 2.26×10^{-7} M, comparable to that of its parent compound, the truncated analogue of pepstatin (1.90×10^{-7} M) and only 10-fold less than that of pepstatin A (0.238×10^{-7} M) (see Supplementary data). This indicates that incorporation of the diazirine or the Cy3 moieties did not significantly alter the binding property of the probe. We next tested the ability of the probe **10** to covalently modify pepsin and renin (also an aspartic protease) in an activity-dependent manner. As shown in Figure 1a, using only 500 nM of the probe, as few as 6 ng of pepsin and 10 ng of renin (corresponding to ~8 nM and 12 nM of the respective enzymes) could be detected following separation of samples by SDS/PAGE and in-gel fluorescence scanning. In the absence of UV photolysis, no labelling of the enzymes was observed (see Supplementary data). A variety of proteins, including aspartic proteases (i.e. cathepsin D, mucorpepsin and HIV-1 protease) and other enzymes, were also tested: only active aspartic proteases could be labelled by our probe, indicating that our strategy is indeed general and specific towards aspartic proteases.

We next assessed whether the labelling of aspartic proteases in our strategy is activity-based. No labelling was seen with heat-inactivated aspartic proteases (data not shown). Furthermore, competitive labelling studies were carried out with pepsin in the presence of its natural inhibitor, pepstatin A. Pepstatin A inhibits a variety of acidic aspartic proteases, including pepsin, renin and cathepsin D, by forming a 1:1 tight complex with the enzyme. As shown in Figure 1b, labelling of pepsin (50 ng) with the probe **10** (500 nM) was completely inhibited by 2 μM of pepstatin A (i.e., 4-fold excess of the probe). In a separate experiment, it was revealed that the same amount of pepsin (50 ng) when labelled with varied concentrations of **10**, reached saturation of labelling with 2 μM of the probe (see Supplementary data), indicating



Scheme 2. Synthesis of TFMPD-Lys-Val-Val-Sta-OH probe **10**. Reagents and conditions: (a) (i) H_2, Pd–C (cat), AcOH; (ii) Cy3-NHS, DIEA, DMF; (b) *N*,*O*-dimethylhydroxylamine hydrochloride, DCC, HOBt, DIEA, DMF; (c) LAH, THF, 0 °C; (d) ethyl acetate, LDA, THF, −78 °C; (e) 20% K_2CO_3, MeOH/H_2O (2:1); (f) (i) **5**, 2 M Cs_2CO_3, EtOH/H_2O (4:1), pH 7; (ii) KI (cat), DMF, 50 °C; (g) (i) TFA; (ii) Boc-Val-OH, HBTU, HOBt, DIEA, DMF; (h) (i) TFA; (ii) **1**, HBTU, HOBt, DIEA, DMF; (i) (i) TFA; (ii) TFMPD, HBTU, HOBt, DIEA, DMF; (j) TFA, TFMSA, thioanisole/EDT (2:1).

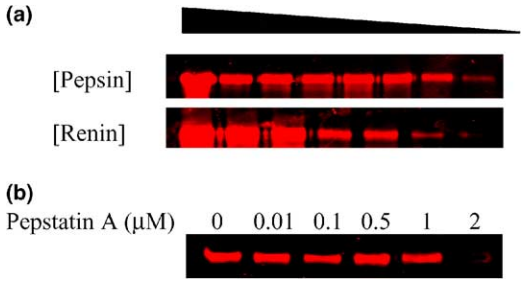


Figure 1. (a) Affinity-based labelling of pepsin and renin. Decreasing amounts of pepsin (L to R: 500, 200, 100, 50, 25, 12 and 6 ng) and renin (L to R: 200, 100, 50, 30, 20, 10 and 5 ng) were labelled with **10** (500 nM). (b) Competitive labelling of pepsin (50 ng) with **10** (500 nM) in the presence of pepstatin A. Labelling was completely inhibited by 2 μM of pepstatin A (4-fold excess over **10**). See Supplementary data for details.

our probe probably binds to the active site of pepsin in a manner similar to pepstatin A. Taken together, these data indicate that probe **10** labels a variety of aspartic proteases specifically, and the labelling, although driven by affinity binding of the probe, is intrinsically dependent on the activity/mechanism of the enzymes, and thus is activity-based.

In order to test whether **10** could be used to profile aspartic proteases in future proteomic experiments, we carried out the labelling of pepsin in the presence of a large excess of other proteins (Fig. 2). We added aliquots of 50 ng of pepsin, with and without pepstatin A (10 μM), to increasing amounts of yeast total cellular lysates (0, 50, 100, 150 and 200 μg). Following incubation with 2.5 μM of the probe and photo cross-linking, the mixtures were resolved by SDS-PAGE and analyzed by in-gel fluorescence and Coomassie staining (see Supplementary data). The result showed that, with the yeast lysate alone (e.g., no pepsin), some fluorescent bands were visible which were unaffected by the addition of pepstatin A, indicating these bands may have arisen from either the nonspecific labelling of endogenous yeast proteins, or endogenous yeast aspartic proteases whose

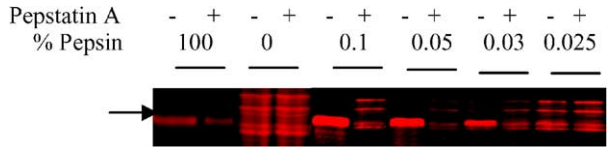


Figure 2. Affinity-based labelling of pepsin present in the yeast proteome. Lanes 1 and 2: pure pepsin (50 ng). Lanes 3–4: yeast lysate only (50 μg). Lanes 5–12: pepsin (50 ng) was added to increasing amounts of yeast total cellular lysate (50, 100, 150 and 200 μg), with and without pepstatin A (10 μM). The mixture was then labelled with probe **10** (2.5 μM) following in-gel fluorescence scanning and Coomassie staining (see Supplementary data). As shown in lanes 9 and 10, as little as 50 ng active pepsin could be detected in a background of 150 μg yeast proteins (corresponding to 0.03% of the total proteins) and this labelling was blocked by pepstatin A (10 μM). Arrow indicates the labelled pepsin.

activity is not inhibited by pepstatin A. However, what is more important is that, the probe was able to detect as little as 0.03% of pepsin present in the yeast crude lysate, which is comparable to the level at which proteases are normally expressed endogenously in a complex proteome.[6b,c] The labelling of pepsin in the yeast lysate was also inhibited by pepstatin A, further confirming the highly specific and activity-based feature of our probe. Our strategy thus lays the framework for the eventual large-scale functional profiling of aspartic proteases in a complex proteome.

In summary, we have successfully designed and synthesized an affinity-based probe which may be useful for the proteomic profiling of aspartic proteases. A solid-phase strategy was developed for the convenient synthesis of this probe, and in future, other analogous probes. We have established optimal conditions for the selective labelling of aspartic proteases over other proteins. Competitive inhibition studies clearly demonstrated the affinity-based strategy is a good complement to existing activity-based profiling approaches,[3] in that it is also suitable to indirectly profile specific subsets of enzymes on the basis of their activities. Equally important, we have shown that the strategy may be used to selectively label aspartic proteases in the presence of a large excess of other proteins, thus making it useful for future pro- teome profiling experiments.

Acknowledgments

Funding support was provided by the National University of Singapore (NUS) and the Agency for Science, Technology and Research (A*STAR) of Singapore. S.Q.Y. is the recipient of the 2002 BMRC Young Investigator Award from the Biomedical Research Council of Singapore.

Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.tetlet.2005.04.015.

References and notes

1. Tyers, M.; Mann, M. *Nature* **2003**, *422*, 193–197.
2. Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.
3. Jessani, N.; Cravatt, B. F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 54–59.
4. (a) Liu, Y. S.; Patricelli, M. P.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14694–14699; (b) Nazif, T.; Bogyo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2967–2972; (c) Wang, G.; Uttamchandani, M.; Chen, G. Y. J.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 737–740; (d) Lo, L. C.; Pang, T. L.; Kuo, C. H.; Chiang, Y. L.; Wang, H. Y.; Lin, J. J. *J. Proteome Res.* **2002**, *1*, 35–40.
5. (a) Adam, G. C.; Sorensen, E. J.; Cravatt, B. F. *Nat. Biotechnol.* **2002**, *20*, 805–809; (b) Zhu, Q.; Girish, A.; Chattopadhaya, S.; Yao, S. Q. *Chem. Commun.* **2004**, 1512–1513.
6. (a) Hagenstein, M. C.; Mussgnug, J. H.; Lotte, K.; Plessow, R.; Brockhinke, A.; Kruse, O.; Sewald, N. *Angew. Chem. Int. Ed.* **2003**, *42*, 5635–5638; (b) Chan, E. W. S.; Chattopadhaya, S.; Panicker, R. C.; Huang, X.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 14435–14446; (c) Saghatalien, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10000–10005.
7. Barrett, A. J.; Rawlings, N. D.; Woessner, J. F. *Handbook of Proteolytic Enzymes*; Academic: San Diego, 1998.
8. Marciniszyn, J.; Huang, J. S.; Hartsuck, J. A.; Tang, J. *J. Biol. Chem.* **1976**, *251*, 7095–7102.
9. McConnell, R. M.; Frizzell, D.; Camp, A.; Evans, A.; Jones, W.; Cagle, C. *J. Med. Chem.* **1991**, *34*, 2298–2300.
10. Rich, D. H.; Sun, E. T.; Boparai, A. S. *J. Org. Chem.* **1978**, *43*, 3624–3626.
11. Synthesis and characterization of probe **10**. The Boc protecting group on **8** was removed using neat TFA (10 mL/g resin) for 1 h. The resin was then collected by filtration and washed extensively with DMF, DMF/DIEA (1:1), DMF, DCM and MeOH and dried in vacuo. 4-(3-Trifluoromethyl-3*H*-diazirin-3-yl)-benzoic acid (4 equiv), HBTU (4 equiv) and HOBt (4 equiv) were dissolved in a minimal amount of DMF. DIEA (8 equiv) was added and the reaction mixture was agitated for 10 min. The pre-activated solution was then added to the deprotected resin and the reaction was allowed to proceed for 4 h. The resin was then collected by filtration and washed with DMF, DCM and MeOH and dried in vacuo to afford resin-bound TFMPD-Lys(Cy3)-Val-Val-Sta-OH **9**. The final product was cleaved from the solid support using standard TFMSA cleavage protocol. To the resin **9** in a rbf was added thioanisole/EDT (2:1) 150 μL and the mixture was cooled to 0 °C. TFA (1 mL) and TFMSA (0.1 mL) were added and the cleavage reaction was allowed to proceed for a further 2 h. The filtered solution was subsequently collected and concentrated in vacuo. The crude mixture was subjected to RP-HPLC purification and the pure product **10** was afforded as a red solid (0.4 mg; 25% yield): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.54 (dd, J = 31.2 Hz, J = 12.5 Hz, 1H), 7.94–7.91 (m, 2H), 7.55–7.27 (m, 10H), 6.43 (d, J = 12.9 Hz, 2H), 4.55–4.51 (m, 1H), 4.21–4.06 (m, 3H), 3.68 (s, 3H), 3.54 (m, 2H), 3.18 (m, 2H), 3.07 (m, 1H), 2.66 (m, 2H), 2.27 (m, 2H), 2.01–1.22 (m, 27H) including 1.76 (s, 12H), 0.93–0.85 (m, 18H); ^{19}F NMR (282 MHz, CD_3OD) δ −0.76; ESI-MS calcd for $C_{62}H_{83}F_3N_9O_8$ $[M-I]^+$ 1138.6. Found 1138.4.
12. Lin, Y.; Fusek, M.; Lin, X.; Hartsuck, J. A.; Kezdy, F. J.; Tang, J. *J. Biol. Chem.* **1992**, *267*, 18413–18418.