



Specific enrichment of nonribosomal peptide synthetase module by an affinity probe for adenylation domains



Fumihiko Ishikawa^{*}, Hideaki Kakeya^{*}

Department of System Chemotherapy and Molecular Sciences, Division of Bioinformatics and Chemical Genomics, Graduate School of Pharmaceutical Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:

Received 21 October 2013

Revised 15 December 2013

Accepted 19 December 2013

Available online 25 December 2013

Keywords:

Nonribosomal peptide synthetase

Polyketide synthase

Adenylation domain

Affinity probe

Gramicidin S

ABSTRACT

We targeted the development of an affinity probe for adenylation (A) domains that can facilitate enrichment, identification, and quantification of A domain-containing modules in nonribosomal peptide synthetase (NRPS)–polyketide synthase (PKS) hybrids and NRPSs. A 5'-O-sulfamoyladenine (AMS) non-hydrolyzable analogue of adenosine monophosphate (AMP) has been reported as a scaffold for the design of inhibitors exhibiting tight binding of adenylation enzymes. Here we describe the application of an affinity probe for A domains. Our synthetic probe, a biotinylated L-Phe-AMS (L-Phe-AMS-biotin) specifically targets the A domains in NRPS modules that activates L-Phe to an aminoacyladenylate intermediate in both recombinant NRPS enzyme systems and whole proteomes.

© 2014 Elsevier Ltd. All rights reserved.

A number of small molecule natural products with significant biological activity such as antimicrobial, anticancer, and immunosuppressant activities belong to a large class of natural products known as polyketide (PK)–nonribosomal peptide (NRP) hybrid molecules and NRPSs.¹ Originating from bacteria and fungi, these peptide-containing bioactive natural products consist not only of the 20 proteinogenic amino acids, but also non-proteinogenic amino acids, aryl acids, and other acids, thus generating highly complex chemical diversity.² The biosynthesis of these small molecules is performed by highly versatile and large multimodular enzymes called nonribosomal peptide synthetase (NRPS)–polyketide synthase (PKS) hybrids and NRPSs.^{3–5} There has been much progress in the understanding biochemical programming and molecular basis of adenylation (A) domain substrate specificity in NRPS and NRPS–PKS hybrid systems. This has facilitated the prediction of structural features of newly discovered NRP and NRP–PK natural products assembled by these systems uncovered by genomic information. Large portions of these compounds often readily correlate to the amino acid specificity of A domains found on their associated modular enzymes.^{6–8} Genome sequencing has revealed that secondary metabolite gene clusters encoding these enzymes are widely dispersed and largely uncharacterized.⁹ In addition, complicated organisms containing symbiotic bacteria are particularly resistant to most genetic methods that rely on

pure, culturable strains. Direct detection of biosynthetic enzymes from bacterial proteomes complements genetic approaches in understanding the activity and dynamics of these enzymes in their native proteomes. By taking advantage of the A domains of NRPS and PKS–NRPS hybrid systems, this work aims to specifically enrich A domain-containing modules from bacterial proteomes with sequenced genomes using small-molecule probes and directly link the chemotypes of expressed peptide-containing natural products to their biosynthetic enzymes. Chemical probes for A domains using amino, aryl, or other acid building blocks that are found in NRP and NRP–PK hybrid molecules would facilitate probe-guided selective enrichment and identification of A domain-containing modules from proteomic samples by LC–MS/MS analysis (Fig. 1). In sequenced producers, such approaches would have applications in monitoring the expression dynamics of NRPS modules and optimizing bacterial culture conditions. In unsequenced organisms, such studies should facilitate discovery of the expressed NRPS/PKS–NRPS gene clusters.^{10,11}

The affinity probe design for the A domains is based on the reaction mechanism for amino acid loading, which is catalyzed by the A domains as depicted in Figure 2a. The A domain recognizes a specific amino, acyl, or another acid from the cellular pool, and catalyzes the formation of a tightly bound acyladenylate intermediate. This in turn transfers the acyl group onto the thiol group of a 4'-phosphopantethein present on a downstream carrier protein (CP) of NRPS assembly line.^{12,13} A 5'-O-sulfamoyladenine (AMS), a non-hydrolyzable analogue of adenosine monophosphate (AMP), has been applied widely in the design of inhibitors that

^{*} Tel.: +81 75 753 4524; fax: +81 75 753 4591.

E-mail addresses: fishika@pharm.kyoto-u.ac.jp (F. Ishikawa), scseigyo-hisyo@pharm.kyoto-u.ac.jp (H. Kakeya).

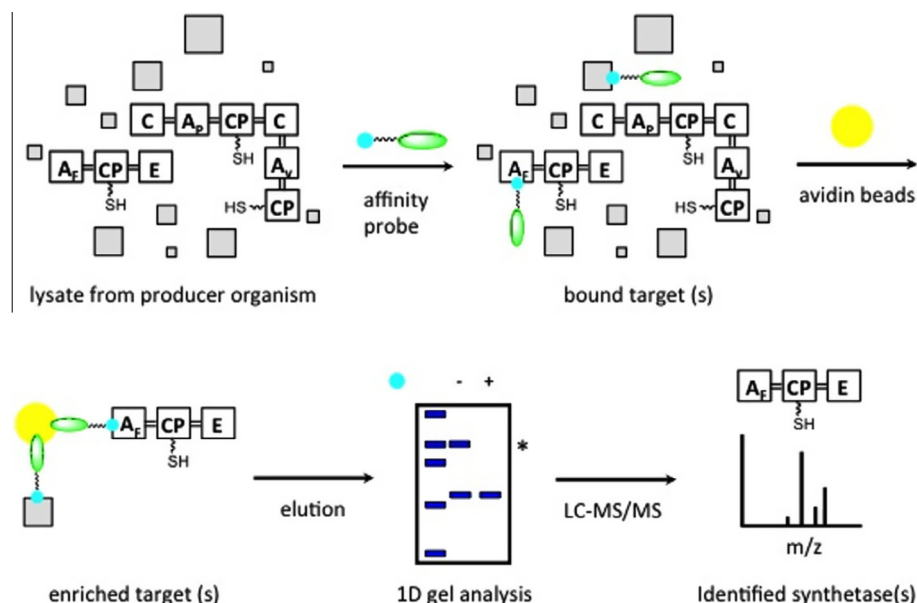


Figure 1. Methods for proteomic analysis of NRPSs using an affinity probe. Modules comprise carrier protein (CP), adenylation (A) [A_F : L-Phe specific; A_P : L-Pro specific; A_V : L-Val specific], epimerase (E), and condensation (C) domains.

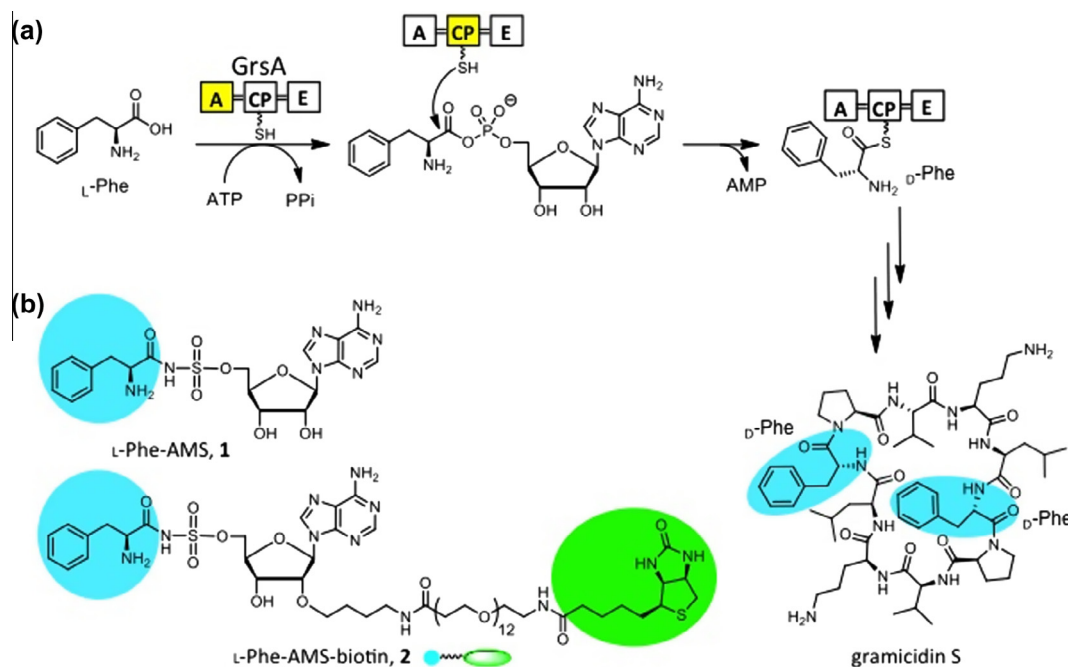


Figure 2. (a) Biosynthesis of the gramicidin S. (b) Structures of L-Phe-AMS **1** and L-Phe-AMS-biotin **2**.

display tight binding of adenylation enzymes and functionally related aminoacyl tRNA synthetases.^{14–16} A 5'-O-[N-(phenylalanyl)-sulfamoyl] adenosine (L-Phe-AMS, **1**, Fig. 2b) has been developed by Marahiel and coworkers wherein the reactive acylphosphate linkage has been replaced by a bioisosteric and non-hydrolyzable acylsulfamate group.¹⁵ Recently, an acylsulfamate-based probe installed a photoreactive group has been reported for profiling of the stand-alone aryl-acid adenylation enzyme, MbtA.¹⁷ A more complete set of probes would be crucial, however, for full module identification. Here, we describe the design, synthesis, and characterization of a biotinylated variant of L-Phe-AMS, referred to as L-Phe-AMS-biotin and highlight its utility as an agent for the

Table 1

The apparent K_i values of **1** and **2** for A domain of the tridomain GrsA

Inhibitors	K_i^{app} (nM)
1	54.9 ± 1.2
2	339 ± 26

* The assays were performed in the presence of 0.0025% Igopal CA-630.

specific enrichment, identification, and quantification of A domain-containing modules.

To expand the acylsulfamate scaffold to identify A domains of modular NRPSs and NRPS-PKS hybrids from proteomic samples,

our studies began with the preparation of a suitably tagged A domain affinity probe (Fig. 2b). We chose a tridomain GrsA as an

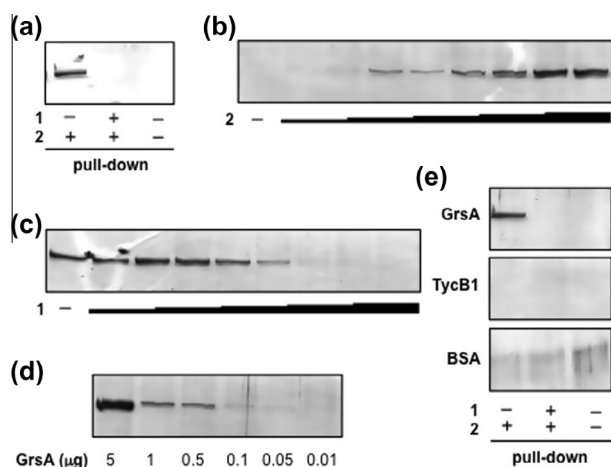


Figure 3. In vitro characterization of L-Phe-AMS-biotin **2** with purified proteins. (a) Pull-down of GrsA and competition with excess **1**. GrsA (7.9 nM) was preincubated in either the absence or presence of 10 μM of **1** and treated with 1 μM of **2**. (b) Concentration dependence of **2**. GrsA (7.9 nM) was incubated with 1 nM–10 μM of **2**. (c) Competitive profiling with GrsA. GrsA (7.9 nM) was incubated with 4.25 μM of **2** in the presence of 0.1 nM–10 μM of **1**. (d) Limit of detection of purified GrsA. GrsA (5–0.01 μg/mL) was incubated with 4.25 μM of **2**. (e) Binding specificity of **2**. GrsA (1 μg/mL), TycB1 (1 μg/mL), and BSA (1 μg/mL) were treated with 1 μM of **2** in either the absence or presence of **1**. Full gels (Fig. S9) and experimental procedures are provided in the SI.

initial proof of concept target protein (Fig. 2a). The probe design was based on the tight-binding bisubstrate inhibitor, L-Phe-AMS **1**. A long PEG spacer arm was selected to minimize steric hindrance between modular enzymes and avidin molecules and permit isolation of low-abundance proteins.¹⁸ To this end, an affinity probe, biotinylated L-Phe-AMS (L-Phe-AMS-biotin, **2**) was synthesized from commercially available adenosine, 4-dibromobutane, Boc-Phe-OSu, and NHS-PEG12-Biotin (Scheme S2). In addition, L-Phe-AMS **1** was prepared according to the literature procedure to provide an analogue of **2** for comparative activity studies (Scheme S1).^{14–16,19,20}

Our biochemical studies began with an examination of the specific activities of these probes against recombinant GrsA derived from gramicidin S biosynthetic enzymes as a model NRPS module. GrsA is a single module peptide synthetase containing the domain structure A (L-Phe)-CP-E and is responsible for the incorporation of D-Phe during gramicidin S biosynthesis (Fig. 2a).^{21,22} For comparative studies, we chose the NRPS module, TycB1. Derived from the tyrocidine biosynthetic pathway, it incorporates L-Pro and contains the domain structure C-A (L-Pro)-CP.⁶ Both recombinant *holo*-GrsA and *holo*-TycB1 were prepared as a C-terminal His₆-tagged fusion proteins according to the literature procedure (Fig. S1).²³ *K_i* values were determined by a coupled hydroxamate-MesG continuous spectrophotometric assay (Fig. S2).²⁴ Probes **1** and **2** displayed tight-binding inhibition against GrsA-catalyzed formation of the aminoacyladenylate, with calculated *K_i^{app}* values of 54.9 ± 1.2 and 339 ± 26 nM, respectively (Table 1 and Fig. S3). Conversely, the TycB1-catalyzed reaction showed no detectable inhibition by **1** (Fig. S4). These experiments here suggest that attachment of

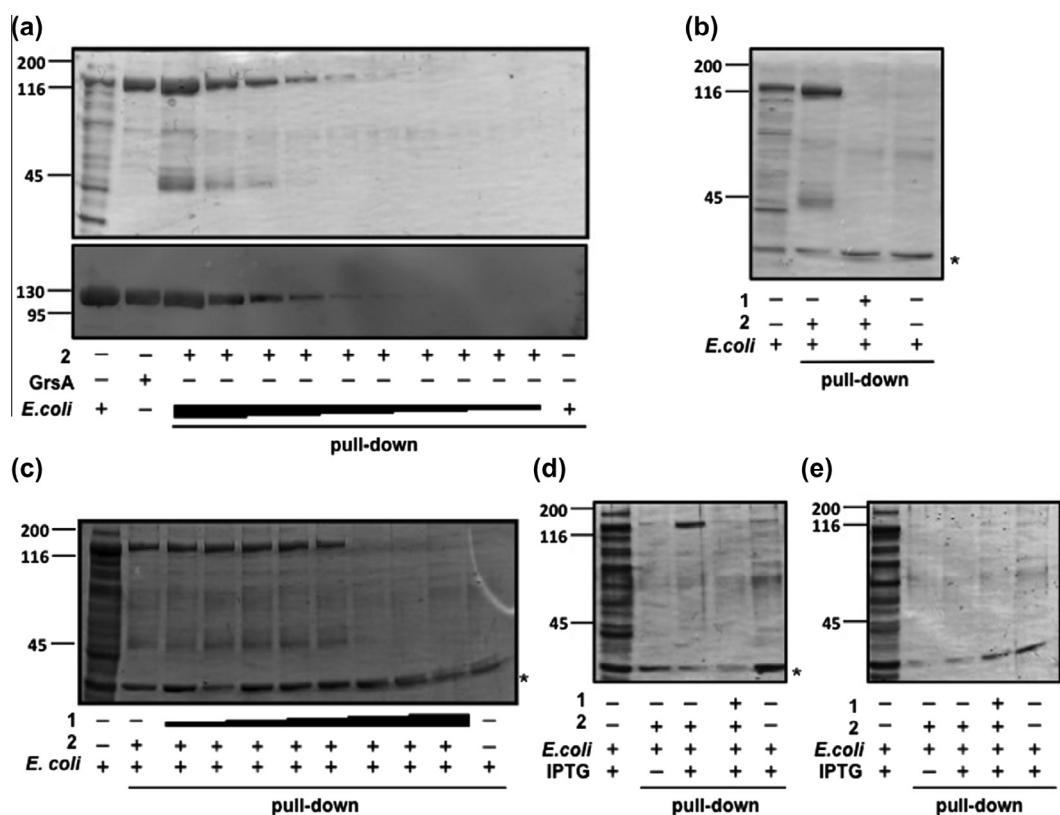


Figure 4. Pull-down assays from *E. coli* cellular lysates. (a) Pull-down of overexpressed GrsA. *E. coli* lysates (1 mg/mL–31.3 μg/mL) were treated with 4.25 μM of **2**. Upper: CBB staining. Lower: Anti-His blot. (b) Competition with excess **1**. *E. coli* lysates (250 μg/mL) with overexpressed GrsA were incubated with 4.25 μM of **2** in either the absence or presence of 100 μM **1**. (c) Concentration dependent competition by **1**. *E. coli* lysates (250 μg/mL) with overexpressed GrsA were preincubated with **1** (0–100 μM) before addition of **2** (4.25 μM). (d, e) Binding specificity of probe **2**. The lysates (250 μg/mL) prepared in the presence or absence of IPTG were incubated with 4.25 μM of **2** in either the absence or presence of 100 μM **1**. *E. coli* was transformed with (d) pQE60-grsA and (e) pQE60-tycB1 plasmids. (*) depicts a non-specifically bound protein. Full gels (Fig. S10) and experimental procedures are provided in the SI.

the pegylated long spacer at the 2'-oxygen atom of **2** preserves the high affinity binding of the L-Phe-activating A domain of GrsA. In addition, no-cross-reactivity was observed between the NRPS modules. This indicates that an L-Phe-AMS warhead is capable of selectively targeting L-Phe-activating A domains.

To evaluate the ability of probe **2** to pull-down recombinant GrsA, GrsA was first incubated with probe **2** for 12 h at 4 °C at pH 8 in the absence of Igepal CA-630 as a detergent. This was accompanied by a weak GrsA band on the SDS-PAGE gel indicating precipitated GrsA (Fig. S5). In contrast, pull-down assays in the presence of the detergent affected probe function by preventing aggregation effects. As shown in Figure 3a, a well-defined band indicating precipitated protein was observed. When excess **1** was incubated with GrsA for 1 h prior to addition of probe **2**, no precipitation band was detected. L-Phe-AMS **1** competitively inhibited the binding of probe **2**. This indicates that the binding of probe **2** occurs at the active site of A domain of GrsA. To investigate the concentration dependence of probe **2**, varying concentrations of **2** were incubated with 1 µg of GrsA. As shown in Figure 3b, GrsA displayed an escalation in precipitation as probe concentration was increased. Additionally, a dose-response curve was created by varying concentrations of **1** with fixed concentrations of GrsA and probe **2**. The precipitated band intensities were plotted against

the concentrations of **1** to calculate an IC₅₀ value of 11.0 ± 4.2 nM for **1** (Figs. 3c and S6). Quantitative analyses with a fixed amount of probe **2** and streptavidin-agarose resin indicated that the limit of detection was estimated above 100 ng of GrsA (Fig. 3d). Finally, to ensure specific pull-down by probe **2**, GrsA, TycB1, and BSA were incubated with 4.25 µM **2**. This resulted in preferential precipitation of GrsA with virtually no detectable precipitation of TycB1 and BSA (Fig. 3e). This result, in addition to knowledge of the differential amino acid specificity of A domains, strongly suggests selective targeting of the A domain is guided by the amino acid moiety attached to the probe molecule.

We next asked whether probe **2** could be used to probe endogenous A domains in proteome. We began with an analysis of the conditions of probe **2** incubated with overexpressed *holo*-GrsA and *holo*-TycB1 in whole *Escherichia coli* cellular lysates. *E. coli* cells were grown either in the presence or absence of IPTG (Fig. S7). The lysates containing overexpressed GrsA were incubated with 4.25 µM **2** for 12 h at 4 °C at pH 8. Bound proteins were isolated by purification with streptavidin-agarose resin. Non-specifically bound proteins were removed by washing with ice-cold RIPA buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, and 0.1% SDS). SDS-PAGE of the eluent revealed a single band of GrsA at 120 kDa with CBB staining and Western

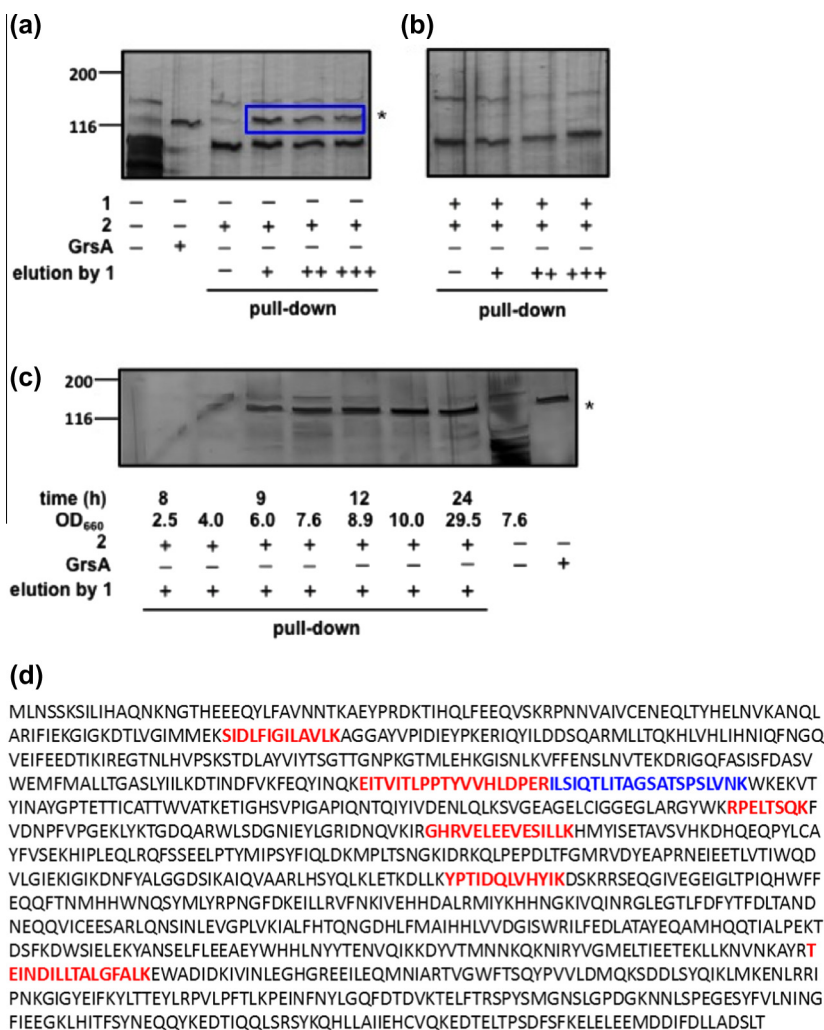


Figure 5. Pull-down assay of endogenous GrsA from *Aneurinibacillus migulanus* ATCC 9999 cellular lysate. (a) *A. migulanus* lysate (1.2 mg/mL) was treated with 9.5 µM **2**. Bound proteins were eluted by treatment with **1** (0, 0.5, 1, and 5 mM). (b) The lysate (1.2 mg/mL) was incubated with 9.5 µM **2** in the presence of 100 µM **1**. Bound proteins were eluted as described above. (c) Dynamics of expression of GrsA. The lysates at each cell concentration (OD₆₆₀ = 2.5–29.5) were incubated with 9.5 µM **2**. Precipitated proteins were eluted by treatment with 0.5 mM **1**. (d) LC-MS/MS analysis identified the isolated band (blue box) as GrsA. Amino acids of the identified peptides are colored either in red or blue. Full gels (Fig. S11) and experimental procedures are provided in the SI.

blotting using an anti 6 × His antibody (Fig. 4a). Pull-down of GrsA was abrogated by treatment of the lysates with vehicle (DMSO) (Fig. 4a). Furthermore, preincubation of the lysates with excess **1** resulted in an absence of GrsA precipitation (Fig. 4b). In contrast, the observed ~35 kDa band (*) is nonspecific, as L-Phe-AMS **1** does not have any effect on the intensity (Fig. 4b). Quantitative competitive assays showed reduced GrsA band intensities with increasing concentrations of **1** (Fig. 4c). Finally, to ensure specific pull-down of GrsA by probe **2** in the proteomic environment, incubation of the lysates with overexpressed GrsA or TycB1 with 4.25 μM **2** were carried out. This resulted in preferential pull-down of GrsA with no detectable pull-down of TycB1 (Fig. 4d and e). In the absence of IPTG, GrsA is not expressed, as evidenced by the lack of precipitated GrsA. Remarkably, the specific pull-down by probe **2** can be completely inhibited by preincubation with excess **1**. This is indicative of the ability of probe **2** to discriminate between different active-site structures of adenylation enzymes in crude cellular lysates. This finding corresponds identically to protein precipitated from recombinant GrsA, TycB1, and BSA by probe **2** (Fig. 3e). Taken together, these experiments indicate that **2** is a true affinity probe that shows no-cross-reactivity with other enzymes. Additionally, the extremely high specificity of **2** shows that the sulfamoyl scaffold coupled with a pegylated long spacer unit is an optimal design for applications involving specific enrichment, identification, and quantification of endogenous NRPS modules from crude cellular lysates.

As an ultimate test of proteomic activity, we evaluated the availability of probe **2** in a proteomic context by applying it to a native system. To isolate the cellular target of probe **2**, the lysate from the gramicidin S producer, *Aneurinibacillus migulanus* ATCC 9999 was incubated with probe **2**, and the precipitated proteins were isolated by purification with a streptavidin-immobilized resin. Proteins were then eluted by treatment with 20 mM Tris–HCl buffer (pH 8) containing 0, 0.5, 1, and 5 mM of **1** (Fig. 5a). A band (*) at approximately 120 kDa appeared after eluting the resin with media containing **1** as a vehicle to release bound proteins (Fig. 5a). More significantly, this band, affinity-purified by probe **2** can be completely inhibited by preincubation with 100 μM **1** (Fig. 5b). A sample of the excised band was submitted for LC–MS/MS protein identification and found to contain the endogenous GrsA (Figs. 5d and S8). We further evaluated probe **2** as a quantification tool to gain an understanding of modular synthase expression at each time point (Fig. 5c). Each of the lysates was incubated with probe **2** for 12 h at 4 °C. Precipitated proteins were then eluted by treatment with 500 μM **1**. The evidence provided in Figure 5c and d indicates that endogenous GrsA production abruptly began within 9 h (OD₆₆₀ = 6.0) after inoculation. Interestingly, at 24 h the amount of GrsA in a proteomic context was on the same level as 12 h, as estimated by the relative band intensity. This indicated the slow disappearance of GrsA in the producer. Probe **2** allowed successful monitoring of the expression dynamics and function of GrsA. Accordingly, these results validate the use of probe **2** as a selective isolation, identification, and quantification tool for A domain-containing modules within a proteomic environment.

In summary, we have demonstrated proteomic applications for the study of adenylation enzymes in NRPS modules using an affinity probe. This probe is capable of selectively targeting A domains with substrate specificity that corresponds to an attached amino acid moiety. It is noteworthy that this probe can attain specific enrichment of the target even from a native cellular lysate. The

synthesis of this probe is straightforward, offering easy access to synthetic derivatives containing different amino acids or other acids for targeting A domains of interest. In addition, we have determined optimal conditions for the specific pull-down of A domain-containing modules and their detection using a competitive elution method and established the activity of this probe both in purified proteins and within proteomic samples. Combinations of this probe with other domain-specific reactive probes such as thioesterase^{11,25} and dehydratase²⁶ should further improve the enrichment and detection of biosynthetic enzymes in proteomic samples.

Acknowledgments

This work was supported by a Grant-in-Aid for Research Activity start-up 24810015 (F.I.) and a research grant from the Japan Society of the Promotion of Science (H.K.). We thank Prof. Mohamed Marahiel (Philipps-Universität Marburg, Germany) for providing the GrsA and TycB1 expression constructs.

Supplementary data

Supplementary data (supporting figures; procedures for the syntheses of **1** and **2**; complete gel images; full experimental details; and copies of ¹H and ¹³C NMR spectra) associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2013.12.082>.

References and notes

- Walsh, C. T. *Science* **2004**, *303*, 1805.
- Hur, G. H.; Vickery, C. R.; Burkart, M. D. *Nat. Prod. Rep.* **2012**, *10*, 1074.
- Fischbach, M. A.; Walsh, C. T. *Chem. Rev.* **2006**, *106*, 3468.
- Smith, S.; Tsai, S. C. *Nat. Prod. Rep.* **2007**, *24*, 1041.
- Boettger, D.; Hertweck, C. *ChemBioChem* **2013**, *14*, 28.
- Mootz, H. D.; Marahiel, M. A. *J. Bacteriol.* **1997**, *179*, 6843.
- van Wageningen, A. M. A.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J. M.; Solenberg, P. J. *Chem. Biol.* **1998**, *5*, 155.
- Schümann, J.; Hertweck, C. *J. Am. Chem. Soc.* **2007**, *129*, 9564.
- Keller, N. P.; Turner, G.; Bennett, J. W. *Nat. Rev. Microbiol.* **2005**, *3*, 937.
- Bumpus, S. B.; Evans, B. S.; Thomas, P. M.; Ntai, I.; Kelleher, N. L. *Nat. Biotechnol.* **2009**, *10*, 951.
- Meier, J. L.; Niessen, S.; Hoover, H. S.; Foley, T. L.; Cravatt, B. F.; Burkart, M. D. *ACS Chem. Biol.* **2009**, *4*, 948.
- Kleinkauf, H.; Gevers, W.; Lipmann, F. *Proc. Natl. Acad. Sci. U.S.A.* **1969**, *62*, 226.
- Lee, S. G.; Lipmann, F. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2343.
- Heacock, D.; Forsyth, C.; Shiba, K.; Musier-Forsyth, K. *Bioorg. Chem.* **1996**, *24*, 273.
- Finking, R.; Neumüller, A.; Solsbacher, J.; Konz, D.; Kretzschmar, G.; Schweitzer, M.; Krumm, T.; Marahiel, M. A. *ChemBioChem* **2003**, *4*, 903.
- Qiao, C.; Gupte, A.; Boshoff, H. I.; Wilson, D. J.; Bennett, E. M.; Somu, R. V.; Barry, C. E., III; Aldrich, C. C. *J. Med. Chem.* **2007**, *50*, 6080.
- Duckworth, B. P.; Wilson, D. J.; Nelson, K. M.; Boshoff, H. I.; Barry, C. E., III; Aldrich, C. C. *ACS Chem. Biol.* **2012**, *7*, 1653.
- Sato, S.; Kwon, Y.; Kamisuki, S.; Srivastava, N.; Mao, Q.; Kawazoe, Y.; Uesugi, M. *J. Am. Chem. Soc.* **2007**, *129*, 873.
- Cisar, J. S.; Ferreras, J. A.; Soni, R. K.; Quadri, L. E. N.; Tan, D. S. *J. Am. Chem. Soc.* **2007**, *129*, 7752.
- Ferreras, J. A.; Stirrett, K. L.; Lu, X.; Ryu, J.-S.; Soll, C. E.; Tan, D. S. *Chem. Biol.* **2008**, *15*, 51.
- Kratzschmar, J.; Krause, M.; Marahiel, M. A. *J. Bacteriol.* **1989**, *171*, 5422.
- Hoyer, K. M.; Mählert, C.; Marahiel, M. A. *Chem. Biol.* **2007**, *14*, 13.
- Stachelhaus, T.; Hüser, A.; Marahiel, M. A. *Chem. Biol.* **1996**, *3*, 913.
- Wilson, D. J.; Aldrich, C. C. *Anal. Biochem.* **2010**, *404*, 56.
- Meier, J. L.; Mercer, A. C.; Burkart, M. D. *J. Am. Chem. Soc.* **2008**, *130*, 5443.
- Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 769.