Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 14 (2004) 6067–6070

Improving the intein-mediated, site-specific protein biotinylation strategies both in vitro and in vivo

Lay-Pheng Tan,[a] Rina Y. P. Lue,[a] Grace Y. J. Chen[a,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 28 August 2004; revised 21 September 2004; accepted 22 September 2004
Available online 18 October 2004

Abstract—One of the critical issues in the generation of a protein microarray lies in the choice of immobilization strategies, which ensure proteins are adhered to the glass surface while properly retaining their native biological activities. We previously developed intein-mediated strategies for protein biotinylation and site-specific protein microarray generation. Herein, we report new findings of these strategies, which improve the biotinylation efficiency of proteins by up to 10-folds.

A major challenge in the post-genome era is to develop analytical techniques, which are capable of high-throughput identification and characterization of a large number of proteins encoded by the human genome. One of the most promising technologies available is the protein microarray, which offers the platform for simultaneous studies of tens of thousands of proteins.[1,2] Currently, an impending issue in protein microarray technologies lies in the successful development of robust strategies, which allow efficient immobilization of proteins onto glass surfaces while maintaining their native biological functions.[2] This is mainly due to the fact that proteins could easily unfold and lose their activity if not properly attached to a suitable surface. However, there are few existing immobilization strategies, which allow for uniform and stable immobilization of proteins in a microarray.[3–8] Zhu et al. used Ni–NTA-coated glass slides for site-specific attachment of $(His)_6$-tagged proteins in their ground-breaking work of the 'yeast proteome array'.[3] The noncovalent $(His)_6$–Ni–NTA interaction however, is not ideal due to its relatively weak binding, and is often susceptible to interference by commonly used chemicals and salts. Consequently, only a limited number of down-stream screening techniques may be compatible with this type of protein arrays.[2] Several groups, including our own, have developed alternative approaches which allow stable, and at the same site-specific, immobilization of proteins.[4–8] Mrksich and co-workers immobilized cutinase-fused proteins onto phosphonate-containing glass surfaces.[4] Johnsson and co-workers developed a site-specific method to covalently immobilize hAGT-fused proteins onto modified glass surfaces.[5] Similarly, Walsh and co-workers used Sfp phosphopantetheinyl transferase to mediate site-specific covalent immobilization of target proteins fused to the peptide carrier protein (PCP) excised from a nonribosomal peptide synthetase (NRPS).[6] We recently developed intein-mediated approaches to site-specifically biotinylate proteins both in vitro and in vivo, and subsequently immobilized them onto avidin-coated slides to generate the corresponding protein array.[7,8] The strategy of protein immobilization using our intein-mediated methods is highlighted by the following key aspects: (1) proteins were site-specifically biotinylated at their C-termini, leading to their subsequent immobilization on avidin-functionalized surfaces in an uniform orientation; (2) unlike other fusion protein approaches, our approaches did not introduce any extra macromolecular tag at the end of protein biotinylation, thus minimizing the potential perturbation to the protein's biological activity; (3) the biotin/avidin interaction is one of the strongest noncovalent interaction known ($K_d \sim 10^{-15}$ M), and is stable under most stringent conditions, making it an ideal method for protein immobilization; (4) our protein biotinylation strategies could be done either in vitro or in vivo, where the in vivo strategy allowed easy access

Keywords: Intein; Protein biotinylation; Protein array; Proteomics.
* Corresponding author. Tel.: +65 68741683; fax: +65 67791691; e-mail: chmyaosq@nus.edu.sg

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.09.083

to desired biotinylated proteins, in the form of crude lysates (e.g., unpurified proteins), for subsequent protein immobilizations and microarray generation. We showed that biotinylated proteins in the cell lysate could be washed away on-chip in an efficient (e.g., protein immobilization and purification are done in a single step) and highly parallel (e.g., thousands of different protein spots could be processed simultaneously on a single glass slide) fashion, resulting in purified proteins immobilized on the microarray.[8] This is true because of the rare occurrence of naturally biotinylated proteins in the cell, and the highly specific and strong nature of biotin/avidin interaction, which can withstand extremely stringent washing/purification conditions otherwise impossible with other affinity tags.[7]

Despite these obvious advantages over other existing methods for site-specific protein immobilization, our intein-mediated approaches were hampered by one key problem—its relatively low protein biotinylation efficiency, both in vitro and (especially) in vivo (e.g., <50% and 0–20%, respectively). This has precluded them from becoming a truly useful tool for large-scale protein expression and protein microarray generation. To this end, we present new findings on our intein-mediated protein biotinylation strategies. In the current work, we studied the efficiency of our biotinylation strategies, both in vitro and in vivo, in respect to different intein fusions used (Fig. 1). We have thus far identified a mini-intein, Mxe intein from *Mycobacterium xenopi*, which improved the in vitro and in vivo efficiency of protein biotinylation by at least 2- and 10-folds, respectively.

Currently, over 100 different inteins have been identified from different organisms.[9] Intein are believed to have evolved to possess differential protein splicing activities based on the context of their host organisms.[10] In our intein-mediated protein biotinylation methods (Fig. 1),[7,8] the target protein is fused to the N-terminus of a suitable intein. Upon expression in a host cell, the resulting fusion protein undergoes intein-catalyzed, intramolecular rearrangement to generate a protein–intein thioester. Upon further treatments with a cysteine-containing biotinylating reagent, cysteine–biotin (Fig. 1 inset), either in vitro or in vivo, the thioester is cleaved, and at the same time biotinylated, at the protein–intein fusion site to generate the final target protein, which contains a biotin 'tag' at its C-terminus. Given the number of different inteins available in the literature, which confer varied catalytic activities, we speculated that, with our intein-mediated protein biotinylation approaches, the biotinylation efficiency of a target protein fused to different intein tags may differ as well. In our previous studies, we have successfully used the 50 kDa Sce VMA intein isolated from *Saccharomyces cerevisiae* to biotinylate proteins, both in vitro and in vivo, with varying degrees of efficiency.[7,8] We speculated that improved protein biotinylation may be achieved by the use of other intein fusions. We were particularly interested in two naturally occurring mini-inteins, Mxe and Mth, isolated from *Mycobacterium xenopi* and *Methanobacterium thermoautotrophicum*, respectively, because of their relatively small sizes (198 and 134 amino acid residues, respectively).[10] Compared with the Sce VMA intein, these two mini-inteins lack the homing endonuclease domain but process the two important terminal regions, which are essential for protein splicing activity. Previous studies indicated that proteins fused to these two mini-inteins undergo splicing efficiently.[10] We therefore compared, in our intein-mediated strategies, the relative biotinylation efficiency of a protein when fused to each of the three different inteins.

Two new EGFP–intein constructs, EGFP–Mxe and EGFP–Mth, which express EGFP as the N-terminal fusions of the two mini-inteins, Mxe and Mth, respectively, were cloned as shown in Figure 2. Together with the original EGFP–Sce construct,[7,8] all three EGFP-expressing genes used in this study are otherwise identical, except their inteins. Chitin resin, pTYBI, pTWIN1 and pTWIN2 expression vectors were purchased from New England Biolabs (USA). pEGFP expression vector was purchased from Clontech (USA). Cysteine–biotin was prepared as previously described.[7] 2-Mercaptoenthanesulfonic acid (MESNA) was purchased from Aldrich (USA). A chitin binding domain (CBD) was fused to the C-terminus of each intein for easy purification of the fusion using chitin columns. The EGFP–Mxe intein and EGFP–Mth intein constructs were generated by cloning the EGFP gene (PCR-amplified from pEGFP vector) into pTWIN1

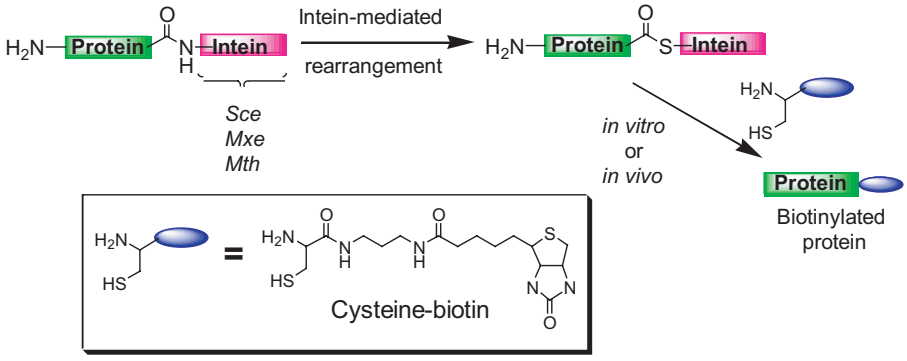

Figure 1. Intein-mediated protein biotinylation.

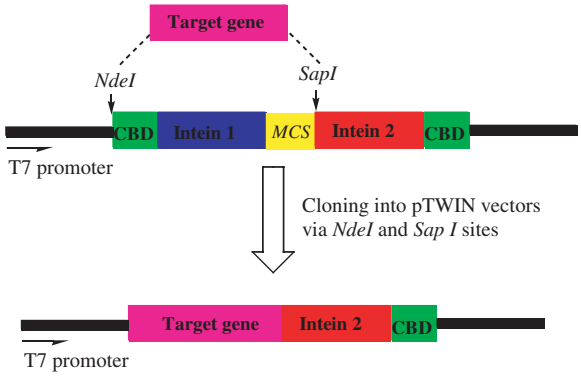

Figure 2. Cloning of EGFP–Mxe and EGFP–Mth from the pTWIN vectors (NEB, USA). The pTWIN vectors were designed for the generation of an N-terminal cysteine residue and/or a C-terminal thioester on the target protein. Both pTWIN1 and pTWIN2 carry the 154 amino acid residues Ssp DnaB mini-intein (Intein 1) for the production of an N-terminal cysteine. pTWIN1 vector differs from pTWIN2 vectors at the intein 2 coding region, with Mth RIR1 intein (134 amino acid residues) in place of Mxe GyrA intein (198 amino acid residues). Cloning of EGFP gene into pTWIN vectors via NdeI and SapI restriction sites resulted in the fusion of EGFP to the N-terminus of each intein.

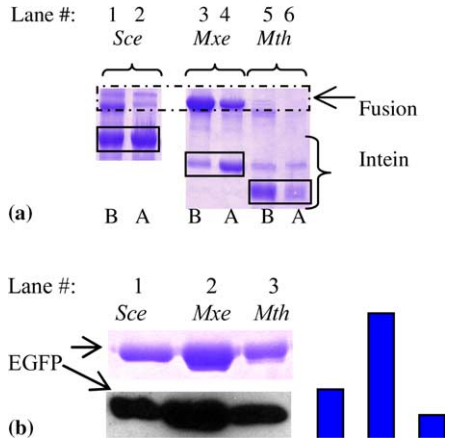

Figure 3. Efficiency of in vitro protein biotinylation of EGFP fused to three different inteins. (a) Yields of fusion proteins and its cleavage efficiency. B: Proteins bound on chitin beads before cysteine–biotin/MESNA cleavage; A: proteins remaining on chitin beads after cysteine–biotin/MESNA cleavage. (b) Yields of eluted/biotinylated EGFP after cysteine/MESNA biotinylation. Lane 1: EGFP–Sce intein–CBD fusion; lane 2: EGFP–Mxe intein–CBD fusion; lane 3: EGFP–Mth intein–CBD. Both coomassie gel and anti-biotin blot were presented. The gel from anti-biotin blot was used to quantitate protein biotinylation efficiency (represented by the bar graph on the right).

and pTWIN2 vectors, respectively, at the two restriction sites, NdeI and SapI, following protocols provided by the vendor. The resulting constructs, EGFP–Mxe intein and EGFP–Mth intein, contain the EGFP gene fused to the 23 kDa Mxe GyrA mini-intein from *Mycobacterium xenopi* and the 17 kDa Mth RIRI mini-intein from *Methanobacterium thermoautotrophicum*. All three constructs were transformed into ER2566 *E. coli.* host strain (NEB) for protein expression. Fusion proteins were biotinylated, either in vitro or in vivo, and subsequently assessed for their biotinylation efficiency as previously described.[8,11]

In order to compare the in vitro-based, on-column protein biotinylation efficiency, several parameters were assessed and they were done by examining (1) EGFP–intein fusions isolated on the chitin column before cleavage, (2) intein tags remained on the column following cysteine–biotin cleavage, as well as (3) the eluted, biotinylated EGFP. Results are shown in Figure 3: in vivo premature cleavage was evident with both Sce and Mth intein fusions, generating large amount of EGFP, which could not be subsequently biotinylated (Fig. 3a; lanes 1 and 5). This inevitably led to low protein cleavage/biotinylation efficiency upon treatments with cysteine–biotin (lanes 2 and 6 in Fig. 3a; lanes 1 and 3 in Fig. 3b). For the EGFP–Mxe intein fusion however, insignificant in vivo cleavage of the fusion was detected (Fig. 3a, lane 3), ensuring majority of the expressed EGFP to be subsequently biotinylated. As a result, significantly higher overall protein biotinylation efficiency was observed with this intein fusion (Fig. 3b, lane 2 vs lanes 1 and 3), giving rise to an estimated >2-fold increase in the overall protein biotinylation efficiency in vitro.

We next assessed the in vivo protein biotinylation efficiency with the three constructs. Cysteine–biotin, together with MESNA, was added to bacterial cells expressing EGFP–Sce intein, EGFP–Mxe intein and EGFP–Mth intein, respectively, and the in vivo biotinylation reaction was incubated further at 4 °C for 24 h, as previously described.[8,11] Upon extensive washings, cells were harvested, lysed and directly analyzed by SDS–PAGE and western blots with anti-biotin antibody (Fig. 4): similar to the in vitro experiments described earlier, significantly improved biotinylation efficiency of EGFP was observed with EGFP–Mxe intein (up to 10-fold increase compared with EGFP–Sce VMA; lane 2 vs 1). Consequently, the overall in vivo protein

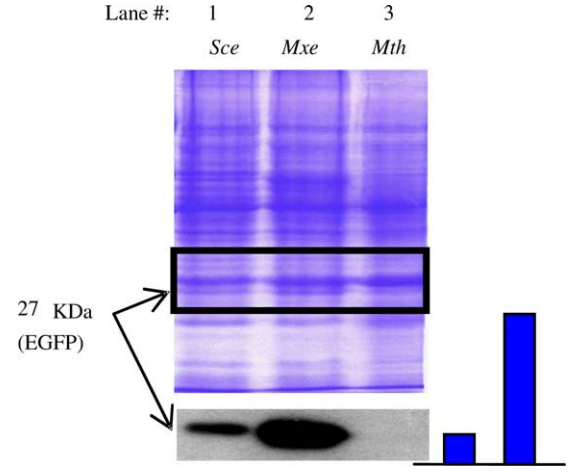

Figure 4. Efficiency of in vivo protein biotinylation of EGFP fused to three different inteins. Coomassie gel (top) of the crude lysate obtained from in vivo biotinylation was shown, with regions corresponding to EGFP highlighted (by solid box). Anti-biotin blot (bottom) was used to confirm the presence of biotinylated EGFP within the cell lysate, and quantitated for biotinylation efficiency (represented by the bar graph on the right). Lane 1: EGFP–Sce intein–CBD fusion; lane 2: EGFP–Mxe intein–CBD fusion; lane 3: EGFP–Mth intein–CBD.

biotinylation efficiency was correspondingly improved to an estimated 50–95% (of all EGFP expressed in the cell). This improvement is impressive when compared results obtained from the original construct (i.e., EGFP–Sce VMA with an estimated 0–20% in vivo biotinylation efficiency), especially if one considers that the only difference between the two constructs were the inteins. The other mini-intein fused protein, EGFP–Mth intein, did not produce any significant amount of biotinylated EGFP (i.e., lane 3), indicating that in vivo biotinylation was greatly reduced, presumably as a result of premature cleavage of the fusion.

Taken together, we showed that the biotinylation efficiency, both in vitro and in vivo, of our intein-mediated strategies are highly dependent upon the intein fusions. Good improvement of the protein biotinylation efficiency may be achieved by the choice of different inteins. This thus lays groundwork for further improvement of our strategies in the future. In summary, our new findings indicate the efficiency of intein-mediated protein biotinylation, both in vitro and in vivo, depends greatly on the intein fused to target protein: as much as 2–10 times improvement in protein biotinylation may be achieved by the simple switch in the intein used. Highlighted in the improved intein-mediated strategies was that, for the first time, we were able to biotinylate the majority of the target protein expressed inside live cells, achieving an estimated 50–95% overall protein biotinylation efficiency. This improvement is critical if one is to consider using the intein-mediated approaches in future for potential large-scale protein expression and protein microarray generation in a truly high-throughput manner (e.g., without protein purifications). So far we have only tested three different inteins. There are well over 100 known inteins, some of which may prove to be even better choices for protein biotinylation in our strategies. With the emerging role of microarray-based technologies in the post-genome era,[1–8] findings described herein may also present useful tools for a variety of other proteomics research.[12–15]

Acknowledgements

Funding support was provided by the National University of Singapore (NUS) and the Agency for Science, Technology and Research (A*STAR) of Singapore.

References and notes

1. MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289*, 1760–1763.
2. Chen, G. Y. J.; Uttamchandani, M.; Lue, R. Y. P.; Lessaicherre, M. L.; Yao, S. Q. *Curr. Top. Med. Chem.* **2003**, *3*, 705–724.
3. Zhu, H. et al. *Science* **2001**, *293*, 2101–2105.
4. Hodneland, C. D.; Lee, Y. S.; Min, D. H.; Mrksich, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5048–5052.
5. Kindermann, N.; George, M.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2003**, *125*, 7810–7811.
6. Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 7754–7755.
7. Lesaicherre, M. L.; Lue, R. Y. P.; Chen, G. Y. J.; Zhu, Q.; Yao, S. Q. *J. Am. Chem. Soc.* **2002**, *124*, 8768–8769.
8. Lue, R. Y. P.; Chen, G. Y. J.; Hu, Y.; Zhu, Q.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 1055–1062.
9. www.neb.com/neb/inteins.html.
10. Xu, M. Q.; Evan, T. C. *Methods* **2001**, *24*, 257–277.
11. Briefly, the transformed ER2566 cells were grown in Luria Bertani (LB) medium supplemented with 100 μg/mL ampicillin at 37 °C in a 250 rpm shaker to an OD_{600} of about ~0.5. Protein expression was induced overnight at room temperature using 0.3 mM isopropyl thiogalactosidase (IPTG). For in vitro-based, on-column biotinylation, cells were harvested and lysed as previously described.[7,8] The clear supernatant was incubated with chitin resin for 30 min at 4 °C with gentle agitation. Subsequently, the resin was washed with the column buffer (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA) followed by incubation with the cleavage buffer (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 50 mM MESNA and 5 mM cysteine–biotin) overnight at 4 °C. Upon resin settlement, the supernatant which contains the eluted, biotinylated protein was collected, together with the column-bound fraction, and analyzed by SDS-PAGE and western blots. For in vivo biotinylation, MESNA and cysteine–biotin (final concentrations: 10 and 5 mM, respectively) were added directly to the cell medium following protein expression. The reaction was allowed to proceed overnight at 4 °C, after which cells were harvested and washed thoroughly with PBS followed by lysis. The lysate was analyzed directly by SDS-PAGE and western blots with anti-biotin antibody.
12. Yeo, S. Y. D.; Srinivasan, R.; Chen, G. Y. J.; Yao, S. Q. *Chem. Eur. J.* **2004**, *126*, 4664–4672.
13. Zhu, Q.; Girish, A.; Chattopadhaya, S.; Yao, S. Q. *Chem. Commun.* **2004**, 1512–1513.
14. Chan, E. W. S.; Chattopadhaya, S.; Panicker, R. C.; Huang, X.; Yao, S. Q. *J. Am. Chem. Soc.*, in press.
15. Hu, Y.; Chen, G. Y. J.; Yao, S. Q. *Angew. Chem., Int. Ed.*, in press.