

Table 1. List of 96 Identified Proteins As Autoantigen Proteins

gene symbol	source <sup>a</sup>	M <sub>w</sub> (kDa)	MRL/lpr <sup>b</sup>	NMS <sup>b</sup>	Gene Ontology <sup>c</sup>	
					cellular location	biological function/process
Agpat3	Chr 10_40	43.3	49.7	2.6	Me	
Bcr	Chr 10_40	35.3	6.4	1.5		K
Unc5b	Chr 10_40	103.7	42.3	3.6	Me	
Pyp	Chr 10_40	33.0	6.3	1.3		
Cdc2a	Chr 10_40	34.1	11.9	3.6	N	K
Cnm2	Chr 19_49	96.6	64.4	4.4	Me	
Thop1	Chr 10_40	78.0	38.9	2.3	C	K, P
Hhex <sup>d</sup>	Chr 19_49	30.0	54.7	1.9	N	Tc
Palm	Chr 10_40	41.6	59.4	4.1	C, Me	
Gnaz	Chr 10_40	40.8	72.2	5.0	Me	
2610028F08Rik	Chr 15_18	28.3	17.9	1.7		
Ank	Chr 15_18	54.3	55.7	1.7	Me	
Nov	Chr 15_18	38.8	20.4	1.5	E	
Neurl	Chr 19_49	36.0	58.0	4.8		
Gpam	Chr 19_49	93.4	54.2	2.6	Mi, Me	
Ncald	Chr 15_18	22.2	5.9	1.1		
Wnt8b	Chr 19_49	40.5	13.8	2.7	E	
Timp3	Chr 10_40	24.2	19.8	2.3	Me, E	K
Osr2	Chr 15_18	30.6	17.2	1.5	N	
Nnp1	Chr 10_40	54.6	51.3	3.7		
Nfic	Chr 10_40	48.8	80.4	3.2	N	Tc
Pfkl	Chr 10_40	85.4	8.3	3.1	C	K
Slc18a2	Chr 19_49	55.8	72.5	2.5	C, Me	
Sgta <sup>d</sup>	Chr 10_40	34.2	23.0	1.3		
Hps6	Chr 19_49	87.3	59.5	3.4		
Sgpl1	Chr 10_40	63.7	90.9	4.0	Me	
Pdxk	Chr 10_40	35.1	8.1	1.0	C	K
Pwp2h	Chr 10_40	102.9	36.3	2.7		
Psap	Chr 10_40	61.1	42.2	2.4	Mi, E	
Fzd6	Chr 15_18	79.1	92.0	2.6	Me	
Ilvbl	Chr 10_40	68.2	45.8	3.0	Me	
Itgb2	Chr 10_40	84.9	41.7	2.9	Me	
Cstb	Chr 10_40	11.0	18.9	1.4	N, C	
Gstt2	Chr 10_40	27.6	25.1	1.3	N, C	
Bsg	Chr 10_40	29.7	20.4	1.3	Me	
Cpn1	Chr 19_49	52.1	6.1	1.0	E	P
Eif3s6	Chr 15_18	52.0	8.7	1.2		
Timm9	Chr 10_40	10.4	14.2	2.2	Mi, Me	
Ndufs7	Chr 10_40	24.7	11.8	2.9	Mi, Me	
Psd	Chr 19_49	12.4	51.6	2.4	Me	
Tfam	Chr 10_40	28.0	17.8	2.0	N, Mi	Tc
Ppap2c	Chr 10_40	31.2	51.2	1.6	Me	
Gpx4	Chr 10_40	22.1	18.1	1.5	N, C, Mi, Me	
Pcbd1	Chr 10_40	12.0	44.4	2.0	N, C	Tc
Ins1	Chr 19_49	12.2	30.1	1.9	E	
Mrpl54	Chr 10_40	15.4	8.5	1.0	Mi	
Oaz1	Chr 10_40	25.1	11.4	1.3		
Cxxc6	Chr 10_40	25.6	32.1	1.5		
Sdc2	Chr 15_18	22.1	42.5	2.0	Me	
Npm3	Chr 19_49	19.0	13.1	1.2	N	Tc
Eif4ebp2	Chr 10_40	12.9	7.6	1.0	Tl	
Ddt	Chr 10_40	13.1	5.1	0.9	C	
Pah	Chr 10_40	51.8	32.1	2.2		
Peo1	Chr 19_49	77.0	6.8	0.9	Mi	
Cabin1	Chr 10_40	65.2	13.8	1.3	N	
Lilrb4	Chr 10_40	37.5	9.0	1.4	Me	
Casp7	Chr 19_49	34.1	7.6	1.3	C	P
Matk	Chr 10_40	53.6	72.9	4.0	C, Me	K
Egr2	Chr 10_40	49.8	74.7	3.2	N	Tc
Slc1a6 <sup>d</sup>	Chr 10_40	60.8	126.0	2.7	Me	
Adn	Chr 10_40	28.1	63.8	3.1		
Gnal1	Chr 10_40	42.0	69.7	4.7		
Tbxa2r	Chr 10_40	37.1	43.9	1.8	Me	
Trhr	Chr 15_18	44.6	188.7	3.5	Me	
Ube2g2	Chr 10_40	33.0	16.0	1.8		U, Tl
Madcam1	Chr 10_40	43.6	33.5	1.8	Me	
Pcdh15	Chr 10_40	129.9	70.6	2.6	C, Me	
Efna2	Chr 10_40	23.6	5.3	1.0	Me	
Sema5a	Chr 15_18	120.3	19.1	1.8	Me	
Aire	Chr 10_40	18.0	8.6	1.2	N, C	Tc, Tl
Fgf8	Chr 19_49	24.7	11.1	1.8	E	
Snrpd2	AA*	13.6	13.4	3.3	N	
Hmgn2	AA	9.4	85.7	2.2	N, C	
Mcrs1	AA	51.7	45.5	3.2	N	
Hnrpa2b1	AA	32.5	19.5	2.1		
Hars	AA	57.4	104.9	4.8	C	Tl
Rpo1-3	AA*	15.1	6.4	4.2		

Table 1. Continued

gene symbol	source <sup>a</sup>	M <sub>w</sub> (kDa)	MRL/lpr <sup>b</sup>	NMS <sup>b</sup>	Gene Ontology <sup>c</sup>	
					cellular location	biological function/process
Hars2	AA*	23.4	7.3	1.6	C, Mi	Tl
Hspca	AA	84.8	5.9	3.1		
Vtn	AA	54.8	5.8	1.6	E	
Snrpd3	AA*	13.9	10.2	2.8	N, C	
Hmgn1	AA*	10.1	35.6	4.5	N, C	Tc
Rnps1	AA	40.8	26.0	1.4	N, C	
Fbl	AA	34.2	24.9	1.9	N	
Npm1	AA	32.6	12.6	1.1	N, C	K
Top3b	AA*	96.9	18.9	2.6		
Coil	AA	62.2	9.8	1.6	N	
Casp8	AA	55.4	27.5	2.8	N, C	P
Ybx1	AA	35.7	27.8	4.7	N, C	Tc
Srpkl	AA	73.1	17.7	3.8	N, C	K
Rpa1	AA	71.4	36.6	3.8	N	
Car9	AA*	47.3	8.9	1.5	Me	
Sag	AA	44.9	11.9	2.2		
Dnahc8	AA	122.1	9.2	1.6	C	
Top3a	AA*	107.0	13.1	2.0		
Fbn2	AA	56.6	5.9	1.4	E	

<sup>a</sup> The source of selected gene done by symbol, is as follows: Chr 10\_40, genetic loci on chromosome 10\_40 cM; Chr 15\_18, genetic loci on chromosome 15\_18 cM; Chr 19\_49, genetic loci on chromosome 19\_49 cM; AA, well-known autoantigen; AA\*, well-known autoantigen homologue. <sup>b</sup> Relative luminescence signals. <sup>c</sup> According to Gene Ontology (GO) Database (<http://www.geneontology.org/>), the proteins were classified by cellular localization and biological function/process, is as follows: Me, Membrane (GO:0016020); N, Nucleus (GO:0005634); C, Cytoplasm (GO:0005737); E, Extracellular region (GO:0005576); Mi, Mitochondrion (GO:0005739); K, Kinase activity (GO:0016301); P, Peptidase activity (GO:0008233); U, Ubiquitin (GO:0005551); Tl, Translation (GO:0006412); Tc, Transcription (GO:0006350). <sup>d</sup> Hhex, lane1; Sgta, lane3; Slc1a6, lane4 in Figure 4.

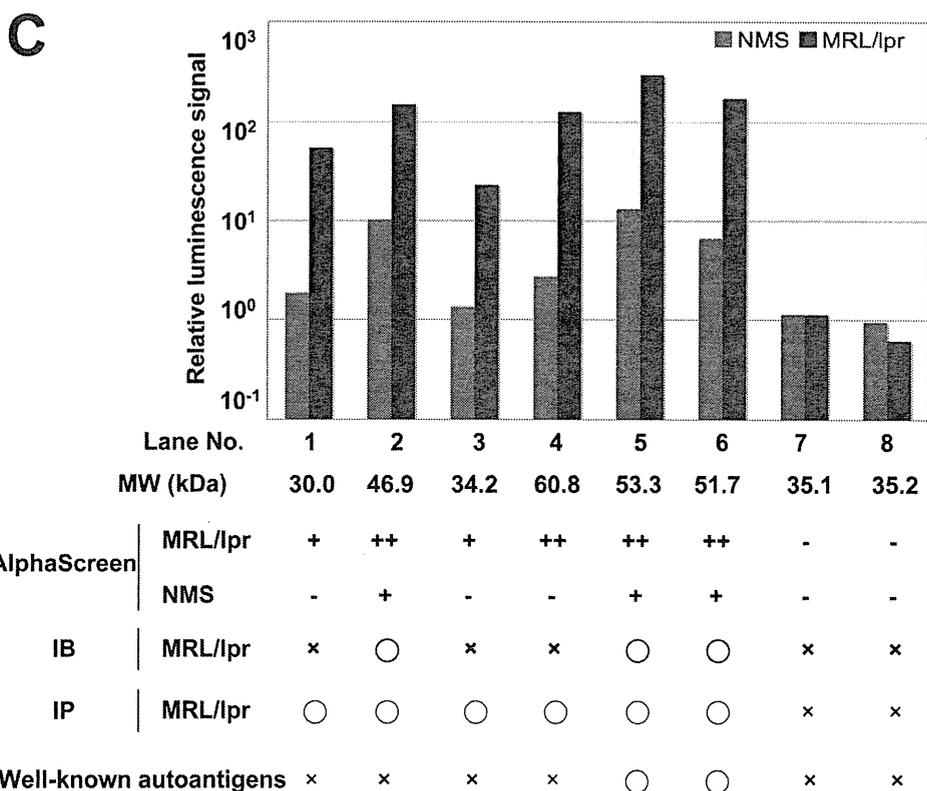
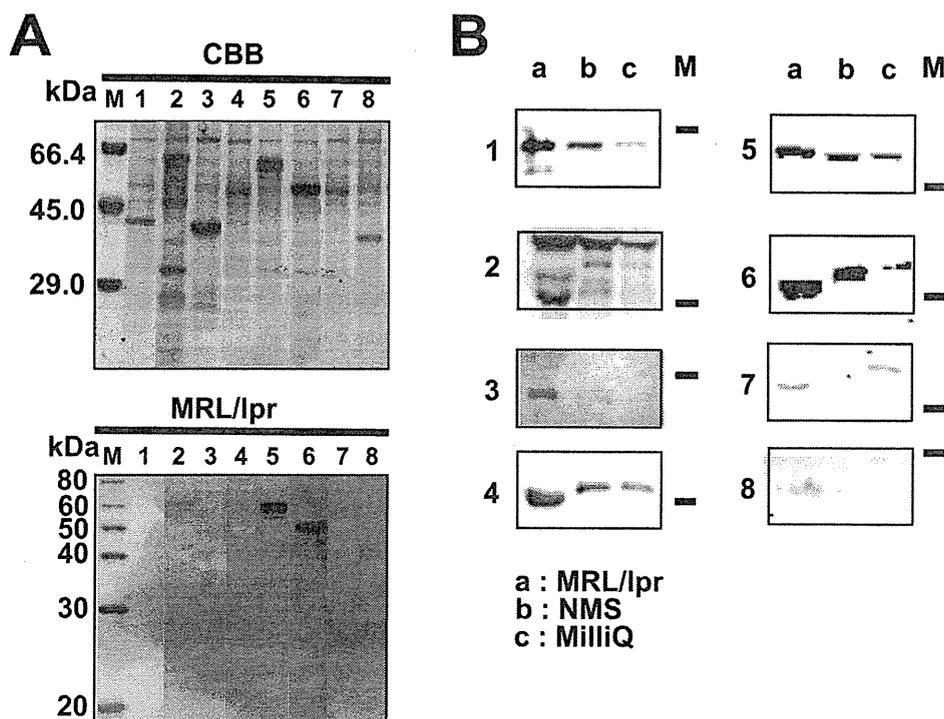
reactive autoantigen proteins. Interestingly, these loci were reported as the susceptibility loci of arthritis. Furthermore, according to Gene Ontology (GO) Database (<http://www.geneontology.org/>), 73 (76%) of the proteins were classified by cellular localization (Figure 3E) and 25 (26%) classified by biological function/process (Figure 3F). The annotated proteins found were classified as localized with Membrane (22%), Membrane/Cytoplasm (4%), Extracellular space (7%), Nucleus (13%), Nucleus/Cytoplasm (13%), Cytoplasm (7%) and Mitochondrion (2%). Also, the annotated proteins were involved in diverse biological functions/processes such as Kinase activity (8%), Kinase activity/peptidase activity (1%), transcription (8%), Transcription/Translation (1%), Translation (3%), Ubiquitin/Translation (1%) and Peptidase activity (3%). Data analysis showed that 26 and 25 proteins were annotated in localization of nucleus and cytoplasm respectively (Figure 3E, Table 1), and that 9 and 9 proteins were related to cellular events of protein phosphorylation and transcription, respectively (Figure 3F, Table 1). Many nuclear proteins were reported as autoantigens.<sup>2</sup> Interestingly, localization of 38 (39.6%) antigen proteins reacting with antibodies in MRL/lpr mouse sera was annotated in membrane and/or extracellular space. These results suggest that the wheat cell-free system is a viable platform to study folded membrane proteins that function as antigens. The data analysis suggests that MRL/lpr autoantigen proteins are represented by a wide variety of biological functions localized in whole cells, rather than just nuclear proteins. Taken together, these results indicate that the BPL-based screening method would be useful for identification of autoantigen proteins.

**Validation of Identified Autoantigen Proteins by Immunoblotting and Immunoprecipitation.** Recent reports have mentioned the possibility that autoantibodies may react with conformational epitopes.<sup>29–31</sup> These data were obtained by liquid phase immunoprecipitation assays using recombinant proteins.<sup>32,33</sup> Under these situations, autoantigen proteins we found were analyzed by immunoblotting and immunoprecipitation. For this analysis, six proteins were randomly selected (see legend in Figure 4). Immunoblot analysis showed that

three proteins reacted with MRL/lpr mouse sera (Lanes 2, 5, and 6 in Figure 4A). Two of these three proteins, LmnB2 (Lane 5) and topoisomerase II alpha (Top2a) (Lane 6), have been well characterized as autoantigen proteins so far.<sup>22,34</sup> Also, six of our identified autoantigen proteins were not detected by immunoblot analysis in the sera of NMS (data not shown). Interestingly, immunoprecipitation analysis revealed antigenicity of all six autoantigen proteins (Figure 4B, C), whereas two proteins randomly selected from nonautoantigen proteins, serving as a negative control, did not show significant reaction to the sera from MRL/lpr mice by both immunoblot analysis and immunoprecipitation (Lanes 7 and 8 in Figure 4A, B). These results suggest that the BPL-based screening method may be useful for identification of autoantigen proteins reacting with autoantibodies recognizing conformational epitopes.

## Discussion

To address high-throughput protein production, we have utilized our wheat germ high-throughput protein synthesis system,<sup>10,16</sup> which can produce large numbers of recombinant proteins using a fully automated robot.<sup>19</sup> To create a library of target autoantigen proteins, full-length human and mouse cDNA resources were provided by the Mammalian Gene Collection (MGC) clones (Mammalian Gene Collection Program, <http://mgc.nci.nih.gov/>) and FANTOM.<sup>27,28</sup> Since the full-length cDNA was provided in plasmids, no additional time-consuming cloning steps were needed for the synthesis of linear DNA templates by PCR for direct entry into the cell-free based protein production system. Additionally, researchers can select and use any appropriate peptide tag for downstream applications, like a bls used in this study, owing to the ease of template construction. In fact, given the advantages of the gateway system and PCR, a recent publication reported successful production of 13 000 His-tagged human proteins by the wheat cell-free system using full-length cDNA resources.<sup>12</sup> Furthermore, because protein purification is a time-consuming-step, an assay system with no purification requirement could dramatically increase the throughput. For that, a specific



**Figure 4.** Detection of identified autoantigen proteins by immunoblotting and immunoprecipitation. (A) Immunoblotting analysis by using recombinant proteins. Purified recombinant proteins were separated by SDS-PAGE and stained with CBB (Upper). Purified recombinant proteins were reacted with serum from MRL/lpr mouse (Lower). (B) Immunoprecipitation analysis using recombinant proteins. Translation mixtures expressing biotinylated proteins were incubated with 1  $\mu$ L of undiluted serum overnight at 4  $^{\circ}$ C. Immobilized Protein A sepharose was added to each sample, and incubated for 60 min at 4  $^{\circ}$ C. After washing, proteins were separated by SDS-PAGE, followed by immunoblotting with Alexa488-STA. M indicates a 45 kDa protein molecular weight marker. (C) Whole data of randomly selected proteins. IB, Immunoblotting; IP, Immunoprecipitation. Relative luminescence signal,  $10^2 \leq ++$ ;  $5 \leq + < 10^2$ ;  $- < 5$ . (A, B, C) Lane 1, Hhex; Lane 2, Tdg; Lane 3, Sgta; Lane 4, Slc1a6; Lane 5, Lmn2; Lane 6, Top2a; Lane 7, Cs; Lane 8, Car4. Lane 7 and 8 were negative controls. Detailed information on individual proteins was indicated in Supplementary Table 1 (Supporting Information).

protein has to be clearly recognized in a homogeneous condition. In this study, we selected biotin as our target protein label due to the highly specific binding of biotin-streptavidin. Commonly, biotinylated proteins are produced via NHS ester-activated biotins. However, this technique requires laborious purification to remove any nonreacted biotin reagent in the reaction mixture. Therefore, we used a BirA biotin-ligase-based labeling system. By addition of BirA and biotin to the wheat cell-free system, highly specific biotin-labeling is available and the biotinylated proteins can be directly used for assaying<sup>3,15,35</sup> without further purification. Taken together, the biotinylated protein library produced by the wheat cell-free system is suitable for autoantigen screening.

Several autoantigen screening methods such as SERPA, SEREX and protein microarray are currently widely used for identification of autoantigen proteins, each of which has inherent limitations. In this study, we demonstrated improved methodologies that overcome the aforementioned limitations using a wheat cell-free based BPL and luminescence assay that allow detection of autoantigen proteins with autoantibodies in sera. The BPL-based screening revealed that specific antibody interaction were detected at subpicogram scale, with a linear response over a 1000-fold range, and appear to be more sensitive than conventional method, such as ELISA.<sup>5,36</sup> It should be noted that protein microarray could also detect autoantigen at picogram scale, in the linear fashion over a 1000-fold range.<sup>6</sup> While the detection sensitivity of the BPL-based autoantigen protein screening method might be equivalent to microarray based autoantigen protein screening, the folded state of the autoantigen proteins differs. The reports which autoantibodies would recognize conformational epitopes<sup>29,30</sup> might contain an important implication for screening of autoantigen proteins. Although conventional methods use denatured or unfolded proteins, such as dehydrated or detergent-treated proteins, the BPL-based method tested in this study better represents the folded, native form as all procedures are carried out in the solution phase without dehydration or detergent treatment. In addition, the BPL-based screening method using serum dilutions of 1:10 000 could produce high luminescent signal in 25  $\mu$ L of reaction volume (Figure 2C). Thus, 50  $\mu$ L of serum would be sufficient to screen 20 000 kinds of human proteins.

Autoimmune diseases were thought to be a complex of both genetic and nongenetic factors influencing susceptibility, severity and response to therapies.<sup>37</sup> Twin and family studies suggest that approximately 60% of susceptibility is due to genetic factors and genes within the HLA locus, particularly HLA-DRB1, which accounts for almost half of the genetic component of susceptibility.<sup>37</sup> Also, genetic analyses identified that other susceptibility locus of RA, SLE, and so on.<sup>38</sup> In this study, we screened proteins encoded by genes on an autoimmune susceptibility loci,<sup>26</sup> and 71 out of 111 clones (Table 1 and Supplementary Table 1, Supporting Information) localized on the chromosomes 10, 15 and 19<sup>26</sup> were found as new autoantigen proteins reacting with the sera of MRL/lpr mice. Additionally, the Gene Ontology (GO) Database may be extremely useful for the screening of autoantigen protein. For example, based on data from the GO Database, localization of 38 (39.6%) autoantigen proteins out of 96 MRL/lpr autoantigen proteins were annotated in membrane and/or extracellular space (Figure 3E). These results suggest that a specific protein library focused on the human autoimmune susceptibility loci and membrane

proteins or extracellular spaces classified according to GO Database may be a good target for screening of autoantigen proteins.

A key obstacle for robust genome-wide screening has been experimentally simple techniques and automated technology. The BPL-based screening method is one of the simplest approaches for identification of autoantigen proteins, because all experimental processes, including construction of DNA templates, and interaction and detection of antigen–autoantibodies reactions, were reduced to mixing steps. Synthesis of the biotinylated protein library was accomplished using a fully automated robot,<sup>19</sup> and the biotinylated proteins can be used in subsequent screening steps without purification. The method described here can be used in developed for use in 96, 384 (Figure 3A) or 1536-well microtiter-plate format through the use of appropriate automated liquid handling robots. Therefore this method is suitable for development of a genome-wide screening platform. In conclusion, the BPL-based screening method has a high potential for identification of autoantigen proteins in human autoimmune diseases.

**Abbreviations:** Alexa488-STA, streptavidin Alexa Fluor 488 conjugate; NMS, normal mouse sera; SEREX, serological expression cloning; SERPA, serological proteome analysis; bls, biotin ligation site; DHFR, dihydrofolate reductase; Lmnb2, lamin B2; Hars, histidyl-tRNA synthetase; Top2a, topoisomerase II alpha; Hhex, hematopoietically expressed homeobox; Tdg, thymine DNA glycosylase; Sgta, small glutamine-rich tetratricopeptide repeat (TPR)-containing, alpha; Slc1a6, solute carrier family 1 (high affinity aspartate/glutamate transporter), member 6; Cs, citrate synthase; Car4, carbonic anhydrase 4; BPL, biotinylated protein library; FANTOM, functional annotation of mouse; CBB, coomassie brilliant blue; RF, rheumatoid factor; hnRNP, heterogeneous nuclear ribonucleoprotein; Sm, Smith; GO, Gene Ontology; NHS, *N*-hydroxysuccinimide; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; MRL/lpr, MRL/Mp-*lpr/lpr*.

**Acknowledgment.** This work was partially supported by the Special Coordination Funds for Promoting Science and Technology by the Ministry of Education, Culture, Sports, Science and Technology, Japan (T.S. and Y.E.). We thank Michael Andy Goren and Dr. Akira Nozawa for proofreading this manuscript.

**Supporting Information Available:** Supplementary Table 1: List of selected 226 mouse genes and primer sequences used in this study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Davidson, A.; Diamond, B. *N. Engl. J. Med.* **2001**, *345* (5), 340–50.
- (2) von Mühlen, C.; Tan, E. *Semin. Arthritis Rheum.* **1995**, *24* (5), 323–58.
- (3) Mewar, D.; Wilson, A. *Biomed. Pharmacother.* **2006**, *60* (10), 648–55.
- (4) Gunawardana, C.; Diamandis, E. *Cancer Lett.* **2007**, *249* (1), 110–9.
- (5) Anderson, K.; Ramachandran, N.; Wong, J.; Raphael, J.; Hainsworth, E.; Demirkan, G.; Cramer, D.; Aronzon, D.; Hodi, F.; Harris, L.; et al. *J. Proteome Res.* **2008**, *7* (4), 1490–9.
- (6) Robinson, W.; DiGennaro, C.; Hueber, W.; Haab, B.; Kamachi, M.; Dean, E.; Fournel, S.; Fong, D.; Genovese, M.; de Vegvar, H.; et al. *Nat. Med.* **2002**, *8* (3), 295–301.
- (7) Hudson, M.; Pozdnyakova, I.; Haines, K.; Mor, G.; Snyder, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (44), 17494–9.
- (8) Sheridan, C. *Nat. Biotechnol.* **2005**, *23* (1), 3–4.
- (9) Zhu, H.; Snyder, M. *Curr. Opin. Chem. Biol.* **2003**, *7* (1), 55–63.

- (10) Sawasaki, T.; Ogasawara, T.; Morishita, R.; Endo, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (23), 14652–7.
- (11) Endo, Y.; Sawasaki, T. *Methods Mol. Biol.* **2005**, *310*, 145–67.
- (12) Goshima, N.; Kawamura, Y.; Fukumoto, A.; Miura, A.; Honma, R.; Satoh, R.; Wakamatsu, A.; Yamamoto, J.; Kimura, K.; Nishikawa, T.; et al. *Nat. Methods* **2008**, *5* (12), 1011–7.
- (13) Sawasaki, T.; Kamura, N.; Matsunaga, S.; Saeki, M.; Tsuchimochi, M.; Morishita, R.; Endo, Y. *FEBS Lett.* **2008**, *582* (2), 221–8.
- (14) Beaudet, L.; Bédard, J.; Breton, B.; Mercuri, R.; Budarf, M. *Genome Res.* **2001**, *11* (4), 600–8.
- (15) Takahashi, H.; Nozawa, A.; Seki, M.; Shinozaki, K.; Endo, Y.; Sawasaki, T. *BMC Plant Biol.* **2009**, *9*, 39.
- (16) Madin, K.; Sawasaki, T.; Ogasawara, T.; Endo, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (2), 559–64.
- (17) Sawasaki, T.; Morishita, R.; Gouda, M.; Endo, Y. *Methods Mol. Biol.* **2007**, *375*, 95–106.
- (18) Sawasaki, T.; Hasegawa, Y.; Tsuchimochi, M.; Kamura, N.; Ogasawara, T.; Kuroita, T.; Endo, Y. *FEBS Lett.* **2002**, *514* (1), 102–5.
- (19) Sawasaki, T.; Gouda, M.; Kawasaki, T.; Tsuboi, T.; Tozawa, Y.; Takai, K.; Endo, Y. *Methods Mol. Biol.* **2005**, *310*, 131–44.
- (20) Stephen, C.; Helminen, P.; Lane, D. *J. Mol. Biol.* **1995**, *248* (1), 58–78.
- (21) Miller, F.; Waite, K.; Biswas, T.; Plotz, P. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87* (24), 9933–7.
- (22) Brito, J.; Biamonti, G.; Caporali, R.; Montecucco, C. *J. Immunol.* **1994**, *153* (5), 2268–77.
- (23) Theofilopoulos, A.; Dixon, F. *Adv. Immunol.* **1985**, *37*, 269–390.
- (24) Spirin, A. *Trends Biotechnol.* **2004**, *22* (10), 538–45.
- (25) Endo, Y.; Sawasaki, T. *Biotechnol. Adv.* **2003**, *21* (8), 695–713.
- (26) Nose, M. *Allergol. Int.* **2007**, *56* (2), 79–86.
- (27) Okazaki, Y.; Furuno, M.; Kasukawa, T.; Adachi, J.; Bono, H.; Kondo, S.; Nikaido, I.; Osato, N.; Saito, R.; Suzuki, H.; et al. *Nature* **2002**, *420* (6915), 563–73.
- (28) Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; et al. *Science* **2005**, *309* (5740), 1559–63.
- (29) Xie, H.; Zhang, B.; Matsumoto, Y.; Li, Q.; Notkins, A.; Lan, M. *J. Immunol.* **1997**, *159* (7), 3662–7.
- (30) Tuomi, T.; Rowley, M.; Knowles, W.; Chen, Q.; McAnally, T.; Zimmet, P.; Mackay, I. *Clin. Immunol. Immunopathol.* **1994**, *71* (1), 53–9.
- (31) Mackay, I.; Rowley, M. *Autoimmun. Rev.* **2004**, *3* (7–8), 487–92.
- (32) Lan, M.; Wasserfall, C.; Maclaren, N.; Notkins, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (13), 6367–70.
- (33) Grubin, C.; Daniels, T.; Toivola, B.; Landin-Olsson, M.; Hagopian, W.; Li, L.; Karlsen, A.; Boel, E.; Michelsen, B.; Lernmark, A. *Diabetologia* **1994**, *37* (4), 344–50.
- (34) Chang, Y. H.; Shiau, M. Y.; Tsai, S. T.; Lan, M. S. *Biochem. Biophys. Res. Commun.* **2004**, *320* (3), 802–9.
- (35) Masaoka, T.; Nishi, M.; Ryo, A.; Endo, Y.; Sawasaki, T. *FEBS Lett.* **2008**, *582* (13), 1795–801.
- (36) Quintana, F. J.; Farez, M. F.; Viglietta, V.; Iglesias, A. H.; Merbl, Y.; Izquierdo, G.; Lucas, M.; Basso, A. S.; Khoury, S. J.; Lucchinetti, C. F.; Cohen, I. R.; Weiner, H. L. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (48), 18889–94.
- (37) Wordsworth, P.; Bell, J. *Ann. Rheum. Dis.* **1991**, *50* (6), 343–6.
- (38) Cornélis, F.; Fauré, S.; Martínez, M.; Prud'homme, J.; Fritz, P.; Dib, C.; Alves, H.; Barrera, P.; de Vries, N.; Balsa, A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (18), 10746–50.

PR9010553



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)

## *In vitro* dissection revealed that the kinase domain of wheat RNA ligase is physically isolatable from the flanking domains as a non-overlapping domain enzyme

Shin-ichi Makino<sup>a,b,1</sup>, Tatsuya Sawasaki<sup>a,b,c</sup>, Yaeta Endo<sup>a,b,c</sup>, Kazuyuki Takai<sup>a,b,c,\*</sup><sup>a</sup> Cell-free Science and Technology Research Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan<sup>b</sup> Venture Business Laboratory, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan<sup>c</sup> Department of Materials Science and Biotechnology, Graduate School of Science and Engineering, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

## ARTICLE INFO

## Article history:

Received 31 May 2010

Available online 10 June 2010

## Keywords:

Multidomain enzyme

Wheat RNA ligase

Domain structure

Cell-free translation

## ABSTRACT

Wheat RNA ligase contains 5'-hydroxyl kinase, 2',3'-cyclic phosphate 3'-phosphodiesterase, and 5'-phosphate 2'-phosphate-3'-hydroxyl RNA ligase activities in a 110-kDa polypeptide. Taking advantage of a wheat cell-free protein production system, we prepared various fragments containing a part of the enzyme. The method allowed us to check the activities of the fragments rapidly, eliminating the time-consuming cloning and sequencing steps for the expression of the fragment proteins. The results showed that each of the three activities can be assigned to a non-overlapping domain that does not require the presence of the other part(s) of the enzyme for its activity. This contrasts to the case of yeast tRNA ligase, in which the central kinase domain has been suggested to require to be tethered to one of the flanking domains for its activity.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Multidomain proteins are more widely found in eukaryotes than in prokaryotes [1,2], providing an intriguing aspect of protein evolution. In order to understand the complicated function of a multidomain enzyme, the protein is often dissected into smaller functional pieces. Segregating a domain from the enzyme would prove the functionality of each domain. Smaller proteins are also useful, in general, for three-dimensional structure determination. However, it is not always possible to dissect an enzyme into discrete functional domains in spite of the laborious experiments, which may include a time-consuming step of molecular cloning for production of many different protein fragments. An easier method for assessing the activities of truncated proteins is desired.

Wheat RNA ligase is known as a tri-functional enzyme that catalyzes various RNA-ligation reactions, such as that in tRNA splicing [3] and *in vitro* circularization of uncapped RNA [4]. It has three activities, 5'-hydroxyl kinase (hereafter abbreviated as "kinase" or symbolized by "K"), 2',3'-cyclic phosphate 3'-phosphodiesterase

("CPD" or "P"), and 5'-phosphate 2'-phosphate-3'-hydroxyl RNA ligase ("ligase" or "L") [5–7]. Another plant RNA ligase from *Arabidopsis* [8] has also been characterized well.

Yeast tRNA ligase [9] also has essentially the same three activities, although it has less than 20% sequence homology with the plant RNA ligases [4,8]. The yeast enzyme has its ligase, kinase, and CPD active sites in this order from the N-terminus to the C-terminus [10]. The border between the ligase and kinase domains has been identified: the ligase domain has its activity when it is physically isolated from the remaining C-terminal fragment, and the C-terminal fragment has its kinase and CPD activities when physically isolated from the ligase domain [10]. It is obvious from sequence comparisons that the plant enzymes also have their ligase, kinase, and CPD active sites in this order from the N-terminus, and the *Arabidopsis* enzyme has been dissected into the ligase and kinase-CPD fusion domains [11,12].

An attempt to locate the border between the kinase and CPD domains of the yeast enzyme was made with complementation tests: some combinations of truncated polypeptide fragments complemented a deletion mutant [10]. However, no autonomously active kinase domain has been isolated in spite of the extensive works, suggesting that the kinase domain needs to be tethered to one of the two flanking domains [10,13]. This raises a question why the fusion to only one of the ligase and CPD domain is sufficient for the kinase activity. It seems possible that the domains assembled into a complex containing all the three domains in the complementation tests, *i.e.*, the domains were not physically

\* Corresponding author at: Cell-free Science and Technology Research Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan. Fax: +81 89 927 9941.

E-mail address: [takai@eng.ehime-u.ac.jp](mailto:takai@eng.ehime-u.ac.jp) (K. Takai).

<sup>1</sup> Present address: Center for Eukaryotic Structural Genomics (CESG), Department of Biochemistry, University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA.

isolated with each other, even though they were present as separate polypeptides. If that was the case, we have to be careful in concluding that the identified kinase-CPD border is a true border that separates the kinase and CPD functions on the primary structure. In addition, not all that is known about the architecture of yeast tRNA ligase applies to plant RNA ligase, considering the narrower substrate specificity in the yeast enzyme than the plant ones [14] as well as the low sequence homology [4,8]. Therefore, a more extensive analysis of domain organization of plant RNA ligase may deepen the understanding of the mechanism of these eukaryotic RNA ligases and evolution of multidomain enzymes.

In this report, we apply a strategy utilizing cell-free protein synthesis aiming at rapidly segregating functional domains from the multidomain enzyme from wheat. We will discuss on possible interactions among the domains in wheat RNA ligase.

## 2. Materials and methods

Details in materials and methods are provided in Supplementary materials. The DNA sequences for many different enzyme fragments were prepared by PCR with a cDNA clone for wheat RNA ligase [4] as a template [15]. They were transcribed and translated in a wheat-germ cell-free translation system, and the products were purified and qualitatively assayed for the enzyme activities. The sequences for the protein fragments corresponding to the domains identified in this paper were subcloned into a cell-free expression vector pEU for large scale cell-free protein production. mRNA preparation and cell-free translation were performed essentially according to the previous paper [15] with some optimization. Enzyme activities were measured using RNA molecules with

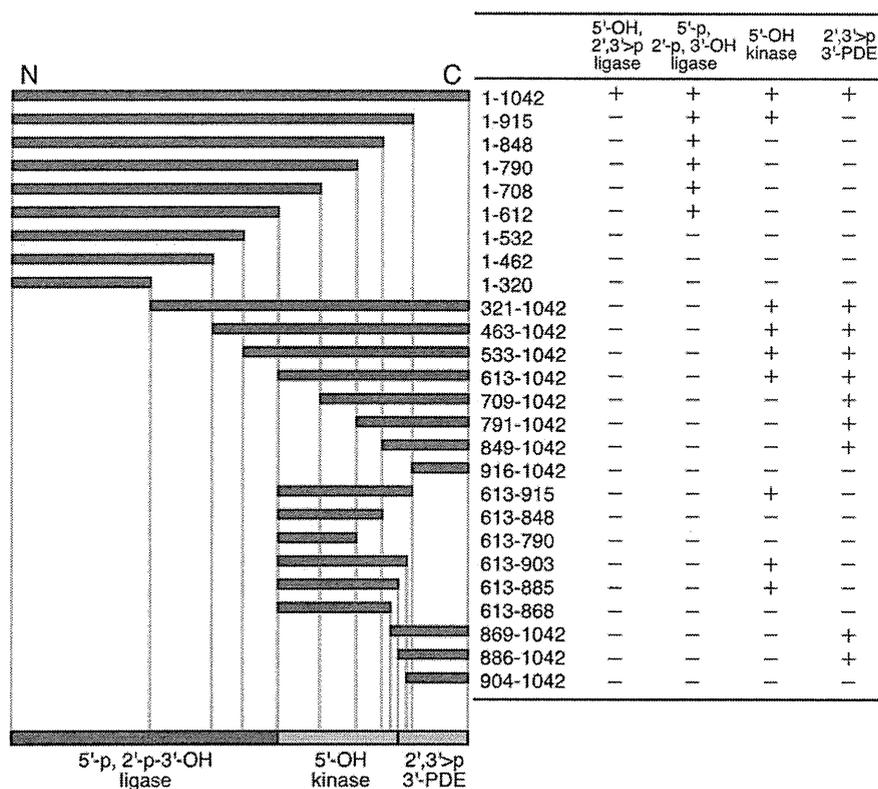
definite nucleotide sequences and terminal phosphate forms in a buffer containing 20 mM HEPES-KOH (pH 7.8), 250 mM NaCl, 80 mM potassium acetate, 2 mM magnesium acetate, 2 mM DTT, 0.01% (w/v) BSA, and 1 mM ATP when required.

## 3. Results

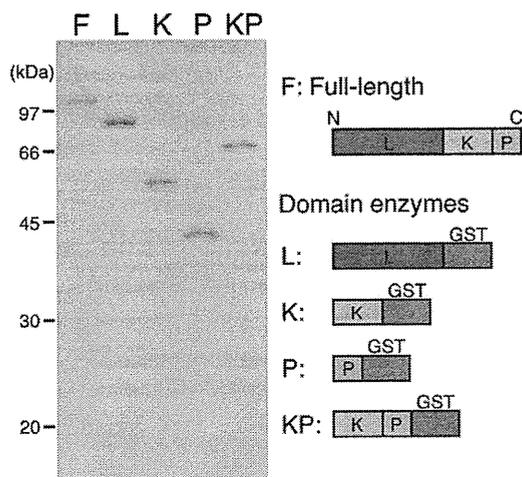
### 3.1. Construction of enzyme fragments and separation of the three activities

In order to locate the three functional domains of the wheat RNA ligase enzyme on its primary structure, we constructed various protein fragments as listed in Fig. 1. Each enzyme fragment was synthesized as an N-terminal GST-fusion by the wheat cell-free translation system without cloning into a plasmid (Fig. S1 in Supplementary materials). Initially we constructed a series of N-terminal and C-terminal truncations. The C-terminal 613–1042 fragment had the kinase and CPD activities. Those activities were detected by a shift in the electrophoretic mobility of an RNA molecule (Fig. S2) and were confirmed by measuring the change in the molecular mass by TOF-MS analyses. This fragment was useful for producing an RNA substrate with the 5'-phosphate and 3'-hydroxyl-2'-phosphate termini, which was used for the detection of the core ligase activity. The remaining N-terminal 1–612 fragment had the core ligase activity. This is consistent with the previous data on yeast and *Arabidopsis* enzymes [11]. The 1–612 and 613–1042 fragments are symbolized in this paper by L and KP, respectively.

In order to dissect the two end-modifying activities of the C-terminal half, we applied the same truncation strategy on KP. As a re-



**Fig. 1.** Enzyme fragments generated from the wheat RNA ligase gene and its activity. A schematic diagram shows each location of the fragments containing a part of the enzyme. Fragments are denoted on the right as a combination of the first and the last residue number from the N-terminus of the full-length enzyme. *In vitro* enzymatic activities of each fragment by purified enzymes shown in Fig. S1 are summarized on the right. The domain organization deduced from the activities is shown on the bottom bar.



**Fig. 2.** Proteins used for the biochemical assays. (left panel) An SDS-PAGE image of the domain proteins, with the positions of the size markers indicated on the left. 0.2  $\mu$ g of each protein was loaded. The gel was stained with CBB. (right panel) Schematic representation of the proteins consisting of the ligase (L), kinase (K), phosphodiesterase (P) domains, and the GST moiety.

sult, we located the CPD activity on the 886–1042 fragment (P), and the kinase activity on the 613–885 fragment (K) (Figs. S1 and S2).

### 3.2. Synthesis of the domain enzymes

We constructed the plasmids for cell-free expression of the individual functional domains described above as C-terminal GST-fusions. Each protein was successfully synthesized and purified (Fig. 2). These enzymes may be denoted in this paper by the fragment symbols as in the figure. The yields of the purified proteins were 0.88 mg/ml reaction (60 h) for L, 0.18 mg/ml reaction (36 h) for K, 0.66 mg/ml reaction (36 h) for P, and 0.45 mg/ml reaction (36 h) for KP. Each domain enzyme showed the same property as the respective N-terminal GST-fusion in the qualitative assays (data not shown). We also produced the full-length enzyme (F) in the same manner with the yield of 0.031 mg/ml reaction (36 h).

### 3.3. Comparisons of activities between the full-length enzyme and the domain enzymes

We then characterized the catalytic activities of the domain enzymes as compared with the full-length enzyme. The ligase activities (Fig. 3A and B) were measured by monitoring the product of the intermolecular ligation between two RNAs of defined sequences with the controlled terminal structures shown in Fig. 3A. The F enzyme showed a lag in the ligation reaction between the 5'-hydroxyl and 2',3'-cyclic phosphate ends, but not between the 5'-phosphate and 3'-hydroxyl-2'-phosphate ends (Fig. 3B). The L domain enzyme had a comparable specific activity to the F enzyme with 80% of substrate RNAs ligated with the equimolar donor/acceptor ratio. Surprisingly, the other domain enzymes had higher specific activities than those of the respective activities of the F enzyme (Fig. 3C and D). We observed an extra band close to that of the 5'-triphosphate RNA in Fig. 3A (center and right), on which we will report elsewhere.

## 4. Discussion

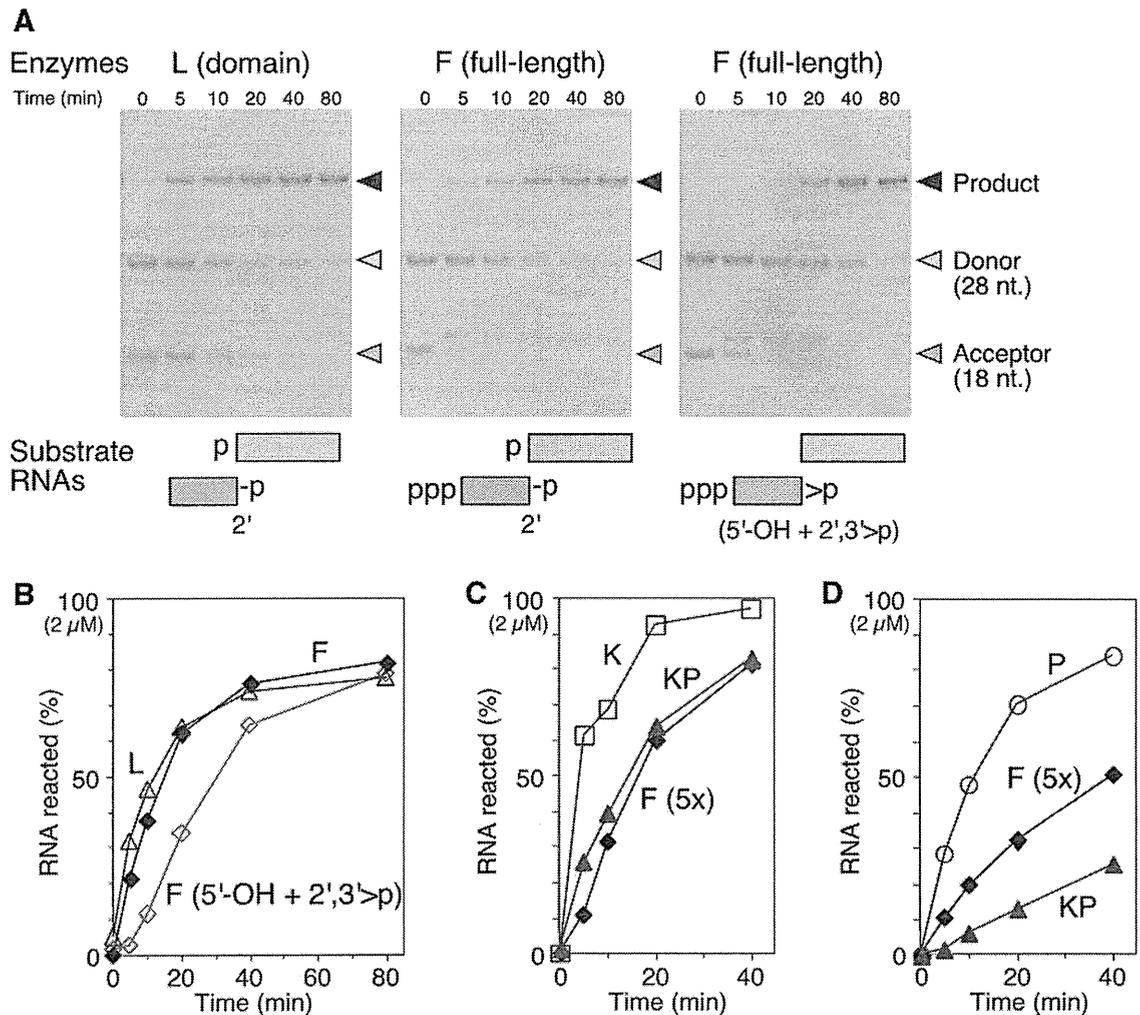
The present method for mapping functional domains on primary structures of multifunctional proteins through preparation of protein fragments is advantageous over other methods in that

it does not include time intensive cloning steps for each truncated construct. Preparation of protein fragments, in general, may suffer from difficulties because some incomplete parts of a domain may cause misfolding of another domain. The use of the wheat cell-free system may also have an advantage in this point, because it seems to have a higher ability to prevent inter-domain misfolding of multidomain proteins than a bacterial system, resulting in a high success rate in soluble protein production [16,17]. Moreover, we noticed another advantage of the wheat cell-free system by the fact that purified enzyme fragments as well as the domain enzymes showed little contamination of ribonucleases so that we were able to check the activity on RNA substrates.

It has been concluded that the kinase domain of yeast tRNA ligase requires to be tethered to at least one of the remaining two flanking domains. *Arabidopsis* RNA ligase has also been separated into an active ligase domain and an active kinase-CPD fusion domain, while it has not been tested if the latter domain can be separated into autonomously active kinase and CPD domains. Our data in Figs. 1 and 3 clearly show for the first time that plant RNA ligase has an autonomous and fully active kinase domain that does not overlap or depend on either the ligase or the CPD domain. In other words, the kinase domain enzyme is physically isolatable from the other two domains: it does not require the presence of either of the two domains for its kinase activity. The boundary between the kinase and CPD domains seems consistent with the previously reported one in yeast tRNA ligase [10]. Therefore, it might be possible that the yeast enzyme can also be dissected into the three isolatable enzymes. The domain activity might be affected by the method for preparation of the protein fragment and even by a single amino acid residue near the boundary of the fragments. On the other hand, our results might also be suggesting that plant RNA ligase is different from yeast tRNA ligase in the tertiary assembly of the domains to form the full-length enzyme. It is an enchanting hypothesis that the difference in the substrate specificity between the yeast and plant enzymes might reflect some difference in the three-dimensional spatial assembly of the domains [14]. It may be possible that selection pressures for the narrower substrate specificity for tRNA and the smaller size in the yeast enzyme have reduced the independence of the kinase domain from the other domains during molecular evolution. More intensive analyses on the yeast enzyme may be required for clearly solving this issue. Some bacteria and a baculovirus have related multifunctional enzymes, while they have a 2',3'-phosphatase activity instead of the CPD activity [18,19]. For none of these enzymes, the 5' and 3' end-modifying activities have been separated physically on the primary structure. T4 polynucleotide kinase, which is considered to provide a 5'-kinase and a 2',3'-phosphatase activities for T4 RNA ligase, is a homotetrameric protein [20] and is likely to be quite different from the kinase-CPD domain of the yeast and plant RNA ligases.

The time course experiments in Fig. 3 suggest that each domain enzyme is no less active than the corresponding domain in the full-length enzyme. Rather, it seems that the activities of the K and P domain enzymes are higher than the corresponding activities of the KP and F enzymes. It is thus likely that the interaction between the K and P domains suppresses the activity of each domain. We have also noticed that the presence of the tethered L domain seems to inhibit the kinase activity weakly. Nandakumar et al. [11] have reported that the ligase domain enzyme of *Arabidopsis* RNA ligase show substantial suppression on the activities of its kinase-CPD enzyme. While they deduced that the inhibition is due to the high affinity of the ligase domain to RNA substrates according to the report, it may be possible that the native full-length enzyme has suppressive interactions among the domains.

Native assembly of the domains composing a multidomain enzyme may, in general, be intuitively expected to provide a synergistic effect on its activity. By contrast, our proper dissection



**Fig. 3.** Activities of the full-length and domain enzymes. (A) Denaturing polyacrylamide gels showing the time courses of the intermolecular ligation reactions catalyzed by the ligase domain and the full-length enzymes. The terminal structures of the substrate RNAs are schematically drawn at the bottom. The substrate RNAs used were 18NTGCIPCPD and GP1PNK (left panel), 18NTGCPD and GP1PNK (center panel), and 18NTG and GP1 (right panel). The concentrations of the enzymes and each RNA substrate were 0.57  $\mu\text{M}$  (F), 0.70  $\mu\text{M}$  (L), and 2  $\mu\text{M}$  (RNA), respectively. The gels were stained with methylene blue. (B–D) The time courses of the reactions catalyzed by the full-length and each domain enzymes. The relative amounts of the substrate and product RNAs for the RNA ligase (B), 5'-kinase (C), and 2',3'-cyclic phosphate-3'-phosphodiesterase (D) reactions are plotted. The initial concentration of the substrate RNAs was 2  $\mu\text{M}$ . The enzyme concentrations were 0.57  $\mu\text{M}$  for F and 0.70  $\mu\text{M}$  for L in B, 0.43  $\mu\text{M}$  for F, 0.18  $\mu\text{M}$  for K, and 0.14  $\mu\text{M}$  for KP in (C), and 0.14  $\mu\text{M}$  for F, 0.076  $\mu\text{M}$  for P, and 0.045  $\mu\text{M}$  for KP in (D). Specific activities in mol (mol of enzyme) $^{-1}$  min $^{-1}$  calculated from the time courses are 0.11 for L and 0.13 for F (5'-p, 2'-p-3'-OH ligase) in (B), 0.80 for K, 0.56 for KP, and 0.15 for F in (C), and 1.3 for P, 0.28 for KP, and 0.28 for F in (D). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

enabled us to realize that the domain assembly into the multidomain RNA ligase caused a suppressive effect on the kinase and CPD domain activities. The embedded kinase and CPD domains may be well-regulated to solely meet the necessity for the following RNA ligation reaction. Our results shed new light on the way multidomain enzymes have been selected.

#### Acknowledgments

We are grateful to Michael Andy Goren, Department of Biochemistry, University of Wisconsin-Madison, for reading the manuscript. This work was in part supported by the Special Coordination Funds for Promoting Science and Technology by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.06.030.

#### References

- [1] W.J. Netzer, F.U. Hartl, Recombination of protein domains facilitated by co-translational folding in eukaryotes, *Nature* 388 (1997) 343–349.
- [2] E.V. Koonin, Y.I. Wolf, G.P. Karev, The structure of the protein universe and genome evolution, *Nature* 420 (2002) 218–223.
- [3] P. Gegenheimer, H.J. Gabius, C.L. Peebles, J. Abelson, An RNA ligase from wheat germ which participates in transfer RNA splicing *in vitro*, *J. Biol. Chem.* 258 (1983) 8365–8373.
- [4] S. Makino, T. Sawasaki, Y. Tozawa, Y. Endo, K. Takai, Covalent circularization of exogenous RNA during incubation with a wheat embryo cell extract, *Biochem. Biophys. Res. Commun.* 347 (2006) 1080–1087.
- [5] L. Pick, J. Hurwitz, Purification of wheat germ RNA ligase. I. Characterization of a ligase-associated 5'-hydroxyl polynucleotide kinase activity, *J. Biol. Chem.* 261 (1986) 6684–6693.
- [6] L. Pick, H. Furneaux, J. Hurwitz, Purification of wheat germ RNA ligase. II. Mechanism of action of wheat germ RNA ligase, *J. Biol. Chem.* 261 (1986) 6694–6704.
- [7] R.C. Schwartz, C.L. Greer, P. Gegenheimer, J. Abelson, Enzymatic mechanism of an RNA ligase from wheat germ, *J. Biol. Chem.* 258 (1983) 8374–8383.
- [8] M. Englert, H. Beier, Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins, *Nucleic Acids Res.* 33 (2005) 388–399.

- [9] C.L. Greer, C.L. Peebles, P. Gegenheimer, J. Abelson, Mechanism of action of a yeast RNA ligase in tRNA splicing, *Cell* 32 (1983) 537–546.
- [10] R. Sawaya, B. Schwer, S. Shuman, Genetic and biochemical analysis of the functional domains of yeast tRNA ligase, *J. Biol. Chem.* 278 (2003) 43928–43938.
- [11] J. Nandakumar, B. Schwer, R. Schaffrath, S. Shuman, RNA repair: an antidote to cytotoxic eukaryal RNA damage, *Mol. Cell* 31 (2008) 278–286.
- [12] L.K. Wang, B. Schwer, M. Englert, H. Beier, S. Shuman, Structure–function analysis of the kinase-CPD domain of yeast tRNA ligase (Trl1) and requirements for complementation of tRNA splicing by a plant Trl1 homolog, *Nucleic Acids Res.* 34 (2006) 517–527.
- [13] Q. Xu, D. Teplow, T.D. Lee, J. Abelson, Domain structure in yeast tRNA ligase, *Biochemistry* 29 (1990) 6132–6138.
- [14] C.L. Greer, D. Söll, I. Willis, Substrate recognition and identification of splice sites by the tRNA-splicing endonuclease and ligase from *Saccharomyces cerevisiae*, *Mol. Cell. Biol.* 7 (1987) 76–84.
- [15] T. Sawasaki, T. Ogasawara, R. Morishita, Y. Endo, A cell-free protein synthesis system for high-throughput proteomics, *Proc. Natl. Acad. Sci. USA* 99 (2002) 14652–14657.
- [16] N. Hirano, T. Sawasaki, Y. Tozawa, Y. Endo, K. Takai, Tolerance for random recombination of domains in prokaryotic and eukaryotic translation systems: limited interdomain misfolding in a eukaryotic translation system, *Proteins* 64 (2006) 343–354.
- [17] N. Goshima, Y. Kawamura, A. Fukumoto, A. Miura, R. Honma, R. Satoh, A. Wakamatsu, J. Yamamoto, K. Kimura, T. Nishikawa, et al., Human protein factory for converting the transcriptome into an *in vitro*-expressed proteome, *Nat. Methods* 5 (2008) 1011–1017.
- [18] A. Martins, S. Shuman, Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5'-kinase, and polynucleotide 3'-phosphatase activities, *J. Biol. Chem.* 279 (2004) 18220–18231.
- [19] A. Martins, S. Shuman, An end-healing enzyme from *Clostridium thermocellum* with 5' kinase, 2',3' phosphatase, and adenylyltransferase activities, *RNA* 11 (2005) 1271–1280.
- [20] H. Zhu, P. Smith, L.K. Wang, S. Shuman, Structure–function analysis of the 3' phosphatase component of T4 polynucleotide kinase/phosphatase, *Virology* 366 (2007) 126–136.

# Biotinylated-sortase self-cleavage purification (BISOP) method for cell-free produced proteins

Satoko Matsunaga<sup>1</sup>, Kazuhiro Matsuoka<sup>1</sup>, Kouhei Shimizu<sup>1</sup>, Yaeta Endo<sup>\*1,2,3,4</sup> and Tatsuya Sawasaki<sup>\*1,2,3,4</sup>

## Abstract

**Background:** Technology used for the purification of recombinant proteins is a key issue for the biochemical and structural analyses of proteins. In general, affinity tags, such as glutathione-S-transferase or six-histidines, are used to purify recombinant proteins. Since such affinity tags often interfere negatively with the structural and functional analyses of proteins, they are usually removed by treatment with proteases. Previously, Dr. H. Mao reported self-cleavage purification of a target protein by fusing the sortase protein to its N-terminal end, and subsequently obtained tag-free recombinant protein following expression in *Escherichia coli*. This method, however, is yet to be applied to the cell-free based protein production.

**Results:** The histidine tag-based self-cleavage method for purifying proteins produced by the wheat cell-free protein synthesis system showed high background, low recovery, and unexpected cleavage between the N-terminally fused sortase and target protein during the protein synthesis. Addition of calcium chelator BAPTA to the cell-free reaction inhibited the cleavage. In order to adapt the sortase-based purification method to the cell-free system, we next used biotin as the affinity tag. The biotinylated sortase self-cleavage purification (BISOP) method provided tag-free, highly purified proteins due to improved recovery of proteins from the resin. The N-terminal sequence analysis of the GFP produced by the BISOP method revealed that the cleavage indeed occurred at the right cleavage site. Using this method, we also successfully purified the E2 heterocomplex of USE2N and USE2v1. The c-terminal src kinase (CSK) obtained by the BISOP method showed high activity in phosphorylating the Src protein. Furthermore, we demonstrated that this method is suitable for automatically synthesizing and purifying proteins using robots.

**Conclusion:** We demonstrated that the newly developed BISOP method is very useful for obtaining high quality, tag-free recombinant proteins, produced using the cell-free system, for biochemical and structural analyses.

## Background

Technology used for purifying a recombinant protein has a significant impact on its biochemical function, structural properties, and other aspects, such as generating an antibody against the protein. Currently available, established purification methods generally attach an affinity tag to the N-terminus or C-terminus end of the target protein, and then recover the target protein by affinity chromatography [1]. Purification tags used today are classified into peptide-tags and protein-tags based on their nature and form. His-tag, a typical and globally the most

used peptide-tag, is highly effective in purifying the tagged protein by using immobilized metal affinity chromatography, such as nickel sepharose [2]. The glutathione-S-transferase (GST) tag, a typical protein-tag, has a high specific binding capacity for glutathione, and is, generally, known to have little or no effect on the activity of the fused recombinant protein [3]. In both cases, however, production of tag-free recombinant protein requires treatment with a protease, such as PreScission or TEV protease. Therefore, in order to purify a tag-free recombinant protein multiple chromatography steps are necessary.

Currently, there are only a few recombinant protein purification methods that combine affinity purification, cleavage, and separation of the fusion partner in one-step. One such one-step purification method was reported by Mao [4], in which the catalytic core of the transpeptidase

\* Correspondence: yendo@eng.ehime-u.ac.jp, sawasaki@eng.ehime-u.ac.jp

<sup>1</sup> The Cell-Free Science and Technology Research Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

<sup>2</sup> The Venture Business Laboratory, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

Full list of author information is available at the end of the article

sortase A (srtA, amino acid residues from 60-206), found in the cell envelope of *Staphylococcus aureus* [5,6], was used for fusion with the target protein. The SrtA enzyme is known to catalytically cleave the Thr-Gly bond of its recognition motif LPXTG (X is any amino acid) in the presence of calcium and triglycine [6-9]. Thus, the purified target protein eluted off the affinity column has only an extra Gly residue on the N-terminus end. This excellent approach was designed for the purification of recombinant proteins expressed in *Escherichia coli* cells. However, as discussed in that report [4], the fusion protein was partially self-cleaved during the expression, probably because of the difficulty in controlling the concentration of calcium in the living cells. Inability to suppress the srtA activity during the expression of the fused recombinant protein in *E. coli* was, therefore, a major limiting factor for using this otherwise excellent approach as a general tool for the production and purification of recombinant proteins.

At present, several types of cell-free protein production systems have been reported as alternative methods for obtaining recombinant proteins [see 10 and 11 for reviews]. In this regard, it is noteworthy that we are also developing wheat embryo based cell-free system for in vitro protein production [12-14]. Cell-free protein production is very flexible because it utilizes only the translational machinery of the cell without other factors, such as DNA replication and metabolic pathways, of the living system. The cell-free system, thus, could simply be modified by the addition or subtraction of reagents. In this study, we have adopted the self-cleavage activity of srtA in the wheat cell-free system for the production of tag-free recombinant proteins, and demonstrated an improved self-cleavage purification method by incorporating biotinylation reagents and calcium chelates in the cell-free synthesis system.

## Results and Discussion

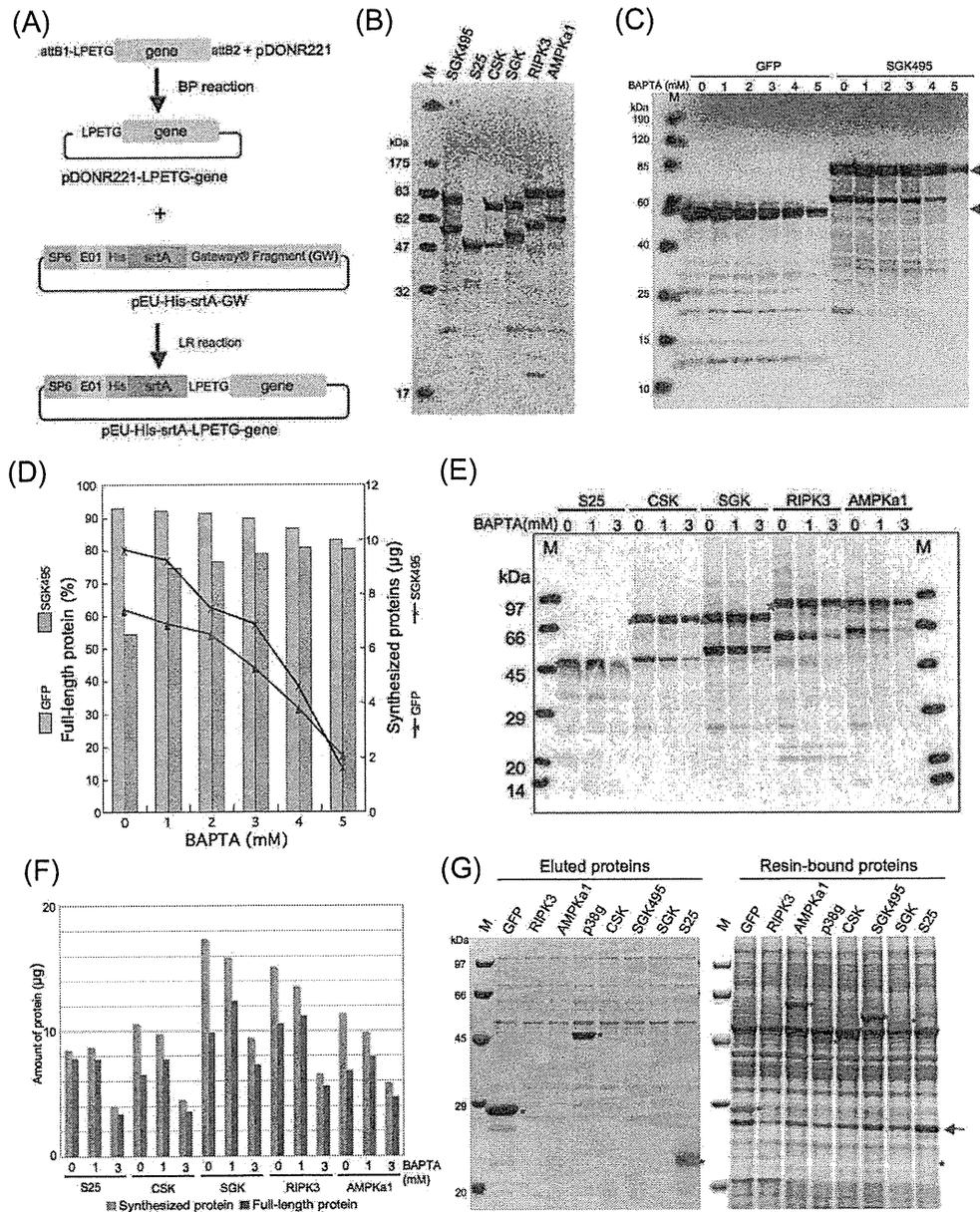
### Self-cleavage activity of srtA during the cell-free protein synthesis

First, based on the previous report, we constructed the expression vector pEU-His-srtA-GW by inserting the DNA fragment required for the Gateway (GW) recombination technology into the previously described wheat germ cell-free expression vector pEU-E01 [13]. The Gateway system allows easy recombination of the targeted genes. The srtA cleavage site, LPETG, was introduced in the forward PCR primer according to the instructions provided with the Gateway system (Figure 1A). To test the self-cleavage purification system, we selected the human protein kinases and malaria vaccine candidate Pfs25 (*Plasmodium falciparum* 25 kDa ookinete surface antigen precursor) as candidate proteins because they are very important proteins for practical use. PCR products

of the coding regions of Pfs25 and six protein kinases were amplified, and each PCR amplified fragment was individually inserted into the pDONR221 vector by BP recombination reaction. Subsequently, the inserted LPETG-gene fragment was cloned into the pEU-His-srtA-GW vector by LR recombination reaction, resulting in a pEU-His-srtA-LPETG-gene plasmid. These pEU-His-srtA-LPETG-gene plasmids were then used for <sup>14</sup>C-Leu-labeled protein synthesis using the wheat cell-free system. Unfortunately, in all cases, 20 to 40% of the synthesized proteins were cleaved during the cell-free synthesis (Figure 1B). The cleavage rate was dependent on the type of the gene used in creating the plasmid construct. For example, cleavage of Pfs25 (S25 lane in Figure 1B) and GFP (Figure 1C) during the protein synthesis process were very low, whereas almost 40% of the synthesized SGK495 protein was cleaved during the cell-free synthesis. Since calcium supplementation is known to induce the srtA activity [4,9], the cell-free system was treated with the calcium chelator, BAPTA (Figure 1C). Treatment with more than 4 mM BAPTA dramatically blocked the inexpedient cleavage of SGK495. However, protein synthesis was inhibited as the BAPTA concentration was increased (Figure 1D). For example, levels of SGK495 and GFP proteins produced in the presence of 5 mM BAPTA were approximately 20% and 30%, respectively, of their respective levels in the absence of BAPTA. In other proteins, the synthesis conditions in the presence of 1 and 3 mM BAPTA were investigated (Figure 1E and 1F). By comparing the results of the BAPTA-concentration dependent blocking of the srtA activity and inhibition of the protein synthesis, we concluded that 1 mM BAPTA is optimal for the cell-free synthesis of srtA-fusion proteins, because at this concentration of BAPTA the recovery of all full-length srtA fusion proteins improved by approximately 10-20% without any major inhibition of the protein synthesis.

### Self-cleavage purification of His-srtA-fusion proteins

A total of eight plasmid constructs, each containing a different gene (generalized here as pEU-His-srtA-LPETG-gene), were used for the cell-free protein production and self-cleavage purification studies, and the results are shown in Figure 1G. Out of eight proteins, three proteins, GFP, p38g and Pfs25, clearly eluted from a nickel-nitrilotriacetic acid (Ni-NTA) sepharose column, whereas other five proteins could not be recovered in the eluted fraction (Figure 1G, left panel). To confirm protein synthesis and self-cleavage, column resins of all samples were analyzed by SDS-PAGE after boiling with SDS-sample buffer (Figure 1G, right panel). Surprisingly, even though AMPKa1, CSK, SGK and SGK495 were synthesized and self-cleaved on the column resin, their cleaved forms were not eluted off the resin. We could not find the RIPK3 protein in the



**Figure 1 Synthesis of srtA-fusion proteins using the wheat germ cell-free system.** **A.** Schematic representation of the pEU-His-srtA-LPETG-Gene plasmids created using the Gateway system. **B.** Autoradiogram of SDS-PAGE of proteins synthesized using the cell-free system in the presence of [<sup>14</sup>C] Leu. Lane M, Protein MW standards labeled by using [<sup>14</sup>C]-containing feld pen. **C.** Autoradiogram of [<sup>14</sup>C] Leu incorporated GFP and SGK495 proteins synthesized by the wheat cell-free system in the presence of the Ca<sup>2+</sup> chelating reagent BAPTA. The number represents concentration (mM) of BAPTA used in the protein synthesis reaction. Arrowheads denote the sizes of the full-length proteins. **D.** Rate of synthesis of the full-length protein and productivity of GFP (pink-colored bar and red-colored line) and SGK495 (purple-colored bar and blue-colored line) in the presence of different concentrations of BAPTA. **E.** Autoradiogram of [<sup>14</sup>C]-Leu incorporated proteins synthesized by the cell-free system in the presence of BAPTA. Asterisk denotes the sizes of the full-length proteins. **F.** Rate of synthesis of the full-length protein and productivity of proteins in the presence of different concentrations of BAPTA. Productivities of total synthesized and full-length proteins indicated as blue and red bars respectively. **G.** Purifications of proteins by the cell-free synthesis using the pEU-His-srtA-LPETG-Gene plasmid constructs. CBB-stained protein bands on the SDS-PAGE gel of the eluted (left panel) and resin-bound (right) target proteins are indicated using asterisk. Arrow represents the cleaved His-tagged srtA. Lane M (both panels): Protein MW standards.

eluted or in the resin-bound fraction, suggesting that this protein was expressed at a very low level. In addition, purity of the protein in the eluted fraction was not high, as there were several similar protein contaminants in every lane. Furthermore, many proteins bound to the Ni-NTA resin (right panel in Figure 1G). These results suggested that further technical improvements were necessary to achieve high quality purified proteins with high efficiency from the cell-free based His-srtA system.

#### **Biotinylated sortase self-cleavage purification (BISOP)**

Recently we successfully adapted the biotin-labeling system of *E. coli* biotin ligase and biotin to the wheat cell-free protein synthesis system [15]. The biotinylation reaction modifies a specific lysine residue at the biotin ligation site (bls: GLNDIFEAQKIEWHE, the underline is the ligation site). The biotinylated proteins could be directly used for an assay without further purification because of very low biotin concentration. Since the His-tag based approach showed many contaminated proteins in the eluted fraction of the Ni-NTA column, we next used the biotin-labeled tag for protein purification. For this purpose, we constructed the pEU-BISOP-LPETG-GFP plasmid based on the pEU-His-srtA-LPETG-GFP vector as shown in Figure 1A and 2A. Next, we compared whether the His-tagged or the biotin-tagged protein could be better purified by the sortase self-cleavage method, processed either manually or using automated robots to eliminate any human error. Clearly, on the CBB-stained SDS-PAGE, a single major protein band (with low background) was found in the eluted fraction following the sortase self-cleavage of the biotin-tagged srtA fusion protein; in contrast, contaminating proteins were found along with the major protein band in the eluted fraction of the self-cleaved His-tagged srtA fusion protein (Figure 2B). Reaction performance of robot for His-srtA-LPETG-GFP or BISOP-LPETG-GFP was 63 or 58% of recovery, 52 or 88% of purity, and 68 or 62  $\mu$ g of yield respectively. These results suggest that the BISOP method is better suited for producing tag-free purified proteins by the cell-free system.

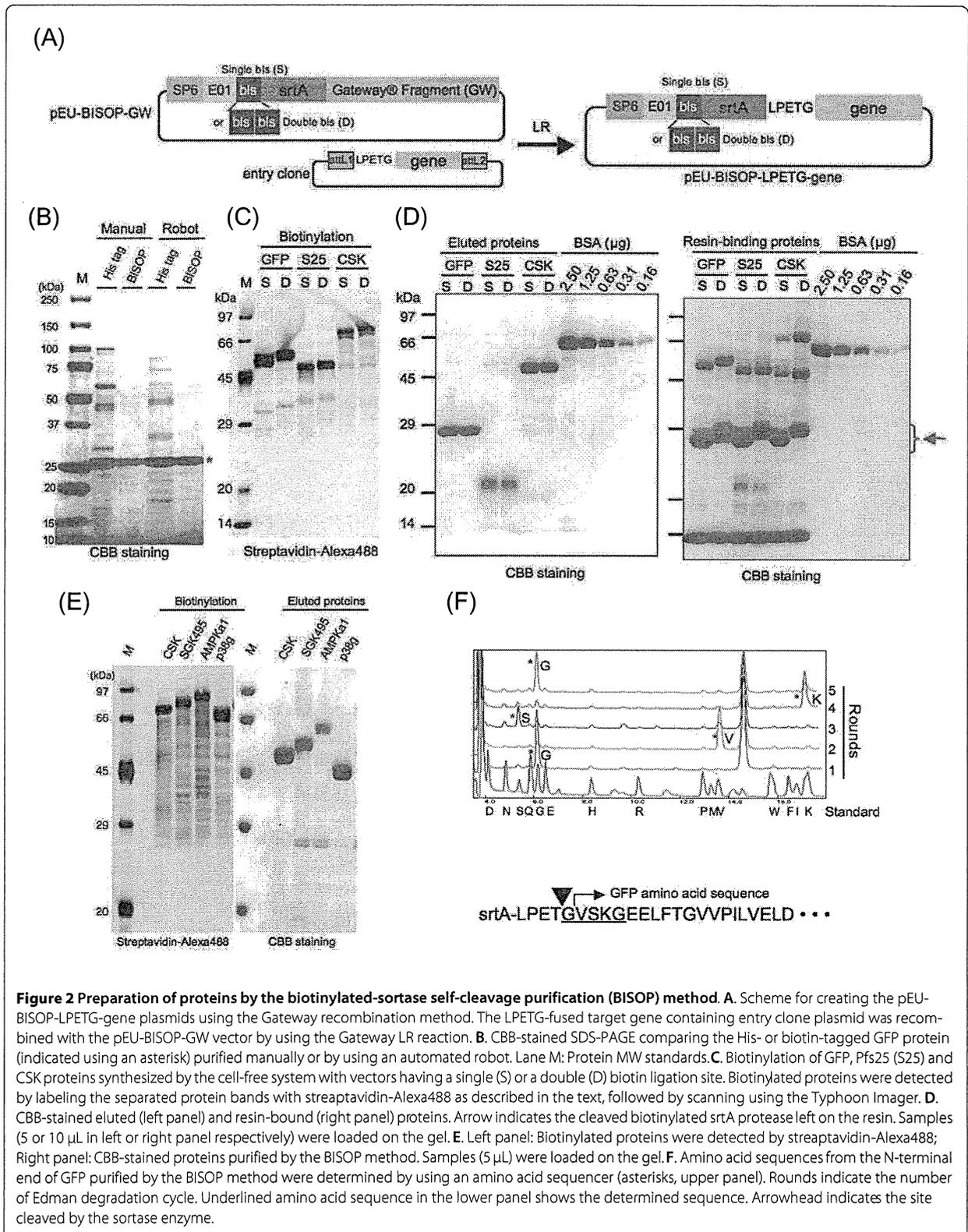
Next, to examine whether the number of bls has any effect on the purification, two vectors, one having a single bls and the other having double bls, were constructed (Figure 2A). Coding regions of GFP, Pfs25 and CSK were individually cloned into each one of these two vectors, and then the resultant recombinant plasmids were used for the protein synthesis using the cell-free system. Staining with Alexa488-labeled streptavidin revealed slight mobility shifts for proteins containing double bls, as compared to those containing single bls (Figure 2C). We did not observe any difference between the single and double bls containing proteins with respect to biotinylation, elution and resin binding characteristics. Similar to the self-

cleavage purification of GFP (shown in Figure 2B), all three eluted proteins were highly purified (left panel in Figure 2D). The higher purity of proteins obtained using the BISOP method might be due to the presence of very few contaminating proteins on the streptavidin-conjugated resin (compare right panel in Figure 2D with Figure 1G). Interestingly, the CSK protein, prepared by the BISOP method was eluted of the resin, whereas with the His-tag based method it was not found in the eluted fraction (Figure 1G). Total amounts of purified GFP, S25 and CSK by the BISOP method were 44, 37, and 55  $\mu$ g per reaction respectively. Also full-length GFP, S25 and CSK proteins remaining on the column were approximately 6, 3 and 3  $\mu$ g respectively, and cleaved S25 and CSK proteins on the column were 3 and 3  $\mu$ g respectively. These data means that rate of target proteins remaining on the column was approximately 15%. Next, the BISOP method was used for purifying several other proteins. For this purpose, four protein kinase genes were individually inserted into the Gateway system vector pEU-BISOP-GW following the procedures described above, and the results are shown in Figure 2E. Both SGK495 and AMPK $\alpha$ 1, which were not recovered earlier from the resin when the His-tag based method was used (Figure 1G), were also purified by the BISOP method (Figure 2E).

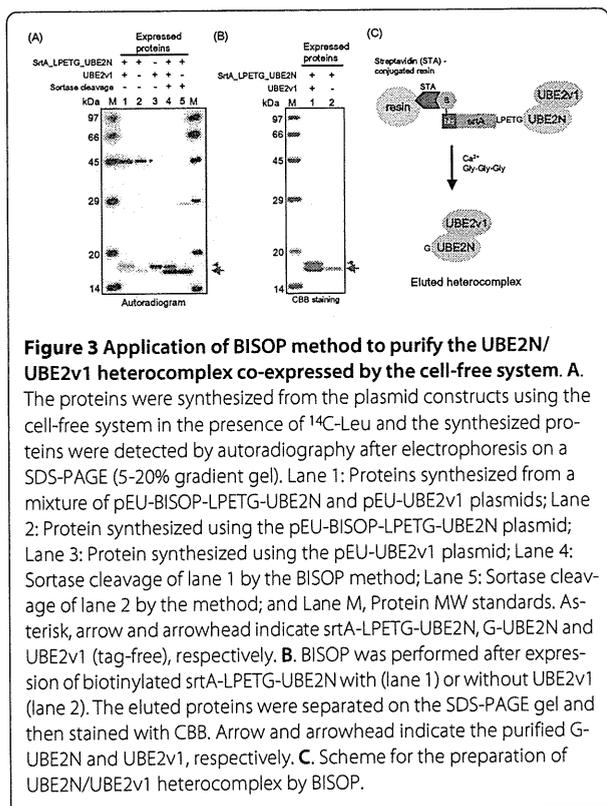
Total amounts of purified CSK, SGK495, AMPK $\alpha$ 1 and p38g by the BISOP method were 55, 42, 24 and 57  $\mu$ g per reaction respectively. In addition, analysis of the N-terminal sequence of the GFP protein purified by the BISOP revealed the expected cleavage of the Thr-Gly bond of the inserted LPETG sortase-recognition site (Figure 2F). These results suggested that the BISOP method would be suitable for the purification of the cell-free produced proteins with high efficiency and purity.

#### **Purification of E2 heterocomplex by BISOP**

Analysis of protein complex is one of important targets for their structural and biochemical analysis. Thus, next we examined whether a protein heterocomplex, co-expressed using the wheat cell-free system, could be purified by the BISOP method. To test this notion, we next co-expressed UBE2N and UBE2v1, two proteins forming the heterodimer complex of the ubiquitin-conjugate (E2) enzyme [16], using the BISOP method. Specific complex formation between these two proteins produced by the cell-free system has already been reported [15]. At first, we confirmed co-expression of the biotinylated srtA-LPETG-UBE2N and UBE2v1 (tag-free form) by incorporating  $^{14}$ C-Leu during the cell-free synthesis (Figure 3A). Notably, we recovered the E2 heterocomplex consisting of UBE2N and UBE2v1 when both UBE2v1 and srtA-LPETG-UBE2N were co-expressed using the cell-free system and then purified by the BISOP method (Figures 3B and 3C). The protein band corresponding to UBE2v1



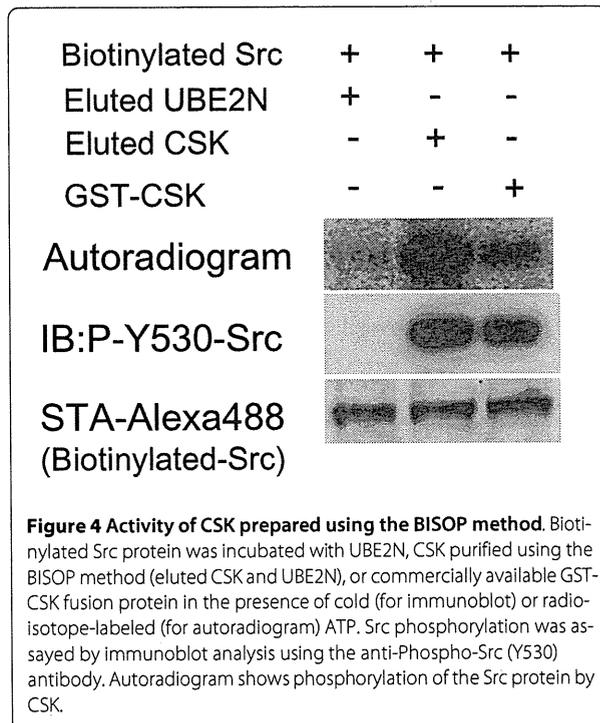
**Figure 2 Preparation of proteins by the biotinylated-sortase self-cleavage purification (BISOP) method.** **A.** Scheme for creating the pEU-BISOP-LPETG-gene plasmids using the Gateway recombination method. The LPETG-fused target gene containing entry clone plasmid was recombined with the pEU-BISOP-GW vector by using the Gateway LR reaction. **B.** CBB-stained SDS-PAGE comparing the His- or biotin-tagged GFP protein (indicated using an asterisk) purified manually or by using an automated robot. Lane M: Protein MW standards. **C.** Biotinylation of GFP, Pfs25 (S25) and CSK proteins synthesized by the cell-free system with vectors having a single (S) or a double (D) biotin ligation site. Biotinylated proteins were detected by labeling the separated protein bands with streptavidin-Alexa488 as described in the text, followed by scanning using the Typhoon Imager. **D.** CBB-stained eluted (left panel) and resin-bound (right panel) proteins. Arrow indicates the cleaved biotinylated sr1A protease left on the resin. Samples (5 or 10  $\mu$ L in left or right panel respectively) were loaded on the gel. **E.** Left panel: Biotinylated proteins were detected by streptavidin-Alexa488; Right panel: CBB-stained proteins purified by the BISOP method. Samples (5  $\mu$ L) were loaded on the gel. **F.** Amino acid sequences from the N-terminal end of GFP purified by the BISOP method were determined by using an amino acid sequencer (asterisks, upper panel). Rounds indicate the number of Edman degradation cycle. Underlined amino acid sequence in the lower panel shows the determined sequence. Arrowhead indicates the site cleaved by the sortase enzyme.



was however not found when the biotinylated srtA-LPETG-UBE2N was expressed alone. Therefore, this result suggests that the BISOP method would be useful for purification of protein complexes produced by the cell-free system.

#### Activity of CSK purified by the BISOP method

It is very important that the purification method provide functionally active protein. We, therefore, investigated whether the CSK protein purified using the BISOP method could specifically phosphorylate the Tyr-530 residue of human Src protein (Swiss-Prot no. P12931) [17]. For this purpose, biotinylated Src was synthesized by the cell-free system and the synthesized protein was partially purified using the magnetic streptavidin-conjugated beads. The bead-bound Src was then treated with [<sup>32</sup>P]-labeled or unlabeled ATP and the CSK protein that was purified by the BISOP method. Both autoradiogram of the SDS-PAGE separated proteins from the reaction mix containing [<sup>32</sup>P]-labeled ATP and immunoblot analysis of the SDS-PAGE separated proteins from the reaction mix containing the unlabeled ATP using the specific anti-phospho-Src antibody showed specific phosphorylation of the Y530 residue of Src by CSK (Figure 4). These results suggest that BISOP would be suitable for the in vitro synthesis of active proteins.



#### Conclusion

The cell-free system is an easy to use method for synthesizing recombinant proteins. In this study, we have developed a new biotinylated-sortase self-cleavage purification (BISOP) method to achieve high quality purification of several proteins, including a protein heterocomplex, without any affinity tag. In addition, we showed that this method could be automated using robots. Results of this study indicate that the combination of the cell-free protein synthesis system and BISOP could provide a simple and easy method for the preparation of high quality recombinant proteins without any affinity tag. Since the cell-free system is suitable for high-throughput protein production, this combined method could also be utilized at the genome-wide level.

#### Methods

##### General

The following procedures have been either described in detail or cited [13,15,18,19]: isolation of the wheat germs and preparation of the extract, generation of the DNA template by polymerase chain reaction (PCR) using split-primers, parallel synthesis of mRNA and protein, estimation of the amount of protein synthesized by densitometric scanning of the Coomassie brilliant blue (CBB)-stained band and autoradiogram of radiolabeled-proteins, and detection of biotinylated proteins using Alexa488-conjugated streptavidin (Invitrogen) and the

Typhoon Imager (GE Healthcare) fitted with 532 nm laser and 526 emission filter.

#### Template genes

The cDNA clone of the malarial parasite *Plasmodium falciparum* 25 kDa ookinete surface antigen precursor (Pfs25) was kindly provided by Dr. Tsuboi (Cell-free Research and Technology Center, Ehime University). cDNAs of GFP, UBE2N (GenBank accession no. [BC003365](#)), and UBE2v1 (GenBank accession no. [BC000468](#)) were described in our previous reports [15,18]. Mammalian gene collection (MGC) cDNA clones of CSK (BC104875), SGK (BC001263), SGK495 (BC007835), AMPK $\alpha$ 1 (BC048980), p38 $\gamma$  (BC015741), RIPK3 (BC062584) and Src (BC011566) were also used in this study.

#### Construction of the srtA-based self-cleavage vector

DNA fragment encoding the mature-form of srtA (corresponding to amino acids 60-206, GenBank accession no. [AF162687](#)) was artificially synthesized and inserted into the pUC57 to create the plasmid pUC57-srtA-EcoRV-SpeI by the GenScript Corporation (Boston). The pUC57-srtA-EcoRV-SpeI plasmid as used as a template to amplify the mature srtA fragment (DraI-His or bls-srtA-EcoRV-SpeI) by PCR using the following pair of primers: M13(M3) (5'-GTAAAACGACGGCCAGT) and DraI-A-His-srtA (5'-GAGATTTAAATGGCCAGCAGC CATCACCATCACCATCATAGCAGCGGCCTGGTGC CGC) or M13(M3) and DraI-A-bls-srtA (5'-GAGATTTAAATGGCCAGCAGCGGCCTGAACGACATCTTC GAGGCCAGAAGATCGAGTGGCACGAAAGCAGC G GCCTGGTGCCGC). The DraI-A-His-srtA primer included 6  $\times$  His-tag codons and a DraI restriction enzyme site, and the DraI-A-bls-srtA primer included a bls (biotin ligase site) recognition sequence and a DraI restriction enzyme site. After digestion with DraI and SpeI enzymes each fragment was inserted into the EcoRV and SpeI sites in pEU-E01-MCS provided by CellFree Sciences, Ltd <http://www.cfsciences.com/eg/index.html> to create pEU-His-srtA-MCS and pEU-BISOP-MCS plasmids, respectively. To create the Gateway'-based plasmids pEU-His-srtA-GW and pEU-BISOP-GW, DNA fragment needed for Gateway' recombination cloning technology (Invitrogen) was inserted into the EcoRV site of the pEU-His-srtA-MCS and pEU-BISOP-MCS plasmids. PCR reaction was performed by PrimeStar enzyme according to instruction (Takara Bio, Otsu, Japan). Nucleotide sequences of the DNA inserts in all plasmid constructs were subsequently confirmed by using the ABI PRISM 310 Genetic Analyzer using the BigDye terminator v1.1 Cycle sequence kit (Applied Biosystems, Foster City, CA).

#### Plasmid construction for the cell-free protein production

We introduced the srtA self-cleavage site (DNA encoding for the amino acid sequence LPETG) into the recombinant plasmid construct for the cell-free production of proteins. The DNA fragments coding the respective protein were amplified by PCR using two gene specific primers: forward primer attB1-LPETG-Gene (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-**CTGCCCGAGACCGG**Catg(n)<sub>19</sub>; uppercase, lower-cased and bold sequences indicated common, gene specific and LPETG sequences, respectively; n represent gene specific sequence) and reverse primer attB2-Gene (5'-GGGGACCACTTTGTACAAGAAAGCTGGTcxxxxnnnnnnnnnnnnnnnnnn; xxx is the complementary sequence of the stop codon; n represent gene specific sequence). PCR reaction was performed using the PrimeStar enzyme (Takara Bio, Otsu, Japan) and following the supplier's instructions. The amplified attB1-LPETG-Gene-attB2 fragments were inserted into the donor vector pDONR221 by BP reaction to generate the entry plasmids. The LPETG-fusion Gene in the entry plasmid was transferred to the pEU-His-srtA-GW or pEU-BISOP-GW by LR reaction to generate the pEU based-plasmid clones. BP and LR reactions were carried out according to the instructions provided by the supplier of the reagents (Invitrogen, Carlsbad, CA). Nucleotide sequences of the DNA inserts in all the plasmid constructs were subsequently confirmed by using the ABI PRISM 310 Genetic Analyzer described method above.

#### Cell-free protein production

For the cell-free protein production, we employed the wheat germ cell-free protein expression system using the bilayer translation method described previously [18,19]. Cell-free protein production was carried out using the ENDEXT' Wheat Germ Expression S Kit and according to the instructions provided by the supplier (CellFree Sciences Co., Ltd., Matsuyama, Japan). Briefly, 250  $\mu$ L of transcriptional mixture [80 mM HEPES-KOH, pH 7.8, 16 mM magnesium acetate, 2 mM spermidine, 10 mM DTT, 2.5 mM NTP mix, 1 U/ $\mu$ L SP6 RNA polymerase (Promega, Madison, WI), 1 U/ $\mu$ L RNase inhibitor, RNasin (Promega), and 100 ng/ $\mu$ L undigested plasmid DNA] was incubated at 37°C for 6 h, and then mixed with 250  $\mu$ L of wheat embryo extract (120 A 260/mL, CellFree Sciences Co., Ltd.) and 1  $\mu$ L of 20 mg/mL creatine kinase (Roche Applied Science, Indianapolis, IN). This mixture, called the translational mixture, was then carefully transferred to the bottom of the well of a 6-well tissue culture plate (Whatman Inc., Clifton, NJ) that already contained 5.5 mL of TSB (30 mM HEPES-KOH, pH 7.8, 100 mM potassium acetate, 2.7 mM magnesium acetate, 0.4 mM spermidine, 2.5 mM DTT, 0.3 mM amino acid mix, 1.2 mM ATP, 0.25 mM GTP and 16 mM creatine phosphate) by

inserting the pipette tip down to the bottom of the well, thereby creating two distinct layers. The plate was then covered with the sealing film, and was incubated at 17°C for 18 hr without shaking. For the calcium chelating experiment, 1 mM BAPTA (1,2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid) (Sigma-Aldrich, St Louis, MO) was added to both the translational mixture and TSB. For biotin labeling of proteins [15,20], 2 µg of biotin protein ligase (BirA, GenBank accession no. NP\_0312927) produced by the wheat cell-free system and 6 µM D-biotin (Nacalai Tesque, Kyoto, Japan) were added to the bottom translational mixture.

For co-expression of the biotinylated srtA-LPETG-UBE2N and UBE2v1, the cell-free production method was slightly modified. Each translation mixture was prepared independently and pre-incubated at 26°C for 30 min, following which they were mixed, and were subsequently used in the bilayer translation reaction as described above.

#### Self-cleavage purification of proteins produced by the cell-free system

Reaction mixture (6 mL) from the cell-free expression system described above was mixed with 100 µL of Ni-NTA sepharose (GE Healthcare) or 100 µL of streptavidin-sepharose (GE Healthcare). The sepharose beads were pre-equilibrated with phosphate buffered saline (PBS) for 4 to 6 hours (His tag) or 30 min (BISOP) at 4°C. Sepharose bead-captured srtA-fusion proteins were collected by centrifugation (3,000 × g) and the beads were washed three times with PBS buffer. Self-cleavage purification of the target protein was performed by incubation of the beads with 100 µL of Elution buffer [20 mM Tris-HCl (pH 7.5), 5 mM Tri-Gly (Sigma-Aldrich), 5 mM CaCb, 150 mM NaCl, 1 mM DTT and 2% glycerol] for 4 hours at 16°C. The buffer-bead mixture was then transferred into a micro spin-column (GE Healthcare) and the eluted fraction was recovered by flash centrifuge (3,000 × g) at 4°C. The N-terminal end of the eluted GFP was determined by amino acid sequence analysis using the Applied Biosystems ABI 473A protein sequencer and according to the instructions provided by Applied Biosystems.

The cell-free protein production and self-cleavage purification were also carried out using an automatic robot, Protomist DTII (CellFree Sciences Co., Ltd.), basically according to manufacturer's instructions. Addition of biotinylation reagents and BAPTA were also carried out as mentioned above.

#### Phosphorylation assay

The phosphorylation assay was mainly performed according to the previous published methods [13,21]. To assay for phosphorylation of the biotinylated Src by CSK, 40 µl of the reaction mixture was mixed with 15 µl of bio-

tin magnetic beads (Promega, MI), and was washed twice with PBS buffer and once with protein kinase (PK) buffer [50 mM Tris-HCl (pH 7.6), 500 mM potassium acetate, 50 mM MgCl<sub>2</sub>, 0.1 mM DTT]. Beads were suspended in 10 µl PK buffer supplemented with cold ATP (for immunoblot) or [ $\gamma$ -<sup>32</sup>P]-ATP (for autoradiogram), and CSK purified using the BISOP method, UBE2N or commercially available GST-CSK fusion protein (purchased from Carna Biosciences Inc., Kobe) was added in the reaction mixture. The mixtures were incubated at 37°C for 30 min, following which they were boiled in the SDS-denaturing buffer and the proteins were separated on 12.5% SDS-polyacrylamide gel. Autoradiogram of Src phosphorylation was analyzed by BAS-2500 (FUJIFILM, Tokyo, Japan). For immunoblot analysis, proteins were transferred from the SDS-PAGE gel to PVDF membrane (Millipore Bedford, MA, USA) following standard procedures. The blots were then processed using the Immobilon Western detection reagents (Millipore) and antibody against phosphorylated Src (Y527) or Src (Cell Signaling Technology, Beverly, MA) according to the manufacturer's procedure. The anti-phospho-Src (Y527) antibody recognizes the phosphorylated Y530 residue in human Src.

#### Abbreviations

srtA: sortase SrtA; GFP: green fluorescent protein; SGK: serum/glucocorticoid regulated kinase 1; CSK: c-src tyrosine kinase; UBE2N: ubiquitin-conjugating enzyme E2N; UBE2v1: ubiquitin-conjugating enzyme E2 variant 1; AMPKa1: 5'-AMP-activated protein kinase alpha 1 catalytic subunit; PRKAA1: protein kinase, AMP-activated: alpha 1 catalytic subunit; MAPK12: mitogen-activated protein kinase 12; p38g: p38 gamma; Pfs25: *Plasmodium falciparum* 25 kDa ookinete surface antigen precursor; BAPTA 1: 2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid; PBS: phosphate buffered saline.

#### Authors' contributions

SM conceived the study and performed some of the experiments; KM and KS performed also participated in performing the experiments; YE conceived the study and supervised the work; TS conceived and designed the study, supervised the work and contributed to writing the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

We greatly thank Naoki Nakagawa for technical assistance. This work was partially supported by the Special Coordination Funds for Promoting Science and Technology by the Ministry of Education, Culture, Sports, Science and Technology, Japan (19657041 and 22310127 to TS).

#### Author Details

<sup>1</sup>The Cell-Free Science and Technology Research Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan, <sup>2</sup>The Venture Business Laboratory, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan, <sup>3</sup>Proteo-Medicine Research Center, Ehime University, Toon, Ehime 791-0295, Japan and <sup>4</sup>RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

Received: 10 January 2010 Accepted: 4 June 2010

Published: 4 June 2010

#### References

1. Nygren PA, Stahl S, Uhlen M: Engineering proteins to facilitate bioprocessing. *Trends Biotechnol* 1994, **12**(5):184-188.

2. Janknecht R, de Martynoff G, Lou J, Hipskind RA, Nordheim A, Stunnenberg HG: **Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus.** *Proc Natl Acad Sci USA* 1991, **88**(20):8972-8976.
3. Smith DB, Johnson KS: **Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.** *Gene* 1988, **67**(1):31-40.
4. Mao H: **A self-cleavable sortase fusion for one-step purification of free recombinant proteins.** *Protein Expr Purif* 2004, **37**(1):253-263.
5. Mazmanian SK, Liu G, Ton-That H, Schneewind O: **Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall.** *Science* 1999, **285**(5428):760-763.
6. Ton-That H, Mazmanian SK, Faull KF, Schneewind O: **Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates.** *J Biol Chem* 2000, **275**(13):9876-9881.
7. Ilangovan U, Ton-That H, Iwahara J, Schneewind O, Clubb RT: **Structure of sortase, the transpeptidase that anchors proteins to the cell wall of Staphylococcus aureus.** *Proc Natl Acad Sci USA* 2001, **98**(11):6056-6061.
8. Ton-That H, Mazmanian SK, Alksne L, Schneewind O: **Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis.** *J Biol Chem* 2002, **277**(9):7447-7452.
9. Mao H, Hart SA, Schink A, Pollok BA: **Sortase-mediated protein ligation: a new method for protein engineering.** *J Am Chem Soc* 2004, **126**(9):2670-2671.
10. Endo Y, Sawasaki T: **Cell-free expression systems for eukaryotic protein production.** *Curr Opin Biotechnol* 2006, **17**(4):373-380.
11. He M: **Cell-free protein synthesis: applications in proteomics and biotechnology.** *N Biotechnol* 2008, **25**(2-3):126-132.
12. Madin K, Sawasaki T, Ogasawara T, Endo Y: **A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes.** *Proc Natl Acad Sci USA* 2000, **97**(2):559-564.
13. Sawasaki T, Ogasawara T, Morishita R, Endo Y: **A cell-free protein synthesis system for high-throughput proteomics.** *Proc Natl Acad Sci USA* 2002, **99**(23):14652-14657.
14. Madin K, Sawasaki T, Kamura N, Takai K, Ogasawara T, Yazaki K, Takei T, Miura K, Endo Y: **Formation of circular polyribosomes in wheat germ cell-free protein synthesis system.** *FEBS Lett* 2004, **562**(1-3):155-159.
15. Sawasaki T, Kamura N, Matsunaga S, Saeki M, Tsuchimochi M, Morishita R, Endo Y: **Arabidopsis HYS protein functions as a DNA-binding tag for purification and functional immobilization of proteins on agarose/DNA microplate.** *FEBS Lett* 2008, **582**(2):221-228.
16. Zhang M, Windheim M, Roe SM, Peggie M, Cohen P, Prodromou C, Pearl LH: **Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex.** *Mol Cell* 2005, **20**(4):525-538.
17. Nada S, Okada M, MacAuley A, Cooper JA, Nakagawa H: **Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of p60c-src.** *Nature* 1991, **351**(6321):69-72.
18. Sawasaki T, Hasegawa Y, Tsuchimochi M, Kamura N, Ogasawara T, Kuroita T, Endo Y: **A bilayer cell-free protein synthesis system for high-throughput screening of gene products.** *FEBS Lett* 2002, **514**(1):102-105.
19. Sawasaki T, Morishita R, Gouda MD, Endo Y: **Methods for high-throughput materialization of genetic information based on wheat germ cell-free expression system.** *Methods Mol Biol* 2007, **375**:95-106.
20. Takahashi H, Nozawa A, Seki M, Shinozaki K, Endo Y, Sawasaki T: **A simple and high-sensitivity method for analysis of ubiquitination and polyubiquitination based on wheat cell-free protein synthesis.** *BMC Plant Biol* 2009, **9**:39.
21. Sawasaki T, Hasegawa Y, Morishita R, Seki M, Shinozaki K, Endo Y: **Genome-scale, biochemical annotation method based on the wheat germ cell-free protein synthesis system.** *Phytochemistry* 2004, **65**(11):1549-1555.

doi: 10.1186/1472-6750-10-42

**Cite this article as:** Matsunaga et al., Biotinylated-sortase self-cleavage purification (BISOP) method for cell-free produced proteins *BMC Biotechnology* 2010, **10**:42

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit



# Regulation of Arabidopsis defense responses against *Spodoptera littoralis* by CPK-mediated calcium signaling

Chidananda Nagamangala Kanchiswamy<sup>†1,3,4</sup>, Hirotaka Takahashi<sup>†2,5</sup>, Stefano Quadro<sup>6</sup>, Massimo E Maffei<sup>4</sup>, Simone Bossi<sup>4</sup>, Cinzia Berteza<sup>4</sup>, Simon Atsbaha Zebelo<sup>4</sup>, Atsushi Muroi<sup>1,3</sup>, Nobuaki Ishihama<sup>7</sup>, Hirofumi Yoshioka<sup>7</sup>, Wilhelm Boland<sup>6</sup>, Junji Takabayashi<sup>3</sup>, Yaeta Endo<sup>5</sup>, Tatsuya Sawasaki<sup>5</sup> and Gen-ichiro Arimura<sup>\*1,3</sup>

## Abstract

**Background:** Plant Ca<sup>2+</sup> signals are involved in a wide array of intracellular signalling pathways after pest invasion. Ca<sup>2+</sup>-binding sensory proteins such as Ca<sup>2+</sup>-dependent protein kinases (CPKs) have been predicted to mediate the signaling following Ca<sup>2+</sup> influx after insect herbivory. However, until now this prediction was not testable.

**Results:** To investigate the roles CPKs play in a herbivore response-signaling pathway, we screened the characteristics of Arabidopsis CPK mutants damaged by a feeding generalist herbivore, *Spodoptera littoralis*. Following insect attack, the *cpk3* and *cpk13* mutants showed lower transcript levels of plant defensin gene *PDF1.2* compared to wild-type plants. The CPK cascade was not directly linked to the herbivory-induced signaling pathways that were mediated by defense-related phytohormones such as jasmonic acid and ethylene. CPK3 was also suggested to be involved in a negative feedback regulation of the cytosolic Ca<sup>2+</sup> levels after herbivory and wounding damage. *In vitro* kinase assays of CPK3 protein with a suite of substrates demonstrated that the protein phosphorylates transcription factors (including ERF1, HsfB2a and CZF1/ZFAR1) in the presence of Ca<sup>2+</sup>. CPK13 strongly phosphorylated only HsfB2a, irrespective of the presence of Ca<sup>2+</sup>. Furthermore, *in vivo* agroinfiltration assays showed that CPK3- or CPK13-derived phosphorylation of a heat shock factor (HsfB2a) promotes *PDF1.2* transcriptional activation in the defense response.

**Conclusions:** These results reveal the involvement of two Arabidopsis CPKs (CPK3 and CPK13) in the herbivory-induced signaling network via HsfB2a-mediated regulation of the defense-related transcriptional machinery. This cascade is not involved in the phytohormone-related signaling pathways, but rather directly impacts transcription factors for defense responses.

## Background

One of the significant factors determining successful plant growth and reproduction is an efficient defense against insect attacks. After herbivore feeding there is a dramatic Ca<sup>2+</sup> influx limited to a few cell layers lining the damage zone [1,2]. Signals induced rapidly by herbivore attack have been found to spread over the leaf, leading to a strong Ca<sup>2+</sup>-dependent transmembrane potential (Vm) depolarization in the damage zone followed by a transient

Vm hyperpolarization in the close vicinity and a constant depolarization at distances greater than 6-7 mm [1]. These initial cues are transmitted within the plant by signal transduction pathways that include phosphorylation cascades, such as mitogen-activated protein (MAP) kinases, and the jasmonic acid (JA) pathway, which play a central and conserved role in promoting resistance to a broad spectrum of insects [3]. However, there is a missing link to downstream signaling and gene regulation for defense responses. In this network, Ca<sup>2+</sup>-binding sensory proteins are of particular interest, since following Ca<sup>2+</sup> influx the sensory proteins may secondarily decode information contained in the temporal and spatial patterns of

\* Correspondence: garimura@ecology.kyoto-u.ac.jp

<sup>1</sup> Global COE Program: Evolution and Biodiversity, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

<sup>†</sup> Contributed equally

Full list of author information is available at the end of the article

