J Biomol NMR (2009) 43:145–150
DOI 10.1007/s10858-008-9296-5

COMMUNICATION

Attachment of an NMR-invisible solubility enhancement tag using a sortase-mediated protein ligation method

Yoshihiro Kobashigawa · Hiroyuki Kumeta · Kenji Ogura · Fuyuhiko Inagaki

Received: 6 October 2008 / Accepted: 5 December 2008 / Published online: 13 January 2009

Abstract Sample solubility is essential for structural studies of proteins by solution NMR. Attachment of a solubility enhancement tag, such as GB1, MBP and thioredoxin, to a target protein has been used for this purpose. However, signal overlap of the tag with the target protein often made the spectral analysis difficult. Here we report a sortase-mediated protein ligation method to eliminate NMR signals arising from the tag by preparing the isotopically labeled target protein attached with the non-labeled GB1 tag at the C-terminus.

Keywords Sortase · Protein ligation · GB1 · Intein · INSET · Solubility enhancement

Abbreviations

SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
INSET	Isotopically invisible solubility/stability enhancement tag

Sample solubility is essential for solution NMR studies. A number of approaches have been developed to address this problem, with buffer-condition screening and site-specific mutagenesis (Bagby et al. 1997; Huang et al. 1996) being widely used. However, these methods rely on a considerable amount of trial and error. The attachment of solubility enhancement tags, such as GB1, MBP and thioredoxin, to target proteins is one of the solutions that has been applied to the NMR analysis of less soluble samples (Zhou et al. 2001; Chang et al. 2006; Safadi and Shaw 2007). However, signal overlap of the tag with the target protein often complicates the spectral analysis. Therefore, incorporation of an unlabeled tag into an isotopically labeled target protein offers a potential solution to this difficulty (Züger and Iwai 2005).

Intein-mediated protein ligation (Muralidharan and Muir 2006) is useful for this purpose. There are two types of intein-mediated ligation: one uses a split intein and the other uses expressed protein ligation (EPL) (Muralidharan and Muir 2006). They were successfully applied for preparation of segmental isotopically labeled protein for NMR studies (Yamazaki et al. 1998; Otomo et al. 1999; Camarero et al. 2002; Xu et al. 1999). In both methods, the target protein should be expressed as a fusion protein with an intein, eventually resulting in lower expression yield and lower solubility (Yamazaki et al. 1998; Otomo et al. 1999; Nyanguile et al. 2003). Moreover, a high concentration of ligation partners are required when using EPL (Camarero et al. 2002; Xu et al. 1999). Thus, both intein-based ligation methods possess difficulties to their practical application.

Sortase, a transpeptidase found in the cell envelope of many Gram-positive bacteria, anchors surface proteins to the peptidoglycan cross bridge of the cell wall (Mazmanian et al. 1999; Novick 2000; Perry et al. 2002). *Staphylococcus aureus* sortase (Sortase A) catalyzes the transpeptidation by cleaving the peptide bond between threonine and glycine at an LPXTG recognition motif and subsequently joining the carboxyl group of the threonine to

Electronic supplementary material The online version of this article (doi:10.1007/s10858-008-9296-5) contains supplementary material, which is available to authorized users.

Y. Kobashigawa · H. Kumeta · K. Ogura · F. Inagaki (✉)
Department of Structural Biology, Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan
e-mail: finagaki@pharm.hokudai.ac.jp

an amino group of the pentaglycine sequence on the cell wall peptideglycan (Mazmanian et al. 1999; Novick. 2000; Perry et al. 2002). Recently, sortase A was used for in vitro protein ligation (Mao et al. 2004). However, this novel protein ligation method has not yet been applied to NMR sample preparation.

Sortase-mediated protein ligation only requires the attachment of an LPXTG motif to the C-terminus of the N-terminal fragment. The attachment of this short sequence does not result in lower expression or lower solubility as often happens in the intein-mediated ligation system (Yamazaki et al. 1998; Otomo et al. 1999; Nyanguile et al. 2003). Since sortase and the ligation fragments are separately prepared and mixed to start the reaction, the optimization of ligation efficiency is possible by biochemical approaches.

Here, we applied the sortase-mediated ligation method to the attachment of a non-labeled GB1-tag to isotopically labeled target proteins to improve sample solubility. In order to improve expression level (Huth et al. 1997; Sugase et al. 2008; Gronenborn et al. 1991) and to avoid the precipitation of the target protein during purification and ligation, we expressed the isotopically labeled target protein fused with a cleavable GB1 tag at its N-terminus (Fig. 1). After ligation with a non-labeled C-terminal GB1 by sortase-mediated ligation, the N-terminal fused GB1 was subsequently removed. We applied this ligation system to the structural analyses of Vav C-terminus SH3 (denoted as VcSH3), which is almost insoluble at physiological pH.

The constructs for VcSH3 and sortase A contained a hexahistidine-tag, and could be purified by affinity chromatography using Ni-NTA followed by gel filtration chromatography. After purification by affinity chromatography using Ni-NTA, the N-terminal hexahistidine-tag of the GB1 construct was cleaved by Tev protease to expose N-terminal tri-glycine residues for sortase-mediated protein ligation (Fig. 1). Protein ligation was carried out at room temperature for three days in 20 mM Tris–HCl (pH 8.0), 10 mM $CaCl_2$, 150 mM NaCl and 2 mM 2-mercaptoethanol. The N-terminus fragment, VcSH3, was mixed with C-terminus fragment GB1 and sortase A at a molar ratio 1–2 to 0.1 where concentration of N-terminus fragments, VcSH3, was 25 μM.

We initially performed the protein ligation by batch method (Fig. 2a, lane 2), and found that the reaction efficiency was less than 20% and most of the VcSH3 was in unligated state. Mao et al. (2004) mentioned that the reaction efficiency was quite low in protein-protein ligation. To improve the efficiency of the sortase-mediated ligation, we tried dialysis of the reaction mixture during ligation to remove the GLEHHHHHH by-product. This by-product also has N-terminus Gly, which may compete with the ligation reaction of the INSET (isotopically invisible solubility enhancement-tag). In fact, a previous report

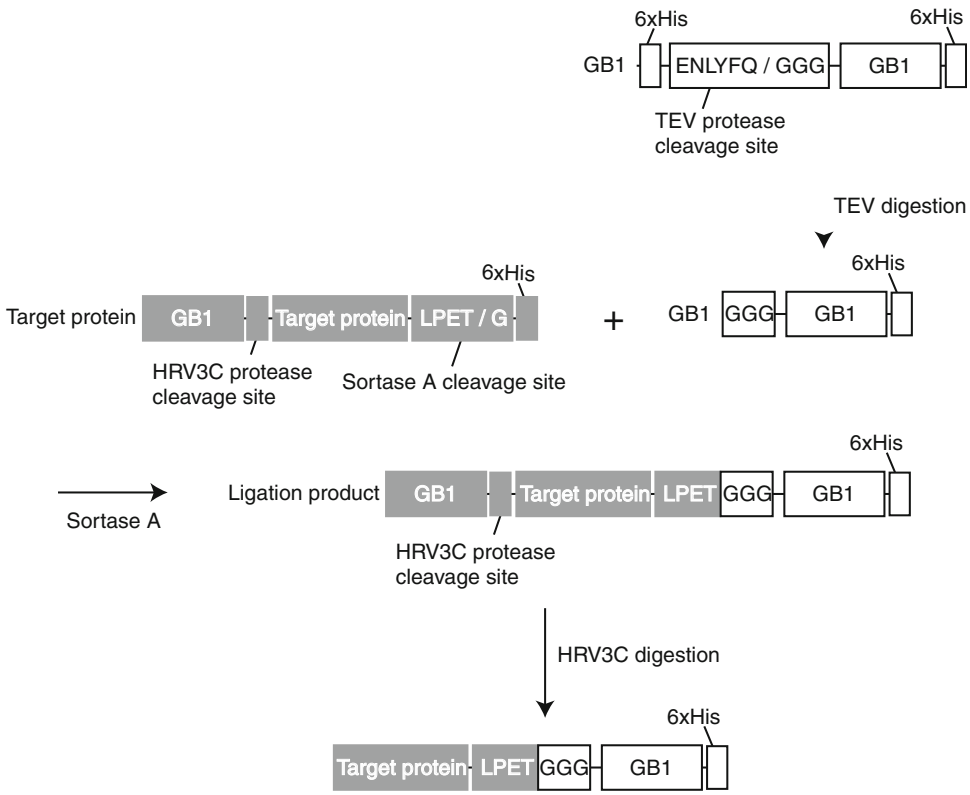

Fig. 1 Schematic representation of the protocol for the preparation of the INSET-attached target protein. Isotopically labeled and non-labeled regions are colored gray and white, respectively

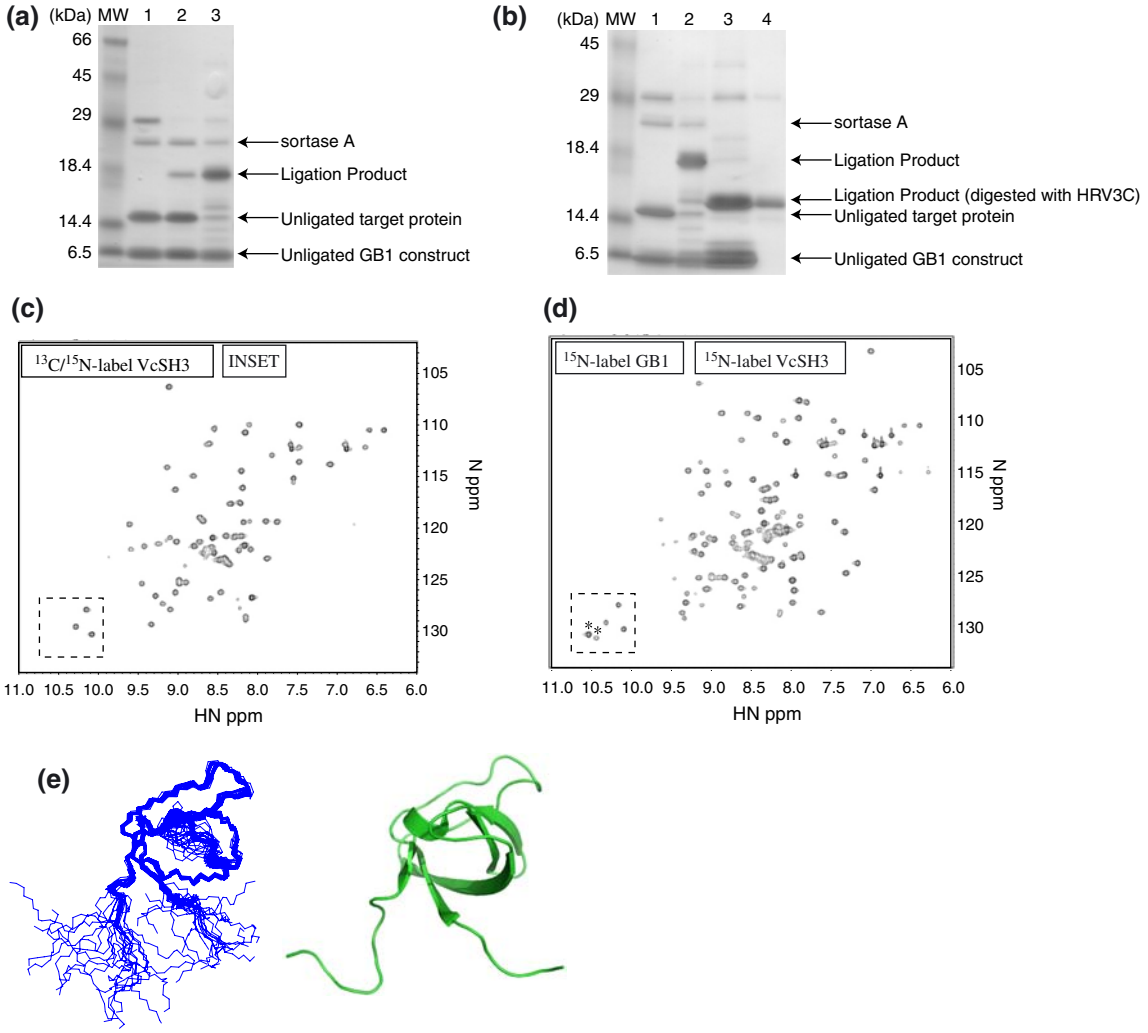

Fig. 2 **a** SDS-PAGE of sortase-mediated protein ligation of VcSH3. Lane 1, before sortase reaction; lane 2, by batch method after reaction for three days; lane 3, by dialysis method after reaction for three days. **b** SDS-PAGE analysis of sortase-mediated protein ligation, and purification of INSET-attached VcSH3. Lane 1, before sortase reaction; lane 2, after reaction for three days; lane 3, after ligation, dialysis and Ni-NTA purification; lane 4, after further purification by gel filtration. **c** The $^{1}H^{-15}$ N HSQC spectrum of $^{13}C/^{15}$ N-labeled INSET-attached VcSH3. **d** The $^{1}H^{-15}$ N HSQC spectra of uniformly ^{15} N-labeled GB1-VcSH3. Signals enclosed by rectangle in **c** and **d** are due to Trp indoles. Asterisk in **d** denotes indole signals from GB1. **e** Overlay of the twenty NMR structures of $^{13}C/^{15}$ N-labeled INSET-attached VcSH3 (*left*) and ribbon representation (*right*)

showed that the single Gly at the N-terminus was sufficient for the sortase-mediated protein ligation (Mao et al. 2004). Protein ligation was carried out at room temperature for three days in 20 mM Tris–HCl (pH 8.0), 10 mM $CaCl_2$, 150 mM NaCl and 2 mM 2-mercaptoethanol dialyzed with the same solution using BioDesignDialysis Tubing 3500 MWCO (BioDesign Inc.). Our experiment revealed that ligation efficiency was markedly improved over that observed in the batch method when dialysis was applied (Fig. 2a, lane 3) and most of the VcSH3 was ligated with GB1. Using dialysis method, we prepared $^{13}C/^{15}$ N-labeled VcSH3 attached with the INSET. Figure 2b shows the results of SDS-PAGE for sortase-mediated protein ligation of VcSH3. Three days after the start of the reaction (lane 1), ligation efficiency reached more than 90%, as judged from the residual amount of unligated VcSH3 construct (lane 2). During the reaction, VcSH3 was not precipitated due to GB1 fusion at N-terminus, but was precipitated within several hours when the N-terminal GB1 was cleaved by HRV3C protease (data not shown). The ligation product was purified by affinity chromatography using Ni-NTA, followed by HRV3C protease digestion (lane 3) and final purification by gel filtration chromatography (lane 4). The final product was concentrated to 0.6 mM in 20 mM MES (pH 7.0), 2 mM DTT and 150 mM NaCl in 90% H_2O/10% $^{2}H_2O$, and was used for NMR analysis. All the NMR experiments were carried out at 25°C on a Varian Inova 800 MHz, 600 MHz or 500 MHz NMR spectrometer as described previously

(Kobashigawa et al. 2007a, b). The solubility of VcSH3 was enhanced more than 10 fold by the attachment of the INSET. Figure 2c and d show the $^{1}H^{-15}N$ HSQC spectra of the INSET-attached $^{13}C/^{15}N$-labeled VcSH3 domain and the uniformly ^{15}N-labeled GB1-VcSH3 fusion protein, respectively. We observed signals derived from VcSH3 alone, but did not observe any GB1 signals (Fig. 2c) in comparison to the GB1-fusion protein (Fig. 2d). For example, we could observe only three signals in the Trp indole region enclosed by rectangle of Fig. 2c, while five peaks were observed in Fig. 2d. In fact, VcSH3 possesses three Trp residues while two in GB1. We determined the solution structure of the VcSH3 domain. The well-defined structure was obtained due to attachment of the INSET (Fig. 2e), confirming that the INSET is a useful solubility enhancement tag that does not complicate the NMR spectrum.

We found that attachment of the INSET by sortase-mediated protein ligation is an efficient and convenient method for the preparation of solubility enhanced protein conjugates without complicating the NMR spectrum. We also found that ligation efficiency was markedly increased by dialysis method. Moreover, it was found that the N-terminus GB1 fusion tag that was ultimately cleaved was useful in preventing precipitation of the target protein during the purification and protein ligation processes.

For wide use of the INSET and the sortase-mediated protein ligation method to NMR studies, establishment of the strategy for optimization of the reaction is helpful. Sortase-mediated protein ligation is a simple enzymatic-reaction, for which the optimization of reaction conditions can be possible to improve ligation efficiency. To confirm this, we monitored ligation efficiency under various experimental conditions using the VcSH3 ligation system. For convenience, ligation efficiency was estimated by monitoring time-dependent decrease in the band of VcSH3 in the SDS-PAGE. Since after longer reaction time, slight degradation of VcSH3, sortase A or ligation product was observed in the SDS-PAGE band, ligation efficiency should be taken as an apparent value. Protein ligation was carried at 12, 25 and 35°C in 20 mM Tris-HCl (pH 8.0), 10 mM $CaCl_2$, 150 mM NaCl and 2 mM 2-mercaptoethanol dialyzed with the same solution by using micro dialyzer TOR-3 K (Genetics). We dialyzed 200 µl of the reaction solution to the 25 ml of the reaction buffer (20 mM Tris-HCl (pH 8.0), 10 mM $CaCl_2$, 150 mM NaCl and 2 mM 2-mercaptoethanol) for seven days at 12 and 25°C, and for five days at 35°C. We first tested the effect of protein concentration on ligation efficiency (Fig. 3a and Supplementary Fig. 2). While maintaining the molar ratios of VcSH3, the GB1 construct and sortase A at 1:2:0.1, the concentration dependence of VcSH3 was tested at 10, 20 and 50 µM. The ligation efficiency was enhanced with increases in protein concentration. Next, we studied the dependence of ligation efficiency on sortase A concentration (Fig. 3b and Supplementary Fig. 3). We fixed the concentration of VcSH3 and GB1 construct at 20 and 40 µM respectively and studied the ligation efficiency at sortase A concentration of 2, 5 and 10 µM. The ligation efficiency was improved with increase in sortase concentration. Since sortase A was shown to be active under 37°C (Mao et al. 2004), we studied the effect of temperature on the ligation reaction at 12, 25 and 35°C (Fig. 3c and Supplementary Fig. 4). We fixed the concentration of VcSH3, GB1 construct and sortase A at 20, 40 and 2 µM respectively. The ligation reaction became efficient at higher temperature and after four days, the apparent reaction efficiency reached at 80% at 35°C. This value is similar to the sortase A concentration of 5 µM at 25°C (Fig. 3b). Thus, at 35°C, the concentration of sortase A could be reduced to less than half that used at 25°C. From these results, we concluded that the ligation efficiency can be optimized by biochemical considerations. Finally, we tested whether ligation proceeded under the lower temperature. We studied ligation efficiency at 12°C with various sortase A concentration (Fig. 3d and Supplementary Fig. 5). We fixed the concentration of VcSH3 and the GB1 construct to 20 and 40 µM respectively at the sortase A concentration of 2, 5, 10 and 25 µM. After seven days, most of the initial VcSH3 disappeared at 25 µM sortase A concentration. This is similar to 5 µM sortase A at 25°C after seven days reaction (Fig. 3b). We confirmed that the ligation proceeds at lower temperature with higher sortase concentration. The ligation condition could be optimized under various temperature range depending on the sortase concentration. Lower temperature together with increases in enzyme concentration might allow for application of sortase-mediated protein ligation to unstable samples.

In this paper, we described the INSET approach using sortase-mediated protein ligation. It provided a convenient method to prepare soluble NMR samples for structural studies. In particular, the cleavable N-terminus GB1-tag was effective in preventing precipitation of the target protein during purification and ligation. We showed that the dialysis method markedly enhanced the ligation efficiency. These approaches would enable us to apply sortase-mediated protein ligation to less soluble proteins and to determine their structures using solution NMR. We also showed that the ligation reaction could be optimized straightforward by biochemical approaches. The present study provides a strategy to optimize the reaction condition for higher ligation efficiency in the sortase-mediated system. The present study will be also helpful for preparation of the segmental labeled proteins. Although the general applicability has yet to be demonstrated, we believe that the sortase-mediated protein ligation method is a

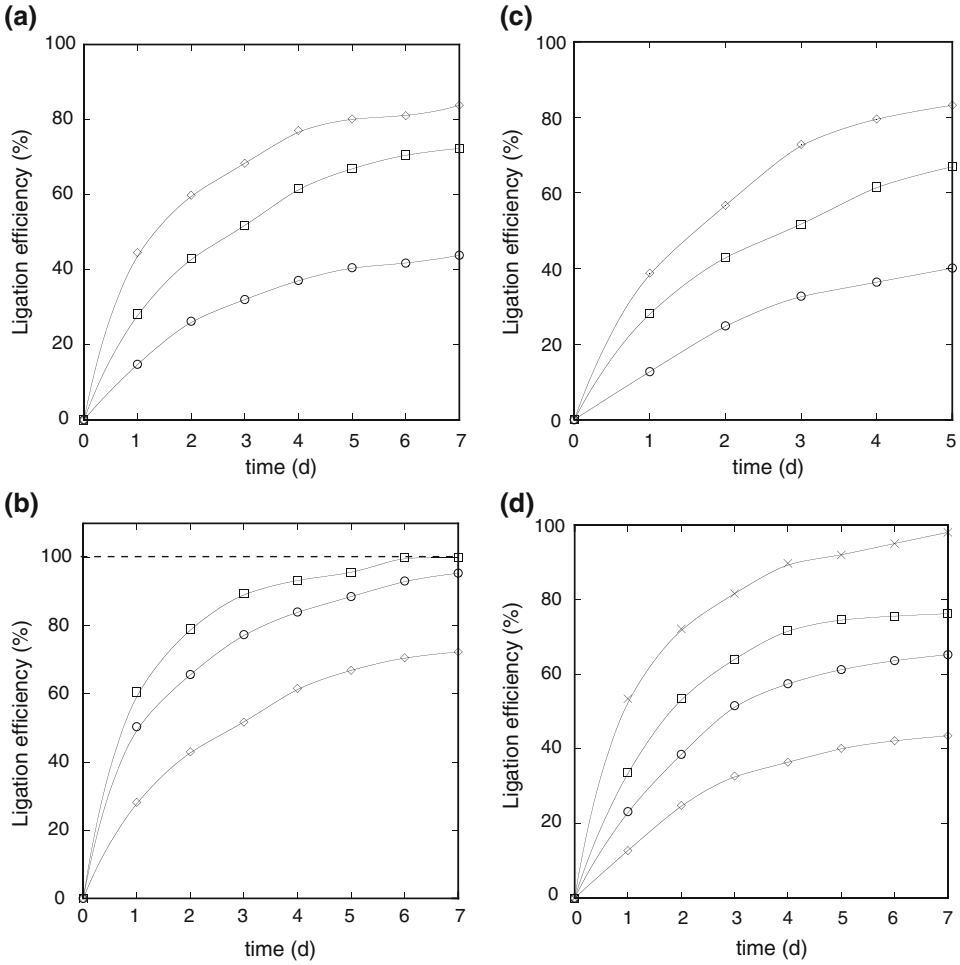

Fig. 3 Time-course of the sortase-mediated protein ligation of VcSH3 under various experimental conditions. **a** Molar ratio of VcSH3, GB1 construct and sortase A was fixed to 1:2:0.1, while protein concentration was varied. VcSH3 concentration at 10 (*circle*), 20 (*rectangle*) and 50 μM (*diamond*) were used in the experiment. Reaction was performed at 25°C. **b** Concentration of VcSH3 and GB1 construct was fixed to 20 and 40 μM respectively, while sortase A concentration was varied at 2 (*diamond*), 5 (*circle*) and 10 μM (*rectangle*). Reaction was performed at 25°C. **c** Concentration of VcSH3, GB1 construct and sortase A was fixed to 20, 40 and 2 μM respectively. Reaction was performed at 12 (*circle*), 25 (*rectangle*) and 35°C (*diamond*). **d** Concentration of VcSH3 and GB1 construct were fixed to 20 and 40 μM respectively, while sortase A concentration was varied at 2 (*diamond*), 5 (*circle*), 10 (*rectangle*) and 25 mM (*cross*). Reaction was performed at 12°C

convenient, efficient and powerful tool for the preparation of segmental labeled proteins in sufficient quantity and purity. Accordingly, the application of NMR can be extended to proteins of larger molecular weight.

Acknowledgements This work was supported by Grants-in-Aid for Scientific Research and the National Project on "Targeted Proteins Research Program" from the Ministry of Education, Science and Culture of Japan.

References

Bagby S, Tong KI, Liu D, Alattia JR, Ikura M (1997) The button test: a small scale method using microdialysis cells for assessing protein solubility at concentrations suitable for NMR. J Biomol NMR 10:279–282

Camarero JA, Shekhtman A, Campbell EA, Chlenov M, Gruber TM, Bryant DA, Darst SA, Cowburn D, Muir TW (2002) Autoregulation of a bacterial σ factor explored by using segmental isotopic labeling and NMR. Proc Natl Acad Sci USA 99:8536–8541

Chang YG, Song AX, Gao YG, Shi YH, Lin XJ, Cao XT, Lin DH, Hu HY (2006) Solution structure of the ubiquitin-associated domain of human BMSC-UbP and its complex with ubiquitin. Protein Sci 15:1248–1259

Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. Science 253:657–661

Huang B, Eberstadt M, Olejniczak ET, Meadows RP, Fesik SW (1996) NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain. Nature 384:638–641

Huth JR, Bewley CA, Jackson BM, Hinnebusch AG, Clore GM, Gronenborn AM (1997) Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. Protein Sci 6:2359–2364

Kobashigawa Y, Naito M, Inagaki F (2007a) An efficient method for protein phosphorylation using the artificially introduced of cognate-binding modules into kinases and substrates. J Biotechnol 131:458–465

Kobashigawa Y, Sakai M, Naito M, Yokochi M, Kumeta H, Makino Y, Ogura K, Tanaka S, Inagaki F (2007b) Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. Nat Struct Mol Biol 14:503–510

Mao H, Hart SA, Schink A, Pollok BA (2004) Sortase-mediated protein ligation: a new method for protein engineering. J Am Chem Soc 126:2670–2671

Mazmanian SK, Liu G, Ton-That H, Schneewind O (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. Science 285:760–763

Muralidharan V, Muir TW (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins. Nat Method 3:429–438

Novick RP (2000) Sortase: the surface protein anchoring transpeptidase and the LPXTG motif. Trends Microbiol 8:148–151

Nyanguile O, Dancik C, Blakemore J, Mulgrew K, Kaleko M, Stevenson SC (2003) Synthesis of adenoviral targeting molecules by intein-mediated protein ligation. Gene Ther 10:1362–1369

Otomo T, Teruya K, Uegaki K, Yamazaki T, Kyogoku Y (1999) Improved segmental isotope labeling of proteins and application to a larger protein. J Biomol NMR 14:105–114

Perry AM, Ton-That H, Mazmanian SK, Schneewind O (2002) Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. III. Lipid II is an in vivo peptidoglycan substrate for sortase-catalyzed surface protein anchoring. PXTG motif. J Biol Chem 277:16241–16248

Safadi SS, Shaw GS (2007) A disease state mutation unfolds the parkin ubiquitin-like domain. Biochemistry 46:14162–14169

Sugase K, Landes MA, Wright PE, Martinez-Yamout M (2008) Overexpression of post-translationally modified peptides in *Escherichia coli* by co-expression with modifying enzymes. Protein Expres Purif 57:108–115

Xu R, Ayers B, Cowburn D, Muir TW (1999) Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. Proc Natl Acad Sci USA 96:388–393

Yamazaki T, Otomo T, Oda N, Kyogoku Y, Uegaki K, Ito N, Ishino Y, Nakamura H (1998) Segmental isotope labeling for protein NMR using peptide splicing. J Am Chem Soc 120:5591–5592

Zhou P, Lugovskoy AA, Wagner G (2001) A solubility-enhancement tag (SET) for NMR studies of poor behaving proteins. J Biomol NMR 20:11–14

Züger S, Iwai H (2005) Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies. Nat Biotechnol 23:736–740