J Biomol NMR (2015) 63:1–8
DOI 10.1007/s10858-015-9981-0

COMMUNICATION

Efficient segmental isotope labeling of multi-domain proteins using Sortase A

Lee Freiburger[1,2] · Miriam Sonntag[1,2] · Janosch Hennig[1,2] · Jian Li[3] · Peijian Zou[1,2,3] · Michael Sattler[1,2,3]

Received: 11 May 2015 / Accepted: 24 August 2015 / Published online: 30 August 2015

Abstract NMR studies of multi-domain protein complexes provide unique insight into their molecular interactions and dynamics in solution. For large proteins domain-selective isotope labeling is desired to reduce signal overlap, but available methods require extensive optimization and often give poor ligation yields. We present an optimized strategy for segmental labeling of multi-domain proteins using the *S. aureus* transpeptidase Sortase A. Critical improvements compared to existing protocols are (1) the efficient removal of cleaved peptide fragments by centrifugal filtration and (2) a strategic design of cleavable and non-cleavable affinity tags for purification. Our approach enables routine production of milligram amounts of purified segmentally labeled protein for NMR and other biophysical studies.

Lee Freiburger and Miriam Sonntag have equally contributed to this work.

The Sortase A construct can be obtained from the authors upon request.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-015-9981-0) contains supplementary material, which is available to authorized users.

✉ Michael Sattler
sattler@helmholtz-muenchen.de

Lee Freiburger
lee.freiburger@tum.de

Miriam Sonntag
miriam.sonntag@mytum.de

Janosch Hennig
janosch.hennig@helmholtz-muenchen.de

Jian Li
lijianzhongbei@163.com

Peijian Zou
peijian.zou@helmholtz-muenchen.de

[1] Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany

[2] Center for Integrated Protein Science Munich (CIPSM) at Department of Chemistry, Technische Universität München, Lichtenbergstr.4, 85747 Garching, Germany

[3] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

Keywords Protein ligation · Sortase A · Multi-domain proteins · Segmental isotope labeling · Protein expression

Nuclear magnetic resonance (NMR) spectroscopy provides unique insight into the structure, dynamics and molecular interactions of biological macromolecules, such as large multi-domain proteins and their complexes. To fully understand their molecular functions it is required to study full-length proteins and the assembly in molecular complexes. However, studying large multi-domain proteins is challenging due to fast transverse relaxation which is associated with line broadening, poor sensitivity, as well as signal overlap due to the increased number of spins (Gobl et al. 2014; Riek et al. 2000). A number of these challenges have been addressed through specific and optimized isotope labeling and deuteration (Gardner and Kay 1998; Sattler and Fesik 1996; Tugarinov et al. 2006) and optimized pulse sequences (Pervushin et al. 1997; Riek et al. 2000; Sprangers et al. 2007; Tugarinov et al. 2004). One important and useful approach for analysis of large multi-domain proteins involves segmental isotope labeling (Cowburn et al. 2004; Muona et al. 2010; Riek et al. 2002; Skrisovska et al. 2010; Zuger and Iwai 2005), whereby a

single domain is labeled specifically with NMR active nuclei, and the remaining domains are expressed with NMR inactive nuclei. This approach drastically reduces the peak overlap as only a single domain in a multi-domain protein is visible. In order to produce these samples the differently labeled domains must be produced under different conditions and then linked together. It is the linking step, which has limited the feasibility of the segmental labeling technique.

Several methods have been proposed to facilitate domain ligation such as expressed protein ligation (EPL), protein trans-splicing (PTS) and native chemical ligation (Cowburn et al. 2004; Muona et al. 2010; Xu et al. 1999; Zuger and Iwai 2005). While these methods have been successfully used in NMR studies, they often suffer from poor efficiency and a requirement for extensive optimization. In the case of in vivo PTS ligation with inteins often significant cross-labeling is observed, i.e. where isotope labeling is observed also for the domain which should be unlabeled due to leaky expression (Muona et al. 2010).

More recently, the transpeptidase Sortase A (SrtA), an enzyme responsible for anchoring proteins to the cell wall in gram positive bacteria, has been introduced for the ligation of two separately expressed protein fragments (Kobashigawa et al. 2009; Levary et al. 2011; Mao et al. 2004). Sortase mediated protein ligation involves the expression and purification of two separate protein chains, one with specific isotope labeling (i.e. ^{15}N, ^{13}C, ^{2}H), and the other unlabeled. The two chains are ligated through the use of recombinantly produced SrtA. The *Staphylococcus aureus* SrtA enzyme specifically recognizes a Leu-Pro-Xxx-Thr-Gly (LPXTG) motif on the N-terminal domain and then ligates a peptide with an N-terminal glycine (Mao et al. 2004; Mazmanian et al. 1999). As each domain is expressed and purified separately each domain can be differentially isotope-labeled without additional optimization. The Sortase-mediated ligation does not require cofactors such as ATP or any non-natural modifications of the linked domains and is performed under mild conditions. Only the recognition motif (LPXTG) is required for ligation and there is no restriction for the amino acid sequence outside this region.

SrtA-mediated ligation has been used for many applications which include N- and C-terminal site-specific labeling (Guimaraes et al. 2013; Williamson et al. 2014), immobilization (Clow et al. 2008), and multi-domain ligation (Levary et al. 2011). Previously published protocols describe methods to label proteins with probes such as fluorophores or biotin. The protocol presented here can achieve the ligation of two large, structured domains.

Previously, the final purified yield of the ligated protein has been reported as $\sim$40 % after 72 h reaction time (Refaei et al. 2011). For the preparation of highly optimized (and costly) isotope labeled samples these yields are not sufficient. There have been other reports of Sortase-mediated ligations with greater yields under optimized conditions. However, the high yields were determined by monitoring the time-dependent disappearance of the precursor bands in gels (Kobashigawa et al. 2009; Levary et al. 2011). Moreover, the optimized conditions mentioned above include long reaction times (up to 3 days) or high temperatures (37–42 °C). Such conditions can lead to unfolding or proteolytic degradation of the protein product, especially when intrinsically disordered regions are present. It is therefore desirable to significantly improve the reaction yields and reduce the reaction time and temperature.

To introduce glycine as a first residue in the C-terminal ligation domain we usually express the C-terminal fragment as a fusion protein, which is connected to the C-terminal domain by a recognition motif for tobacco etch virus (TEV) or small ubiquitin-like modifier (SUMO) proteases. After cleavage, these proteases leave glycine as first residue in the C-terminal fragment. The fusion protein is chosen to enhance solubility and for affinity purification to aid in the initial purification of the component. A challenge with applying the Sortase strategy has been to obtain high yields due to the complete reversibility of the ligation reaction. Upon ligation, a small peptide fragment is cleaved, which is able to back react to produce the original reactants. We have overcome this limitation through modifications of the reaction protocol as detailed below.

Two critical optimizations to the Sortase protocol enable highly efficient segmental isotope labeling of multi-domain proteins (detailed protocols are provided in the Supporting Information):

1. The reaction is set up to remove the cleaved peptide fragment throughout the reaction by transferring the reactants to a centrifugal concentrator with a molecular weight cut-off that allows separation of the cleaved fragment (Fig. 1a). Continuous concentration and replacement of the buffer thereby allows for efficient removal of the fragment while retaining all other components. Centrifugal concentration is employed between 1250 and 2000 g to avoid that proteins are concentrated too quickly and thus might precipitate. Moreover, the ligation reaction is performed at room temperature as many proteins are unstable at higher temperatures.
2. The ligated product is purified by the strategic use of cleavable and non-cleavable His-tags (Fig. 1b). For this purpose, the N-terminal domain is designed with His-tags at both the N- and C-terminus. Initial purification with a Ni–NTA column separates the N-terminal domain and the ligated product from the other components, which lack a His-tag. The

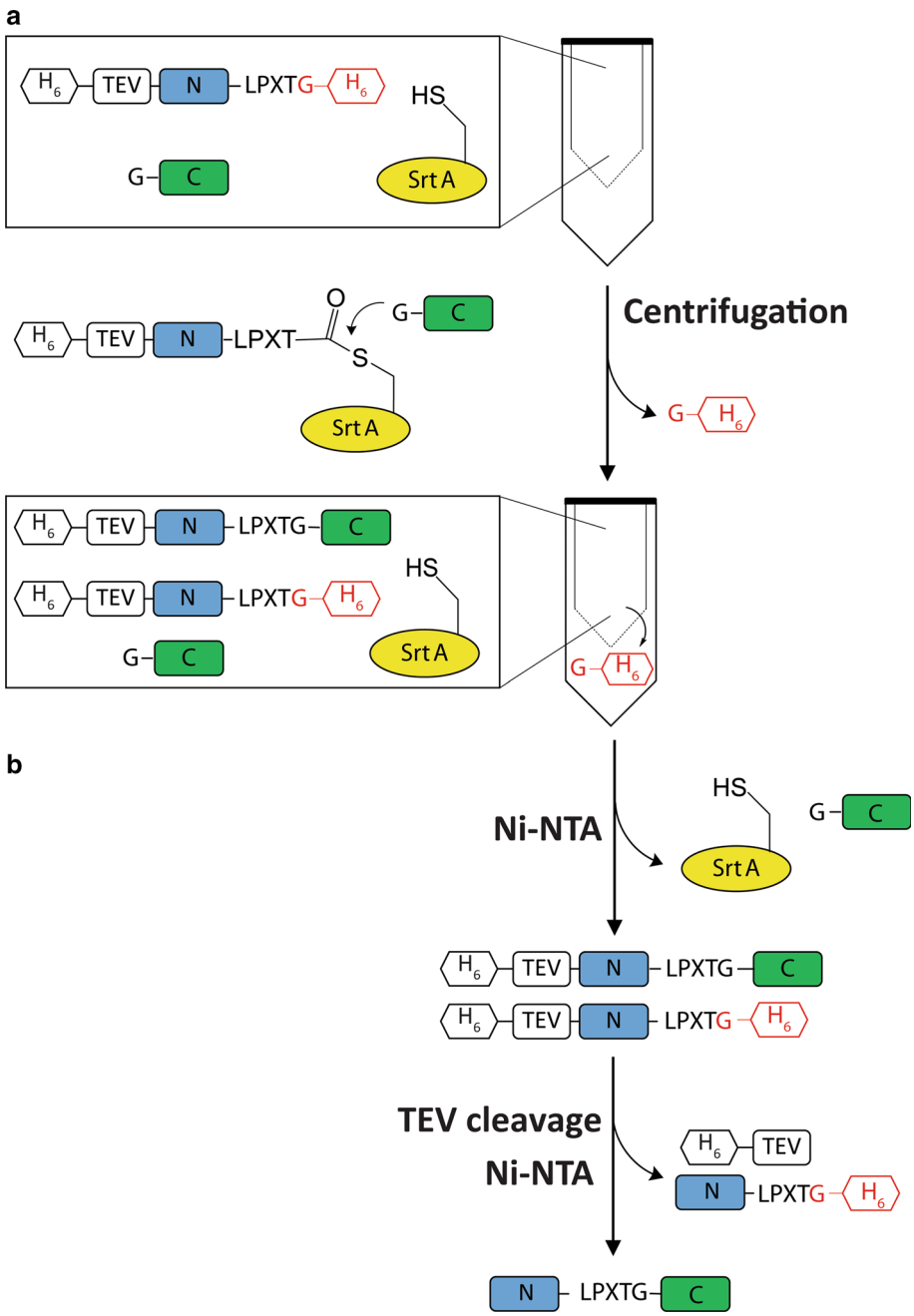

Fig. 1 Schematic representation of the improved ligation protocol. **a** Sortase ligation reaction. The N- and C-terminal domains are incubated in the presence of SrtA in a centrifugal concentrator with a molecular weight cut-off that allows separation of the cleaved fragment. During centrifugation, SrtA cleaves the Thr-Gly bond in the LPXTG recognition motif. The C-terminal domain which contains an N-terminal Gly residue restores the Thr-Gly peptide bond by resolving the formed intermediate which results in the ligated product. Continuous concentration with an optimal molecular-weight cut-off of the filter unit allows the elimination of the cleaved fragment. Thus only the ligated product and the educts are retained in the centrifugal concentrator. **b** Purification of the ligated product. Upon completion of the ligation reaction, the ligated product is purified from the still unligated C-terminal and SrtA enzyme via a Ni–NTA column. After the Ni–NTA column only the unligated N-terminal domain and the ligated product are left. TEV protease cleavage and a subsequent Ni–NTA column separates the unligated N-terminal domain form the ligated product

N-terminal His-tag is then removed by cleavage with TEV protease followed by purification with a second Ni–NTA column. The C-terminal His-tag is not cleavable by TEV but is removed during the ligation reaction. Therefore, the non-ligated N-terminal domain will still retain a His-tag at the C-terminus and thus be separated from the ligated product at this stage. This approach allows straightforward purification of the desired ligated protein with high efficiency.

To determine the efficiency of the ligation reaction we measured the final protein concentration by UV absorbance and compared it to the initial concentration of our limiting protein component. We have developed and successfully tested this protocol using the optimized variant of Sortase A from *S. aureus* reported previously (Chen et al. 2011). Notably, we obtained final purified yields of up to 70 % compared to the initial amount of the limiting protein component. The optimized protocol consistently provides 40–100 % more ligated protein compared to the dialysis method under otherwise identical conditions as shown below. To demonstrate the improvements of our protocol we applied this to two multi-domain proteins that are

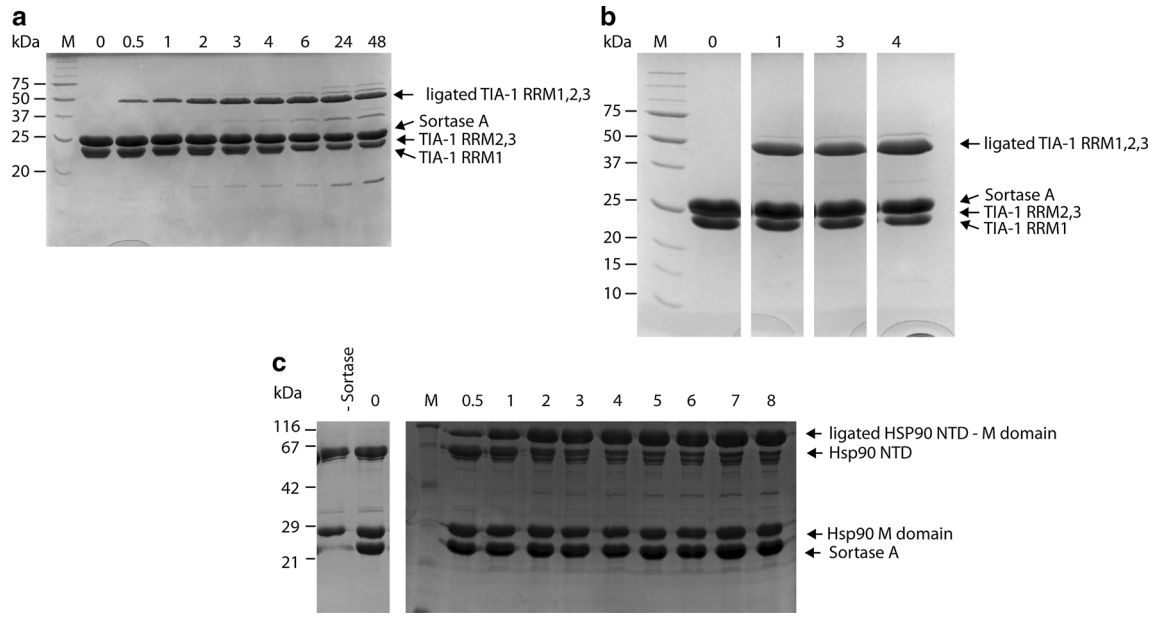

Fig. 2 Monitoring the reaction of the ligation by SDS-PAGE. The reaction times are given in hours on top. **a** An attempted ligation of TIA-1 using dialysis to remove the cleaved fragment. The appearance of the product is slow and the reaction does not reach completion. After 24 h degradation of the ligated product starts. **b** Ligation reaction of TIA-1 using the concentration method. The reaction reaches completion within 3–4 h. **c** SDS-PAGE analysis demonstrating successful ligation of the NTD and MD domains of Hsp90 using SrtA. The first lane is a negative control for the N- and C-terminal fragments without the addition of SrtA. Upon addition of SrtA product can be observed after a short time and almost a complete disappearance of the N-terminal domain precursor occurs between 4 and 6 h

representative for typical application of segmental labeling (1) the splicing factor T cell-restricted intracellular antigen-1 (TIA-1) and (2) the heat shock protein Hsp90. These two proteins exhibit very distinct features with large differences in size and charge.

TIA-1 consists of three RNA recognition motif (RRM) domains, which are connected by short 10–11 amino acids flexible linkers (Wang et al. 2014). In order to dissect the contribution of the individual RRM domains for RNA binding and study the structure of this multi-domain protein we prepared segmentally labeled samples where either RRM1 (N-terminal domain) or RRM2–RRM3 (C-terminal domain) are selectively isotope labeled and linked together via SrtA to a final size of 31.6 kDa (Fig. 2a, b). We have also demonstrated the utility of our segmental labeling protocol for Hsp90 (Lorenz et al. 2014) by ligating the N-terminal (NTD) and middle (MD) domains using differential segmental isotope labeling for a final protein product of 60 kDa molecular weight (Fig. 2c). Both domains of Hsp90 are relatively large, 25 and 32 kDa, respectively, and are connected by a long highly charged linker. The NTD precursor also contains a MBP fusion protein which thus corresponds to a molecular weight of 60 kDa in the ligation reaction. The final ligated product therefore has a molecular weight of $\sim$90 kDa before TEV cleavage. The availability of two complementary domain-selectively isotope-labeled proteins, i.e. $^{2}H,^{15}N$–NTD, ^{1}H-MD and ^{1}H-NTD, $^{2}H,^{15}N$–MD was crucial to identify and map the interaction site of Hsp90 with a client protein. This system demonstrates that this protocol can be used with very large systems, consisting of multiple independent domains.

For segmental isotope labeling of TIA-1, we initially attempted an intein-based approach using protein trans-splicing (PTS) in vivo (Muona et al. 2010). However, we were unable to optimize conditions to obtain sufficient amounts of appropriate segmentally labeled protein. Moreover, leaky expression lead to undesired cross-labeling of the different domains. We then attempted ligation using SrtA, where the two fragments to be ligated are expressed and purified separately so that cross-labeling can be safely excluded. An advantage of this approach is also that previously optimized expression and purification protocols for individual domains can be employed. Initially we eliminated undesired fragments and side products via dialysis as was proposed previously (Kobashigawa et al. 2009; Refaei et al. 2011). This approach was in principle functional (Fig. 2a), however the final yields of the purified ligated product were not satisfying. We tried to enhance the ligation yield by optimizing the ligation time and concentration of the precursors. In order to determine the optimal ligation time, we incubated all three reactants up to 2 days.

Plotting the normalized intensity of the ligated protein band over the ligation time (Supplementary Fig. 3) showed, that the intensity of the ligated product band increases within the first 24 h of ligation starting 30 min after incubation. Incubating the reactants longer than 24 h results in hydrolysis of the ligated product band by SrtA (Mao et al. 2004; Ton-That et al. 2000), indicated by the drop in intensity of the ligated product band. Hydrolysis of the LPXTG motif by SrtA in the absence of an amino group nucleophile leads to a dead end product containing only an LPXT motif. After 6 h of ligation, we observed degradation of the precursor bands for TIA-1 (Fig. 2a). Therefore, all subsequent ligation reactions were quenched between 5 and 6 h to yield sufficient amounts of ligated product but prevent degradation of the precursors. Screening different precursor concentrations and molar ratios of the reactants showed that the ligation works best with 1:1:2 molar ratios for the N-, C-terminal domain and Sortase A, respectively, at 25 μM concentration of the N-terminal domain (Supplementary Fig. 1). Increasing the ligation temperature as suggested by Levary et al.(2011) resulted in degradation of the precursors. Therefore, we performed all subsequent ligation reactions at room temperature.

We then introduced the two critical optimizations mentioned above. One challenge with TIA-1 is the fact that the C-terminal fragment has a molecular weight comparable to SrtA. Although, following the reaction by monitoring the disappearance of the C-terminal fragment is not possible, one can clearly observe the appearance of a new prominent protein band at around 50 kDa after 1 h (Fig. 2b). We observed the same pattern as for the dialysis approach suggesting similar ligation kinetics (Supplementary Fig. 3). Within the first 24 h of ligation the intensity of the ligated protein band increases before it starts to decrease. We also observe degradation of the precursors starting after 6 h of ligation. Therefore, all subsequent reactions were stopped after 5–6 h to yield sufficient amounts of ligated product and prevent degradation. We repeated these reactions at least three times for both systems. Final yields of the purified ligated product of up to 70 % were obtained. A direct comparison of the dialysis and concentrator approach with TIA-1, where otherwise identical conditions and concentrations were used shows that final purified ligation yields are increased by 43 % using our optimized protocol (Supplementary Fig. 2a, b). Moreover, the direct comparison shows that the ligation reaction is significantly faster when using the concentrator approach (Supplementary Fig. 3).

When applying the improved protocol for segmental isotope labeling to Hsp90 (Supplementary Fig. 2c, 4) we

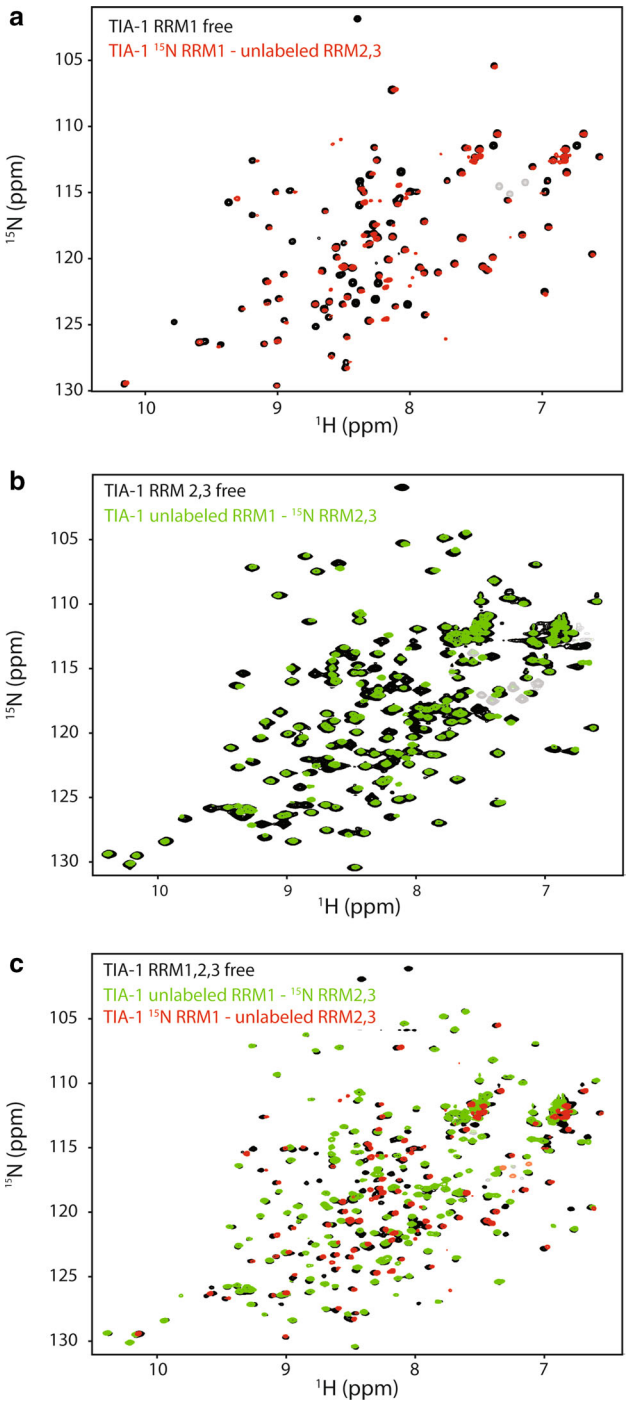

Fig. 3 **a** ^{1}H,^{15}N HSQC spectra of RRM1 (*black*) and **b** RRM2,3 (*black*) of TIA-1 compared to the respective segmentally labeled RRM1,2,3 (*red* and *green*, respectively). **c** Spectra overlay of segmentally labeled RRM1 (*red*) and RRM2,3 (*green*) to uniformly labeled TIA-1 RRM1,2,3 (*black*). The NMR spectra of the segmentally labeled proteins superimpose closely to the corresponding individual domains and full length protein. This confirms successful ligation both N- and C-terminal domain labeling and indicates that the structural integrity is not affected

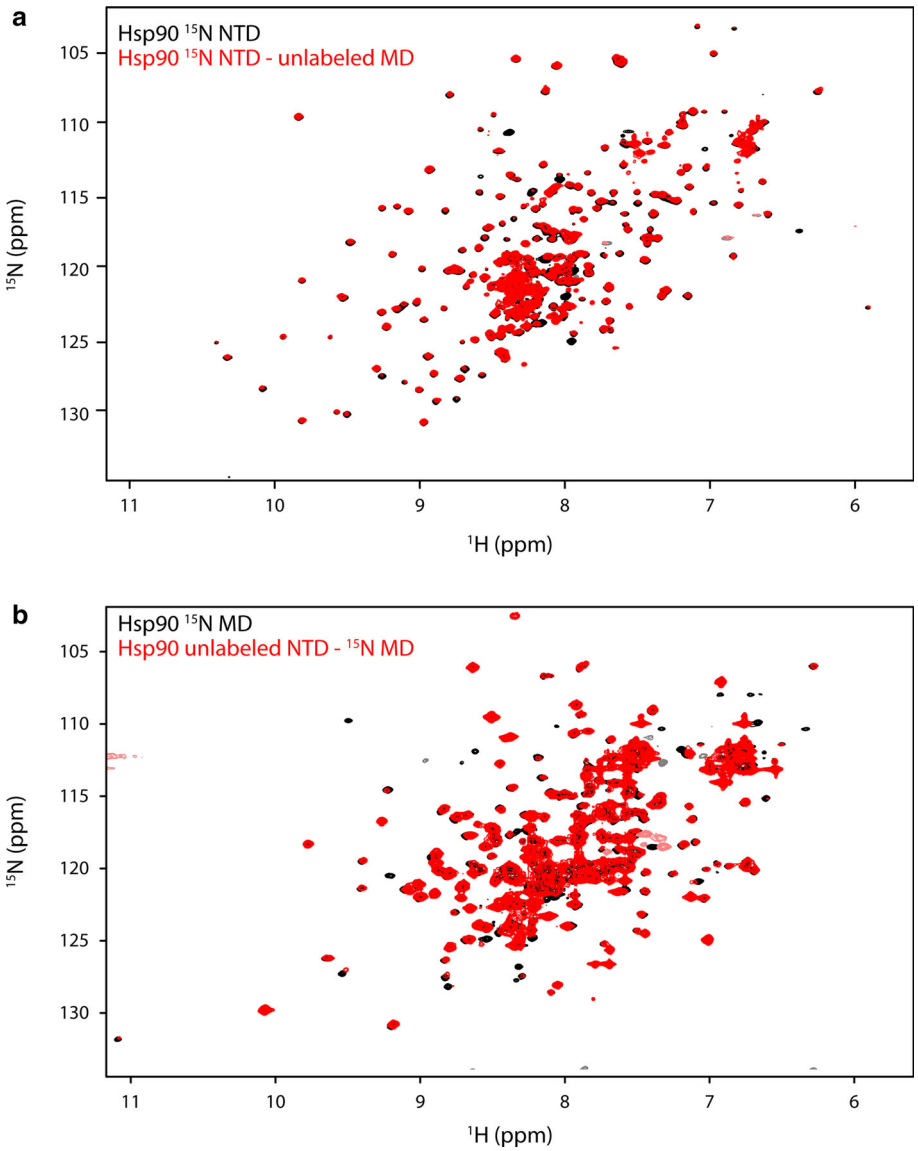

Fig. 4 Spectral overlay of ^{1}H,^{15}N HSQC experiments for **a** ^{15}N-labeled Hsp90 NTD (residues 1–258, *black*) with segmentally NTD-^{15}N-labeled Hsp90 NTD-MD protein (residues 1–529, *red*). **b** Spectral overlay of the Hsp90 middle domain (residues 259–529, *black*) with MD-^{15}N-labeled Hsp90 NTD-MD (residues 1–529, *red*)

typically obtain purified product yields between 45 and 60 % and occasionally up to 70 %. In case of Hsp90, we observe a significant reduction of the limiting component in under 6 h. In all tested cases the purified ligation yields were increased by approximately twofold when using the concentrator approach compared to the dialysis method (Supplementary Fig. 2c, 5). The reduction in reaction time minimizes the risk of dead end side product production and thus increases the efficiency of the ligation reaction and simplifies the purification. We have observed this protocol to be robust and consistently obtain high yields from our ligation reactions.

In a last step, we confirmed the final segmentally labeled protein using NMR spectroscopy. Superposition of ^{1}H, ^{15}N correlation spectra of ligated proteins with ^{15}N-labeling in either RRM1 or RRM2-RRM3 with a corresponding spectrum of uniformly labeled full-length protein shows almost identical chemical shifts thus confirming the effectiveness of the ligation for TIA-1 (Fig. 3). For Hsp90 NMR spectra of the NTD alone and in the Hsp90 NTD-MD segmentally labeled protein show some notable chemical shift differences, while spectra of the MD compared to the Hsp90 NTD-MD segmentally labeled protein shows a high degree of overlap with a few peaks showing a reduction in intensity (Fig. 4). This indicates that there are some effects from the ligation of the NTD-MD indicating that the NTD and MD may interact.

Our improved protocol is not only useful for the efficient production of segmentally isotope-labeled proteins for NMR studies, but for a wide range of additional biophysical experiments, such as small-angle neutron scattering (SANS), electron paramagnetic resonance (EPR) or

fluorescence-based methods. Reducing the number of isotope-labeled amino acids by segmental labeling greatly simplifies NMR spectra. This can allow for the residue-specific analysis of NMR parameters, which may be obscured by signal overlap in a uniformly isotope-labeled sample. The method will be most useful when working with large multi-domain systems, i.e. where flexible linkers connect structurally independent domains.

The potential for the efficient preparation of segmentally ^{2}H-labeled proteins opens also novel opportunities for SANS. It has been demonstrated that combining SANS data with NMR data can provide unique complementary information for structural analysis of large multi-domain proteins (Hennig et al. 2013, 2014; Lapinaite et al. 2013; Madl et al. 2011). A unique information is provided by SANS experiments when using contrast matching experiments where deuterated and non-deuterated proteins can be discriminated by adjusting the D_2O concentration of the measurement buffer. Thereby, it is possible to render specific components invisible depending on their isotope composition. As the contrast match point is different for protonated and deuterated proteins segmental deuteration, i.e. of individual domains, can thus provide information about relative domain arrangements in a multi-domain protein. SANS contrast matching has been employed in a number of recent studies involving differentially deuterated multi-protein complexes (Hennig et al. 2014; Lapinaite et al. 2013; Ramakrishnan 1986). The efficient segmental isotope-labeling protocol presented here will enable SANS analysis of single chain multi-domain proteins in the future.

The improved protocol will also be useful for the introduction of domain-selective labeling for other biophysical methods, which require the attachment of a chemical probe to specific residues or regions in a protein often via a cysteine residue. For residue-specific labeling only a single reactive cysteine residue should be available. Therefore, typically all other native cysteine residues have to be replaced by alanine or serine using site-directed mutagenesis. However, this may perturb the structural integrity of the protein. With segmental labeling, one domain can be specifically modified with the probe, whereas other domains remain unmodified. Given the importance of these methods to characterize the interactions and conformational dynamics of (large) proteins in solution, the availability of an efficient segmental labeling protocol will enable and greatly enhance such studies in the future.

The optimized protocol presented here increases the purified ligation yield at least up to 40 % and reduces reaction times down to 6 h. In all cases tested we observed a substantial increase in ligated protein yield compared to that of the standard dialysis method. The protocol is robust and straightforward to implement and does not require extensive optimization. It routinely enables the production of segmentally labeled protein with excellent yields and quality and therefore greatly expands the application and utility of NMR and other biophysical methods that benefit from protein ligation to more complex systems.

Acknowledgments We thank Iren Wang for discussions and TIA-1 expression protocols and Arie Geerlof at the Protein Expression and Purification Facility of the Helmholtz Center Munich for discussion. This work was supported by *The Deutsche Forschungsgemeinschaft* (DFG, SFB1035 and GRK1721 to M.S.), the Center for Integrated Protein Science Munich (CIPSM). L.F. acknowledges support by an European Molecular Biology Organization (EMBO) Longterm Fellowship and Marie Curie FP7 International Incoming Fellowship. J. H. acknowledges the EMBO for a long-term fellowship (ALTF-276-2010) and the Swedish Research Council (Vetenskapsrådet) for a postdoc fellowship.

References

Chen I, Dorr BM, Liu DR (2011) A general strategy for the evolution of bond-forming enzymes using yeast display. Proc Natl Acad Sci USA 108:11399–11404. doi:10.1073/pnas.1101046108

Clow F, Fraser J, Proft T (2008) Immobilization of proteins to biacore sensor chips using *Staphylococcus aureus* sortase A. Biotechnol Lett 30:1603–1607. doi:10.1007/s10529-008-9718-1

Cowburn D, Shekhtman A, Xu R, Ottesen JJ, Muir TW (2004) Segmental isotopic labeling for structural biological applications of NMR. Methods Mol Biol 278:47–56. doi:10.1385/1-59259-809-9:047

Gardner KH, Kay LE (1998) The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct 27:357–406. doi:10.1146/annurev.biophys.27.1.357

Gobl C, Madl T, Simon B, Sattler M (2014) NMR approaches for structural analysis of multidomain proteins and complexes in solution. Prog Nucl Magn Reson Spectrosc 80:26–63. doi:10.1016/j.pnmrs.2014.05.003

Guimaraes CP, Witte MD, Theile CS, Bozkurt G, Kundrat L, Blom AEM, Ploegh HL (2013) Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. Nat Protoc 8:1787–1799. doi:10.1038/nprot.2013.101

Hennig J, Wang I, Sonntag M, Gabel F, Sattler M (2013) Combining NMR and small angle X-ray and neutron scattering in the structural analysis of a ternary protein-RNA complex. J Biomol NMR 56:17–30. doi:10.1007/s10858-013-9719-9

Hennig J et al (2014) Structural basis for the assembly of the Sxl-Unr translation regulatory complex. Nature 515:287–290. doi:10.1038/nature13693

Kobashigawa Y, Kumeta H, Ogura K, Inagaki F (2009) Attachment of an NMR-invisible solubility enhancement tag using a sortase-mediated protein ligation method. J Biomol NMR 43:145–150. doi:10.1007/s10858-008-9296-5

Lapinaite A, Simon B, Skjaerven L, Rakwalska-Bange M, Gabel F, Carlomagno T (2013) The structure of the box C/D enzyme reveals regulation of RNA methylation. Nature 502:519–523. doi:10.1038/nature12581

Levary DA, Parthasarathy R, Boder ET, Ackerman ME (2011) Protein-protein fusion catalyzed by sortase A. PLoS One 6:e18342. doi:10.1371/journal.pone.0018342

Lorenz OR et al (2014) Modulation of the Hsp90 chaperone cycle by a stringent client protein. Mol Cell 53:941–953. doi:10.1016/j.molcel.2014.02.003

Madl T, Gabel F, Sattler M (2011) NMR and small-angle scattering-based structural analysis of protein complexes in solution. J Struct Biol 173:472–482. doi:10.1016/j.jsb.2010.11.004

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. doi:10.1021/ja039915e

Mazmanian SK, Liu G, Hung TT, Schneewind O (1999) Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. Science 285:760–763. doi:10.1126/science.285.5428.760

Muona M, Aranko AS, Raulinaitis V, Iwai H (2010) Segmental isotopic labeling of multi-domain and fusion proteins by protein trans-splicing in vivo and in vitro. Nat Protoc 5:574–587. doi:10.1038/nprot.2009.240

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371. doi:10.1073/pnas.94.23.12366

Ramakrishnan V (1986) Distribution of protein and RNA in the 30S ribosomal subunit. Science 231:1562–1564. doi:10.1126/science.3513310

Refaei M et al (2011) Observing selected domains in multi-domain proteins via sortase-mediated ligation and NMR spectroscopy. J Biomol NMR 49:3–7. doi:10.1007/s10858-010-9464-2

Riek R, Pervushin K, Wuthrich K (2000) TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. Trends Biochem Sci 25:462–468. doi:10.1016/S0968-0004(00)01665-0

Riek R, Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K (2002) Solution NMR techniques for large molecular and supramolecular structures. J Am Chem Soc 124:12144–12153. doi:10.1021/ja026763z

Sattler M, Fesik SW (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. Structure 4:1245–1249. doi:10.1016/S0969-2126(96)00133-5

Skrisovska L, Schubert M, Allain FHT (2010) Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. J Biomol NMR 46:51–65. doi:10.1007/s10858-009-9362-7

Sprangers R, Velyvis A, Kay LE (2007) Solution NMR of supramolecular complexes: providing new insights into function. Nat Methods 4:697–703. doi:10.1038/nmeth1080

Ton-That H, Mazmanian SK, Faull KF, Schneewind O (2000) 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 275:9876–9881. doi:10.1074/jbc.275.13.9876

Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. Annu Rev Biochem 73:107–146. doi:10.1146/annurev.biochem.73.011303.074004

Tugarinov V, Kanelis V, Kay LE (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc 1:749–754. doi:10.1038/nprot.2006.101

Wang I, Hennig J, Jagtap PK, Sonntag M, Valcarcel J, Sattler M (2014) Structure, dynamics and RNA binding of the multi-domain splicing factor TIA-1. Nucleic Acids Res 42:5949–5966. doi:10.1093/nar/gku193

Williamson DJ, Webb ME, Turnbull WB (2014) Depsipeptide substrates for sortase-mediated N-terminal protein ligation. Nat Protoc 9:253–262. doi:10.1038/nprot.2014.003

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. doi:10.1073/pnas.96.2.388

Zuger S, Iwai H (2005) Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies. Nat Biotechnol 23:736–740. doi:10.1038/nbt1097