NIH Public Access
Author Manuscript
Methods Enzymol. Author manuscript; available in PMC 2010 April 29.

Published in final edited form as:
Methods Enzymol. 2009 ; 462: 151–175. doi:10.1016/S0076-6879(09)62008-5.

Segmental Isotopic Labeling of Proteins for Nuclear Magnetic Resonance

Liu Dongsheng, **Rong Xu**, and **David Cowburn***
New York Structural Biology Center, 89 Convent Avenue, New York, NY 10027, USA

Abstract

Nuclear Magnetic Resonance (NMR) spectroscopy has emerged as one of the principle techniques of structural biology. It is not only a powerful method for elucidating the 3D structures under near physiological conditions, but also a convenient method for studying protein-ligand interactions and protein dynamics. A major drawback of macromolecular NMR is its size limitation caused by slower tumbling rates and greater complexity of the spectra as size increases. Segmental isotopic labeling allows specific segment(s) within a protein to be selectively examined by NMR thus significantly reducing the spectral complexity for large proteins and allowing a variety of solution-based NMR strategies to be applied. Two related approaches are generally used in the segmental isotopic labeling of proteins: expressed protein ligation and protein *trans*-splicing. Here we describe the methodology and recent application of expressed protein ligation and protein *trans*-splicing for NMR structural studies of proteins and protein complexes. We also describe the protocol used in our lab for the segmental isotopic labeling of a 50 kDa protein Csk (C-terminal Src Kinase) using expressed protein ligation methods.

Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy has emerged as one of the principle techniques of structural biology. It is not only a powerful method for elucidating the 3D structures under near physiological conditions, but also a convenient method for studying protein-ligand interactions as well as protein dynamics. As molecular weight increases, the size limitations of NMR become apparent – a slower tumbling rate reduces resolution, and adds greater complexity of the spectra. Slower tumbling of the macromolecule means a faster transverse relaxation rate resulting in poor signal to noise (s/n) of the spectrum. Larger size implies more signals appearing in the spectrum and this complicates the assignment of individual signals. In recent years, higher magnetic fields NMR machines have been constructed to increase resolution and signal to noise. Cryogenic probes are widely used to increase the signal to noise significantly. The development of multinuclear multidimensional experiments based on sophisticated pulse schemes has greatly enhanced the utility of NMR for the study of macromolecular structure and dynamics (Mittermaier and Kay, 2006; Wuthrich, 2003).

In the sample preparation aspect, the structural investigation of proteins by NMR has heavily depended on the incorporation of stable isotopes by increased (mainly ^{15}N, ^{13}C) or decreased (mainly ^{1}H) occurrence of an NMR-active isotope. The combination of ^{15}N, ^{13}C triple-resonance spectroscopy with deuteration, however, holds the most promise for addressing the molecular weight limitations currently imposed on structure determination by solution NMR

*Correspondence should be addressed to: David Cowburn, New York Structural Biology Center, 89 Convent Avenue, New York, NY 10027 Telephone (212) 939-0660 *9115 dcadmin@nysbc.org.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(Mittermaier and Kay, 2006). Although using deuterium labeling and TROSY type experiments (Wider, 2005; Wuthrich, 2003) addresses the issue of line width, NMR methods have still most routinely been applied to proteins under approximately 20 kDa to avoid the issues of spectral overlap and broad line widths (Card and Gardner, 2005). The segmental labeling of proteins, in which segments of the sequence with arbitrary isotopic composition are recombined, represents a significant step to reduce the spectra complexity.

The principal applications of segmental labeling (Cowburn and Muir, 2001) may be identified as (i) assignment and structure determination of large proteins and their complexes by reducing their spectra overlap, (ii) formation of labeled proteins from readily producible fragments when the complete product is impractical, or cytotoxic, for expressions and/or labeling, (iii) segmental labeling of a segment known to be a major epitope for interaction with a ligand in a fuller sequence context, (iv) observation of domain orientation and ligand perturbation in large systems using residual dipolar couplings and/or relaxation methods, (v) introduction of non-natural or modified amino acids (such as phosphotyrosine) which cannot be included in normal expression methods, and (vi) introduction of NMR unobservable tags (such as solubility enhancement tags, or lanthanide binding tags). Although there are other methods for making segmental isotopic labeled protein or RNA samples (Kim et al., 2002; Varadan et al., 2004) the main focus in this paper is expressed protein ligation and protein *trans*-splicing methods.

Expressed protein ligation and protein *trans*-splicing have been applied to a large number of different protein engineering problems. The successful implementations in segmental isotopic labeling for NMR are still somewhat limited. This may arise from several factors. If the overall aim is the reduction of spectral complexity, several issues must be addressed. First, a project involving segmental labeling will require more time and reagents than a conventional approach. If the target protein is composed of a single domain, normally a refolding procedure is required, no matter whether *trans*-splicing or the expressed protein ligation method is used. If the target protein contains several independent folding domains, the ligation can be conducted at native conditions; a refolding procedure is potentially not required. Compared with protein-peptide ligation, protein-protein ligation is generally slow. Because of this, purification methods are required remove the unligated precursors (Harris, 2006; Shi and Muir, 2005; Vitali et al., 2006; Zuger and Iwai, 2005).

Segmental isotopic labeling using expressed protein ligation

Overview

Expressed protein ligation is a protein engineering approach that allows recombinant and synthetic polypeptides to be chemoselectively and regioselectively joined together (Figure 1A) (Muir et al., 1998). It is a modified version of native chemical ligation (NCL) (Dawson et al., 1994) in which at least one of the building blocks are made by recombinant DNA methods. The first step of this process is the chemoselective transthioesterification of an unprotected protein C-terminal α-thioester with an N-terminal Cys of a second segment. The so-formed thioester spontaneously undergoes an S→N-acyl transfer to form a native peptide bond and the resulting peptide product is obtained. The reaction can proceed in aqueous conditions at neutral pH and internal Cys residues within both peptide segments are permitted. Although this procedure is relatively simple, segmental labeling using expressed protein ligation had not been applied until the introduction of the IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag) system (New England Biolabs) (Chong et al., 1997; Southworth et al., 1999). The IMPACT System contains a series of expression vectors (pTXBx, pTYBx, pTWINx, etc) which allow the fusion of the cleavable intein tag to either the C-terminus (C-terminal fusion) or N-terminus (N-terminal fusion) of a target protein (Xu and Evans, 2001). It provides a practical way to generate the reactive C-terminal α-thioester and N-terminal Cys

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

segments. The combination of IMPACT system with NCL provides a powerful method for the segmental labeling of proteins for the NMR studies (Table 1).

Ligation site

For segmental labeling of a protein required folding, a ligation site can be selected within the domain(Ottesen et al., 2008). For a ligation without refolding process, the ligation point is always chosen to be located within the linker region that connects the two domains (Vitali et al., 2006). In both cases, the residues flanking the ligation site are of special significance. One of these is the N-terminal Cys residue of the C-terminal segment, the other is the C-terminal residue forming the α-thioester for the N-terminal portion (Figure 1A). Since the Cys residue at the junction is generally regarded as obligatory for EPL, the primary consideration is to choose a ligation site before a Cys residue. In the absence of Cys at the desired ligation site, mutation usually is introduced at sites having similar residues such as Ala or Ser. In comparing a homologous family of similar proteins, it may be desirable to about mutating highly conserved residues.

Attention should also be paid to the residue at the C-terminus of the α-thioester segment (X position). Efficiency of the protein-intein thiolysis step varies and different inteins have their own set of preferences at this position (Chong et al., 1998). For example, if the target protein is fused with *Mxe* GyrA intein from pTWIN1 vector, C-terminal Asp, Pro, Gly and Ala should be avoided. If the target protein is fused with *Sce* VMA intein from pTYB1, pTYB3 vector, C-terminal Asp and Arg should be avoided because of the in vivo cleavage; Asn, Cys and Pro should also be avoided because of the possible block of cleavage. If the *Mth* RIR1 intein from pTWIN2 vector was fused to the C-terminus of the target protein, it is recommended that a Gly or Ala be present at the C-terminus of the target protein while Asp and Pro are not recommended (Xu and Evans, 2001).

As an additional consideration is that the ligation reaction rate highly relies on the molar concentration of both precursors. For the segmental labeling of protein, the ligation speed is then critical since the molar concentration of both segments cannot be very high. Although all 20 amino acids are potentially suitable for ligation, Val, Ile, and Pro represent bad choices for the α-thioester because of slow ligation rates. Rapid ligation rates are observed when the α-thioester residue X is His, Cys, or Gly (Hackeng et al., 1999).

Synthesis of a Segment with C-terminal α-thioester

Although chemical synthetic methods are feasible (e.g. (Romanelli et al., 2004)), the recombinant intein fusion expression method is the most widely used in the generation of C-terminal α-thioester for segmental labeling, because isotopic labeling is more easily obtained, and expression is more feasible for longer segments . Currently, three modified inteins from the IMPACT system can be utilized for the generation of segment C-terminal α-thioester: *Mxe* GyrA(198 aa, from pTWIN1, pTXB1,3) (Evans et al., 1998), *Mth* RIR1(134 aa, from pTWIN2) (Evans et al., 1999), and *Sce* VMA(454 aa, from pTYB1,2,3,4) (Chong and Xu, 1997). Cloning a target gene into these vectors results in the fusion of the C-terminus of the target protein to the N-terminus of an intein, which is mutated to undergo thiol-induced cleavage at its N terminus (Xu and Evans, 2001). As noted above, the C-terminal residue of the target protein is critical because the *in vivo* and thiol-induced cleavage at the intein can be dramatically affected by it.

Synthesis of a Segment with N-terminal Cysteine

Three methods generally used in the preparation of segments with N-terminal Cys are chemical synthesis, cleavage of a precursor sequence expressed with the Cys residue adjacent to a cleavage site, and intein-mediated recombinant methods (Muralidharan and Muir, 2006).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Several cleavage methods can be utilized: (1) Methionyl aminopeptidase (Camarero et al., 2001; Gentle et al., 2004; Iwai and Pluckthun, 1999); (2) Factor Xa protease (Camarero et al., 2002; Romanelli et al., 2004; Xu et al., 1999) (3) TEV protease (Tolbert et al., 2005; Tolbert and Wong, 2004); (4) Cyanogen bromide (Macmillan and Arham, 2004); (5)SUMO protease (Weeks et al., 2007); (6)Leader peptidase (Hauser and Ryan, 2007); (7) EnteroKinase (Hosfield and Lu, 1999); (8) Thrombin. Recently we found that the leader sequence LVPRC (modified from the optimal thrombin cleavage sequence LVPRGS) can also be cleaved by thrombin with lower efficiency. Thus thrombin cleavage can also be used in the preparation of protein segments with C-terminal Cys(Liu et al., 2008). Each cleavage method has its own limitations. For example, nonspecific proteolysis can be observed, when there are sequences within the protein similar to the cleavage site; Cyanogen bromide cleavage requires denaturation and no other Met residues in the target protein; Endogenous Methionyl aminopeptidase is often inefficient and results in low yields of the desired material.

N-terminal fusion with an intein can also be exploited to prepare N-terminal Cys segments. The modified intein *Ssp* DnaB (Mathys et al., 1999), *Mth* RIR1 (Evans et al., 1999), *Mxe* GyrA (Southworth et al., 1999) have all been engineered for the generation of N-terminal Cys proteins. pTWIN1 and pTWIN2 vector from IMPACT system can be utilized for this purpose using the *Ssp* DnaB N-terminal fusion. A target protein sequence was fused in-frame to the C-terminus of the intein. Following expression in *E. coli* , the fusion protein was bound to a chitin resin and the cleavage of intein fusion is induced in a pH and temperature-dependent fashion. Note that sometimes the protein undergoes significant *in vivo* cleavage, causing the loss of the affinity tag (Mathys et al., 1999).

Ligation protocol

The isolated α-thioester and N-terminal Cys segments can be ligated in the presence of many additives such as denaturents, chaotropes, or detergents. The optimum pH of the ligation reaction is 7–8 and the presence of a moderate concentration of the thiol reagent such as MESNA (2-mercaptoethanesulfonic acid) is always required. The thiol-induced cleavage and ligation can also be carried out simultaneously by adding both thiol and N-terminal Cys segment (Camarero et al., 2002; Vitali et al., 2006) (Table 1). A reactive C-terminal thioester can be generated on the target protein when an appropriate thiol compound is used to induce cleavage of the intein. MESNA is an attractive reagent since it provides higher ligation efficiency, and is odorless and extremely soluble in aqueous solutions (Ayers et al., 1999; Evans et al., 1998). Johnson and Kent have compared the use of a number of thiol compounds find a highly effective and practical catalyst MPAA(4-carboxylmethyl)thiophenol) (Johnson and Kent, 2006). Ligation in the presence of MPAA showed an order of magnitude (i.e. ~10-fold) rate increase over the standard catalyst mix used for ligation. The authors also suggested that MPAA will be used as a much more effective alternative to MESNA for use in the semi-synthesis of protein by expressed protein ligation. Recently a study suggests reaction of DTT-thioester protein and Tris(tris(hydroxymethyl) aminomethane) buffer can takes place, and a Tris adduct is formed, which is fused to the C-terminus of the protein by a stable amide bond (Peroza and Freisinger, 2008). Although the reactions of other thioesters/thiol systems with Tris are not reported, caution is appropriate in the use of Tris buffer.

Segmental labeling using protein *Trans*-splicing

The ligation of proteins segments can also be accomplished by *trans*-splicing which relies on the high affinity and catalytic activity by the two halves of an intein to ligate the two extein sequence (Figure 1B). The N- and C-terminal pieces individually have no splicing activity but that when combined; associate non-covalently to give a functional protein splicing element. The *trans*-splicing reaction is initiated by binding of the intein fragments together, where expressed protein ligation reaction lacks this favorable binding. Several split inteins have been

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

studied including native split *Ssp* DnaE (Evans et al., 2000; Martin et al., 2001), artificially split PI-*Pfu*I (Otomo et al., 1999b; Yagi et al., 2004; Yamazaki et al., 1998), *Mtu* RecA (Lew et al., 1998; Mills and Perler, 2005; Shingledecker et al., 1998), *Psp* Pol-1 (Southworth et al., 1998), *Ssp* DnaB (Brenzel et al., 2006), *Sce* VMA (Mootz et al., 2003; Schwartz et al., 2007). Among these split inteins, *Ssp* DnaB, *Sce* VMA, *Ssp* DnaE are reported to be active in protein *trans*-splicing under native conditions, while others require a denaturation-refolding treatment of the mixed intein halves (Otomo et al., 1999a; Otomo et al., 1999b; Yagi et al., 2004; Yamazaki et al., 1998). The naturally occurring *Ssp* DnaE intein has been shown to mediate protein *trans*-splicing *in vivo* under native conditions and provide a segmental isotope method in the living cells (Zuger and Iwai, 2005). It has been recommended to include several residues flanking the target protein sequence to enhance the ligation efficiency (Xu and Evans, 2001). Table 2 summarizes the typical use of protein *trans*-splicing methods for the segmental labeling for NMR sample preparations.

Multiple segment assembly

For extremely large systems, segmental labeling involving more than two domains may be necessary. Both expressed protein ligation (Blaschke et al., 2000; Cotton et al., 1999; Ottesen et al., 2008)and protein *trans*-splicing (Otomo et al., 1999a; Shi and Muir, 2005) methods can be utilized in this task. Using expressed protein ligation, a protecting group on N-terminal Cys of a central segment should be introduced to avoid self-ligation and/or cyclization when the central segment(s) is/are ligated with the C-terminal segment. The protection group can be a protease cleavable peptide. After the central segment is ligated with the C-terminal segment, the protecting group is removed and the resulted protein ligated with N-terminal segment α-thioester. For the *trans*-splicing protocol, two separate inteins, which should be orthogonal to each other, may be used to ligate N- and C-terminal segments to both the ends of middle segment.

Segmental Labeling of C-terminal Src Kinase(Csk)

Overview

Csk has been shown to be important in the regulation of neural development, T-cell development and regulation and cyto-skeletal organization (Cole et al., 2003; Roskoski, 2005). Topographically, Csk is similar to the Src family kinase, having a SH3 and SH2 domain pair followed by a catalytic domain. A proline-rich sequence from the protein tyrosine phosphatases (murine PEP, (Cloutier and Veillette, 1996), human *PTPN22*/Lyp (Bottini et al., 2004)) can binds to the Csk-SH3 domain.Csk is a 50 kDa protein that is not good target for traditional NMR studies because of its size and self association. Segmental labeling is a promising approach to overcoming the size issue. From preliminary studies, we knew the refolding of Csk is very difficult, so we decided use expressed protein ligation do the segmental labeling. We chose pTWIN1 expression vector (New England Biolabs) for the expression of both of the segments based on these considerations with a view to obtaining segmentally labeled, full length Csk. At the same time, this approach provides SH32 and kinase domain separately and conveniently. We also need a fusion system to overcome the poor expression and folding of the kinase domain; pTWIN1 vectors are designed for protein purification or for the isolation of proteins with an N-terminal cysteine and/or a C-terminal thioester. It has a modified *Ssp* DnaB intein as N-terminal fusion (intein 1) and *Mxe* GyrA intein as C-terminal fusion (intein 2). The presence of the N- and C- terminal chitin binding domain (CBD) facilitates purification.

Csk can be divided into three domains -- SH3, SH2 and kinase domain (Figure 2). The ligation point is chosen to be located within the SH2-kinase linker region that connects the two domains, residues V172-E194 (Holtrich et al., 1991;Partanen et al., 1991). Since there is no Cys in the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

linker region, we have to find a similar residue that can be mutated to Cys with minimal change. Ser and Ala are good choices for mutation to Cys. A178, A179 and S186 can be considered as potential mutation sites. A178 and S186 were not used because the proceeding residues, V177 and R185 respectively, would likely result in low efficiency for the ligation (Hackeng et al., 1999). So we chose A179 to mutate to Cys and A178 to form an α-thioester. In the crystal structure of Csk, A178 and A179 are located at the N-terminal of a small helix αBC (Ogawa et al., 2002) and it was anticipated that the mutations would lead to minimal structural perturbations. In the following steps, the SH32 and kinase domain sequence were cloned into the expression vector, which allowed the generation of SH32-intein2-CBD and CBD-intein1-kinase fusion proteins, respectively.

Cloning Csk SH32 and Kinase gene to expression vector

The Csk SH32 domain gene, containing residues M1-A178 was amplified from human full-length Csk gene by polymerase chain reaction (PCR) using primers 5'-GGT GGT CAT ATG TCA GCA ATA CAG GCC (SH3_intein_fw) and 5'- GGT GGT TGC TCT TCC GCA CGC CAC TGT GCC CTC CAT (SH2_intein_rev) with Nde I and Sap I cleavage sites. Note that a stop codon should not be included in the reverse primer for the sequential expression of *Mxe* GyrA intein (intein2). The Sap I site will be lost after cloning so that no additional residue was introduced between SH32 and intein2. All of the PCR reactions were in 50 μl final volume with pfu polymerase (Novagen), 0.5 mM of all four dNTPs (Sigma), 0.2 μM of each oligonucleotide and 50 ng plasmid DNA. PCR was carried out in a Mastercycler (Eppendorf) machine and DNA was amplified by 30 cycles under standard conditions with annealing at 60° C and enlongation at 72°C. The PCR product was purified, digested with Nde I and Sap I, and ligated to the similarly digested pTWIN1 vector. After transformation of the plasmid into *E.coli* DH5a, the clone was validated by DNA sequencing using T7 universal primer (5'-TAA TAC GAC TCA CTA TAG GG). The vector was named pTWIN1-SH32. The corresponding protein product was named SH32-intein2-CBD (Figure 3).

The Csk kinase domain gene containing residues A179-L450 with an A179C mutation, with or without C-terminal His-tag, was amplified from full-length Csk gene by PCR using primers containing Sap I and BamH I sites. For the cloning of kinase without His-tag, primers: 5'-GGT GGT TGC TCT TCC AAC TGC CAG GAT GAG TTC TAC CGC (kinase_intein_fw) and 5'-GGT GGT GGA TCC TTA CAG GTG CAG CTC GTG GGT (kinase_intein_rev) were used. For the cloning of kinase with C-terminal His-tag, primers: kinase_intein_fw and 5'-GGT GGT GGA TCC TTA ATG GTG ATG GTG ATG GTG CAG GTG CAG CTC GTG GGT (kinase_intein_his_rev) were used. Note that a stop codon should be included in the reverse primer. The Sap I site is lost after cloning so that no additional residue is introduced between *Ssp* DnaB intein (intein1) and the kinase domain. The amplified gene was cloned into the expression vector pTWIN1 and the sequences were confirmed by DNA sequencing using *Ssp* DnaB forward primer (5´-ACT GGG ACT CCA TCG TTT CT) and T7 terminator reverse primer (5'- TAT GCT AGT TAT TGC TCA G). The designed fusion proteins without and with a C-terminal His-tag are named pTWIN1-kinase and pTWIN1-kinase-His, respectively (Figure 3). The corresponding protein products of pTWIN1-kinase and pTWIN1-kinase-his are named CBD-intein1-kinase and CBD-intein1-kinase-his, respectively. Both of the fusion proteins contain A179C mutations.

Expression and purification Csk SH32 with C-terminal *Mxe* GyrA intein (intein2)

For the expression of SH32-intein2-CBD protein, the plasmid pTWIN1-SH32 was transformed into BL21-CondonPlus (DE3)-RIL Component Cell (Stratagene, 230245). A fresh cell harboring pTWIN-SH32 was inoculated in 20 ml of LB medium containing 100 μg/ml ampicillin and cultured at 37°C until the optical density (A600) reached 0.8. The whole culture was then transferred to 1 L of the same medium and cultured at 37°C until the optical density

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(A600) reached 0.6. The cells were induced with 0.5 mM IPTG at 30°C. For the isotopic labeling, M9 media with 0.1% $^{15}NH^4Cl$ and/or 0.2% [U-^{13}C] glucose was used. After induction at 30°C overnight, the cell pellet was collected by centrifugation at 4000 g for 20 min at 4°C, and then resuspended in 35 ml of Buffer A (50 mM Tris-HCl, pH 7.5 200 mM NaCl) and stored at −80°C for future use. No *in vivo* cleavage of the SH32-intein2-CBD was found during the processing procedure above. As a practical note, in systems that are susceptible to *in vivo* degradation we recommend that the cell growth and protein purification steps be performed in the same day to improve yields.

Cells containing the SH32-intein2-CBD protein were passed through a French press cell twice to break the cells. The cell lysate was centrifuged at 12,000 g for 20 min at 4°C. The clarified cell extract was loaded to 6 ml of Chitin Beads (New England Biolabs S6651L) which pre-equilibrated with Buffer A. The column was washed with 50 ml of Buffer A to remove the unbound protein. The column bound with protein SH32-intein2-CBD was stored at 4°C for the ligation with kinase domain. The fusion protein was stable for at least four days at 4°C with no thiol reducing agent added. The longtime storage should be at −80 °C. For the purification of the SH32 domain, the cleavage of the intein-tag was induced by equilibrating the chitin beads with 50 mM DTT, 50 mM KPi, pH 7.2. After 24 hr of on-column cleavage at room temperature, the target protein was eluted from the column by 40 ml of Buffer B (20 mM Tris-HCl, pH 8.0) and loaded onto a Mono Q column equilibrated with Buffer B at flow rate 1 ml/min (Figure 4A). The elution was conducted by linear gradient to 60% Buffer C (20 mM Tris-HCl, pH 8.0, 1.0 M NaCl) in 60 min. The major peak near 34% of Buffer C was collected and identified as SH32.

Expression and purification of Csk kinase with N-terminal *Ssp* DnaB intein (CBD-intein1-kinase)

For the expression of CBD-intein1-kinase or CBD-intein1-kinase-his, the plasmid pTWIN1-kinase or pTWIN1-kinase-his was transformed into BL21-CondonPlus (DE3)-RIL Component Cell (Stratagene, 230245). A fresh cell harboring pTWIN-SH32 was inoculated in 20 ml of LB medium containing 100 μg/ml ampicillin and cultured at 37°C until the optical density (A600) reached 0.8. The whole culture was then transferred to 1L of the same medium and cultured at 37°C until the optical density (A600) reached 0.6. The cells were induced with 0.2 mM IPTG at 15°C for 16 hr (for LB or non-deuterated M9 media) or 40 hr (when deuteration applied). After centrifugation at 4,000 g for 20 min, the cells were resuspended in 35 ml of Buffer A and stored at −80°C for future use.

Partial degradation was found for the two fusion protein after the induction. Two bands 60 kDa and 30 kDa appeared on the SDS-PAGE gel corresponding to the fusion and degradation products respectively. The calculated molecular weight for the CBD-intein1-kinase, CBD-intein1, and kinase were 60, 30, and 30 kDa respectively. Since the cell growth in D_2O is much slower than the H_2O media, the degradation was more severe. This makes it impossible to use chitin beads to purify the kinase protein. We used CBD-intein1-kinase-his protein in the ligation experiments. Although the degradation still remains, both non-degraded fusion-kinase and degraded kinase domain can bind the metal affinity beads.

For the purification of the C-terminal kinase protein, cells containing the CBD-intein1-kinase-his and its degradation products were passed through a French pressure cell twice to break the cells at 4°C. The lysate was centrifuged at 12,000 g for 20 min at 4°C to remove the cell debris. The clarified cell extract was loaded to 10 ml of TALON beads charged with Co^{2+} and pre-equilibrated with Buffer A. The column was washed with 50 ml of Buffer A to remove the unbound protein whereas the degraded and intein1 fusion kinase-his can bind to the column. The bound protein was eluted by Buffer A with 150 mM imidozole and the total elusion volume was about 10 ml. The purified protein can be stored at 4°C for several weeks. For long time

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

storage, glycerol was added to final 25% (v/v) and stored at −80°C. In the case of kinase used in ligation, no additional purification procedures were required. If the kinase was directly used in the NMR experiments without ligation, additional purification procedures are needed. For the purification of the kinase domain, the pH of elusion protein (mixture of degraded kinase-his and fusion CBD-chitin-kinase-his protein) was adjusted to 7.0 with 1 M Tris-HCl (pH 6.5) buffer and stored at 15°C for 2 days to let the fusion protein degrade. The degraded protein mixture was passed through a chitin column equilibrated with Buffer A. CBD-intein1 bind to the column and the flow through was collected, exchanged to Buffer B and loaded onto a Mono Q column equilibrated with Buffer B at flow rate 1 ml/min. The elution was conducted by linear gradient to 50% Buffer C in 50 min. The peak near 15.5% of Buffer C was collected and identified as kinase-his with an N-terminal Cysteine (Figure 4B).

Ligation of SH32 domain testing peptide

An easy way to test whether the α-thioester can formed after the MESNA thiolysis is to use a fluorescent peptide. A model peptide NH_2-CGRGRGRK[fluorescein]-$CONH_2$ was chemically synthesized by New England Peptide Inc. The observed molecular weight was 1246 Da. This peptide was used to monitor reactivity of SH32 (Ayers et al., 1999; Xu et al., 1999). The peptide was designed to be highly soluble and the fluorescein probe was incorporated to provide a convenient way to monitoring the progress of the ligation reaction by using SDS-PAGE and other methods. In preliminary ligation studies, we investigated whether this peptide could be ligated with Csk SH32. Nearly quantitative ligation of the synthetic peptide to the corresponding protein domain with C-terminal thioester was observed when a large access of peptide was added (molar ratio 1:10) as indicated by SDS-PAGE. Typically, 20 μl of chitin beads with loaded SH32-intein2-CBD was used. To these beads was added a solution of synthetic peptide (1 mg/ml) in Buffer A along with 3% MESNA. The suspension was then gently agitated at room temperature overnight, the supernatant was removed and analysis performed by SDS-PAGE. The ligation of the synthetic peptide to the recombinant protein was evidenced by the production of a highly fluorescent 21-kDa protein band on SDS-PAGE. These studies thus established that the C-terminal thioester was successfully formed during the MESNA thiolysis. After 48 h of storage at 15°C, the SH32 α-thioester failed to ligate with the synthetic peptide.

Ligation of SH32 with kinase domain

The kinase mixture (CBD-intein1-kinase-his, kinase-his) from the TALON beads was concentrated to about 1 mM using Amicon Ultra centrifugal filter (Millipore, UFC901024) at 4°C. The kinase protein mixture was flushed to the chitin column bound with SH32-intein2-CBD in the presence of Buffer A. The molar ratio of SH32 α-thioester and N-terminal Cysteine kinase was estimated to be 1:1 from the SDS-PAGE of the solution. 0.5 mM EDTA and 200 mM of MESNA were added to the system and the column was stored at 15°C for 72 hr. During this time, three reactions proceed: thiolysis of SH32-intein2-CBD to form SH32 α-thioester; degradation of the fusion CBD-intein1-kinase-his to form kinase-his; and reaction of SH32 α-thioester with kinase-his with an N-terminal Cysteine. SDS-PAGE was used to monitor the ligation process. The speed of the MESNA induced cleavage is dependent on the concentration of MESNA and the period of incubation. We increased the MESNA to 200 mM to increase the cleavage speed since the Ala at the C-terminal of the SH32 is generally not recommended in the pTWIN1 C-terminal fusion. On-column ligation has been reported to increase the ligation efficiency (Vitali et al., 2006). During 3 days ligation, about 20–25% of the SH32 and kinase protein form the ligated segmental labeled protein.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Purification of ligation product

Because of the low ligation efficiency, removing the unligated protein from the ligated sample is critical. Figure 5 shows the combination of Co^{2+} column and Q column for the purification procedure. After 72 h ligation, the ligation mixture was eluted from the chitin column using Buffer A. The amount of protein was monitored using the Coomassie Plus-200 protein assay reagent (Pierce, 23238). Elution was continued until no more protein was detected – typically about 30 ml. The eluate was dialysised against 1.0 L Buffer A at 4°C for 4 hr, twice, to completely remove the MESNA and EDTA. The solution was then loaded to 10 ml of TALON beads charged with Co^{2+} and pre-equilibrated with Buffer A. The column was washed with 50 ml of Buffer A to remove the unbound protein (mainly inteins and SH32). The peptide PEP-3BP1 is used to reduce the self-association of the SH3 domain (Borchert et al., 1994;Ghose et al., 2001). The protein tyrosine phosphatase PEP proline-rich peptide (PEP-3BP1, residues 605–629, SRRTDDEIPPPLPERTPESFIVVEE) was expressed and purified as described (Ghose et al., 2001). The column was washed with another 10 ml Buffer A with 1 mM PEP-3BP1 peptide in it to remove the SH32 domain bind with the ligation product due to the dimerization of SH3 domain. The bound protein (mainly SH32-kinase-his, kinase-his) was eluted by Buffer A with 150 mM imidozole and the total elusion volume was about 10 ml. 40 ml Buffer B was added to the elusion to reduce the salt concentration and the total 50 ml elusion was loaded onto a Mono Q column equilibrated with Buffer B at flow rate 1 ml/min (Figure 4C). The elution was conducted by linear gradient to 60% Buffer C (20 mM Tris-HCl, pH 8.0, 1.0 M NaCl) in 60 min. The major peak near 34% of Buffer C was collected and identified as SH32-kinase-his ligation product.

NMR Spectroscopy

^{1}H-^{15}N HSQC spectra were acquired on the segmental labeled Csk sample at 25°C, 800 MHz. Figure 6(A) is the ^{1}H-^{15}N HSQC of [^{15}N-SH32]-Kinase, about 160 amide peaks can be identified from this spectrum. Figure 6(B) is ^{1}H-^{15}N HSQC SH32-[^{15}ND-Kinase], about 240 amide peaks can be identified from this spectrum. All NMR experiments were performed on Bruker 800 MHz, US^{2} spectrometers, which were equipped with triple-resonance cryo-probes. Protein solutions were prepared in the following buffer conditions: 50 mM Tris-HCl (pH 7.5), 10 mM DTT, 1.0 mM EDTA, 0.01% (w/v) NaN3, 5% D2O, 0.1 mM DSS (4,4-dimethyl-4-silapentane-1-sulfonate). All the NMR samples were sealed in the NMR tube under a nitrogen atmosphere. Pulsed-field gradient techniques with a WATERGATE pulse sequence were used for all H_2O experiments, resulting in a good suppression of the solvent signal. The ^{1}H chemical shifts were referenced to internal DSS. In the 2D ^{1}H-^{15}N HSQC experiments, 512 complex points were collected in the ^{1}H dimension and 128 complex points in the 15N dimension. The corresponding sweep widths were 14 and 36 ppm in the ^{1}H and ^{15}N dimensions, respectively. ^{15}N dimension were multiplied by a cosine-bell window function and zero-filled to 512 points before Fourier transformation using Topspin 1.4. The ^{15}N chemical shift was referenced indirectly using the $^{1}H/^{15}N$ frequency ratios of the zero point 0.101329118 (^{15}N) (Live et al., 1984; Wishart et al., 1995).

Acknowledgments

We are grateful to Dr. Michael Goger, Dr. Kaushik Dutta, and Dr. Shibani Bhattacharya for their assistance in the NMR instrumentation and Professor Tom Muir for advice and discussion. Supported by NIH grants EB001991 and GM47021.

References

Ayers B, et al. Introduction of unnatural amino acids into proteins using expressed protein ligation. Biopolymers 1999;51:343–354. [PubMed: 10685045]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Blaschke UK, et al. Synthesis of multi-domain proteins using expressed protein ligation: Strategies for segmental isotopic labeling of internal regions. Tetrahedron 2000;56:9461–9470.

Borchert TV, et al. The Crystal-Structure of Human Csksh3 - Structural Diversity near the Rt-Src and N-Src Loop. Febs Lett 1994;341:79–85. [PubMed: 7511113]

Bottini N, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nature Genetics 2004;36:337–338. [PubMed: 15004560]

Brenzel S, et al. Engineering artificially split inteins for applications in protein chemistry: Biochemical characterization of the split Ssp DnaB intein and comparison to the split Sce VMA intein. Biochemistry 2006;45:1571–1578. [PubMed: 16460004]

Camarero JA, et al. Rescuing a destabilized protein fold through backbone cyclization. Journal of Molecular Biology 2001;308:1045–1062. [PubMed: 11352590]

Camarero JA, et al. Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. Proceedings of the National Academy of Sciences of the United States of America 2002;99:8536–8541. [PubMed: 12084914]

Card PB, Gardner KH. Identification and optimization of protein domains for NMR studies. Methods in Enzymology 2005;394:3–16. [PubMed: 15808215]

Chong SR, et al. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 1997;192:271–281. [PubMed: 9224900]

Chong SR, et al. Modulation of protein splicing of the Saccharomyces cerevisiae vacuolar membrane ATPase intein. Journal of Biological Chemistry 1998;273:10567–10577. [PubMed: 9553117]

Chong SR, Xu MQ. Protein splicing of the Saccharomyces cerevisiae VMA intein without the endonuclease motifs. Journal of Biological Chemistry 1997;272:15587–15590. [PubMed: 9188443]

Cloutier JF, Veillette A. Association of inhibitory tyrosine protein kinase p50(csk) with protein tyrosine phosphatase PEP in T cells and other hemopoietic cells. Embo Journal 1996;15:4909–4918. [PubMed: 8890164]

Cole PA, et al. Protein tyrosine kinases Src and Csk: a tail's tale. Current Opinion in Chemical Biology 2003;7:580–585. [PubMed: 14580561]

Cotton GJ, et al. Insertion of a synthetic peptide into a recombinant protein framework: A protein biosensor. Journal of the American Chemical Society 1999;121:1100–1101.

Cowburn D, Muir TW. Segmental isotopic labeling using expressed protein ligation. Methods in Enzymology 2001;339:41–54. [PubMed: 11462824]

Dawson PE, et al. Synthesis of Proteins by Native Chemical Ligation. Science 1994;266:776–779. [PubMed: 7973629]

Evans TC, et al. Semisynthesis of cytotoxic proteins using a modified protein splicing element. Protein Science 1998;7:2256–2264. [PubMed: 9827992]

Evans TC, et al. The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. Journal of Biological Chemistry 1999;274:18359–18363. [PubMed: 10373440]

Evans TC, et al. Protein trans-splicing and cyclization by a naturally split intein from the dnaE gene of Synechocystis species PCC6803. Journal of Biological Chemistry 2000;275:9091–9094. [PubMed: 10734038]

Gentle IE, et al. Direct production of proteins with N-terminal cysteine for site-specific conjugation. Bioconjugate Chemistry 2004;15:658–663. [PubMed: 15149194]

Ghose R, et al. A novel, specific interaction involving the Csk SH3 domain and its natural ligand. Nature Structural Biology 2001;8:998–1004.

Grant JE, et al. The N terminus of GTP gamma S-activated transducin alpha-subunit interacts with the C terminus of the cGMP phosphodiesterase gamma-subunit. Journal of Biological Chemistry 2006;281:6194–6202. [PubMed: 16407279]

Hackeng TM, et al. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. Proceedings of the National Academy of Sciences of the United States of America 1999;96:10068–10073. [PubMed: 10468563]

Harris TK. Discovering new drug-targeting sites on flexible multidomain protein kinases: Combining segmental isotopic and site-directed spin labeling for nuclear magnetic resonance detection of interfacial clefts. Methods Molecular Biology 2006;316:199–225.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Hauser PS, Ryan RO. Expressed protein ligation using an N-terminal cysteine containing fragment generated in vivo from a pe1B fusion protein. Protein Expression and Purification 2007;54:227–233. [PubMed: 17493830]

Holtrich U, et al. 2 Additional Protein-Tyrosine Kinases Expressed in Human Lung - 4th Member of the Fibroblast Growth-Factor Receptor Family and an Intracellular Protein-Tyrosine Kinase. Proceedings of the National Academy of Sciences of the United States of America 1991;88:10411–10415. [PubMed: 1720539]

Hosfield T, Lu Q. Influence of the amino acid residue downstream of (Asp)(4)Lys on enterokinase cleavage of a fusion protein. Analytical Biochemistry 1999;269:10–16. [PubMed: 10094769]

Iwai H, Pluckthun A. Circular beta-lactamase: stability enhancement by cyclizing the backbone. Febs Lett 1999;459:166–172. [PubMed: 10518012]

Johnson ECB, Kent SBH. Insights into the mechanism and catalysis of the native chemical ligation reaction. Journal of the American Chemical Society 2006;128:6640–6646. [PubMed: 16704265]

Kim I, et al. NMR study of 100 kDa HCV IRES RNA using segmental isotope labeling. Journal of the American Chemical Society 2002;124:9338–9339. [PubMed: 12167005]

Lew BM, et al. Protein splicing in vitro with a semisynthetic two-component minimal intein. Journal of Biological Chemistry 1998;273:15887–15890. [PubMed: 9632632]

Liu DS, et al. N-terminal cysteinyl proteins can be prepared using thrombin cleavage. Febs Lett 2008;582:1163–1167. [PubMed: 18331839]

Live DH, et al. Long-Range Hydrogen-Bond Mediated Effects in Peptides - N-15 Nmr-Study of Gramicidin-S in Water and Organic-Solvents. Journal of the American Chemical Society 1984;106:1939–1941.

Macmillan D, Arham L. Cyanogen bromide cleavage generates fragments suitable for expressed protein and glycoprotein ligation. Journal of the American Chemical Society 2004;126:9530–9531. [PubMed: 15291543]

Martin DD, et al. Characterization of a naturally occurring trans-splicing intein from Synechocystis sp PCC6803. Biochemistry 2001;40:1393–1402. [PubMed: 11170467]

Mathys S, et al. Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. Gene 1999;231:1–13. [PubMed: 10231563]

Mills KV, Perler FB. The mechanism of intein-mediated protein splicing: Variations on a theme. Protein and Peptide Letters 2005;12:751–755. [PubMed: 16305544]

Mittermaier A, Kay LE. Review - New tools provide new insights in NMR studies of protein dynamics. Science 2006;312:224–228. [PubMed: 16614210]

Mootz HD, et al. Conditional protein splicing: A new tool to control protein structure and function in vitro and in vivo. Journal of the American Chemical Society 2003;125:10561–10569. [PubMed: 12940738]

Muir TW, et al. Expressed protein ligation: A general method for protein engineering. Proceedings of the National Academy of Sciences of the United States of America 1998;95:6705–6710. [PubMed: 9618476]

Muralidharan V, Muir TW. Protein ligation: an enabling technology for the biophysical analysis of proteins. Nature Methods 2006;3:429–438. [PubMed: 16721376]

Ogawa A, et al. Structure of the carboxyl-terminal Src kinase, Csk. Journal of Biological Chemistry 2002;277:14351–14354. [PubMed: 11884384]

Otomo T, et al. NMR observation of selected segments in a larger protein: Central-segment isotope labeling through intein-mediated ligation. Biochemistry 1999a;38:16040–16044. [PubMed: 10587426]

Otomo T, et al. Improved segmental isotope labeling of proteins and application to a larger protein. Journal of Biomolecular Nmr 1999b;14:105–114. [PubMed: 10427740]

Ottesen JJ, et al. An amalgamation of solid phase peptide synthesis and ribosomal peptide synthesis. Biopolymers 2008;90:406–414. [PubMed: 17636509]

Partanen J, et al. Cyl Encodes a Putative Cytoplasmic Tyrosine Kinase Lacking the Conserved Tyrosine Autophosphorylation Site (Y416src). Oncogene 1991;6:2013–2018. [PubMed: 1945408]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Peroza EA, Freisinger E. Tris is a non-innocent buffer during intein-mediated protein cleavage. Protein Expression and Purification 2008;57:217–225. [PubMed: 18032065]

Romanelli A, et al. Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. Proceedings of the National Academy of Sciences of the United States of America 2004;101:6397–6402. [PubMed: 15087498]

Roskoski R. Src kinase regulation by phosphorylation and dephosphorylation. Biochemical and Biophysical Research Communications 2005;331:1–14. [PubMed: 15845350]

Schwartz EC, et al. Post-translational enzyme activation in an animal via optimized conditional protein splicing. Nature Chemical Biology 2007;3:50–54.

Shi JX, Muir TW. Development of a tandem protein trans-splicing system based on native and engineered split inteins. Journal of the American Chemical Society 2005;127:6198–6206. [PubMed: 15853324]

Shingledecker K, et al. Molecular dissection of the Mycobacterium tuberculosis RecA intein: design of a minimal intein and of a trans-splicing system involving two intein fragments. Gene 1998;207:187–195. [PubMed: 9511761]

Skrisovska L, Allain FH. Improved segmental isotope labeling methods for the NMR study of multidomain or large proteins: application to the RRMs of Npl3p and hnRNP L. Journal of Molecular Biology 2008;375:151–164. [PubMed: 17936301]

Southworth MW, et al. Control of protein splicing by intein fragment reassembly. Embo Journal 1998;17:918–926. [PubMed: 9463370]

Southworth MW, et al. Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. Biotechniques 1999;27:110–120. [PubMed: 10407673]

Tolbert TJ, et al. A new strategy for glycoprotein synthesis: ligation of synthetic glycopeptides with truncated proteins expressed in E-coli as TEV protease cleavable fusion protein. Bioorganic & Medicinal Chemistry 2005;13:909–915. [PubMed: 15653356]

Tolbert TJ, Wong CH. Conjugation of glycopeptide thioesters to expressed protein fragments: semisynthesis of glycosylated interleukin-2. Methods Molecular Biology 2004;283:255–266.

Varadan R, et al. Solution conformation of Lys(63)-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. Journal of Biological Chemistry 2004;279:7055–7063. [PubMed: 14645257]

Vitali F, et al. Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. Embo Journal 2006;25:150–162. [PubMed: 16362043]

Weeks SD, et al. Ligation independent cloning vectors for expression of SUMO fusions. Protein Expression and Purification 2007;53:40–50. [PubMed: 17251035]

Wider G. NMR techniques used with very large biological macromolecules in solution. Methods in Enzymology 2005;394:382–398. [PubMed: 15808229]

Wishart DS, et al. H-1, C-13 and N-15 Chemical-Shift Referencing in Biomolecular Nmr. Journal of Biomolecular Nmr 1995;6:135–140. [PubMed: 8589602]

Wuthrich K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). Journal of Biomolecular Nmr 2003;27:13–39. [PubMed: 15143746]

Xu MQ, Evans TC. Intein-mediated ligation and cyclization of expressed proteins. Methods 2001;24:257–277. [PubMed: 11403575]

Xu R, et al. Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies. Proceedings of the National Academy of Sciences of the United States of America 1999;96:388–393. [PubMed: 9892643]

Yagi H, et al. Conformational change of H+ATPase beta monomer revealed on se gmental isotope labeling NMR spectroscopy. Journal of the American Chemical Society 2004;126:16632–16638. [PubMed: 15600369]

Yamazaki T, et al. Segmental isotope labeling for protein NMR using peptide splicing. Journal of the American Chemical Society 1998;120:5591–5592.

Zuger S, Iwai H. Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies. Nature Biotechnology 2005;23:736–740.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

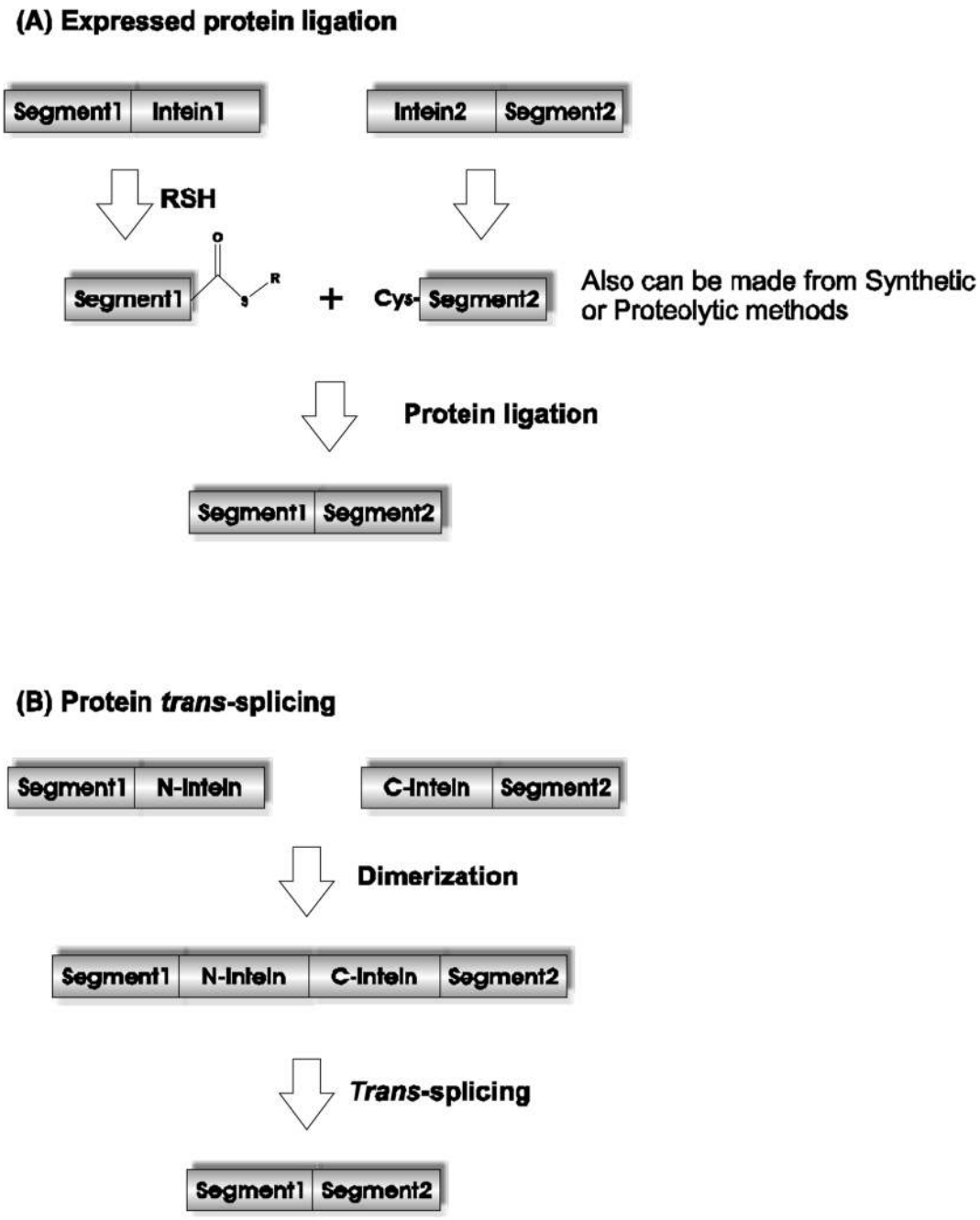

Figure 1.

(A), Segmental isotopic labeling using expressed protein ligation. Intein-based fusion approaches allow the straightforward introduction of α-thioester and N-terminal Cys. N-terminal Cys can also be made by cleavage of the leader sequence and chemical synthesis methods. (B), Segmental labeling using protein *trans*-splicing. The N- and C-terminal pieces (referred to in the figure as N-intein and C-intein) individually have no activity but that when combined, associate noncovalently to give a functional protein splicing element. Conserved residues around the splicing junction are: a Cys/Ser at the N-terminus of intein, a His-Asn sequence at the C-terminus of intein, and a Ser/Thr/Cys at the N-terminus of C-terminal extein (segment 2)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

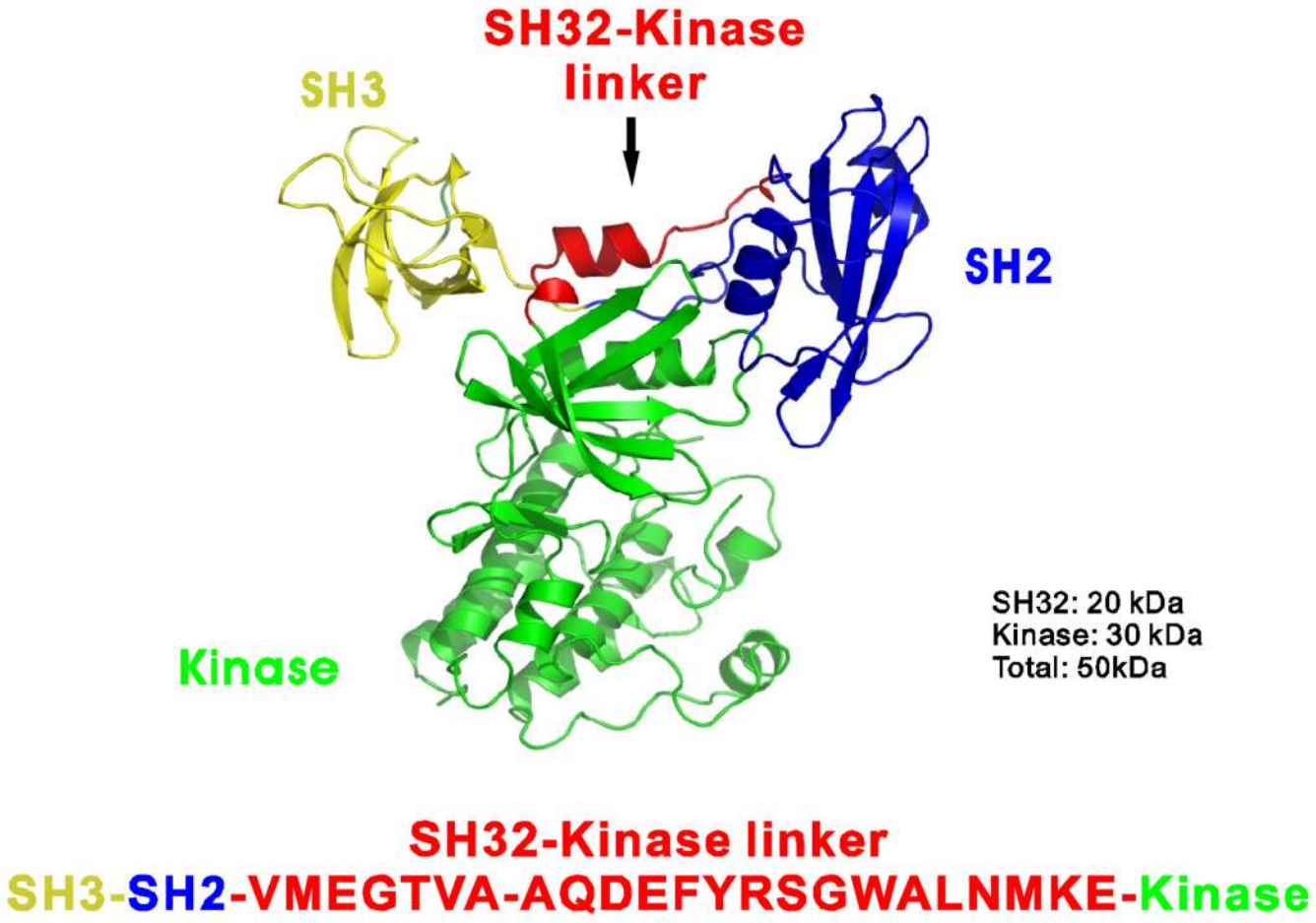

Figure 2.
Structure of full-length Csk and the sequence of SH32-kinase linker (modeled from 1K9A (Ogawa et al., 2002)). The SH3, SH2, and kinase domain indicated. A178 was chosen to form an α-thioester and A179 to mutate into Cys. In the structure of Csk, A178 and A179 are located at the N-terminal of a small helix αBC.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

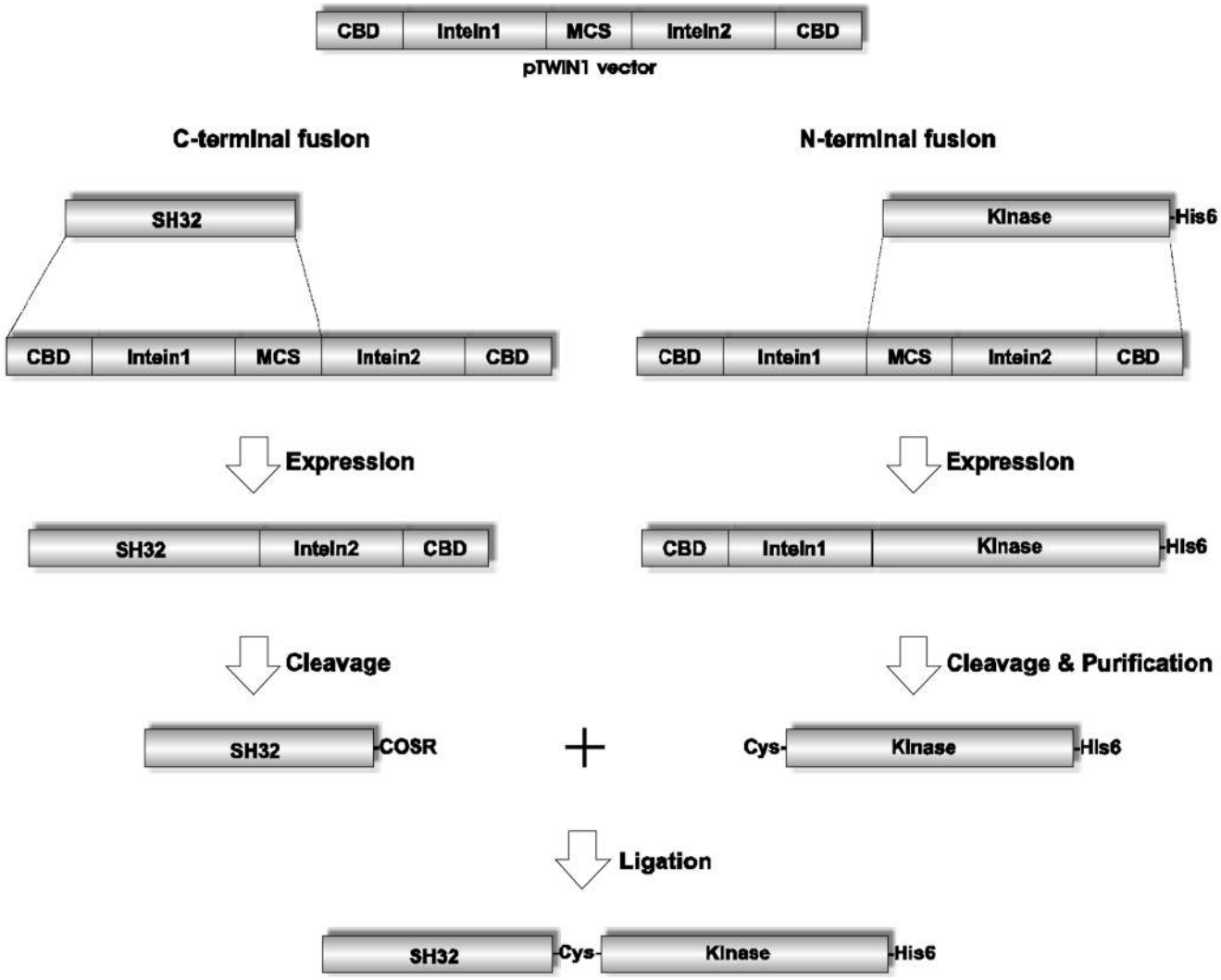

Figure 3.

Segmental labeling of human Csk (50 kDa) protein. Genes of SH32 or kinase domain are cloned into pTWIN1 vector as C-terminal fusion and N-terminal fusion, respectively. Intein2-CBD was fused to the C-terminal of SH32, allows the isolation of SH32 using the thio-inducible cleavage activity of intein2. A reactive C-terminal thioester can be generated on SH32 when MESNA is used as the thiol reagent to induce cleavage of the intein. Kinase domain was fused with CBD-intein1 to permit the preparation of kinase with an N-terminal Cysteine. His-tag was added to facilitate the purification because of the *in vivo* cleavage. The final ligation of SH32 with α-thioester and kinase with N-terminal Cysteine form the segmental labeled full length Csk.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

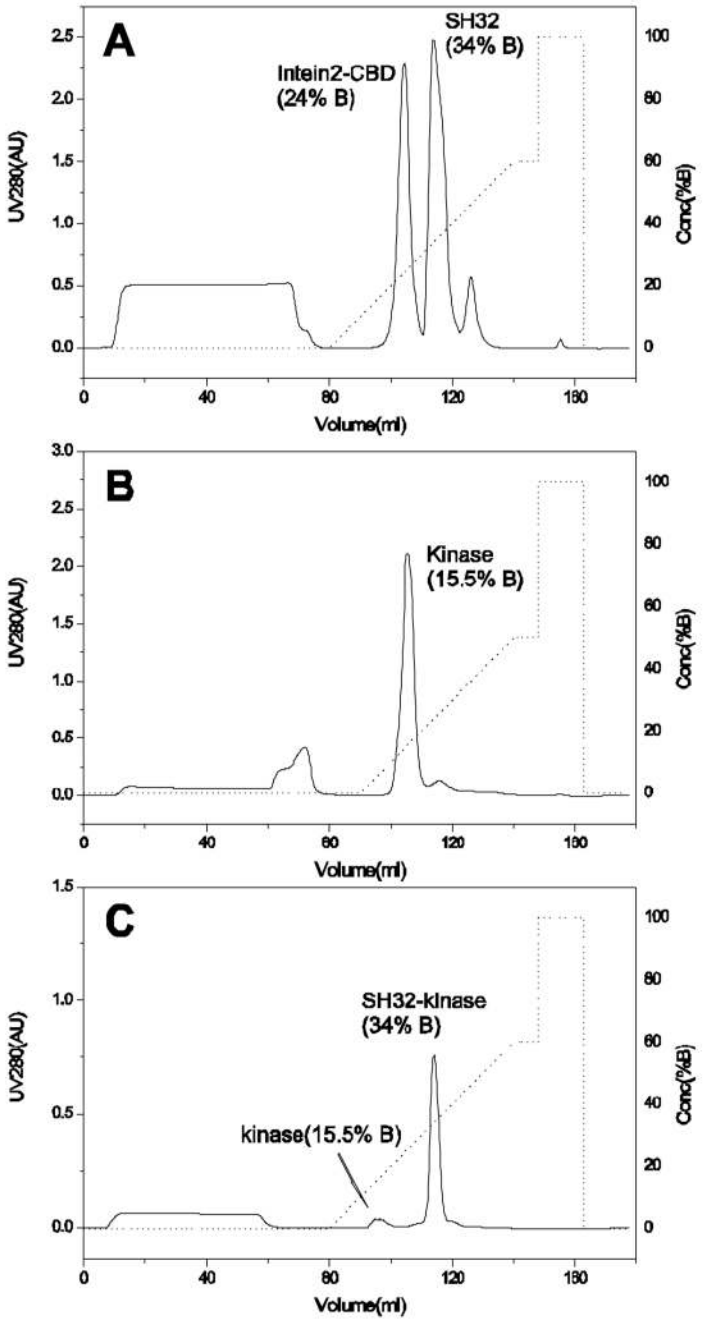

Figure 4.

Momo Q elusion profile of selected product. Loading buffer: 20 mM Tris-HCl, pH 8.0 (Buffer B) Elusion buffer: 20 mM Tris-HCl, pH 8.0, 1.0 M NaCl (Buffer C), flow rate 1 ml/min. The elution was conducted by linear gradient to 50–60% Buffer C. (A) SH32 domain. (B) Kinase domain. (C) Final purification step of SH32-kinase ligation product.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

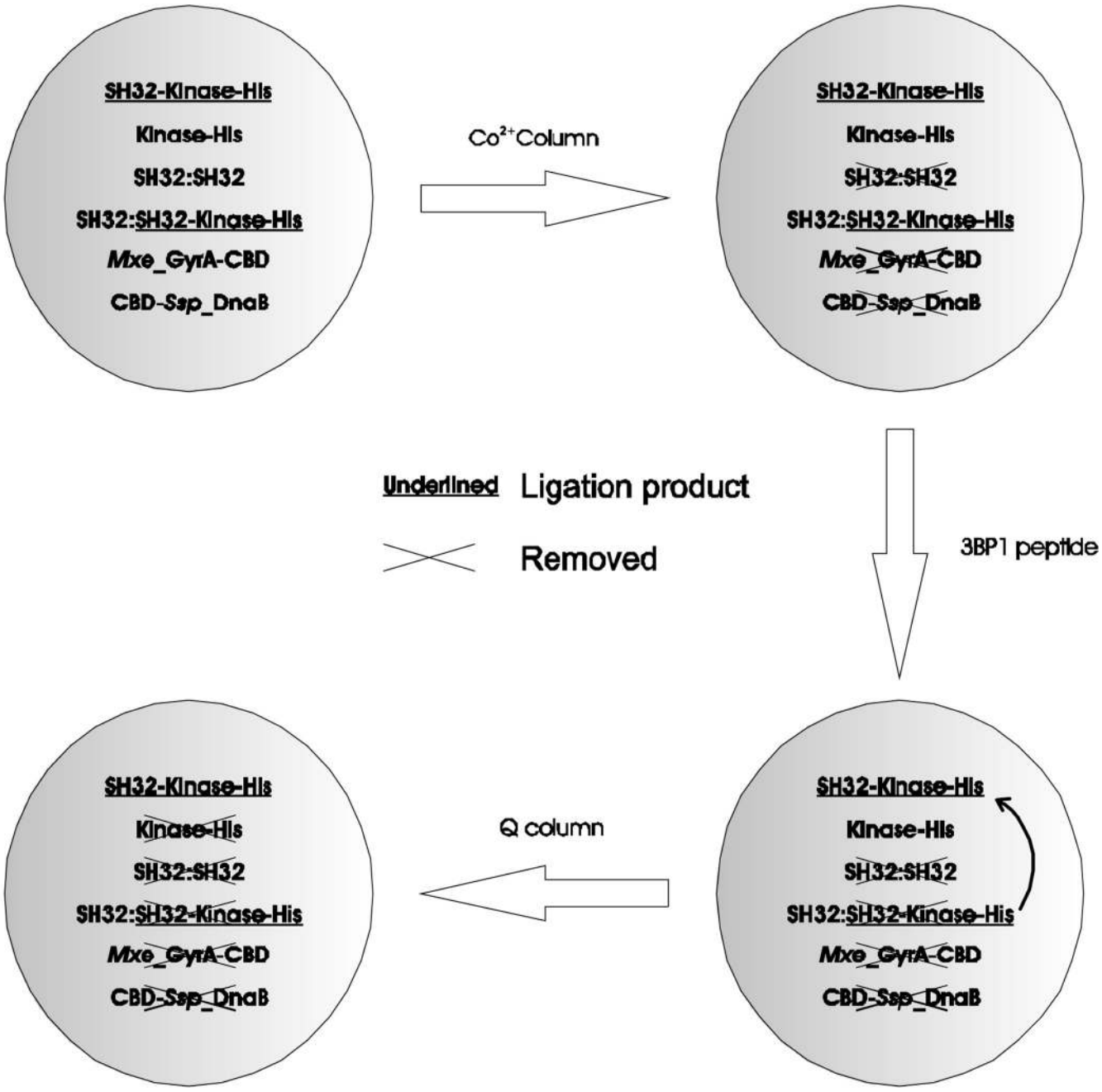

Figure 5.

Purification procedure used in the isolation of the ligated product. The molecule in the first circle is the total products after the ligation experiment. After a Co^{2+} iron column, the protein with a C-terminal His-tag are bound to the column. The proteins without His-tag are removed. One problem come from the dimerization of the SH32, although SH32 can not bind to the Co^{2+} column, it can bind to the ligated sample which is bound to the column. PEP-3BP1 peptide was used to disrupt the dimerization via SH3 domain. In the final purification step, since the SH2 domain is highly negative charged, it can bind to the Q column at high salt concentrations (See Figure 4).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

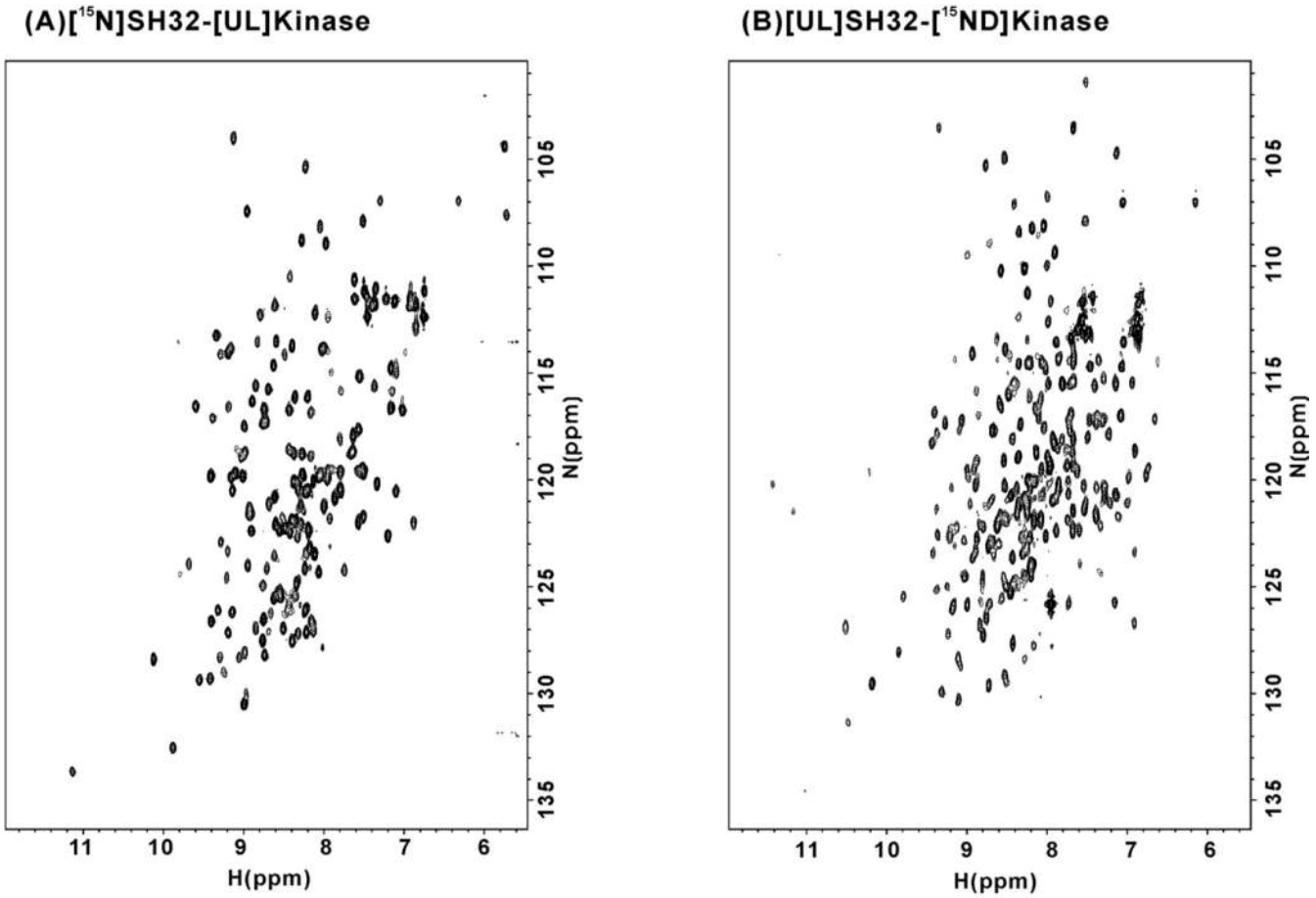

Figure 6.

^{1}H-^{15}N HSQC spectra at 800 MHz of (A): [^{15}N-SH32]-Kinase; (B): SH32-[^{15}N, ^{2}H-kinase]. All of the samples are concentrated to 100 μM. The buffer contains 200 μM non-labeled PEP-3BP1 peptide, 50 mM Tris-HCl (pH 7.5), 10 mM DTT, 1.0 mM EDTA, 0.01% (w/v) NaN_3, 5% D_2O, 0.1 mM DSS. NMR spectra were taken at 25°C.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Selected examples of segmental isotopic labeling using expressed protein ligation methods

Name/Total Size	Linker	Method for α-thioester	Method for N-Cys	Comments
Abl SH3-SH2 domain/18 kDa	G-C	*Sce* VMA from pTYB2	Factor Xa cleavage	First segmental labeled sample using expressed protein ligation methods. Thiophenol and benzyl mercaptan were used as catalysts. ^{1}H-^{15}N HSQC acquired (Xu et al., 1999)
σ factor σ^{A} and Δ1.1-σ^{A}/47 and 32 kDa	G-CG	*Mxe* GyrA from pTXB1	Factor Xa cleavage	Cleavage and ligation on the chitin beads. Ligation reaction was rapid and efficiency was high (Camarero et al., 2002)
Mxe GyrA inteine/22 kDa	F-C	Peptide synthesis	Factor Xa cleavage	Ligation buffer: (6 M Gdm Cl/0.1 M NaPi at pH 8) containing 3% MESNA and 2% ethanethiol. Room temperature 5 h. ^{1}H-^{15}N HSQC, 2D H(N)CO used (Romanelli et al., 2004)
G protein α subunit/40 kDa	K-C	*Mxe* GyrA from pTXB3	Peptide synthesis	On column cleavage and ligation with 1% MESNA. 4°C, 24 h (Grant et al., 2006)
PTB RRM3-RRM4/24 kDa	G-C	*Mxe* GyrA from pTWIN1	*Ssp* DnaB from pTWIN1	Ligation on column at 37°C for 16 h. 130 interdomain NOEs were unambiguously identified (Vitali et al., 2006)
Npl3p RM1-RRM2/21 kDa	Y-C Or G-C	*Mxe* GyrA from pTWIN1	*Ssp* DnaB from pTWIN1	Ligation conducted at native or denatured conditions. ^{1}H-^{15}N HSQC suggested no domain interaction (Skrisovska and Allain, 2008)
hnRNP L RRM3-RRM4/25 kDa	F-C	*Mxe* GyrA from pTWIN1	*Ssp* DnaB from pTWIN1	Refolding required before ligation. On column ligation. High ligation efficiency. Interdomain NOE obtained (Skrisovska and Allain, 2008).
DHFR/19 kDa	L-C	Peptide synthesis	Met aminopeptidase	A single ^{13}C residue in an active site loop was generated. Ligation conducted at denatured state (Ottesen et al., 2008)
c-Crk-II/34 kDa	G-C G-C	*Mxe* GyrA from pTXB1 *Sce* VMA from pTYB2	Factor Xa Cleavage	Sequential ligation of three recombinant polypeptides. The incorporation of isotopes in the central domain of the protein(Blaschke et al., 2000; Ottesen et al., 2008).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Selected examples of segmental isotopic labeling using protein *trans*-splicing methods

Name/Total Size	Linker	Intein	Comments
αC/10 kDa	GGG-TG	PI-*Pfu*I	First protein to be segmental isotope labeled. Denatured precursors refolding. *Trans*-splicing at 70°C (Yamazaki et al., 1998) . Ligation yield was improved by some modification of the protocol (Otomo et al., 1999b). ^{1}H-^{15}N HSQC, ^{15}N edited NOESY used.
MBP/42 kDa	GGG-TG	PI-*Pfu*I	Denatured precursors refolding. *Trans*-splicing at 70°C. Ligation protocol optimized. 2D ^{1}H-^{15}N HSQC used (Otomo et al., 1999b)
MBP/42 kDa	GGG-TG TNP-CGE	PI-*Pfu*I PI-*Pfu*II	Central-segment isotope labeling through protein *trans*-splicing using two separate inteins (Otomo et al., 1999a)
H^{+}-ATPase β subunit/52 kDa	GGG-TG	PI-*Pfu*I	Denatured precursors refolding. *Trans*-splicing at 70°C, 1 hr. Triple resonance 3D, RDC experiments used (Yagi et al., 2004)
GB1-CBD/15 kDa	GS-CFNKGT	*Ssp* DnaE	Ligation *in vivo*. Based on a dual expression system that allows sequential expression of two precursor fragments in different media. ^{1}H-^{15}N HSQC, HNCACB, 2D-[^{1}H-^{15}N]-HNCO used (Zuger and Iwai, 2005)