Biomol NMR Assign (2009) 3:41–43
DOI 10.1007/s12104-008-9137-1

ARTICLE

# NMR resonance assignment of DnaE intein from *Nostoc punctiforme*

**Kimmo Heinämäki · Jesper S. Oeemig · Janica Djupsjöbacka · Hideo Iwaï**

Received: 29 October 2008 / Accepted: 19 December 2008 / Published online: 13 January 2009


**Abstract** DnaE intein from *Nostoc punctiforme* (*Npu*) is one of naturally occurring split inteins, which has robust protein splicing activity. Highly efficient *trans*-splicing activity of *Npu*DnaE intein could widen various biotechnological applications. However, structural basis of the efficient protein splicing activity is poorly understood. As a first step toward better understanding of protein *trans*-splicing mechanism, we present the backbone and side-chain resonance assignments of a single chain variant *Npu*DnaE intein as determined by triple resonance experiments with [$^{13}$C,$^{15}$N]-labeled protein.

**Keywords** Protein splicing · Intein · Protein *trans*-splicing

## Biological context

In protein splicing, an intervening sequence (intein) in a host protein catalyzes ligation of the disrupted host protein sequences and concomitant self-excision from the host protein as a posttranslational modification (Paulus 2000). Protein splicing could also take place in *trans*, when the precursor protein containing an intein is split into two fragments within the intein sequence. Naturally split DnaE inteins have also been discovered that catalyze protein splicing in *trans* (Wu et al. 1998). Protein ligation using protein *trans*-splicing has become an important tool for protein engineering in vitro as well as in vivo, as it has opened new possibilities to manipulate proteins with a covalent peptide bond as a posttranslational modification (Xu and Evans 2005). It has been used for biotechnological applications such as protein cyclization, site-specific modifications, and segmental isotopic labeling (Iwai et al. 2001; Scott et al. 1999; Xu and Evans 2005; Muona et al. 2008; Züger and Iwai 2005). Although the chemical reactions involved in protein splicing are generally accepted (Paulus 2000), structural features influencing the protein splicing reaction remain elusive. Particularly, for biotechnological applications such as segmental isotopic labeling, it is of importance to elucidate the structural basis of how the junction sequences modulate the protein splicing activity (Iwai et al. 2006). As a first step toward understanding the chemical and structural basis of protein splicing process at atomic resolution, here we report the NMR assignments of a single chain variant of naturally split DnaE intein from *Nostoc punctiforme* (*Npu*), bearing robust protein splicing activity.

## Materials and methods

### Protein design

We have constructed a single chain variant (137 residues) of naturally split *Npu*DnaE intein by fusing the N- and C-terminal split intein fragments genetically. In addition, we introduced a mutation of C1A and removed the C-terminal extein sequence to avoid the protein splicing reaction occurring during NMR measurements. The single chain *Npu*DnaE intein was created by assembly PCR of the amplified PCR products of the N- and C-terminal

J. S. Oeemig and K. Heinämäki contributed equally to this work.

K. Heinämäki · J. S. Oeemig · J. Djupsjöbacka · H. Iwaï (✉)
Research Program in Structural Biology and Biophysics, Institute of Biotechnology, University of Helsinki, P.O. Box 65, Helsinki 00014, Finland
e-mail: hideo.iwai@helsinki.fi

fragments of *Npu*DnaE intein from the genomic DNA (Iwai et al. 2006). The PCR product of the single chain *Npu*DnaE intein was ligated into pHYRSF53 resulting in pJDJRSF05, which contains an N-terminal hexahistidine tag and a yeast Smt3 protein as a fusion tag for protein purification. Two glycine residues remained after the removal of His-tagged Smt3 domain. The final protein contains the following primary structure:

```
 0 1         11         21         31         41
GG ALSYETEILT VEYGLLPIGK IVEKRIECTV YSVDNNGNIY TQPVAQWHDR
   GEQEVFEYCL EDGSLIRATK DHKFMTVDGQ MLPIDEIFER ELDLMRVDNL
   PNIKIATRKY LGKQNVYDIG VERDHNFALK NGFIASN
```

## Expression and purification

The protein was expressed in *E. coli* ER2566 with M9 medium supplemented with $^{15}NH_4Cl$ and either 100% $^{13}C_6$-glucose or 20% $^{13}C_6$-glucose as sole nitrogen and carbon sources (Iwai and Fiaux 2007). The cells were disrupted by ultrasonification and purified by HisTrap column (GE Healthcare). The fusion protein of the N-terminally his-tagged Smt3 protein was digested by yeast Ulp1 protease. The Smt3 domain, Ulp1 and undigested fusion proteins were removed by HisTrap column. The flow-through fractions were collected and dialyzed against 10 mM sodium phosphate, pH 8.0 and further purified with MonoQ 5/50 GL ion-exchange column (GE Healthcare). The protein was dialyzed against 10 mM sodium phosphate, pH 6.0 and concentrated by Amicon Ultra 4 centrifugal filter unit (Millipore). The sample was concentrated to 250 µl and transferred into a Shigemi micro NMR tube.

## NMR measurements

NMR measurements were performed at the $^1$H frequency of 600 MHz on Varian Innova spectrometer equipped with a triple resonance cryogenic probe head or on Varian Innova 800 MHz spectrometer equipped with triple resonance probe head. The experiments were performed at 298 K. For the sequence specific assignments, the following experiments were used. [$^{15}$N,$^1$H]-HSQC, [$^{13}$C,$^1$H]-HSQC, HNCA, HNCACB, HNCO, CBCA(CO)NH, HN(CA)CO, HN(CO)CA and CC(CO)NH (Sattler et al. 1999). The side-chain $^1$H and $^{13}$C resonance assignments were based on HBHA(CO)NH, HNHB, CC(CO)NH, HCC(CO)NH, H(C)CH-TOCSY, HCCH-COSY, $^{15}$N-resolved [$^1$H,$^1$H]-TOCSY and $^{13}$C-resolved [$^1$H,$^1$H]-TOCSY experiments. The assignments of aromatic side-chains were based on the spectra of CBCGCDHD, CBCGCDCEHE, and ct-[$^{13}$C,$^1$H]-HSQC.

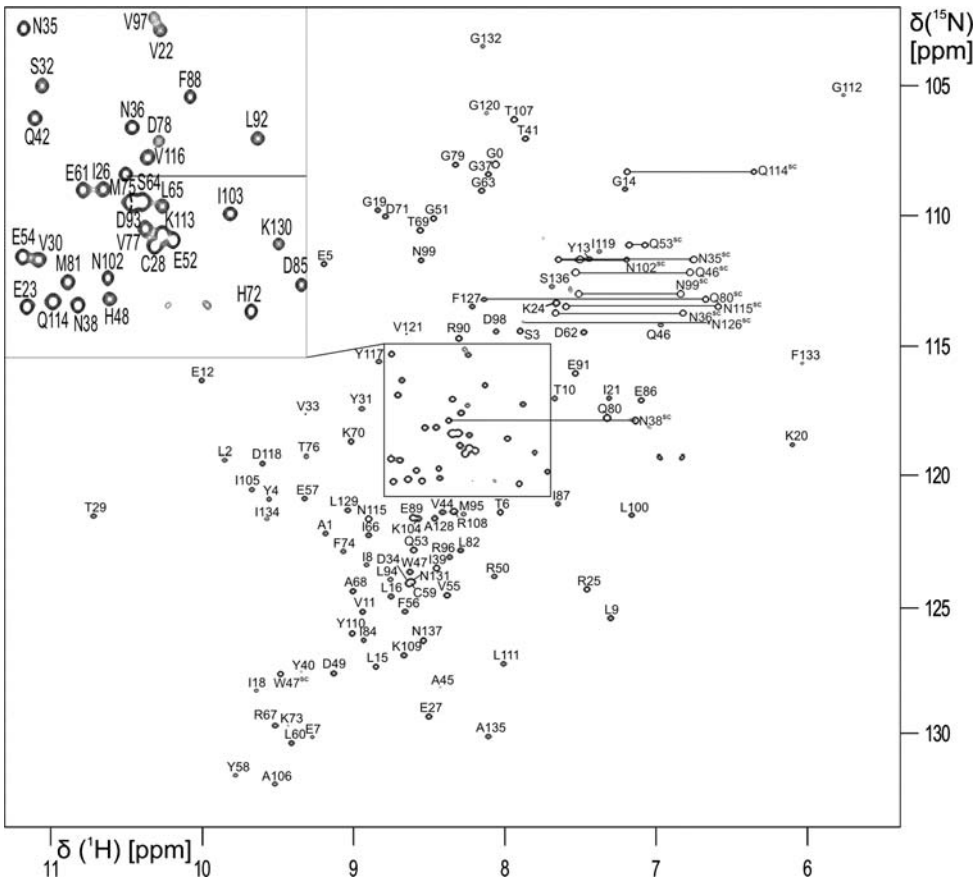


**Fig. 1** Two-dimensional [$^{15}$N,$^1$H]-HSQC spectrum of 2 mM single chain *Npu*DnaE intein in 10 mM sodium phosphate buffer at pH 6.0 at 298 K. The spectrum was recorded at 600 MHz $^1$H frequency. A total of 100 ($\omega_1$) × 683 ($\omega_2$) complex points were collected for this spectrum, with $t_{1,max} = 18.2$ ms, $t_{2,max} = 32.0$ ms. The assignments are indicated by the number and one character codes for 20 amino acids. The side-chain resonances are marked by "sc". The side-chain amide groups of asparagines and glutamines are connected by horizontal lines

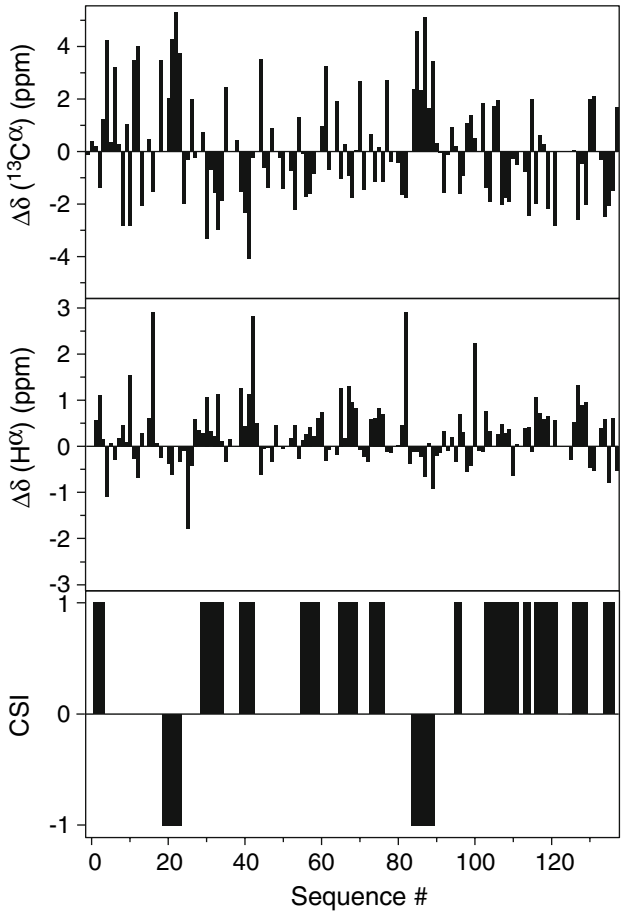


**Fig. 2** Plots of the secondary chemical shifts for $H^{\alpha}$ and $^{13}C^{\alpha}$ and the chemical shift index (CSI). The values for glycine residues are excluded for $H^{\alpha}$. CSI was calculated by the program CSI with $H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and C′ chemical shifts (Wishart and Sykes 1994)

## Assignment and data deposition

In Fig. 1, the backbone resonance assignments of the single chain *Npu*DnaE intein are illustrated indicating high quality of the NMR data. The resonance assignments were obtained except for the first residue (numbered as −1) and residues 122–124. 96.4% of the backbone $H^N$, $H^{\alpha}$, $^{13}C$, $^{13}C^{\alpha}$, $^{15}N$ atoms, and 96.2% of the side-chain atoms has been assigned. This includes stereospecific assignments of diastereoscopic methyl groups in leucines and valines based on a fractional $^{13}C$-labeled sample. When possible, $H^{\beta}$-atoms were stereospecifically assigned with $^{15}N$–$^{1}H^{\beta}$ coupling constant and $H^N$–$H^{\beta}$, $H^{\alpha}$–$H^{\beta}$ NOEs. The chemical shift deviations of $H^{\alpha}$, $^{13}C^{\alpha}$ from the random coil values and the chemical shift index predicting the secondary structures are plotted in Fig. 2. The chemical shifts of the single chain *Npu*DnaE intein have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 16009.

**Acknowledgments** The authors thank Olga Valvuk and Kimmo Pääkkönen for their contributions at the early stage of the project. This work is supported by the grants from the Academy of Finland (118385), Sigrid Jusélius Foundation, and the Biocentrum Helsinki.

## References

Iwai H, Fiaux J (2007) Use of biosynthetic fractional $^{13}C$-labeling for backbone NMR assignment of proteins. J Biomol NMR 37:187–193. doi:10.1007/s10858-006-9124-8

Iwai H, Lingel A, Plückthun A (2001) Cyclic green fluorescent protein produced in vivo using an artificially split PI-PfuI intein from *Pyrococcus furiosus*. J Biol Chem 276:16548–16554. doi:10.1074/jbc.M011639200

Iwai H, Züger S, Jin J, Tam PH (2006) Highly efficient protein *trans*-splicing by a naturally split DnaE intein from *Nostoc punctiforme*. FEBS Lett 580:1853–1858. doi:10.1016/j.febslet.2006.02.045

Muona M, Aranko AS, Iwai H (2008) Segmental isotopic labelling of a multi-domain protein by protein ligation using protein *trans*-splicing. ChemBioChem 9:2958–2961. doi:10.1002/cbic.200800604

Paulus H (2000) Protein splicing and related forms of protein autoprocessing. Annu Rev Biochem 69:447–496. doi:10.1146/annurev.biochem.69.1.447

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog NMR Spectroc 34:93–158. doi:10.1016/S0079-6565(98)00025-9

Scott CP, Abel-Santos E, Wall M, Wahnon DC, Benkovic SJ (1999) Production of cyclic peptides and proteins in vivo. Proc Natl Acad Sci USA 24:13638–13643. doi:10.1073/pnas.96.24.13638

Wishart DS, Sykes BD (1994) The $^{13}C$ chemical shift index. A simple method for the identification of protein secondary structure using $^{13}C$ chemical shift data. J Biomol NMR 4:171–180. doi:10.1007/BF00175245

Wu H, Hu Z, Liu XQ (1998) Protein *trans*-splicing by a split intein encoded in a split DnaE gene of *Synechocystis sp.* PCC6803. Proc Natl Acad Sci USA 95:9226–9231. doi:10.1073/pnas.95.16.9226

Xu MQ, Evans TC Jr (2005) Recent advances in protein splicing: manipulating proteins in vitro and in vivo. Curr Opin Biotechnol 16:440–446. doi:10.1016/j.copbio.2005.06.012

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. doi:10.1038/nbt1097