



Published in final edited form as:

Biomol NMR Assign. 2008 December ; 2(2): 111–113. doi:10.1007/s12104-008-9098-4.

¹H, ¹³C, and ¹⁵N NMR assignments of an engineered intein based on *Mycobacterium tuberculosis* RecA

Zhenming Du[§], Yangzhong Liu[†], Scott McCallum[§], John Dansereau[‡], Victoria Derbyshire[‡], Marlene Belfort[‡], Georges Belfort[§], Patrick Van Roey[‡], and Chunyu Wang^{§,*}

[§] Biology Department, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180

[†] Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, P.R.China 230026

[‡] Wadsworth Center, New York State Department of Health, 150 New Scotland Avenue, Albany, NY 12208

Keywords

intein; assignments; NMR; RecA; *Mycobacterium tuberculosis*

Biological context

Protein splicing is a precise post-translational process in which an in-frame protein fusion, called an intein, is removed from the precursor sequence with the concomitant ligation of the two flanking fragments, the N- and C-exteins (1). The majority of the inteins identified to date consist of two structurally and functionally independent domains: the homing endonuclease domain and the splicing domain (2). Although the basic steps of intein splicing has been well documented (for a review, see (3)), the atomic details of the splicing mechanism is still lacking. Advancing the understanding of the mechanism of intein splicing can have two major impacts. First, inteins have found extensive applications in protein engineering and biotechnology and therefore are an indispensable tool for biomedical research. Second, inteins are potential targets for antibiotic development because only unicellular organisms such as *Mycobacterium tuberculosis* have inteins that are vital for their survival.

Many structural studies of inteins have been carried out (4–7). All intein splicing domains have the Hint (hedgehog intein) fold, shared by the Hedgehog C-terminal autoprocessing domain (4–7). Romanelli et al. (7) have used NMR and segmental isotopic labeling to obtain an unusual J-coupling for the scissile peptide bond at the N-terminal splicing junction, supporting the “ground-state destabilization” model as the mechanism for the initiation of splicing. Additionally, Margaret et al. (5) has solved the solution NMR structure of a splicing-deficient precursor of KlbA intein from *Methanococcus jannaschii*. In contrast, Van Roey et al. (6) has obtained several crystal structures of engineered *Mtu* RecA inteins and showed that an F-block aspartate D422 is flexible and interacts with both both N- and C- terminal splicing junction, suggesting a likely regulatory role for D422.

Despite the abundance of structural information, the precise roles of conserved residues in the splicing mechanism remain poorly defined. In order to carry out mechanistic studies at atomic

*To whom correspondence should be addressed: wangc5@rpi.edu.

resolution using solution NMR, we have determined the backbone and sidechain resonance assignments of a RecA mini-intein $\Delta\Delta I_{hh}$ -CM. The $\Delta\Delta I_{hh}$ -CM intein is a 139-residue intein engineered from the 440-amino acid *Mycobacterium tuberculosis* RecA intein, through several minimization steps. The central endonuclease domain of the 440-residue *Mycobacterium tuberculosis* RecA intein was removed to generate a mini-intein, ΔI (8), which was further minimized to 139 amino acid residues ($\Delta\Delta I_{hh}$) by replacing a loop with the corresponding loop of Hedgehog (hh) (9). The cleavage mutant (CM), $\Delta\Delta I_{hh}$ -CM, with the D422G mutation favors C-terminal cleavage over splicing (9).

Methods and Experiments

Uniformly isotopically labeled (U- ^{15}N) or U- ^{13}C , ^{15}N) $\Delta\Delta I_{hh}$ -CM was expressed in JM101 cells in M9 medium and purified by affinity chromatography using chitin-beads (New England Biolabs, Ipswich, MA, USA) followed by DTT-catalyzed release of intein from the CBD fusion protein. NMR samples (0.6 mM $\Delta\Delta I_{hh}$ -CM) were prepared in 20 mM sodium phosphate, 100 mM sodium chloride, 1 mM sodium azide in 90% H_2O /10% D_2O or 99.9% D_2O at pH 7.0.

All NMR experiments were carried out at 25°C on a Bruker 800MHz or 600MHz (^1H) spectrometer equipped with a cryogenic probe. Spectra were processed with nmrPipe software and analyzed using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, USA). The ^1H chemical shifts were referenced relative to DDS and the ^{15}N and ^{13}C chemical shifts were referenced indirectly.

The sequence specific backbone $^1\text{H}^{\text{N}}$, $^{13}\text{C}^{\alpha}$, ^{15}N and side chain $^{13}\text{C}^{\beta}$ assignments for $\Delta\Delta I_{hh}$ -CM were obtained by standard triple resonance NMR methods, such as HNCOC, HNCACOC, HNCA, HNCOCAC, HNCACB, HNCOCACB, ^{15}N NOESY and ^{15}N TOCSY experiments. Assignments for aliphatic side chains were achieved by standard NMR techniques, including HC(CC)(CO)NH, CC(CO)NH, ^{13}C NOESY, HCCH-TOCSY, ^{13}C COSY, ^{15}N NOESY and ^{15}N TOCSY. The (H β)C β (C γ C δ)H δ and aromatic- ^{13}C NOESY experiments were employed for the assignments of aromatic resonances.

Extents of assignments and data deposition

$\Delta\Delta I_{hh}$ -CM contains 139 amino acids. Nearly complete assignments of the backbone (>98%) $^1\text{H}^{\text{N}}$, ^{15}N , $^{13}\text{C}^{\alpha}$, $^1\text{H}^{\alpha}$, and $^{13}\text{C}'$ resonances were obtained with the exception of C1, G63, and H128, $^{13}\text{C}'$ of D10, K27, R43, T70, and L112 which are before proline residues, and the ^{15}N resonances of the 5 proline residues, P11, P28, P44, P71, and P113. All the $^{13}\text{C}^{\beta}$ and $^1\text{H}^{\beta}$ resonances were assigned except for those of K27. Aliphatic side chain ^1H and ^{13}C assignments are nearly complete (97%). Unassigned resonances include the sidechains of K27, R115 and some of the ^1H and ^{13}C resonances for the aromatic rings. The chemical shifts have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under accession number 15560.

Acknowledgments

The work was supported by NIH grants GM081408 (C.W.) and GM44844 (M.B.).

References

1. Perler FB, Davis EO, Dean GE, Gimble FS, Jack WE, Neff N, Noren CJ, Thorner J, Belfort M. Nucleic acids research 1994;22:1125–1127. [PubMed: 8165123]
2. Perler FB. Nucleic Acids Research 2002;30:383–384. [PubMed: 11752343]
3. Paulus H. Annual Review of Biochemistry 2000;69:447–496.

4. Moure CM, Quioco FA. *Nucleic Acids and Molecular Biology* 2005;16:257–271.
5. Johnson MA, Southworth MW, Herrmann T, Brace L, Perler FB, Wuthrich K. *Protein Science* 2007;16:1316–1328. [PubMed: 17586768]
6. Van Roey P, Pereira B, Li Z, Hiraga K, Belfort M, Derbyshire V. *Journal of Molecular Biology* 2007;367:162–173. [PubMed: 17254599]
7. Romanelli A, Shekhtman A, Cowburn D, Muir TW. *PNAS* 2004;101:6397–6402. [PubMed: 15087498]
8. Derbyshire V, Wood DW, Wu W, Dansereau JT, Dalgaard JZ, Belfort M. *Proceedings of the National Academy of Sciences of the United States of America* 1997;94:11466–11471. [PubMed: 9326633]
9. Hiraga K, Derbyshire V, Dansereau JT, Van Roey P, Belfort M. *Journal of Molecular Biology* 2005;354:916–926. [PubMed: 16288917]
10. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. *Journal of Biomolecular NMR* 1995;6:277–293. [PubMed: 8520220]

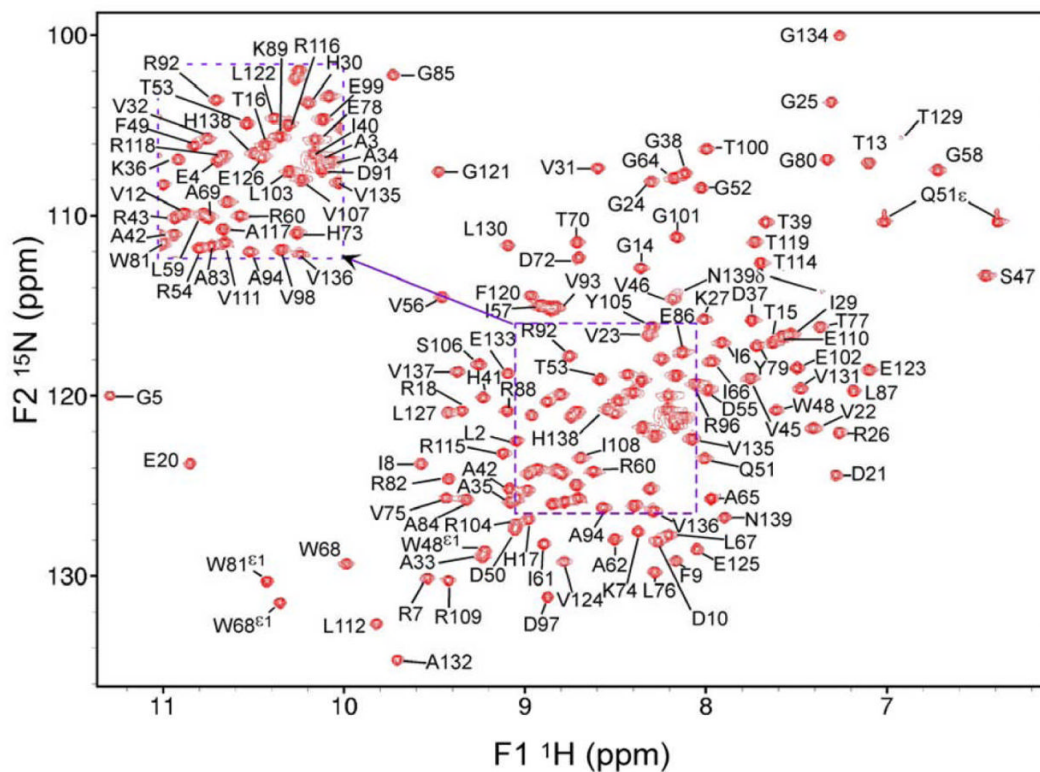


Figure 1. ^1H - ^{15}N HSQC spectrum of $\Delta\Delta\text{I}_{\text{hh}}$ -CM with assignments labeled. The inset at upper left is for labeling peaks in the crowded central regions of the HSQC. The two side chain NH_2 (Q51^ϵ and N139^δ) groups have been assigned. Spectrum is recorded at 600MHz and 25°C.