Mutational Analysis of Splicing Activities of Ribonucleotide Reductase α Subunit Protein from Lytic Bacteriophage P1201

Shu-Chen Kan · Liang-Kun Yu · Jiau-Hua Chen · Hui-Yu Hu · Wen-Hwei Hsu

Abstract A CP1201 RIR1 intein is found in the ribonucleotide reductase alpha subunit (RNR α subunit) protein of lytic bacteriophage P1201 from *Corynebacterium glutamicum* NCHU 87078. This intein can be over-expressed and spliced in *Escherichia coli* NovaBlue cells. Mutations of C539, the N-terminal residue of the C-extein in the CP1201 RIR1 protein, led to the changes of pattern and level of protein-splicing activities. A G392S variant was found to be a temperature-sensitive protein with complete splicing activity at 17 and 28°C but not at 37°C or higher. We also found that the cleavage at the CP1201 RIR1 intein C-terminus of the double mutant G392S/C539G was blocked, but other cleavage activities could be efficiently performed at 17°C. G392S/C539G variant possessed the properties of low-temperature-induced cleavage at the intein N-terminus.

S.-C. Kan · W.-H. Hsu (✉)
Institute of Molecular Biology, National Chung Hsing
University, Taichung 402, Taiwan
e-mail: whhsu@dragon.nchu.edu.tw

L.-K. Yu
Department of Life Science, National Chung Hsing University,
Taichung 402, Taiwan

J.-H. Chen
Department of Food Science and Technology, Chia Nan
University of Pharmacy and Science, Tainan 717, Taiwan

H.-Y. Hu (✉)
Department of Food Science and Technology,
Hung Kuang University, Taichung 433, Taiwan
e-mail: huiyu@sunrise.hk.edu.tw

Introduction

Protein splicing is a post-translational process, in which an intervening segment, called intein, catalyzes its own excision from a protein precursor, and ligation of the flanking regions, called extein, by peptide bond formation [1]. Little overall sequence similarity is found in inteins; however, comparative analysis of intein sequences reveals the motifs used to identify and characterize inteins from different organisms [2, 3]. Over 375 inteins have been found in eukaryotes, eubacteria, and archaebacteria [4], suggesting a wide distribution.

The chemistry of protein splicing has been discussed in some cases tested experimentally [5–9]. Protein splicing is initiated by an N–S or N–O acyl rearrangement of the intein N-terminal Cys, Ser, or Ala to form a linear thioester or ester with the C-terminal residue of the N-extein. A transesterification reaction shifts the N-extein onto the side chain of the Cys/Ser/Thr at the first residue of the C-extein. The resultant branched (thio) ester intermediate is resolved by the cyclization of the C-terminal Asn of the intein, leading to excision of intein and ligation of the exteins by peptide bond. Intein-contained gene has been modified to change the cleavage activities at the N- or C-terminal splice junction, and used to develop intein-based protein purification systems [10], with target protein as the N-extein or the C-extein and an affinity domain as the C-extein or the N-extein.

We have experimentally identified the RIR1 intein of the bacteriophage P1201 from *Corynebacterium glutamicum* NCHU 87078 [11]. This intein that interrupts the RNR α subunit (RIR1) can be over-expressed in *Escherichia coli* [11]. In this report, site-directed mutagenesis of the N-terminal residue (C539) of the C-extein in the CP1201 RIR1 protein was performed, and the extent and pattern of protein-splicing activities were determined. We

demonstrated that the cleavage at the intein C-terminus is blocked in the double mutant, while other cleavage activities can be performed at lower temperature.

Materials and Methods

Bacteria, Phage, Plasmids, and Media

Corynebacterium glutamicum NCHU 87078, a glutamic acid hyper producing strain, was grown in CM medium (1% peptone, 1% yeast extract, 1% glucose, and 0.25% NaCl; pH 7.5) and TYG medium (1% tryptic peptone, 0.6% yeast extract, 0.5% glucose, and 0.1 M MgSO$_4$·7H$_2$O; pH 7.2) at 30°C. *Escherichia coli* NovaBlue was used as host for recombinant plasmids. Plasmid pQE30 (Qiagen, Valencia, CA) was used for gene expression. Cells were cultivated in Luria–Bertani medium (LB; 1% trypton, 1% NaCl, 0.5% yeast extract; pH 7.5). Ampicillin (100 μg/ml) was added when necessary. Isolation of phage P1201 from *C. glutamicum* NCHU 87078 and purification of phage DNA were performed essentially using the methods described previously [11].

Construction of RIR1 Variants

The open reading frame (ORF) of putative RNR α subunit gene (*rir*1) containing the CP1201 RIR1 intein sequence was amplified from phage P1201 DNA using gene-specific primers ORF56-F (5′-CGGGATCCATGACTAAAGATAC CTAC–3′) and ORF56-R (5′-AACTGCAGCTACAAAGC GCAGCTTAG–3′). The amplified PCR products were cloned into *Bam*HI and *Pst*I sites of pQE30 to generate plasmid pRNR.

The ORF of *rir*1 gene was subjected to site-directed mutagenesis. Site-directed mutagenesis was performed essentially according to the method of overlap extension PCR [12, 13] using plasmid pRNR as the template. The following primers contain the desired mutations: RIR 1C539N-F (5′-AAC GGT TTG GTA ACG GGA NNN TGC TCG GAG-3′) and RIR1C539N-R (5′-CTG GAA AAT CTC CGA GCA NNN TCC CGT TAC-3′), that encoded each of the 20 amino acids (NNN) at position 539. The Gly392Ser substitution was amplified using primers: RIR1G392S-F (5′-CCT GAC CTC CAG GAG CTA CCG TCG TGC-3′) and G392S-R (5′-GCA CGA CGG TAG CTC CTG GAG GTC AGG-3′). All mutations were confirmed by DNA sequencing of the whole ligated PCR fragment.

Expression of the CP1201RIR1 Variants

The *rir*1 gene on plasmid pRNR was expressed in *E. coli* NovaBlue. Recombinant *E. coli* cells were grown at 37°C with shaking to OD$_{600}$ of approximately 0.7, induced with

1 mM isopropyl-1-thio-β-ᴅ-galactopyranoside (IPTG), and incubated at 17, 28, and 37°C for an additional 0.5–24 h with shaking. Cells were harvested by centrifugation, resuspended in Tris–HCl buffer (pH 8.0) containing 100 mM NaCl, and disrupted by sonication. The crude cell extract was mixed with 2× sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer without ᴅʟ-1,4-dithiothreitol, boiled for 10 min, and then subjected to SDS-PAGE analysis. Gels were stained with Coomassie Brilliant Dye (Sigma).

Results and Discussion

Influence of the Identity of the N-terminal Residue of the C-extein on the Cleavage Reaction.

Amino acid sequence analysis of RIR1 protein showed that the CP1201 RIR1 intein, positioned between residue 380 and 538 of the protein, had conserved motifs and conserved Cys at the N-terminal splice site and Asn at the downstream splice junction (Fig. 1a). The protein-splicing activity of the CP1201 RIR1 intein had been proved by cloning of the ORF of *rir*1 gene into the *Bam*HI and *Xho*I sites of pET28b and expressing in *E. coli* BL21 (DE3) [11]. SDS-PAGE analysis revealed that apparent molecular masses of extein and intein are approximately 79 and 18 kDa, respectively (Fig. 1b, lane C).

To determine the effect of the difference in the N-terminal residue of the C-extein on the cleavage reaction, all the 19 protein variants with the mutation at Cys539 of RIR1 were constructed. The mutant *rir*1 genes were cloned into pQE30 and expressed in *E. coli* NovaBlue. The crude cell extract of IPTG-induced *E. coli* cells in the cleavage mixture were incubated at 28°C for 6 h and then analyzed by SDS-PAGE. Since the cloned gene is transcribed by a powerful T5 promoter, the majority of total protein in the crude cell extract is from the cloned gene. Protein bands corresponding to the sizes of precursor fusion protein (P, 97 kDa), extein (E, 78 kDa), intein (I, 19 kDa), N-terminal extein (N, 42 kDa), C-terminal extein (C, 36 kDa), and C-terminal extein fused intein (I–C, 55 kDa) were observed (Fig. 1b). Based on the band pattern, splicing activities of CP1201 RIR1 variants could be classified into three types (Table 1, Fig. 2). The residue of wild-type CP1201 RIR1 protein in position 539 is cysteine, which belongs to splicing type 1 in our cleavage analysis. It has been reported that complete splicing is observed when the first N-terminal amino acid of C-extein was Cys, and Ser (Fig. 1b, lanes C and S) [14–16]. Variants of C539F, C539Y, C539 W, and C539T performed type 2 splicing, to which N-terminal and C-terminal exteins produced by the cleavage of the precursor proteins could not be linked by forming peptide bond (Fig. 1b, lanes F, Y, W and T). When
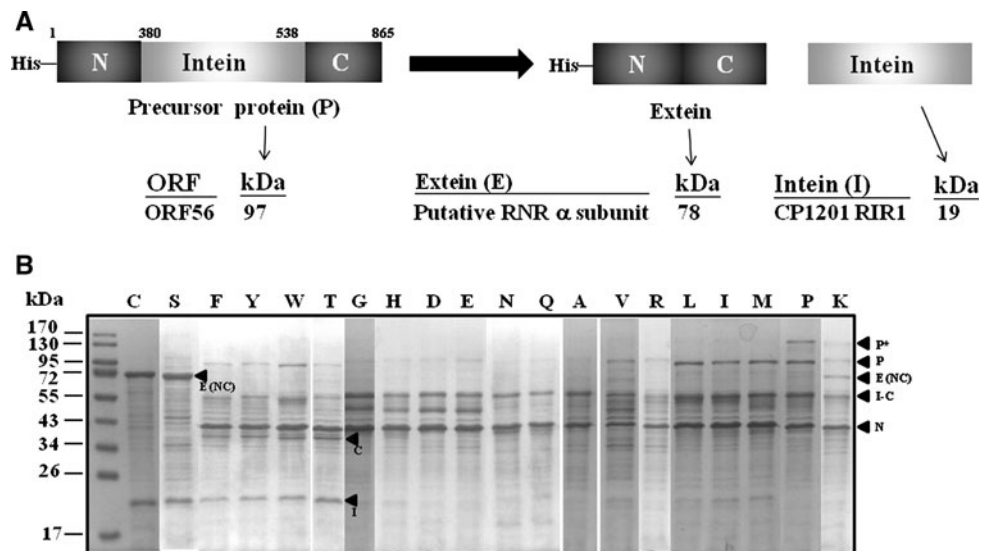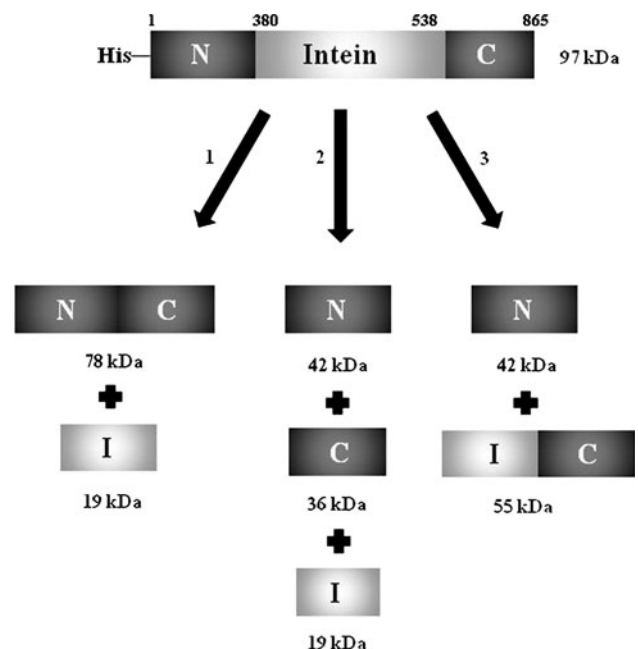
Fig. 1 Protein splicing of CP1201 RIR1 intein. **a** schematic illustration of the precursor protein containing N-terminal extein (N), intein (I), and C-terminal extein (C). **b** SDS-PAGE analysis of cleavage of the C-extein as a function of the identity of the N-terminal residue of the C-extein. The one-letter amino acid code of the N-terminal residue of the C-extein for each protein is given above the corresponding *lane*. Each variant was incubated under the cleavage conditions as described in materials and methods for 6 h at 28°C. For each sample, 5 μg crude cell extract was loaded. The bands labeled P*, P, E(NC), I, N, C. and I–C correspond to branched intermediate, uncleaved precursor protein, extein, intein, N-terminal extein, C-terminal extein, and C-terminal extein fused intein, respectively

Table 1 Effect of the first C-extein residue on cleavage at the C-terminus of the CP1201RIR1 intein

The first C-extein Residue	Splicing type	Completed splicing	N-cleavage	C-cleavage	Ligated extein
C, S	1	+	+	+	+
F, Y, W, T	2	−	+	+	−
G, H, D, E, N, Q, A	3	−	+	−	−
V, R, L, I, M, P, K					

the amino acid in Cys539 position was replaced with Gly, His, Asp, Glu, Asn, Gln, Ala, Val, Arg, Leu, Ile, Met, Pro, and Lys, the type 3 splicing occurred, in which C-terminal cleavage and extein ligation were not observed. (Fig. 1b, lanes G, H, D, E, N, Q, A, V, R, L, I, M, P, and K). In types 2 and 3 splicing, the C-terminal cleavage was most efficient for Phe, Thr, Trp, and Tyr, and no C-terminal cleavage activity for other substituted residues. However, in *Saccharomyces cerevisiae* VMA gene containing *Sce* VMA intein, the C-terminal cleavage was the most efficient for the first N-terminal amino acid of C-extein is Ala, Gln, or Met, but less efficient for Asp, Arg, Glu, His, Ile, Lys, and Val [17]. In type 3 splicing, the N-terminal cleavage has the following preference: Asp, Glu, Gly, His > Ala, Asn, Gln > Arg, Ile, Leu, Met, Val > Lys, Pro. A branched intermediate with apparent molecular mass (140 kDa) higher than its unspliced precursor (97 kDa) was found when the Cys539 was replaced with Lys and Pro (Fig. 1b, lanes K and L). The accumulation of branched intermediate is also reported from the splicing of N454A/C455S double mutant of *S. cerevisiae* VMA protein [18]. Our data indicated that protein-splicing activities could be determined


Fig. 2 A schematic representation of the three major splicing types of CP1201RIR1 variants

by the N-terminal residue of C-extein, and the level of cleavage activities is largely depends on the difference of the intein-contained proteins.
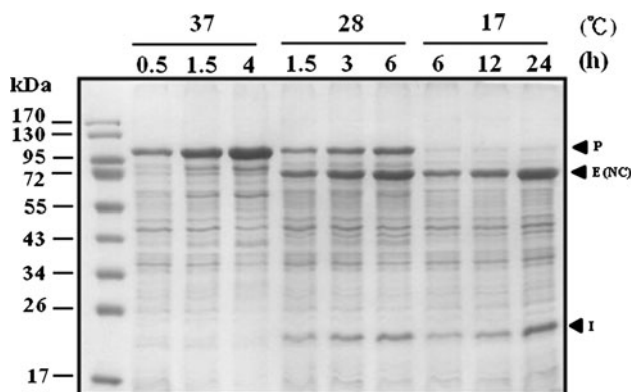
Fig. 3 SDS-PAGE analysis of N- and C-terminal cleavages of G392S variant. The variants were incubated at 17, 28, and 37°C. In each *lane*, 5 μg of total crude cell extract was loaded and SDS-PAGE was performed
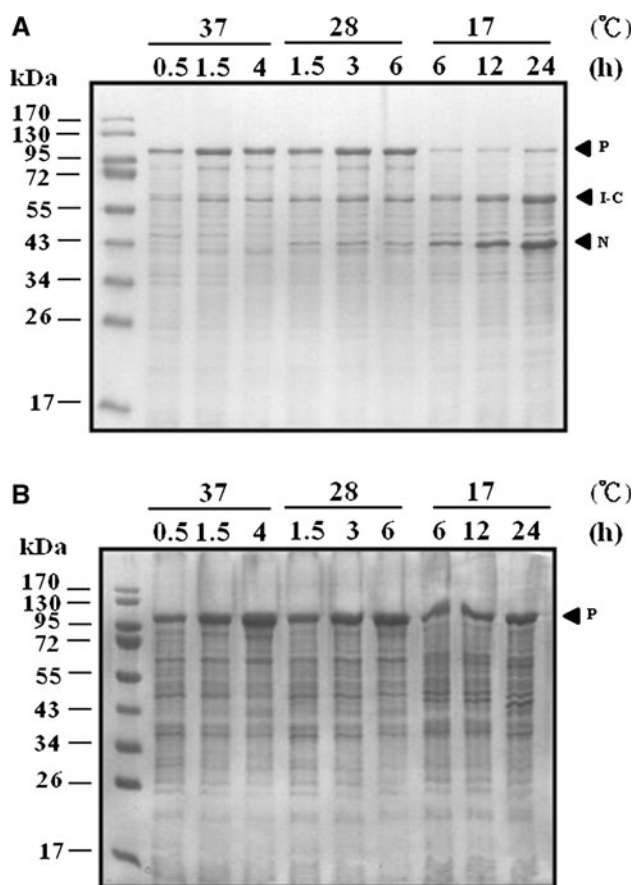

Fig. 4 SDS-PAGE analysis of the cleavage of G392S/C539G variant (**a**) and G392S/C539I variant (**b**). The variants were incubated at 17, 28, and 37°C. In each *lane*, 5 μg of total crude cell extract was loaded and SDS-PAGE was performed

Temperature-Dependent Splicing of G392S Variant

A G392S variant of RIR1 protein was obtained unexpectedly from the site-directed mutagenesis of the N-terminal residue of the C-extein by overlap PCR method. No splicing activity was detected at 37°C, whereas protein bands corresponding to the precursor protein, extein, and intein were found in the cleavage mixture incubated at 17 and 28°C (Fig. 3). Apparently, G392S variant was a temperature-sensitive protein with the splicing activity at lower temperature. Temperature-induced protein splicing is also found in DNA polymerase II (Pol II) DP2 subunit from *Pyrococcus abyssi.* [19]. In contrast to G392S variant, unspliced precursor fusion protein of the *P. abyssi* Pol II DP2 subunit in *E. coli* is found at 20°C, and the protein-splicing activity of precursor fusion protein is undergone in vitro at 40°C or higher.

If temperature-sensitive (G392S) variant is designed for the protein purification using cleavage at the intein N-terminus, then the C-terminal cleavage activity must need to be blocked. Since variants C539G and C539I have no C-terminal cleavage activity (Table 1), two double mutants of G392S/C539G and G392S/C539I were constructed. Variant G392S/C539G was spliced at 28°C and more efficient at 17°C, showing protein bands corresponding to the sizes of the precursor protein (97 kDa), N-terminal extein (42 kDa), and C-terminal extein-fused intein (55 kDa) (Fig. 4a). However, G392S/C539I variant was expressed as inclusion bodies at tested temperatures (Fig. 4b). The temperature shift experiment for variant
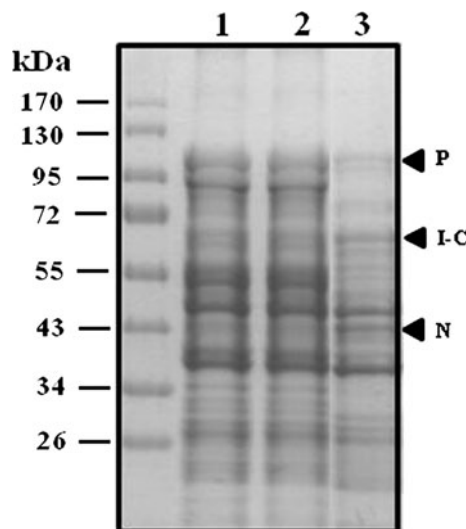
Fig. 5 Temperature shift assay of G392S/C539G variant. An overnight culture of *E. coli* NovaBlue cells carrying G392S/C539G RIR1 variant was inoculated into 10 ml of LB medium containing ampicillin (100 μg/ml) and then grown at 37°C until an OD_{600} of 0.7 was reached. The cells were induced with 1 mM IPTG at 37°C for 4 h (*lane 1*) and subsequently shift to 37°C (*lane 2*) and 17°C (*lane 3*) for 16 h. For each sample, 5 μg crude cell extract was loaded

G392S/C539G was carried out using the recombinant *E. coli* cells which have been induced for *rir1* gene expression at 37°C for 4 h. It was found that the G392S/C539G protein could be synthesized at 37°C (Fig. 5, lane 1) thereby being inhibited from splicing, and would subsequently undergo splicing when shifted to 17°C (Fig. 5, lane 3).

In summary, we demonstrated that the C539 residue is important in the splicing activities of CP1201RIR1 and substitution of C539 with other residues leads to the changes of pattern and level of protein splicing. The G392S/C539G-variant protein could be synthesized at 37°C and spliced at 17°C. However, the mechanism of low-temperature-induced splicing remains to be elucidated by 3D structure analysis.

References

1. Paulus H (2000) Protein splicing and related forms of protein autoprocessing. Annu Rev Biochem 69:447–496
2. Pietrokovski S (1994) Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins. Protein Sci 3:2340–2350
3. Gogarten JP, Senejani AG, Zhaxybayeva O, Olendzenski L, Hilario E (2002) Inteins: structure, function, and evolution. Annu Rev Microbiol 56:263–287
4. Perler FB (2002) Inbase: the intein database. Nucleic Acids Res 30:383–384
5. Hodges RA, Perler FB, Noren CJ, Jack WE (1992) Protein splicing removes intervening sequences in an archaea DNA polymerase. Nucleic Acids Res 20:6153–6157
6. Xu MQ, Southworth MW, Mersha FB, Hornstra LJ, Perler FB (1993) In vitro protein splicing of purified precursor and the identification of a branched intermediate. Cell 75:1371–1377
7. Wallace CJ (1993) The curious case of protein splicing: mechanistic insights suggested by protein semi synthesis. Protein Sci 2:697–705
8. Clarke ND (1994) A proposed mechanism for the self-splicing of proteins. Proc Natl Acad Sci USA 91:11084–11088
9. Cooper AA, Stevens TH (1995) Protein splicing: self-splicing of genetically mobile elements at the protein level. Trends Biochem Sci 20:351–356
10. Xu MQ, Paulus H, Chong S (2000) Fusions to self-splicing inteins for protein purification. Methods Enzymol 326:376–418
11. Chen CL, Pan TY, Kan SC, Kuan YC, Hong LY, Chiu KR, Sheu CS, Yang JS, Hsu WH, Hu HY (2008) Genome sequence of the lytic bacteriophage P1201 from *Corynebacterium glutamicum* NCHU 87078: evolutionary relationships to phages from Corynebacterineae. Virology 378:226–232
12. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77:51–59
13. Chiu WC, You JY, Liu JS, Hsu SK, Hsu WH, Shih CH, Hwang JK, Wang WC (2006) Structure–stability–activity relationship in covalently cross-linked *N*-carbamoyl D-amino acid amidohydrolase and *N*-acylamino acid racemase. J Mol Biol 359:741–753
14. Belfort M, Roberts RJ (1997) Homing endonucleases: keeping the house in order. Nucleic Acids Res 25:3379–3388
15. Jurica MS, Stoddard BL (1999) Homing endonucleases: structure, function and evolution. Cell Mol Life Sci 55:1304–1326
16. Chen L, Benner J, Perler FB (2000) Protein splicing in the absence of an intein penultimate histidine. J Biol Chem 275:20431–20435
17. Chong S, Montello GE, Zhang A, Cantor EJ, Liao W, Xu MQ, Benner J (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. Nucleic Acids Res 26:5109–5115
18. Chong S, Shao Y, Paulus H, Benner J, Perler FB, Xu MQ (1996) Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. J Biol Chem 271:22159–22168
19. Mills KV, Connor KR, Dorval DM, Lewandowski KT (2006) Protein purification via temperature-dependent, intein-mediated cleavage from an immobilized metal affinity resin. Anal Biochem 356:86–93