Enzyme and Microbial Technology 118 (2018) 6–12

Contents lists available at ScienceDirect

Enzyme and Microbial Technology

journal homepage: www.elsevier.com/locate/enzmictec

Intermolecular disulfide bonds between unpaired cysteines retard the C-terminal trans-cleavage of Npu DnaE

Yanran Xu[a], Lei Zhang[a], Buyong Ma[c], Lifu Hu[a], Huili Lu[a], Tonglu Dou[a], Junsheng Chen[a,*], Jianwei Zhu[a,b,*]

[a] Engineering Research Center of Cell and Therapeutic Antibody, Ministry of Education, School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200240, China
[b] Jecho Laboratories, Inc., Frederick, MD, USA
[c] Basic Science Program, Leidos Biomedical Research, Inc. Cancer and Inflammation Program, National Cancer Institute, Frederick, MD, 21702, USA

ARTICLE INFO

Keywords:
Split intein
Redox trap
C-terminal cleavage
Npu DnaE

ABSTRACT

Npu DnaE is a naturally occurred split intein possessing robust trans-splicing activity and could be engineered to perform rapid C-terminal cleavage module by a single mutation D118G. Unfortunately, however, for this modified selfcleaving module, reducing agents were needed to trigger the rapid cleavage, which prevents the utilization in purification of disulfide bonds containing recombinant proteins. In this study, we demonstrated that the unpaired cysteine residues in *Npu* DnaE tend to form disulfide bonds, and contributed to the reduction of the cleavage under non-reducing conditions. This redox trap can be disrupted by site-directed mutation of these unpaired cysteines. The results further indicated that the position 28 and 59 may play certain roles in the correct folding of the active conformation.

1. Introduction

Inteins are self-catalytic enzymes that can excise themselves from precursor proteins and concomitantly ligate the two exteins together with a peptide bond [1–4]. Many inteins have been engineered to self-cleaving elements which could perform N/C cleavage under the control of pH or reducing reagents. By combining these self-cleaving elements with affinity tags such as chitin-binding domain (CBD) [5], maltose-binding protein (MBP) [6], and elastin like protein (ELP) [7–9], a number of self-cleaving purification tags were developed. These intein based self-cleaving purification tags are invaluable to the purification of a variety recombinant protein as they can facilitate convenient affinity purification and efficient removal of purification tags without expensive protease.

Intein based controllable C-terminal cleavage have been applied to the tagless purification of recombinant proteins, including therapeutic proteins containing disulfide bonds [10]. Wu [11] used a self-cleaving ΔI-CM for the tagless purification of recombinant human antibody fragments expressed in *E. coli* cells. The recombinant protein was secreted to the periplasmic space and correctly disulfide bonded, the tag was completely removed by ΔI-CM. Shi [12] employed a novel soluble tag Zbasic together with ΔI-CM to express and taglessly purify recombinant human interleukin-15. Human epidermal growth factor (hEGF) was purified with correct disulfide bonds patterns by using *Ssp* DnaB mini intein [13]. In most of these cases, the main drawback was the uncontrollable premature cleavage of the cleaving intein during the expression of the recombinant protein. One solution to alleviate the premature cleavage is to use the split inteins [7,14,15], which are active only when the complementary two fragments associated.

An important recent advancement in intein based self-cleaving tag has been the establishment of the split intein mediated ultra-rapid purification of tagless protein (SIRP). In this case, *Npu* DnaE, a naturally occurred ultra-fast split intein, was engineered into a C-terminal cleavage element by introducing a single mutation D118 G, the resultant could achieve ~80% C-terminal cleavage within 3 h at room temperature [16]. This cleavage mutant was further developed into the SIRP method by placing affinity tags in the C-terminus rather than the traditional N-terminus of the N segment. This system could perform almost complete C-terminal cleavage within 30 min at room temperature in the presence of reducing agent, thus a promising tool for protein purification [17]. Although the SIRP system is highly effective and inexpensive, the need of reducing agents as inducer makes it not ideal for the purification of proteins containing disulfide bonds, especially the therapeutic proteins such as monoclonal antibodies (mAbs) and human cytokines.

Unpaired cysteine residues present in an enzyme may be

* Corresponding authors at: Engineering Research Center of Cell and Therapeutic Antibody, Ministry of Education, School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200240, China.
E-mail addresses: js.chan@sjtu.edu.cn (J. Chen), jianweiz@sjtu.edu.cn (J. Zhu).

https://doi.org/10.1016/j.enzmictec.2018.06.013
Received 31 December 2017; Received in revised form 15 June 2018; Accepted 27 June 2018
Available online 03 July 2018
0141-0229/

problematic, as they may improperly form disulfide bonds leading to misfolding and subsequently the decrease in activity. Therefore, it is desirable to decrease the number of unpaired cysteines either by substitution or deleting. However, the unpaired cysteines may contribute to the folding or activity to the enzyme, thus the substitution or deletion may affect the activity of enzymes. Therefore it is necessary to investigate the involvement of the unpaired cysteine in the formation of undesired disulfide bonds and its role in the folding and activity to the enzyme, prior to the development of the cysteine-deficient mutants.

Here in this report, we probed the role of the unpaired cysteine residues in *Npu* DnaE to the thiol enhancing effect to the C-terminal cleavage. By site directed mutagenesis of the unpaired cysteine residues in *Npu* DnaE, we demonstrated that disulfide bonds tend to form between these unpaired cysteines and thus retarding the C-cleavage, this redox trap can be disrupted by substitution of these unpaired cysteines. Additionally, these results indicated the Cys28 and Cys59 may play certain roles in forming correct conformation. Our results have useful implications for designing of thiol-independent C-cleavage DnaE inteins and for understanding the structural and dynamical traits of DnaE inteins.

2. Materials and methods

2.1. Bacterial strains and reagents

E. coli strains DH5α and BL21 (DE3) (TransGen Biotech, Beijing, China) were used for general cloning and recombinant protein expression respectively. Restriction enzymes were from Takara (Dalian, China). Site-directed mutagenesis kit was from Vazyme (Nanjing, China). Ni sepharose 6 fast flow (GE Healthcare, NJ, USA) were used to purify the His-tagged recombinant protein. Chemical reagents were purchased from Sinopharm Chemical Reagent Co. Ltd (Shanghai, China). Gene encoding the *cis Npu* DnaE was synthesized by GenScript Inc (Nanjing, China).

2.2. Plasmid construction

The protein constructs and numberings are shown in Table 1. Sequences of primers used in plasmid construction were listed in Table 2. Primers 1 and 2 were used to amplify the *Npu* DnaE N gene with the mutation of C1A from the synthesized *cis Npu* DnaE construct, the PCR products were then digested by *Nde*I and *Not*I and inserted into pET28a to yield the construct 1. The constructs of N segment bearing C28S, C59S and C28S/C59S mutations were generated by site-directed mutagenesis using construct **1** as template, with primer pairs 3/4 and 5/6. The gene encoding the *Npu* DnaE C segment extended with three native C-extein amino acids "CFN" was amplified with primers 9 and 10 and then joined with eGFP gene (amplified with primers 11 and 12) via overlap extension PCR and cloned into pET30a between *Nde*I and *Xho*I sites, the D118 G mutation was then introduced by site directed mutagenesis using the primers 7 and 8 to give the construct C*-CFN-eGFP. The C*-AFN-eGFP mutant was obtained by site-directed mutagenesis from the construct C*-CFN-eGFP using primers 14 and 15. The contiguous *Npu* DnaE variants containing one or two amino acid substitutions and the native C-extein "CFN" (constructs 5–8) were obtained by site-directed mutagenesis from the synthesized *cis Npu* DnaE construct using primer pairs 3/4, 5/6 and 7/8 and cloned into pET28a between *Nde*I and *Not*I. All of the constructed plasmids were confirmed by sequencing.

Table 1
Protein constructs used in this study.

Construct	Short name	Protein sequence	Molecular weight (kDa)
1	N_{C1A}	6xH-*Npu* DnaE N_{C1A}	12.6
2	$N_{C1A/C28S}$	6xH-*Npu* DnaE $N_{C1A/C28S}$	12.6
3	$N_{C1A/C59S}$	6xH-*Npu* DnaE $N_{C1A/C59S}$	12.6
4	$N_{C1A/C28S/C59S}$	6xH-*Npu* DnaE $N_{C1A/C28S/C59S}$	12.6
5	N_{C1A}-C*	6xH-*Npu* DnaE N_{C1A} -C_{D118G}	16.7
6	$N_{C1A/C28S}$-C*	6xH-*Npu* DnaE $N_{C1A/C28S}$ -C_{D118G}	16.7
7	$N_{C1A/C59S}$-C*	6xH-*Npu* DnaE $N_{C1A/C59S}$ -C_{D118G}	16.7
8	$N_{C1A/C28S/C59S}$-C*	6xH-*Npu* DnaE $N_{C1A/C28S/C59S}$-C_{D118G}	16.7
9	C*-CFN-eGFP	*Npu* DnaE C_{D118G}-CFN-eGFP-6xH	35.0
10	C*-AFN-eGFP	*Npu* DnaE C_{D118G}-AFN-eGFP-6xH	35.0

2.3. Protein expression and purification

All recombinant protein expressions were carried out by *E. coli* BL21 (DE3). Single colonies bearing appropriate plasmids were picked and cultured in 5 mL liquid LB supplemented with 50 μg/mL kanamycin overnight at 37°C. Then the culture was inoculate into 500 mL LB containing the same antibiotics and cultured at 37 °C until the OD_{600} reached 0.6. Protein expression was then induced by addition of 0.5 mM IPTG at 25 °C for 20 h. After expression, cells were pelleted by centrifugation and stored at −80 °C for later use.

All *Npu* DnaE-N constructs and *Npu* DnaE contiguous constructs contain a 6-His tag at the N-terminal while all *Npu* DnaE C-eGFP constructs have the 6-His tag attached at the C-terminal. For recombinant protein purification, pelleted cells were re-suspended in Buffer A (20 mM Tris − HCI, 150 mM NaCl, pH7.6) and lysed by passing through the AH1500 high-pressure homogenizer (ATS). The suspension was centrifuged at 16,000*g* for 30 min at 4 °C to remove the insoluble cell debris. The clear lysate was then loaded onto Nickel column by gravity (GE Healthcare), and washed extensively with Buffer A supplemented with 50 mM imidazole, and then the recombinant proteins were eluted with Buffer A containing 150 mM imidazole. The fractions containing intein precursors were pooled and exchanged against cleavage buffer (20 mM Tris − HCI, 150 mM NaCl, 2 mM EDTA, pH7.6) by dialysis.

2.4. In vitro C-terminal cleavage experiments

Purified N segments and C precursors were mixed in an equal molar ratio and incubated at a specified temperature, with or without reducing agents. For the characterization of the cleavage kinetics, aliquots were removed from the reaction at different time points after the initiation of the reaction and immediately quenched by boiling with SDS-PAGE sampler buffer at 95 °C for 10 min. Samples were analyzed on 12% SDS-PAGE gels and stained with Coomassie brilliant blue. The scanned image was analyzed by ImageJ for quantification and the percentage of cleavage was estimated from the molar ratio of the cleavage product and the original precursor. All experiments were performed in triplicates.

2.5. Molecular dynamics simulations

The structure of *trans*-splicing *Npu* DnaE intein (PDB: 4LX3) was downloaded from RCSB. Cys28 and Cys59 were substituted by Ser residues in PyMOL 1.5.03 and the mutated structure was named as 'Ser28, 59'. The simulated structures were solvated in a TIP3P water box with a margin of at least 10 Å, and then sodium and chloride ions were added to the system up to a concentration of 0.10 M. AMBER ff12SB force field was used to perform efficient simulations with a constant pressure of 1 atm and a temperature of 298 K. The time step was 2 fs with an SHAKE constraint on all bonds with hydrogen atoms. Long-range electrostatic interactions were calculated with the Particle Mesh Ewald method in Amber 14. The MD runs were performed after 6000 steps of minimization, 50 ps heating, and 100 ps equilibration at

Table 2
Primers used in this study.

Number	Name	5'-3' SEQUENCE
1	NpuN(C1 A)-F	AAACATATGGCATTAAGCTATGAAACG
2	NpuN-R	AAAGCGGCCGCTCAATTCGGCAAATTATC
3	C28S-F	CGCATCGAAAGCACTGTTTATAGCGTTGATAATAATGGAAATATTTATAC
4	C28S-R	ATAAACAGTGCTTTCGATGCGCTTTTCTACAATTTTACCAATCGGTAATA
5	C59S-F	TTTGAGTATAGCTTGGAAGATGGTTCATTGATTCGGGCAACAAAAGACCA
6	C59S-R	ATCTTCCAAGCTATACTCAAACACCTCTTGTTCTCCGCGATCGTGCCATT
7	D118G-F	AATGTCTATGGAATTGGAGTTGAGCGCGACCATAATTTTGCA
8	D118G-R	AACTCCAATTCCATAGACATTTTGTTTGCCTAAATATTTACG
9	NpuC-F	AAA CATATG ATGATCAAAATAGCCACACGTAAATAT
10	NpuC(eGFP)-R	CAGCTCCTCGCCCTTGCTCATGCTTCCATTGAAACAATTAGAAGC
11	eGFP-F	GCTTCTAATTGTTTCAATGGAAGCATGAGCAAGGGCGAGGAGCTG
12	eGFP-R	AAA CTCGAG CTTGTACAGCTCGTCCATGCCGAGAGT
13	NpuC-R	AAAGCGGCCGCTCAATTGAAACAATTAGAAGCTATGAAGCCATT
14	C + 1 A-F	GCTTCTAATGCTTTCAATATGAGCAAGGGCGAGGAGCTGTTC
15	C + 1 A-R	CATATTGAAAGCATTAGAAGCTATGAAGCCATTTTTGAGTGC

298 K. Cpptraj of AMBER was used to analyze trajectories [18]. The RMSD (root mean squared deviation) and RMSF (root mean square fluctuations) analysis of the dynamic trajectories was obtained using starting crystal structures as the reference. Both MD simulations were accelerated with the CUDA version of PMEMD in GPU cores of NVIDIA® Tesla K20 [19].

3. Results

3.1. Generation and purification of Npu DnaE variants

Recently, it was reported that a single mutation D118 G could fully abolish the *trans*-splicing activity and confer an elevated C-terminal cleavage activity to the *Npu* DnaE intein. However the reaction was still thiol-dependent, making it unsuitable for the purification of proteins containing disulfide bonds. According to the primary sequence of the *Npu* DnaE, there exist 4 cysteines. Among which, Cys1 and Cys + 1 were required for the splicing activity but not essential for C-terminal cleavage, while the Cys28 and Cys59 were located in the *Npu* DnaE N segment. To investigate the potential influence of the cysteines on the thiol-enhancing effect of the D118 G cleavage mutant, various fusion proteins were prepared as listed in Table 1. For the *Npu* DnaE N variants, a plasmid derived N-terminal hexahistidine was included for purification. For the *Npu* DnaE C variants, the enhanced green fluorescent protein (eGFP) was used as the C-extein, and a hexahistidine tag was attached at the C-terminal for purification. The native flanking residue "CFN" downstream the *Npu* DnaE C was kept for the construct C*-CFN-eGFP, while in the construct C*-AFN-eGFP it was replaced to "AFN". For the contiguous *Npu* DnaE constructs, the native extein "CFN" was kept and a hexahistidine tag was placed at the N terminal. All *Npu* DnaE N and C variants as well as *Npu* DnaE contiguous constructs could be expressed in soluble form in *E. coli* BL21 (DE3) and purified via Ni sepharose affinity chromatography. For all the four *Npu* DnaE N segment variants (construct 1–4), the SDS-PAGE analysis showed that a 13 kDa band could be observed (Fig. 1A), which met the theoretical molecular mass of the *Npu* DnaE N segment monomer. For the construct N_{C1A}, $N_{C1A/C28S}$ and $N_{C1A/C59S}$ that contained cysteine residues, additional bands could be seen between 25 kDa and 35 kDa under non-reducing conditions. These bands disappeared as the 13 kDa monomer band became stronger under reducing conditions, which indicates these additional bands were the dimer of each *Npu* DnaE N variants. There were no bands other than the monomer observed in $N_{C1A/C28S/C59S}$ under non-reducing conditions (Fig. 1A). For the *Npu* DnaE C-eGFP fusions, monomer as well as dimer of C*-CFN-eGFP could be observed in the non-reducing SDS-PAGE, while the construct C*-AFN-eGFP was in monomer form only (Fig. 1C). To investigate the potential interaction between Cys + 1 and cysteines residues in the N segment, a series of *Npu* DnaE contiguous variants were constructed. For the variants containing cysteines in N segment (constructs 5–7), besides the expected 17 kDa monomer band, bands moved slightly faster than the monomer were also observed, which disappeared in the reducing SDS-PAGE (Fig. 1B). This may due to the intramolecular disulfide bonds formed between C + 1 and C28/C59 which resulting a tighter conformation with a faster migration rate. Several additional bands could be observed and were speculated as the different types of the disulfide bond linked oligomer, as they disappeared on the reducing SDS-PAGE. Only monomer and dimer could be found in $N_{C1A/C28S/C59S}$-C* (Fig. 1B). These results indicate that, the Cys28, Cys59 and Cys + 1 may facilitate the formation of intermolecular disulfide bonded N—N homodimer and N—C heterodimer (Fig. 1D).

3.2. Influence of cysteines on C-terminal cleavage

It was previously reported that the reaction between *Npu* DnaE N (C1 A) and *Npu* DnaE C (D118 G) was thiol-activated but not thiol-dependent [16], therefore, we first test if thiol could enhance the C-terminal cleavage. Purified N_{C1A} and C*-CFN-eGFP were mixed at an equimolar in the cleavage buffer, in the presence or absence of 2 mM DTT and incubate at 22 °C for 12 h. The SDS-PAGE showed that the C-terminal cleavage reaction could occur in the absence of DTT but with a low yield and rate and the addition of DTT greatly enhanced the reaction (Fig. 2A, Lane 1 compared with Lane5), which was consistent with previous reports [16]. To test if the cysteines in the *Npu* DnaE and its flanking C-extein were involved in the redox trap of the C-terminal cleavage, purified N and C variants were subjected to cleavage experiments, either with or without DTT. For the reaction between the cysteines containing N segment and C*-CFN-eGFP, the addition of DTT obviously promote the generation of the cleavage product compared with the reaction without DTT (Fig. 2A), indicating that the cysteines may hamper the cleavage reaction by forming disulfide bonds and this redox trap could be disrupted by the addition of DTT. Notably, in the reaction between $N_{C1A/C28S/C59S}$ and C*-CFN-eGFP, addition of DTT resulted more C-terminal cleavage than the reaction without DTT (Fig. 2A, Lane 4 compared with Lane 8). This enhancing effect of DTT may due to the breaking of the disulfide bonds between two C*-CFN-eGFP, as dimers could also be observed during the expression (Fig. 1C). To exclude the influence of Cys + 1, another set of cleavage was carried out with the C*-AFN-eGFP. Similar to C*-CFN-eGFP, the reaction of C*-AFN-eGFP with N variants containing cysteines were boosted by DTT, while caused no difference to the reaction with $N_{C1A/C28S/C59S}$. The cleavage product yield of reaction between $N_{C1A/C28S/C59S}$ and C*-AFN-eGFP in the absence of DTT was comparable to the reactions supplemented with DTT (Fig. 2B, Lane 12 compared with Lane 16). These results indicate that the cysteine residues were involved in the redox

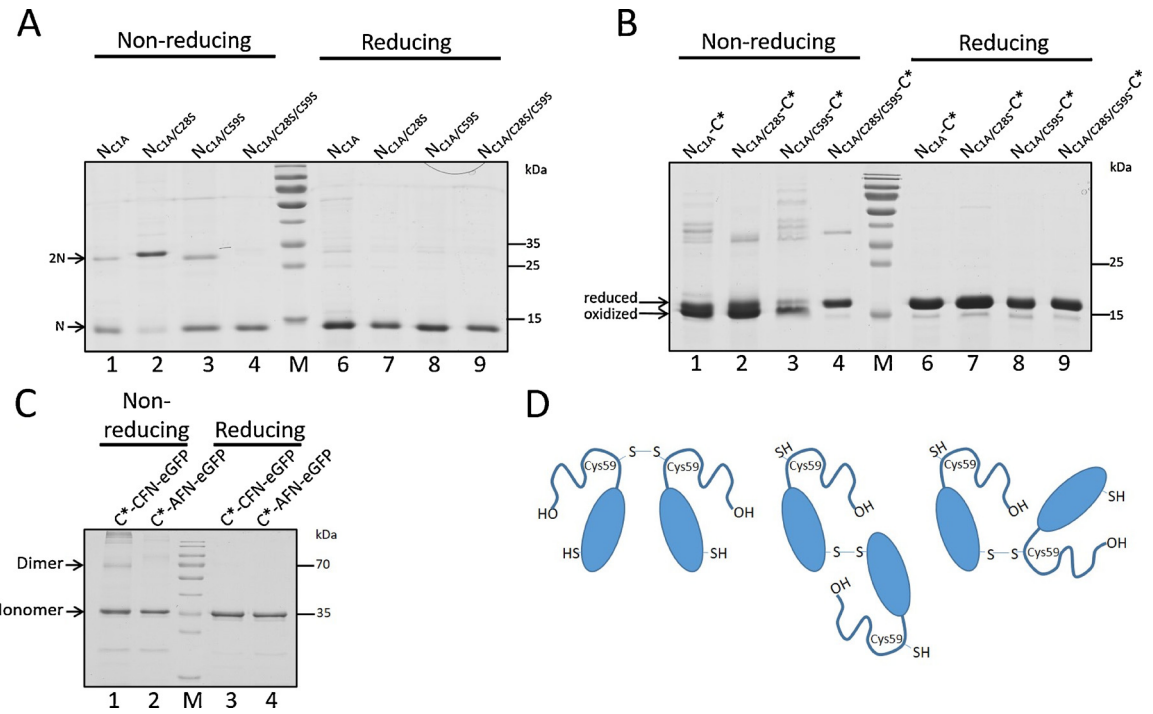

Fig. 1. SDS-PAGE analysis of proteins purified from E. coli BL21 (DE3) to exam the involvement of unpaired *Npu* DnaE Cys residues in disulfide bond formation, under reducing or non-reducing conditions. (A) N segment variants, (B) Variants of *Npu* DnaE in contiguous form, with the +1 residue kept as Cys, (C) C segment variants, (D) Schematic depicting the formation of disulfide bonds between unpaired cysteines. Mutations to each unpaired Cys residue were indicated above the lanes.

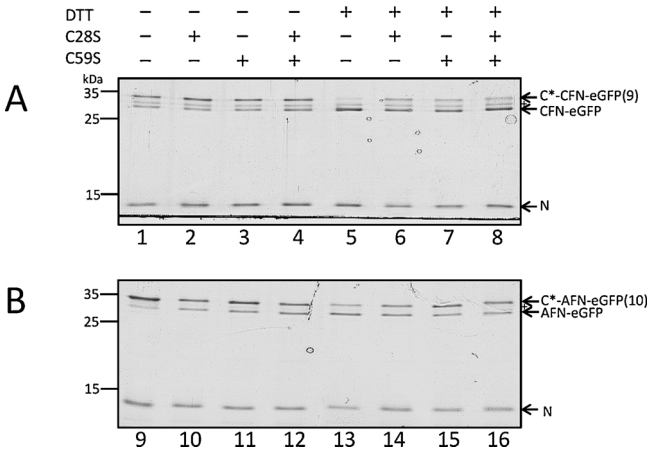

Fig. 2. Influence of the unpaired Cys residues in *Npu* DnaE to the *in vitro* C-terminal cleavage. After purification, N segments variants (constructs 1–4) were separately mixed either with C*-CFN-eGFP (construct 9) (A) or C*-AFN-eGFP (construct 10) (B) at an equimolar and incubated at 22 °C for 12 h in the presence or absence of 2 mM DTT. "+" denotes the unidentified band. The reactions were then analyzed by reducing SDS-PAGE. Mutations to each unpaired Cys residue were indicated above the lanes.

trap of the C-terminal cleavage reaction, and the cysteine-free *Npu* DnaE N and C precursors could facilitate the cleavage reaction.

3.3. Time dependent analysis of in vitro C-cleavage reaction

We have constructed the cysteine-free C-terminal cleavage mutant and demonstrated that the mutant could perform the cleavage in the absence of reducing agent. To further test whether the cleavage of cysteine-free mutant could be boosted by reducing agent, we then tested the C-terminal cleavage kinetics of the cysteine-free intein pairs at different temperatures in the presence or absence of DTT (Fig. 3A, B). The amount of cleavage products as well as the reaction rate increased as the temperature increased. The reaction rate remained unaffected in the absence of DTT at specified temperature. The highest cleavage efficiency was observed at 37°C, with about 50% cleavage achieved in 5 h (Fig. 3C), which is lower than that of the reaction between N_{C1A} and C*-CFN-eGFP supplemented with DTT (with ~80% cleavage in 3 h). Taken together, these results demonstrated that the cysteine residues in *Npu* DnaE were not required for the C-terminal cleavage, and the removal of these cysteines did not affect the cleavage efficiency. The independence of reducing agent indicates the potential application in the purification of disulfide containing protein as a self-cleaving tag.

3.4. Cys-Ser mutations decrease intein's catalytic conformation dynamics

In order to understand the differences of intein's activity caused by Cys-Ser mutations, we performed two 100 ns molecular dynamics simulations as described in Materials and Methods. As can be seen in Fig. 5, the Cys-Ser mutations not only caused the dynamics changes near mutated residues, they also led to large perturbation of protein dynamics throughout intein, especially around B-strand 108–125, which sits beneath the active site. It is interesting to see the sharp decreased dynamics of reside 109–113 and slightly increased fluctuation of residues 115–125. Moreover, the β-sheets around reside 52–54 and 109–113 were corrupted due to the mutation related perturbation. Previously, it was also illustrated that the protein dynamics differentiate intein's activity [20]. Our current experimental and computational results further support the correlation of intein's activity and its conformational dynamics.

4. Discussion

In this report, it has been demonstrated that in *Npu* DnaE, the unpaired non-catalytic cysteine residues tend to form redox trap to the C-terminal cleavage, and mutation of these unpaired cysteine disrupted this retardation.

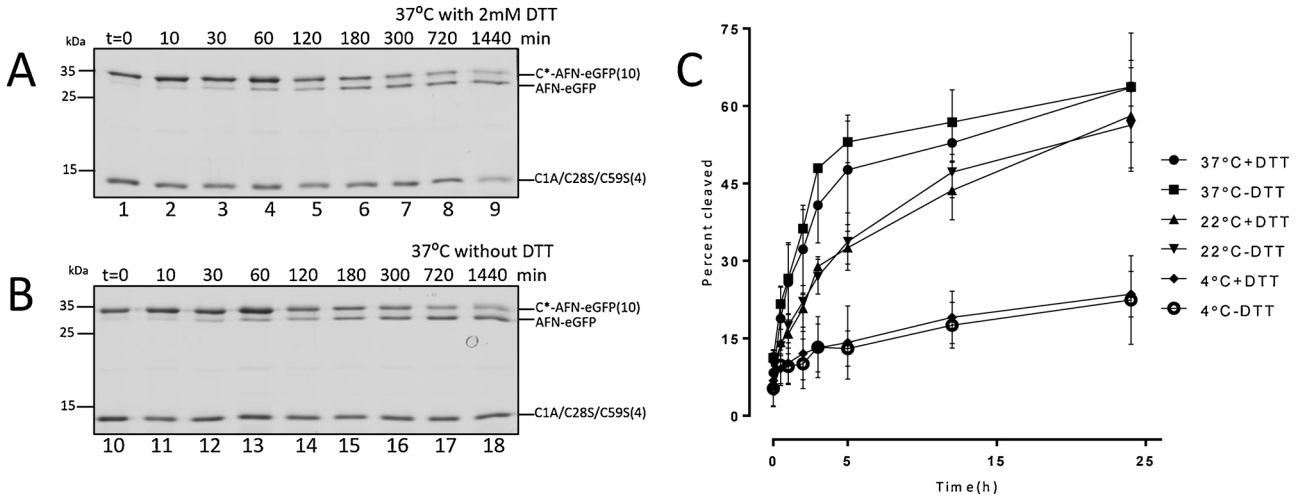

Fig. 3. Influence of the reducing agent DTT to the C-terminal cleavage kinetics of the Cysteine-free *Npu* DnaE cleavage variants. (A) and (B): SDS-PAGE analysis of the reaction between C1 A/C28S/C59S (construct 4) and C*-AFN-eGFP (construct 10) in the presence or absence of DTT at 37 °C. (C) Time course of the appearance of the AFN-eGPF at specified temperatures. The error bar represents the SD from three independent experiments.

Redox traps have been shown to suppress protein splicing with other inteins, either naturally occurred [21–23] or artificially engineered [24–28], by blocking catalytic cysteine residues with disulfide bonds. These redox traps could be utilized to fulfill intein based conditional splicing or cleavage. *Npu* DnaE, whose splicing activity was strictly thiol-dependent, was engineered into a C-terminal cleaving module. Although the catalytic cysteine residues, which may contribute to the redox trap, were mutated to confer the rapid C-terminal cleavage, the cleavage still exhibited as DTT dependent. It limited the utilization of this rapid C-terminal cleavage module in the purification of disulfide containing proteins. We proposed that the retardation under non-reducing conditions may due to the disulfide bridged dimer or oligomer complexes formed during segment preparation or fragment association. To test this hypothesis, we first investigated the involvement of the unpaired cysteines in the redox trap. For the native N segment, there are two unpaired cysteine residues, the Cys28 and Cys59, thus raising the chance of forming different types of internal or intramolecular disulfide bonds. As expected, dimers were observed in the expression of $N_{C1A/C28S}$ and $N_{C1A/C59S}$. These dimers were speculated as disulfide linked homodimers, as there was only one Cys in each intein monomer. These homodimers were further confirmed by the expression of $N_{C1A/C28S/C59S}$, which contained no cysteine residues and only monomer could be found on non-reducing SDS-PAGE. For the expression of N_{C1A}, only monomer and dimer could be seen under oxidized condition. No

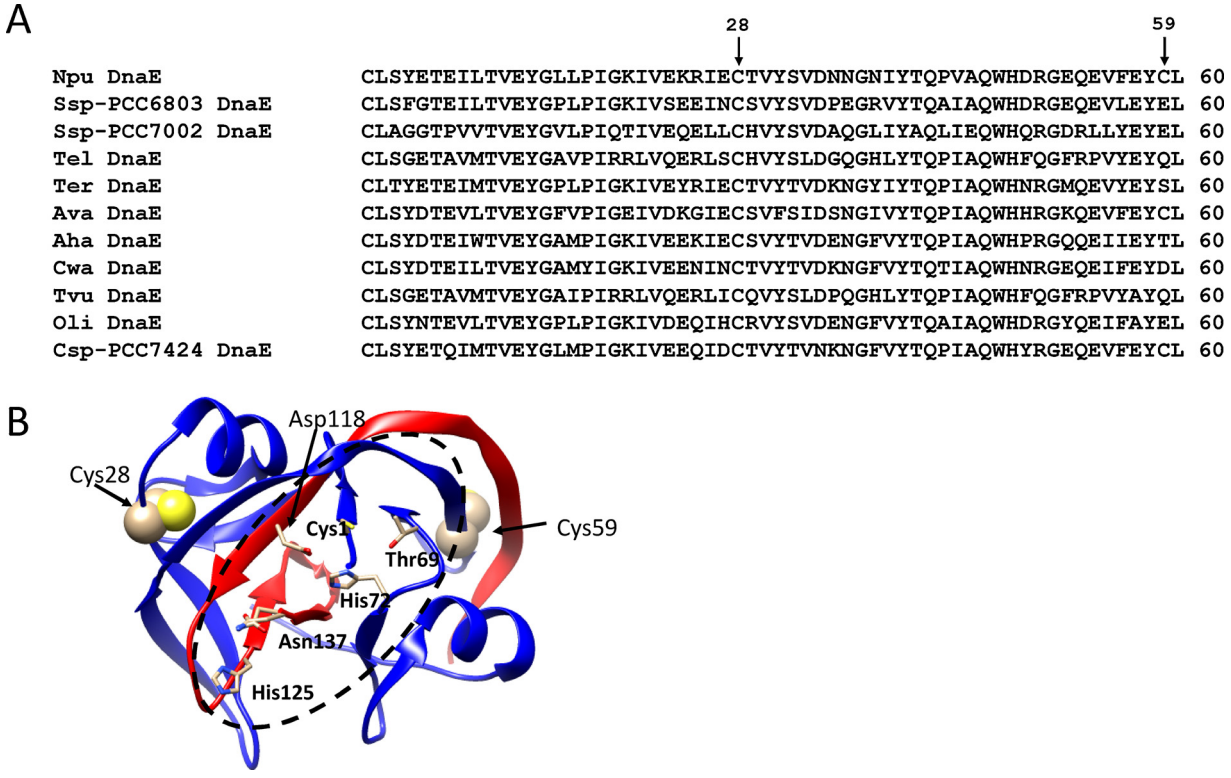

Fig. 4. Presentation of the location of the Cys28 and Cys59. (A) Multiple sequences alignment of split DnaE inteins primary sequences. (B) The crystal *trans*-structure of *Npu* DnaE intein (PDB: 4LX3) highlighting catalytic residues (sticks, marked with a dashed ellipse) and unpaired Cys28 and Cys59 (spheres).

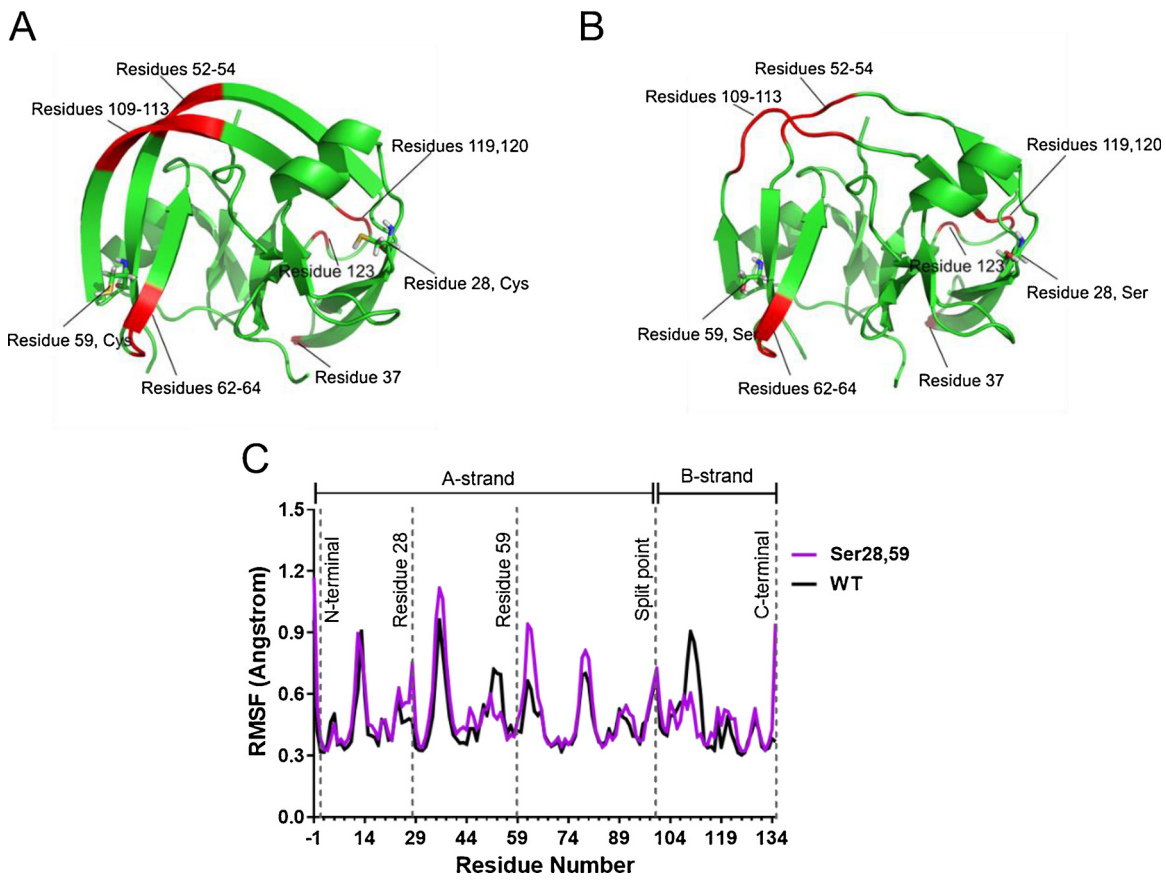

Fig. 5. Molecular dynamics simulation of *trans*-splicing *Npu* DnaE inteins. Structures of wild-type (A) and Cys-Ser mutated (B) molecules of *Npu* DnaE intein were simulated. The residues perturbed by the mutation were pointed out and marked in red color. (C) RMSF of the backbone Cα atoms for each amino acid in the simulated wild-type (WT) and mutated (Ser28, 59) inteins (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

significant difference in the migration rate of monomer could be observed whether on reducing or non-reducing SDS-PAGE, thus excluding the possibility of the internal disulfide bonds formed between Cys28 and Cys59. For the expression of *Npu* DnaE C segments, homodimers could be seen in construct C*-CFN-eGFP, which bearing the native C-extein "CFN", while the replacement of the Cys+1 with Ala resulted monomer only, indicating the C-extein containing unpaired cysteines may form disulfide bond bridged dimer or oligomer which may trap the cleavage. This speculation was further supported by the cleavage reaction between the cysteine-free $N_{C1A/C28S/C59S}$ and C*-CFN-eGFP, as the addition of DTT enhanced the C-terminal cleavage. It was predicted that, due to the close proximity, Cys+1 was ready to form intramolecular disulfide bonds with Cys1 when the two complementary segments associated, preventing further splicing reaction [22]. However, our results indicated that Cys+1 has chance to form internal disulfide bonds with Cys28 or Cys59 when the *Npu* DnaE was expressed in contiguous form. The assembly of *Npu* DnaE was reported to follow the "Capture and Collapse" mechanism, in which the association between the NpuN and NpuC was triggered by the electrostatic interaction between the disorder NpuC and NpuN2 (amino acid 51–102) [29]. This mechanism suggests that, upon the association there are chances for the split intein to form misfolded conformation which may lead to the retardation of splicing or cleavage. Additionally, these disulfide bonds were ready to form, as it formed when expressed in the reducing *E. coli* Bl21 (DE3) cells.

We next test whether the presence of the disulfide bonded homodimers affects the C-terminal cleavage of *Npu* DnaE. Our results showed that DTT could augment the cleavage when cysteine presented in *Npu* DnaE but had no enhancing effect to the reaction between $N_{C1A/C28S/C59S}$ and C*-AFN-eGFP, the cysteine-free pairs. Furthermore, cleavage kinetics study of the reaction of $N_{C1A/C28S/C59S}$ with C*-AFN-eGFP showed no difference either in the presence or absence of DTT, indicating the redox trap of cleavage could be disrupted by substitution of these unpaired cysteines, with cleavage kinetics unaffected.

Eliminating the unpaired cysteines could disrupt the redox-trap, thus avoiding reducing agents that may affect the POI. But on the other hand, lacking of this redox-trap will also lead the reaction uncontrolled, making the cleavage mutant not suitable in protein purification, especially in the case of on-column cleavage. To address this, other switchers should be employed together to confer the controllability to the thiol-independent *Npu* DnaE cleavage variant. For example, ELP (Elastin-like polypeptides) could be fused to the cleavage module to perform the column-free purification through ITC (Inverse transition cycling). Also, Zn^{2+} ions could be supplemented to suppress the C-terminal cleavage, and this suppression could be removed by adding EDTA.

It is also notable that the double mutation of Cys28 and Cys59 to serine leads to a reduction of cleavage both in rate and overall yield. This is interesting since according to the mechanism of intein splicing reaction, these two cysteine residues are not catalytic and structurally far away from catalytic site (Fig. 4B). It was also reported that mutation of these two cysteine residues to serine cause no negative effect to the expression [30]. This reduction in cleavage activity may due to the difference between serine and cysteine. Although serine is similar to cysteine in structure and size, they are chemically different in side chain, and cysteine is relatively more hydrophobic than serine. Substitution of these two cysteines with serine may cause subtle conformational changes or global conformational dynamics perturbations

that further affect the cleavage. The MD simulation further supported this speculation. In another word, these results indicate that both of these two positions may participate in intein fold stability or dynamics. Mutations distant from catalytic sites affecting intein activity was also observed with other intein and was described as “ripple effect” previously [31]. For example, in the *Mtu* RecA mini intein, mutation V67 L was reported minimally affect the crystal structure [32] but enhance the splicing by improving the stabilization of the conformation [33]. In the fast split DnaE inteins, several residues distant from the splicing sites contribute to the global stabilization or the correct orientation of the catalytic residues [34]. A variant of *Npu* DnaE engineered to recognize the noncanonical SGV extein accumulated several distal mutations [35]. Our findings also demonstrated another example of the “ripple effect”. However, the rationale of the influence from Cys28 and Cys59 to the cleavage activity of *Npu* DnaE remains elusive and systematic substitutions are necessary to identify optimized amino acids to these positions to restore full cleavage activity.

Although only *Npu* DnaE was studied in this report, the involvement of the unpaired non-catalytic cysteine residue in the redox trap is most likely common in split DnaE inteins, as the multiple sequence alignment of DnaE inteins showed that Cys28 was highly conserved (Fig. 4A).

5. Conclusions

In summary, we have demonstrated that in *Npu* DnaE the unpaired cysteines mediated the formation of intermolecular disulfide bonds between intein segments and further retard the cleavage. This redox trap could be avoided by eliminating these unpaired cysteines. Our results also indicate that the Cys28 and Cys59 may participate in the formation of correct conformation of *Npu* DnaE. Elimination of DTT as a cleavage inducer would unambiguously broaden the compatibility to target proteins containing disulfide bonds.

Author contribution

Junsheng Chen and Jianwei Zhu developed the concept and designed the experiments. Yanran Xu, Lei Zhang and Tonglu Dou performed the experiments. Buyong Ma gave advices on the molecular dynamics. Lifu Hu and Huili Lu performed the molecular dynamics simulations. Yanran Xu, Junsheng Chen and Buyong Ma analyzed the data and drafted the manuscript which was revised by all authors.

Acknowledgment

This work was supported by National Natural Science Foundation of China [Grant numbers 81773621 and 81473127].

References

[1] F.B. Perler, InBase: the intein database, Nucleic Acids Res. 30 (1) (2002) 383–384.
[2] H. Paulus, Protein splicing and related forms of protein autoprocessing, Annu. Rev. Biochem. 69 (1) (2000) 447–496.
[3] J. Peter Gogarten, Alireza G. Senejani, Olga Zhaxybayeva, Lorraine Olendzenski, E. Hilario, Inteins: structure, function, and evolution, Annu. Rev. Microbiol. 56 (1) (2002) 263–287.
[4] L. Han, J. Chen, K. Ding, H. Zong, Y. Xie, H. Jiang, B. Zhang, H. Lu, W. Yin, J. Gilly, J. Zhu, Efficient generation of bispecific IgG antibodies by split intein mediated protein trans-splicing system, Sci. Rep. 7 (1) (2017) 8360.
[5] T.C. Evans, J. Benner, M.-Q. Xu, The in vitro ligation of bacterially expressed proteins using an intein from methanobacterium thermoautotrophicum, J. Biol. Chem. 274 (7) (1999) 3923–3926.
[6] S. Mathys, T.C. Evans, I.C. Chute, H. Wu, S. Chong, J. Benner, X.-Q. Liu, M.-Q. Xu, 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 231 (1) (1999) 1–13.
[7] C. Shi, Q. Meng, D.W. Wood, A dual ELP-tagged split intein system for non-chromatographic recombinant protein purification, Appl. Microbiol. Biotechnol. 97 (2) (2013) 829–835.
[8] L. Tian, S.S.M. Sun, A cost-effective ELP-intein coupling system for recombinant protein purification from plant production platform, PLoS One 6 (8) (2011) e24183.
[9] M.R. Banki, L. Feng, D.W. Wood, Simple bioseparations using self-cleaving elastin-like polypeptide tags, Nat. Methods 2 (9) (2005) 659–662.
[10] B.A. Fong, W.-Y. Wu, D.W. Wood, The potential role of self-cleaving purification tags in commercial-scale processes, Trends Biotechnol. 28 (5) (2010) 272–279.
[11] W.-Y. Wu, K.D. Miller, M. Coolbaugh, D.W. Wood, Intein-mediated one-step purification of Escherichia coli secreted human antibody fragments, Protein Expr. Purif. 76 (2) (2011) 221–228.
[12] S. Shi, H. Chen, H. Jiang, Y. Xie, L. Zhang, N. Li, C. Zhu, J. Chen, H. Luo, J. Wang, L. Feng, H. Lu, J. Zhu, A novel self-cleavable tag Zbasic–ΔI-CM and its application in the soluble expression of recombinant human interleukin-15 in Escherichia coli, Appl. Microbiol. Biotechnol. 101 (3) (2017) 1133–1142.
[13] R.S. Esipov, V.N. Stepanenko, L.A. Chupova, U.A. Boyarskikh, M.L. Filipenko, A.I. Miroshnikov, Production of recombinant human epidermal growth factor using Ssp dnaB mini-intein system, Protein Expr. Purif. 61 (1) (2008) 1–6.
[14] G. Volkmann, W.C. Sun, X.Q. Liu, Controllable protein cleavages through intein fragment complementation, Protein Sci. 18 (11) (2009) 2393–2402.
[15] Y.F. Li, Self-cleaving fusion tags for recombinant protein production, Biotechnol. Lett. 33 (5) (2011) 869–881.
[16] M. Ramirez, N. Valdes, D. Guan, Z. Chen, Engineering split intein DnaE from Nostoc punctiforme for rapid protein purification, Protein Eng. Des. Sel. 26 (3) (2013) 215–223.
[17] D. Guan, M. Ramirez, Z. Chen, Split intein mediated ultra-rapid purification of tagless protein (SIRP), Biotechnol. Bioeng. 110 (9) (2013) 2471–2481.
[18] H.F. Chen, Binding induced intrinsically disordered protein folding with molecular dynamics simulation, Adv. Exp. Med. Biol. 827 (2015) 111–121.
[19] M.U. Rahman, H. Liu, A. Wadood, H.F. Chen, Allosteric mechanism of cyclopropylindolobenzazepine inhibitors for HCV NS5B RdRp via dynamic correlation network analysis, Mol. Biosyst. 12 (11) (2016) 3280–3293.
[20] M. Cronin, M.J. Coolbaugh, D. Nellis, J.W. Zhu, D.W. Wood, R. Nussinov, B.Y. Ma, Dynamics differentiate between active and inactive inteins, Eur. J. Med. Chem. 91 (2015) 51–62.
[21] M.C. Nicastri, K. Xega, L.Y. Li, J. Xie, C.Y. Wang, R.J. Linhardt, J.N. Reitter, K.V. Mills, Internal disulfide bond acts as a switch for intein activity, Biochemistry 52 (34) (2013) 5920–5927.
[22] W. Chen, L.Y. Li, Z.M. Du, J.J. Liu, J.N. Reitter, K.V. Mills, R.J. Linhardt, C.Y. Wang, Intramolecular disulfide bond between catalytic cysteines in an intein precursor, J. Am. Chem. Soc. 134 (5) (2012) 2500–2503.
[23] K.V. Mills, B.M. Lew, S.Q. Jiang, H. Paulus, Protein splicing in trans by purified N- and C-terminal fragments of the Mycobacterium tuberculosis RecA intein, Proc. Natl. Acad. Sci. U. S. A. 95 (7) (1998) 3543–3548.
[24] B.P. Callahan, N.I. Topilina, M.J. Stanger, P. Van Roey, M. Belfort, Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications, Nat. Struct. Mol. Biol. 18 (2011) 630.
[25] L. Saleh, M.W. Southworth, N. Considine, C. O'Neill, J. Benner, J.M. Bollinger, F.B. Perler, Branched intermediate formation is the slowest step in the protein splicing reaction of the Ala1 KlbA intein from Methanococcus jannaschii, Biochemistry 50 (49) (2011) 10576–10589.
[26] B.P. Callahan, M. Stanger, M. Belfort, A redox trap to augment the intein toolbox, Biotechnol. Bioeng. 110 (6) (2013) 1565–1573.
[27] C.X. Cui, W.T. Zhao, J.L. Chen, J.J. Wang, Q.Q. Li, Elimination of in vivo cleavage between target protein and intein in the intein-mediated protein purification systems, Protein Expr. Purif. 50 (1) (2006) 74–81.
[28] J.X. Shi, T.W. Muir, Development of a tandem protein trans-splicing system based on native and engineered split inteins, J. Am. Chem. Soc. 127 (17) (2005) 6198–6206.
[29] N.H. Shah, E. Eryilmaz, D. Cowburn, T.W. Muir, Naturally split inteins assemble through a “Capture and Collapse” mechanism, J. Am. Chem. Soc. 135 (49) (2013) 18673–18681.
[30] M. Vila-Perello, Z.H. Liu, N.H. Shah, J.A. Willis, J. Idoyaga, T.W. Muir, Streamlined expressed protein ligation using split inteins, J. Am. Chem. Soc. 135 (1) (2013) 286–292.
[31] K. Hiraga, I. Soga, J.T. Dansereau, B. Pereira, V. Derbyshire, Z. Du, C. Wang, P. Van Roey, G. Belfort, M. Belfort, Selection and structure of hyperactive inteins: peripheral changes relayed to the catalytic center, J. Mol. Biol. 393 (5) (2009) 1106–1117.
[32] P. Van Roey, B. Pereira, Z. Li, K. Hiraga, M. Belfort, V. Derbyshire, Crystallographic and mutational studies of Mycobacterium tuberculosis recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue, J. Mol. Biol. 367 (1) (2007) 162–173.
[33] A.S. Zwarycz, M. Fossat, O. Akanyeti, Z. Lin, D.J. Rosenman, A.E. Garcia, C.A. Royer, K.V. Mills, C. Wang, V67L mutation fills an internal cavity to stabilize RecA mtu intein, Biochemistry 56 (21) (2017) 2715–2722.
[34] N.H. Shah, G.P. Dann, M. Vila-Perello, Z.H. Liu, T.W. Muir, Ultrafast protein splicing is common among cyanobacterial Split inteins: implications for protein engineering, J. Am. Chem. Soc. 134 (28) (2012) 11338–11341.
[35] S.W. Lockless, T.W. Muir, Traceless protein splicing utilizing evolved split inteins, Proc. Natl. Acad. Sci. U. S. A. 106 (27) (2009) 10999–11004.