ISSN 0006-2979, Biochemistry (Moscow), 2025, Vol. 90, No. 6, pp. 818-827 © Pleiades Publishing, Ltd., 2025.
Russian Text © The Author(s), 2025, published in Biokhimiya, 2025, Vol. 90, No. 6, pp. 884-895.

Designing a Thermostable Mini-Intein for Intein-Mediated Purification of Recombinant Proteins and Peptides

Andrey A. Karanov[1,2,a]*, Evgeniy A. Zayats[1], Maria A. Kostromina[1], Yulia A. Abramchik[1], Aleksandra R. Sharafutdinova[1], Maria S. Surkova[1], Andrey A. Zamyatnin, Jr.[2], and Roman S. Esipov[1,2,b]

[1]*Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russia*
[2]*Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 119234 Moscow, Russia*
[a]*e-mail: andrey-karanov2000@mail.ru* [b]*e-mail: refolding@mail.ru*

Received February 7, 2025
Revised May 23, 2025
Accepted May 23, 2025

Abstract—This paper reports the design of a thermostable temperature-activated mini-intein based on the full-length intein DnaE1 from *Thermus thermophilus* HB27 (*Tth*DnaE1). We performed rational design of three mini-inteins *Tth*DnaE1 Δ272, Δ280, and Δ287 through deletion mutations in the full-length intein sequence. Two mini-inteins (Δ272 and Δ280) were capable of efficient protein splicing at temperatures above 50°C. The most active mini-intein with the Δ280 deletion was selected as a platform for further design of a self-cleaving carrier of affinity tags through single-point mutagenesis. Three mutations – C1A, D405G, and the combined C1A/D405G – were introduced to inhibit N-terminal extein cleavage and extein ligation. As a result, the mini-intein Δ280 with double mutation C1A/D405G displayed the highest efficiency of C-terminal extein cleavage with temperature optimum around 60°C. Thus, we constructed thermostable temperature-activated mini-intein capable of efficient protein splicing or cleavage of the C-terminal extein. The engineered *Tth*DnaE1 Δ280 C1A/D405G mini-intein can serve as a basis for the development of new expression system for intein-mediated production of pharmaceutically relevant recombinant proteins and peptides.

DOI: 10.1134/S0006297925600358

Keywords: recombinant proteins and peptides, mini-inteins, endonuclease domain, thermostability

INTRODUCTION

Development of highly efficient methods for producing pharmaceutically relevant recombinant peptides and proteins is a key challenge of medical biotechnology. Production of these compounds in prokaryotic expression system meets several difficulties. First, translation in prokaryotes introduces an N-terminal formylmethionine residue, which often needs to be removed to obtain a biologically active molecule [1]. Second, the target product may undergo intracellular proteolysis *in vivo*. This phenomenon is often observed during peptide production [2].

One of the solutions to overcome these issues involves the production of target peptides as a structural part of hybrid proteins followed by their release through cleavage [3]. Conventionally cleavage of hybrid proteins is carried out using site-specific proteases such as TEV protease, SUMO protease, enterokinase, or factor Xa. However, this approach has some drawbacks: the need to use costly enzymes, limited efficiency of cleavage, laborious optimization of the proteolysis conditions, and the need to separate the unnecessary part of the cleaved hybrid protein [3].

One of the promising alternatives is the construction of self-cleaving hybrid proteins based on inteins [4, 5]. In this approach the phenomenon of protein splicing is utilized. Protein splicing is a post

* To whom correspondence should be addressed.

translational process involving autocatalytic excision of an internal part of the precursor protein (intein) with simultaneous ligation of the flanking regions (exteins) [6]. Point mutations may block extein cleavage and ligation resulting in mutant forms of inteins that function as autocatalytically cleaving proteins carrying affinity tags. This approach proved beneficial. First, it allows to simplify the isolation and purification of the hybrid protein. Second, it reduces costs by eliminating the need for using proteases.

Despite these advantages, intein-mediated isolation methods face limitations. A common issue is premature intracellular cleavage of the hybrid protein [7-9]. There is also a risk that the efficiency of hybrid protein cleavage may be insufficient [10]. Isoelectric point of a hybrid protein may coincide with the pH range optimal for intein activity thus initiating irreversible aggregation [11]. To address these limitations, it could be useful to expand the repertoire of expression constructs based on new inteins.

Conditionally-controlled protein splicing/cleavage of exteins has been a focal point in recent intein research. Of special interest in this respect are those inteins that are thermoactivated and thermostable. Enzymatic activity of these inteins could be controlled by temperature. This approach includes the synthesis of an inactive intein-containing fusion protein and further enzyme activation via structure restoration. For instance, Shen et al. used thermostable and thermoactivated intein to regulate the activity of a xylanase for lignocellulose biomass processing [12]. Wang et al. employed a thermoregulated intein to control the activity of a thermostable DNA polymerase from *Pseudomonas fluorescens*, potentially enhancing PCR techniques [13].

In this study we suggest a new application of thermal activation and thermostability in inteins for use in purification of recombinant proteins and peptides. High activation temperature of intein may prevent premature *in vivo* cleavage. Furthermore, thermal stability enables initial purification via temperature-mediated precipitation of ballast proteins.

It is necessary to consider the domain organisation of inteins during the development of intein-based biotechnological tools. Inteins are generally classified as either full-length or mini-inteins. Full-length inteins consist of two domains. The HINT-domain (Hedgehog-Intein) responsible for protein splicing is formed by the intein's N- and C-terminal regions (designated HgN and HgC). The second domain with endonuclease activity is located in the middle part of intein, thus splitting the HINT-domain in two parts. The structure of mini-inteins includes only the HINT-domain, while the endonuclease domain is absent. Mini-inteins can be either natural or artificial. Small size of mini-inteins is advantageous, as well as lack of endonuclease domain, which may negatively affect solubility of the hybrid protein. Artificial mini-inteins derived from the DnaB of *Synechocystis* sp. and RecA of *Mycobacterium tuberculosis* have been shown to be the most useful tools for isolation and purification of proteins and peptides [14-19].

This study aims at the development of expression system for isolation of pharmaceutically relevant peptides and proteins based on the thermoactivated and thermostable DnaE1 intein of *Thermus thermophilus* (*Tth*DnaE1). Shen et al. have shown that *Tth*DnaE1 is a thermostable full-length intein capable of performing protein splicing at 60°C [12]. The goal of this study is the rational design of modified mini-inteins based on *Tth*DnaE1 through deletion of endonuclease domain and introducing point mutations at catalytic residues, as well as evaluation of their practical applicability.

MATERIALS AND METHODS

Generation of strains producing recombinant mini-inteins. To construct genes of artificial mini-inteins, two fragments of the *dna*E gene were PCR-amplified using chromosomal DNA of *Thermus thermophilus* strain HB27 as a template. One fragment encoded N-terminal (upstream of the deletion) region of the intein. The other fragment encoded C-terminal sequence downstream of the deletion. Primers for the N-terminal fragment included recognition sites for the restriction endonucleases *Nde*I and *Esp*3I, while primers for the C-terminal fragment included *Esp*3I and *Bpu*1102I sites. Amplification was carried out using Q5 High-Fidelity DNA Polymerase (New England Biolabs, USA). Products of amplification were digested with *Esp*3I restriction endonuclease and ligated with the help of a bacteriophage T4 ligase (Thermo Fisher Scientific, USA). The ligated fragment was amplified and cloned into a pET16b vector at recognition sites of the restriction endonucleases *Nde*I and *Bpu*1102I. The resulting sequences of mini-inteins were verified by Sanger sequencing (Evrogen, Russia).

Producer strains were obtained by transforming engineered plasmids into *E. coli* NiCo21 (DE3) cells (New England Biolabs, USA). Cultivation was performed in LB medium (10 g/l trypton, 5 g/l yeast extract, 10 g/l NaCl) supplemented with 100 μg/ml ampicillin. Cell cultures were grown in a thermostated shaker Certomat SII (Sartorius Stedim Biotech, Germany) at 37°C and 180 rpm until optical density OD_{600} 0.8 followed by addition of IPTG to final concentration 0.4 mM and further incubation for 4 h at 37°C. All target proteins were expressed in a soluble form.

Isolation and purification of recombinant mini-inteins. Cell biomass was resuspended at a ratio 1 : 10 (*w/v*) in buffer A (50 mM Tris-HCl, 200 mM NaCl,

5 mM EDTA (Merck Millipore, Germany), 1 mM PMSF (Sigma-Aldrich, USA), pH 8.0) and disrupted using an ultrasonic disintegrator VCX 130 (Sonics Materials, USA). Ammonium sulfate (Panreac, Spain) was added to the cell supernatant to 20% saturation, followed by centrifugation for 20 min at 10,000*g*. Precipitate containing ballast proteins was discarded. To precipitate the target proteins, ammonium sulfate was added to the prepared supernatant to 40% saturation. The resulting pellet was resuspended in buffer B (50 mM Tris-HCl, 200 mM NaCl, pH 8.0) and clarified by centrifugation. The clarified solution was loaded onto a XK16/20 column packed with Chelating Sepharose Fast Flow resin (Cytiva, USA). The column was washed with buffer C (50 mM Tris-HCl, 200 mM NaCl, 50 mM imidazole, pH 8.0) to elute non-target proteins. In the second stage a target protein was eluted with buffer D (50 mM Tris-HCl, 200 mM NaCl, 250 mM imidazole, pH 8.0). EDTA was added to the eluted fraction to final concentration 5 mM.

Fractions containing the target protein were pooled and concentrated to 5 ± 1 mg/ml using a YM-10 membrane (Merck Millipore). The final purification step involved gel-filtration chromatography on a HiLoad 16/60 Superdex 75 pg column (Cytiva) in buffer E (1 mM NaH_2PO_4, 20 mM Na_2HPO_4, 200 mM NaCl, 5 mM EDTA, pH 8.0). Fractions containing a target protein were combined.

Evaluation of mini-intein protein splicing efficiency. Efficiency of protein splicing was investigated in a pH range from 6.0 to 9.0 at 60°C. For this purpose, purified mini-intein samples were diluted with buffer E to final concentration of 0.3 mg/ml, and pH was adjusted. Temperature dependence of splicing was evaluated at pH 6.0 after 14-h incubation in the temperature range from 20 to 80°C. Splicing efficiency was assessed by electrophoretic analysis.

Analytical methods. Protein concentrations were determined using the Bradford protein assay [20]. Electrophoretic analysis was carried out according to the Laemmli technique in a 15% SDS-PAGE [21]. Densitometric analysis of gels was performed in three replicates to calculate average splicing efficiencies. To obtain electropherograms with enhanced resolution in the peptide range of molecular mass, a Tris-tricine buffer system was used for electrophoresis in a 10% SDS-PAGE [22].

RESULTS

Rational design of thermostable mini-inteins. The key point in rational design of mini-inteins is identification of a boundary between the endonuclease domain and HINT-domain in the full-length intein. The structure of *Tth*DnaE1 [including natural exteins each containing 5 amino acid residues (aa)] was modeled using AlphaFold 3. Superposition of the obtained model with the structure of the artificial mini-intein *Ssp*DnaB (PDB ID: 1MI8) enabled identification

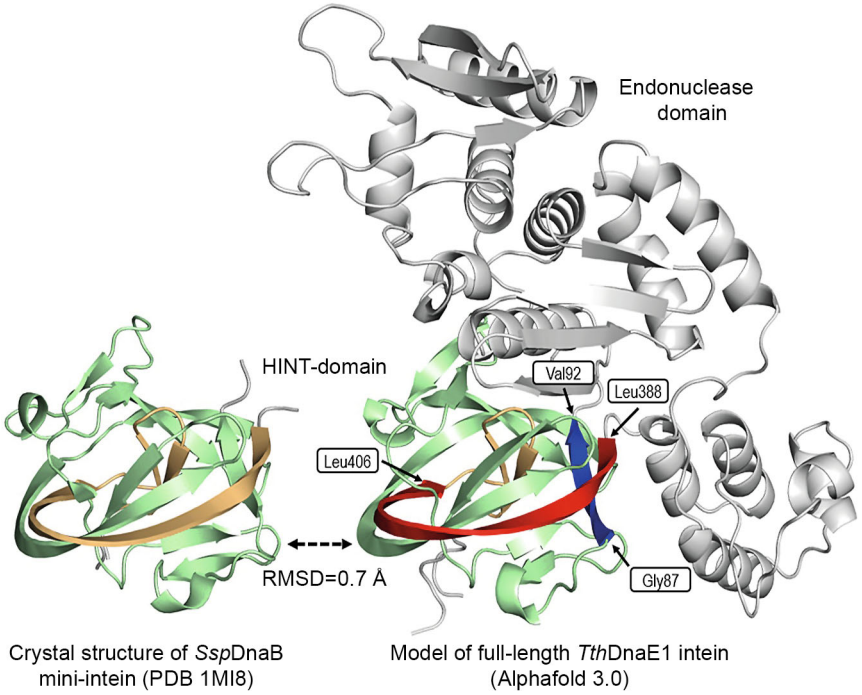

Fig. 1. Comparison of the structures of *Tth*DnaE1 and artificial mini-intein *Ssp*DnaB. N- and C-terminal regions of the polypeptide chain forming HINT-domain are shown in green and orange, respectively. β-Strands formed by amino acid residues G87-V92 and L388-L406 are shown in blue and red, respectively.

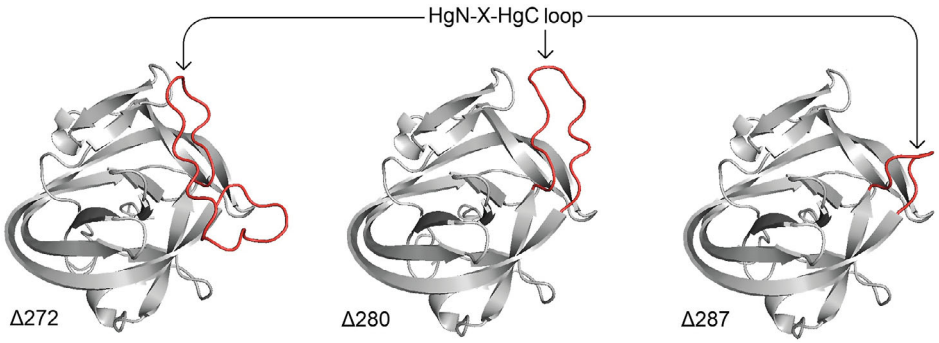

Fig. 2. Structure models of mini-inteins *Tth*DnaE1 Δ272 (a), Δ280 (b), and Δ287 (c).

of the HINT-domain in *Tth*DnaE1 (Fig. 1). We illustrated the similarity between the structure of the HINT-domain in the *Tth*DnaE1 model and crystal structure of *Ssp*DnaB: root mean square deviation of the coordinates of carbon atoms in the peptide backbone was around 0.7 Å. β-Strands formed by the amino acid residues G87-V92 and L388-L406 were the structural elements closest to the *Tth*DnaE1 endonuclease domain.

First, we analyzed the previously reported examples of rational design of mini-inteins. There are at least 11 examples of full-length inteins for which removal of the endonuclease domain without loss of splicing activity was shown: *Mtu*RecA, *Ssp*DnaB, *Npu*DnaB, *Rma*DnaB, *Ssp*DnaX, *Ssp*GyrB, *Ter*ThyX, *Ter*RIR1, *Sce*Vma, Pi-PFUI, and *Pho*VMA [15, 23-26]. One intein of particular interest was *Dna*E2 of *Trichodesmium erythraeum* (*Ter*DnaE2). Removal of the endonuclease domain from this intein resulted in the significant decrease in splicing efficiency [25, 27]. Nevertheless, this intein is interesting since it is an ortholog of *Tth*DnaE1.

Amino acid sequences of the *Tth*DnaE1 and of the abovementioned inteins were compared using multiple sequence alignment with the Clustal Omega (Fig. S1 in the Online Resource 1) [28]. The obtained alignment was used to generate a phylogenetic tree (Fig. S2 in in the Online Resource 1) using Mega program (UPGMA method) [29]. The inteins DnaB of *Synechocystis* sp. (*Ssp*DnaB), *Nostoc punctiforme* (*Npu*DnaB), and *Rhodothermus marinus* (*Rma*DnaB) were shown to be the closest relatives to the investigated intein. Hence, in order to select deletions, it is most feasible to compare the intein investigated in our study with these inteins. It should be also mentioned that among those the *Ssp*DnaB intein is especially notable due to its wide application for intein-mediated isolation of proteins and peptides [6, 14-17].

We decided to select three deletions, which could theoretically result in formation of loops by the remaining residues of endonuclease domain with length difference of approximately 8 aa. We considered deletion variants that are homologous to those that resulted in active artificial mini-inteins derived from DnaB. Three deletions in *Tth*DnaE1 correspond to this criterium, which are termed according to the number of deleted residues: Δ272, Δ280, and Δ287. The mini-intein *Tth*DnaE1 obtained as a result of the Δ272 deletion includes N- and C-terminal fragments of the amino acid sequence homologous to the fragments of polypeptide chain of the *Ssp*DnaB mini-intein (Table S1 in in the Online Resource 1). The variants containing deletions Δ280 and Δ287 include the fragments homologous to the N- and C-terminal fragments of the two artificial mini-inteins *Npu*DnaB (Table S1 in the Online Resource 1). The structures of the suggested mini-inteins [with natural exteins each containing 5 amino acid residues (aa)] were generated with AlphaFold 3.0 [30]. It was shown that the structure of HINT-domain in all three mini-inteins is identical, while the remaining residues of the endonuclease domain form disordered loop HgN-X-HgC with size of 23, 15, and 8 aa in the *Tth*DnaE1 Δ272, Δ280, and Δ287, respectively (Fig. 2; see also domain scheme in Fig. S3 in the Online Resource 1).

Structural stability of the HINT-domain of the artificial mini-inteins was investigated *in silico* through molecular dynamics simulations (MD) with Gromacs 2024 and the Amber ff99SB-ILDN force field (detailed description of the used MD methods is presented in the Online Resource 1) [31, 32]. Each model was placed into rhombic dodecahedron unit cells filled with tip3p water molecules and Na^+ and Cl^- ions to neutralize charge. After equilibration at 60°C and pressure 1 Bar by modeling in the NVT- and NPT-ensembles, productive MD simulation with duration of 500 ns was performed.

To examine stability of the structures of the mini inteins during MD simulation we calculated Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) of the atomic positions of the peptide backbone and the radius of gyration of the protein molecules. It should be mentioned that structure instability is expected for the exteins and loops

connecting HgN and HgC. In this regard, the RMSD values and radius of gyration were calculated not only for the entire protein molecule, but also for the HINT-domain separately. The obtained results are presented in Fig. S4 in the Online Resource 1.

Analysis of the RMSD values and radius of gyration indicated high stability of the HINT-domain in all three mini-inteins in the course of MD simulation. The profile of RMSF values indicates that exteins and HgN-X-HgC loop are the sites of instability in the polypeptide chain. Furthermore, the HgN-X-HgC loop in the mini-intein Δ280 was significantly more dynamic in comparison with two other variants.

One other factor that should be considered in the design of mini-inteins is the effect of the HgN-X-HgC loop on the protein solubility. This loop can include fragments of the polypeptide chain that forms the core of the endonuclease domain globule in the full-length intein. In this case hydrophobicity of amino acid residues of the unordered loop could negatively affect the protein solubility.

Therefore, we analyzed the potential effects of the sequence and spatial structure of the designed mini-inteins on solubility of the proteins using CamSol 2.0 web server [33]. The approach of the prediction of such effects of the fragments of polypeptide chain is based on: 1) analysis of amino acid sequence based on physicochemical properties of amino acid residues; 2) correction of the obtained results considering protein spatial structure. Structure models of the suggested mini-inteins were analyzed using CamSol 2.0. It was shown that the loops HgN-X-HgC do not contain amino acid residues which could affect solubility negatively (Fig. S5 in the Online Resource 1, a-c).

Besides the design of mini-inteins, another important challenge is the selection of the lengths of the extein fragments extein for further investigation of protein splicing *in vitro*. Intein activity could be significantly affected by the amino acid sequence of an extein. That is why it is reasonable to investigate protein splicing with natural exteins. On the other hand, use of inteins for protein purification is especially relevant in the case of peptides with unordered structure. Therefore, we decided to investigate protein splicing by using hybrid constructs containing mini-inteins with fragments of natural exteins up to 20 aa in length.

Certain fragments of natural exteins could participate in formation of the globule of natural DnaE protein. Presence of hydrophobic sites may negatively affect solubility. To resolve this issue the size of extein fragments was optimized with the help CamSol 2.0 web server. Structure model of the full-length intein with fragments of natural exteins with length of 20 aa was obtained by using of neural network AlphaFold 3.0. Analysis of the obtained model with the help of CamSol 2.0 showed that the exteins larger than 10 aa at the N-end could significantly reduce the protein solubility (Fig. S5 in the Online Resource 1, d).

The final step of the design was selection of point mutations in order to change the intein capability of protein splicing. The sequence of *Tth*DnaE1 was compared with the sequence of *Mtu*RecA intein, for which key amino acids participating in protein splicing have been determined (based on the previously performed alignment of amino acid sequences) [23]. This allowed to identify homologous residues in *Tth*DnaE1: Cys1, Asp405, His422, and Asn423. Positions of these residues in the model structure of *Tth*DnaE1 Δ280 intein are shown in Fig. S6 in in the Online Resource 1.

Protein splicing occurs due to rearrangement with participation of hydroxyl/thiol groups of the first amino acid of intein and the first amino acid of the C-end of extein (in the case of *Tth*DnaE1 these are Cys1 and Ser+1). The Cys1Ala mutation disables cleavage of the N-terminal extein. It was shown that the highly conserved Asp residue is significant for protein splicing (Asp405 in *Tth*DnaE1) [23]. Substitution of this Asp residue with Gly was shown to significantly accelerate the cleavage of C-terminal extein. Hence, the D405G mutation is practical for the *Tth*DnaE1 both as a single mutation and in combination with C1A. Two last residues in the intein (His422, Asn423 in *Tth*DnaE1) participate in the cleavage of the C-terminal extein. Mutations in these positions are not practical because cleavage of the C-terminal extein is most significant for the development of the approach to intein-mediated isolation and purification of proteins.

Hence, modeling of the protein spatial structure with the help AlphaFold 3.0, MD simulations in Gromacs, and predicting effects of amino acid on the protein solubility allowed us to perform rational design of three mini-inteins: *Tth*DnaE1 Δ272, Δ280, and Δ287. Analysis of functional amino acids in the intein allowed us to select the C1A, D405G, and C1A/D405G mutations that prevent cleavage of the N-terminal extein.

Preparation and investigation of properties of artificial mini-inteins. Expression constructs encoding mini-inteins *Tth*DnaE1 Δ272, Δ280, and Δ287 flanked with the fragments of natural exteins (N-extein with length 10 aa and C-extein with length 16 aa) and N-terminal affinity tag 6x-His were created, and the corresponding *E. coli* producer strains were generated. Recombinant proteins were isolated using a two-stage protocol including purification via affinity and gel-filtration chromatography.

Temperature and pH dependence of splicing efficiency (range 20-80°C at pH 6.0; range 6.0-9.0 at 60°C, respectively) was evaluated based on the results of incubation of the reaction mixtures for 14 h. Four possible products were expected (in accordance with

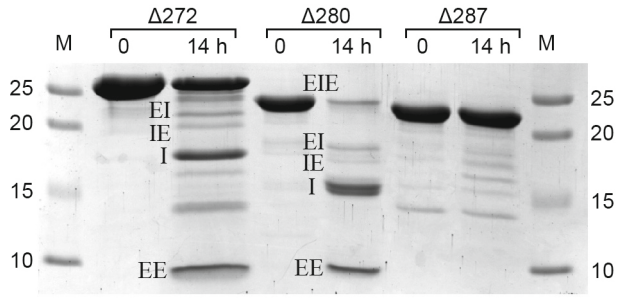

Fig. 3. Electrophoretic analysis of the products of splicing of hybrid proteins containing mini-intein *Tth*DnaE1 with Δ272, Δ280, and Δ287 deletions (15% SDS-PAAG). Designations: EIE, non-cleaved mini-intein; EI, residual protein without C-terminal exteins; IE, residual protein without N-terminal extein; I, mini-intein without exteins; EE, product of ligation of exteins.

the calculated molecular mass, from smallest to largest): 1) product of extein ligation (designated as EE; 6.2 kDa); 2) mini-intein without exteins (product I; 16.4 kDa for Δ272, 15.5 kDa for Δ280); 3) residual protein without N- terminal extein (product IE; 18.3 kDa for Δ272, 17.5 kDa for Δ280); 4) residual protein without C-terminal extein (product EI; 20.5 kDa for Δ272, 19.7 kDa for Δ280).

The results of electrophoretic analysis of the mini-inteins splicing products after incubation at 60°C for 14 h at pH 6.0 are presented in Fig. 3. Cleavage of the proteins with deletion Δ272 and Δ280 (EIE bands) resulted in formation of side products in addition to the main products (EE and I).

The *Tth*DnaE1 Δ280 intein was found to be the most active: almost complete cleavage of the corresponding hybrid protein was observed after incubation. The Δ272 deletion variant exhibited significantly lower activity, accompanied by high level of proteolysis of both hybrid protein and of the cleavage products observed. The variant Δ287 showed no detectable splicing activity.

These results confirm the success of the rational design approach, yielding a mini-intein variant (Δ280) capable of efficient, temperature-dependent protein splicing. This construct was therefore selected for further development of the self-cleaving carrier protein.

Next, site-directed mutagenesis of the gene of mini-intein *Tth*DnaE1 Δ280 was performed to introduce point mutations C1A, D405G, and C1A/D405G. The resulting hybrid proteins were expressed using the same method as for wild type constructs. Splicing activity of the modified mini-inteins was examined according to the abovementioned protocol.

Comparison of the products of cleavage of the hybrid proteins based on the generated mini-intein Δ280 and on its mutant variants C1A, D405G, and C1A/D405G is presented in Fig. 4. Introduction of all three mutations resulted in significant changes in composition and relative abundance of the formed splicing products.

Splicing efficiency in all three mutant constructs was strongly dependent on both temperature and pH (Fig. 5). Optimal range of pH (6.0-6.5) for cleaving was determined for all tested variants.

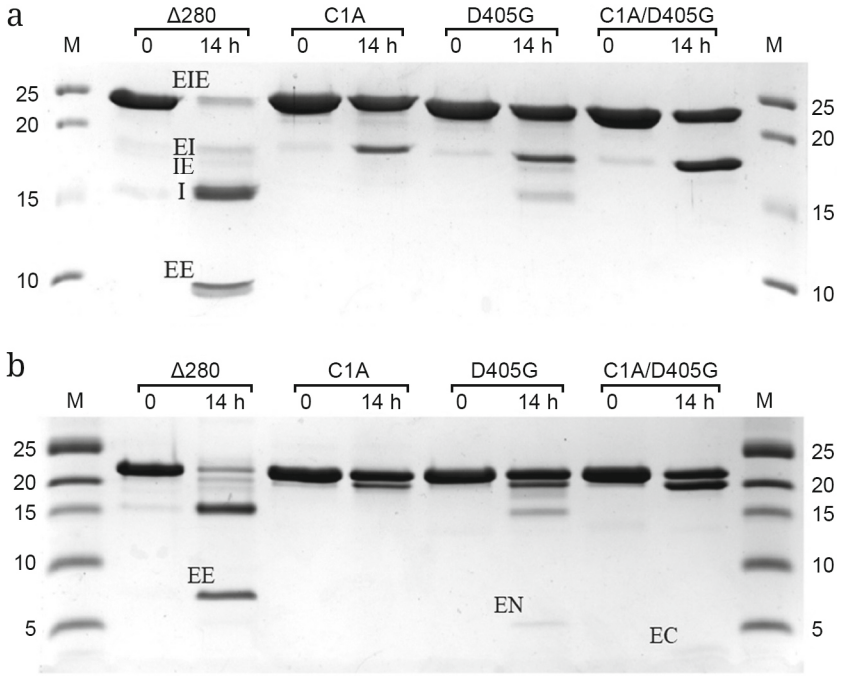

Fig. 4. Electrophoretic analysis of the products of splicing of hybrid proteins based on the mini-intein *Tth*DnaE1 Δ280 without mutations and with point substitutions C1A, D405G, and C1A/D405G. a) 15% SDS-PAAG; b) 10% tricine SDS-PAAG. Designations are the same as in Fig. 3; EN, N-terminal extein; EC, C-terminal extein.

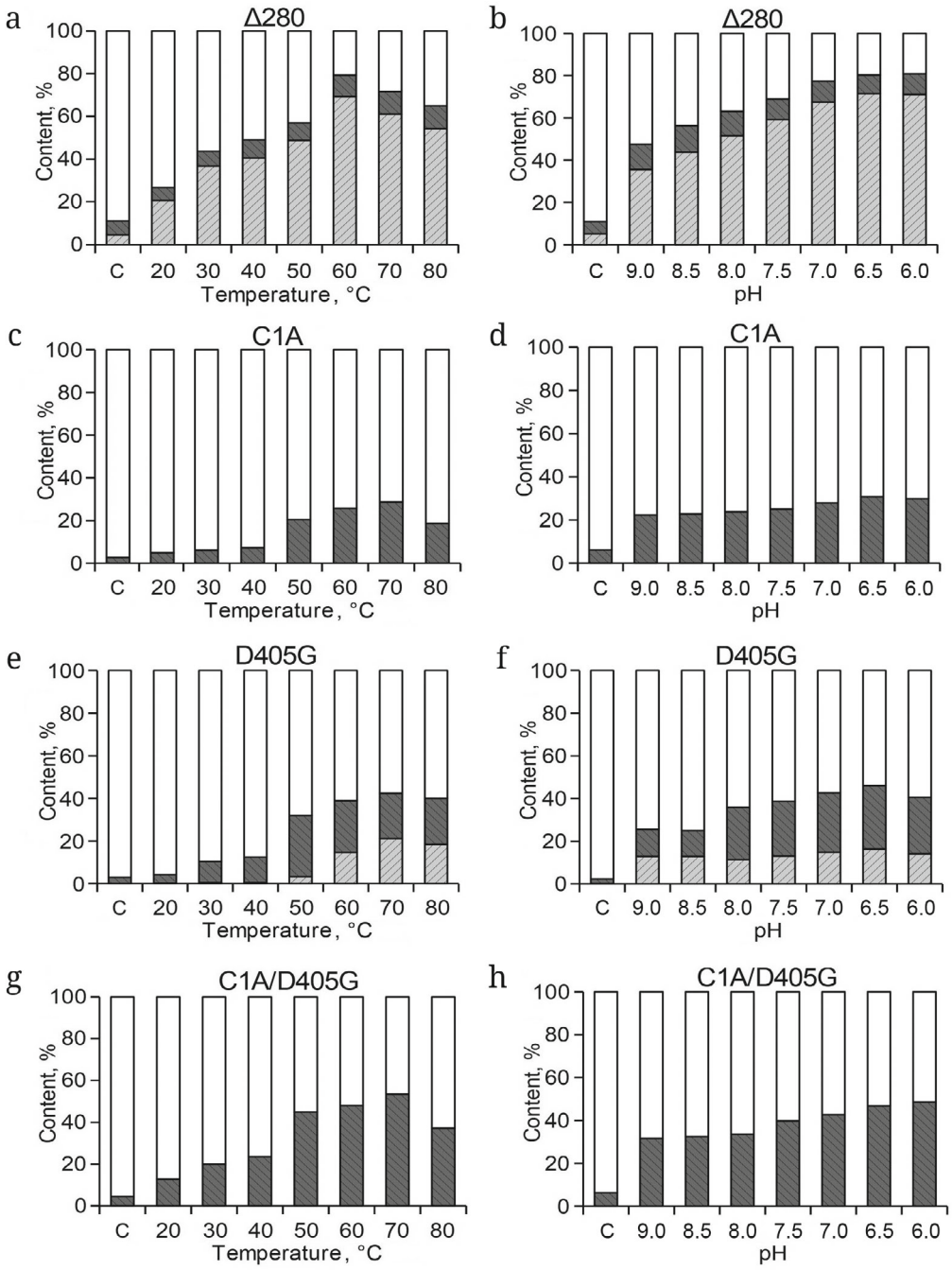

Fig. 5. Products of cleavage of the hybrid proteins based on the *Tth*DnaE1 Δ280 mini-intein (a, b) and its mutant forms C1A (c, d), D405G (e, f), and C1A/D405G (g, h) depending on the incubation conditions. a, c, e, g) Content of products formed at pH 6.0 and temperature in the range 20-80°C; b, d, f, h) at 60°C and pH in the range 6.0-9.0. Initial hybrid protein marked with white color, product I – with light gray, EI product – with dark gray. The data are obtained using densitometry of the 15% SDS+PAAGs. C, control, hybrid protein before incubation.

Increasing the incubation temperature enhanced efficiency of cleavage significantly: hybrid protein based on the mini-intein *Tth*DnaE1 Δ280 exhibited 79% (at 60°C and pH 6.0) versus 27% (at 20°C). Similar results were obtained for the mutant forms. When incubated at 70°C and pH 6.0-6.5, the cleavage efficiency increased notably: from 4% to 31% for the C1A mutant, from 5% to 42% for the D405G mutant, from 11% to 54% for the C1A/D405G mutant. Optimal incubation time of 14 h was selected based on the data of the cleavage kinetics. After 14-h incubation accumulation of the products slows down; moreover, prolonged incubation may lead to proteolytic degradation of the products (Fig. S7 in the Online Resource 1).

All expected cleavage products were detected for protein carrying D405G mutation: cleavage of both exteins and cleavage of C- or N-terminal extein separately (product EN, 4.2 kDa; formation of the EC product was detected while increasing the gel load). At the same time no extein ligation was observed. Cleavage of the C- and N-terminal exteins was confirmed by electrophoretic analysis in the Tris-tricine buffer system (Fig. 5b). Hence, the D405G mutation significantly increases efficiency of cleavage of the C-terminal extein and blocks ligation of exteins.

Cleavage of only C-terminal extein (yielding products EI and EC) was observed for C1A and C1A/D405G variants. Hence, this mutation resulted in the desirable effect – prevention of cleavage of the N-terminal extein. Notably, the hybrid protein with double mutation exhibited the highest cleavage efficiency.

DISCUSSION

In this study, we successfully performed the rational design of the thermoactivated mini-intein based on the full-length DnaE1 intein from *Thermus thermoplilus*.

An approach that includes amino acid sequence analysis, structure modeling with AlphaFold 3.0 and MD simulations allowed us to select three deletion variants (Δ272, Δ280, and Δ287) for removal of the endonuclease domain from the initial *Tth*DnaE1 intein. Experimental investigation of the engineered mini-inteins showed that the medium-sized deletion was the least disruptive for protein splicing – the *Tth*DnaE1 Δ280 mini-intein exhibited the highest activity. The reasons behind the reduced activity of the Δ272 and Δ287 mini-inteins require further investigation. It should be mentioned that AlphaFold structure modeling and MD simulation in Gromacs are not *ab initio* approaches. It is possible that the negative effects of the Δ272 and Δ287 deletions could be explained by their effects exactly on the processes of folding of mini-inteins, which are challenging to predict.

Next, we have investigated the effects of point mutations at the key positions responsible for protein splicing. Two targets were selected for mutagenesis: Cys1 and Asp405. The ability of cleavage of the N-terminal extein was blocked by the C1A mutation. The single D405G mutation resulted in prevention of ligation of the cleaved exteins. The obtained results indicate significance of the Cys1 and Asp405 residues for different stages of protein splicing.

The capability of C-terminal extein cleavage is valuable for the intein-mediated isolation of proteins and peptides. All other activities of intein may result in formation of non-desired side products. The C1A/D405G mini-intein variant demonstrated the desirable result: the only product was C-terminal extein while side products formation was not observed. Hence, it was shown that the mini-intein *Tth*DnaE1 Δ280 C1A/D405G is capable of effective cleavage of the C-terminal extein, and constitutes a promising object for further studies.

It was shown that temperature optimum of the mini-intein engineered in this study is in the range of 60-70°C. Therefore, the efficiency of the cleavage of the C-terminal extein is reduced at lower temperatures (20-40°C), which corresponds to the standard cultivation temperatures of the *E. coli* producer strains. This feature could be useful to overcome the premature intracellular cleavage of hybrid proteins.

The design of an expression system based on the *Tth*DnaE1 Δ280 C1A/D405G mini-intein and demonstration of its applicability for production of a number of model peptides seems to be a logical next step of the study.

CONCLUSIONS

Development of new approaches for the isolation and purification of recombinant proteins and peptides is one of the key challenges of biotechnology. Construction of hybrid proteins that include the target peptide sequence provides a solution for many technological problems characteristic to prokaryotic systems such as presence of N-terminal formylmethionine or intracellular degradation. Inteins were successfully proven themselves to be useful as affinity tag carriers and simultaneously autocatalytic proteases. Design of novel inteins exhibiting new methods of control of protein splicing and cleavage of exteins helps to solve the key challenges of this technology.

In this study we engineered a thermostable and thermoactivated mini-inteins based on the full-length DnaE1 intein from *Thermus thermophilus* HB27. Properties of such inteins have some significant advantages for their use in intein-mediated protein purification. High temperature optimum for the cleavage of exteins could prevent cleavage of this hybrid protein *in vivo* at temperatures optimal for cultivation of mesophilic producer strains. In future studies we plan to develop an expression system using the *Tth*DnaE1 Δ280 mini-intein with double mutation C1A/D405G and test this system for the production of pharmaceutically significant peptides.

Abbreviations. MD, molecular dynamics.

Supplementary information. The online version contains supplementary material available at https://doi.org/10.1134/S0006297925600358.

Contributions. R. S. Esipov and A. A. Zamyatin – concept and supervision of the study; A. A. Karanov,

E. A. Zayats, A. R. Sharafutdinova, and M. S. Surkova – conducting experiments; A. A. Karanov, E. A. Zayats, M. A. Kostromina, and Yu. A. Abramchik – discussion of the results of the study; A. A. Karanov and E. A. Zayats – writing text of the paper; R. S. Esipov, M. A. Kostromina, and Yu. A. Abramchik – editing text of the paper.

Ethics approval and consent to participate. This work does not contain any studies involving human and animal subjects.

Conflict of interest. The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Ben-Bassat, A. (1991) *Purification and Analysis of Recombinant Proteins* (Seetharam, R., and Sharma, S. K., eds), Marcel Dekker, N.Y., pp. 147-151.
2. Enfors, S. O. (1992) Control of *in vivo* proteolysis in the production of recombinant proteins, *Trends Biotechnol*, **10**, 310-315, https://doi.org/10.1016/0167-7799(92)90256-u.
3. Mahmoodi, S., Pourhassan-Moghaddam, M., Wood, D. W., Majdi, H., and Zarghami, N. (2019) Current affinity approaches for purification of recombinant proteins, *Cogent Biol.*, **5**, 1665406, https://doi.org/10.1080/23312025.2019.1665406.
4. Eskandari, A., Leow, T. C., Rahman, M. B. A., and Oslan, S. N. (2024) Utilization and prospect of purification technologies in natural proteins, peptides and recombinant proteins, *J. Proteins Proteom.*, **15**, 233-257, https://doi.org/10.1007/s42485-024-00139-7.
5. Esipov, R. S., Stepanenko, V. N., Chupova, L. A., Boyarskikh, U. A., Filipenko, M. L., and Miroshnikov, A. I. (2008) Production of recombinant human epidermal growth factor using *Ssp* dnaB mini-intein system, *Protein Expr. Purif.*, **61**, 1-6, https://doi.org/10.1016/j.pep.2008.05.009.
6. Volkmann, G., and Mootz, H. D. (2013) Recent progress in intein research: from mechanism to directed evolution and applications, *Cell Mol. Life Sci.*, **70**, 1185-1206, https://doi.org/10.1007/s00018-012-1120-4.
7. Yuan, H., Prabhala, S. V., Coolbaugh, M. J., Stimple, S. D., and Wood, D. W. (2024) Improved self-cleaving precipitation tags for efficient column free bioseparations, *Protein Expr. Purif.*, **224**, 106578, https://doi.org/10.1016/j.pep.2024.106578.
8. Lahiry, A., Fan, Y., Stimple, S. D., Raith, M., and Wood, D. W. (2018) Inteins as tools for tagless and traceless protein purification, *J. Chem. Technol. Biotechnol.*, **93**, 1827-1835, https://doi.org/10.1002/jctb.5415.
9. Cui, C., Zhao, W., Chen, J., Wang, J., and Li, Q. (2006) Elimination of *in vivo* cleavage between target protein and intein in the intein-mediated protein purification systems, *Protein Expr. Purif.*, **50**, 74-81, https://doi.org/10.1016/j.pep.2006.05.019.
10. Hiraga, K., Derbyshire, V., Dansereau, J. T., Van Roey, P., and Belfort, M. (2005) Minimization and stabilization of the *Mycobacterium tuberculosis* recA intein, *J. Mol. Biol.*, **354**, 916-926, https://doi.org/10.1016/j.jmb.2005.09.088.
11. Esipov, R. S., Beirakhova, K. A., Chupova, L. A., Likhvantseva, V. G., Stepanova, E. V., and Miroshnokov, A. I. (2012) Recombinant fragment of pigment epithelium-derived factor (44-77) prevents pathological corneal neovascularization, *Bioorg. Khim.*, **38**, 1-8.
12. Shen, B., Sun, X., Zuo, X., Shilling, T., Apgar, J., Ross, M., Bougri, O., Samoylov, V., Parker, M., Hancock, E., Lucero, H., Gray, B., Ekborg, N. A., Zhang, D., Johnson, J. C. S., Lazar, G., and Raab, R. M. (2012) Engineering a thermoregulated intein-modified xylanase into maize for consolidated lignocellulosic biomass processing, *Nat. Biotechnol.*, **30**, 1131-1136, https://doi.org/10.1038/nbt.2402.
13. Wang, Y., Shi, Y., Hellinga, H. W., and Beese, L. S. (2023) Thermally controlled intein splicing of engineered DNA polymerases provides a robust and generalizable solution for accurate and sensitive molecular diagnostics, *Nucleic Acids Res.*, **51**, 5883-5894, https://doi.org/10.1093/nar/gkad368.
14. Yan, S.-S., Yan, J., Shi, G., Xu, Q., Chen, S.-C., and Tian, Y.-W. (2005) Production of native protein by using *Synechocystis* sp. PCC6803 DnaB mini-intein in *Escherichia coli*, *Protein Expr. Purif.*, **40**, 340-345, https://doi.org/10.1016/j.pep.2004.12.021.
15. Wu, H., Xu, M.-Q., and Liu, X.-Q. (1998) Protein trans-splicing and functional mini-inteins of a cyanobacterial dnaB intein, *Biochim. Biophys. Acta*, **1387**, 422-432, https://doi.org/10.1016/s0167-4838(98)00157-5.
16. Malone, C. L., Boles, B. R., and Horswill, A. R. (2007) Biosynthesis of *Staphylococcus aureus* autoinducing peptides by using the *Synechocystis* DnaB mini-intein, *Appl. Environ. Microbiol.*, **73**, 6036-6044, https://doi.org/10.1128/AEM.00912-07.
17. Tian, L., and Sun, S. S. (2011) A cost-effective ELP-intein coupling system for recombinant protein purification from plant production platform, *PLoS One*, **6**, e24183, https://doi.org/10.1371/journal.pone.0024183.
18. Gangopadhhyay, J. P., Jiang, S.-Q., van Berkel, P., and Paulus, H. (2003) *In vitro* splicing of erythropoietin by the *Mycobacterium tuberculosis* RecA intein without substituting amino acids at the splice junctions, *Biochim. Biophys. Acta*, **1619**, 193-200, https://doi.org/10.1016/s0304-4165(02)00495-6.
19. Banki, M. R., Feng, L., and Wood, D. W. (2005) Simple bioseparations using self-cleaving elastin-like polypeptide tags, *Nat. Methods*, **2**, 659-661, https://doi.org/10.1038/nmeth787.

20. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.*, **72**, 248-254, https://doi.org/10.1016/0003-2697(76)90527-3.
21. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, **227**, 680-685, https://doi.org/10.1038/227680a0.
22. Schägger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem.*, **166**, 368-379, https://doi.org/10.1016/0003-2697(87)90587-2.
23. Van Roey, P., Pereira, B., Li, Z., Hiraga, K., Belfort, M., and Derbyshire, V. (2007) Crystallographic and mutational studies of *Mycobacterium tuberculosis* recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue, *J. Mol. Biol.*, **367**, 162-173, https://doi.org/10.1016/j.jmb.2006.12.050.
24. Aranko, A. S., Oeemig, J. S., Zhou, D., Kajander, T., Wlodawer, A., and Iwaï, H. (2014) Structure-based engineering and comparison of novel split inteins for protein ligation, *Mol. Biosyst.*, **10**, 1023-1034, https://doi.org/10.1039/c4mb00021h.
25. Lin, Y., Li, M., Song, H., Xu, L., Meng, Q., and Liu, X. Q. (2013) Protein trans-splicing of multiple atypical split inteins engineered from natural inteins, *PLoS One*, **8**, e59516, https://doi.org/10.1371/journal.pone.0059516.
26. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene*, **192**, 271-281, https://doi.org/10.1016/s0378-1119(97)00105-4.
27. Perler, F. B. (2002) InBase, the intein database, *Nucleic Acids Res.*, **30**, 383-384, https://doi.org/10.1093/nar/30.1.383.
28. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.*, **7**, 539, https://doi.org/10.1038/msb.2011.75.
29. Hall, B. G. (2013) Building phylogenetic trees from molecular data with MEGA, *Mol. Biol. Evol.*, **30**, 1229-1235, https://doi.org/10.1093/molbev/mst012.
30. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C. C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, **630**, 493-500, https://doi.org/10.1038/s41586-024-07487-w.
31. Van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free, *J. Comput. Chem.*, **26**, 1701-1718, https://doi.org/10.1002/jcc.20291.
32. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field, *Proteins*, **78**, 1950-1958, https://doi.org/10.1002/prot.22711.
33. Sormanni, P., Aprile, F. A., and Vendruscolo, M. (2015) The CamSol method of rational design of protein mutants with enhanced solubility, *J. Mol. Biol.*, **427**, 478-490, https://doi.org/10.1016/j.jmb.2014.09.026.

Publisher's Note. Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. AI tools may have been used in the translation or editing of this article.