Supplementary Information

Creation of A Caspese-3 Sensing System Using A Combination of Split-GFP and Split-Intein

Seiji Sakamoto,* Mika Terauchi, Anna Hugo, Tanner Kim, Yasuyuki Araki and Takehiko Wada*
Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan
e-mail: sakamoto@tagen.tohoku.ac.jp; hiko@tagen.tohoku.ac.jp

I. Experimental Section

General Materials. All chemicals and solvents were of reagent or HPLC grade. All materials were obtained from Sigma-Aldrich unless otherwise noted. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. KOD Plus Neo2 DNA polymerase and dNTPs were obtained from Toyobo (Tokyo, Japan). Isopropyl-*β*-D(-)-thiogalactopyranoside (IPTG), DL-dithiothreitol (DTT), tris(2-carboxyethyl)phosphine (TCEP), and PURESYSTE classic II were purchased from Wako Pure Chemical Industries (Osaka, Japan). A caspase-3 inhibitor (Ac-DNLD-CHO) was obtained from Peptide Institute, Inc. (Osaka, Japan). Recombinant active caspase-3, and -9 were purchased from Medical & Biological Laboratories (Nagoya, Japan). One unit of the recombinant caspase-3 or -9 is defined as the amount of enzyme that cleaves 1 nmol of each caspase substrate [DEVD-pNA or LEHD-pNA (pNA: pnitroanaline)] per hour at 37 °C in 50 mM HEPES buffer (pH 7.2) containing 100 mM NaCl, 10 mM DTT, 10 mM EDTA, 0.1% CHAPS and 5% glycerol. MALDI-TOF MS was measured on a Bruker autoflex III mass spectrometer by using 3,5-dimethoxy-4-hydroxycinnamic acid as a matrix.

Plasmid Construction and Expression of GFPN. An *E. coli* strain XL1-BLUE was used as the bacterial host for the construction of all plasmids. A DNA sequence encoding a N-terminal fragment (1–214) of GFP OPT (GFPN) was prepared by PCR amplification from full-length GFP OPT as a

template using primers, 5'-GC ATA TCG GAT CCG AGT AAA GGA GAA GAA CTT TTC-3' and 5'- T CGA ATT CTC GAG TCA CTT TTC GTT GGG ATC TTT C -3'. The PCR product was ligated into pET22b(+) vector *via* the BamHI/XhoI restriction sites. The NdeI/BamHI site of the obtained plasmid was replaced with the DNA fragment ATG CAC CAT CAT CAT CAC CAT GGC TCT TCG coding a 6 × His tag (MHHHHHHGSS). Sequence of the construct was verified by dye-terminator sequencing.

The GFPN protein was expressed in BL21(DE3) at 25 °C overnight under the control of T7 promoter with 0.5 mM IPTG. The proteins were purified with Ni-NTA (Qiagen) and Sephadex G-25 (GE Healthcare) columns. The purity and MW were confirmed by SDS PAGE analysis (Figure S2). The DNA and amino acid sequences of GFPN are shown in Figure S3.

Plasmids Construction of DnaB Split-Inteins for Preparing Cyclic GFP C-Terminal Fragments, cM4(DEVD) and cM4(DEVG). The plasmid construct encoding DnaB split-intein was prepared by PCR using a pTWIN vector (New England BioLabs) as a template with appropriate primers. The PCR product encoding an intein C-terminal fragment (residues 106–154) was cloned into a modified pET22b vector (pET22b-6H-MCS-6H, Figure S4), which contains NsiI and EagI restriction sites in its MCS, *via* the NdeI/NsiI restriction sites. Subsequently, the PCR product coding an intein N-terminal fragment (1–105) was ligated *via* the EagI/XhoI sites to render pET22b-sDnaB. The DNA insert for cpM4(DEVD) was created with Klenow fragment DNA polymerase I using synthetic oligo DNAs, 5'- CAC AAC TCG GCC GCT GGT ATC ACC GGT **GAT GAG GTG GAC** GGT CGC GAT CAC ATG -3' and 5'- GCC AGA GAT GCA TAC ATA TTC GTG CAG AAC CAT GTG ATC GCG ACC GTC CAC CTC -3'. The reaction product was ligated into the pET22b-sDnaB plasmid *via* EagI/NsiI sites to render a pET22b-sDnaB-cpM4(DEVD) plasmid. The DNA sequence encoding DEVD in the pET22b-sDnaB-cpM4(DEVD) was replaced with one encoding DEVG by overlap-extension PCR using primers 5'- GAT GAG GTG GgC GGT CGC GAT CAC ATG -3' and 5'- GCG ACC GcC CAC CTC ATC ACC GGT GAT -3' to confer pET22b-sDnaB-cpM4(DEVG). The intein-disabled version of construct, pET22b-sDnaB-cpM4(DEVD)-T70A/H73A, was created by overlap extension PCR using primers 5'- GCA **GCA** GCA AAT **GCT** AGA TTT TTA ACT ATT

GAT GG -3' and 5'- TCT **AGC** ATT TGC **TGC** TGC CTT GAT AGT TCT ACC TA -3'. Sequences of constructs were verified by dye-terminator sequencing. The DNA and amino acid sequences of split-intein constructs are shown in Figure S5-S7.

SDS-PAGE Analysis of Split-Intein Fusions Expressed in *E. coli*. All split-intein fusions were expressed in BL21(DE3) at 25 °C overnight under the control of T7 promoter with 0.5 mM IPTG. The cells were pelleted by centrifugation, and frozen at -20 °C overnight. Cells were resuspended in a bind buffer [20mM Tris HCl (pH 7.9), 0.5 M NaCl, 5 mM imidazole, 0.5 mM TCEP] on ice bath and lysed using standard sonication protocols. The soluble fractions were loaded onto Ni-NTA agarose beads (Qiagen) pre-equilibrated with a bind buffer. After washing the column with 10 volumes of bind buffer and 6 volumes of wash buffer [20mM Tris HCl (pH 7.9), 0.5 M NaCl, 30 mM imidazole, 0.5 mM TCEP], the proteins were eluted with 6 volumes of elution buffer [20mM Tris HCl (pH 7.9), 0.5 M NaCl, 1 M imidazole, 0.5 mM TCEP]. Each eluted fraction was analyzed by SDS-PAGE to check the processing status of split-intein constructs.

Preparation and Isolation of Cyclic GFP C-terminal Fragments Using PURESYSTEM Classic II. In vitro transcription and translation of the intein construct was carried out using a pET22b-sDnaB-cpM4(DEVD), pET22b-sDnaB-cpM4(DEVD) or pET22b-sDnaB-cpM4(DEVD)-T70A/H73A plasmid (500 ng) in PURESYSTEM classic II (BioComber) in a total volume of 50 μL at 37 °C for 1 hr. The reaction mixture was diluted with 50 μL of water and incubated with 10 μL of a Ni-NTA agarose slurry (QIAGEN) at 4 °C for 3 hr to remove the His-tagged proteinous factors. The mixture was applied on an ultrafiltration membrane unit (Vivacon 500 100K, Sartorius stedim) and centrifuged at 4 °C and 1500 g for 1 hr for removal of ribosome. The flow-through was lyophilized and the residue was dissolved in 5% acetonitrile/water containing 0.1% trifluoroacetic acid. This was further cleaned using MonoTip C18 (GL Science) and analyzed by MALDI-TOF MS. cM4(DEVD), *m/z* 2354.39 $[(M+H)^+]$ (Calcd. = 2354.15); cM4(DEVG), *m/z* 2296.56 $[(M+H)^+]$ (Calcd. = 2296.51).

Caspase Sensing Using A Split-Intein Mediated Cyclic Peptide, cM4(DEVD). In vitro transcription and translation of cM4(DEVD) was carried out using a pET22b-sDnaB-cpM4(DEVD) plasmid (500 ng) in PURESYSTEM classic II in a total volume of 50 μL at 37°C for 1 hr. The 10 μL of reaction solution was mixed with a GFP N-terminal fragment (final protein concentration was 1 μM) in 50 mM HEPES buffer (pH 7.2) containing 100 mM NaCl, 10 mM DTT, 10 mM EDTA, 0.1% CHAPS and 5% glycerol with and without a caspase-3 inhibitor, Ac-DNLD-CHO. The proteolysis reaction was started by the addition of caspase-3 (1 u) in a total volume of 200 μL. The samples were subsequently incubated at 25 °C. Fluorescence spectra were measured on a JASCO FP-6500 fluorescence spectrophotometer at 25 °C. The concentration of cM4(DEVD) was estimated by comparing the maximum fluorescence intensity of the reconstituted split-GFP with that of an original split-GFP sample of known concentration.

II. Supporting Figures

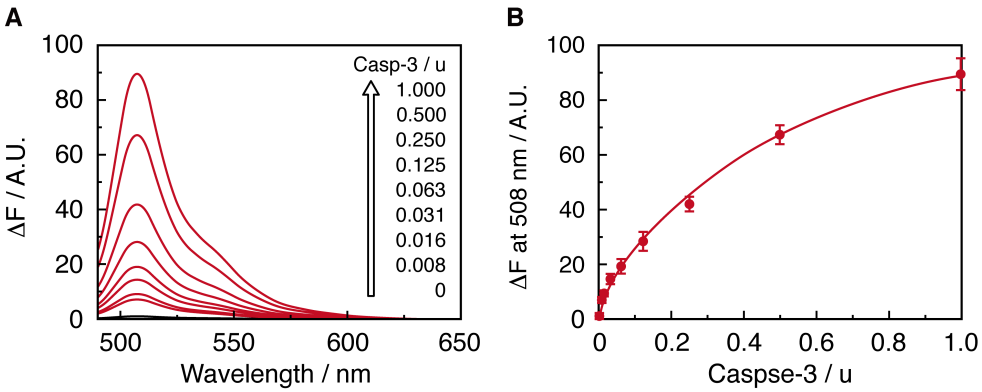

Figure S1. (A) Fluorescence spectra of the GFPN/cM4(DEVD) mixture after 6 h incubation at 25 °C with various concentrations of caspase-3. (B) A plot of fluorescence intensity change (ΔF) at 508 nm of the GFPN/cM4(DEVD) mixture as a function of caspase-3 concentration. The reconstitution reaction and the fluorescence measurement were performed as in Figure 3. $\Delta F = F - F_0$, where F and F_0 denote fluorescent intensities after and before incubation, respectively. Arbitrary fluorescence units (A.U.).

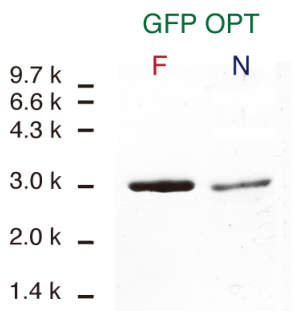

Figure S2. SDS-PAGE of the purified proteins. 15% Polyacrylamide gel was used for the analysis. A gel was stained with Coomassie Brilliant Blue.

pET22b-6H-GFPN

```
          NdeI
             1/1                                        BamHI
cat atg CAC CAT CAT CAT CAc cat ggC TCT TCg gat ccg AGT AAA GGA GAA GAA CTT TTC ACT
    MET His His His His His His Gly Ser Ser Asp Pro Ser Lys Gly Glu Glu Leu Phe Thr
                                                    -> GFPN (2-214)
             61/21
             GGA GTT GTC CCA ATT CTT GTT GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT TCT GTC
             Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val

             121/41
             CGT GGA GAG GGT GAA GGT GAT GCA ACA ATC GGA AAA CTT ACC CTT AAA TTT ATT TGC ACT
             Arg Gly Glu Gly Glu Gly Asp Ala Thr Ile Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr

             181/61
             ACT GGA AAA CTA CCT GTT CCA TGG CCA ACA CTT GTC ACT ACT CTG ACT TAT GGT GTT CAA
             Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln

             241/81
             TGC TTT TCC CGT TAT CCG GAT CAC ATG AAA CGG CAT GAC TTT TTC AAG AGT GCC ATG CCC
             Cys Phe Ser Arg Tyr Pro Asp His MET Lys Arg His Asp Phe Phe Lys Ser Ala MET Pro

             301/101
             GAA GGT TAT GTA CAG GAA CGC ACT ATA TCT TTC AAA GAT GAC GGG AAA TAC AAG ACG CGT
             Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Lys Tyr Lys Thr Arg

             361/121
             GCT GTA GTC AAG TTT GAC GGT GAT ACC CTT GTT AAT CGT ATC GAG TTA AAA GGT ACT GAT
             Ala Val Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Thr Asp

             421/141
             TTT AAA GAA GAT GGA AAC ATT CTC GGA CAC AAA CTC GAA TAC AAC TTT AAC TCA CAC AAT
             Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Phe Asn Ser His Asn

             481/161
             GTA TAC ATC ACG GCA GAC AAA CAA AAG AAT GGA ATC AAA GCT AAC TTC ACT GTT CGC CAC
             Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Thr Val Arg His

             541/181
             AAC GTT GAA GAT GGC TCC GTT CAA CTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC
             Asn Val Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly

             601/201
             GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA CAA ACT GTT CTT TCG AAA
             Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Thr Val Leu Ser Lys

             661/221                  XhoI
             GAT CCC AAC GAA AAG tga ctc gag
             Asp Pro Asn Glu Lys ***
                                 STOP
```

Figure S3. The DNA and amino acid sequence of GFPN on pET22b-6H-GFPN. Mutations (S30R, Y39I, F64L, S65T, F99S, N105K, E111V, I128T, Y145F, M153T, V163A, K166T, I167V, I171V, S205T, and A206V) are shown in Red. The linker and 6 × His Tag are shown in green and blue, respectively.

pET22b-6H-MCS-6H

```
                      BglII
GCG TAG AGG ATC aga tct CGA TCC CGC GAA ATT AAT ACG ACT CAC TAT AGG GGA ATT GTG
                                            T7 Promoter       Lac Operator
                            XbaI                                rbs          NdeI
AGC GGA TAA CAA TTC CCC tct aga AAT AAT TTT GTT TAA CTT TAA GAA GGA GAT ATA cat

1/1                     Nco I          EagI   BamH I  EcoRI   SacI      SalI
ATG CAC CAT CAT CAT CAc cat ggC TCT GGc ggc cgg gat ccg aat tcg agc tcC gtc gac
MET His His His His His His Gly Ser Gly Gly Arg Asp Pro Asn Ser Ser Ser Val Asp
                  6 x His Tag
61/21
HindIII   NsiI            Xho I                          STOP
aag ctt atg cat GCT AGC ctc gag CAC CAC CAC CAC CAC CAC TAA TAA TGA CTA GTC AGC
Lys Leu Met His Ala Ser Leu Glu His His His His His His *** *** ***
                           6 x His Tag

181/61
TGA TCC GGC TGC TAA CAA AGC CCG AAA GGA AGC TGA GTT GGC TGC TGC CAC CGC TGA GCA

241/81
ATA ACT AGC ATA ACC CCT TGG GGC CTC TAA ACG GGT CTT GAG GGG TTT TTT GCT GAA AGG
```

Figure S4. The DNA and amino acid sequence of a modified pET22b(+) vector, pET22b-6H-MCS-6H.

pET22b-sDnaB-cpM4(DEVD)

```
        NdeI
           1/1
       cat atg TCA CCA GAA ATA GAA AAG TTG TCT CAG AGT GAT ATT TAC TGG GAC TCC ATC GTT TCT
           MET Ser Pro Glu Ile Glu Lys Leu Ser Gln Ser Asp Ile Tyr Trp Asp Ser Ile Val Ser
              -> Ssp DnaBC Intein (106-154)
           61/21
           ATT ACG GAG ACT GGA GTC GAA GAG GTT TTT GAT TTG ACT GTG CCA GGA CCA CAT AAC TTT
           Ile Thr Glu Thr Gly Val Glu Glu Val Phe Asp Leu Thr Val Pro Gly Pro His Asn Phe

           121/41                                        EagI
           GTC GCG AAT GAC ATC ATT GTA CAC AAC Tcg gcc gCT GGT ATC ACC GGT GAT GAG GTG GAC
           Val Ala Asn Asp Ile Ile Val His Asn Ser Ala Ala Gly Ile Thr Gly Asp Glu Val Asp
                                             -> GFPC M4 11-16          Casp3 Substrate (DEVD)
           181/61                                        NsiI
           GGT CGC GAT CAC ATG GTT CTG CAC GAA TAT GTa tgc atC TCT GGC GAT AGT CTG ATC AGC
           Gly Arg Asp His MET Val Leu His Glu Tyr Val Cys Ile Ser Gly Asp Ser Leu Ile Ser
              -> GFPC M4 1-10                            -> Ssp DnaB N-Intein (1-105)
           241/81
           CTG GCT AGC ACA GGA AAA AGA GTT TCT ATT AAA GAT TTG TTA GAT GAA AAA GAT TTT GAA
           Leu Ala Ser Thr Gly Lys Arg Val Ser Ile Lys Asp Leu Leu Asp Glu Lys Asp Phe Glu

           301/101
           ATA TGG GCA ATT AAT GAA CAG ACG ATG AAG CTA GAA TCA GCT AAA GTT AGT CGT GTA TTT
           Ile Trp Ala Ile Asn Glu Gln Thr MET Lys Leu Glu Ser Ala Lys Val Ser Arg Val Phe

           361/121
           TGT ACT GGC AAA AAG CTA GTT TAT ATT CTA AAA ACT CGA CTA GGT AGA ACT ATC AAG GCA
           Cys Thr Gly Lys Lys Leu Val Tyr Ile Leu Lys Thr Arg Leu Gly Arg Thr Ile Lys Ala

           421/141
           ACA GCA AAT CAT AGA TTT TTA ACT ATT GAT GGT TGG AAA AGA TTA GAT GAG CTA TCT TTA
           Thr Ala Asn His Arg Phe Leu Thr Ile Asp Gly Trp Lys Arg Leu Asp Glu Leu Ser Leu

           481/161                                                          XhoI
           AAA GAG CAT ATT GCT CTA CCC CGT AAA CTA GAA AGC TCC TCT TTA CAA TTG CTC GAG CAC
           Lys Glu His Ile Ala Leu Pro Arg Lys Leu Glu Ser Ser Ser Leu Gln Leu Leu Glu His

           541/181                        SpeI
           CAC CAC CAC CAC CAC TAA TAA TGA CTA GTC AGC TGA TCC GGC TGC TGC TAA CAA AGC CCG
           His His His His His *** *** ***
```

Figure S5. The DNA and amino acid sequence of sDnaB-cpM4(DEVD) on pET22b- sDnaB-cpM4(DEVD). DnaB C-intein (106-154), cpM4, caspase-3 substrate (DEVD), and DnaB N-intein (1-105) sequences are shown in cyan, red, pink, and blue, respectively.

pET22b-sDnaB-cpM4(DEVG)

```
        NdeI
           1/1
       cat atg TCA CCA GAA ATA GAA AAG TTG TCT CAG AGT GAT ATT TAC TGG GAC TCC ATC GTT TCT
           MET Ser Pro Glu Ile Glu Lys Leu Ser Gln Ser Asp Ile Tyr Trp Asp Ser Ile Val Ser
                -> Ssp DnaBC Intein (106-154)
           61/21
           ATT ACG GAG ACT GGA GTC GAA GAG GTT TTT GAT TTG ACT GTG CCA GGA CCA CAT AAC TTT
           Ile Thr Glu Thr Gly Val Glu Glu Val Phe Asp Leu Thr Val Pro Gly Pro His Asn Phe

           121/41                                  EagI
           GTC GCG AAT GAC ATC ATT GTA CAC AAC Tcg gcc gCT GGT ATC ACC GGT GAT GAG GTG GgC
           Val Ala Asn Asp Ile Ile Val His Asn Ser Ala Ala Gly Ile Thr Gly Asp Glu Val GLY
                                           -> GFPC M4 11-16          Casp3 Substrate (DEVG)
           181/61                                  NsiI
           GGT CGC GAT CAC ATG GTT CTG CAC GAA TAT GTa tgc atC TCT GGC GAT AGT CTG ATC AGC
           Gly Arg Asp His MET Val Leu His Glu Tyr Val Cys Ile Ser Gly Asp Ser Leu Ile Ser
              -> GFPC M4 1-10                      -> Ssp DnaB N-Intein (1-105)
           241/81
           CTG GCT AGC ACA GGA AAA AGA GTT TCT ATT AAA GAT TTG TTA GAT GAA AAA GAT TTT GAA
           Leu Ala Ser Thr Gly Lys Arg Val Ser Ile Lys Asp Leu Leu Asp Glu Lys Asp Phe Glu

           301/101
           ATA TGG GCA ATT AAT GAA CAG ACG ATG AAG CTA GAA TCA GCT AAA GTT AGT CGT GTA TTT
           Ile Trp Ala Ile Asn Glu Gln Thr MET Lys Leu Glu Ser Ala Lys Val Ser Arg Val Phe

           361/121
           TGT ACT GGC AAA AAG CTA GTT TAT ATT CTA AAA ACT CGA CTA GGT AGA ACT ATC AAG GCA
           Cys Thr Gly Lys Lys Leu Val Tyr Ile Leu Lys Thr Arg Leu Gly Arg Thr Ile Lys Ala

           421/141
           ACA GCA AAT CAT AGA TTT TTA ACT ATT GAT GGT TGG AAA AGA TTA GAT GAG CTA TCT TTA
           Thr Ala Asn His Arg Phe Leu Thr Ile Asp Gly Trp Lys Arg Leu Asp Glu Leu Ser Leu

           481/161                                                          XhoI
           AAA GAG CAT ATT GCT CTA CCC CGT AAA CTA GAA AGC TCC TCT TTA CAA TTG CTC GAG CAC
           Lys Glu His Ile Ala Leu Pro Arg Lys Leu Glu Ser Ser Ser Leu Gln Leu Leu Glu His

           541/181                       SpeI
           CAC CAC CAC CAC CAC TAA TAA TGA CTA GTC AGC TGA TCC GGC TGC TGC TAA CAA AGC CCG
           His His His His His *** *** ***
```

Figure S6. The DNA and amino acid sequence of sDnaB-cpM4(DEVG) on pET22b- sDnaB-cpM4(DEVG). DnaB C-intein (106-154), cpM4, DEVG, and DnaB N-intein (1-105) sequences are shown in cyan, red, pink, and blue, respectively.

pET22b-sDnaB-cpM4(DEVD)-T70A/H73A

```
         NdeI
            1/1
        cat atg TCA CCA GAA ATA GAA AAG TTG TCT CAG AGT GAT ATT TAC TGG GAC TCC ATC GTT TCT
            MET Ser Pro Glu Ile Glu Lys Leu Ser Gln Ser Asp Ile Tyr Trp Asp Ser Ile Val Ser
                -> Ssp DnaBC Intein (106-154)
            61/21
            ATT ACG GAG ACT GGA GTC GAA GAG GTT TTT GAT TTG ACT GTG CCA GGA CCA CAT AAC TTT
            Ile Thr Glu Thr Gly Val Glu Glu Val Phe Asp Leu Thr Val Pro Gly Pro His Asn Phe

            121/41                                    EagI
            GTC GCG AAT GAC ATC ATT GTA CAC AAC Tcg gcc gCT GGT ATC ACC GGT GAT GAG GTG GAC
            Val Ala Asn Asp Ile Ile Val His Asn Ser Ala Ala Gly Ile Thr Gly Asp Glu Val Asp
                                              -> GFPC M4 11-16          Casp3 Substrate (DEVD)
            181/61                                    NsiI
            GGT CGC GAT CAC ATG GTT CTG CAC GAA TAT GTa tgc atC TCT GGC GAT AGT CTG ATC AGC
            Gly Arg Asp His MET Val Leu His Glu Tyr Val Cys Ile Ser Gly Asp Ser Leu Ile Ser
                 -> GFPC M4 1-10                      -> Ssp DnaB N-Intein (1-105)
            241/81
            CTG GCT AGC ACA GGA AAA AGA GTT TCT ATT AAA GAT TTG TTA GAT GAA AAA GAT TTT GAA
            Leu Ala Ser Thr Gly Lys Arg Val Ser Ile Lys Asp Leu Leu Asp Glu Lys Asp Phe Glu

            301/101
            ATA TGG GCA ATT AAT GAA CAG ACG ATG AAG CTA GAA TCA GCT AAA GTT AGT CGT GTA TTT
            Ile Trp Ala Ile Asn Glu Gln Thr MET Lys Leu Glu Ser Ala Lys Val Ser Arg Val Phe

            361/121
            TGT ACT GGC AAA AAG CTA GTT TAT ATT CTA AAA ACT CGA CTA GGT AGA ACT ATC AAG GCA
            Cys Thr Gly Lys Lys Leu Val Tyr Ile Leu Lys Thr Arg Leu Gly Arg Thr Ile Lys Ala

            421/141
            gCA GCA AAT gcT AGA TTT TTA ACT ATT GAT GGT TGG AAA AGA TTA GAT GAG CTA TCT TTA
            ALA Ala Asn ALA Arg Phe Leu Thr Ile Asp Gly Trp Lys Arg Leu Asp Glu Leu Ser Leu

            481/161                                                          XhoI
            AAA GAG CAT ATT GCT CTA CCC CGT AAA CTA GAA AGC TCC TCT TTA CAA TTG CTC GAG CAC
            Lys Glu His Ile Ala Leu Pro Arg Lys Leu Glu Ser Ser Ser Leu Gln Leu Leu Glu His

            541/181                          SpeI
            CAC CAC CAC CAC CAC TAA TAA TGA CTA GTC AGC TGA TCC GGC TGC TGC TAA CAA AGC CCG
            His His His His His *** *** ***
```

Figure S7. The DNA and amino acid sequence of sDnaB-cpM4(DEVD)-T70A/H73A on pET22b-sDnaB-cpM4(DEVD)-T70A/H72A. DnaB C-intein (106-154), cpM4, caspase-3 substrate (DEVD), and DnaB N-intein (1-105) sequences are shown in cyan, red, pink, and blue, respectively.