# **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: <a href="mailto:sakamoto@tagen.tohoku.ac.jp">sakamoto@tagen.tohoku.ac.jp</a>; <a href="mailto:hiko@tagen.tohoku.ac.jp">hiko@tagen.tohoku.ac.jp</a>; <a href="mailto:hiko@tagen.tohoku.ac.jp">hiko@tagen.tohoku.ac.jp</a>

10

# 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 <sup>15</sup> 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 <sup>20</sup> 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-<sup>25</sup> 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 TC<u>G GAT CC</u>G AGT AAA GGA GAA GAA CTT TTC-3' and 5'- T CGA ATT <u>CTC GAG</u> 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 <sup>5</sup> 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). <sup>10</sup> 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 <sup>15</sup> 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 intern 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 <sup>20</sup> 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 25 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

15

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-sDnaBcpM4(DEVD), pET22b-sDnaB-cpM4(DEVD) or pET22b-sDnaB-cpM4(DEVD)-T70A/H73A plasmid <sup>20</sup> (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 <sup>25</sup> 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)<sup>+</sup>] (Calcd. = 2354.15); cM4(DEVG), *m/z* 2296.56 [(M+H)<sup>+</sup>] (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  $\mu$ L at 37°C for 1 hr. The 10  $\mu$ L of reaction solution was mixed with a GFP N-terminal fragment (final protein concentration was 1  $^{5}$   $\mu$ M) 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  $\mu$ L. 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 ( $\Delta$ F) at 508 nm <sup>10</sup> 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<sub>0</sub>, where F and F<sub>0</sub> denote fluorescent intensities after and before incubation, respectively. Arbitrary fluorescence units (A.U.).

15

20



**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

	Nde	εI																				
5		1/1									E	BamHI	Ι									
	cat	atg	CAC	CAT	CAT	CAT	CAc	cat	ggC	TCT	TC <u>g</u>	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	
		61/2	21											-> (	FPN	(2-2	(14)					
10		GGA	GTT	GTC	CCA	ΔΨΨ	СТТ	GTT	GAA	ጥጥል	GAT	GGT	GAT	GTT	ΔΔΤ	GGG	CAC	ΔΔΔ	ጥጥጥ	тст	GTC	
10		Glv	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Glv	Asp	Val	Asn	Glv	His	Lvs	Phe	Ser	Val	
		- 1									- 1	- 1	- 1			- 1		2 -				
		121/	41																			
		CGT	GGA	GAG	GGT	GAA	GGT	GAT	GCA	ACA	ATC	GGA	AAA	CTT	ACC	CTT	AAA	TTT	ATT	TGC	ACT	
15		Arg	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Ile	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	
		1.91	/61																			
		ACT	GGA	AAA	ста	ССТ	GTT	CCA	ТGG	CCA	ACA	СТТ	GTC	ACT	ACT	CTG	ACT	ТАТ	GGT	GTT	САА	
		Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	
20			_	-					-									-	_			
		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	Hıs	ME'l'	Lys	Arg	Hıs	Asp	Phe	Phe	Lys	Ser	Ala	ME'l'	Pro	
25		301/	/101																			
25		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					~ ~ -			~			~ ~ -		~ ~ ~					~	
30		GCT	G'I'A	GTC	AAG	TTT	GAC	GGT	GA'I'	ACC	CTT	GTT	AA'I'	CGT	ATC	GAG	'I''I'A	AAA	GGT	AC'I'	GA'I'	
		Ald	Val	Val	цуs	Pne	Giù	GIY	Asp	.T.U.T.	Leu	Val	ASI	Arg	ITE	Giù	Leu	цуs	GIY	TUL	Asp	
		421/	/141																			
		TTT	AAA	GAA	GAT	GGA	AAC	ATT	CTC	GGA	CAC	AAA	CTC	GAA	TAC	AAC	Τ <mark>Τ</mark> Τ	AAC	TCA	CAC	AAT	
35		Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	Phe	Asn	Ser	His	Asn	
			10.00																			
		481/	161	лщо	» C C	CC7	CAC	~ ~ ~	C 7 7	7 7 C	7 7 M	CC7	лщо	~ ~ ~	COM	770	mmc	7 <b>C</b> III	Cmm	000	CAC	
		Val	Tvr	TIP	Thr	Ala	Asn	LVS	Gln	LVS	Asn	GUV	TIP	LVS	Ala	AAC	Phe	Thr	Val	Ara	His	
40		var	- <u>y</u> -	TTC		1110	1100	цус	0111	шуU	11011	OLY	110	шуU	min	11011	1110	1111	var	g	1110	
		541/	/181																			
		AAC	$_{\rm GTT}$	GAA	GAT	GG <mark>C</mark>	TCC	$\operatorname{GTT}$	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	
45		601	/201																			
45		GAT	GGC	CCT	GTC	CTT	ጥጥል	CCA	GAC		CAT	TAC	CTG	TCC			∆ ⊂ T	GTT	CTT	TCG		
		Asp	Glv	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tvr	Leu	Ser	Thr	Gln	Thr	Val	Leu	Ser	Lvs	
		12	1						12			- 1 -									-10	
		661/	/221					Xho	DI													
50		GAT	CCC	AAC	GAA	AAG	tga	ctc	gag													
		Asp	Pro	Asn	Glu	Lys	***	_														
							STOR	2														

**Figure S3.** The DNA and amino acid sequence of GFPN on pET22b-6H-GFPN. Mutations (S30R, <sup>55</sup> 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

5 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 10 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 Gly Ser Gly Gly Arg Asp Pro Asn Ser Ser Val Asp 15 6 x His Tag 61/21 NsiI HindIII Xho I STOP aag ett atg cat GCT AGC ete gag CAC CAC CAC CAC CAC CAC TAA TAA TG<u>A CTA GT</u>C AGC Lys Leu Met His Ala Ser Leu Glu His His His His His His \*\*\* \*\*\*  $^{***}$ 20 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 25 241/81 ATA ACT AGC ATA ACC CCT TGG GGC CTC TAA ACG GGT CTT GAG GGG TTT TTT GCT GAA AGG

30

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

## pET22b-sDnaB-cpM4(DEVD)

5															
	deI														
	1/1														
	t atg TCA CCA GAA ATA GAA AAG TTG TCT CAG AGT GAT ATT TAC TGG GAC TCC ATC G	TT TCT													
	MET Ser Pro Glu Ile Glu Lys Leu Ser Gln Ser Asp Ile Tyr Trp Asp Ser Ile V	al Ser													
10	-> 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 A	AC TTT													
	Ile Thr Glu Thr Gly Val Glu Glu Val Phe Asp Leu Thr Val Pro Gly Pro His A	.sn Phe													
15	121/41 <b>EagI</b>														
	GTC GCG AAT GAC ATC ATT GTA CAC AAC Teg gee gCT GGT ATC ACC GGT GAT GAG G	TG GAC													
	Val Ala Asn Asp Ile Ile Val His Asn Ser Ala Ala Gly Ile Thr Gly Asp Glu V	al Asp													
	-> GFPC M4 11-16 Casp3 Substr	ate (DEVD)													
	181/61 Nsil														
20	GGT CGC GAT CAC ATG GTT CTG CAC GAA TAT GT <mark>a tgc at</mark> C TCT GGC GAT AGT CTG A	.TC AGC													
	Gly Arg Asp His MET Val Leu His Glu Tyr Val Cys Ile Ser Gly Asp Ser Leu I	le 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 T	TT GAA													
25	Leu Ala Ser Thr Gly Lys Arg Val Ser lle Lys Asp Leu Leu Asp Glu Lys Asp P	he Glu													
	201/101														
	ATA TGG GCA ATT AAT GAA CAG ACG ATG AAG CTA GAA TCA GCT AAA GTT AGT CGT G	TA TIT A Dhe													
20	THE TTP ATA THE ASH GIU GIN THE MET LYS LEU GIU SEE ATA LYS VAL SEE AFG V	al Phe													
30	261/121														
	J01/121 TET JOT CCC ANA ANG CTA CTT TAT ATT CTA ANA ACT CCA CTA CCT ACA ACT ATC A														
	Cus Thr Clu Lus Lus Lus Val Tur The Lau Lus Thr Ara Cau Clu Ara Thr The L	vs Ala													
	Cyo mi Giy Lyo Lyo Lee var fyr fre Lee Lyo mi my Lee Cry my mi fre L	yo mia													
35	421/141														
55	ACA GCA AAT CAT AGA TTT TTA ACT ATT GAT GGT TGG AAA AGA TTA GAT GAG CTA T	СТ ТТА													
	Thr Ala Asn His Arg Phe Leu Thr Ile Asp Gly Trp Lvs Arg Leu Asp Glu Leu S	er Leu													
	481/161 <b>Xhoi</b>														
40	AAA GAG CAT ATT GCT CTA CCC CGT AAA CTA GAA AGC TCC TCT TTA CAA TTG CTC G	AG CAC													
	Lys Glu His Ile Ala Leu Pro Arg Lys Leu Glu Ser Ser Leu Gln Leu $\overline{\text{Leu G}}$	lu His													
	541/181 SpeI														
	CAC CAC CAC CAC TAA TAA TGA CTA GTC AGC TGA TCC GGC TGC TGC TAA CAA A	.GC CCG													
45	His His His His $^{***}$														

**Figure S5.** The DNA and amino acid sequence of sDnaB-cpM4(DEVD) on pET22b- sDnaB-<sup>50</sup> 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)

5	Nde	I																			
		1/1																			
	cat	ata	тса	CCA	GAA	АТА	GAA	AAG	TTG	тст	CAG	AGT	GAT	АТТ	TAC	TGG	GAC	тсс	ATC	GTT	ТСТ
		MET	Ser	Pro	Glu	TIe	Glu	LWS	Leu	Ser	Gln	Ser	Asp	TIe	Tvr	Trp	Asp	Ser	TIe	Val	Ser
			->	Sen	Dnal			$\frac{1}{1}$	16-1F	54)	0111	DCT	пор	110	+ <u>y</u> +	ттр	1100	DCL	110	var	001
10		61/2	1	SSP	Dilai	JC 11	ICETI	1 (1)	/0 1.	,-,											
10		01/2	100	CAC	7 O III	<b>CCA</b>	ста	<b>C D D</b>	C7 C	CUUM		<b>C M</b>	mmc		CILC	007	CC7	007	<b>0 M m</b>		
		ATT	ACG	GAG	ACT	GGA	GTC	GAA	GAG	GTT	TIT	GAT	TTG	ACT	GTG	CCA	GGA	CCA	CAT	AAC	TTT
		TTe	Thr	GIU	Thr	GTÀ	vai	GIU	GIU	vai	Pne	Asp	Leu	Tur	val	Pro	GТХ	Pro	HIS	Asn	Pne
		121/	41								Ţ	TagT									
15		GTC	606	ידע ב	GAC	ATC	<u>አ</u> ጥጥ	GTTA	CAC	AAC	Tea	acc	сСШ	CCT	Ътс	ACC	CCT	СЪТ	CAC	GTG	GaC
15		Val	Ala	Asn	Asp	TIO	TIO	Val	Hie	Acn	Ser	Ala	<u>y</u> or Ala	Glv	TIO	Thr	Glv	Aen	G111	Val	GLY
		var	mu	11011	nsp	110	110	var	1113	ASI	-> (		M/ 1	0 ± y 1 <b>1 - 1</b> 4	s C	THT	Car	-n3p	lubei	-rate	(DEVC)
		181/61 Vert Mait																			
		GGT	CGC	GAT	CAC	ATG	GTT	CTG	CAC	GAA	ТАТ	GTa	tac	atc	тст	GGC	GAT	AGT	CTG	ATC	AGC
20		Glv	Ara	Asn	His	MET	Val	Len	His	Glu	Tvr	Val	Cvs	TIP	Ser	Glv	Asn	Ser	Leu	Tle	Ser
20		013	-> (	TEPC	м4 .	1-10	vur	шеч		Oru	- <u>y</u> -	var	-> <	Sen I	)naB	N-Tr	ntop	1 - (1)	-105	110	DCI
	-/ GFPC M4 1-10 -/ SSP DNAB N-INTEIN (1-105) 241/81																				
		CTTC	CCT	ACC	707	CCA	<u>א</u> א	ACA	CTT	ΨĊΨ	አጥጥ	~~~	CAT	TTC	ጥጥአ	CAT	GAA	אאא	CAT	<b>TTT</b>	CAA
		LOU		Sor	Thr	Clu	Tue	AGA	Val	Sor	TIO	Tue	Jen	LOU	LOU	Jen	Clu	Tue	Jen	Dho	GLU
25		шец	пта	Der	THT	Gry	цуз	ALQ	var	Der	TTC	цуз	лэр	шец	шец	лэр	Gru	цуз	лэр	THE	Giù
25		301/	101																		
		201/	TOT	CC7	7 mm	~~~	C 7 7	CAC	ACC	лтс	770	CILLA	C 7 7	men	COM	ההה	CIIII	A CIT	CCT	CIIIA	mmm
		AIA Tla	TGG	GCA	AII Tla	AAI	GAA	CAG	ACG	AIG	AAG	Tau	GAA	ICA	GCI 71-	AAA	GII	AGI	CGI	GIA	III Dhe
		TTe	тгр	Ald	TTe	ASII	Giù	GIII	THE	MET	цуѕ	ьец	Giù	Ser	Ald	цуѕ	Val	Set	Arg	Val	Plie
30		361/	121																		
		TGT	ACT	GGC	ΔΔΔ	AAG	СТА	GTT	тат	ΑͲͲ	СТА	ААА	ACT	CGA	СТА	GGT	AGA	ACT	ATC	AAG	GCA
		Cvs	Thr	Glv	Lvs	Lvs	Len	Val	Tvr	Tle	Len	Lvs	Thr	Ara	Len	Glv	Ara	Thr	Tle	Lvs	Ala
		010			210	210	200		-1-		200	210		9	200		9		110	210	
		421/	141																		
35		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
		401/	1 C 1																	-	
		401/101 XAOI													<b>G A G</b>						
		AAA -	GAG	CAT	A'I''I'	GCT	CTA	CCC	CGT	AAA	CTA	GAA	AGC	TCC	TCT	TTA	CAA	TTG	CTC	GAG	CAC
40		Lуs	GLU	HIS	тте	Ala	Leu	Pro	Arg	Lуs	Leu	Glu	Ser	Ser	Ser	Leu	Gln	Leu	Leu	Glu	Hls
		541/	181						5	SpeJ											
		CAC	CAC	CAC	CAC	CAC	ТАА	ТАА	TGA	СТА	GTC	AGC	тga	TCC	GGC	TGC	TGC	ТАА	CAA	AGC	CCG
		His	His	His	His	His	* * *	***	***	2	<u> </u>		1 011	100	200	100	100		J		
45																					
-																					

**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 <sup>50</sup> shown in cyan, red, pink, and blue, respectively.

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

5	Nde	εL																			
		1/1																			
	cat	atσ	TCA	CCA	GAA	ATA	GAA	AAG	TTG	TCT	CAG	AGT	GAT	ATT	TAC	TGG	GAC	TCC	ATC	GTT	TCT
		MET	Ser	Pro	G111	Tle	Glu	Lvs	Len	Ser	Gln	Ser	Asp	Tle	Tvr	Trp	Asp	Ser	Tle	Val	Ser
			-> 9	Sen I	DnaB	Tni	tein	(10)	5-154	1)	0.1.11	001	110 P		- 1 -	P	11010	001			001
10		61/2	21	oop .				(10)		-/											
10		ΔTT		GAG	ACT	CCA	CTTC	CVV	GNG	CTT	ጥጥጥ	CAT	TTC	ACT	CTC	CCA	CCA	CCA	CAT	AAC	The second se
		TIO	Thr	Glu	Thr	Cly	Val	Clu	Glu	Val	Dho	Aen	Tou	Thr	Wal	Dro	Clw	Pro	Hie	Aen	Dho
		TTE	TIIT	Gru	THIT	GTĀ	var	Gru	Gru	var	The	лэр	шец	THT	var	110	GTA	110	111.5	ASII	1 ne
		121	/ / 1									TagT									
16			CCC	አአጥ	CAC	λΨC	አመመ	CTUN	CAC	AAC	Tag	agi	а С Ш	CCT	λΨC	ACC	CCT	CAT	CAC	CTTC	CAC
15		Wal	Ala	AAI	Jan	TIO	TIO	Wol	UAC No.	AAC	200	315	<u>y</u> cı <u></u> 12	Clu	TIO	Thr	Clu	Jan 1	Clu	Wol	Jan
		Val	Ald	ASII	Asp	TTe	TTe	Val	nis	ASII	Ser	ALA	MAI	GIY	r Te	THT	GLY	Asp	Gru	var	ASP
		101	(61								-~ (	JE PC	M4 I	L L - L ( F	5		Cas	sps a	Subsi	-rate	e (DEVD)
		TOT/	CCC	CAM	CAC	лпс	Cmm	CILC	CAC	C 7 7	mam	Cm-	NSI	L 	mom	~~~~	сл m	л <b>с</b> ш	ome	3.00	3.00
		GGI	CGC	GAI	CAC	AIG	GII	CIG	CAC	GAA	TAI	GI <b>a</b>	Lgc	al.	TUT	GGC	GAT	AGT	CIG	ATC	AGC
20		GΙΫ	Arg	Asp	HIS	MET	vai	Leu	HIS	GIU	Tyr	vai	Cys	e	ser	GTĀ	Asp	ser	Leu	ITE	Ser
		0.4.1	->	GFP	C M4	1-10	J						-> 8	ssp I	InaB	N-11	nterr	n (1-	-105)	£	
		241/	81																		
		CTG	GCT	AGC	ACA	GGA	AAA	AGA	G'I''I'	TCT	A'1''1'	AAA	GA'I'	'T'TG	'T'TA	GA'I'	GAA	AAA	GA'I'	TTT	GAA
		Leu	Ala	Ser	Thr	Gly	Lys	Arg	Val	Ser	lle	Lys	Asp	Leu	Leu	Asp	GLu	Lys	Asp	Phe	Glu
25																					
		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
30		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																		
35		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/																Xho	JI	
		AAA	GAG	CAT	ATT	GCT	CTA	CCC	CGT	AAA	CTA	GAA	AGC	TCC	TCT	TTA	CAA	TTG	CTC	GAG	CAC
40		Lys	Glu	His	Ile	Ala	Leu	Pro	Arg	Lys	Leu	Glu	Ser	Ser	Ser	Leu	Gln	Leu	Leu	Glu	His
		541/	181/							Spel	τ										
		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	* * *	* * *	* * *												
45																					

**Figure S7.** The DNA and amino acid sequence of sDnaB-cpM4(DEVD)-T70A/H73A on pET22bsDnaB-cpM4(DEVD)-T70A/H72A. DnaB C-intein (106-154), cpM4, caspase-3 substrate (DEVD), <sup>50</sup> and DnaB N-intein (1-105) sequences are shown in cyan, red, pink, and blue, respectively.