Supplemental Data

Table S1. Data collection and refinement statistics*.

Space group	$P4_32_12$
Unit cell parameters (Å)	$a = b = 148.32$, $c = 116.77$

Data collection

Synchrotron beam line	NW12 (PF, Japan)	SERCAT (APS, USA)
Anomalous scatter	None	Se
Data set	Native	Peak
Resolution (Å)	30-2.6 (2.69-2.6)§	50-3.2 (3.31–3.2)
Reflections (Unique)	38590	20843
Redundancy	6.4 (3.0)	8.1 (4.0)
I/σ(I)	32.6 (2.8)	24.3 (2.5)
R_{merge} (%)	5.7 (49.8)	7.0 (45.7)
Completeness (%)	97.1 (93.4)	98.5 (97.3)

Phasing statistics

Space group	$P4_32_12$
Resolution (Å)	40.0 - 3.2
Number of Se atoms	6
Figure of merit	0.37

Refinement

Space group	$P4_3$¶
Resolution (Å)	10.0 - 2.6 (2.69 – 2.6)‖
Reflections used	72792
R_{factor} (%)	0.226 (0.271)
R_{free} (%)	0.257 (0.290)

Overall B-factor/RMSD (Å^2)	42.5/1.6
Number of protein atoms	7758
Number of water molecules	333

Model quality

RMSD bond length (Å)	0.015
RMSD bond angles (°)	1.79
RMSD improper angles (°)	0.83

Ramachadran plot (Regions)	Number of residues (%)
Most favorable	87.1
Allowed	12.9
Generously allowed	0.0
Disallowed	0.0

$R_{merge}=\Sigma_{hkl}\Sigma_j \mid I_j(hkl) - \langle I(hkl)\rangle \mid / \Sigma_{hkl}\Sigma_j \langle I(hkl)\rangle$, where $I_j(hkl)$ and $\langle I(hkl)\rangle$ are the intensity of measurement *j* and the mean intensity for the reflection with indices *hkl*,

respectively. $R_{factor,\ free}=\Sigma_{hkl}||F_{calc}(hkl)| - |F_{obs}(hkl)||/\Sigma_{hkl}|F_{obs}|$, where the crystallographic R-factor is calculated including and excluding reflections in the refinement. The free reflections constituted 5% of the total number of reflections. RMSD – root mean square deviation. $I/\sigma(I)$ – ratio of mean intensity to a mean standard deviation of intensity. *The data for the native protein and its SeMet derivative were collected at 100K at the synchrotron beam lines NW12 (Photon Factory, Japan) and SERCAT (Argonne, USA), respectively, and processed with the HKL2000 program package (Otwinowski and Minor, 1997). The initial phases were obtained by the single anomalous dispersion technique using the MLPHARE program and improved by solvent flattening using the DM program (Collaborative Computational Project, 1994). The model was built manually using the O program (Jones et al., 1991) and refined with the CNS program (Brunger et al., 1998) to the final R-factor/R-free of 0.251/0.266 at 2.6Å resolution. §The data for the highest resolution shell are shown in brackets. ¶The crystals belong to $P4_3$ space group with a perfect (50%) merohedral twinning mimicking $P4_32_12$ space group. As the data processing showed the nearly identical R-merge values (0.4% difference) for the $P4_3$ and $P4_32_12$ space groups consistent with concept of the perfect twinning, the data were processed in $P4_32_12$ space group and were expanded to $P4_3$ space group for refinement to remove the random statistical error that would be introduced in the intensities of the reflections related by the twinning operator if the data are processed in the $P4_3$ space group. The refinement was carried out in the $P4_3$ space group using the twinning option of the CNS program and resulted in ~6% drop of R-free at the very first rigid body step while disrupting the crystallographic $P4_32_12$ symmetry by ~1.5^o. In contrast, no R-free convergence and disruption of the

crystallographic symmetry was achieved when the twinning option was switched off during the rigid body refinement in the $P4_3$ space group. ‖The large native crystal the data from which were used in refinement contained a small crystal attached to its surface during the data collection. The data processing showed that the low resolution intensities (~30 – 10Å) were severely affected by the overlap with the reflections obtained from the small crystal. At higher resolution, the spots obtained from these two crystals were largely resolved, whereas those overlapping reflections could be rejected during the merging procedure without loss of the overall completeness given the high redundancy of the data set collected. Therefore the low resolution was limited to 10Å during the refinement to avoid the effect of the poorly measured reflections.

References

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S. and al., e. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905-921.

Collaborative Computational Project (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D*, **50**, 760-763.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110-119.

Otwinowski, Z. and Minor, W. (1997) Processing X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.

Table S2. Plasmids used in this work

Name	Description	Source or note
GreB EXPRESSION VECTORS		
pIA576	*E. coli greB* under T7 ϕ10 promoter; encodes N-terminal HMK & C-terminal His_6 tags	this work
pIA577	*E. coli greB* under T7 ϕ10 promoter; encodes the C-terminal His_6 tag	(Perederina et al., 2006)
pIA689	GreB D41N in pIA577	this work
pIA697	GreB W148A in pIA576	this work
pIA698	GreB V139D in pIA576	this work
pIA720	GreB H101A in pIA576	this work
pIA721	GreB I118A in pIA576	this work
pIA722	GreB M124A in pIA576	this work
pVS110	GreB S122E in pIA576	this work
pVS111	GreB M124E in pIA576	this work
pVS112	GreB I118E in pIA576	this work
CORE RNAP EXPRESSION VECTORS		
pVS10	P_{T7}–*rpoA*–*rpoB*–*rpoC*^{His6}; *rpoZ*	http://www.osumicrobiology.org/faculty/documents/ExpressionvectorforE.coliRNApolymerase.htm
pAS4	β' 672,673AA in pVS10	this work
pVS50	β' 672,673DD in pVS10	this work
pIA665	β' 672,673RR in pVS10	this work
pVS51	deletions of the β' C-terminal CC residues 665-671 and 674-680 in pVS10	this work

Perederina, A. A., Vassylyeva, M. N., Berezin, I. A., Svetlov, V., Artsimovitch, I., and Vassylyev, D. G. (2006). Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from Escherichia coli and Gfh1 from Thermus thermophilus. Acta Crystallogr. F *62*, 44-46.

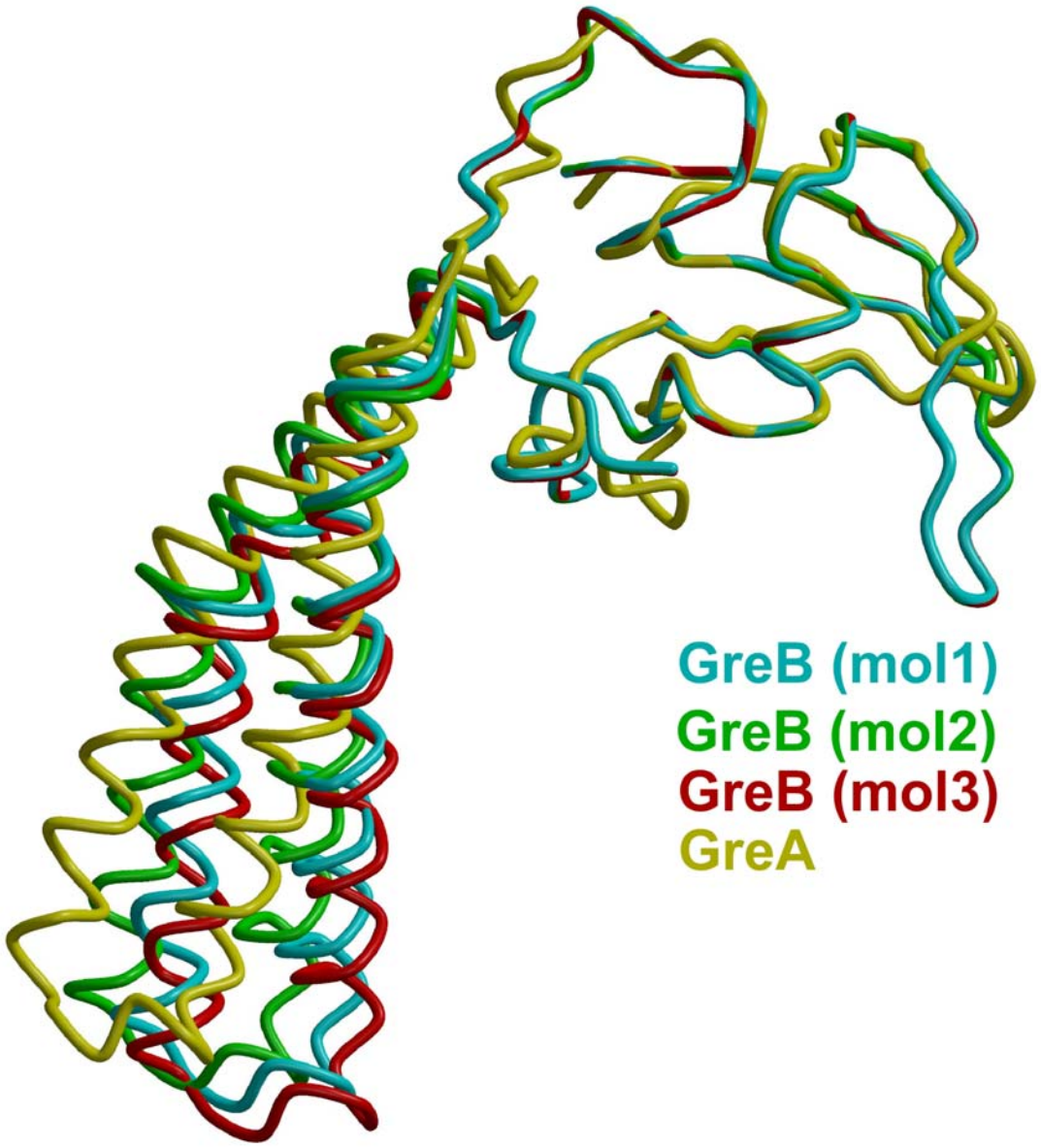

Figure S1. Superposition of the three independent GreB molecules in the asymmetric unit of the crystals and GreA structure by the C-domains.

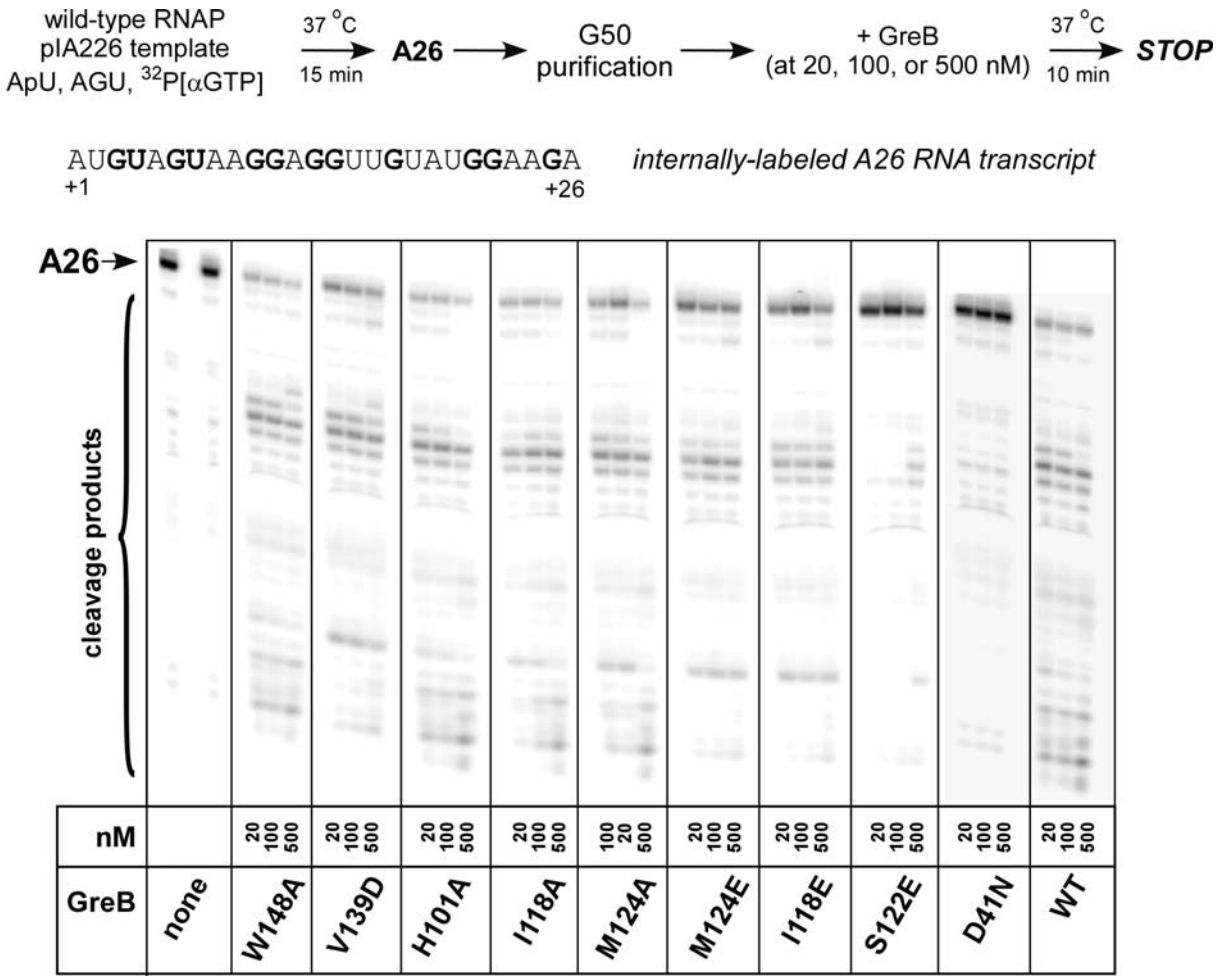

Figure S2 Side-by-side comparison of the effects of GreB substitutions on the nascent RNA cleavage in halted radiolabeled A26 TECs. The reaction schematic and the A26 RNA sequence are shown on top. GreB variants were added to A26 TECs formed with the wild-type RNAP (at ~20 nM) to a final concentration indicated below each lane. The reactions were incubated at 37 ^{o}C for 10 min, quenched by the addition of an equal volume of the STOP buffer, and analyzed on 12% denaturing urea/acrylamide (19:1) gels. The leftmost lane was loaded with A26 RNA before 37 ^{o}C incubation, the adjacent lane - with the A26 incubated with the GreB storage buffer. Positions of the A26 transcript and cleavage products are indicated; both the 5' and the 3' products are seen because the halted A26 RNA is internally labeled at 10 positions throughout its entire length (shown in bold above the gel panel). The assay was repeated at least twice for each GreB variant. The substitution of Ser122 for Glu produced the most pronounced cleavage defect, substitutions of other three C-domain residues (Ile118, Met124, Val13) for negatively charged residues (see Supplementary Figure 3 for the rate measurements) were less detrimental.

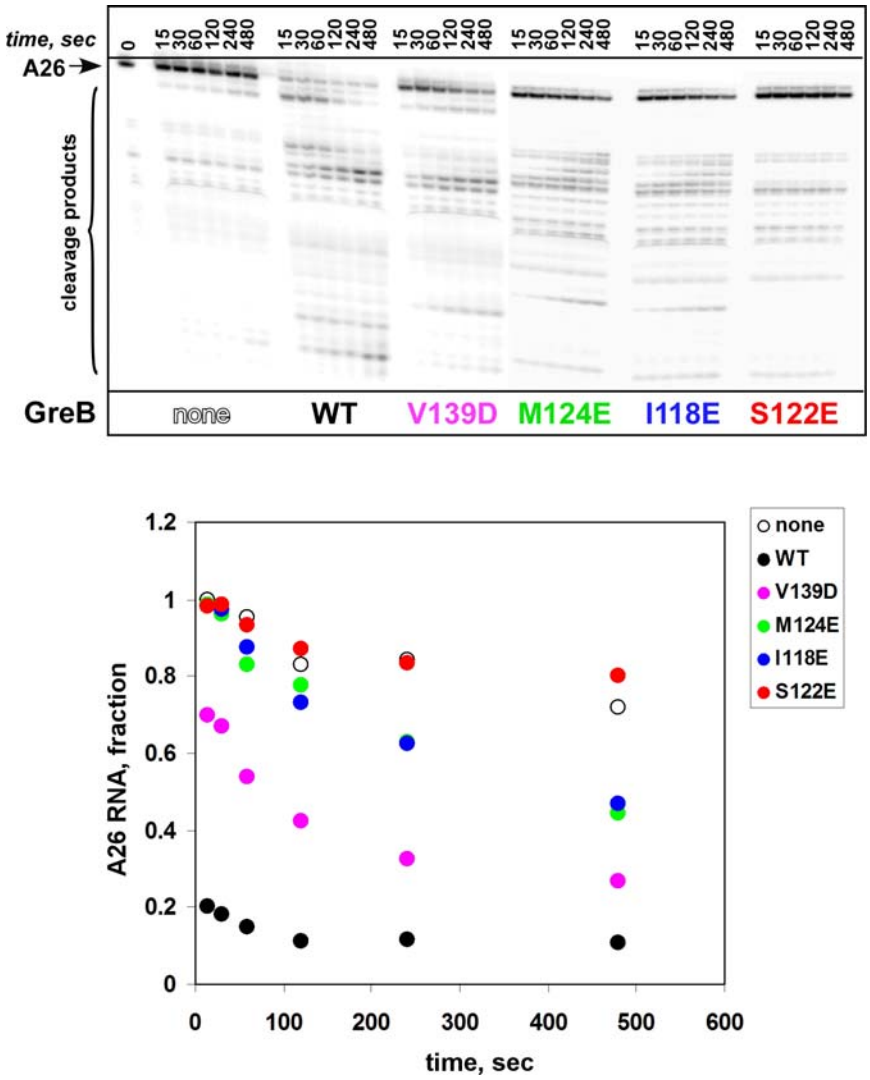

Figure S3 The effects of GreB substitutions on the rate of RNA cleavage in the A26 TECs (formed as shown in Supplementary Figure 2). The halted complexes were incubated with 100 nM GreB (or storage buffer) for 15, 30, 60, 120, 240, 480 seconds at 37 ^{o}C, quenched by the addition of an equal volume of the STOP buffer, and analyzed on 12% denaturing urea/acrylamide (19:1) gels. Quantification of the A26 RNA remaining after cleavage is shown below. With the wild-type GreB protein, nearly all susceptible to cleavage A26 RNA was shortened by at least two nt within the first 15 seconds, and the internally-labeled nascent transcript was further degraded during the course of the assay. The remaining complex (~10%) was resistant to cleavage even upon prolonged incubation (at least 30 min, data not shown); presumably, in these Gre-resistant TECs the RNA is not extruded into the secondary channel and thus cannot serve as a substrate for GreB. Substitutions of the C-domain residues conferred defects in the Gre-assisted cleavage, ranging from an essentially complete lack of cleavage enhancement by the S122E protein to less dramatic phenotypes displayed by others. These results show that even those Gre variants that display moderate defects upon long incubation (*e.g.*, V139D; Supplementary Figure 2) confer significant defects in the *rate* of the nascent RNA cleavage.

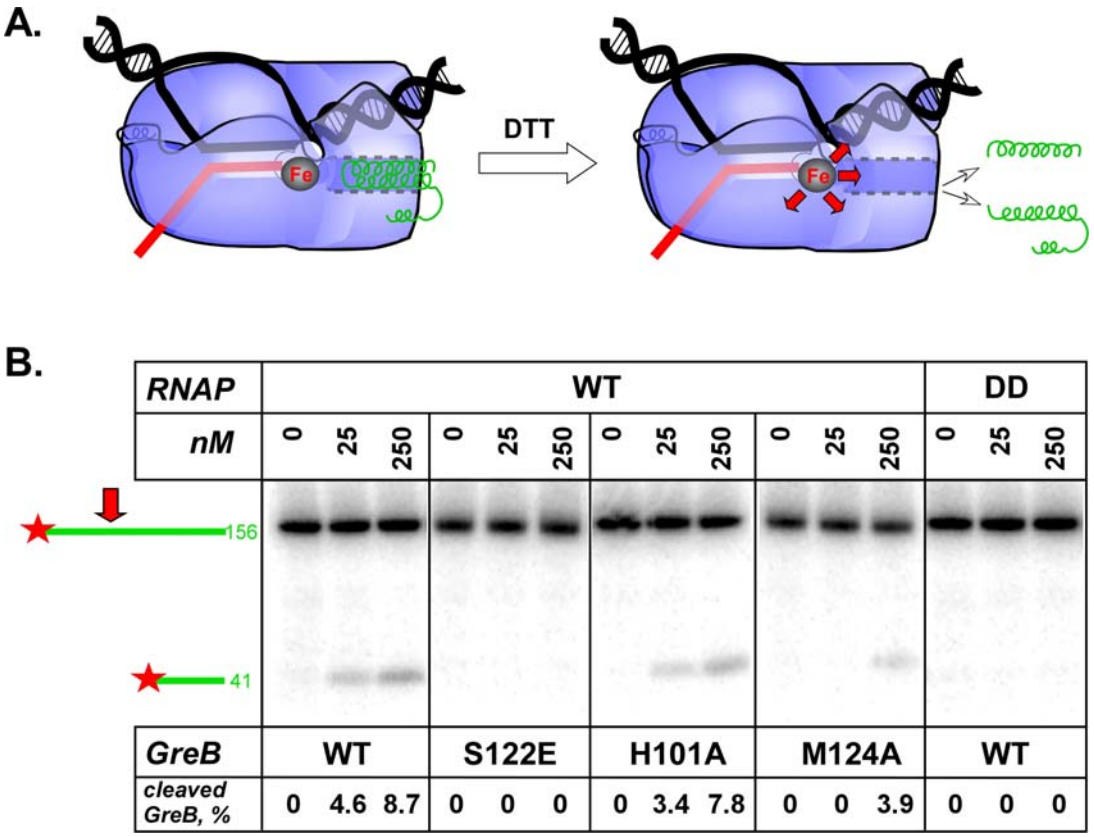

Figure S4 Fe-mediated cleavage assays.

A. Schematic depiction of the assay. Upon addition of DTT, the Fe^{2+} ion bound with a high affinity to the RNAP active site induces hydroxyl-radical cleavage of nearby ligands. When GreB binds to RNAP, its coiled-coil N-terminal domain is extended into the RNAP secondary channel, and the tip of the CC (around residue 41) is cleaved in the presence of iron and DTT.

B. Fe^{2+} mediated cleavage of selected GreB variants. To visualize cleavage products, GreB with an N-terminal HMK tag (RRASV) were ^{32}P-labeled with protein kinase A (NEB). Labeled proteins were purified using G50 spin columns (GE Health). For Fe^{2+} -mediated cleavage assays, RNAP (at 25 or 250 nM, or storage buffer) was mixed with 10 μg of acetylated BSA and $GreB_{HMK}$ (at 100 nM) in 13 μl of 20 mM MOPS, pH7.5, 100 mM NaCl, and incubated at 30 ^{o}C for 30 min. One μl of freshly made 750 μM $(NH_4)_2$ $Fe(SO_4)_2$ and 15 mM DTT each were then added, followed by a 15 min incubation at 30 ^{o}C. Reactions were stopped by addition of 1/3 volume of 4X NuPAGE loading dye, heated for 2 min at 85 ^{o}C and loaded onto a 12% Bis-Tris NuPAGE gel (Invitrogen). After separation, the gels were dried and exposed to a Phosphor Storage screen (GE Health). Positions of the full-length GreB and the N-terminal cleavage product are indicated, the fraction of the N-terminal fragment is indicated below each lane.

This is a qualitative assay that can be used to infer the position of the GreB CC in the secondary channel; the fraction of cleaved GreB is not directly proportional to its concentration; we never observed more than 15% protein cleaved under these conditions.