Supplementary materials of

Shotgun proteomics revealed preferential degradation of misfolded *in vivo* obligate GroE substrates by Lon protease in *Escherichia coli.*

Tatsuya Niwa, Yuhei Chadani, and Hideki Taguchi*

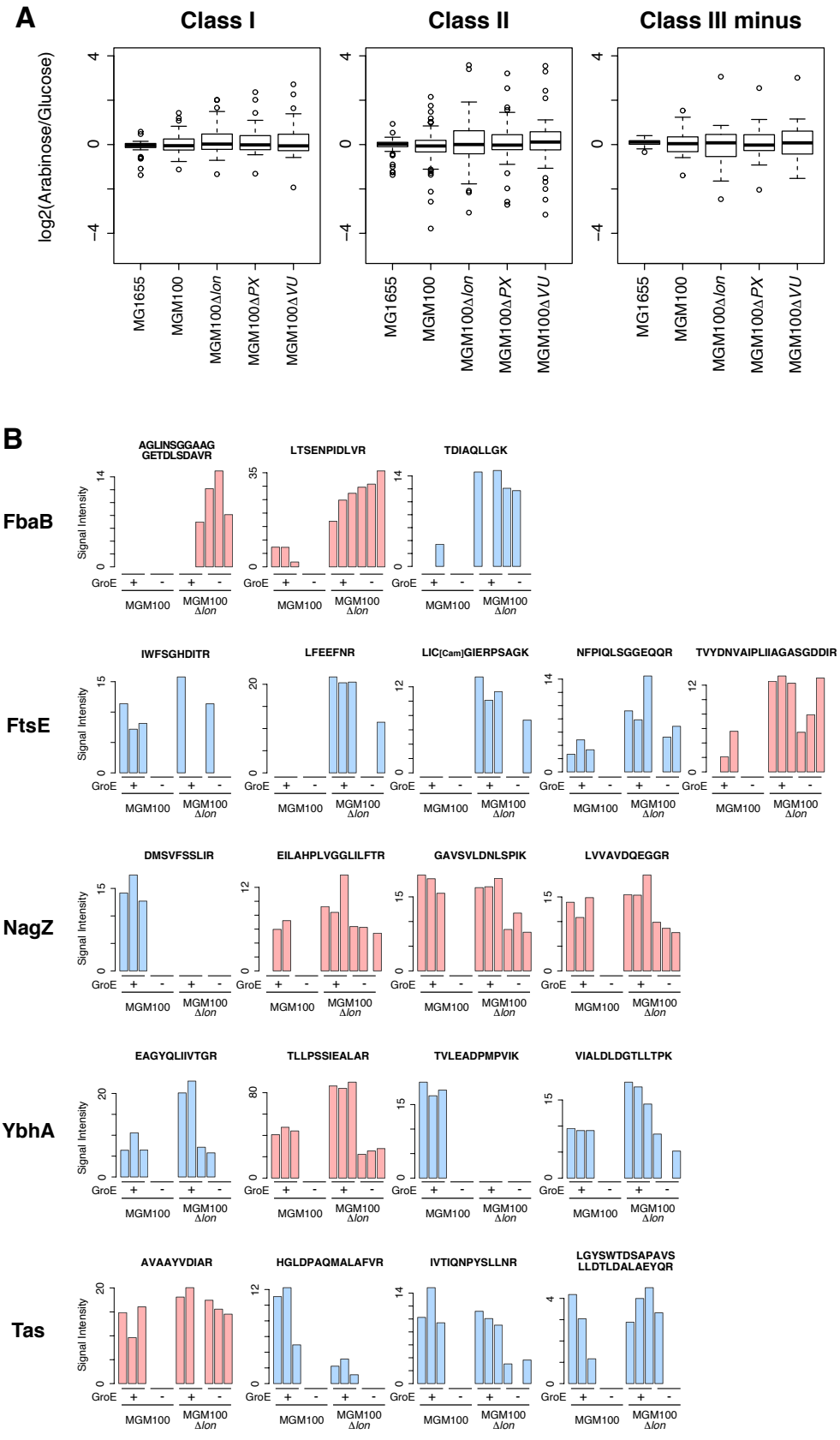
Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan

This document contains the following information:

- Supplementary Figure S1: Additional data of proteome analysis of MGM100 and MGM100-derived strains.
- Supplementary Figure S2: Additional data of proteome analysis of Δ*dnaKJ* and Δ*dnaKJ*Δ*lon* strains.
- Supplementary Table S1: Abundance changes of known Lon substrates in MGM100 and MGM100Δ*lon* strains.
- Supplementary Table S2: *E. coli* strains used in this study.
- Supplementary Table S3: Parameters for the SWATH acquisition.

Supplementary datasets are available as separated excel files:

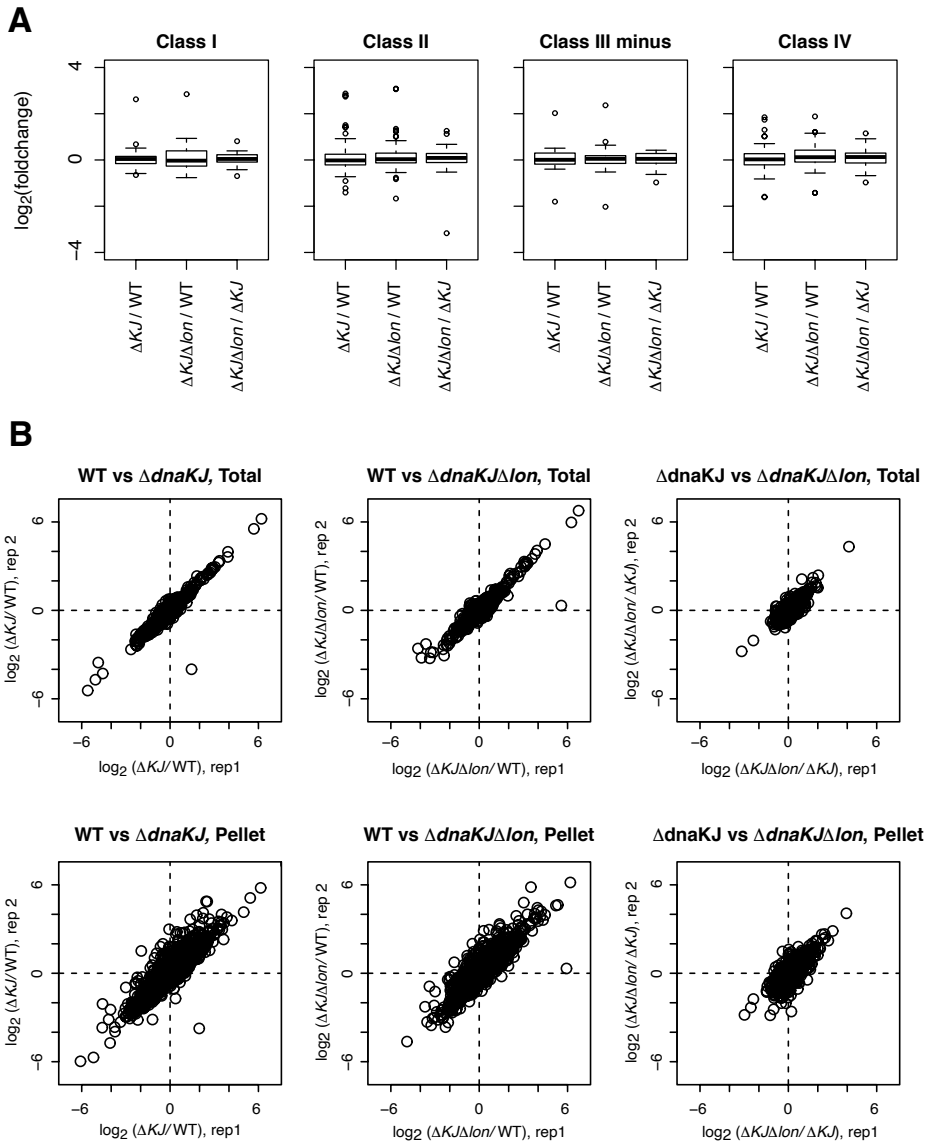
- Supplementary Dataset S1 : Proteome data for the MGM100 and MGM100-derived strains.
- Supplementary Dataset S2 : Proteome data for the Δ*dnaKJ* and Δ*dnaKJ*Δ*lon* strains.
- Supplementary Dataset S3 : List of proteins with cluster numbers.
- Supplementary Dataset S4 : List of proteins up- or down-regulated by the Lon deletion.


A
Class I
Class II
Class III minus
log2(Arabinose/Glucose)
MG1655
MGM100
MGM100Δlon
MGM100ΔPX
MGM100ΔVU
B
FbaB
AGLINSGGAAG GETDLSDAVR
LTSENPIDLVR
TDIAQLLGK
FtsE
IWFSGHDITR
LFEEFNR
LIC[Cam]GIERPSAGK
NFPIQLSGGEQQR
TVYDNVAIPLIIAGASGDDIR
NagZ
DMSVFSSLIR
EILAHPLVGGLILFTR
GAVSVLDNLSPIK
LVVAVDQEGGR
YbhA
EAGYQLIIVTGR
TLLPSSIEALAR
TVLEADPMPVIK
VIALDLDGTLLTPK
Tas
AVAAYVDIAR
HGLDPAQMALAFVR
IVTIQNPYSLLNR
LGYSWTDSAPAVS LLDTLDALAEYQR
Signal Intensity
GroE
MGM100
MGM100 Δlon

Supplementary Figure S1. Additional data of proteome analysis of MGM100 and its derivative strains.

(A) Distribution of the foldchanges of GroE Class I, II, and III minus substrates represented as boxplots (by R software). The center bar in the box represents the median, and the upper and lower portion of the box represent 75% and 25% quantile, respectively. The numbers of the evaluated substrates were ~36 for Class I, ~77 for Class II, and ~29 for Class III minus.

(B) Intensities of the peptides corresponding to the five proteins identified in the GroE-normal cells but not in the GroE-depleted cells in the MGM100 strain (see Table 1). Only the intensities of MGM100 and MGM100Δ*lon* strains were shown. The graphs depicted in red indicate the peptides whose intensities under the GroE depletion condition in MGM100Δ*lon* were detected in all three technical replicates, *i.e.* the peptides assumed to be present under this condition.



Supplementary Figure S2. Additional data of proteome analysis of Δ*dnaKJ* and Δ*dnaKJ*Δ*lon* strains.

(A) Distribution of the foldchanges of GroE Class I, II, III minus, and IV substrates represented as boxplots (by R software). The center bar in the box represents the median, and the upper and lower portion of the box represent 75% and 25% quantile, respectively. The numbers of the evaluated substrates were ~35 for Class I, ~81 for Class II, ~28 for Class III minus, and ~38 for Class IV.

(B) Reproducibility of the proteome changes by the DnaK/DnaJ deletion. Upper panels showed the reproducibility of the fold changes between two biological replicates in the total fraction, and lower panels showed those in the pellet fraction.

Supplementary Table S1. Abundance changes of known Lon substrates in MGM100 and MGM100Δ*lon* strains.

Protein Name	Gene name	Reference	Number of Detection MGM100 Ara	Number of Detection MGM100 Glc	Number of Detection MGM100 Δ*lon* Ara	Number of Detection MGM100 Δ*lon* Glc	Foldchange Glc/Ara, in MGM100	Foldchange Glc/Ara, in MGM100 Δ*lon*	Foldchange MGM100Δ*lon* / MGM100, Ara	Foldchange MGM100Δ*lon* / MGM100, Glc
SulA	*sulA*	*Proc. Natl. Acad. Sci. U S A.* 1983 [1]	0	0	0	0	(N. D.)	(N. D.)	(N. D.)	(N. D.)
RcsA	*rcsA*	*J. Bacteriol* . 1991 [2]	0	0	0	**3**	(N. D.)	(N. D.)	(N. D.)	(N. D.)
λXis	*xis*	*J. Bacteriol.* 1998 [3]	0	0	0	0	(N. D.)	(N. D.)	(N. D.)	(N. D.)
Homoserine succinyl transferase (HTS)	*metA*	*Mol. Microbiol.* 2000 [4]	1	**3**	2	**3**	(N. D.)	(N. D.)	(N. D.)	**2.96**
SoxS	*soxS*	*Mol. Microbiol.* 2004 [5]	0	0	0	0	(N. D.)	(N. D.)	(N. D.)	(N. D.)
MarA	*marA*		0	0	**3**	**3**	(N. D.)	**1.39**	(N. D.)	(N. D.)
RecA	*recA*	*Mol. Cell.* 2006 [6]	**3**	**3**	**3**	**3**	**1.15**	**0.98**	**1.54**	**1.32**
RuvB	*ruvB*		**3**	**3**	**3**	**3**	**0.84**	**1.34**	**1.48**	**2.36**
IbpA	*ibpA*	*Mol. Microbiol.* 2010 [7]	**3**	**3**	**3**	**3**	**2.51**	**2.59**	**1.81**	**1.86**
IbpB	*ibpB*		**3**	**3**	**3**	**3**	**1.62**	**2.65**	**1.3**	**2.13**
CspD	*cspD*	*Mol. Microbiol.* 2011 [8]	**3**	**3**	**3**	**3**	**0.69**	**0.4**	**1.03**	**0.59**
FdoH	*fdoH*	*J. Biol. Chem.* 2012 [9]	**3**	**3**	**3**	2	**0.62**	(N. D.)	**0.78**	(N. D.)
MetR	*metR*	*Proteomics* . 2018 [10]	1	**3**	0	**3**	(N. D.)	(N. D.)	(N. D.)	**1.27**
CysB	*cysB*		**3**	**3**	**3**	**3**	**0.95**	**1.47**	**0.98**	**1.52**
PuuB	*puuB*		0	0	0	0	(N. D.)	(N. D.)	(N. D.)	(N. D.)
YceA	*yceA*		**3**	**3**	**3**	**3**	**0.87**	(N. D.)	**0.64**	(N. D.)
FrsA	*frsA*		**3**	**3**	**3**	**3**	**1.29**	**1.06**	**1.57**	**1.29**
TldD	*tldD*		**3**	1	**3**	1	(N. D.)	(N. D.)	**1.68**	(N. D.)
LipA*	***lipA***		**3**	**3**	**3**	**3**	**0.18**	**0.41**	**1.23**	**2.85**
SrlD	*srlD*		**3**	0	**3**	0	(N. D.)	(N. D.)	**3.77**	(N. D.)
YghW	*yghW*		0	0	0	0	(N. D.)	(N. D.)	(N. D.)	(N. D.)
NuoB	*nuoB*		**3**	2	**3**	**3**	(N. D.)	**0.16**	**0.85**	(N. D.)
FlgE	*flgE*		**3**	0	**3**	0	(N. D.)	(N. D.)	**0.73**	(N. D.)
ProV	*proV*		**3**	**3**	**3**	**3**	**1.17**	**0.71**	**0.91**	**0.55**
YrdA	*yrdA*		**3**	**3**	**3**	**3**	**1.47**	**1.14**	**1.08**	**0.84**
InaA	*inaA*		0	0	**3**	**3**	(N. D.)	**0.71**	(N. D.)	(N. D.)

* *In vivo* obligate GroE substrate

Supplementary Table S2. *E. coli* strains used in this study.

Strain	Description	Source
BW25113	F$^-$, λ$^-$, Δ(*araD-araB*)567, Δ*lacZ*4787(::*rrnB*-3), *rph*-1, Δ(*rhaD-rhaB*)568, *hsdR*514	Laboratory stock
JW0013-KC	BW25113 Δ*dnaK*::FRT-Km^r-FRT	Baba *et al.*, 2006 [11]
JW0014-KC	BW25113 Δ*dnaJ*::FRT-Km^r-FRT	
JW0427-KC	BW25113 Δ*clpP*::FRT-Km^r-FRT	
JW0428-KC	BW25113 Δ*clpX*::FRT-Km^r-FRT	
JW0429-KC	BW25113 Δ*lon*::FRT-Km^r-FRT	
JW3902-KC	BW25113 Δ*hslU*::FRT-Km^r-FRT	
JW3903-KC	BW25113 Δ*hslV*::FRT-Km^r-FRT	
ECY210	BW25113 Δ*lon*::FRT-Cm^r-FRT	This study
ECY262	BW25113 Δ*dnaKJ*::FRT-Km^r-FRT	
ECY289	BW25113 Δ*hslUV*::FRT-Km^r-FRT	
ECY290	BW25113 Δ*clpXP*::FRT-Km^r-FRT	
ECY292	BW25113 Δ*hslUV*::FRT-Cm^r-FRT	
ECY293	BW25113 Δ*clpXP*::FRT-Cm^r-FRT	
MGM100	F$^-$, λ$^-$, *rph-1, araBADp-groE*, *zje-2335*::*kan*	McLennan & Masters, 1999 [12]
MGM100Δ*lon*	MGM100 Δ*lon*::FRT-Cm^r-FRT	This study
MGM100Δ*clpXP*	MGM100 Δ*clpXP*::FRT-Cm^r-FRT	
MGM100Δ*hslUV*	MGM100 Δ*hslUV*::FRT-Cm^r-FRT	
MG1655	F$^-$, λ$^-$, *ilvG*$^-$, *rfb*-50, *rph*-1	Laboratory stock
MG1655Δ*dnaKJ*	MG1655 Δ*dnaKJ*::FRT-Km^r-FRT	This study
MG1655Δ*dnaKJ*Δ*lon*	MG1655 Δ*dnaKJ*::FRT-Km^r-FRT, Δ*lon:*:FRT-Cm^r-FRT	

Supplementary Table S3. Parameters for SWATH acquisition.

	TripleTOF 4600	TripleTOF 6600
Acetonitrile gradient (in the presence of 0.1 % formic acid)	10 - 40 %, 67 min	10 - 30 %, 60 min + 30 - 40 %, 10 min
m/z range of MS1	400 - 1250 Da	350 - 1500 Da
Accumulation time for MS1	100 msec	100 msec
Accumulation time for MS2	60 msec	45 msec
Number of SWATH window	40	50
m/z of SWATH window (variable window optimized for *E. coli* cell lysate)	399.5 - 424.2 423.2 - 445.1 444.1 - 462.1 461.1 - 477.5 476.5 - 491.8 490.8 - 505.6 504.6 - 518.2 517.2 - 530.9 529.9 – 543.0 542.0 - 553.9 552.9 - 564.9 563.9 - 575.9 574.9 - 585.8 584.8 - 596.3 595.3 - 606.7 605.7 - 616.6 615.6 - 627.1 626.1 – 637.0 636.0 - 646.9 645.9 - 657.3 656.3 - 667.8 666.8 - 678.2 677.2 - 688.1 687.1 - 699.1 698.1 - 709.6 708.6 - 720.6 719.6 - 732.1 731.1 - 744.2 743.2 - 756.9 755.9 - 770.1 769.1 - 784.4 783.4 - 800.3	349.5 - 378.7 377.7 - 405.7 404.7 - 425.8 424.8 - 441.9 440.9 - 456.3 455.3 - 470.1 469.1 - 482.7 481.7 - 494.2 493.2 - 505.7 504.7 - 516.7 515.7 - 527.0 526.0 - 537.4 536.4 - 547.2 546.2 - 556.4 555.4 - 565.6 564.6 - 574.8 573.8 - 583.4 582.4 - 592.0 591.0 - 600.6 599.6 - 609.3 608.3 - 617.9 616.9 - 627.1 626.1 - 635.7 634.7 - 644.9 643.9 - 654.1 653.1 - 663.9 662.9 - 674.2 673.2 - 684.0 683.0 - 694.4 693.4 - 705.3 704.3 - 716.2 715.2 - 728.3

	799.3 - 816.8 815.8 – 835.0 834.0 - 854.8 853.8 - 876.8 875.8 - 901.5 900.5 - 929.6 928.6 - 960.9 959.9 - 999.4	727.3 - 740.4 739.4 - 753.0 752.0 - 766.2 765.2 - 780.6 779.6 - 796.1 795.1 - 812.2 811.2 - 829.5 828.5 - 847.9 846.9 - 867.4 866.4 - 887.6 886.6 - 908.8 907.8 - 932.4 931.4 - 958.3 957.3 - 987.0 986.0 - 1022.1 1021.1 - 1070.4 1069.4 - 1126.2 1125.2 - 1199.8

References

1. Mizusawa, S.; Gottesman, S. Protein Degradation in *Escherichia Coli*: The Lon Gene Controls the Stability of SulA Protein. *Proc Natl Acad Sci U S A* **1983**, *80*, 358–362, doi:10.1073/PNAS.80.2.358.
2. Stout, V.; Torres-Cabassa, A.; Maurizi, M.R.; Gutnick, D.; Gottesman, S. RcsA, an Unstable Positive Regulator of Capsular Polysaccharide Synthesis. *J Bacteriol* **1991**, *173*, 1738–1747, doi:10.1128/JB.173.5.1738-1747.1991.
3. Leffers, J.; Gottesman, S. Lambda Xis Degradation in Vivo by Lon and FtsH. *J Bacteriol* **1998**, *180*, 1573–1577, doi:10.1128/JB.180.6.1573-1577.1998.
4. Biran, D.; Gur, E.; Gollan, L.; Ron, E.Z. Control of Methionine Biosynthesis in *Escherichia Coli* by Proteolysis. *Mol Microbiol* **2000**, *37*, 1436–1443, doi:10.1046/J.1365-2958.2000.02097.X.
5. Griffith, K.L.; Shah, I.M.; Wolf, R.E. Proteolytic Degradation of *Escherichia Coli* Transcription Activators SoxS and MarA as the Mechanism for Reversing the Induction of the Superoxide (SoxRS) and Multiple Antibiotic Resistance (Mar) Regulons. *Mol Microbiol* **2004**, *51*, 1801–1816, doi:10.1046/J.1365-2958.2003.03952.X.
6. Neher, S.B.; Villén, J.; Oakes, E.C.; Bakalarski, C.E.; Sauer, R.T.; Gygi, S.P.; Baker, T.A. Proteomic Profiling of ClpXP Substrates after DNA Damage Reveals Extensive Instability within SOS Regulon. *Mol Cell* **2006**, *22*, 193–204, doi:10.1016/J.MOLCEL.2006.03.007.
7. Bissonnette, S.A.; Rivera-Rivera, I.; Sauer, R.T.; Baker, T.A. The IbpA and IbpB Small Heat-Shock Proteins Are Substrates of the AAA+ Lon Protease. *Mol Microbiol* **2010**, *75*, 1539–1549, doi:10.1111/J.1365-2958.2010.07070.X.
8. Langklotz, S.; Narberhaus, F. The *Escherichia Coli* Replication Inhibitor CspD Is Subject to Growth-Regulated Degradation by the Lon Protease. *Mol Microbiol* **2011**, *80*, 1313–1325, doi:10.1111/J.1365-2958.2011.07646.X.
9. Westphal, K.; Langklotz, S.; Thomanek, N.; Narberhaus, F. A Trapping Approach Reveals Novel Substrates and Physiological Functions of the Essential Protease FtsH in *Escherichia Coli*. *J Biol Chem* **2012**, *287*, 42962–42971, doi:10.1074/JBC.M112.388470.
10. Arends, J.; Griego, M.; Thomanek, N.; Lindemann, C.; Kutscher, B.; Meyer, H.E.; Narberhaus, F. An Integrated Proteomic Approach Uncovers Novel Substrates and Functions of the Lon Protease in *Escherichia Coli*. *Proteomics* **2018**, *18*, doi:10.1002/PMIC.201800080.
11. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.A.; Tomita, M.; Wanner, B.L.; Mori, H. Construction of *Escherichia Coli* K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection. *Mol Syst Biol* **2006**, *2*, doi:10.1038/MSB4100050.
12. McLennan, N.; Masters, M. GroE Is Vital for Cell-Wall Synthesis. *Nature* **1998**, *392*, 139, doi:10.1038/32317.