REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
11-04-2017	Final Report	1-Sep-2013 - 28-Feb-2017

4. TITLE AND SUBTITLE
Final Report: A Novel Stress-Induced Translation Machinery as a Unique Mechanism for bacterial cell death and survival

5a. CONTRACT NUMBER
W911NF-13-1-0371

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Hanna Engelberg-Kulka

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Hebrew University of Jerusalem
Authority for R&D
Edmond J Safra Campus, Givat Ram
-

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
63696-LS.8

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
E. coli mazEF is a stress-induced toxin-antitoxin module. MazF is a stable endoribonuclease that cleaves at ACA sites in RNAs. Here we further studied the Stress-induced Translation Machinery (STM) generated by MazF under stress. STM is composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs. We found that MazF cleaves only ACA sites located in the open reading frames of processed mRNAs, while out-of-frame ACAs are resistant. This in-frame ACA cleavage of MazF depends on MazF binding to an Extracellular Death Factor (EDF)-like element in ribosomal protein bS1 (bacterial S1), apparently causing

15. SUBJECT TERMS
E. coli, mazEF, MazF, stress-induced toxin, antitoxin, endoribonuclease, RNA, stress-induced Translation Machinery

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Hanna Engelberg-Kulka
UU	UU	UU	UU		19b. TELEPHONE NUMBER 972-267-5825

Report Title

Final Report: A Novel Stress-Induced Translation Machinery as a Unique Mechanism for bacterial cell death and survival

ABSTRACT

E. coli mazEF is a stress-induced toxin-antitoxin module. MazF is a stable endoribonuclease that cleaves at ACA sites in RNAs. Here we further studied the Stress-induced Translation Machinery (STM) generated by MazF under stress. STM is composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs. We found that MazF cleaves only ACA sites located in the open reading frames of processed mRNAs, while out-of-frame ACAs are resistant. This in-frame ACA cleavage of MazF depends on MazF binding to an Extracellular Death Factor (EDF)-like element in ribosomal protein bS1 (bacterial S1), apparently causing MazF to be part of STM ribosomes. Furthermore, due to the in-frame MazF cleavage of ACAs, under stress, a bias occurs in the reading of the genetic code causing the amino acid threonine to be encoded only by its synonym codons ACC, ACU, or ACG (which are resistant to MazF), instead of by ACA.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received		Paper
04/04/2017	6	Adi Oron-Gottesman, Martina Sauert, Isabella Moll, Hanna Engelberg-Kulka. A Stress-Induced Bias in the Reading of the Genetic Code in Escherichia coli, mBio, ( ): 1. doi: 1,038,778.00
04/10/2017	9	Sathish Kumar, Hanna Engelberg-Kulka. Quorum Sensing Peptides Mediating interspecies Bacterial Cell Death as a Novel Class of Antimicrobial Agents, Current Opinion in Microbiology, ( ): 22. doi: 1,038,976.00
07/24/2014	1	Ariel Erental, Ziva Kalderon, Ann Saada, Yoav Smith and Hanna Engelberg-Kulka. Apoptotic-Like Death (ALD): An Extreme SOS Response in Escherichia coli, mBio, (06 2014): 1. doi: 327,576.00
08/05/2015	2	Hanna Engelberg-Kulka, Sathish Kumar. Yet another way that phage Lamda manipulates Escherichia coli host: Lamda-rexB is involved in the lysogenic - lytic switch , Molecular Microbiology, (05 2015): 1. doi: 10.1111/mmi.12969 361,885.00
08/05/2015	3	Sathish Kumar, Hanna Engelberg-Kulka, Finbarr Hayes, Ziva Kalderon. The SOS Response is Permitted in Escherichia coli Strains Deficient in the Expression of the mazEF Pathway, PLoS ONE, (12 2014): 1. doi: 10.1371/journal.pone.0114380 361,893.00
08/18/2016	4	Sathish Kumar, Ilana Kolodkin-Gal, Oliver Vesper, Nawsad Alam, Ora Schueler-Furman, Isabella Moll, Hanna Engelberg-Kulka. Escherichia coli Quorum-Sensing EDF, A Peptide Generated by Novel Multiple Distinct Mechanisms and Regulated by trans-Translation, mBio, ( ): 1. doi: 1,014,662.00
TOTAL:	**6**	

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

I. A lecture in the 2017 Meeting of Molecular Genetics of Bacteria and Phage.
University of Wisconsin-Madison, USA
The lecture: Prof. Hanna Engelberg-Kulka,
II. A seminar in Bar-ilan University, Tel-Aviv, Israel
The lecture: Prof. Hanna Engelberg-Kulka
III. A seminar in Hebrew University-Medical School, Israel.
The lecture: Prof. Hanna Engelberg-Kulka
IV. The lecture on PhD thesis, in student seminar Hebrew University-Medical School, Israel.
The lecture: Adi Oron-Gottesman

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper	
11/01/2016	5	Hanna Engelberg-Kulka. Escherichia coli Quorum-Sensing EDF, a Peptide Generated by, Conference of Molecular Genetics of Bacteria and Phages. 08-AUG-16, Madison, Wisconsin. : ,

TOTAL: **1**

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received Paper

TOTAL:

Number of Manuscripts:

Books

Received Book

TOTAL:

Received	Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

2014: Honorarium PhD to the PI of this grant, Professor Hanna Engelberg-Kulka from the University of Vienna for her contribution to the Vienna University and to Science in general

Graduate Students

NAME	PERCENT_SUPPORTED	DISCIPLINE
Sathish Kumar	100	Microbiology
Adi Oron-Gottesman	100	Microbiology
FTE Equivalent:	**2.00**	
Total Number:	**2**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	DISCIPLINE
Kati Hakim	100	Microbiology
Miri Vinogradov	100	Microbiology
FTE Equivalent:	**2.00**	
Total Number:	**2**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 2.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 2.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:..... 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME
Sathish Kumar
Adi Oron-Gottesman
Total Number: **2**

Names of other research staff

NAME PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

E. coli mazEF is a stress-induced toxin-antitoxin module. MazF is a stable endoribonuclease that cleaves at ACA sites in RNAs. Here we further studied the Stress-induced Translation Machinery (STM) generated by MazF under stress. STM is composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs. Here, we further studied the E. coli STM system by constructing a Green Fluorescent Protein (GFP) reporter molecule consisting of a leaderless GFP mRNA that is expressed upon the induction of MazF. Surprisingly, GFP was expressed in spite of the existence of 17 ACA sites in the GFP leaderless mRNA. However we noticed that all of them are located out of the open reading frame of GFP. In contrast, inserting an ACA site in the open reading frame of GFP prevented its expression after mazF induction. We also showed, by direct experiments, that MazF cleaves ACA sites only when they are located in the open reading frame (to be called here frame 0) of the leaderless mRNAs; they were never cleaved when they were located out-of-frame (designated here frame +1, and frame +2). In addition, the in-frame MazF cleavage of leaderless GFP mRNA was dependent on MazF binding to NNW, an Extracellular Death Factor (EDF)-like sequence in the ribosomal protein bS1 (bacterial S1). EDF is the E.coli extracellular quorum sensing penta-peptide NNWNN which is mediating bacterial cell death by inducing MazF.
The genetic code is a universal characteristic of all living organisms. It defines the set of rules by which nucleotide triplets specify which amino acid will be incorporated into a protein. Our results represent the first existing report on a stress- induced bias in the reading of the genetic code. We found that In E.coli, under stress, the amino acid threonine is encoded only by its synonym codons ACC, ACU, or ACG (which are resistant to MazF), instead of by ACA. This because under stress. MazF generates a Stress induced Translation Machinery (STM) in which MazF cleaves in-frame ACA sites of the processed mRNAs. The results of these study were published in Oron-Gottesman,A., Sauert, M., Moll, I., and Engelberg-Kulka,H. Stress-induced bias in the reading of the genetic code in Escherichia coli. Mbio 7(6):e01855, 2016.

Technology Transfer

Final Progress Report Engelberg-Kulka,H

ABSTRACT

E. coli mazEF is a stress-induced toxin-antitoxin module. MazF is a stable endoribonuclease that cleaves at ACA sites in RNAs. Here we further studied the Stress-induced Translation Machinery (STM) generated by MazF under stress. STM is composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs. We found that MazF cleaves only ACA sites located in the open reading frames of processed mRNAs, while out-of-frame ACAs are resistant. This in-frame ACA cleavage of MazF depends on MazF binding to an Extracellular Death Factor (EDF)-like element in ribosomal protein bS1 (bacterial S1), apparently causing MazF to be part of STM ribosomes. Furthermore, due to the in-frame MazF cleavage of ACAs, under stress, a bias occurs in the reading of the genetic code causing the amino acid threonine to be encoded only by its synonym codons ACC, ACU, or ACG (which are resistant to MazF), instead of by ACA.

Such a seminal discovery is not an every day event. It is therefore not surprising that our results were published in the prestigious journal mBio. When the manuscript was accepted for publications the Editor wrote us : "Thank you for submitting us such an outstanding research".

Introduction

E. coli mazEF is a toxin -antitoxin system that was discovered by us as being responsible for Programmed Cell Death (PCD) (Aizenman et al., 1996), and is since extensively studied by us (about 30 articles) and by others. Among other mechanisms and cellular phenomena related to MazF, we have previously detected that: a) MazF cleaves at ACA sites at or closely upstream to AUG-start codons of specific mRNAs, and thereby generating leaderless mRNAs belonging to a novel "Leaderless regulon"; and b) MazF targets the 16S rRNA within the 30S ribosomal subunit at the decoding centre, thereby removing 43 nucleotides from the 3'-terminus. Since these 43 nucleotides include the anti-SD (Shine-Dalgarno) region, these deficient ribosomes, that we call "stress ribosomes", are selectively able to translate the generated leaderless mRNAs (Vesper et al., 2011; Moll and Engelberg-Kulka, 2012). Thus, under stressful conditions, MazF is induced, which leads to the generation of a novel "leaderless regulon" that is translated by the novel "stress ribosomes", producing a Stress-induced Translation Machinery (STM).

Here we further studied STM and discovered a stress-induced hidden secret of the genetic code generated in *E.coli* .

Specific Aims

We undertook the following main direction:
We asked:
Are ACAs located in the STM translated leaderless mRNAs resistant to MazF cleavage?

RESULTS

GFP reporter systems of the Stress-induced Translation Machinery (STM) and their expression dependency on MazF induction

To study the STM, we constructed a GFP reporter molecule: the first ATG of *gfp* sequence was preceded by AC, generating an ACATG sequence that would potentially enable MazF to cleave at ACA, thus generating a leaderless GFP mRNA. In addition, this ACATG sequence is preceded by a stem and loop structure that interfered with the Shine-Delgarno (SD) recognition sequence (Fig. 1G). Furthermore, since *gfp* has 17 ACA sites (Fig. 1B), we modified all the ACA sites without changing the original amino acid (Fig. 1A). We inserted this reporter molecule into plasmid pUH-C, which we used to transform *E. coli* MG1655 (WT) or its derivative MG1655Δ*mazEF*. At logarithmic phase, we induced *mazF* by adding Nalidixic Acid (NA). Adding NA led to a significant increase in GFP expression in the WT strain MG1655 (Figs 1C and F2), but not in MG1655Δ*mazEF* (Figs 1E and F4), confirming that our constructed GFP molecule was indeed a reporter for MazF-induced STM. The induced MazF produced a leaderless GFP mRNA, and also generated deficient ribosomes that lacked the last 43 nucleotides of the 16S rRNA, including the anti-SD sequence (Vesper et al., 2011).

We also studied MazF-dependent *gfp* expression using a STM reporter carrying the WT *gfp* sequence harboring 17 out-of frame ACA sites (Fig. 1B). We expected that the presence of these ACA sites would cause the *gfp* leaderless mRNA to be cleaved by MazF, induced by the addition of NA, leading to less expression of *gfp* than in the untreated, control culture. We were surprised to observe that in this case the level of GFP was higher than the untreated control culture (Fig. 1D). Furthermore, similar levels of GFP were obtained by *mazF* induction of the two STM reporter systems. One, containing the WT *gfp* sequence with ACAs (Fig 1D) and the other containing the *gfp* sequence with no ACAs (Figs 1C and F2). Using the strains MG1655 and MG1655Δ*mazEF*, we found that the expression of the WT *gfp* reporter was dependent on *mazF* induction because there was no increase in GFP levels in strain MG1655Δ*mazEF* (data not shown).

ACA sites located in frame 0 of the *gfp* sequence interferes with MazF-induced STM-GFP expression

Finding that, in spite of the presence of 17 ACA sites, the WT GFP reporter was resistant to MazF cleavage led us to inspect these sites more carefully: each of the 17 ACA sites was located in the +1 frame and not one was in frame 0. In response, we asked: Would MazF act similarly if the ACA sites were located in frame 0 of the *gfp* sequence in the STM reporter? To this aim, into the *gfp* sequence of our STM reporter (that harbor no ACAs, Fig. 1A), we inserted an ACA in each of five different frame 0 sites. We selected five locations (Fig 2A circled 1-5) in which we modified the sequence so that an ACA, now in frame 0, would be adjacent to one of the frame +1 ACAs of the WT *gfp* sequence. We found that, for each of these artificially introduced frame 0 STM reporters the level of GFP expression was reduced in the WT strain (Fig. 2Ca-e) while unaffected in the Δ*mazEF* derivative (Fig S2). In contrast, when we used a STM-GFP reporter with a *gfp* sequence that has no ACAs, the addition of NA led to an increase in the level of MazF-mediated GFP expression (Fig. 2A). By quantitative comparison, we clearly show that rather than an increase of about 50% of the MazF-mediated expression of the STM reporter with no ACAs (Fig. 2DB), there was a decrease (in the range of 10-30%) in the MazF-

mediated expression of the STM reporters carrying an ACA site in each of the different five frame 0 locations (Fig. 2Da-e). When we used an STM reporter without any ACA sites, MazF induction led to a 50% (about) increase of GFP expression; in contrast, when we used an STM reporter that included one ACA in frame 0 of the *gfp* sequence, MazF induction led to a decrease in GFP expression. Thus, we found that the presence of a single, in-frame ACA triplet caused an additive effect of an about 60-80% decrease in GFP expression (compare the first red column with each of the other red columns in Fig. 2D). We understand that this decrease in STM-GFP expression was caused by MazF-cleaving those individual ACA sites located in frame 0 of the *gfp* ORF in the STM system**.**

Direct determination of a MazF cleavage at in-frame ACA sites of a leaderless mRNA

To confirm our indirect results using STM-GFP reporters, we developed a method to directly determine MazF cleavage at in-frame ACA sites in a GFP leaderless mRNA. In this method we extracted RNA from MazF induced and uninduced MG1655 (WT) cells harboring either (i) a STM-GFP reporter with only one in-frame ACA site (see location 1 in Fig. 2A), or (ii) the WT GFP reporter that carries 17 out of frame ACA sites. Using the extracted RNAs, we prepared corresponding cDNA samples which we amplified by a PCR reaction. In this reaction, we determined the site of MazF cleavage using two different forward primers (PF Long and PF Short) and one reverse primer (PR) that we designed to start from the end of the GFP reporter sequence (Fig. 3A). We designed PF Long to start from the beginning of the sequence of the leaderless GFP reporters, and PF Short to start directly after the in-frame ACA site that we generated in location 1 (Fig. 2A). We hypothesized that, with the addition of NA to induce MazF expression, if MazF were to cleave at this in-frame ACA site of the leaderless STM-GFP mRNA, using PF Long would not lead to a RT-PCR reaction. Indeed, we observed almost no RT-PCR product when the MazF-induced sample was amplified by PF Long (Fig. 3B lane 4). In contrast, this RT-PCR product was observed in the absence of NA, when no MazF expression was induced (Fig. 3B lane 2). We suggest that the minimal amount of RT-PCR product seen in the *mazF*-induced sample using PF Long (Fig. 3B lane 4) probably represents incomplete MazF cleavage at the in-frame ACA site of the reporters. To control for the quality and integrity of the RNA samples used after *mazF*- induction, we used PF Short, designed to start immediately downstream from the in-frame ACA site. Using PF Short, we obtained similar amounts of RT-PCR products in MazF-induced (Fig. 3B lane 3) and un-induced samples (Fig. 3B, lane 1). Note that similar results to these obtained here (Fig. 3B) were also obtained while using a leaderless STM-GFP reporter harboring an in-frame ACA site in location 2 (data not shown).

To support the evidence of MazF cleavage at the in-frame ACA cutting site, we designed an additional, close forward primer (PF Close) that ends immediately before the ACA site, upstream from the MazF cleavage site (Fig. 3C). Using the STM-GFP reporter with the in-frame ACA site in location 1, as with PF Long, with PF Close, we expected to observe a decrease in the PCR product after the addition of NA. Indeed, we found much less expression after the addition of NA (Fig. 3D lane 2) than without the addition of NA (Fig. 3D lane 1), indicating that when the endoribonuclease MazF was expressed, it cut at the in-frame ACA site. Note that when we use close primers that were designed right upstream to the ACA cutting site, we see a decrease in band intensity . However, when we use the short forward primer designed right downstream to this cutting site (as describes in Fig. 3A), we

did not observed this decrease. Thus, we were able to zoom into the ACA cutting site, and to confirm that indeed MazF cleaves at this in-frame ACA site.

Using this same technique, with the WT GFP reporter carrying 17 out of frame ACA sites (Fig. 1B), we also confirmed that out-of-frame ACAs were resistant to cleavage by MazF. Though (+1-frame) ACA sites were present, when we used PF Long, inducing MazF did not lead to a decrease in the amount of the PCR product obtained (Fig. 3B, compare lanes 6 and 8). Since all the ACA sites in the WT GFP reporter were in reading frame +1, we asked if ACA sites in a +2 frame would also be resistant to MazF cleavage. We constructed two different STM-GFP reporters carrying an ACA site in frame +2 (data not shown). When we used PF Long, and after MazF induction, as we found for +1-frame ACAs, when we used +2-frame ACAs, we observed no reduction in the amount of the RT PCR product obtained.

Together, our results confirmed that in the leaderless GFP mRNA, MazF did not cleave ACA when they were in +1-frame or +2-frame, but only when were in frame 0.

The EDF-like element in S1 is involved in the MazF in-frame ACA cleavage of the STM-GFP reporter

The Extracellular Death Factor (EDF), pentapeptide NNWNN, binds to MazF and is involved in its activity (Belitsky, et al 2010). We were surprised to find NNW, an EDF-like sequence, in the C-terminal domain of the ribosomal protein bS1. Previously, we showed that MazF binds to bS1 through the NNW sequence, and that a W->**A** mutation in this sequence prevents the binding of MazF to bS1 (Kumar et al, submitted). Since here, we found that cleavage by MazF was dependent on an in-frame ACA, we wondered if there might be a connection to MazF binding to bS1. We asked: Will a W->A mutation in the EDF-like element of bS1 prevent in-frame ACA MazF cleavage? First, we studied the effect of this mutation on the expression of the STM-GFP reporter, carrying an in-frame ACA site in location 1 (Fig. 2A). We found that in MG1655 WT cells harboring the multi-copy pACYC plasmid carrying the bS1 gene (*rpsA*) with a W->A mutation, the expression of GFP was increased by about 35% after MazF induction (Fig. 4Aa and Fig 4Ab), this, in spite of the existence of an in-frame ACA site. Moreover, these results were unlike those of a similar experiment we performed in which, instead of using the mutant bS1, we used the pACYC plasmid carrying the gene encoding for WT bS1. In our experiments here, after MazF induction, GFP expression was severely reduced, by about 75% (Fig. 4Ab and Fig. 4Bb). Thus, our quantitative analysis comparing the results with bS1 WT and mutant bS1 revealed that, in the presence of the WT bS1, the induction of MazF led to an additive reduction of GFP expression of about 110% (Fig. 4B, compare red columns). We were able to support the role of the EDF-like element of bS1 in the reduction of the expression of the STM-GFP reporter after MazF induction by the results of additional experiments (data not shown) in which we used four other in-frame ACA sites, generated each in a different location of the *gfp* sequence (Fig. 2A).

Finally, we confirmed the involvement of the EDF-like element of bS1 in the MazF in-frame ACA cleavage using the method that we developed for directly determining ACA cleavage. Once again, we used the STM-GFP reporter with one in-frame ACA site in location 1 (Fig. 2A). Recall that, in this assay, if an ACA site would be cleaved, the PCR product would not be obtained by the use of the PF Long (Fig. 3). We observed the absence of this PCR product only in *mazF* induced cells harboring a pACYC plasmid carrying the bS1 WT (Fig. 4C, lane 4). On the other

hand, this PCR product was obtained without MazF induction (Fig. 4C lane 2) and also in a MazF-induced sample in cells harboring a pACYC plasmid carrying the mutant bS1 (Fig 4C lane 8). This is indicating that MazF does not cleave at the in-frame ACA site in the presence of a mutant bS1.

Our combined results suggest that, in STM, the EDF-like element in bS1 was involved in the MazF in-frame ACA cleavage, probably because, through bS1, MazF becomes a part of the ribosome in the stress induced translation machinery.

In leaderless mRNAs to be translated by the STM system, all the ACA triplets are located out-of-frame

Previously, we characterized MazF-induced small (below 20 kDa) stress proteins that would be translated by STM (Amitai et al, 2009). Among these were EF-P, DeoC, SoxS, RbfA, and AhpC. Here, we found that ACA sites located out-of frame within leaderless mRNAs are resistant to MazF cleavage. So, we asked: In the leaderless mRNAs, are the ACA triplets that specify for these proteins located in-frame or out-of-frame? We found that all of the ACA triplets were situated out-of-frame in *efp* (Fig 5A),*deoC* (Fig. 5B), *soxS*, *rbfA*, and *ahpC* mRNAs (data not shown). Note that in a few cases, as for yfiD and yfbU, we did find some ACAs located in-frame (data not shown). We hypothesize that these represent minor examples, which might point to the existence of an additional mechanism(s) that may resist cleavage by MazF.

Moreover, we also observed another characteristic related to the genetic code. It is well known that ACA is encoding for the amino acid threonine. Since, the presence of ACAs in the open reading frame of leaderless mRNA is not permitting its translation under stress induced MazF, threonine is encoded by the synonym codons ACC, ACU and ACG which are resistant to MazF cleavage (Zhang et al.,2003) (Fig. 5A-D).

Furthermore, our recent studies, in which we identified the *E. coli* MazF leaderless regulon (Sauert et al. 2016), permitted us, here, to identify mRNAs encoding for larger proteins, including *rpsA*, that encodes the ribosomal bS1 protein (Fig. 5D) and *groEL* (Fig. 5C). As can be seen *rpsA* has 21 out of frame ACA sites, and *groEL* has 13 out of frame ACA sites. Furthermore, in both cases, the synonym threonine codons ACC, ACU and ACG are located in frame 0 of the open reading frame. The *rpsA* mRNA carries 12 ACCs, 12ACUs, and one ACG (Fig. 5D) and the *groEL* mRNA carries 25 ACCs, and 8 ACUs (Fig. 5C).

We studied the case of *groEL* in depth, finding that, when we induced MazF by adding NA, inserting even one in-frame ACA site into the *groEL* sequence caused a reduction in GroEL translation. As can be seen in Fig. 6, when we studied *groEL* expression after MazF induction by NA in the MG1655 (WT) strain, having introduced an in-frame ACA site in any one of three different locations led to reduced expression of *groEL* (Fig. 6B). We observed no such reduction in the MG1655Δ*mazEF* derivative strain (Fig. 6B).

CONCLUSIONS

Here, we further studied the *E. coli* STM system by constructing a Green Fluorescent Protein (GFP) reporter molecule consisting of a leaderless GFP mRNA that is expressed upon the induction of MazF. Surprisingly, GFP was expressed in spite of the existence of 17 ACA sites in the GFP leaderless mRNA. However we noticed that all of them are located out of the open reading frame of GFP. In

contrast, inserting an ACA site in the open reading frame of GFP prevented its expression after *mazF* induction. We also showed, by direct experiments, that MazF cleaves ACA sites only when they are located in the open reading frame (to be called here frame 0) of the leaderless mRNAs; they were never cleaved when they were located out-of-frame (designated here frame +1, and frame +2). In addition, the in-frame MazF cleavage of leaderless GFP mRNA was dependent on MazF binding to NNW, an *Extracellular Death Factor* (EDF)-like sequence in the ribosomal protein bS1 (bacterial S1). EDF is the *E.coli* extracellular quorum sensing penta-peptide NNWNN which is mediating bacterial cell death by inducing MazF (Kolodkin-Gal et al., 2007). In the first part of this research report we uncovered the mechanism of its formation.

The genetic code is a universal characteristic of all living organisms. It defines the set of rules by which nucleotide triplets specify which amino acid will be incorporated into a protein. Our results represent the first existing report on a stress-induced bias in the reading of the genetic code. We found that In *E.coli*, under stress, the amino acid threonine is encoded only by its synonym codons ACC, ACU , or ACG, instead of by ACA. This because under stress**.** MazF generates a Stress induced Translation Machinery (STM) in which MazF cleaves in-frame ACA sites of the processed mRNAs.

The results of these study were published in Oron-Gottesman,A., Sauert, M., Moll, I., and Engelberg-Kulka,H. Stress-induced bias in the reading of the genetic code in Escherichia coli**.** Mbio 7(6):e01855, 2016 (See attached).

References

1) Aizenman, E., Engelberg-Kulka, H. and Glaser, G. An *Escherichia coli* chromosomal "addiction module" regulated by ppGpp: A model for programmed cell death. PNAS 93:6059-6063, **1996.**
2) Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A., and Engelberg Kulka,H. (2009). *Escherichia coli* MazF leads to the simultaneous selective synthesis of both"death proteins" and "survival proteins". PLoS Genet. *5*, e1000390.
3) Belitsky, M., Avshalom, H., Erental, A., Yelin, I., Kumar, S., London, S., Sperber, M., Schueler-Furman, O., and Engelberg-Kulka, H.(2011). The *Escherichia coli* Extracellular Death Factor EDF induces the Endoribonucleolytic Activities of MazF and ChpBK. Mol. Cell *41* 625-635, 2011.
4) Kolodkin-Gal, I., Hazan, R., Gaathon, A., Carmeli, S., and Engelberg-Kulka, H. (2007). A linear penta-peptide is a quorum sensing factor required for mazEF-mediated cell death in *Escherichia coli*. *Science* 318, 652-655.
5) Sauert, M., Wolfinger, M, Vesper, O., Byrgazov, K., Moll, I. (2016). The MazF-regulon: a toolbox for the post-transcriptional stress response in *Escherichia coli*. Nucl. Acids Res. doi: 10.1093/nar/gkw115.
6) Vesper, O., Amitai, S., Belitsky, M., Kaberdina, A.,C., Engelberg-Kulka, H., and Moll, I. (2011). Selective translation of leaderless mRNAs by specialized-ribosomes generated by MazF in *Escherichia coli*. Cell, *147*, 1-11.
7) Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Quing, G., and Inouye, M. (2003). MazF cleaves cellular mRNA specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell *12*, 913-923.

Figure Legends

Figure 1. Construction of GFP STM reporters and the dependency of their GFP expression on MazF induction.
(A) *gfp* sequence in which ACA sites were here changed by us without changing the original amino acid (highlighted in blue). **(B**) Sequence of WT *gfp* (ACA sites are highlighted in blue). (**C)** GFP expression in *E. coli* strain MG1655 transformed with plasmid pUH-C carrying a GFP-STM reporter with no ACA's in the *gfp* sequence. FU stands for fluorescence units, **(D)** As in (C), but the WT GFP-STM reporter includes ACAs. **(E)** As in (C), but in derivative strain MG1655Δ*mazEF*. These results shown are the averages of the results from three repeated experiments for each condition. **(F)** Quantitative comparison of GFP expression in MazF induced samples (red columns) versus uninduced samples (blue columns). These data were calculated as percentages based on the results shown in C, D, and E. For each assay, 100% represents the results for the untreated sample. **(G**) Schematic presentation of the GFP-STM reporter molecule that we constructed. Immediately after the first ATG, the GFP sequence is either the WT *gfp* that includes ACAs (B), or the sequence that we modified so that it no longer included any ACAs (A).

Figure 2. Introducing ACAs in reading frame 0 of STM-GFP reporter led to reduced MazF-induced GFP expression.
(A) Locations of five different in-frame ACA sites in the *gfp* sequence of the STM reporter. We used the STM-GFP reporter with no ACA sites as a platform (Fig 1A). Each ACA was inserted to a different location, generating only one ACA site in frame 0. The modified triplets are highlighted in yellow and indicated by circles with the numbers 1 to 5. The various sites of the triplets modified to be ACA were selected to be in place of or adjacent to the original out-of frame ACA sites (highlighted in blue in the WT *gfp* sequence (Fig 1B) that we had originally modified (Fig 1A). **(B)** The expression of the GFP-STM reporter with no ACAs in the *gfp* sequence (as in Fig. 1B). **(Ca-e)** The expression of the GFP-STM reporter with one of the five ACA sites at five different locations in the 0 frame of the *gfp* sequence (**a-e** correspond to locations 1-5 as described in A). The assays were carried out as described in Fig 1. **(D**) A quantitative comparison (in percentage) of the results in (B) and (Ca-e) was carried out as described in Fig 1F.

Figure 3. A molecular approach to study MazF cleavage at in-frame ACAs of the leaderless mRNA GFP reporters.
(A) Illustration of the locations of primers designed for PCR amplification. The green line represents the STM-GFP reporter sequence including one in-frame ACA site. The "long primer forward" is marked "PF long" (blue arrow). The "short primer forward" is marked "PF short" (red arrow). The "reverse primer" is marked PR (blue arrow). **(B)** Agarose gel with samples of PCR products of GFP reporters. Lane 1: the in- frame ACA GFP reporter with PCR amplification by a short primer. Lane 2 as in land 1, but PCR amplification by a long primer. Lane 3: as in 1 but with the addition of NA. Lane 4: as in 2 but with the addition of NA. Lanes 5-6: as in lanes 1-4 but with the WT GFP reporter. **(C)** As in (A), but with a different forward primer; the close forward primer is marked close PF (purple arrow). **(D**) As in (B) but the GFP reporter has an in-frame ACA in location 1. Lane 1: PCR amplification with a close primer. Lane 2: as in lane 1 but with the addition of NA.

Figure 4. In GFP-STM reporter, MazF-induced cleavage of in-frame ACA is dependent on EDF-like sequence in bS1.
(Aa) The effect of MazF induction on the expression of the GFP-STM reporter in cells harboring a plasmid carrying the bS1 mutation W->A in the EDF-like sequence (mutation W444A). **(Ab**) As in (Aa) but in cells harboring a plasmid carrying WT bS1. These experiments were carried out as described in Fig 2C; MazF expression was induced by the addition of NA. **(B)** Quantitative comparison of GFP expression in MazF induced samples (red columns) versus uninduced samples (blue columns). Data were calculated as percentages from the results of (Aa) and (Ab). In each assay, the value of 100% was given to the results for the untreated sample.
(**C)** Agarose gel with samples of PCR products of the GFP-STM reporter carrying an in-frame ACA site. Lanes 1-4: samples obtained from cells harboring a constitutive pACYC plasmid carrying the WT bS1. Lanes 5-8: sample obtained from cells harboring a constitutive pACYC plasmid carrying the bS1 mutation W->A in the EDF-like sequence (mutation W444A). Lanes 1 and 5: PCR amplification with a short primer. Lanes 2 and 6: PCR amplification with a long primer. Lanes 3 and 7: as in lanes 1 and 5 with the addition of NA. Lanes 4 and 8: as in lanes 2 and 6 with the addition of NA.

Figure 5. Locations of the synonym threonine codons in genes specifying for MazF-induced regulon products.
The DNA sequences specifying for **(A**) EF-P, **(B)** DeoC, **(C)** GroEL and **(D**) **b**S1. synonym threonine codons are :ACAs (highlighted in yellow), ACCs (in blue), ACT(in grey), and ACGs(in magenta).

Figure 6. MazF- induced GroEL expression is reduced when ACAs were introduced in frame 0 of *groEL.*
(A) Locations of three different in-frame ACA sites in the leaderless *groEL* sequence. Each ACA was inserted to a different location, generating only one ACA site in frame 0. The sites are numbered in circles 1-3 and highlighted in yellow. The different inserted ACAs were selected to be instead and nearby the original out- of frame ACA sites in the normal *groEL* sequence. **(B)** Western blot assays for the expression of the three in-frame STM-GroEL reporters in MazF induced MG1655 (WT) cells by NA. Bands are marked in 1, 2 and 3, corresponding to location of the in-frame ACA *groEL* sequence used. **(C**) As in B with the use of Δ*mazEF* derivative cells of *E-coli* MG165.

Figure 7. Model for frame dependent MazF cleavage in the STM system under stress conditions.
MazF becomes a part of the stress induced ribosome through its attachment to the bS1 ribosomal protein in the 30S ribosomal subunit. The translation is performed according to the open reading frame. Movement of the ribosome is illustrated by the black arrows .When the ribosome reaches an out -of frame ACA (marked in red), translation is not interrupted. When it reaches an in-frame ACA (marked in bold red), MazF cleaves the mRNA, and translation is prevented.

MAJOR ACCOMPLISHMENTS

During this last period of our research we had a major accomplishments. We studied the Stress-induced Translation Machinery (STM) generated by MazF under

stress, and composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs. We found that MazF cleaves only ACA sites located in the open reading frames of processed mRNAs, while out-of-frame ACAs are resistant. This in-frame ACA cleavage of MazF depends on MazF binding to an Extracellular Death Factor (EDF)-like element in ribosomal protein bS1 (bacterial S1), apparently causing MazF to be part of STM ribosomes. Furthermore, due to the in-frame MazF cleavage of ACAs, under stress, a bias occurs in the reading of the genetic code causing the amino acid threonine to be encoded only by its synonym codons ACC, ACU, or ACG (which are resistant to MazF), instead of by ACA.

To our knowledge this is the first report showing a stress-induced bias in the reading of the genetic code. Further studies will reveal if the described bias in the reading of the genetic code is specific only for *E. coli*, or if it may take place in other prokaryotes, or indeed, in some eukaryotes. We already show in this study that in the eukaryote jellyfish, the mRNA of the gene coding for GFP carries 17 ACAs, all of which are located out-of-frame. And in the open reading frame, the jellyfish GFP carries the threonine synonym codons, 17 ACCs, and one ACU. Since both this eukaryotic organism and the prokaryote *E. coli* appear to have the same bias in the reading of the genetic code, our results deals for the first time with a novel conceptual aspect of the genetic code: a new way of reading the genetic code under stress.

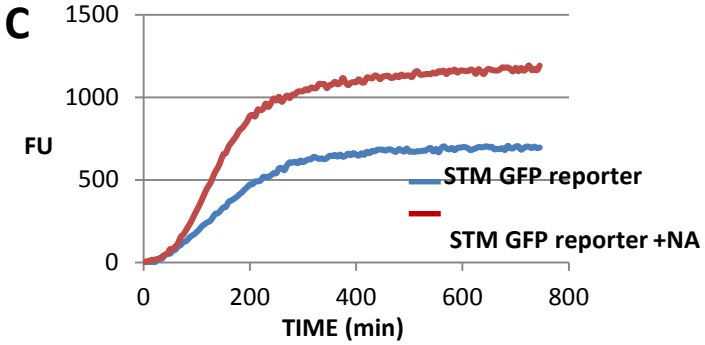
Fig 1

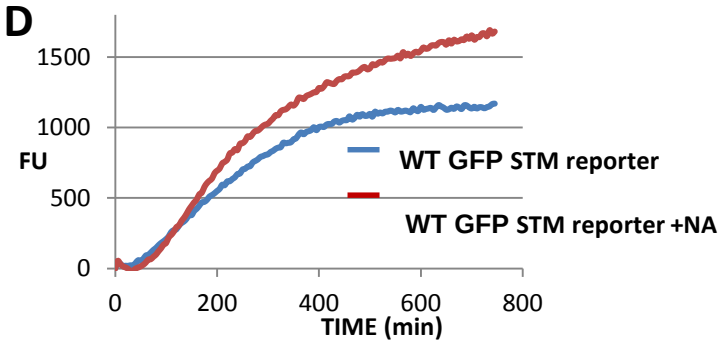
A ***gfp* sequence with no ACAs**

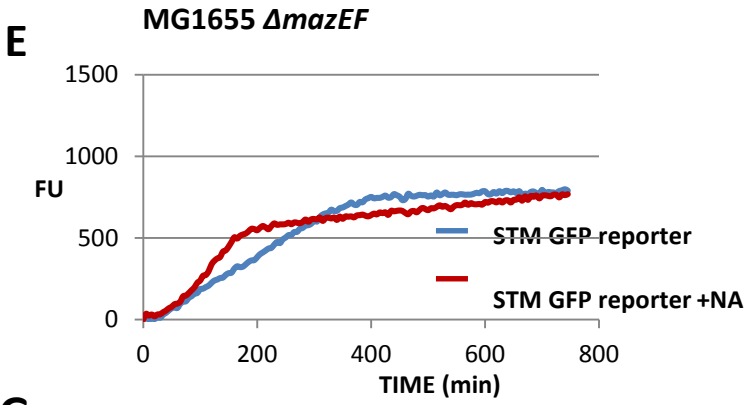
```
ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC
GAG CTG GAC GGC GAC GTA AAC GGC CAT AAG TTC AGC GTG TCC GGC GAG GGC
GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC
GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC TTG ACC TAC GGC
GTG CAG TGC TTC GCC CGC TAC CCC GAC CAT ATG AAG CAG CAC GAC TTC TTC
AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG
GAC GAC GGC AAC TAT AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAT ACC
CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAT
ATC CTG GGG CAT AAG CTG GAG TAT AAC TAT AAT AGC CAT AAG GTC TAT ATC
ACC GCC GAT AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ACC CGC CAT
AAT ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAT ACC
CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAT AAC CAC TAC CTG AGC ACC
CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAT ATG GTC CTG
CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAT
AAG TAA
```

B **WT *gfp* sequence**

```
ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC
GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC
GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC
GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC TTG ACC TAC GGC
GTG CAG TGC TTC GCC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC
AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG
GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC
CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC
ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAG GTC TAT ATC
ACC GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ACC CGC CAC
AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC
CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC
CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG
CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC
AAG TAA
```







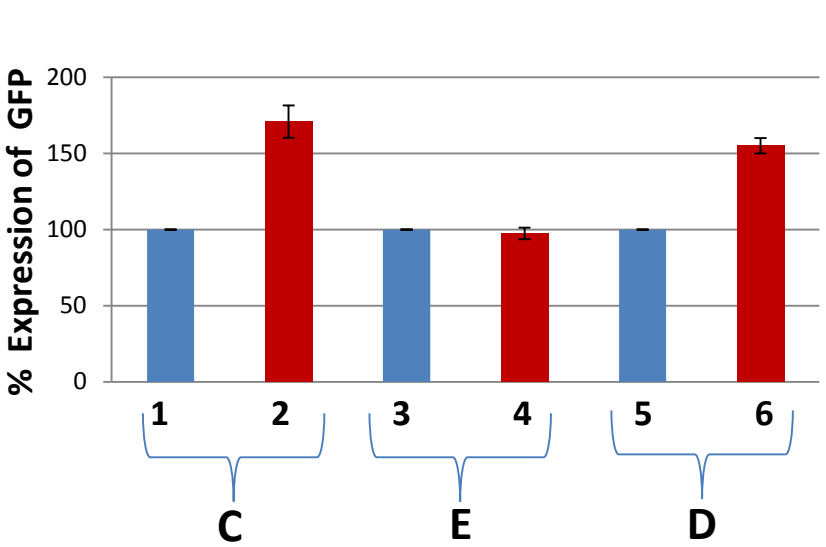


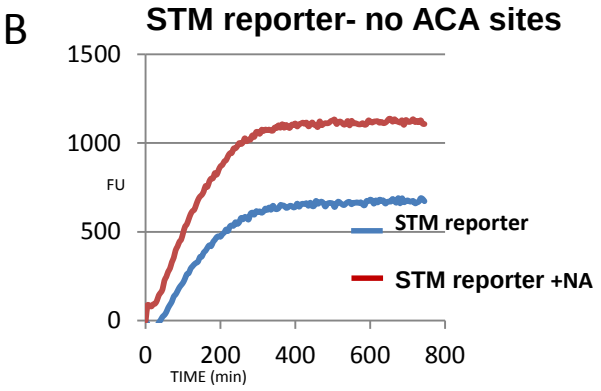
Fig 2

A

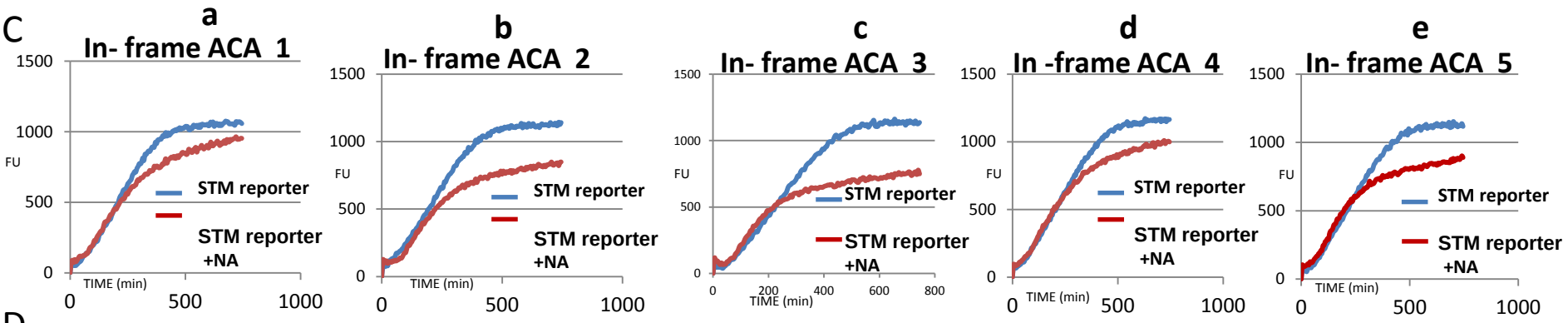
```
ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAT AAG TTC AGC GTG TCC GGC
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC TTC
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC CCC
TTG ACC TAC GGC GTG CAG TGC TTC GCC CGC TAC CCC GAC CAT ACA AAG
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAT ACA ACC CGC GCC GAG
GTG AAG TTC GAG GGC GAT ACA CTG GTG AAC CGC ATC GAG CTG AAG GGC
ATC GAC TTC AAG GAG GAC GGC AAT ATC CTG GGG CAT AAG CTG GAG TAT
AAC TAT ACA AGC CAT AAG GTC TAT ATC ACC GCC GAT AAG CAG AAG AAC
GGC ATC AAG GTG AAC TTC AAG ACC CGC CAT AAT GAC GAG GAC GGC AGC
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAT ACA GGC ATC GGC GAC GGC
CCC GTG CTG CTG CCC GAT AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAT ATG GTC CTG CTG GAG TTC
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAT AAG TAA
```

1 2 3 4 5

B



C



D

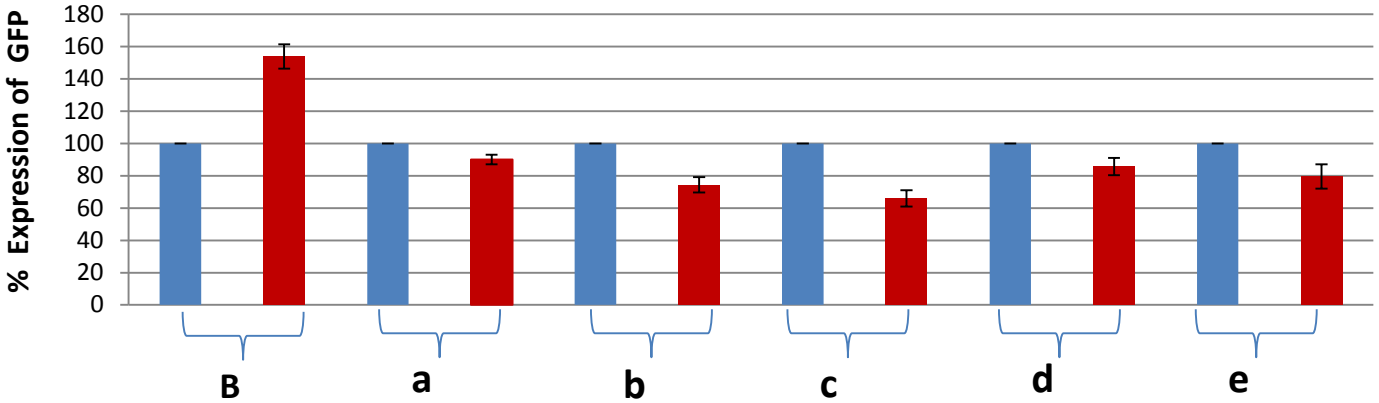
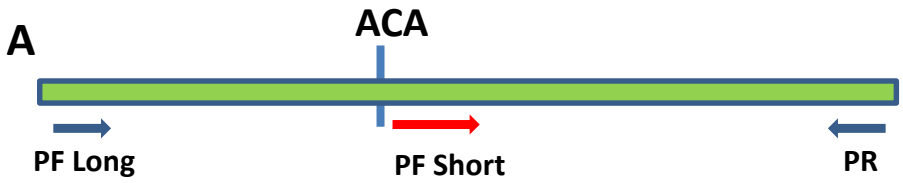
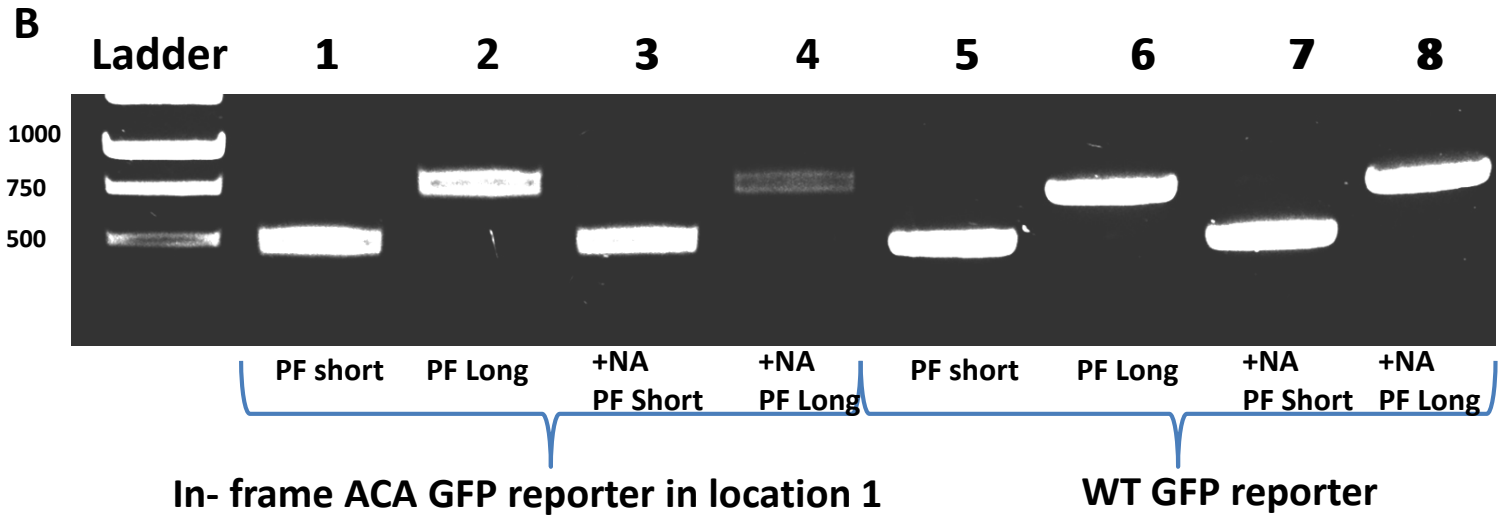
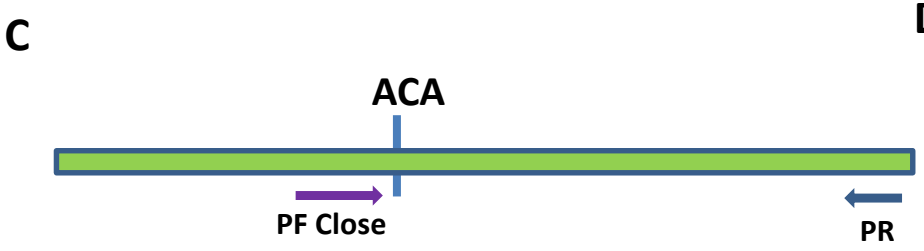


Fig 3


A
ACA
PF Long
PF Short
PR


B
Ladder
1
2
3
4
5
6
7
8
1000
750
500
PF short
PF Long
+NA PF Short
+NA PF Long
PF short
PF Long
+NA PF Short
+NA PF Long
In- frame ACA GFP reporter in location 1
WT GFP reporter


C
ACA
PF Close
PR

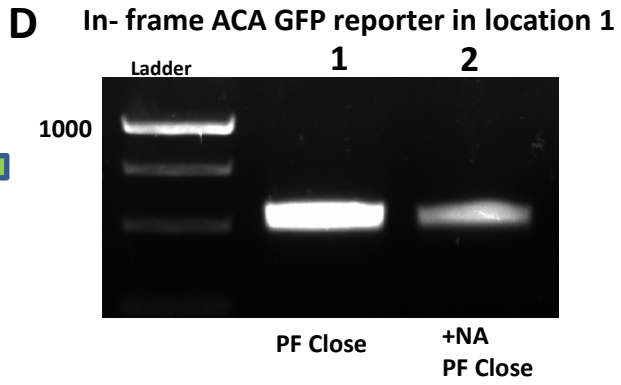

D
In- frame ACA GFP reporter in location 1
Ladder
1
2
1000
PF Close
+NA PF Close

Fig 4

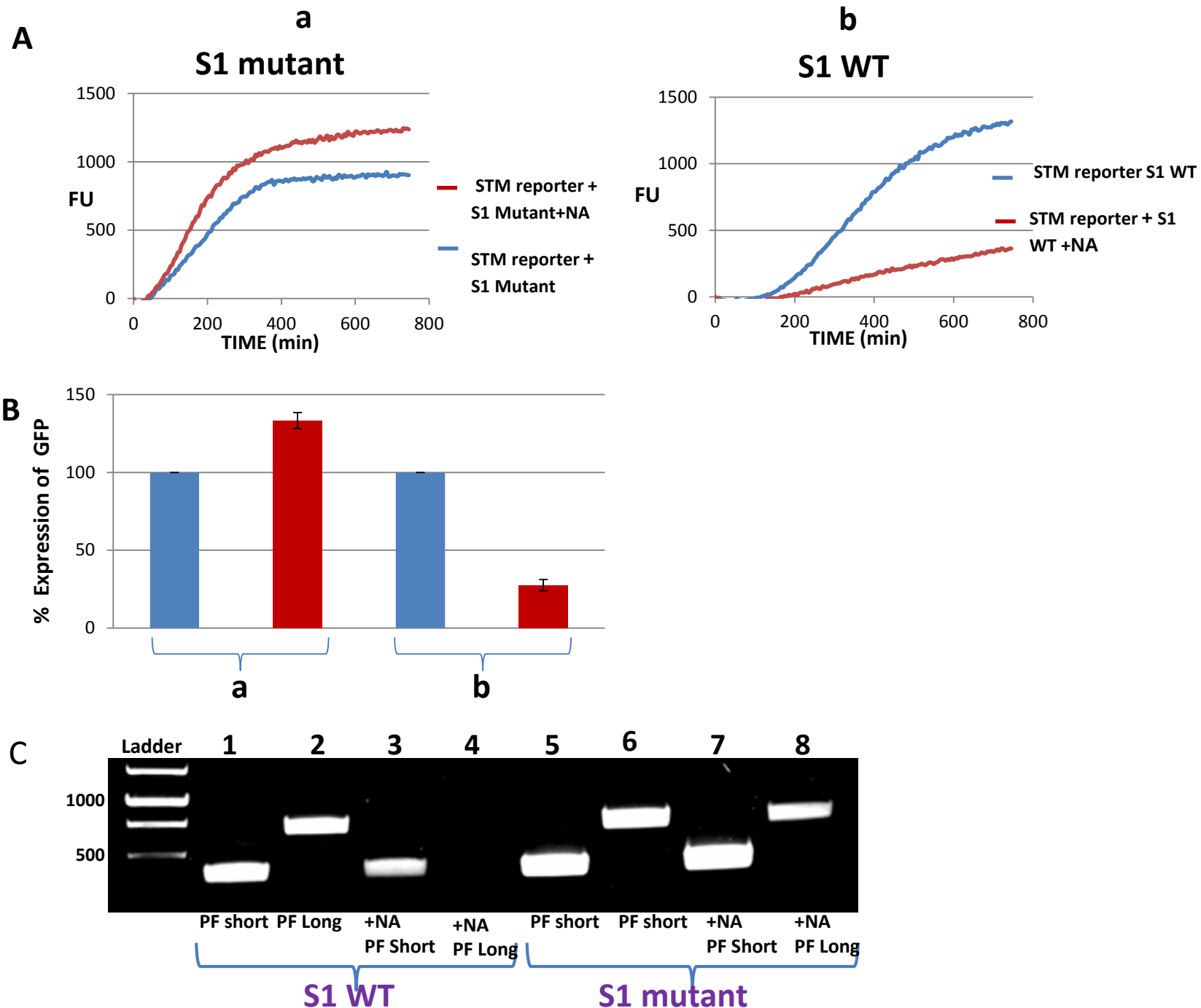

A
a
S1 mutant
1500
1000
500
0
FU
0
200
400
600
800
TIME (min)
STM reporter + S1 Mutant+NA
STM reporter + S1 Mutant
b
S1 WT
1500
1000
500
0
FU
0
200
400
600
800
TIME (min)
STM reporter S1 WT
STM reporter + S1 WT +NA
B
150
100
50
0
% Expression of GFP
a
b
C
Ladder
1
2
3
4
5
6
7
8
1000
500
PF short
PF Long
+NA PF Short
+NA PF Long
PF short
PF short
+NA PF Short
+NA PF Long
S1 WT
S1 mutant

Fig 5

A *efP*

TAACAAATTTCAGAGGGCCTT Atg GCA ACG TAC TAT AGC AAC GAT TTT CGT GCT GGT
CTT AAA ATC ATG TTA GAC GGC GAA CCT TAC GCG GTT GAA GCG AGT GAA TTC GTA
AAA CCG GGT AAA GGC CAG GCA TTT GCT CGC GTT AAA CTG CGT CGT CTG CTG ACC
GGT ACT CGC GTA GAA AAA ACC TTC AAA TCT ACT GAT TCC GCT GAA GGC GCT GAT
GTT GTC GAT ATG AAC CTG ACT TAC CTG TAC AAC GAC GGT GAG TTC TGG CAC TTC
ATG AAC AAC GAA ACT TTC GAG CAG CTG TCT GCT GAT GCA AAA GCA ATT GGT GAC
AAC GCT AAA TGG CTG CTG GAT CAG GCA GAG TGT ATC GTA ACT CTG TGG AAT GGT
CAG CCG ATC TCC GTT ACT CCG CCG AAC TTC GTT GAA CTG GAA ATC GTT GAT ACC
GAT CCG GGC CTG AAA GGT GAT ACC GCA GGT ACT GGT GGC AAA CCG GCT ACC CTG
TCT ACT GGC GCT GTG GTT AAA GTT CCG CTG TTT GTA CAA ATC GGC GAA GTC ATC
AAA GTG GAT ACC CGC TCT GGT GAA TAC GTC TCT CGC GTG AAG taa

B *deoC*

CGACAAGCCAGGAGAATGAAatg ACT GAT CTG AAA GCA AGC AGC CTG CGT GCA CTG AAA
TTG ATG GAC CTG ACC ACC CTG AAT GAC GAC GAC ACC GAC GAG AAA GTG ATC GCC
CTG TGT CAT CAG GCC AAA ACT CCG GTC GGC AAT ACC GCC GCT ATC TGT ATC TAT
CCT CGC TTT ATC CCG ATT GCT CGC AAA ACT CTG AAA GAG CAG GGC ACC CCG GAA
ATC CGT ATC GCT ACG GTA ACC AAC TTC CCA CAC GGT AAC GAC GAC ATC GAC ATC
GCG CTG GCA GAA ACC CGT GCG GCA ATC GCC TAC GGT GCT GAT GAA GTT GAC GTT
GTG TTC CCG TAC CGC GCG CTG ATG GCG GGT AAC GAG CAG GTT GGT TTT GAC CTG
GTG AAA GCC TGT AAA GAG GCT TGC GCG GCA GCG AAT GTA CTG CTG AAA GTG ATC
ATC GAA ACC GGC GAA CTG AAA GAC GAA GCG CTG ATC CGT AAA GCG TCT GAA ATC
TCC ATC AAA GCG GGT GCG GAC TTC ATC AAA ACC TCT ACC GGT AAA GTG GCT GTG
AAC GCG ACG CCG GAA AGC GCG CGC ATC ATG ATG GAA GTG ATC CGT GAT ATG GGC
GTA GAA AAA ACC GTT GGT TTC AAA CCG GCG GGC GGC GTG CGT ACT GCG GAA GAT
GCG CAG AAA TAT CTC GCC ATT GCA GAT GAA CTG TTC GGT GCT GAC TGG GCA GAT
GCG CGT CAC TAC CGC TTT GGC GCT TCC AGC CTG CTG GCA AGC CTG CTG AAA GCG
CTG GGT CAC GGC GAC GGT AAG AGC GCC AGC AGC TAC taa

C *groEL*

GCGCACGACACTGAACATACGAATTTAAGGAATAAAGATA atg GCA GCT AAA GAC GTA AAA
TTC GGT AAC GAC GCT CGT GTG AAA ATG CTG CGC GGC GTA AAC GTA CTG GCA GAT
GCA GTG AAA GTT ACC CTC GGT CCA AAA GGC CGT AAC GTA GTT CTG GAT AAA TCT
TTC GGT GCA CCG ACC ATC ACC AAA GAT GGT GTT TCC GTT GCT CGT GAA ATC GAA
CTG GAA GAC AAG TTC GAA AAT ATG GGT GCG CAG ATG GTG AAA GAA GTT GCC TCT
AAA GCA AAC GAC GCT GCA GGC GAC GGT ACC ACC ACT GCA ACC GTA CTG GCT CAG
GCT ATC ATC ACT GAA GGT CTG AAA GCT GTT GCT GCG GGC ATG AAC CCG ATG GAC
CTG AAA CGT GGT ATC GAC AAA GCG GTT ACC GCT GCA GTT GAA GAA CTG AAA GCG
CTG TCC GTA CCA TGC TCT GAC TCT AAA GCG ATT GCT CAG GTT GGT ACC ATC TCC
GCT AAC TCC GAC GAA ACC GTA GGT AAA CTG ATC GCT GAA GCG ATG GAC AAA GTC
GGT AAA GAA GGC GTT ATC ACC GTT GAA GAC GGT ACC GGT CTG CAG GAC GAA CTG
GAC GTG GTT GAA GGT ATG CAG TTC GAC CGT GGC TAC CTG TCT CCT TAC TTC ATC
AAC AAG CCG GAA ACT GGC GCA GTA GAA CTG GAA AGC CCG TTC ATC CTG CTG GCT
GAC AAG AAA ATC TCC AAC ATC CGC GAA ATG CTG CCG GTT CTG GAA GCT GTT GCC
AAA GCA GGC AAA CCG CTG CTG ATC ATC GCT GAA GAT GTA GAA GGC GAA GCG CTG
GCA ACT CTG GTT GTT AAC ACC ATG CGT GGC ATC GTG AAA GTC GCT GCG GTT AAA
GCA CCG GGC TTC GGC GAT CGT CGT AAA GCT ATG CTG CAG GAT ATC GCA ACC CTG
ACT GGC GGT ACC GTG ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCA ACC
CTG GAA GAC CTG GGT CAG GCT AAA CGT GTT GTG ATC AAC AAA GAC ACC ACC ACT
ATC ATC GAT GGC GTG GGT GAA GAA GCT GCA ATC CAG GGC CGT GTT GCT CAG ATC
CGT CAG CAG ATT GAA GAA GCA ACT TCT GAC TAC GAC CGT GAA AAA CTG CAG GAA
CGC GTA GCG AAA CTG GCA GGC GGC GTT GCA GTT ATC AAA GTG GGT GCT GCT ACC
GAA GTT GAA ATG AAA GAG AAA AAA GCA CGC GTT GAA GAT GCC CTG CAC GCG ACC
CGT GCT GCG GTA GAA GAA GGC GTG GTT GCT GGT GGT GGT GTT GCG CTG ATC CGC
GTA GCG TCT AAA CTG GCT GAC CTG CGT GGT CAG AAC GAA GAC CAG AAC GTG GGT
ATC AAA GTT GCA CTG CGT GCA ATG GAA GCT CCG CTG CGT CAG ATC GTA TTG AAC
TGC GGC GAA GAA CCG TCT GTT GTT GCT AAC ACC GTT AAA GGC GGC GAC GGC AAC
TAC GGT TAC AAC GCA GCA ACC GAA GAA TAC GGC AAC ATG ATC GAC ATG GGT ATC
CTG GAT CCA ACC AAA GTA ACT CGT TCT GCT CTG CAG TAC GCA GCT TCT GTG GCT
GGC CTG ATG ATC ACC ACC GAA TGC ATG GTT ACC GAC CTG CCG AAA AAC GAT GCA
GCT GAC TTA GGC GCT GCT GGC GGT ATG GGC GGC ATG GGT GGC ATG GGC GGC ATG
ATG taa

D *rpsA*

GGACGTTAAATATAAACCTGAAGATTAAAC atg ACT GAA TCT TTT GCT CAA CTC TTT GAA
GAG TCC TTA AAA GAA ATC GAA ACC CGC CCG GGT TCT ATC GTT CGT GGC GTT GTT
GTT GCT ATC GAC AAA GAC GTA GTA CTG GTT GAC GCT GGT CTG AAA TCT GAG TCC
GCC ATC CCG GCT GAG CAG TTC AAA AAC GCC CAG GGC GAG CTG GAA ATC CAG GTA
GGT GAC GAA GTT GAC GTT GCT CTG GAC GCA GTA GAA GAC GGC TTC GGT GAA ACT
CTG CTG TCC CGT GAG AAA GCT AAA CGT CAC GAA GCC TGG ATC ACG CTG GAA AAA
GCT TAC GAA GAT GCT GAA ACT GTT ACC GGT GTT ATC AAC GGC AAA GTT AAG GGC
GGC TTC ACT GTT GAG CTG AAC GGT ATT CGT GCG TTC CTG CCA GGT TCT CTG GTA
GAC GTT CGT CCG GTG CGT GAC ACT CTG CAC CTG GAA GGC AAA GAG CTT GAA TTT
AAA GTA ATC AAG CTG GAT CAG AAG CGC AAC AAC GTT GTT GTT TCT CGT CGT GCC
GTT ATC GAA TCC GAA AAC AGC GCA GAG CGC GAT CAG CTG CTG GAA AAC CTG CAG
GAA GGC ATG GAA GTT AAA GGT ATC GTT AAG AAC CTC ACT GAC TAC GGT GCA TTC
GTT GAT CTG GGC GGC GTT GAC GGC CTG CTG CAC ATC ACT GAC ATG GCC TGG AAA
CGC GTT AAG CAT CCG AGC GAA ATC GTC AAC GTG GGC GAC GAA ATC ACT GTT AAA
GTG CTG AAG TTC GAC CGC GAA CGT ACC CGT GTA TCC CTG GGC CTG AAA CAG CTG
GGC GAA GAT CCG TGG GTA GCT ATC GCT AAA CGT TAT CCG GAA GGT ACC AAA CTG
ACT GGT CGC GTG ACC AAC CTG ACC GAC TAC GGC TGC TTC GTT GAA ATC GAA GAA
GGC GTT GAA GGC CTG GTA CAC GTT TCC GAA ATG GAC TGG ACC AAC AAA AAC ATC
CAC CCG TCC AAA GTT GTT AAC GTT GGC GAT GTA GTG GAA GTT ATG GTT CTG GAT
ATC GAC GAA GAA CGT CGT CGT ATC TCC CTG GGT CTG AAA CAG TGC AAA GCT AAC
CCG TGG CAG CAG TTC GCG GAA ACC CAC AAC AAG GGC GAC CGT GTT GAA GGT AAA
ATC AAG TCT ATC ACT GAC TTC GGT ATC TTC ATC GGC TTG GAC GGC GGC ATC GAC
GGC CTG GTT CAC CTG TCT GAC ATC TCC TGG AAC GTT GCA GGC GAA GAA GCA GTT
CGT GAA TAC AAA AAA GGC GAC GAA ATC GCT GCA GTT GTT CTG CAG GTT GAC GCA
GAA CGT GAA CGT ATC TCC CTG GGC GTT AAA CAG CTC GCA GAA GAT CCG TTC AAC
AAC TGG GTT GCT CTG AAC AAG AAA GGC GCT ATC GTA ACC GGT AAA GTA ACT GCA
GTT GAC GCT AAA GGC GCA ACC GTA GAA CTG GCT GAC GGC GTT GAA GGT TAC CTG
CGT GCT TCT GAA GCA TCC CGT GAC CGC GTT GAA GAC GCT ACC CTG GTT CTG AGC
GTT GGC GAC GAA GTT GAA GCT AAA TTC ACC GGC GTT GAT CGT AAA AAC CGC GCA
ATC AGC CTG TCT GTT CGT GCG AAA GAC GAA GCT GAC GAG AAA GAT GCA ATC GCA
ACT GTT AAC AAA CAG GAA GAT GCA AAC TTC TCC AAC AAC GCA ATG GCT GAA GCT
TTC AAA GCA GCT AAA GGC GAG taa

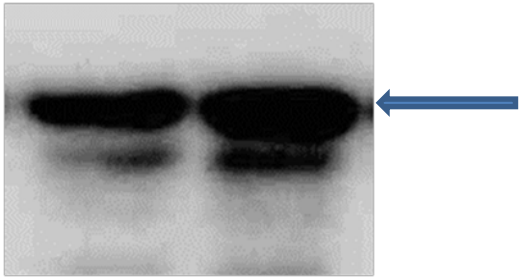
Fig 6

A

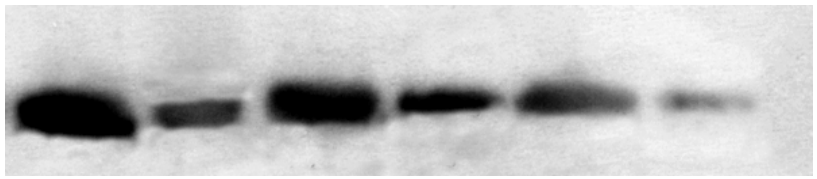
groEL

ATG GCA GCT AAA GAC GTA AAA TTC GGT AAC GAC GCT CGT GTG AAA
ATG CTG CGC GGC GTA AAC GTA CTG GCA GAT GCA GTG AAA GTT ACC
CTC GGT CCA AAA GGC CGT AAC GTA GTT CTG GAT AAA TCT TTC GGT
GCA CCG ACC ATC ACC AAA GAT GGT GTT TCC GTT GCT CGT GAA ATC
GAA CTG GAA GAC AAG TTC GAA AAT ATG GGT GCG CAG ATG GTG AAA
GAA GTT GCC TCT AAA GCA AAC GAC GCT GCA GGC GAC GGT ACC ACC
ACT GCA ACC GTA CTG GCT CAG GCT ATC ATC ACT GAA GGT CTG AAA
GCT GTT GCT GCG GGC ATG AAC CCG ATG GAC CTG AAA CGT GGT ATC
GAC AAA GCG GTT ACC GCT GCA GTT GAA GAA CTG AAA GCG CTG TCC
GTA CCA TGC TCT GAC TCT AAA GCG ATT GCT CAG GTT GGT ACC ATC
TCC GCT AAC TCC GAC GAA ACC GTA GGT AAA CTG ATC GCT GAA GCG
ATG GAC AAA GTC GGT AAA GAA GGC GTT ATC ACC GTT GAA GAC GGT
ACC GGT Location 1 GAA CTG GAC GTG GTT GAA GGT ATG CAG TTC
GAC CGT TCT CCT TAC TTC ATC AAC AAG CCG GAA ACT
GGC GCA GTA GAA CTG GAA AGC CCG TTC GCT GAC AAG
AAA ATC TCC ACA ATC CGC GAA ATG CTG Location 2 GAA GCT GTT
GCC AAA GCA GGC AAA CCG CTG CTG ATC ATC GCT GAA GAT GTA GAA
GGC GAA GCG CTG GCA ACT CTG GTT GTT AAT ACA ATG CGT GGC ATC
GTG AAA GTC GCT GCG GTT AAA GCA CCG GGC TTC GGC GAT CGT CGT
AAA GCT ATG CTG CAG GAT ATC GCA ACC CTG ACT GGC GGT ACC GTG
ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCA ACC CTG GAA
GAC CTG GGT CAG GCT AAA CGT GTT GTG ATC AAC AAA GAC ACC ACC
ACT ATC ATC GAT GGC GTG GGT GAA GAA GCT GCA ATC CAG GGC CGT
GTT GCT CAG ATC CGT CAG CAG ATT GAA GAA GCA ACT TCT GAC TAC
GAC CGT GAA AAA CTG CAG GAA CGC GTA GCG AAA CTG GCA GGC GGC
GTT GCA GTT ATC AAA GTG GGT GCT GCT ACC GAA GTT GAA ATG AAA
GAG AAA AAA GCA CGC GTT GAA GAT GCC CTG CAC GCG ACC CGT GCT
GCG GTA GAA GAA GGC GTG GTT GCT GGT GGT GTT GCG CTG ATC
CGC AAA CTG GCT GAC CTG CGT GGT CAG AAC GAA GAC
CAG Location 3 ATC AAA GTT GCA CTG CGT GCA ATG GAA GCT CCG
CTG CGT CAG ATC GTA TTG AAC TGC GGC GAA GAA CCG TCT GTT GTT
GCT AAT ACA GTT AAA GGC GGC GAC GGC AAC TAC GGT TAC AAC GCA
GCA ACC GAA GAA TAC GGC AAC ATG ATC GAC ATG GGT ATC CTG GAT
CCA ACC AAA GTA ACT CGT TCT GCT CTG CAG TAC GCA GCT TCT GTG
GCT GGC CTG ATG ATC ACC ACC GAA TGC ATG GTT ACC GAC CTG CCG
AAA AAC GAT GCA GCT GAC TTA GGC GCT GCT GGC GGT ATG GGC GGC
ATG GGT GGC ATG GGC GGC ATG ATG taa

B



C



C

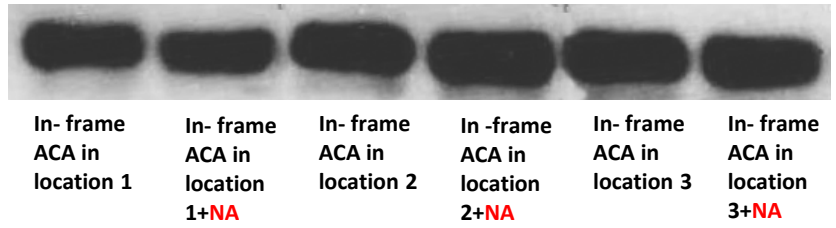
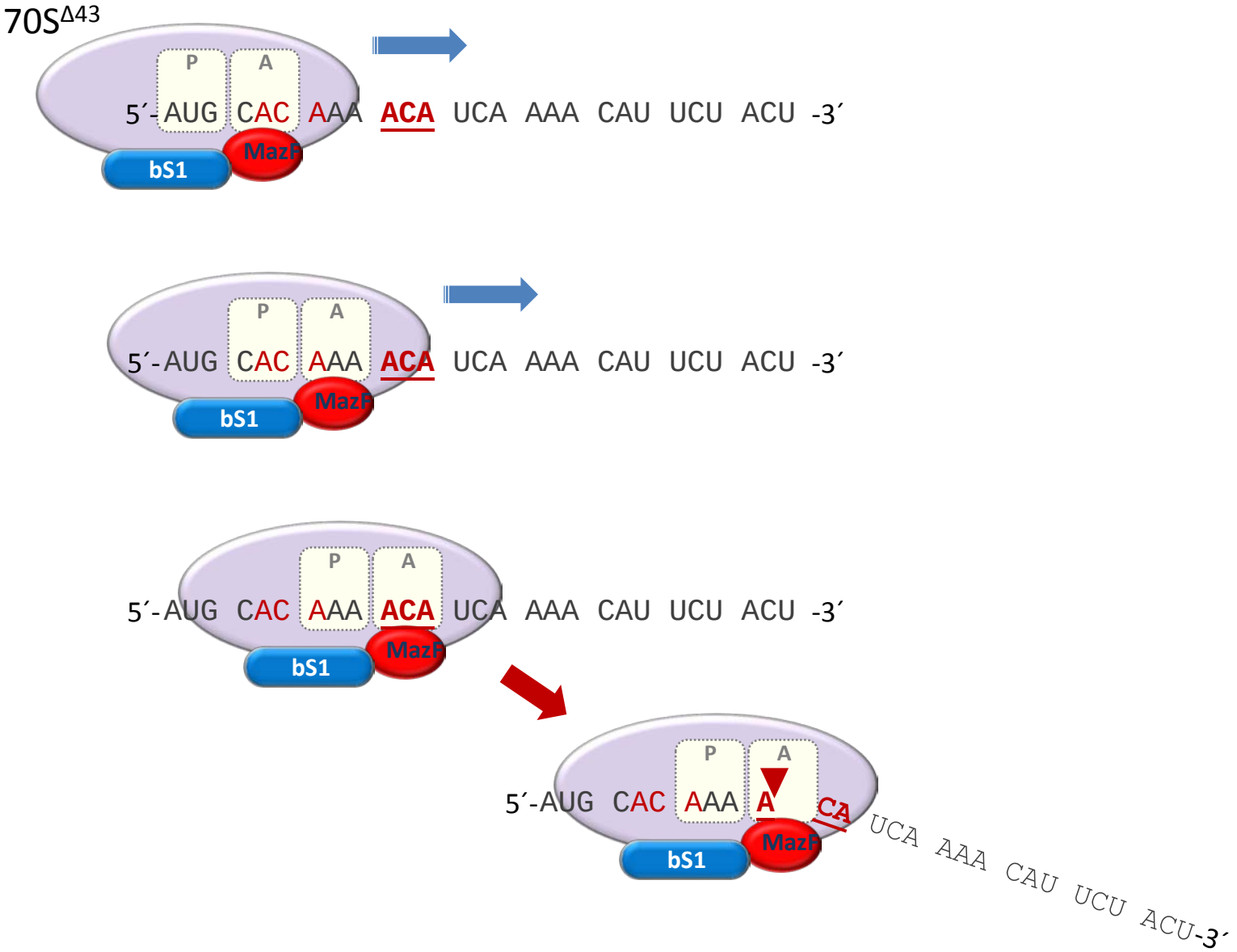


Fig 7


70S$^{\Delta 43}$
P
A
5´-AUG CAC AAA ACA UCA AAA CAU UCU ACU -3´
bS1
MazF
5´-AUG CAC AAA ACA UCA AAA CAU UCU ACU -3´
5´-AUG CAC AAA ACA UCA AAA CAU UCU ACU -3´
5´-AUG CAC AAA A CA UCA AAA CAU UCU ACU-3´