A Bias in the Reading of the Genetic Code of *Escherichia coli* is a Characteristic for Genes that Specify Stress-induced MazF-mediated Proteins

Akanksha Nigam¹, Adi Oron-Gottesman¹ and Hanna Engelberg-Kulka^{1,*}

¹Department of Microbiology and Molecular Genetics, IMRIC, The Hebrew University, Hadassah Medical School, Jerusalem 91120, Israel

Abstract: *Background: Escherichia coli* (*E. coli*) *mazEF*, a stress-induced toxin-antitoxin (TA) system, has been studied extensively. The MazF toxin is an endoribonuclease that cleaves RNAs at ACA sites. Thereby, under stress, the induced MazF generates a Stress-induced Translation Machinery (STM), composed of MazF processed mRNAs and selective ribosomes that specifically translate the processed mRNAs.

ARTICLE HISTORY

Received: October 29, 2019 Revised: May 18, 2020 Accepted: May 18, 2020

DOI: 10.2174/1389202921999200606215305 *Materials and Methods*: Based on the data from the EcoCyc website of the National Center for Biotechnology Information (NCBI), the sequence of all *E. coli* MG1655 genes were scanned for ACA sites upstream from the initiation codons. Among these sequences, the fuzznuc program of the "European Molecular Biology Open Software Suite" (EMBOSS) was used to find the "ACA" pattern. The distribution of the ACA threonine codon, both in-frame and out-of-frame, was determined by using the HTML Script Program (Supplementary Material).

Results: Here it is reported that for most of the *E. coli* proteins mediated by stress-induced MazF, the ACA threonine codon in their mRNAs is not in-frame but rather out-of-frame; in these same RNAs, the three synonymous threonine codons, ACG, ACU, and ACC, are in-frame. In contrast, for proteins translated by the canonical translation system, in the majority of mRNAs, the ACA codon is located in-frame.

Conclusion: The described bias in the genetic code is a characteristic of *E. coli* genes specifying for stress-induced MazF-mediated proteins.

Keywords: Genetic code, Toxin-Antitoxin (TA) modules, E. coli, MazF toxin, ACA sites, genes.

1. INTRODUCTION

E. coli mazEF, abundant in the chromosomes of most bacteria [1-4], is the most studied chromosomal toxinantitoxin (TA) module [1, 2]. E. coli mazF specifies for the stable toxin MazF, and *mazE* specifies for the labile antitoxin, MazE, which is degraded by the ATP-dependent ClpAP serine protease [1]. Various stressful conditions can prevent the production of the antitoxin MazE: the absence of the anti-toxin permits the uninterrupted activity of the toxin MazF [5-7]. MazF is a sequence-specific endoribonuclease that preferentially cleaves single-stranded mRNAs either at the 3' or at the 5' side of the first A in ACA sequences [8, 9]. Moreover, under stress, by cleaving ACA sites immediately adjacent to or within the span of 100 nucleotides upstream from the AUG start codons of any given mRNA, the induced MazF generates leaderless mRNAs [10]. Concomitantly, MazF targets 16S rRNA within 30S ribosomal subunits at the decoding center, consequently removing 43 nucleotides

from the 3' terminus [10]. Since this region comprises the anti-Shine-Dalgarno (aSD) region, the generated deficient ribosomes are selectively able to translate the generated, processed mRNAs [10-12]. This <u>Stress-induced Translation</u> <u>Machinery</u> (STM) is responsible for the selective synthesis of specific proteins resulting from MazF induction [13].

Previously, when the STM system was characterized [14], it was found that MazF cleaves ACA sites located in frame 0 of the coding region of the processed mRNAs, while out-of-frame ACAs are resistant to such cleavage. Moreover, under stressful conditions, when MazF cleaves in-frame ACAs, a bias in the reading of the genetic code is caused. The result of this bias is that the amino acid threonine is no longer encoded by ACA, but rather by its synonym codons ACC, ACU, or ACG [14]. Recently, a proteomic study was carried out [13] in which the stress-induced MazF-mediated proteins of E. coli were identified. It was found that the mRNAs of nearly all the identified proteins are characterized by the presence of an ACA site within a span of 1 to 100 nucleotides upstream from the AUG initiator [13]. Thus, under stressful conditions, the induced MazF processes the mRNAs that are translated by the STM. Here it is shown that in half of these proteomic identified proteins, the threonine

^{*}Address correspondence to this author at the Department of Microbiology and Molecular Genetics, IMRIC, The Hebrew University, Hadassah Medical School, Jerusalem 91120, Israel; Tel: +97226758250; E-mail: hanita@cc.huji.ac.il

codon ACA is not located in-frame, while its synonymous codons, ACG, ACU and ACC, are mostly in-frame. In contrast, in the majority of mRNAs of the proteins translated by the canonical translation system, an ACA is mostly located in-frame. Thus, it is suggested that *E. coli* genes specifying for stress-induced MazF-mediated proteins are mostly characterized by the described bias in the reading of the genetic code.

Recently, Aoi *et al.* reported that, in *Nitrospira* strain ND1, the MazF toxin specifically recognizes the motifs AACU, AACG, and AAUU [15]. However, since, in this case, MazF target is not a coding triplet, a bias in the reading of the genetic code should not be characteristic for genes that specify stress-induced MazF mediated proteins of *Nitrospira* strain ND.

2. METHODS

2.1. Computational Analysis

The distribution of the in-frame and out-of-frame threonine codon (ACA) was determined using the HTML Script Program (Supplementary Materials).

The sequences of all *E. coli* MG1655 genes were scanned for ACA sites upstream from the initiation codons using the *E. coli* K-12 sub-strain MG1655 genomic sequence and the EcoCyc database. For this aim, gene annotations were downloaded from the NCBI Genome database (ftp://ftp.ncbi.nlm. nih.gov/genomes/all/GCF/000/005/845/GCF_000005845.2_ ASM584v2/). Gene annotations were downloaded from the Genome database of the National Center for Biotechnology Information (NCBI) (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/ GCF/000/005/845/GCF_000005845.2_ASM584v2/).

All protein coding gene coordinates and strand information were scanned to find genes that had a coding-free region of at least 100 non-coding nucleotides upstream from their translation start site. Among these sequences, The EM-BOSS-*fuzznuc* function of the *The European Molecular Biology Open Software Suite* (EMBOSS) was used to search for the pattern "ACA" [16].

3. RESULTS

3.1. The Distribution of the Threonine Synonymous Codons ACA, ACG, ACU, and ACC in Relation to the Open Reading Frames in mRNAs of Stress-induced MazFmediated *E. coli* Proteins

The amino acid threonine is encoded by four synonymous codons: ACA, ACG, ACU, and ACC. The locations of all four of these codons, in the mRNAs of stress-induced MazF-mediated *E. coli* proteins (Table 1), were identified. Note that because all these mRNAs carry the MazF cleavage site ACA within the region of 1 to 100 nucleotides upstream from the AUG initiator codon, under stressful conditions, they can be translated by the STM system [10, 11]. As shown in Table 1, about 50% of these mRNAs (18 out of 36) include no in-frame ACA site. However, the three synonymous codons ACG, ACU, and ACC, which are not cleaved by MazF, are located in-frame. Moreover, while they carry no in-frame ACA codons, these mRNAs, do carry out-offrame ACA codon(s).

3.2. The Distribution of the Four Synonymous Threonine Codons, ACA, ACG, ACU, and ACC, in Relation to the Open Reading Frames in the mRNAs of *E. coli* Proteins that are not Mediated by MazF

The locations of the four synonymous threonine codons, ACA, ACG, ACU, and ACC, in mRNAs of E. coli proteins that are not mediated by stress-induced MazF (Table 2), were identified. Because none of these mRNAs carry an ACA MazF cleavage site within a span of 1 to 100 nucleotides upstream from the AUG initiator codon, they were not considered to be translated by the Stress-induced Translation Machinery (STM), but rather by the "regular" canonical translation system. As shown, 88% (44 out of 50) of the described mRNAs do carry at least one in-frame ACA site. This value is significantly higher than the corresponding value of 50% that was obtained in the mRNAs of genes, which are stress-induced MazF-mediated (Table 1). Fig. (2) shows the positions of in-frame (red) and out-of-frame (yellow) ACA sites in some specific proteins (randomly selected from Table 2) that are not mediated by MazF.

4. DISCUSSION

Here the distribution of the four synonymous threonine codons ACA, ACG, ACU, and ACC, in relation to the open reading frames in mRNAs of the stress-induced MazFmediated E. coli proteins, was studied. The results clearly show that in the mRNAs of 50% of these proteins, the threonine codon ACA is not located in-frame, while its synonymous codons, ACG, ACU and ACC are in-frame (Table 1). However, all (100%) of these mRNA molecules carry outof-frame ACAs. In addition, the mRNAs of each of these proteins carry an ACA codon within a span of 100 nucleotides upstream from the AUG initiator (Table 1), thus permitting translation by the Stress-induced Translation Machinery (STM). In contrast, in the majority of mRNAs of the proteins translated by the canonical translation system, ACAs are located in-frame (Table 2). Furthermore, none of them carry an ACA site within a span of 100 nucleotides upstream from the AUG initiator (Table 2).

The frequencies that were found for the four synonymous threonine codons in *E. coli* are ACA: 17%, ACU: 19%, ACC: 40%, and ACG: 25% (Table **3**, column B). In the mRNAs of *E. coli* proteins not mediated by MazF, similar frequencies were found: ACA: 12%, ACU: 19. 5%, ACC: 44%, and ACG: 24.5% (Table **3** column C). In contrast, in the mRNAs of MazF-mediated *E. coli* proteins, it was observed that the frequency of ACA was reduced from 12% to 5% (Table **3**, column D). It is suggested that this significant reduction of more than 50% is a confirmation of the bias against ACA in the genetic code of *E. coli* genes specifying for stress-induced MazF-mediated proteins.

It is emphasized that although most of the mRNAs of the stress-induced MazF-mediated proteins carried out-of-frame ACA codons were found, 50% of these mRNAs also carried

| # | Gene Names | Protein Product | # ACA In-frame | # ACA Out of Frame | # ACG In-frame | # ACC In-frame | # ACU In-frame | Distance of ACA from AUG Initiator |
|----|---------------|--|-------------------|--------------------------|-------------------|-------------------|-------------------|--|
| 1 | bola | DNA-binding transcriptional regulator | 0 | 4 | 1 | 2 | 4 | 29 |
| 2 | dnaK | Chaperone protein | 0 | 22 | 4 | 23 | 17 | 33 |
| 3 | ftsH | ATP dependent zinc metallo- protease ftsH | 0 | 18 | 8 | 10 | 9 | 1 |
| 4 | galF | | 0 | 9 | 3 | 3 | 5 | 15 |
| 5 | groL | 60 kDa chaperonin | 0 | 13 | 0 | 25 | 8 | 25 |
| 6 | groS | 10 kDa chaperonin | 0 | 3 | 0 | 2 | 1 | 32 |
| 7 | grpE | Protein GrpE | 0 | 2 | 4 | 1 | 3 | 90 |
| 8 | hslU | ATP-dependent protease ATPase subunit | 0 | 14 | 0 | 16 | 5 | 19 |
| 9 | IbpA | Small heat shock protein | 0 | 3 | 0 | 2 | 0 | 30 |
| 10 | ihfB | Integration host factor | 0 | 1 | 1 | 3 | 2 | 40 |
| 11 | metK | S-adenosyl methionine syn- thase | 0 | 16 | 1 | 15 | 8 | 40 |
| 12 | osmE | Osmotically- inducible lipo- protein E | 0 | 7 | 1 | 6 | 2 | 14 |
| 13 | pflB | Formate acetyltransferase 1 | 0 | 34 | 1 | 27 | 18 | 1 |
| 14 | RpoD | RNA polymerase sigma factor | 0 | 22 | 7 | 29 | 2 | 35 |
| 15 | sra | Stationary phase induced ribosome associated protein | 0 | 1 | 0 | 2 | 0 | 18 |
| 16 | ybeZ | phoH- like protein | 0 | 16 | 3 | 10 | 4 | 20 |
| 17 | ydfG | NADP dependent3-hydroxy acid dehydrogenase | 0 | 5 | 9 | 7 | 3 | 7 |
| 18 | yeeX | Uncharacterized protein | 0 | 5 | 0 | 1 | 1 | 19 |
| 19 | def | Peptide deformylase | 1 | 4 | 1 | 1 | 0 | 15 |
| 20 | dps | DNA protection during star- vation protein | 1 | 9 | 0 | 8 | 3 | 20 |
| 21 | galU | UTP-glucose-1-phosphate uridylyltransferase | 1 | 5 | 4 | 6 | 2 | 18 |
| 22 | lon | Lon protease | 1 | 22 | 6 | 20 | 5 | 50 |
| 23 | osmY | Osmotically-inducible protein Y | 1 | 7 | 0 | 14 | 4 | 13 |
| 24 | pckA | Phosphoenolpyruvate carbox- ykinase | 1 | 20 | 2 | 27 | 14 | 23 |
| 25 | rho | Transcription termination factor | 1 | 12 | 3 | 13 | 6 | 7 |
| 26 | rpmA | <i>rpmA</i> 50 S ribosomal protein 1 2 0 | | 0 | 1 | 1 | 89 | |

Table 1. The locations of ACA sites in mRNAs of stress-induced MazF-mediated E. coli proteins.

(Table 1) contd....

| Nigam | et | al. |
|-------|-----|-----|
| | ••• | |

| # | Gene Names | Protein Product | # ACA In-frame | # ACA Out of Frame | # ACG In-frame | # ACC In-frame | # ACU In-frame | Distance of ACA from AUG Initiator |
|----|---------------|--|-------------------|--------------------------|-------------------|-------------------|-------------------|--|
| 27 | rraB | Regulator of ribonuclease activity | 1 | 6 | 1 | 3 | 1 | 18 |
| 28 | ybeD | Uncharacterized protein | 1 | 3 | 1 | 2 | 4 | 1 |
| 29 | clpX | ATP-dependent Clp protease ATP-binding subunit ClpX | 2 | 14 | 6 | 10 | 1 | 20 |
| 30 | infC | Translation initiation factor IF-3 | 2 | 0 | 3 | 0 | 0 | 18 |
| 31 | osmC | peroxiredoxin | 2 | 3 | 4 | 5 | 0 | 2 |
| 32 | ygaU | Uncharacterized protein | 2 | 4 | 2 | 2 | 3 | 37 |
| 33 | clpB | Chaperone protein ClpB | 3 | 25 | 9 | 11 | 9 | 86 |
| 34 | dksA | RNA polymerase binding transcription factor | 3 | 2 | 2 | 1 | 0 | 1 |
| 35 | tyrB | Aromatic-amino-acid ami- notransferase | 3 | 7 | 4 | 4 | 4 | 74 |
| 36 | ycaC | Uncharacterized protein | 3 | 5 | 3 | 8 | 2 | 93 |

A ftsH

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

B osmE

atg AAC AAG AAT ATG GCA GGA ATT CTG AGT GCA GCG GCG GTA TTA ACC ATG CTG GCG GGT TGT ACG GCT TAT GAT CGT ACC AAA GAC CAG TTT GTA CAG CCT GTG GTG AAA GAC GTC AAA AAA GGC ATG AGC CGG GCG GAG GTT GCA CAA ATT GCG GGT AAA CCT TCG TCT GAA GTG AGC ATG ATC CAT GCT CGC GGT ACT TGC CAG ACC TAC ATC CTG GGT CAA CGT GAT GGT AAA GCA GAA ACC TAC TTT GTC GCG TTA GAT GAT ACC GGA CAT GTC ATC AAC TCC GGT TAT CAG ACC TGT GCT GAA TAC GACACT GAT CCA CGT GGT CAA GT

C ibpA

atg CGT AAC TTT GAT TTA TCC CCG CTT TAC CGT TCT GCT ATT GGA TTT GAC CGT TTG TTT AAC CAC TTA GAA AAC AAC CAG AGC CAG AGT AAT GGC GGC GCC CCT CCG TAT AAC GTT GAA CTG GTA GAC GAA AAC CAT TAC CGC ATT GCT ATC GCT GTG GTG GTG GTT TTT GCT GAG AGC GAA CTG GAA ATT ACC GCC CAG GAT AAT CTG CTG GTG GTG AAA GGT GCT CAC CAC GAC GAA CTG GAA AAA GAG CGC ACC TAT CTG TAC CAG GGC ATC GCT GAA CGC AAC TTT GAA CGC AAA TTC CAG TTA GCT GAG AAC ATT CAT GTT CGT GGT GCT AAC CTG GTA AAT GGT TTG CTG TAT ATC GAT CTC GAA CGC GTG ATT CCG GAA GCG AAA AAA CCG CGC CGT ATC

D

groS

GAA ATC AAC taa

atg AAT ATT CGT CCA TTG CAT GAT CGC GTG ATC GTC AAG CGT AAA GAA GTT GAA ACT AAA TCT GCT GGC GGC ATC GTT CTG ACC GGC TCT GCA GCG GCT AAA TCC ACC CGC GGC GAA GTG CTG GCT GTC GGC AAT GGC CGT ATC CTT GAA AAT GGC GAA GTG AAG CCG CTG GAT GTG AAA GTT GGC G<mark>AC A</mark>TC GTT ATT TTC AAC GAT GGC TAC GGT GTG AAA TCT GAG AAG ATC G<mark>AC A</mark>AT GAA GAA GTG TTG ATC ATG TCC GAA AGC G<mark>AC A</mark>TT CTG GCA ATT GTT GAA GCC baa

Fig. (1). The locations of ACA sites in the mRNAs that encode stress-induced MazF-mediated proteins in *E. coli*. DNA sequences: (A) *ftsH*, (B) *osmE*, (C) *ibpA* and (D) *groS*. Out-of-frame ACA codons are highlighted in yellow. Sequences were taken from ecocyc.org (data-base for *E. coli*, Strain K-12). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

| Table 2. | Absence of ACA sites upstream to the AUG initiator and the locations of the threonine synonymous codons (ACA, ACG, |
|----------|--|
| | ACU, and ACC) in mRNAs of encoding proteins that are not stress-induced MazF-mediated E. coli proteins. |

| # | Gene Name | Protein Product | # ACA # ACA # AC In-frame Out-of- In-fra Frame | | # ACG In-frame | # ACC In-frame | # ACU In-frame | # ACA Up- stream AUG Initiator* |
|----|-----------|--|--|----|-------------------|-------------------|-------------------|---------------------------------------|
| 1 | mrdA | Peptidoglycan DD-transpeptidase MrdA | 7 | 20 | 10 | 19 | 2 | 0 |
| 2 | speF | ornithine decarboxylase, degradative | 5 | 23 | 6 | 8 | 5 | 0 |
| 3 | mrcA | peptidoglycan glycosyltransferase/ pepti- doglycan DD-transpeptidase MrcA | 5 | 27 | 12 | 22 | 6 | 0 |
| 4 | pgrR | DNA-binding transcriptional repressor | 5 | 5 | 5 | 3 | 4 | 0 |
| 5 | bamA | outer membrane protein assembly factor BamA | 5 | 27 | 9 | 36 | 7 | 0 |
| 6 | cynR | DNA-binding transcriptional dual regula- tor CynR | 4 | 12 | 5 | 4 | 2 | 0 |
| 7 | dnaG | DNA primase | 4 | 19 | 12 | 13 | 3 | 0 |
| 8 | mocA | molybdenum cofactor cytidylyltransferase | 4 | 4 | 0 | 6 | 3 | 0 |
| 9 | nemA | N-ethylmaleimide reductase | 4 | 6 | 5 | 5 | 3 | 0 |
| 10 | yneK | protein YneK | 4 | 17 | 5 | 5 | 4 | 0 |
| 11 | anmK | anhydro- <i>N</i> -acetylmuramic acid kinase 4 13 | | 4 | 11 | 3 | 0 | |
| 12 | dmlA | D-malate/3-isopropylmalate dehydrogen- ase (decarboxylating) | 3 | 7 | 3 | 6 | 3 | 0 |
| 13 | rlmF | 23S rRNA m ⁶ A1618 methyltransferase | 3 | 4 | 3 | 5 | 0 | 0 |
| 14 | csdA | cysteine sulfinate desulfinase 3 | | 9 | 5 | 6 | 6 | 0 |
| 15 | hdeD | acid-resistance membrane protein | 3 | 2 | 2 | 0 | 0 | 0 |
| 16 | melA | α-galactosidase | 3 | 11 | 11 | 11 | 4 | 0 |
| 17 | envY | DNA-binding transcriptional activator 3 EnvY | | 9 | 6 | 12 | 5 | 0 |
| 18 | gfcB | lipoprotein GfcB | 3 | 7 | 5 | 4 | 4 | 0 |
| 19 | nadC | quinolinate phosphoribosyltransferase (decarboxylating) | 3 | 7 | 6 | 7 | 4 | 0 |
| 20 | speC | ornithine decarboxylase, biosynthetic | 3 | 17 | 9 | 10 | 5 | 0 |
| 21 | ybfK | uncharacterized protein YbfK | 2 | 2 | 1 | 0 | 0 | 0 |
| 22 | ycdU | uncharacterized protein YcdU | 2 | 5 | 5 | 2 | 4 | 0 |
| 23 | hemA | glutamyl-tRNA reductase | 2 | 10 | 6 | 5 | 5 | 0 |
| 24 | trmL | tRNA (cytidine/uridine-2'- <i>O</i>)-ribose me- thyltransferase | 2 | 5 | 2 | 4 | 1 | 0 |
| 25 | aslB | putative anaerobic sulfatase maturation enzyme | 2 | 12 | 7 | 3 | 3 | 0 |
| 26 | gstB | glutathione S-transferase GstB | 2 | 2 | 3 | 2 | 2 | 0 |
| 27 | cirA | ferric dihyroxybenzoylserine outer mem- brane transporter | 2 | 23 | 17 | 26 | 5 | 0 |
| 28 | yaiC | diguanylate cyclase DgcC | 1 9 7 11 3 | | 0 | | | |

(Table 2) contd....

| # | Gene Name | Protein Product | # ACA # ACA # In-frame Out-of- In Frame | | # ACG In-frame | # ACC In-frame | # ACU In-frame | # ACA Up- stream AUG Initiator* |
|----|-----------|--|---|----|-------------------|-------------------|-------------------|---------------------------------------|
| 29 | Mog | molybdopterin adenylyltransferase | 1 | 3 | 7 | 3 | 3 | 0 |
| 30 | allR | DNA-binding transcriptional repressor AllR | 1 | 9 | 6 | 5 | 1 | 0 |
| 31 | msyB | acidic protein that suppresses heat sensitiv- ity of a <i>secY</i> mutant | 1 | 4 | 1 | 2 | 0 | 0 |
| 32 | yeaK | mischarged aminoacyl-tRNA deacylase | 1 | 3 | 1 | 3 | 2 | 0 |
| 33 | rcsA | DNA-binding transcriptional activator RcsA | 1 | 7 | 6 | 7 | 1 | 0 |
| 34 | sdaB | L-serine deaminase II | 1 | 13 | 1 | 13 | 6 | 0 |
| 35 | hslV | HslV hexamer | 1 | 6 | 2 | 5 | 4 | 0 |
| 36 | zapA | cell division protein ZapA | 1 | 5 | 0 | 2 | 4 | 0 |
| 37 | dapF | diaminopimelate epimerase | 1 | 2 | 2 | 7 | 1 | 0 |
| 38 | sdiA | DNA-binding transcriptional dual regula- tor SdiA | 1 | 4 | 4 | 3 | 3 | 0 |
| 39 | oxyR | DNA-binding transcriptional dual regula- tor OxyR | 1 | 7 | 2 | 6 | 3 | 0 |
| 40 | caiA | crotonobetainyl-CoA reductase | 1 | 13 | 3 | 13 | 2 | 0 |
| 41 | pflB | pyruvate formate-lyase (inactive) | 1 | 34 | 1 | 28 | 18 | 0 |
| 42 | osmB | osmotically-inducible lipoprotein OsmB | 1 | 1 | 2 | 3 | 1 | 0 |
| 43 | sdiA | DNA-binding transcriptional dual regula- tor SdiA | 1 | 4 | 4 | 3 | 3 | 0 |
| 44 | proQ | RNA chaperone ProQ | 1 | 9 | 3 | 3 | 0 | 0 |
| 45 | uxaC | D-glucoronate/D-galacturonate isomerase | 0 | 16 | 1 | 14 | 9 | 0 |
| 46 | ackA | acetate kinase | 0 | 19 | 1 | 14 | 6 | 0 |
| 47 | Pgk | phosphoglycerate kinase | 0 | 11 | 0 | 13 | 7 | 0 |
| 48 | sugE | quaternary ammonium compound efflux pump | 0 | 0 | 4 | 2 | 2 | 0 |
| 49 | fimE | regulator for fimA | 0 | 0 | 2 | 5 | 4 | 0 |
| 50 | roxA | ribosomal protein-arginine oxygenase | 0 | 9 | 1 | 3 | 2 | 0 |

*100 nucleotides upstream to AUG initiator.

in-frame ACA codons (Table 1). Therefore, it is suggested that, in such cases, it may provide a stress-operating mechanism for processing mRNAs by MazF within the coding sequence, thereby reducing high levels of proteins that are probably produced because of elevated levels of transcription. Finally, might an ACA codon located out-of-frame have a regulatory function for the synthesis of a protein under stressful conditions? If stress also induces frame-shifting, the out-of-frame MazF ACA cleavage may serve as a mechanism to prevent STM translation through frame-shifting. Accordingly, the destructive effect of frame-shifting due to extensive stress locates the originally out-of-frame ACAs into the open reading frame, thereby exposing them to MazF cleavage. Thus, the out-of-frame ACAs serve as safeguards for accurate translation under stress.

Genomic screening is generally used to detect *cis* and *trans* elements as well as mutations. This is the first report showing the use of genomic screening for the detection of proteins involved in a metabolic state, specifically, stress-induced MazF-mediated proteins in *E. coli*. The presence of an ACA MazF cleavage site within the span of 100 nucleo-tides upstream from the AUG initiator, along with the described bias in the genetic code, will probably also be found in additional bacteria in which MazF cleaves ACA sites.

A rlmF

atg TCC GCC CAG AAA CCG GGG TTG CAT CCG CGC AAC CGT CAT CAC AGC CGC TAC GAT CTC GCC ACG CTT TGT CAG GTC AAT CCT GAA CTC AGG CAA TTC CTC ACG CTT ACA CCC GCC GGG GAG CAA AGC GTA GAC TTT GCC AAT CCG CTG GCG GTG AAG GCG CTC AAT AAG GCG TTG CTG GCC CAT TTT TAC GCC GTA GCG AAC TGG GAT ATC CCC GAC GGT TTT CTC TGC CCA CCA GTA CCG GGC CGG GCG GAT TAT ATT CAT CAC CTT GCC GAT TTA CTG GCA GAA GCA AGC GGA ACA ATT CCA GCG AAT GCC AGC ATT CTG GAT ATC GGC GTT GGT GCG AAC TGT ATT TAT CCG CTG ATT GGC GTA CAT GAA TAT GGC TGG CGT TTT ACC GGT AGT GAA ACC AGC AGC CAG GCG TTA AGC AGT GCG CAG GCG ATT ATC AGT TCT AAT CCG GGG CTT AAC CGC GCC ATT CGT CTG CGT CGG CAA AAA GAG AGT GGG GCG ATT TTT AAC GGC ATC ATC CAT AAA AAC GAG CAA TAC GAC GCG ACC TTG TGT AAC CCG CCA TTC CAC GAT TCC GCC GCT GCG GCA CGG GCA GGT AGT GAG CGT AAA CGC CGT AAC CTG GGG CTG AAC AAA GAC GAT GCA CTG AAC TTT GGC GGC CAG CAA GAG TTG TGG TGT GAA GGC GGT GAA GTC ACC TTT ATC AAA AAG ATG ATT GAA GAG AGC AAA GGC TTC GCG AAG CAG GTG ATG TGG TTT TA TA CTG GTA TCT CGT GGT GAA AAC TTA CCG CCG TTG TAT CGT GCC CTG ACG GAC GTG GGC GCG GTG AAG GTG GTT AAA AAA GAG ATG GCC CAG GGG CAA AAG CAG AGT CGC TTT ATT GCC TGG ACC TTT ATG AAC GAC GAG CAG CGC CGC CGT TTT GTC AAT CGC CAG CGT taa

^B nemA

atg TCA TCT GAA AAA CTG TAT TCC CCA CTG AAA GTG GGC GCG ATC ACG GCG GCA AAC CGT ATT TTT ATG GCA CCG CTG ACG CGT CTG CGC AGT ATT GAA CCG GGT GAC ATT CCT ACC CCG TTG ATG GCG GAA TAC TAT CGC CAA CGT GCC AGT GCC GGT TTG ATT AGT GAA GCC ACG CAA ATT TCT GCC CAG GCC AAA GGA ATA TGCA GGT GCG CCT GGC ATC CAT AGT CGC GAG CAA ATT GCC CGC GGC AAA GGA CAT GCC GGT TGG ATT ATT AGT GAA GCC ACG CAA ATT TCT GCC CAG GCC AAA GGA AAA ATC ACC CGC GGC CAT CCT CAT GCT GCA GCC TG CAA CTT GCC CGG CGC CGC CGC CGC CAT CCT CAC GCC GGC CCT GCA CCT GCC AGC CCG GTA GCG CCT TCC AC GCT GCC CTG CAC CCT GCA CCT GCG CGC CGG GTA GCG CCT TCA GCC AGC CTG CAA CCT GCC GGC CGC CAA CGG GCA GCG CCT TCC GCC GCC CGC CCT CCAC CCT GCC GCC GAA CGG GCA GCG CCT TCA GCC GTC CAAT GAT TTC GCC CGT GCG CTT GAA CTG GAA GAG GCT TCC GT CGT GTA GAA MC TCC GC GGT GCC TTG GAA CGG GCT GAA GCC GGT TTT GAT CGT CAAT GAT TTC CGT CAG GCC ATT GCT AAC GCG GCT GAA GCC GGT TTT GAT CGT CAT GAT CTT CCT CCT CCC GCT GAA CGG GCT GAA ACCC GGT TTT GAT CGT CAAT GAT TTC GCT CAC GGT CAT CCC GGG ATT GAA AAA TCG GCT GCC GAT CGC ATT GGC ATT CGC CAT CGC GAT CCC GGG ATT GAA GAA TGG GGT GCC GAT CGC ATT GGC ATT GCC GAT CAC GGC ACC TTT CCAC GGC GGG GGT GAA CGC GGA TGCA GAG GGC GAC CTG TAT CTG ATT GAA CTG GGG GGT GAA CGC GGA TGCC GAT GCC GAT GCC GAT GCC GGG ATT GGA CGC GGG GGT GAA CGC GGC ATT GCT TAT CTG CAT ATG TCA GAA CCA GAT TGG GCG GGG GGA ACC GGC ATT CGC ATT CGC CAT TC GGT ACT TTC CAC GGT CGG ATA AC GGC CGA AT ACT GAT GCG TTC CGC GAA AAA GCT GAA CGC CGA TTC CGC GAT CGC GAT CGC GAT GCA ATA GAA AAA GCT GAA ACG CTG ATC GGC AAA GGT TTA ACT GAC GGC GGG GGA AAA GCT GAC GAC CTG ATC CGC GAT GCG GAT CAC GGT GCA TTT GGT CGT GAC TGG ATT CCG CGT GCG GAT CGC GAA AGT TTC CGC GCC CGT TTC CAC GGT GAA CCG GGT GAA AGT TTC GCC CCC CTTG CAG CGC AAA GCT GAC CGA CTG ATC CGA GCT GCT GAC CGG GGG GGA AGT TTA CTG GCT GAC CGA CTG GGC TAT ACC GAT TAC CGA GCT GCT GAC CGG AAA GGT TTA CGC GCC GAA AGT TTC CGC GCC GAA AGT TTC CGC GCC GAA AGT TTC CCC ACC GGT GAA CGT GAC CGG GAA GGT TTA ACC GGC GAA AGT TTA CCG ACC GAA AGT TTC CCCA

C mocA

atg TCA GCC ATC GAC TGT ATA ATT ACC GCT GCC GGA TTA TCA TCA AGA ATG GGG CAA TGG AAA ATG ATG TTA CCC TGG GAA CAG GGA CAT ATT CTT GAT MCAGATATC AAA AAT GCG TTG CAG GTT TGT AGC CGA ATT ATT TTA GTC ACC GGC TAT CGT GGT AAT GAA CAG CAC GAA CGT TAT GCG AAC CAG AGC AAT ATC ACT ATT ATT CAC AAC CCA GAT TAT GCG CAG GGT TTA CTA CGA CAG AGC AAT ATC ACT ATT ATT CAC AAC CCA GAT TAT GCG CAG GGT TTA CTA CGA TCC GTC AAG GCC GCA GTA CCC GCA GTG CAA CAT GAA CAT TGT TTT CTC ACC CAC GGT GAC ATG CCA ACC CTC ACC ATC GAT ATT GTC AAA ATC TGG TCA GTA CGA AAT GAT GCC GCA ATA CTG CCA CTC CAT AAT GGC ATC CCC GGC CAT CCG ATT TTA GTG TCA AAA CCA TGC CTG ATG CAG GCA ATC CAA CGA CCC AAT GAT ACC AAT GAT GCC GCA ATA CTG CCA CTC CAT AAT GGC ATC CCA GGA CCC AAT GCC ACC AAT GAT GCC CAA AAA CCA TGC CTG ATG CAG GCA ATC CAA CGA CCC AAT GAT AAA AAT GCA GCA ATA ATT TTA GAT ACT CCG GAC GAC CTT ATT ATA GAA AAT GCA GAA ATA ATT TA GAT ATT GAT ACT CCC GAC CAT CTG AAA AAT GCA GAA AAG TAT ACT GAA ATT TAG

D envY

atg CAA TTG AGC AGC AGT GAA CCT TGC GTG GTG ATC CTG ACC GAA AAA GAG GTA GAG GTA AGC GTC AAT AAC CAT GCT ACG TTT ACC CTT CCG AAA AAC TAC CTG GCC GCC TTC GCG TGC AAC AAC AAC ACG GTC ACG TTT ACC CTT CCG AAA AAC TAC CTG GCC GCC TTC GCG TGC AAC AAC AAC GTC ATT GAA CTC TCA ACG TTA ATT CAC GTA TTA ATC ACC CAC ATC AAC GTC AAT GAA CTC AC GAT TAT CTG TTG TTT TTA AAT AAG AAC TTA ACC CGT GTG GAA GCC TGG CGG GCG ACC CCG GTT ATC GCT TGT CAT ACC CGT GTG CGG CAA GTG TTC CGG CTA GCC GCC AAC CAC AGC AAG CAA GCA ACC AGC AGA CCC TGG CGA GCG GGG GTG AGC GCC AAC CAC AGC AG CAA CCC AGC AGA CCC GCG GAA GTG TTC CGG GTA GCC GCC AAC CTG CTT TTA ACC TTA TCT CAT ACC CGT TTC TTG GAC CAA CGC GGG GTT ATT GCC CTA CTG ATG TAT ATC TTA CGC AGC AGC GCC GCC GAC ACC CGC GCC ATC ACT CTG ATG ATA TTC CTA CGC AGC AGC AGC GTC GC GCC ATT ATT CAA AGC GAT ATT CAG CAT TAC TGG AAT CTG CGA ATT ACC AGT TCG CTA TGT TTA AGC CCC AGC CTG CTC AAA AAG AGA ATT AAA AAC GAA AAT ACC AGC TAT AGC CAG ATT GTC GGG GTG CGC CAA TTA TAG AGA CAAT ACC AGC TAT AGC CAG ATT GTC GGG GTG CGC CAA TTA TTG GGC TTA AGC AGC AGG TAC ATT ATC CGC GAG CGT CGC GCG CAA TTA TG GGC TAT AGC AGC AGG TAC TTA ATC ACT CTG GTA AGG GGG TTA ATT CGG CTA TGC TTA AGC CAG ATT GTC GGG GTG CGC CAA TTA TG GGC TAT AGC AGC AGG TAC TTA ATC ATC TCT GTT TTA AGG CGG CTA TAT TG GGC TAT AGC AGC AGG TAC TTA ATC TTC TGT GTT TTA AGG CGG CTA TAGC AGC AGG TAC TTA TCT CTG CTA TTA TAG AGG GTT TA TTA CAA AGC AAT ACC TTA ATC ACT CAG GTG CGC CAA TTA TG GGC TAT AGC AGC ACG TAC TTA TCT CTC GTT TTA AGG CGG TTA TA CGG CGG CTG TA AGC CAG TAC TTA ATC TCT CTG GTA AAA AAA AAA AAA GTG ATA ACC GGC CTG TG AAT TAT CTC CGC AAA CAA

E gfcB

gtg CGC CCT CTT ATT TTA TCG ATT TTC GCA CTA TTT CTT GCG GGA TGT ACG CAC ÅGC CAG CAA AGT ATG GTC GAT ATT TC GCA CAG CCT TTC CGA AAT CAG GAT ATC ACC GTA GCG GAT CAG CAG ATC CAG GCG TTG CCT TAT TCC ACT ATG TAT TTA CGC CTT AAT GAA GGG CAA CGA ATC TTT GTG GTA CTG GGA TAT ATA GAA CAA GAA CAA AGC AAA TGG TTA TCC CAG GAT AAC GCC ATG CTG GTT ACC CAC AAT GGA CGT CTT TTA AAA ACC GTC AAA CTT AAT AAT CAT CTG CTG GAA GTG AGT ATT CC GGG CAG CAC CTT CTG CGA TAT AAT AAT CTG CTG GAA GTG GCG ACT CTT TA TAAA ACC GTC AAC CTT ATT AAT AAT CTG CTG GAA GTG GCG CGT CGT CTT TA TAAT ACT CTT GGA GAT GAC CAT TTT CGC TCT GCG ACC CTG AGG CGC GAT ATT CTC TGG AGT GAA GAC CAT TTT CGC TCT GCG ACC CTG AGT TCT ACT TTT TCC TTT GCT GGA TTA GAG ACG CTG AAT ATT GCG GGT CGC AAT GTG CTG TGT AAT GTC TGG CAG GAA GAG GTG ACT TCC ACG CGG CCA GAA AAG CAG TGG CCA AAC CTT TT CCC ACG CGG CCA GAA AAG CAG TGG CAA AAC CTT TC TGG GC CAG GAA GAG GTG ACT TCC ACG CGG CCA GAA AAG CAG TGG CAA AAC CTT TAT CGG GCC GAT CGG CCA GTT GCT CAG CAG GTT CGT CAA AGT CGG CAA AAC CCG CGG CCG GGT GTT ATT CCC GTA AAT GCG GTT CGT CAA AGT CGG CAA ACG CCA TGG GTA ATT CCC GGA GTT CTT CTA AAA CCC GCA CCA tga

Fig. (2). The locations of ACA sites in mRNAs encoding proteins that are not mediated by MazF in *E. coli*. DNA sequences: (A) *rlmF*, (B) *nemA*, (C) *mocA*, (D) *envY* and (E) *gfcB*. In-frame ACAs are highlighted in red. Out-of-frame ACAs are highlighted in yellow. Sequences were taken from ecocyc.org (data base *for E. coli*, Strain K-12). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

| Table 3. | Frequency of each synonymous codons for threonine: ACA, ACG, ACU, and ACC in MazF-mediated mRNAs compared |
|----------|---|
| | with non MazF-mediated mRNAs. |
| | |

| А | A B C | | D |
|---------------|--|-------|---|
| Codon Triplet | Codon Triplet Published data* mRNAs of proteins not mediated by MazF (Table 2) | | mRNAs of MazF-mediated proteins (Table 1) |
| ACA | 17% | 12% | 5.0% |
| ACU | 19% | 19.5% | 25.4% |
| ACC | 40% | 44% | 53.8% |
| ACG | 25% | 24.5% | 15.8% |

* Published data from www.genscript.com/tools/codon-frequency-table.

CONCLUSION

This study has revealed that in the mRNAs of most of the stress-induced MazF-mediated *E. coli* proteins translated by the Stress-induced Translation Machinery (STM), the threonine codon ACA is not located in-frame but rather out-offrame, while its synonymous codons, ACG, ACU, and ACC are in-frame. In contrast, in most of the mRNAs of the proteins translated by the canonical translation system, an inframe ACA was found. Thus, it is suggested that this bias in the reading of the genetic code is characteristic of *E. coli* genes specifying for stress-induced MazF-mediated proteins. This is the first report showing the use of genomic screening for the detection of proteins involved in a metabolic state, stress-induced MazF-mediated proteins in *E. coli*.

AUTHORS' CONTRIBUTIONS

Experiments were performed by A.N. and A.O.G. The paper was written jointly by A.N., A.O.G., and H.E.K. The project was directed by H.E.K.

ETHICS APPROVAL AND CONSENT TO PARTICI-PATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

This research, carried out in the laboratory of Hanna Engelberg-Kulka, was supported by grant number 1030/19 from the Israel Science Foundation (ISF), administrated by the Israel Academy of Sciences and Humanities.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

We thank F.R. Warshaw-Dadon (Jerusalem, Israel) for her critical reading of the manuscript. We thank Arthyom Philipovich (Kazan Russia) for developing the HTML script and Alexander Mager (Jerusalem, Israel) for providing us with this script.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

- [1] Aizenman, E.; Engelberg-Kulka, H.; Glaser, G. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine [corrected] 3',5'-bispyrophosphate: a model for programmed bacterial cell death. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*(12), 6059-6063. http://dx.doi.org/10.1073/pnas.93.12.6059 PMID: 8650219
- [2] Engelberg-Kulka, H.; Glaser, G. Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu. Rev. Microbiol.*, 1999, 53, 43-70. http://dx.doi.org/10.1146/annurev.micro.53.1.43 PMID: 10547685

- [3] Mittenhuber, G. Occurrence of mazEF-like antitoxin/toxin systems in bacteria. J. Mol. Microbiol. Biotechnol., 1999, 1(2), 295-302. PMID: 10943559
- [4] Pandey, D.P.; Gerdes, K. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.*, 2005, 33(3), 966-976.
- http://dx.doi.org/10.1093/nar/gki201 PMID: 15718296

 Hazan, R.; Sat, B.; Engelberg-Kulka, H. *Escherichia coli* mazEFmediated cell death is triggered by various stressful conditions. *J. Bacteriol.*, 2004, 186(11), 3663-3669.
http://dx.doi.org/10.1128/JB.186.11.3663-3669.2004 PMID: 15150257

[6] Godoy, V.G.; Jarosz, D.F.; Walker, F.L.; Simmons, L.A.; Walker, G.C. Y-family DNA polymerases respond to DNA damageindependent inhibition of replication fork progression. *EMBO J.*, 2006, 25(4), 868-879.

http://dx.doi.org/10.1038/sj.emboj.7600986 PMID: 16482223

[7] Sat, B.; Hazan, R.; Fisher, T.; Khaner, H.; Glaser, G.; Engelberg-Kulka, H. Programmed cell death in *Escherichia coli*: some antibiotics can trigger mazEF lethality. *J. Bacteriol.*, 2001, 183(6), 2041-2045.

http://dx.doi.org/10.1128/JB.183.6.2041-2045.2001 PMID: 11222603

- Zhang, Y.; Zhang, J.; Hoeflich, K.P.; Ikura, M.; Qing, G.; Inouye, M.; Maz, F. MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in Escherichia coli. *Mol. Cell*, 2003, *12*(4), 913-923. http://dx.doi.org/10.1016/S1097-2765(03)00402-7 PMID: 14580342
- [9] Zhang, Y.; Zhang, J.; Hara, H.; Kato, I.; Inouye, M. Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. J. Biol. Chem., 2005, 280(5), 3143-3150. http://dx.doi.org/10.1074/jbc.M411811200 PMID: 15537630
- [10] Vesper, O.; Amitai, S.; Belitsky, M.; Byrgazov, K.; Kaberdina, A.C.; Engelberg-Kulka, H.; Moll, I. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *E-scherichia coli. Cell*, 2011, 147(1), 147-157.
- http://dx.doi.org/10.1016/j.cell.2011.07.047 PMID: 21944167 [11] Sauert, M.; Wolfinger, M.T.; Vesper, O.; Müller, C.; Byrgazov, K.;
- [11] Saderi, M., wohniget, M.T., Vespel, O., Muller, C., Bylgazov, K., Moll, I. The MazF-regulon: a toolbox for the post-transcriptional stress response in *Escherichia coli*. Nucleic Acids Res., 2016, 44(14), 6660-6675. http://dx.doi.org/10.1093/nar/gkw115 PMID: 26908653
- [12] Moll, I.; Engelberg-Kulka, H. Selective translation during stress in Escherichia coli. Trends Biochem. Sci., 2012, 37(11), 493-498. http://dx.doi.org/10.1016/j.tibs.2012.07.007 PMID: 22939840
- [13] Nigam, A.; Ziv, T.; Oron-Gottesman, A.; Engelberg-Kulka, H. Stress-induced MazF-mediated proteins in *Escherichia coli*. *MBio*, 2019, 10(2), e00340-e19. http://dx.doi.org/10.1128/mBio.00340-19 PMID: 30914510

 [14] Oron-Gottesman, A.; Sauert, M.; Moll, I.; Engelberg-Kulka, H. A Stress-induced bias in the reading of the genetic code in *E-scherichia coli*. *MBio*, **2016**, 7(6), e01855-e16. http://dx.doi.org/10.1128/mBio.01855-16 PMID: 27935840

- [15] Aoi, R.; Miyamoto, T.; Yokota, A.; Ota, Y.; Fujitani, H.; Tsuneda, S.; Noda, N.; Maz, F. MazF endoribonucleolytic toxin conserved in *Nitrospira* specifically cleaves the AACU, AACG, and AAUU motifs. *Toxins (Basel)*, **2020**, *12*(5), E287. http://dx.doi.org/10.3390/toxins12050287 PMID: 32365819
- [16] Rice, P.; Longden, I.; Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, 2000, *16*(6), 276-277. http://dx.doi.org/10.1016/S0168-9525(00)02024-2 PMID: 10827456