What causes antibiotic resistance?

Authors: Anna Rybina, Daria Nikanorova

Introduction

Antibiotic medications are widely used in the treating and averting bacterial infections. However, as time passes, the number of bacterial pathogens that developed mechanisms against antibacterial agents increases. Resistance to antimicrobial compounds has become one of the major challenges in the healthcare community, threatening mankind's battle against infectious diseases. There are several main mechanisms of resistance: “limiting uptake of a drug, modification of a drug target, inactivation of a drug, and active efflux of a drug” (Reygaert 2018).

It is essential to understand biochemical and genetic basis underlying resistance against certain antibiotics to modulate therapy and optimize treatment. One of the approaches is to compare the genome of resistant bacterial strain with the reference not resistant one to elucidate significant single nucleotide polymorphisms (SNPs) in genes that could be associated with antibiotic targets.

In this work, we aimed to identify SNPs that might provide ampicillin resistance to *Escherichia coli* strain, propose possible mechanisms of this resistance and suggest alternative treatment.

Methods

Shotgun sequencing of the genome of *E. coli* strain resistant to ampicillin was performed on Illumina platform in paired-end mode. Raw sequencing data in fastq format were manually inspected for format correctness, number of reads and reads length using bash **(Supplementary materials, section 2)**. Quality of reads was checked via FastQC v0.11.9 (Andrews, S) (Supplementary materials, section 3). Using Trimmomatic v.0.39 (Bolduc, n.d.), reads were filtered and trimmed in paired end mode and Phred-33 scale with the quality score of 20 and 30 **(See parameters in supplementary materials, section 4)**. Changes in quality of processed reads was additionally verified using FastQC.

Reference genome sequence (in fasta format) and genome annotation (in gff format) of *E. coli* str. K-12 substr. MG1655 were downloaded from GenBank (assembly accession: GCA_000005845.2) **(Supplementary materials, section 1)**. Both raw and two types of trimmed sequencing data were subjected to downstream analysis. To align reads to the reference genome, we used BWA tool v0.7.17-r1188 (H. Li and Durbin 2009)) **(Supplementary materials, section 5)**. Reference genome was indexed with default parameters. Reads were mapped implementing the BWA-MEM algorithm with default parameters. Resulting alignment in SAM format was compressed by converting it to BAM format, then sorted and indexed using SAMtools v1.9 (H. Li et al. 2009). Basic statistics and summary of the coverage of mapped reads at a single base pair resolution were also obtained using SAMtools **(See commands and parameters in Supplementary materials, section 5)**. VarScan v2.4.1 (Koboldt, Larson, and Wilson 2013) was applied for variant calling specifying minimum variant frequency equal to either 50 % or 10 % and reporting significant variants above the threshold in vcf format. Obtained BAM and VCF files along

with GFF annotation were visualized and explored in IGV browser (Robinson et al. 2011) **(Supplementary materials, section 6)**. Genes with revealed SNPs were searched for association with antibiotic resistance in UniProt (https://www.uniprot.org/), EcoliWiki (https://ecoliwiki.org/colipedia/index.php/Welcome_to_EcoliWiki) and PubMed (https://pubmed.ncbi.nlm.nih.gov/). To understand the possible role of identified substitutions in ensuring drug resistance, structures of respective proteins were examined in RCSB PDB database.

Results

Quality control, trimming and alignment

In this study, results of sequencing of 455876 Illumina pair-end reads were obtained. According to manual and FastQC analysis of raw sequencing data, average read length was 101 bp. For both fastq files (forward and reverse reads), FastQC output reported warnings in per base sequence content and per sequence GC content. Per tile sequence quality was reported as a warning for reverse reads and was failed in case of forward reads. It might be due to transient problems with the run (bubbles into a flow cell for example). Test for per base sequence quality failed for both samples as might be explained by degrading quality of calls while run progresses (**Fig. S1**). Although our reads have bad quality (fall into red area on graph) only at the end (at length > 95 bp), we decide to trim nucleotides from both ends of our reads in order to improve quality.

Table 1 Number of reads in analysis

Reads	**Number of reads**	**Reads after trimming (% of total)**		**Aligned reads (% of total)**	
		quality score 20	*quality score 30*	*quality score 20*	*quality score 30*
All	911752	892518 (97.89%)	752680 (82.55%)	891649 (97.76%)	752200 (82.48%)
Forward	455876	446259	376340	446038	376080
Reverse	455876	446259	376340	445205	375982

After trimming with the quality score of 20 and 30, 97.89% (446259) and 82.55% (376340) of reads were kept for downstream analysis, respectively (**Table 1**). Trimming with a quality score of 20 resulted in better per base sequence quality: mean quality score increased to 28 and in case of trimming with quality score of 30 even higher (**Fig. S2, S3**). However, sequence length distribution became skewed, varying from 20 to 101 where value of 101 was likely to be mode. Filtering has improved the overall quality of the read sequence data. 99.92% (752200) and 99.87% (891649) of reads were mapped after trimming by quality score of 30 and 20, respectively. Without the prior trimming, 99.87% of reads (910940) were aligned to the reference genome.

Variant calling

In order to distinguish real mutations from the sequencing errors we selected only those bases that differ from a reference sequence in more than 50% of reads mapped on this site. We obtained the same 5 SNPs in all 3 cases: after trimming with a quality score of 20, after trimming with a quality score of 30 and without trimming (**Table 2**).

Table 2 SNPs in *E. coli* ampicillin-resistant strain identified in the study

Position	Base	Coding DNA	Template DNA	mRNA	Amino acid	Gene
93043	C -> G	GCC->GGC	GGC->GCC	GCC->GGC	A -> G	*ftsI*
482698	T -> A	CTG->CAG	CTG->CAG	CAG->CUG	Q -> L	*acrB*
852762	A -> G					*rybA*
3535147	A -> C	GTA->GGA	TAC-> TCC	GUA->GGA	V -> G	*envZ*
4390754	G -> T	GCC->GCA	GGC->TGC	GCC->GCA	A -> A	*rsgA*

Mutation in *rsgA* leads to a synonymous substitution in protein and does not alter its function. Mutation in *rybA* affects a small non-coding RNA rybA. The small non-coding RNA RybA regulates the biosynthesis of aromatic amino acids under peroxide stress in *E. coli* and presumably does not provide antibiotic resistance (Gerstle et al. 2012). Missense mutations in *ftsI*, *acrB*, *envZ* genes result in amino acid substitution in corresponding proteins and may be associated with increased ampicillin resistance.

Protein structure examination

We inspected structures of three proteins, AcrB, FtsI and EnvZ (**Table 1**) using RCSB PDB database (Berman et al. 2000). In case of AcrB protein, we found substitution Q569L in the loop preceding an α-helix (Fig. 1). Based on Kobylka et al. 2020 work (see Fig.S2 from Kobylka et al. 2020), we suggested that substitution Q569L is situated in the periplasmic loop right before one of the α-helices of so-called porter domain.

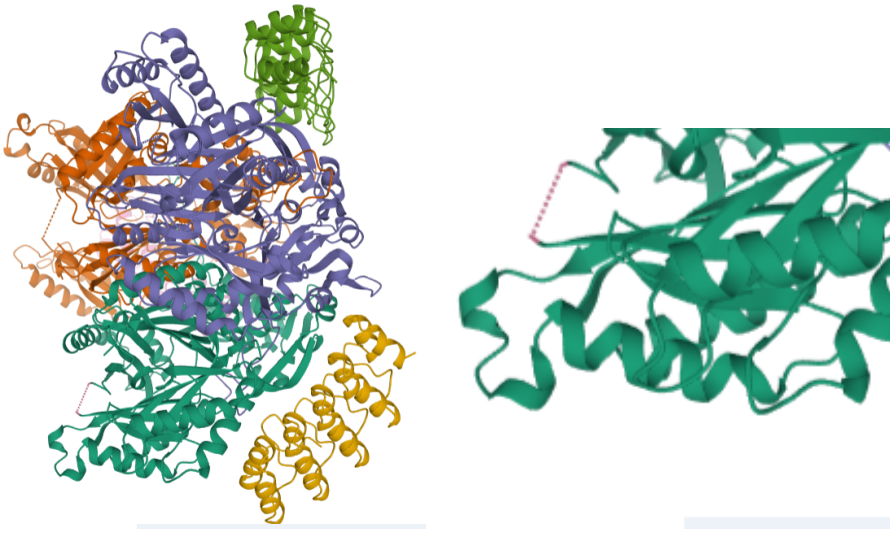

Figure 1. Structure of bacterial efflux pump AcrB from *E. coli* (PDB ID: 5EN5 (Sjuts, Ornik, and Pos 2016). Q569 marked pink is found in the loop. Image was created using Mol* from RCSB PDB database (Berman et al. 2000).

As FtsI protein structure of *E. coli* with ligand ampicillin was not found in the RCSB PDB database, we examined the available resolved structure of FtsI interacting with penicillin (Fig. 2). We assumed that structurally similar ampicillin and penicillin should have similar binding sites in FtsI. We observed that the position of A544G substitution was close to residues responsible for binding ligand penicillin. One may suggest that the identified mutation could change the functioning of the FtsI and affect the binding of FsI to ampicillin.

Figure 2. Crystal structure of transpeptidase domain from *Escherichia coli* FtsI protein in complex with penicillin (PDB ID: 6I1I, (D. Bellini et al. 2019). A544 marked pink is juxtaposed to amino acid residues involved in binding penicillin. Image was created using Mol* from RCSB PDB database (Berman et al. 2000).

According to our findings, the substitution V241G in EnvZ protein is located in the histidine kinase domain, specifically within the α-helix. The position 241 corresponds to the region that probably plays an important role in the EnvZ dimerization (Kishii et al. 2007).

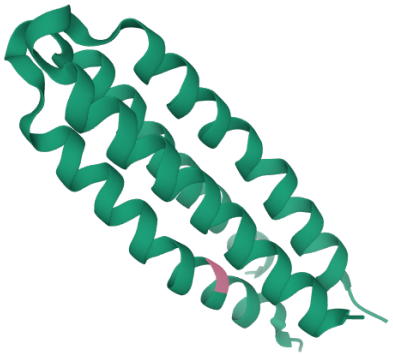

Figure 3. Histidine-containing phosphotransfer (DHp) domain structure of EnvZ from *Escherichia coli.* (PDB ID: 5B1N, (Okajima et al. 2016)). V241 marked pink is located within α-helix. Image was created using Mol* from RCSB PDB database (Berman et al. 2000).

Discussion

In order to investigate specific molecular mechanisms of bacterial resistance to ampicillin, we mapped results of sequencing of ampicillin resistant *E. coli* strain (pair-end reads) on the reference sequence of not-resistant *E. coli* strain, and obtained 5 SNPs (Table 2). Among them only three may cause resistance to ampicillin.

1. acrB

Gene *acrB* is annotated in EcoliWiki as a multi-drug efflux pump gene. Antimicrobial resistance can be provided by the production of efflux systems capable of pumping a toxic molecule out of the bacterial cell (Munita and Arias 2016). Being a component of the AcrAB-TolC system, AcrB is a transporter protein in the inner membrane (Du et al. 2018). RND pumps are able to transport a wide array of molecules: tetracyclines, chloramphenicol, some β-lactams, novobiocin, fusidic acid and fluoroquinolones (Munita and Arias 2016).

Mutation in *acrB* may lead to higher activity of AcrAB-TolC system or more specific binding and pumping out ampicillin. This mechanism reduces the amount of ampicillin in bacteria, thereby reducing its killing effect.

It is important to notice that three domains could be distinguished in AcrB: funnel domain (or docking domain) responsible for AcrB trimerization, transmembrane domain (in charge of energy transduction for drug efflux) and the porter domain, containing binding pockets and facilitating substrate uptake and translocation (Kobylka et al. 2020). In this study we detected Q569L which probably occurred in the loop prior to the porter domain (Fig. 1). It should be noted that as a rule polar amino acid residue glutamine prefers to be in aqueous environment (for example, periplasm, which is consistent with our suggestion about position of Q569) while non-polar hydrophobic leucine tends to be localized in hydrophobic protein core or/and within α-helices (Barnes and Gray 2003). Thus, we assume that the substitution Q569L may cause critical conformational changes in the porter domain and juxtaposed periplasmic loop resulting in elevated ability to bind ampicillin and effective transportation of the antibiotic out of the cell.

2. FtsI

According to EcoliWiki, the product of gene *ftsI* is an essential cell division membrane protein FtsL (also referred to as PBP3 - penicillin bind protein 3). It was shown that FtsL binds not only with penicillin, but also with ampicillin (Tormo et al. 1986). This protein is a part of a divisome - a large molecular machine, which carries out the process of division (Söderström and Daley 2017). FtsL is one of transpeptidases found in *E. coli*, and it catalyzes the cross-linking of cell wall peptidoglycans during cell division. (Donachie 1993).

We assume that mutation in FtsL may prevent the binding of ampicillin to it, causing a resistance to this antibiotic. Identified SNP in *ftsI* gene corresponds to substitution A544G in FtsI which is localized in its penicillin binding protein transpeptidase domain (Dom Bellini et al. 2019). Unfortunately, FtsI protein structure of *E. coli* with ligand ampicillin was not found in RCSB PDB database (Berman et al. 2000). Assuming that structurally similar ampicillin and penicillin have similar binding sites in FtsI, we examined the available resolved structure of FtsI interacting with penicillin (Fig. 2). The position of A544G substitution is close to residues responsible for binding ligand penicillin. It is known that glycine has more conformational flexibility than alanine (Barnes and Gray 2003). Displacement to glycine could probably affect conformational properties of FtsI transpeptidase domain leading to either decreased affinity of ampicillin binding or even loss of this ability and providing bacteria with resistance to the antibiotic.

3. EnvZ

Gene product of *envZ* is an osmolarity sensory histidine kinase/phosphatase EnvZ. In response to environmental osmolarity changes, EnvZ controls the phosphorylation state of the transcription factor OmpR. Phosphorylated OmpR regulates the expression levels of outer membrane porin proteins OmpF and OmpC (Cai and Inouye 2002). Porins OmpF and OmpC allow hydrophilic molecules such as β-lactams (penicillin, ampicillin) to cross the outer membrane. Thus the intracellular level of these antibiotics is particularly affected by changes in permeability of the outer membrane (Munita and Arias 2016). It was reported that OmpF facilitates penetration for some antibiotics, while OmpC could either promote antibiotic entering or impede it (Choi and Lee 2019)Increased amount of phosphorylated OmpR might result in inhibiting ompF expression and activating ompC expression (Harlocker, Bergstrom, and Inouye 1995).

We found a SNP in gene *envZ* in the genome of *E.coli* drug-resistant strain, which results in missense substitution (V -> G) in the dimerization region of the histidine kinase domain (Okajima et al. 2016)). This mutation could affect the functioning of EnvZ leading to an increased phosphorylation of OmpR and, consequently, to a reduced number of OmpF in the outer membrane. As a result, decreased permeability of the outer membrane might prevent ampicillin from reaching its periplasmic target.

Importantly, changes in permeability usually provide only low-level resistance and are often associated by other mechanisms, such as increased number of efflux pumps (Munita and Arias 2016). Similarly, mutation only in the *ftsI* gene did not provide antibiotic resistance, whereas *ftsI* and *envZ* gene mutations increased the resistance of bacteria greatly (Adler et al. 2016). In recent study sequencing of the genome of drug-resistant strain of *E. coli* showed that its genome contained non-synonymous SNPs in the genes of *frdD, ftsI, acrB, OmpD, marR, VgrG,* and *envZ (M. Li et al. 2019)*. Although the mutations of genes f*tsI, acrB* and *envZ* were placed in different positions, these results are consistent with ours.

Under antibiotic attack, bacteria can defend itself by various mechanisms. Our results show that *E.coli* drug-resistant strain may combine different strategies: target site alteration (*ftsI*), activation of efflux system (*acrB*) and reducing permeability of the outer membrane (*envZ*). Such a variety makes it difficult to overcome ampicillin resistance. A possible way is to use alternative antibiotics with different targets, such as inhibitors of protein and nucleic acids biosynthesis. The fluoroquinolones (FQ) inhibit bacterial DNA gyrase, the enzyme, which prevents double-stranded DNA from supercoiling (Kapoor, Saigal, and Elongavan 2017). Unfortunately, fluoroquinolones can be pumped out of the cell through the AcrAB-TolC system. As far as we determined a mutation that altered its component AcrB in studied *E.coli* ampicillin resistant strain, we can assume that this mutation may also increase resistance to fluoroquinolones. Another promising group of antibiotics is tetracyclines. They alter the conserved sequences of the 16S r-RNA, preventing binding of tRNA to the 30S ribosomal subunit. In this study we found only synonymous substitution in a gene, which product is associated with translation (*rsgA*). Thus we may suggest tetracyclines as an alternative way to struggle with the infection caused by studied strain of *E. coli*.

References

Adler, Marlen, Mehreen Anjum, Dan I. Andersson, and Linus Sandegren. 2016. "Combinations of Mutations in envZ, ftsI, mrdA, acrB and acrR Can Cause High-Level Carbapenem Resistance in Escherichia Coli." *The Journal of Antimicrobial Chemotherapy* 71 (5): 1188–98.

Barnes, Michael R., and Ian C. Gray. 2003. *Bioinformatics for Geneticists*. Wiley.

Bellini, D., L. Koekemoer, H. Newman, and C. G. Dowson. 2019. "Crystal Structure of TP Domain from Escherichia Coli Penicillin-Binding Protein 3 in Complex with Penicillin." https://doi.org/10.2210/pdb6i1i/pdb.

Bellini, Dom, Lizbé Koekemoer, Hector Newman, and Christopher G. Dowson. 2019. "Novel and Improved Crystal Structures of H. Influenzae, E. Coli and P. Aeruginosa Penicillin-Binding Protein 3 (PBP3) and N. Gonorrhoeae PBP2: Toward a Better Understanding of β-Lactam Target-Mediated Resistance." *Journal of Molecular Biology* 431 (18): 3501–19.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. "The Protein Data Bank." *Nucleic Acids Research* 28 (1): 235–42.

Bolduc, Benjamin. n.d. "Quality Control of Reads Using Trimmomatic (Cyverse) v1 (protocols.io.ewbbfan)." *Protocols.io*. https://doi.org/10.17504/protocols.io.ewbbfan.

Cai, Sheng Jian, and Masayori Inouye. 2002. "EnvZ-OmpR Interaction and Osmoregulation in Escherichia Coli." *The Journal of Biological Chemistry* 277 (27): 24155–61.

Choi, Umji, and Chang-Ro Lee. 2019. "Distinct Roles of Outer Membrane Porins in Antibiotic Resistance and Membrane Integrity in." *Frontiers in Microbiology* 10 (April): 953.

Donachie, W. D. 1993. "The Cell Cycle of Escherichia Coli." *Annual Review of Microbiology* 47: 199–230.

Du, Dijun, Xuan Wang-Kan, Arthur Neuberger, Hendrik W. van Veen, Klaas M. Pos, Laura J. V. Piddock, and Ben F. Luisi. 2018. "Multidrug Efflux Pumps: Structure, Function and Regulation." *Nature Reviews. Microbiology* 16 (9): 523–39.

Gerstle, Kirstin, Kristin Klätschke, Ulrich Hahn, and Nicolas Piganeau. 2012. "The Small RNA RybA Regulates Key-Genes in the Biosynthesis of Aromatic Amino Acids under Peroxide Stress in E. Coli." *RNA Biology* 9 (4): 458–68.

Harlocker, S. L., L. Bergstrom, and M. Inouye. 1995. "Tandem Binding of Six OmpR Proteins to the ompF Upstream Regulatory Sequence of Escherichia Coli." *The Journal of Biological Chemistry* 270 (45): 26849–56.

Kapoor, Garima, Saurabh Saigal, and Ashok Elongavan. 2017. "Action and Resistance Mechanisms of Antibiotics: A Guide for Clinicians." *Journal of Anaesthesiology, Clinical Pharmacology* 33 (3): 300–305.

Kishii, Ryuta, Liliana Falzon, Takeshi Yoshida, Hiroshi Kobayashi, and Masayori Inouye. 2007. "Structural and Functional Studies of the HAMP Domain of EnvZ, an Osmosensing Transmembrane Histidine Kinase in Escherichia Coli." *The Journal of Biological Chemistry* 282 (36): 26401–8.

Koboldt, Daniel C., David E. Larson, and Richard K. Wilson. 2013. "Using VarScan 2 for Germline Variant Calling and Somatic Mutation Detection." *Current Protocols in Bioinformatics*. https://doi.org/10.1002/0471250953.bi1504s44.

Kobylka, Jessica, Miriam S. Kuth, Reinke T. Müller, Eric R. Geertsma, and Klaas M. Pos. 2020. "AcrB: A Mean, Keen, Drug Efflux Machine." *Annals of the New York Academy of Sciences* 1459 (1): 38–68.

Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (14): 1754–60.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.

Li, Mengchen, Qiaoli Liu, Yanli Teng, Liuyang Ou, Yuanlin Xi, Shuaiyin Chen, and Guangcai Duan. 2019. "The Resistance Mechanism of Escherichia Coli Induced by Ampicillin in Laboratory." *Infection and Drug Resistance*. https://doi.org/10.2147/idr.s221212.

Munita, Jose M., and Cesar A. Arias. 2016. “Mechanisms of Antibiotic Resistance.” *Microbiology Spectrum* 4 (2). https://doi.org/10.1128/microbiolspec.VMBF-0016-2015.
Okajima, T., Y. Eguchi, N. Tochio, Y. Inukai, R. Shimizu, S. Ueda, S. Shinya, et al. 2016. “DHp Domain Structure of EnvZ from Escherichia Coli.” https://doi.org/10.2210/pdb5b1n/pdb.
Reygaert, Wanda C. 2018. “An Overview of the Antimicrobial Resistance Mechanisms of Bacteria.” *AIMS Microbiology* 4 (3): 482–501.
Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. 2011. “Integrative Genomics Viewer.” *Nature Biotechnology* 29 (1): 24–26.
Sjuts, H., A. R. Ornik, and K. M. Pos. 2016. “Apo Structure of Bacterial Efflux Pump.” https://doi.org/10.2210/pdb5en5/pdb.
Söderström, Bill, and Daniel O. Daley. 2017. “The Bacterial Divisome: More than a Ring?” *Current Genetics* 63 (2): 161–64.
Tormo, A., J. A. Ayala, M. A. de Pedro, M. Aldea, and M. Vicente. 1986. “Interaction of FtsA and PBP3 Proteins in the Escherichia Coli Septum.” *Journal of Bacteriology* 166 (3): 985–92.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]