



# Genome-Wide Transcriptional Response of *Mycobacterium smegmatis* MC<sup>2</sup>155 to G-Quadruplex Ligands BRACO-19 and TMPyP4

Egor Shitikov\*, Dmitry Bespiatykh, Maja Malakhova, Julia Bespyatykh, Ivan Bodoev, Tatiana Vedekhina, Marina Zaychikova, Vladimir Veselovsky, Ksenia Klimina, Elena Iliina and Anna Varizhuk

Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

## OPEN ACCESS

### Edited by:

Daisuke Miyoshi,  
Konan University, Japan

### Reviewed by:

Guanhui Wu,  
Shanghai Roche Pharmaceutical  
Co., Ltd., China  
Nihar Ranjan,  
National Institute of Pharmaceutical  
Education and Research, India

### \*Correspondence:

Egor Shitikov  
egorshtkv@gmail.com

### Specialty section:

This article was submitted to  
Antimicrobials, Resistance and  
Chemotherapy,  
a section of the journal  
Frontiers in Microbiology

Received: 17 November 2021

Accepted: 15 February 2022

Published: 04 March 2022

### Citation:

Shitikov E, Bespiatykh D, Malakhova M, Bespyatykh J, Bodoev I, Vedekhina T, Zaychikova M, Veselovsky V, Klimina K, Iliina E and Varizhuk A (2022) Genome-Wide Transcriptional Response of *Mycobacterium smegmatis* MC<sup>2</sup>155 to G-Quadruplex Ligands BRACO-19 and TMPyP4. *Front. Microbiol.* 13:817024. doi: 10.3389/fmicb.2022.817024

G-quadruplexes (G4s) are non-canonical DNA structures that could be considered as potential therapeutic targets for antimicrobial compounds, also known as G4-stabilizing ligands. While some of these ligands are shown *in vitro* to have a stabilizing effect, the precise mechanism of antibacterial action has not been fully investigated. Here, we employed genome-wide RNA-sequencing to analyze the response of *Mycobacterium smegmatis* to inhibitory concentrations of BRACO-19 and TMPyP4 G4 ligands. The expression profile changed (FDR < 0.05, log<sub>2</sub>FC > |1|) for 822 (515↑; 307↓) genes in *M. smegmatis* in response to BRACO-19 and for 680 (339↑; 341↓) genes in response to TMPyP4. However, the analysis revealed no significant ligand-induced changes in the expression levels of G4-harboring genes, genes under G4-harboring promoters, or intergenic regions located on mRNA-like or template strands. Meanwhile, for the BRACO-19 ligand, we found significant changes in the replication and repair system genes, as well as in iron metabolism genes which is, undoubtedly, evidence of the induced stress. For the TMPyP4 compound, substantial changes were found in transcription factors and the arginine biosynthesis system, which may indicate multiple biological targets for this compound.

**Keywords:** mycobacteria, transcriptome, RNA, G4, ligands, BRACO-19, TMPyP4

## INTRODUCTION

The emergence and rapid dissemination of drug-resistant bacterial pathogens are becoming a global health challenge. In 2019, an estimated 10 million people worldwide developed tuberculosis (TB), caused by *Mycobacterium tuberculosis* complex bacteria. In total, nearly half a million people have developed rifampicin-resistant TB (RR-TB), 78% of whom had multidrug-resistant TB (MDR-TB; resistance to both first-line drugs rifampin and isoniazid). The global treatment success rate for MDR/RR-TB was estimated at 57%, compared to an overall success rate of 85% for all cases. Moreover, the treatment of people with resistant TB forms takes longer and requires more expensive and toxic drugs (Seung et al., 2015; Global Tuberculosis Report 2020, 2020). All of these factors encourage the development of new anti-TB agents with new mechanisms of action.

One of the recently emerged trends in the design of antimicrobial drugs suggests targeting quadruplex structures (G4s) in guanine-rich regions of the bacterial genome (Yadav et al., 2021). Such structures comprise planar arrangements of guanine tetrads held together by

Hoogsteen base pairing and  $\pi$ - $\pi$  interactions (Spiegel et al., 2020). G4s have been found in eukaryotes, viruses, and multiple bacterial species (Holder and Hartig, 2014; Ding et al., 2018; Bartas et al., 2019; Marsico et al., 2019), including *M. tuberculosis* (Perrone et al., 2017). Their conservation and non-random genomic distribution (Bartas et al., 2019) are likely indicative of important biological functions. In bacteria, G4s are often found in virulence-related genes (Harris and Merrick, 2015) or their promoter regions and provide a control mechanism involved in the regulation of biological pathways such as transcription and translation (Holder and Hartig, 2014; Wu et al., 2015; Shao et al., 2020). Due to the activity of specific helicases, genomic G4s do not normally persist in cells (Sauer and Paeschke, 2017; Saha et al., 2019). Some G4-resolving helicases (e.g., the PIF1 family) are conserved from bacteria to humans, highlighting the importance of efficient G4 processing (Byrd and Raney, 2017). Unresolved G4s may induce DNA breaks and damage response, recruit transcription factors, or mechanically hamper polymerase passage, affecting transcription and replication (Saranathan and Vivekanandan, 2019).

All the foregoing advocates the use of G4-stabilizing ligands to repress or fine-tune gene expression. An extensive effort is being devoted to designing new small-molecule G4 binders and elucidating the basis for antimicrobial properties of the known ones (Yadav et al., 2021). However, despite recent advances in controlling bacterial growth with exogenous ligands, the underlying mechanisms are poorly understood. Recently, it has been reported that two known G4 ligands, BRACO-19 and c-exNNDI2, exhibit anti-TB activity (Perrone et al., 2017); besides, their possible genomic targets have been identified using general sequence rules and G4 prediction algorithms of relatively low stringency (Beaudoin et al., 2014). Subsequently, the stabilizing activity of TMPyP4 ligand for virulence-related G4 targets has been reported (Mishra et al., 2019). These reports encourage the future development of G4-affecting agents. However, additional studies are needed to verify the role of G4s and determine the affected metabolic pathways.

In the present study, we performed the transcriptomic analysis of the *Mycobacterium smegmatis* MC<sup>2</sup>155 strain in the presence of an inhibitory concentration of two G4 ligands, namely BRACO-19 and TMPyP4. Given the known ability of G4 ligands to stabilize G4s, we have identified putative G4 quadruplex sequences in *M. smegmatis* genome and analyzed G4-associated genes during the transcriptomic response to G4 ligands. In addition, we examined genes involved in replication and reparation, with a particular focus on helicases that may be involved in the unwinding of G4-quadruplexes.

## MATERIALS AND METHODS

### Bacterial Strain, Growth Conditions, and Inhibition Assay

*Mycobacterium smegmatis* MC<sup>2</sup>155 strain was used in this study. Middlebrook 7H11 Agar (HiMedia, India) and Middlebrook 7H9 broth (HiMedia) both supplemented with 0.5% glycerol and 10% Middlebrook OADC Growth Supplement (HiMedia) were used as solid and liquid media, respectively. A frozen

stock of *M. smegmatis* strain was cultured on a solid media for 1 day to obtain a sufficient number of cells for inoculation of broth culture.

To determine MICs for BRACO-19 (PubChem CID no. 9808666) and TMPyP4 (PubChem CID no. 135398505; Sigma-Aldrich, United States), bacterial cells were cultured overnight in Middlebrook 7H9 broth, with the subsequent dilution in the fresh 7H9 medium (1:200). The diluted culture was dispensed (196  $\mu$ l) to the wells of the CELLSTAR 96 Well Cell Culture Plates (Greiner Bio-One GmbH, Germany) and 4  $\mu$ l of serial 2-fold dilutions of the tested compounds in DMSO were added to the wells to final concentrations of 2.5–80  $\mu$ M BRACO-19 and 0.25–16  $\mu$ M for TMPyP4. The plates were incubated for 64 h at 37°C in a humid atmosphere with 5% CO<sub>2</sub> and shaken at 250 rpm. The MICs were determined by an optical density measurement at 570 nm on xMark Microplate Absorbance Spectrophotometer (Biorad, United States) and defined as the lowest concentration of the compound at which no growth was observed. Negative control samples were treated with the same volume of DMSO; the experiments were performed in three biological replicates.

For the transcriptomic analysis, *M. smegmatis* cells were grown up to the mid-exponential phase (OD<sub>570</sub> = 0.47) and transferred to the 5 ml tubes (NUOVA APTACA, Italy). Ligands were added to the tubes to a final concentration corresponding to 1  $\times$  MIC (10  $\mu$ M for BRACO-19 and 4  $\mu$ M for TMPyP4). The same volume of DMSO was added to the control samples (1% v/v). Bacterial cells were incubated for 4 h (cell division time; Logsdon and Aldridge, 2018) at 37°C in a humid atmosphere with 5% CO<sub>2</sub> and shaken at 250 rpm. The experiments were carried in three biological replicates.

### Total RNA Extraction and RNA-seq

Bacterial cells were harvested by centrifugation (8,000  $\times$ g, 10 min, 4°C) and subsequently washed twice with a phosphate-buffered saline (10 ml). A double volume of RNAProtect Bacteria Reagent (Qiagen, United States) was added to the pellet. The mixture was incubated at room temperature for 5 min and then centrifuged (8,000  $\times$ g, 10 min, 4°C). Samples were homogenized for 30 s using an automatic bead homogenizer (MagNalyzer, Roche). RNA was extracted using the MagMAX mirVana Total RNA Isolation Kit (Thermo Fisher Scientific, Lithuania) on the KingFisher Flex Purification System (Thermo Fisher Scientific, United States) according to the manufacturer's instructions. RNA was treated with DNase using Turbo DNA-Free Kit (Thermo Fisher Scientific) in the 50  $\mu$ l volume, and further purified using Agencourt RNAClean XP (Beckman Coulter, United States) according to the manufacturer's instructions. The total amount of RNA was measured with the Quant-iT Ribogreen RNA assay kit (ThermoFisher Scientific), and the quality of extracted RNA was checked by Agilent Bioanalyzer on Agilent RNA 6000 Pico Chips (Agilent Technologies, United States).

Total RNA (300 ng) was used for library preparation. The Ribo-Zero Plus rRNA Depletion Kit (Illumina, United States) was used to remove rRNA, and the NEBNext Ultra II Directional

RNA Library Prep Kit (NEB) was used for library preparation. RNA cleanup was performed with the Agencourt RNA Clean XP kit (Beckman Coulter). Final cleanup was done using the Agencourt AMPure XP system (Beckman Coulter) after which the libraries' size distribution and quality were assessed using a high sensitivity DNA chip (Agilent Technologies). Libraries were subsequently quantified with Quant-iT High Sensitivity dsDNA Assay Kit (Thermo Fisher Scientific). The libraries were sequenced on an Illumina HiSeq 2500 (50 bp single-end reads), with 12 pM loading concentration, and the PhiX control library (Illumina) spiked in at 5%. The RNA-seq data generated in this study have been deposited to the NCBI Sequence Read Archive under accession number PRJNA765512.<sup>1</sup>

## Inferring Putative G4 Quadruplex Sequences

To infer PQSs in *M. smegmatis* MC<sup>2</sup>155 (GenBank accession no. CP000480.1) genome G4-iM Grinder (v1.6.1; Belmonte-Reche and Morales, 2020) package for R (v4.0.5; R Core Team, 2021) was used. The following parameters were used: Complementary=TRUE, BulgeSize=0, RunComposition="G," MaxRunSize=4, MinRunSize=3, MaxNRuns=0, MinNRuns=4, MaxIL=0, MaxPQSSize=50, MinPQSSize=15, MaxLoopSize=15, MinLoopSize=0. Size independent quadruplexes without overlapping (method 3a) with a score  $\geq 40$  were used for further analysis. BEDtools (v2.30.0; Quinlan and Hall, 2010) and custom Python scripts were used to assign PQSs to *M. smegmatis* MC<sup>2</sup>155 genes and intergenic regions (**Supplementary Table S1**). When PQS was located in both gene and intergenic region, it was assigned to both categories. Intergenic PQSs were assigned without regard to strand orientation. For the analysis, only PQSs located upstream (downstream for the template strand) of genes were used. To assign PQSs to promoter regions, a list of previously published *M. smegmatis* MC<sup>2</sup>155 transcription start sites (TSSs) was used (Li et al., 2017). Promoters were considered to be located 50 bp upstream or downstream of TSSs, depending on the strandedness of the gene.

## Circular Dichroism Spectroscopy and Melting Assay

Oligodeoxyribonucleotides (ODNs) were obtained from LITEKH, Russia. To verify G4 folding, circular dichroism (CD) spectra of the ODN samples in 10 mM potassium-phosphate buffer, pH 7, were registered at 15°C using Chirascan spectrophotometer (Applied Biophysics, United Kingdom). The ODN samples were annealed rapidly (heated to 90°C and then snap-cooled on ice) to facilitate intramolecular folding prior to all measurements. Moderate salt concentration was used to avoid extremely high thermal stability of the high-scoring G4s, which would hamper accurate assessment of ligands' effects. Thermal stability ( $T_{1/2}$ ) of each G4 was calculated as an average value of the melting ( $T_m$ ) and annealing ( $T_a$ ) temperatures estimated based on the annealing and melting curves, respectively. The melting and annealing curves were registered using Chirascan

spectrophotometer by CD monitoring at G4-specific maxima (260 nm for parallel-stranded/mixed and 295 nm for antiparallel-stranded/mixed G4s) with a temperature ramp rate of 1°C/min in the presence and in the absence of the ligands. G4 and ligand concentrations were 2.5  $\mu$ M.

## Differential Gene Expression Analysis

The sequenced reads were mapped to the reference *M. smegmatis* MC<sup>2</sup>155 genome (GenBank accession no. CP000480.1) using HISAT2 (v2.2.1; Kim et al., 2015). SAMtools (v1.11; Li et al., 2009) software was used to sort and convert SAM files to BAM, and their subsequent indexing. Mapping quality and coverage along genes were assessed with QualiMap (v2.2.2; Okonechnikov et al., 2016), individual reports were merged with MultiQC (v1.9; Ewels et al., 2016). Mapped reads were assigned to genes with featureCounts (v2.0.1; Liao et al., 2014). Differential gene expression analysis was performed using the edgeR (v3.30.3; Robinson et al., 2010) package for R. Genes with a false discovery rate (FDR) cutoff of 0.05 and with a fold change ( $\log_2$ FC) threshold of  $|1|$  (i.e.,  $\geq 2$ -fold change) were considered to be differentially expressed. For intersample comparison, counts were normalized using gene length corrected trimmed mean of M-values (GeTMM; Smid et al., 2018). Further functional enrichment analysis of the GO terms and KEGG pathways for differentially expressed genes was performed using clusterProfiler (v3.18.1; Yu et al., 2012) package, categories were considered enriched with  $p_{adj.} \leq 0.05$ . In addition, differentially expressed genes were classified into functional categories based on the PATRIC database (Wattam et al., 2014). Clusters of orthologous genes (COG) were annotated by mapping *M. smegmatis* MC<sup>2</sup>155 proteome to NCBI COG database (v2020)<sup>2</sup> using DIAMOND (v2.0.9; Buchfink et al., 2014), and with eggNOG-mapper (v2; Huerta-Cepas et al., 2019). Plots were generated within R using ggplot2 (v3.3.2; Wickham, 2009), ggforce (v0.3.3; Pedersen, 2021), ggvenn (v0.1.9; Yan, 2021), ggsignif (v0.6.1; Ahlmann-Eltze and Patil, 2021), lemon (v0.4.5; Edwards, 2020), and cowplot (v1.1.0; Wilke, 2020) packages.

## Statistical Analysis

All statistical analyses were performed using RStudio (v1.4.1106). Statistical evaluation was performed by the two-sided Wilcoxon rank-sum test and the Student's *t*-test; value of  $p < 0.05$  was considered to be statistically significant.

## RESULTS

### Differential Gene Expression of *Mycobacterium smegmatis* MC<sup>2</sup>155 in the Presence of BRACO-19 and TMPyP4 G4 Ligands

To understand how G4-stabilizing compounds affect gene expression levels, the transcriptomic profiles of *M. smegmatis* MC<sup>2</sup>155 strain exposed to BRACO-19 and TMPyP4 for 4 h

<sup>1</sup><https://www.ncbi.nlm.nih.gov/sra>

<sup>2</sup><https://ftp.ncbi.nlm.nih.gov/pub/COG/COG2020/data/>

at  $1\times$ MIC were examined (**Supplementary Table S1**). According to dilution-based antimicrobial susceptibility testing, BRACO-19 and TMPyP4 had *in vitro* minimum inhibitory concentration (MIC) values of 10 and  $4\mu\text{M}$ , respectively (**Supplementary Figure S1**).

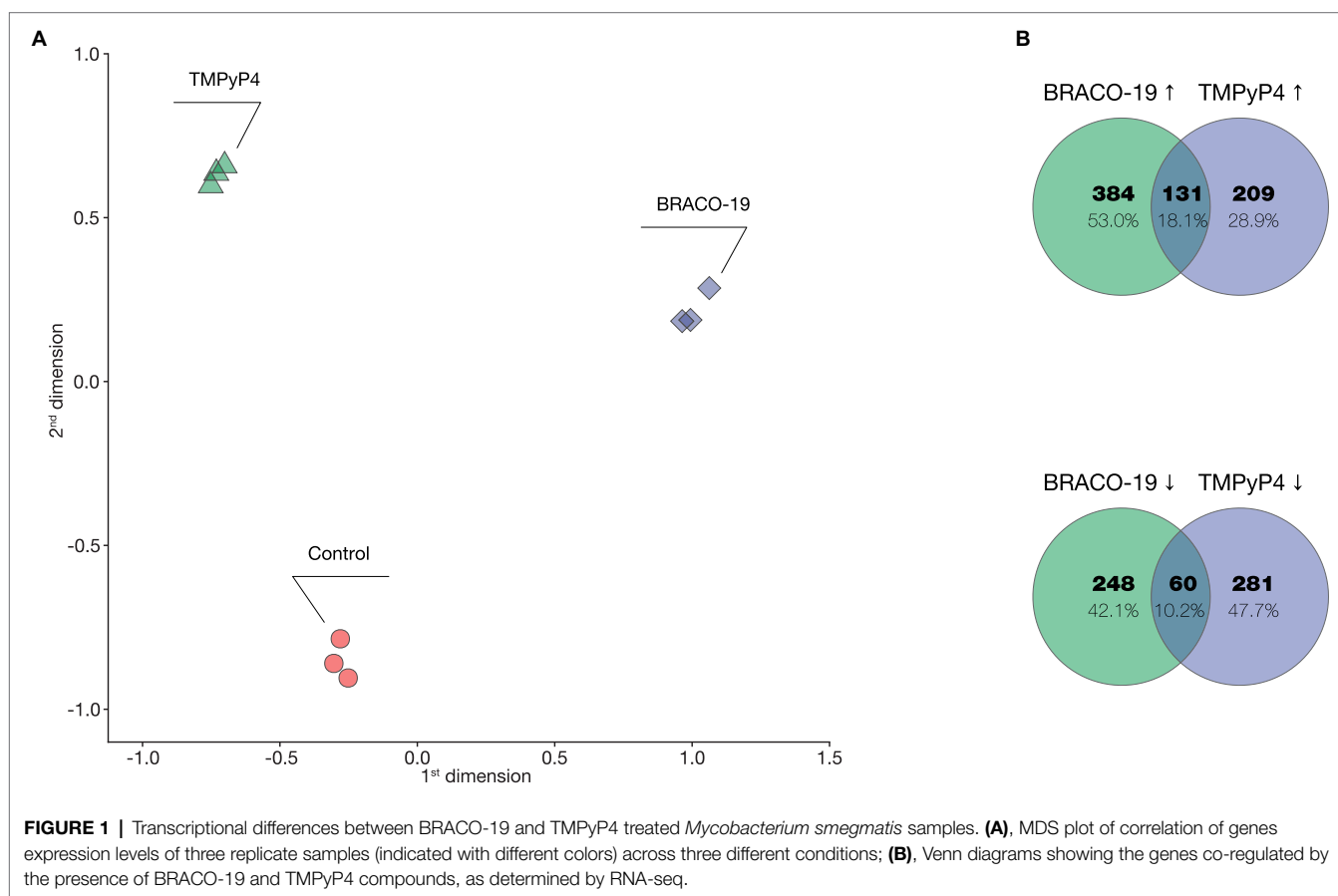
Multidimensional scaling (MDS) of normalized RNA-seq data (three biological replicates and three different conditions) showed a clear separation of samples by condition (Control, BRACO-19, and TMPyP4; **Figure 1A**). Expression significantly changed ( $\text{FDR} < 0.05$ ,  $\log_2\text{FC} > |1|$ ) for 822 (515 $\uparrow$ ; 307 $\downarrow$ ) genes in *M. smegmatis* in response to BRACO-19 and for 680 (339 $\uparrow$ ; 341 $\downarrow$ ) genes in response to TMPyP4. Among the 191 differentially expressed genes (DEGs) in response to both BRACO-19 and TMPyP4, 131 genes were upregulated, and 60 genes were downregulated (**Figure 1B**).

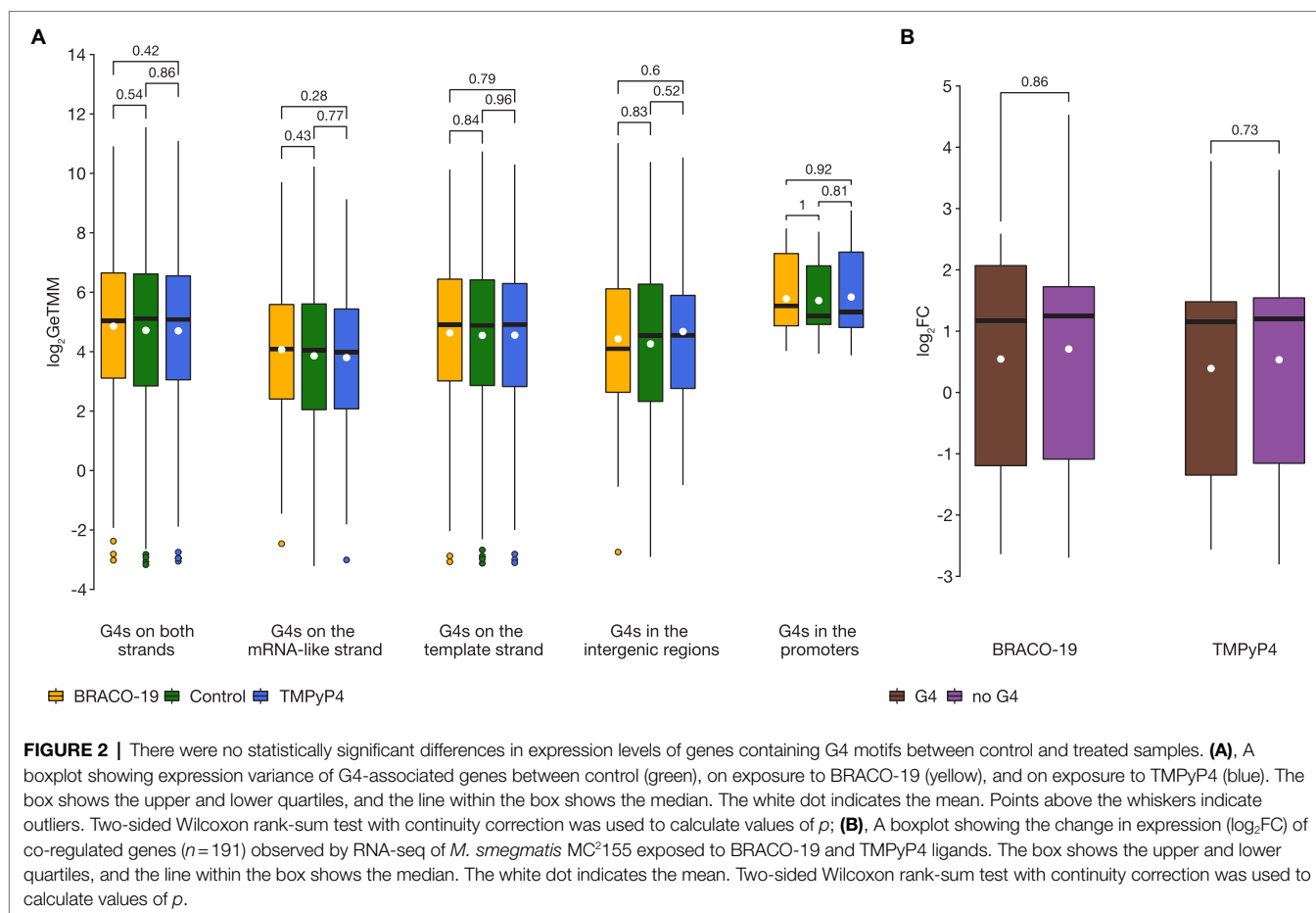
## The Effect of Ligands on the Expression of Genes Linked to G4s

Potential G-quadruplex forming sequences (PQSs) in the *M. smegmatis* MC<sup>2</sup>155 genome were identified using G4-iM Grinder. In total, 53 (14/53 were valid for two genes) PQSs were found in intergenic regions, 17 PQSs were located in promoter regions, 267 PQSs were situated in the mRNA-like strand of the 255 genes, and 398 PQSs were located in the template strand of the 360 genes (**Supplementary Table S2**).

The results of RNA-seq analysis showed that there are no significant differences in expression levels of the bulk of the genes associated with G4s (i.e., those harboring PQSs; **Supplementary Figure S2**). Further, analyses for PQSs located on mRNA-like strand, template strand, intergenic regions, and in promoters between the control sample and samples exposed to G4 ligands, showed no statistically significant differences (two-sided Wilcoxon test  $p > 0.05$ ; **Figure 2A**). Besides, there was no difference in genes expression ( $\log_2\text{FC}$ ) relative to G4 strandedness (mRNA-like or template) in intergenic regions, neither in BRACO-19 (Student's *t*-test  $p = 0.519$ ) nor in TMPyP4 (Student's *t*-test  $p = 0.645$ ) treated samples. Furthermore, no significant differences were observed in expression levels of genes co-regulated by both ligands, between genes associated with G4s ( $n = 14$ ), and without G4s ( $n = 177$ ; two-sided Wilcoxon test  $p > 0.05$ ; **Figure 2B**).

To verify our RNA-seq results and G4-stabilizing effect of the ligands, we investigated 10 high-scoring PQS fragments of randomly selected genes that showed: enhanced expression levels in response to ligand treatment ( $n = 4$ ), no significant changes ( $n = 2$ ), or decreased expression levels ( $n = 4$ ) (**Supplementary Table S3**). The effect of G4-ligands on PQSs stability was determined using thermal denaturation monitored by circular dichroism spectroscopy (**Supplementary Figure S3**). In the presence of BRACO-19 and TMPyP4, G4 stabilization





was observed, nonetheless these results did not correlate with the changes in corresponding genes expression levels.

## Changes in the Expression of DNA Replication and Repair Systems Mediated by G4 Ligands

Additionally, the genes of the replication and repair system were considered (**Supplementary Table S4**), since G4 stabilizing ligands can damage DNA and cause genome instability (Saranathan and Vivekanandan, 2019). In total, 13 and four genes exhibited more than 2-fold expression level changes under the action of BRACO-19 and TMPyP4, respectively (**Table 1**).

Genes encoding DNA polymerases DinB and DnaE2, assessor proteins ImuA and ImuB, DNA gyrases GyrA and GyrB, proteins of alkylation repair system AlkA and Ogt, as well as Hely and RecQ helicases were overexpressed on exposure to BRACO-19. Most of these genes are required for translesion DNA synthesis. Polymerase DnaE2 (C-family probable translesion polymerase) is reportedly involved in adaptive mutagenesis in response to DNA damage (Boshoff et al., 2003). Accessory proteins ImuA and ImuB facilitate DnaE2-dependent translesion synthesis<sup>43</sup>, thus contributing to bacterial SOS response. Although the role of DinB1, DinB2 [duplicated in *M. smegmatis*; MSMEG\_2294/MSMEG\_1014

(both genes had zero total read counts)], and DinB3 (specific for *M. smegmatis*) Y-family polymerases is still not fully understood, it has been reported that DinB1 and DinB3 are typical DNA-dependent DNA polymerases, while DinB2 misincorporates deoxyribonucleotides and ribonucleotides during templated synthesis and lesion bypass (Ordóñez and Shuman, 2014). Notably, according to previous reports, neither *dinB1*- nor *dinB2*-encoded PolIV homolog is up-regulated during the mycobacterial damage response (Rand et al., 2003). The up-regulated DNA gyrases (type II DNA topoisomerases) contribute to DNA synthesis by catalyzing the introduction of negative supercoiled turns in an ATP-dependent manner, while double-stranded DNA is unwound by lengthening the RNA polymerase or helicase in front of the progressive replication fork (Aubry et al., 2006). Among genes responsible for alkylation repair, *ogt* and *alkA* had the highest expression levels. Methyltransferase Ogt and AlkA DNA-glycosylase are the major proteins of repair systems; they counteract DNA alkylation damage, thus preventing cell death (Yang et al., 2011).

In TMPyP4-treated *M. smegmatis*, *ssbB* (MSMEG\_4701), which encodes the RecA accessory factor required for recombination repair during stress (Singh, 2018), was the only up-regulated repair system gene ( $\log_2FC=1.24$ ), and this gene was also

**TABLE 1** | Differentially expressed genes of the DNA replication and repair systems in *M. smegmatis* exposed to BRACO-19 and TMPyP4.

Locus tag	<i>M. tuberculosis</i> H37Rv homologue	Gene name	Gene product	BRACO-19		TMPyP4	
				log <sub>2</sub> FC	FDR	log <sub>2</sub> FC	FDR
MSMEG_4925	Rv1317c	<i>alkA</i>	Transcriptional regulator, Ada family protein/DNA-3-methyladenine glycosylase II	<b>3.15</b>	<b>9.23E-13</b>	0.62	<b>0.000607</b>
MSMEG_4928	Rv1316c	<i>ogt</i>	Possible 3-methyladenine DNA glycosylase Mpg	<b>3.31</b>	<b>9.49E-13</b>	0.44	<b>0.013258</b>
MSMEG_5422	Rv1021	<i>mazG</i>	Nucleoside triphosphate pyrophosphohydrolase	0.66	<b>6.09E-06</b>	<b>-1.13</b>	<b>2.6E-08</b>
MSMEG_1633	Rv3370c	<i>dnaE2</i>	Error-prone DNA polymerase	<b>1.34</b>	<b>1.71E-08</b>	0.90	<b>3.88E-06</b>
MSMEG_3172	Rv1537	<i>dinB1</i>	DNA polymerase IV	<b>1.20</b>	<b>1.97E-07</b>	0.18	0.259968
MSMEG_6443		<i>dinB3</i>	DNA polymerase IV	<b>1.52</b>	<b>1.96E-07</b>	0.33	0.104471
MSMEG_2442	Rv2902c	<i>rnhB</i>	Ribonuclease HII	<b>1.098</b>	<b>6.33E-08</b>	0.45	<b>0.001482</b>
MSMEG_6896	Rv0054	<i>ssbA</i>	Single-stranded DNA-binding protein	<b>-1.43</b>	<b>3.92E-09</b>	-0.39	<b>0.003882</b>
MSMEG_4701	Rv2478c	<i>ssbB</i>	Hypothetical protein	0.93	<b>0.00116</b>	<b>1.24</b>	<b>8.09E-05</b>
MSMEG_1327	Rv0630c	<i>recB</i>	Exodeoxyribonuclease V subunit beta	0.31	0.020787	<b>-1.16</b>	<b>2.72E-07</b>
MSMEG_5397		<i>recQ</i>	ATP-dependent DNA helicase RecQ	<b>2.15</b>	<b>0.00002</b>	<b>1.54</b>	<b>0.0044</b>
MSMEG_0006	Rv0006	<i>gyrA</i>	DNA gyrase subunit A	<b>1.21</b>	<b>1.85E-09</b>	-0.87	<b>1.37E-07</b>
MSMEG_0005	Rv0005	<i>gyrB</i>	DNA gyrase subunit B	<b>1.20</b>	<b>8.75E-10</b>	-0.76	<b>3.44E-07</b>
MSMEG_1620	Rv3395c	<i>imuA</i>	Hypothetical protein	<b>1.52</b>	<b>0.0006</b>	-0.12	0.825962
MSMEG_1622	Rv3394c	<i>imuB</i>	DNA repair polymerase	<b>1.28</b>	<b>0.000145</b>	0.96	<b>0.002957</b>
MSMEG_5397			ATP-dependent DNA helicase RecQ	<b>2.15</b>	<b>0.000105</b>	<b>1.54</b>	<b>0.004401</b>
MSMEG_3885	Rv2092c	<i>helY</i>	DEAD/DEAH box helicase	<b>1.77</b>	<b>3.26E-11</b>	-0.17	0.113436

Values in bold meet the corresponding selection criteria ( $FDR < 0.05$ ,  $\log_2FC > |1|$ ).

upregulated in BRACO-19 case ( $\log_2FC = 0.93$ ). On the contrary, *ssbA* (MSMEG\_6896), which encodes another RecA cofactor (Singh, 2018), was downregulated for both ligands ( $\log_2FC = -1.43$  for BRACO-19, and  $\log_2FC = -0.3$  for TMPyP4). Such a multi-directional change in the level of expression of the aforementioned genes has been described previously for ultraviolet and hypoxia-induced stress in *M. smegmatis* (Singh et al., 2018).

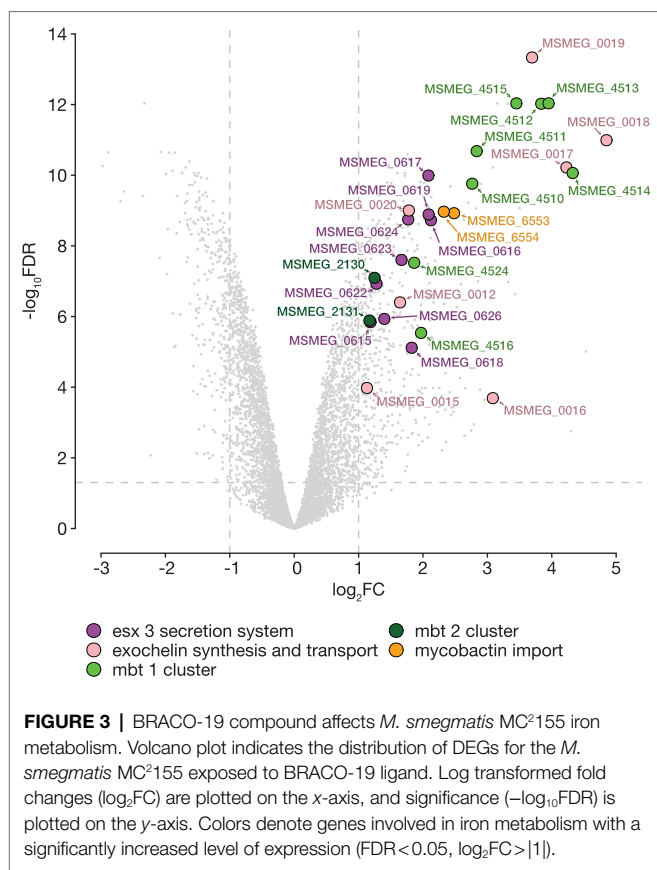
Additional analysis of previously described helicases (DinG, UvrD1, and UvrD2; Thakur et al., 2014; Saha et al., 2019), responsible for the unwinding of the G4-quadruplexes, showed slightly increased expression levels, although only significant ( $FDR < 0.05$ ) for BRACO-19 (Supplementary Table S1). Besides, DNA helicase RecQ was up-regulated under BRACO-19 and TMPyP4. RNA-activated ATPase/dATPase and 3'-5' helicase *helY* gene was upregulated on exposure to BRACO-19. In the TMPyP4 case, the ATP-dependent RNA helicase MSMEG\_1540 was downregulated ( $\log_2FC = -2.08$ ).

## Functions and Pathways of *Mycobacterium smegmatis* MC<sup>2</sup>155 Genes Responding to Treatment With G4 Ligands

Since there was no strong evidence that BRACO-19 and TMPyP4 ligands affect the expression of G4-carrying genes in *M. smegmatis*, further analysis of transcriptomic data was performed. Differentially expressed genes were subjected to functional classification and pathway enrichment analysis based on the

Kyoto Encyclopaedia of Genes and Genomes (KEGG) and gene ontology (GO) databases (Supplementary Table S5).

For BRACO-19, only one KEGG pathway—msm01053 “Biosynthesis of siderophore group nonribosomal peptides” ( $p_{adj} = 1.28e-05$ ), was enriched for upregulated genes. This pathway includes *mbt* cluster genes (Figure 3; Supplementary Table S6), which are responsible for the synthesis of the most important iron-chelating compounds—siderophores (mycobactin and carboxymycobactin in mycobacteria; Sritharan, 2016). Analysis of the exochelin locus, which is a specific siderophore for non-virulent mycobacteria, revealed similar changes. Also, overexpression of genes involved in siderophore transport and acquisition was observed. Among these genes were the iron-dependent transporters *irtAB* and ESX-3 operon. The ATP binding cassette transporter IrtAB imports iron-bound siderophores across the inner membrane, while the type VII secretion system Esx-3 plays a crucial role in iron acquisition through secretion of a pair of PE-PPE family proteins. In turn, iron storage genes *bfrA* and *bfrB*, encoding a heme-containing bacterioferritin (BfrA) and a ferritin-like protein (BfrB), were slightly downregulated. It is worth noting that almost all the aforementioned genes are controlled by the IdeR (iron-dependent regulator; MSMEG\_2750; Sritharan, 2016). The *ideR* gene was slightly upregulated ( $\log_2FC = 0.65$ ) compared to the control, notably, there is a high-scored ( $x = 57$ ) quadruplex with four G-tetrads (GGGG ATGGGGTTGCCGAACGGGGAGGTGGTGGGG) located on the template strand of the gene.



In the case of TMPyP4-exposed bacteria, the following biological process GO terms were enriched: GO:0006351 “transcription, DNA-templated” ( $p_{adj.} = 0.002$ ) for upregulated genes, and GO:0006526 “arginine biosynthetic process” ( $p_{adj.} = 1.21e-05$ ) for downregulated genes. The KEGG pathways associated with upregulated genes in the TMPyP4-treated sample, were msm00920 “Sulfur metabolism” ( $p_{adj.} = 0.02$ ), and msm02010 “ABC transporters” ( $p_{adj.} = 0.02$ ), and one pathway associated with down-regulated genes was msm00220 “Arginine biosynthesis” ( $p_{adj.} = 0.03$ ). Primarily, GO:0006351 term included potential transcription factor genes with increased expression that may reflect significant changes in metabolism in response to ligand action. KEGG pathway msm02010 included almost all the genes in msm00920 pathway (7/8) and consisted of the genes responsible for sugar transport, phosphates, and sulfur-containing amino acids. At the same time, genes involved in the synthesis of the amino acid arginine (MSMEG\_3046, MSMEG\_3047, MSMEG\_3769-MSMEG\_3776) were downregulated in TMPyP4-treated *M. smegmatis*. Of these, carbamoyl phosphate synthetases units A (MSMEG\_3046) and B (MSMEG\_3047) are involved in pyrimidine and arginine biosynthesis, catalyzing the conversion of glutamine to carbamoyl phosphate, while products of the *argDBJC* (MSMEG\_3773-MSMEG\_3776) genes are part of a five-enzyme pathway that uses glutamate to produce ornithine. Both products are used in the urea cycle to produce arginine under the action of ArgG (MSMEG\_3770) in conjunction with ArgF (MSMEG\_3772) and ArgH (MSMEG\_3769) enzymes.

## DISCUSSION

Here, we employed RNA-seq analysis to demonstrate the impact of G4-ligands, BRACO-19 and TMPyP4, on *M. smegmatis* transcription and elucidate their mechanism of action. *Mycobacterium smegmatis* was used in this study because it is considered as a valuable model organism for mycobacterium study due to the close association with *M. tuberculosis* regarding biochemical properties and genetic information (Joseph Antony Sundarsingh et al., 2020), and, importantly, has a similar GC composition. Both ligands used are best known as stabilizers of human telomeric G4s, which implies cancer-specific antiproliferative activity (Rha et al., 2000; Incles et al., 2004). However, they also exhibit broad-spectrum G4 stabilizing activity and, along with other pan-quadruplex ligands, are being tested for their therapeutic potential against various bacterial species, including *M. tuberculosis*. In the case of *M. smegmatis* (this study) and *M. tuberculosis* (Beaudoin et al., 2014; Mishra et al., 2019), both compounds inhibited the growth at submicromolar concentrations, but TMPyP4 was slightly more active for *M. smegmatis*.

Both BRACO-19 and TMPyP4 have been reported to stabilize synthetic G4s from *M. tuberculosis*. TMPyP4 has also been shown to reduce transcription of PQS-containing genes (Beaudoin et al., 2014; Mishra et al., 2019). Based on our findings, there was no strong evidence of significant up- or down-regulation of genes associated with PQSs caused by G4-ligands. Notably, we analyzed ligands’ effects both separately and in conjunction to account for specific DEGs and the co-regulated ones, respectively. However, even for co-regulated genes, we were unable to find concrete evidence of ligand-mediated G4-stabilization. The examples of DEGs that are co-regulated by both compounds and harbor/neighbor PQSs include MSMEG\_4727↓, MSMEG\_1680↓, MSMEG\_4689↓, MSMEG\_3948↓, MSMEG\_2242↑, MSMEG\_3121↑, and MSMEG\_2148↑ (contain PQSs on the template strand of the gene); MSMEG\_1488↑, MSMEG\_0868↑, MSMEG\_6223↑, MSMEG\_5397↑, and MSMEG\_4202↑ (contain PQSs on the mRNA-like strand); MSMEG\_6567↑ (contains one PQS on each strand); MSMEG\_3917↑ (contains a PQS on the template strand in the intergenic region), and MSMEG\_4643↓ (contains a PQS on the mRNA-like strand in the intergenic region). Moreover, additional CD assays showed stabilization of PQSs by BRACO-19 and TMPyP4, which did not correlate with the results of transcriptome analysis.

We hypothesize that the ligands act primarily during replication, when the majority of G4s are formed (as has been proven for eukaryotic cells; Rodriguez et al., 2012), and transcription changes stem partly from the disruption of replication (Lerner and Sale, 2019). It is possible that the observed transcription changes illustrate compensatory response rather than primary ligands’ effects. G4 stabilization upon replication may cause polymerase stalling and SOS-response (Maslowska et al., 2019). Enhanced recruitment of the translesion synthesis machinery or timely ligand dislodging and G4 unwinding by helicases would then be required for the tolerance of ligand-induced toxicity. In the present study, under the action of BRACO-19, we identified modest upregulation of DnaE2, DinB1, and DinB3 translesion polymerases; however,

no significant changes in expression levels of SOS-regulation genes (*recA* and *lexA*) were registered. Meanwhile, treatment with BRACO-19 and TMPyP4 resulted in overexpression of *M. smegmatis* specific gene MSMEG\_5397 encoding RecQ helicase (Table 1). The importance of RecQ DNA helicases in G4 unwinding in bacteria has been previously demonstrated in the case of *Escherichia coli*, although RecQ protein in *M. smegmatis* is not homologous to *E. coli* (Huber et al., 2006; Singh, 2017). In this respect, it is not possible to affirm the relationship between the overexpression of the MSMEG\_5397 and the action of ligands. In addition, BRACO-19 increased the expression of the *hely* gene. In *M. smegmatis* HelyY unwinds 3'-tailed RNA duplexes and RNA-DNA hybrids, but is unable to unwind a 3'-tailed duplex in which the loading strand is DNA. It is possible that HelyY participates in the unwinding of RNA G4 quadruples; however, the eukaryotic homologs Ski2 and Mtr4 do not possess such properties (Uson et al., 2015).

BRACO-19-induced overexpression of iron metabolism-related genes may be an additional indication of the increased need for metalloproteins involved in replication and DNA repair (Expert et al., 2008; Puig et al., 2017), i.e., the Fe/S cluster-containing enzymes (Puig et al., 2017). Interestingly, similar expression patterns for exochelin biosynthesis, exochelin uptake, and mycobactin synthesis genes have been reported for *M. smegmatis* grown in iron deficiency conditions (Ojha and Hatfull, 2007) and under both oxidative and nitrosative stress (Namouchi et al., 2016).

The observed changes in iron metabolism in response to BRACO-19 treatment may, obviously, have a more straightforward explanation. They could all be driven by a single BRACO-19-sensitive (G4-harboring) master gene. We have considered such a possibility, and the iron-dependent regulator (*ideR*) was a primary candidate. However, *ideR* exhibited only minor expression changes and only in BRACO-19-treated samples. This finding goes against the straightforward explanation. Instead of primary effects, we are probably witnessing an adaptation to G4-related general replication stress.

It is reasonable to assume that the global replication stress arises from direct ligand interactions with G4 DNA. Interactions with G4 RNA have not been considered herein but should probably be the subject of future studies. Persistent mRNA G4s are expected to impair translation and tend to be depleted in prokaryotes (Guo and Bartel, 2016). Recent studies advocate the existence of transient G4s in both eukaryotic and prokaryotic RNA (Kwok et al., 2018; Shao et al., 2020). Importantly, G4 folding modulates the function of non-coding RNA (ncRNA), and stabilization of the ncRNA G4s by the ligands may eventually affect the transcription profile (Shao et al., 2020). Thus, a combination of G4-seq and rG4-seq (Marsico et al., 2019) may prove helpful to further elucidate the key targets of G4 ligands and the mechanisms behind their antibacterial activity.

Notably, nucleoid proteins, transcription factors, and probably some other factors affect G4 stability *in vivo* (Parekh et al., 2019), arguing that chromatin immunoprecipitation assays with G4-specific antibodies (G4-ChIP-seq) are needed to *bona fide* confirm G4 formation. The actual number of G4 targets may be substantially lower than that predicted by bioinformatics tools, which is the case in eukaryotes (Hänsel-Hertsch et al., 2018).

Moreover, regarding the transient nature of G4 structures (Yadav et al., 2021), the effects of the ligands should ideally be monitored in dynamics. We only analyzed a single time point and emphasize that the presented results, although valid for selected conditions, are insufficient for comprehensive characterization of the G4 ligands.

In conclusion, we provided a transcriptome-wide analysis of *M. smegmatis* under the action of two G4 ligands. Although there have been significant changes in the transcriptional profile of the bacterium, neither BRACO-19 nor TMPyP4 treatment induced a significant change in G4 associated genes. In turn, for the BRACO-19 compound, we found significant changes in the replication and repair system genes, which is, undoubtedly, evidence of the induced stress. Changes in iron metabolism genes also support this hypothesis. For the TMPyP4 compound, we did not find changes in the repair and replication system, however, significant changes in transcription factors and the arginine biosynthesis system may indicate multiple biological targets for this compound, which does not eliminate direct interaction with DNA.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

ES, MZ, DB, and AV wrote the main manuscript text. DB, JB, KK, VV, and IB prepared figures and tables. MM, TV, MZ, and IB cultivated mycobacteria and performed susceptibility tests. DB, KK, and VV conducted RNA-seq analysis. ES, EL, MM, JB, and AV designed the experiment. All authors reviewed the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the Russian Science Foundation (RSF, grant number 19-75-10109).

## ACKNOWLEDGMENTS

The authors thank the Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency for providing the WGS platform.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.817024/full#supplementary-material>



## REFERENCES

- Ahlmann-Eltze, C., and Patil, I. (2021). ggsignif: Significance Brackets for “ggplot2.” Available at: <https://CRAN.R-project.org/package=ggsignif> (Accessed June, 2021).
- Aubry, A., Mark Fisher, L., Jarlier, V., and Cambau, E. (2006). First functional characterization of a singly expressed bacterial type II topoisomerase: The enzyme from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* 348, 158–165. doi: 10.1016/j.bbrc.2006.07.017
- Bartas, M., Čutová, M., Brázda, V., Kaura, P., Štátný, J., Kolomazník, J., et al. (2019). The presence and localization of G-Quadruplex forming sequences in the domain of bacteria. *Molecules* 24:1711. doi: 10.3390/MOLECULES24091711
- Beaudoin, J.-D., Jodoin, R., and Perreault, J.-P. (2014). New scoring system to identify RNA G-quadruplex folding. *Nucleic Acids Res.* 42, 1209–1223. doi: 10.1093/NAR/GKT904
- Belmonte-Reche, E., and Morales, J. C. (2020). G4-iM grinder: when size and frequency matter. G-Quadruplex, i-motif and higher order structure search and analysis tool. *NAR Genom. Bioinform.* 2:lqz005. doi: 10.1093/NARGAB/LQZ005
- Boshoff, H. I. M., Reed, M. B., Barry, C. E., and Mizrahi, V. (2003). DnaE2 polymerase contributes to In vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell* 113, 183–193. doi: 10.1016/S0092-8674(03)00270-8
- Buchfink, B., Xie, C., and Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Byrd, A. K., and Raney, K. D. (2017). Structure and function of Pif1 helicase. *Biochem. Soc. Trans.* 45, 1159–1171. doi: 10.1042/BST20170096
- Ding, Y., Fleming, A. M., and Burrows, C. J. (2018). Case studies on potential G-quadruplex-forming sequences from the bacterial orders Deinococcales and Thermales derived from a survey of published genomes. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-33944-4
- Edwards, S. M. (2020). Lemon: Freshing Up your “ggplot2” plots. Available at: <https://CRAN.R-project.org/package=lemon> (Accessed June, 2021).
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/BIOINFORMATICS/BTW354
- Expert, D., Boughammoura, A., and Franza, T. (2008). Siderophore-controlled iron assimilation in THE Enterobacterium *Erwinia chrysanthemi*: Evidence For The Involvement Of bacterioferritin and the Suf iron-sulfur cluster assembly Machinery. *J. Biol. Chem.* 283, 36564–36572. doi: 10.1074/JBC.M807749200
- Global Tuberculosis Report 2020 (2020). Available at: <http://apps.who.int/bookorders> (Accessed September 27, 2021).
- Guo, J. U., and Bartel, D. P. (2016). RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* 353:aaf5371. doi: 10.1126/science.aaf5371
- Hänsel-Hertsch, R., Spiegel, J., Marsico, G., Tannahill, D., and Balasubramanian, S. (2018). Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* 13, 551–564. doi: 10.1038/nprot.2017.150
- Harris, L. M., and Merrick, C. J. (2015). G-Quadruplexes in pathogens: a common route to virulence control? *PLoS Pathog.* 11:e1004562. doi: 10.1371/JOURNAL.PPAT.1004562
- Holder, I. T., and Hartig, J. S. (2014). A matter of location: influence of G-Quadruplexes on *Escherichia coli* gene expression. *Chem. Biol.* 21, 1511–1521. doi: 10.1016/j.chembiol.2014.09.014
- Huber, M. D., Duquette, M. L., Shiels, J. C., and Maizels, N. (2006). A conserved G4 DNA binding domain in RecQ family helicases. *J. Mol. Biol.* 358, 1071–1080. doi: 10.1016/j.jmb.2006.01.077
- Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/NAR/GKY1085
- Incles, C. M., Schultes, C. M., Kempfski, H., Koehler, H., Kelland, L. R., and Neidle, S. (2004). A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. *Mol. Cancer Ther.* 3, 1201–1206. doi: 10.1158/1535-7163.1201.3.10
- Joseph Antony Sundarsingh, T., Ranjitha, J., Rajan, A., and Shankar, V. (2020). Features of the biochemistry of *Mycobacterium smegmatis*, as a possible model for *Mycobacterium tuberculosis*. *J. Infect. Public Health* 13, 1255–1264. doi: 10.1016/j.jiph.2020.06.023
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kwok, C. K., Marsico, G., and Balasubramanian, S. (2018). Detecting RNA G-quadruplexes (rG4s) in the transcriptome. *Cold Spring Harb. Perspect. Biol.* 10:a032284. doi: 10.1101/CSHPERSPECT.A032284
- Lerner, L. K., and Sale, J. E. (2019). Replication of G Quadruplex DNA. *Genes* 10:95. doi: 10.3390/GENES10020095
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/BIOINFORMATICS/BTP352
- Li, X., Mei, H., Chen, F., Tang, Q., Yu, Z., Cao, X., et al. (2017). Transcriptome landscape of *Mycobacterium smegmatis*. *Front. Microbiol.* 8:2505. doi: 10.3389/FMICB.2017.02505
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/BIOINFORMATICS/BTT656
- Logsdon, M. M., and Aldridge, B. B. (2018). Stable regulation of cell cycle events in mycobacteria: insights from inherently heterogeneous bacterial populations. *Front. Microbiol.* 9:514. doi: 10.3389/FMICB.2018.00514
- Marsico, G., Chambers, V. S., Sahakyan, A. B., McCauley, P., and Boutell, J. M., Antonio, M. di Antonio, M. D., and Balasubramanian, S. (2019). Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.* 47, 3862–3874. doi:10.1093/NAR/GKZ179.
- Maslowska, K. H., Makiela-Dzubska, K., and Fijalkowska, I. J. (2019). The SOS system: A complex and tightly regulated response to DNA damage. *Environ. Mol. Mutagen.* 60, 368–384. doi: 10.1002/EM.22267
- Mishra, S. K., Shankar, U., Jain, N., Sikri, K., Tyagi, J. S., Sharma, T. K., et al. (2019). Characterization of G-quadruplex motifs in espB, espK, and cyp51 genes of *Mycobacterium tuberculosis* as potential drug targets. *Mol Ther Nucleic Acids* 16, 698–706. doi: 10.1016/j.omtn.2019.04.022
- Namouchi, A., Gómez-Muñoz, M., Frye, S. A., Moen, L. V., Rognes, T., Tønjum, T., et al. (2016). The *Mycobacterium tuberculosis* transcriptional landscape under genotoxic stress. *BMC Genomics* 17, 1–13. doi: 10.1186/S12864-016-3132-1
- Ojha, A., and Hatfull, G. F. (2007). The role of iron in mycobacterium smegmatis biofilm formation: the exochelin siderophore is essential in limiting iron conditions for biofilm formation but not for planktonic growth. *Mol. Microbiol.* 66, 468–483. doi: 10.1111/J.1365-2958.2007.05935.X
- Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32, btv566–btv294. doi: 10.1093/BIOINFORMATICS/BTV566
- Ordóñez, H., and Shuman, S. (2014). *Mycobacterium smegmatis* DinB2 misincorporates deoxyribonucleotides and ribonucleotides during templated synthesis and lesion bypass. *Nucleic Acids Res.* 42, 12722–12734. doi: 10.1093/NAR/GKU1027
- Parekh, V. J., Niccum, B. A., Shah, R., Rivera, M. A., Novak, M. J., Geinguenaud, F., et al. (2019). Role of Hfq in genome evolution: instability of G-Quadruplex sequences in *E. coli*. *Microorganisms* 8:28. doi: 10.3390/MICROORGANISMS8010028
- Pedersen, T. L. (2021). ggforce: Accelerating “ggplot2.” Available at: <https://CRAN.R-project.org/package=ggforce> (Accessed June, 2021).
- Perrone, R., Lavezzo, E., Riello, E., Manganelli, R., Palù, G., Toppo, S., et al. (2017). Mapping and characterization of G-quadruplexes in mycobacterium tuberculosis gene promoter regions. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-05867-z
- Puig, S., Ramos-Alonso, L., Romero, A. M., and Martínez-Pastor, M. T. (2017). The elemental role of iron in DNA synthesis and repair. *Metallomics* 9, 1483–1500. doi: 10.1039/C7MT00116A
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/BIOINFORMATICS/BTQ033
- R Core Team (2021). R: A Language and Environment for Statistical Computing. Available at: <https://www.R-project.org/> (Accessed June, 2021).
- Rand, L., Hinds, J., Springer, B., Sander, P., Buxton, R. S., and Davis, E. O. (2003). The majority of inducible DNA repair genes in *Mycobacterium*

- tuberculosis* are induced independently of RecA. *Mol. Microbiol.* 50, 1031–1042. doi: 10.1046/J.1365-2958.2003.03765.X
- Rha, S. Y., Izbicka, E., Lawrence, R., Davidson, K., Sun, D., Moyer, M. P., et al. (2000). Effect of telomere and telomerase interactive agents on human tumor and normal cell lines. *Clin. Cancer Res.* 6, 987–993.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/BIOINFORMATICS/BTP616
- Rodriguez, R., Miller, K. M., Forment, J. V., Bradshaw, C. R., Nikan, M., Britton, S., et al. (2012). Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* 8, 301–310. doi: 10.1038/nchembio.780
- Saha, T., Shukla, K., Thakur, R. S., Desingu, A., and Nagaraju, G. (2019). *Mycobacterium tuberculosis* UvrD1 and UvrD2 helicases unwind G-quadruplex DNA. *FEBS J.* 286, 2062–2086. doi: 10.1111/FEBS.14798
- Saranathan, N., and Vivekanandan, P. (2019). G-Quadruplexes: more Than just a kink in microbial genomes. *Trends Microbiol.* 27, 148–163. doi: 10.1016/J.TIM.2018.08.011
- Sauer, M., and Paeschke, K. (2017). G-quadruplex unwinding helicases and their function in vivo. *Biochem. Soc. Trans.* 45, 1173–1182. doi: 10.1042/BST20170097
- Seung, K. J., Keshavjee, S., and Rich, M. L. (2015). Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harb. Perspect. Med.* 5:a017863. doi: 10.1101/cshperspect.a017863
- Shao, X., Zhang, W., Umar, M. I., Wong, H. Y., Seng, Z., Xie, Y., et al. (2020). RNA G-quadruplex structures mediate gene regulation in bacteria. *mBio* 11, e02926–e02919. doi: 10.1128/MBIO.02926-19
- Singh, A. (2017). Guardians of the mycobacterial genome: A review on DNA repair systems in mycobacterium tuberculosis. *Microbiology* 163, 1740–1758. doi: 10.1099/MIC.0.000578
- Singh, A. (2018). Expanding the horizon of recombination repair in mycobacteria: identification and characterization of novel proteins. *FASEB J.* 32(Suppl. 1):lb1-925.9. doi: 10.1096/FASEBJ.2018.32.1\_SUPPLEMENT.LB1
- Singh, A., Vijayan, M., and Varshney, U. (2018). Distinct properties of a hypoxia specific paralog of single stranded DNA binding (SSB) protein in mycobacteria. *Tuberculosis* 108, 16–25. doi: 10.1016/J.TUBE.2017.10.002
- Smid, M., van den Braak, R. R. J. C., van de Werken, H. J. G., van Riet, J., van Galen, A., de Weerd, V., et al. (2018). Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* 19:236. doi: 10.1186/S12859-018-2246-7
- Spiegel, J., Adhikari, S., and Balasubramanian, S. (2020). The structure and function of DNA G-Quadruplexes. *Trends Chem.* 2, 123–136. doi: 10.1016/J.TRECHM.2019.07.002
- Sritharan, M. (2016). Iron homeostasis in *Mycobacterium tuberculosis*: mechanistic insights into siderophore-mediated iron uptake. *J. Bacteriol.* 198, 2399–2409. doi: 10.1128/JB.00359-16
- Thakur, R. S., Desingu, A., Basavaraju, S., Subramanya, S., Rao, D. N., and Nagaraju, G. (2014). *Mycobacterium tuberculosis* DinG is a structure-specific helicase that unwinds G4 DNA: implications for targeting G4 DNA as a novel therapeutic approach. *J. Biol. Chem.* 289, 25112–25136. doi: 10.1074/JBC.M114.563569
- Uson, M. L., Ordonez, H., and Shuman, S. (2015). Mycobacterium smegmatis HelY is an RNA-activated ATPase/dATPase and 3'-to-5' helicase that unwinds 3'-tailed RNA duplexes and RNA:DNA hybrids. *J. Bacteriol.* 197, 3057–3065. doi: 10.1128/JB.00418-15
- Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., et al. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* 42, D581–D591. doi: 10.1093/NAR/GKT1099
- Wickham, H. (2009). *ggplot2*. Springer New York.
- Wilke, C. O. (2020). cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2.” Available at: <https://CRAN.R-project.org/package=cowplot> (Accessed June, 2021).
- Wu, R., Zheng, K., Zhang, J., Hao, Y., and Tan, Z. (2015). Formation of DNA:RNA hybrid G-Quadruplex in bacterial cells and its dominance over the Intramolecular DNA G-Quadruplex in mediating transcription termination. *Angew. Chem. Int. Ed.* 54, 2447–2451. doi: 10.1002/ANIE.201408719
- Yadav, P., Kim, N., Kumari, M., Verma, S., Sharma, T. K., Yadav, V., et al. (2021). G-quadruplex structures in bacteria: biological relevance and potential as an antimicrobial target. *J. Bacteriol.* 203:e0057720. doi: 10.1128/JB.00577-20
- Yan, L. (2021). ggvenn: Draw Venn Diagram by “ggplot2.” Available at: <https://CRAN.R-project.org/package=ggvenn> (Accessed June, 2021).
- Yang, M., Aamodt, R. M., Dalhus, B., Balasingham, S., Helle, I., Andersen, P., et al. (2011). The ada operon of mycobacterium tuberculosis encodes two DNA methyltransferases for inducible repair of DNA alkylation damage. *DNA Repair* 10, 595–602. doi: 10.1016/J.DNAREP.2011.03.007
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters 16, 284–287. doi: 10.1089/OMI.2011.0118

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shitikov, Bespiatykh, Malakhova, Bespyatykh, Bodoev, Vedekhina, Zaychikova, Veselovsky, Klimina, Ilna and Varizhuk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.