



# Exploring Computational and Biophysical Tools to Study the Presence of G-Quadruplex Structures: A Promising Therapeutic Solution for Drug-Resistant *Vibrio cholerae*

Uma Shankar<sup>1</sup>, Neha Jain<sup>1</sup>, Prativa Majee<sup>1</sup>, Prashant Kodgire<sup>1</sup>, Tarun Kumar Sharma<sup>2</sup> and Amit Kumar<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Sandeep Tiwari,  
Federal University of Minas Gerais,  
Brazil

### Reviewed by:

Satyabrata Bag,  
National Centre for Disease Control  
(NCDC), India  
Pramod Kumar,  
National Centre for Disease Control  
(NCDC), India  
Eiji Arakawa,  
National Institute of Infectious  
Diseases (NIID), Japan

### \*Correspondence:

Amit Kumar  
amitk@iiti.ac.in;  
amitkdb@gmail.com

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 18 May 2020

**Accepted:** 27 July 2020

**Published:** 25 September 2020

### Citation:

Shankar U, Jain N, Majee P,  
Kodgire P, Sharma TK and Kumar A  
(2020) Exploring Computational  
and Biophysical Tools to Study  
the Presence of G-Quadruplex  
Structures: A Promising Therapeutic  
Solution for Drug-Resistant *Vibrio*  
*cholerae*. *Front. Genet.* 11:935.  
doi: 10.3389/fgene.2020.00935

<sup>1</sup> Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore, India, <sup>2</sup> Translational Health Science and Technology Institute, Faridabad, India

*Vibrio cholerae*, a gram-negative bacterium that causes cholera, has already caused seven major pandemics across the world and infects roughly 1.3–4 million people every year. Cholera treatment primarily involves oral rehydration therapy supplemented with antibiotics. But recently, multidrug-resistant strains of *V. cholerae* have emerged. High genomic plasticity further enhances the pathogenesis of this human pathogen. Guanines in DNA or RNA assemble to form G-quadruplex (GQ) structures which have begun to be seen as potential drug targeting sites for different pathogenic bacteria and viruses. In this perspective, we carried out a genome-wide hunt in *V. cholerae* using a bio-informatics approach and observed ~85 G-quadruplex forming motifs (VC-PGQs) in chromosome I and ~45 putative G-quadruplexes (PGQs) in chromosome II. Ten putative G-quadruplex forming motifs (VC-PGQs) were selected on the basis of conservation throughout the genus and functional analysis displayed their location in the essential genes encoding bacterial proteins, for example, methyl-accepting chemotaxis protein, orotate phosphoribosyl transferase protein, amidase proteins, etc. The predicted VC-PGQs were validated using different bio-physical techniques, including Nuclear Magnetic Resonance spectroscopy, Circular Dichroism spectroscopy, and electrophoretic mobility shift assay, which demonstrated the formation of highly stable GQ structures in the bacteria. The interaction of these VC-PGQs with the known specific GQ ligand, TMPyP4, was analyzed using ITC and molecular dynamics studies that displayed the stabilization of the VC-PGQs by the GQ ligands and thus represents a potential therapeutic strategy against this enteric pathogen by inhibiting the PGQ harboring gene expression, thereby inhibiting the bacterial growth and virulence. In summary, this study reveals the presence of conserved GQ forming motifs in the *V. cholerae* genome that has the potential to be used to treat the multi-drug resistance problem of the notorious enteric pathogen.

**Keywords:** G-quadruplex, multi drug-resistant, *Vibrio cholerae*, therapeutic strategy, gram-negative

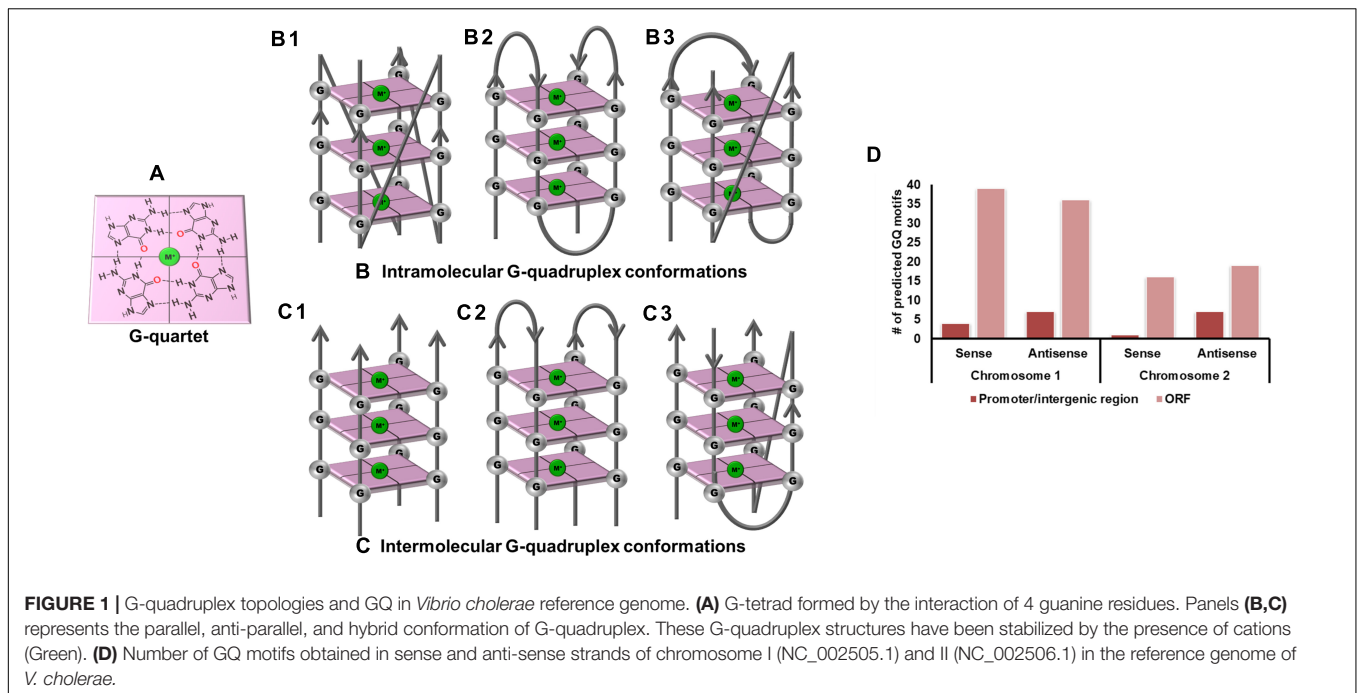
## INTRODUCTION

Cholera, one of the common bacterial diseases, has had a catastrophic effect on the human race. The disease has resulted in seven major pandemic situations since 1817, and is known to be endemic in more than 47 countries, especially in developing countries (Hall, 2018). Cholera is caused by a gram-negative, curved-shaped, flagellated bacterium called *Vibrio cholerae* (*V. cholerae*; Reidl and Klose, 2002). A high infectious dose ( $10^3$ – $10^8$ ) and a very fast replication time of 17 min of *V. cholerae* leads to its higher pathogenesis (Hun Yoon and Waters, 2019). Approximately 1.3–4 million people are infected by this bacterium every year, taking the lives of 21,000–143,000 individuals worldwide (Islam et al., 2017). This disease can be treated if early detection and subsequent treatment is done, but the disease causes major concern as it largely impacts young children in the age group of 1–5 years. *V. cholerae* is usually transmitted through contaminated food and water, and as the bacteria enter the human body, it attaches itself into the epithelial membrane and secretes the cholera toxin. The bacterial infection leads to excessive watery diarrhea which results in severe dehydration followed by shock and, in due course, causes death. There are more than 220 serogroups of *V. cholerae* prevailing in the environment, but the O1 serogroup was mainly accountable for the seven pandemics that have occurred so far (Shimada et al., 1994). Out of this, the first six pandemics were caused by the classical biotype of O1 while the last pandemic was a result of the E1 Tor biotype of *V. cholerae* O1 (Cho et al., 2010). Cholera patients are usually supplemented with antibiotics along with oral rehydration therapy to minimize the severity of the disease and maintain the fluid level of the body. But the use of these antibiotics comes with malediction in the form of resistance. The most commonly used antibiotics against the bacterium includes tetracycline, erythromycin, quinolones, sulfonamides, bleomycin, chloramphenicol, and aminoglycosidase, but due to the extraordinary genomic plasticity, around 90% of the recent *V. cholerae* isolates have attained antimicrobial resistance against them (Verma et al., 2019). The contributing factors towards this increasing antibiotic resistance include chromosomal mutations, export of the antibiotics outside the cell through efflux pumps, and attaining genetic resistance through the exchange of plasmids, transposons, SXT elements, and autonomously replicating and integrating plasmids or integrons (Kumar et al., 2009; Kitaoka et al., 2011a; Verma et al., 2019). Integrating conjugative elements (ICE) and superintegrons from closely or distinctly related bacteria are the major causes of antimicrobial resistance in *V. cholerae* (Verma et al., 2019; Das et al., 2020). Various strategies have been enforced to combat these MDR strains, including the usage of various efflux pump inhibitors and quorum sensing inhibitors, but the problem remains unsolved (Hema et al., 2016; Lawrence et al., 2019). This multi-drug resistant strains stands as a barrier to the treatment of cholera disease and demands the attention of the scientific community.

Research in the field of short secondary-structure forming motifs has recently begun to gain momentum. One such type of non-canonical structures which has been extensively studied is G-Quadruplex (GQ) structures. These specialized structures are

formed when two or more guanine residues interspersed by some nucleotides self-assemble to form quartet structures (Figure 1). G-quadruplex shows structural polymorphism based on guanine repeats, loop sequences, loop length, physiological conditions, cations available, etc., thus opening up a plethora of motifs for drug designing. The GQ structures have been extensively reported to be found in the human genome, especially in the telomeric end of the chromosome (Patel et al., 2007) and in the promoter regions of the proto-oncogenes like *c-Myc*, *c-Kit*, *bcl2*, and *KRAS* (Cogoi and Xodo, 2006; Dai et al., 2006a,b; Yang and Hurley, 2006; Huppert and Balasubramanian, 2007). The relevance of GQ structures were found in all the cellular processes, including DNA recombination, RNA transcription, protein translation, and in recombination, which are imperative for a cell's survival (Kim, 2019). Both the nucleic acids, DNA and RNA, are capable of forming these specialized structures. The GQ structures forming in a specific part of the genome hold a unique role to play; those in the promoter region, for example, modulates the transcription of the genes (Siddiqui-Jain et al., 2002; Rigo et al., 2017; Kong et al., 2018) while those present in the open reading frame (ORF) region affect the replication, transcription, and translation of genes harboring the GQ (Endoh et al., 2013b). Even GQ structures in the untranslated region (UTR) hold significant functions, including miRNA binding (Rouleau et al., 2017), cis-eQTLs, and RNA-binding protein (RBP) interactions (Bugaut and Balasubramanian, 2012; Lee et al., 2020).

Genome mining of GQ forming sequences in different organisms has opened up a new avenue for drug-targeting against the pathogenic organisms (Saranathan and Vivekanandan, 2018). With new reports of GQ structures being involved with different human diseases, development of ligands specifically targeting these GQ structures have also accelerated, and a few have even gone through clinical trials (Yang and Okamoto, 2010; Vy Thi Le et al., 2012; Asamitsu et al., 2018). Various pathogenic organisms, including bacteria, virus, protozoa (Bhartiya et al., 2016), and fungi (Hershman et al., 2008), are reported to bear these unique secondary structures in their genome. For instance, *Mycobacterium tuberculosis* was reported to bear GQ structures in the ORF region of the *espB*, *espK*, and *cyp51* genes (Mishra et al., 2019b) as well as in its gene promoter region (Perrone et al., 2017b). G-quadruplex structures have been shown to be associated with the antigenic variation mechanism of *pilE* gene in *Neisseria gonorrhoeae* (Cahoon and Seifert, 2009; Zhang et al., 2011), *tprK* gene in *Treponema pallidum* (Giacani et al., 2012), and *vlsE* gene in *Borrelia burgdorferi* (Walia and Chaconas, 2013). Interestingly, three highly conserved GQ forming motifs were recognized in the *hsdS*, *recD*, and *pmrA* genes which could act as potential drug targets against pathogenic *Streptococcus pneumoniae* (Mishra et al., 2019a). A recent analysis revealed the presence of GQ motifs in various genes of *Klebsiella pneumoniae* that influenced the gene expression (Shankar et al., 2020). Likewise, several human infecting viruses including Zika (Fleming et al., 2016), Nipah (Majee et al., 2020a), Herpes Simplex virus (Biswas et al., 2016), Human Papillomavirus (Tlučková et al., 2013), and Human Adenovirus (Majee et al., 2020b) are reported to possess potential GQ forming sequences. G-quadruplex sequences in the L gene of Ebola virus and the



core gene of Hepatitis C Virus were found to be novel antiviral drug targets as stabilization of these structures by GQ binding ligands like TMPyP4 [5,10,15,20-Tetrakis-(N-methyl-4-pyridyl) porphine] and PDP negatively regulated the expression of the genes harboring the GQ structures (Wang et al., 2016a,b). Thus, there is an enormous amount of evidence proving the GQ structures to be promising therapeutic targets.

With this perspective, we tried to explore the *V. cholerae* genome for potential GQ forming sequences that could unlock new anti-bacterial strategies for treating cholera. Evaluation of the most recent pandemic causing strain of *V. cholerae*, O1 biovar El Tor str. N16961, revealed the presence of the ten most conserved VC-PGQ (*V. cholerae* – Putative G quadruplex) sequences. As the complete genome of *V. cholerae* is coded by two chromosomes, the large chromosome consists of seven conserved VC-PGQs in the essential bacterial genes, like methyl accepting chemotaxis protein coding gene, *rtxA* gene, GGDEF family protein coding genes, etc., while the other three VC-PGQs are present in the small chromosome. All the ten VC-PGQs of both the chromosomes were found to form stable GQ structures when evaluated through different biophysical techniques like NMR, CD spectroscopy, and EMSA and was also found to have very high binding affinity for the GQ specific ligand, TMPyP4.

## MATERIALS AND METHODS

### Exploring the Available *Vibrio cholerae* Genome Sequences for PGQ Prediction

All the 68 available complete genome sequences of *V. cholerae* were retrieved from the NCBI genome database<sup>1</sup>. In this study,

<sup>1</sup><http://www.ncbi.nlm.nih.gov/genomes>

*V. cholerae* O1 biovar El Tor str. N16961 was used as the reference genome. Each chromosome of the bacteria was individually evaluated for the PGQ prediction. A total of 68 chromosome I and 65 chromosome II of the bacterial strains were checked for the presence of GQ forming sequence through our in-house available GQ prediction tool (Mishra et al., 2016). The PGQ prediction tools make use of the following notion:

$$G \geq 3N_{1-20}G \geq 3N_{1-20}G \geq 3N_{1-20}G \geq 3$$

where G represents to the guanine residues and N represents any of the four nucleotides, including guanine. While the Guanine repeat was set to a minimum of 3, the loop length was set between 1 and 20. The tools scrutinized both the sense and anti-sense strand of the bacterial genome for the PGQ prediction. The PGQ sequences were re-confirmed by using the other online available tool, QGRS Mapper (Kikin et al., 2006). Further, the PGQ sequences were aligned to generate the consensus motifs which depicted the degree of conservation among the nucleotides.

### Determination of Genomic Locations and the Related Functions of the PGQ Sequences

The genomic location and the associated function of the predicted PGQs were determined using the NCBI nucleotide database. The Graphics mode of the GenBank database was further explored for this purpose to analyze whether the PGQ is located in the promoter region, in the ORF, or in the intergenic region.

### Sample Preparation

All the oligonucleotide sequences used in the study were procured from Sigma Aldrich (Bangalore, India), and 100  $\mu$ M

of stock solutions were prepared by adding the required amount of MiliQ water according to the manufacturer's specification. Then the required dilutions were prepared in Tris-Cl buffer (10mM, pH 7.4) containing 50 mM of either of the four different cations, namely  $K^+$ ,  $Na^+$ ,  $Li^+$ , and  $Mg^{2+}$ . The samples were thermally denatured by heating at 92°C for 10 min and then slowly cooled down to room temperature before performing each set of experiments.

## 1D $^1H$ Nuclear Magnetic Resonance Studies

The 1D  $^1H$  nuclear magnetic resonance (NMR) experiments were performed using the AVANCE III 400 MHz Ascend Bruker BioSpin International AG, Switzerland model furnished with the 5 mm broadband inverse (BBI) probe. While Sodium trimethylsilylpropanesulfonate (DSS) was taken as the reference compound, the oligonucleotide sample was prepared in 90/10%  $H_2O/D_2O$  solution containing 50 mM potassium phosphate buffer at 298 K. The data obtained were further processed and analyzed by using Topspin 4.0 (academic license) software.

## Circular Dichroism (CD) Spectrophotometer Experiments

All the circular dichroism (CD) spectra and CD melting experiments were carried out on the Jasco J-815 Spectropolarimeter (Jasco Hachioji, Tokyo, Japan) accompanied with a temperature controlling Peltier junction. Nitrogen gas was continuously circulated within the instrument to avoid condensation around the cuvette. Quartz cuvette of 1mm path length was used for the experiments. The CD spectra profile for the each PGQ was recorded over the wavelength range of 220–320 nm with a scanning speed of 20 nm/min at 25°C. The final concentration of PGQ DNA was maintained at 20  $\mu M$  in the 10 mM Tris buffer containing either of the four cations ( $K^+$ ,  $Na^+$ ,  $Li^+$ , and  $Mg^{2+}$ ). Buffer correction was done for each set of experiments, while each of the spectra was recorded in triplicate. The data obtained from the CD spectrophotometer was then plotted using the SigmaPlot software.

For the CD melting experiments, the temperature range was set between 25 and 98°C with a gradual increase at a rate of 1°C/min. The CD melting plot was collected at a particular wavelength throughout the experiment in the presence of the same Tris-Cl buffer.

## Electrophoretic Mobility Shift Assay

Electrophoretic mobility shift assay (EMSA) was performed with the help of 25% native polyacrylamide gel in the presence of 1X Tris-Borate-EDTA (TBE) buffer. The PGQ sequences were evaluated in four different buffer conditions; therefore, no cation was added into the running TBE buffer (in tank), rather individual oligonucleotide (20  $\mu M$  concentration) were dissolved in the respective buffer (containing either  $K^+$ ,  $Na^+$ ,  $Li^+$ , or  $Mg^{2+}$ ) before loading into the gel. Standard GQ forming sequences like *cKit21* was used as the positive control and linear DNA oligonucleotides of the equal length as that of PGQ as

the negative control. The gel assays were performed in a Bio-Rad Mini protean Tetra Vertical Electrophoresis unit at 4°C, 90 V. The gels were then soaked in Ethidium Bromide solution for staining and further visualized in ImageQuant LAS4000 (GE Healthcare, Biosciences Ltd., Sweden).

## Isothermal Titration Calorimetry Studies of VC-PGQs With TMPyP4

Isothermal titration calorimetry (ITC) studies between the DNA and ligand was accomplished using the MicroCal iTC200 isothermal titration calorimeter (GE Healthcare, United States) at a constant temperature of 25°C. The oligonucleotide and the ligand solution were prepared in the same buffer, i.e., 1X potassium phosphate buffer (pH 7.4). Total solution of TMPyP4 was titrated into the PGQ containing solution through 21 injections of 1.80  $\mu L$  each. The duration of each injection was 3.6 s while the spacing between each successive injection was kept at 120 s. The heat of dilution was determined in each set of experiments and further subtracted from the binding isotherms for the purpose of buffer correction. The data obtained were fitted through the “two mode of binding sites” model using the MicroCal Origin software.

## G-Quadruplex Structure Prediction

Two VC-PGQs, VC-I-PGQ-7, and VC-II-PGQ-1 were modeled by exploiting comparative structure modeling using ModeRNA tool (Rother et al., 2011). For structure prediction, PDB ID: 2M27 was used as a template for VC-I-PGQ-7, and PDB ID: 3SC8 was used for VC-II-PGQ-1. The resultant RNA models were converted to DNA using Discovery Studio v4.0. The image processing was performed using Open source PyMOL visualization tool.

## Molecular Docking of GQs With Potassium ( $K^+$ ) and TMPyP4

AutoDock tools were used for interaction analysis of GQs with  $K^+$  and TMPyP4, respectively. TMPyP4 was downloaded from PubChem (CID: 135442972). Receptor (GQ) and ligand (TMPyP4) were prepared using the available AutoDockTools with the MGLTools package and converted to PDBQT files. For VC-I-PGQ-7, a grid was created by using grid points  $x = 44$ ,  $y = 52$ ,  $z = 46$  with a grid spacing of 0.61. While for VC-II-PGQ-1, a grid was prepared with grid points  $x = 56$ ,  $y = 52$ , and  $z = 46$  with spacing of 0.61. Docking results were analyzed by running docking log files (.dlg) for the respective complexes.

## Molecular Dynamic Simulation Analysis

Molecular dynamic simulation (MDS) was performed using NAMD Suite (Melo and Bernardi, 2018). Input files for simulation analysis was performed by utilizing CHARMM 36m force fields (Huang et al., 2017) using CHARMM-GUI server (Jo et al., 2008). The modeled GQs were solvated in a rectangular water box with an edge distance of 10.0 with periodic boundary conditions. To neutralize the charges, 0.15 M KCl was added using the Monte-Carlo method of ion placement. The grid was constructed using Particle mesh Ewald – Fast

Fourier transformation (PME-FFT). Structures were minimized for 10,000 steps using steepest descent algorithm. Thereafter, the system was equilibrated in isochoric equilibration (NVT) and isobaric conditions (NPT). Nose-Hoover Langevin algorithm was used for pressure control. After this, a production run of 10 ns was performed at constant temperature of 310 K. The trajectory analysis was performed using CPPTraj tool (Roe and Cheatham, 2013).

For analyzing the global fluctuations in the system, root mean square deviation (RMSD) was calculated. For RMSD calculation, the first frame (initial structure at TS = 0 ps) was used as a reference, and rest frames in the simulation trajectory were compared. For global packaging analysis, radius of gyration (Rg) was generated for all the time frames.

## RESULTS

### Mining of Conserved PGQ Sequences From the *Vibrio cholerae* Genome

*Vibrio cholerae* interestingly possesses its complete genome in two circular chromosomes, a large chromosome bearing most of the essential genes required for cell growth and pathogenicity and a small chromosome that contains most of the hypothetical protein coding genes and the supporting genes helping the bacteria to survive different environmental conditions. This bipartite genome was evaluated for the probable GQ forming sequences and, due to the bi-directional mode of replication, both sense and anti-sense strands were taken into consideration. In GQ structure formation, G-tract and loop length plays a major role and it is considered that a minimum G-tract of 3 and loop length in the range of 1–20 nucleotides can form stable GQ conformation *in vitro* and *in vivo*. Therefore, both the chromosomes of *V. cholerae* strains were analyzed for GQ motif prediction by utilizing G-tract = 3 and loop length 1–20 using our in-house developed GQ prediction tool (Version 2.0) (Mishra et al., 2016). A total of 68 complete genomes available in NCBI Genome database till January 31, 2020 were analyzed for GQ prediction (Supplementary Table 1). G-quadruplex prediction analysis revealed an average ~85 G-quadruplex forming motifs (VC-PGQs) in chromosome I and ~45 PGQs in chromosome II (Supplementary Dataset S2). For the reference genome, *V. cholerae* O1 biovar El Tor str. N16961 strain, 86 PGQs with 43 each in a sense, and anti-sense strands were observed for chromosome I (Figure 1D and Supplementary Dataset 2). Out of these, four PGQs in the sense strand and seven PGQs in the anti-sense strands were observed in the promoter regions of various genes while the rest were observed to be located in the open reading frames. Similarly, for chromosome II, 17 PGQs (1 in the promoter region, 16 in ORFs) in the sense and 26 PGQs (7 in promoter and 19 in ORFs) in the anti-sense strands were predicted that satisfied the GQ prediction parameters (Figure 1D and Supplementary Dataset 2). Further sorting on the basis of conservation, ten unique PGQs, seven in chromosome I and three in chromosome II, were retrieved with maximum percentage of conservation among all the available genomes (Table 1 and Supplementary Figure 1). These conserved motifs were also

predicted using QGRS Mapper that provided a good G-score that strengthens the propensity of GQ formation by these motifs (Supplementary Table 2).

### NMR Spectroscopy Displays the Distinctive Imino Peaks of the GQ Specific Hoogsteen Bonds in VC-PGQs

Nuclear Magnetic Resonance spectroscopy is one of the most quintessential techniques used to explore GQ formation. The 1D proton NMR spectra performed for the VC-PGQ sequences dissolved in the potassium containing buffer revealed the distinctive chemical shift signal in the 10–12 ppm region which showcases the imino protons involved in the Hoogsteen hydrogen bond formation of a GQ structure (Feigon et al., 1995; Figure 2). The GQ structures are formed by stacking of G-quartets, which are in turn made up of four guanine residues interconnected by the help of Hoogsteen bonds. These Hoogsteen hydrogen bonds differ from the canonical hydrogen bonds involved in the Watson-Crick base pairing as they link the amino and imino protons of one guanine base to the nitrogen (N7) and oxygen (O6) atoms of the adjacent guanine, respectively. The imino protons involved in the G-tetrad formation are very stable and exchange very slowly with the solvent molecules, thereby allowing their detection through the NMR technique for a long duration of time (Adrian et al., 2012). All the ten VC-PGQs present in both *V. cholerae* chromosomes clearly displayed the peaks in the region between 10 and 12 ppm, which established the fact that the predicted VC-PGQs form stable GQ structures (Figure 2). Along with the GQ chemical shift, three PGQs, VC-I-PGQ-5, VC-I-PGQ-7, and VC-II-PGQ-3, also revealed the formation of small peaks in the 12–14 ppm region depicting the formation of an additional secondary structure with canonical Watson-crick bonding along with the GQ conformation (Figure 2). Presence of a chemical shift in both the regions simultaneously depicts the presence of GQ-hairpin-duplex conformation similar to the one observed in *Escherichia coli* (Kaplan et al., 2016).

### Prediction of the Topology Exhibited by VC-PGQs Structure by Circular Dichroism Spectroscopy in the Presence of Various Cations

The GQ structures demonstrate structural polymorphism due to factors such as their sequence composition and loop length physiological environment. Preliminary ideas regarding the structural conformation of the GQ structures can be obtained through the Circular Dichroism spectroscopy, as it provides us with the signature pattern of spectra related to the relative strand orientation and folding during the GQ formation in a particular physiological condition. A large body of evidence suggests that parallel topology of GQ is confirmed when a positive peak at ~265 nm and a negative peak at ~240 nm is observed in the CD spectra whereas the anti-parallel topology is indicated when the positive peak appears in ~290 nm and the negative peak at the ~260 nm. More often, a hybrid or mixed topology may arise during GQ formation exhibiting two simultaneous positive peaks

**TABLE 1** | List of 10 conserved VC-PGQs along with their locations, strands, and genes harboring them in chromosome I (NC\_002505.1) and II (NC\_002506.1) in the reference genome of *Vibrio cholerae*.

VC-PGQs	GQ Sequence	Start location	End location	GQ strand	Gene harboring GQ motifs
VC-I-PGQ-1	GAAAT- <b>GGG</b> ATT <b>GGG</b> TATT <b>GGGG</b> AC <b>GGG</b> -CATCT	223986	224007	Sense	ORF region of methyl-accepting chemotaxis protein (Involved in chemotaxis)
VC-I-PGQ-2	TTTTA- <b>GGGG</b> A <b>GGG</b> CTA <b>GGGT</b> <b>GGGG</b> -TATTA	366170	366188	Anti-sense	N-acetyl-anhydromuranmyl-L-alanine amidase (801 bp upstream of 3' end)
VC-I-PGQ-3	TTCTA- <b>GGGG</b> A <b>GGG</b> TCAG <b>GGT</b> <b>GGGG</b> -TTACC	2742628	2742646	Anti-sense	Orotate phosphoribosyltransferase (149 bp upstream of 5' end)
VC-I-PGQ-4	CATGT- <b>GGGG</b> GA <b>AGG</b> GAAT <b>GCTT</b> <b>GGG</b> CGCTGAT <b>GGG</b> -CGGCT	1552798	1552827	Sense	ORF region of RTX toxin <i>rtxA</i>
VC-I-PGQ-5	GTCTA- <b>GGG</b> TCTA <b>GGG</b> TCTA <b>GGG</b> TCTA <b>GGG</b> -TCTAG	2493993	2494016	Anti-sense	GTPase ObgE (108bp upstream of 5' end)
VC-I-PGQ-6	AGAGT- <b>GGG</b> TGCTACT <b>GGG</b> CTTAG <b>GGT</b> <b>GGG</b> -TGGCG	1470345	1470369	Sense	ORF region of GGDEF family protein
VC-I-PGQ-7	CGT- <b>GGGG</b> A <b>TACC</b> CG <b>CG</b> <b>GGGG</b> TAGTGAA <b>GGG</b> AC <b>GGG</b> -CTT	73929	73960	Anti-sense	<i>envZ</i> osmolarity sensor protein
VC-II-PGQ-1	CGACA- <b>GGG</b> A <b>GGGG</b> C <b>GGG</b> AA <b>TAGT</b> <b>GGG</b> -CAGAT	89846	89867	Anti-sense	Hypothetical protein (120bp upstream of 5' end) VCA1032
VC-II-PGQ-2	TTTTA- <b>GGGG</b> A <b>GGG</b> TTAG <b>GGT</b> <b>GGGG</b> -TATTT	400058	400076	Anti-sense	Cobryic acid synthase (283bp upstream of 5' end) VCA0727
		508817	508835	Anti-sense	SecA-like protein (313bp upstream of 5' end) VCA0628
VC-II-PGQ-3	TGATC- <b>GGG</b> TCTGTT <b>GGGG</b> ATACAA <b>GGGG</b> GATCC <b>GGG</b> -CTTTT	434411	434443	Anti-sense	ORF region of glucose-1-phosphate adenyltransferase <i>glgC</i> (NP_233086.1)

Five bases are added to the GQ motifs and these sequences along with upstream and downstream bases were used for biophysical studies. The guanines participating in the G-quartet formation are colored red.

at 260nm and 290 nm with a single negative peak at 240 nm (Kyrp et al., 2009; Vorlíčková et al., 2012). Evaluation of VC-PGQs in the presence of different cations through CD spectroscopy revealed that in the presence of  $K^+$  ion, all VC-PGQs other than VC-I-PGQ-2 and VC-I-PGQ-3 showed mixed/hybrid topology while the latter two exhibited parallel topology with a single sharp positive peak at 260nm (Figures 3, 4A). Mixed topology in the VC-I-PGQ-1 and VC-I-PGQ-4 depicted the predominance of anti-parallel topology in the presence of 50 mM  $K^+$  ion and a small population of parallel GQ conformation. Simultaneously in the presence of  $Na^+$  containing buffer, all the VC-PGQs displayed the hybrid topology. Potassium and sodium are largely known cations that stabilize the GQ structures but Magnesium ( $Mg^{2+}$ ), a divalent cation, and Lithium ( $Li^+$ ) show less or neutral effect on the GQ formation which is evident from the CD spectra results obtained for all the VC-PGQs (Bhattacharyya et al., 2016; Figures 3, 4A).

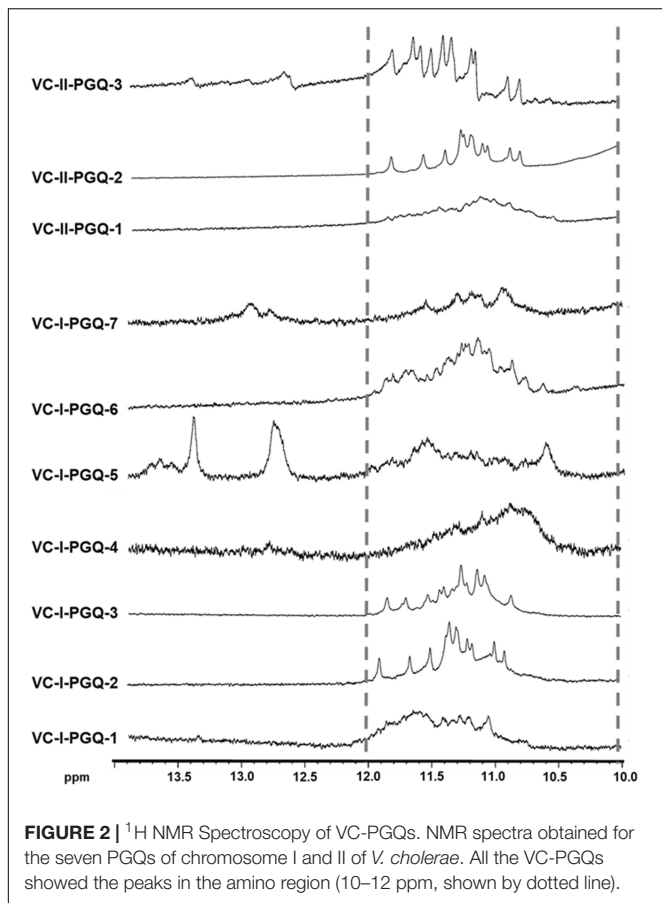
We moved a step ahead on evaluating the CD spectra pattern on the increasing concentration of potassium ion, and we observed a distinct increase in the peak intensity with the increasing concentration of  $K^+$  ion in all the sequences (Figure 4B and Supplementary Figure 2). While in most cases only the peak intensity was increased on the addition of  $K^+$  ions due to higher stabilization of the GQ structures, in the case of VC-I-PGQ-7, a transition from hybrid topology at  $K^+ = 0$  mM to the parallel topology at  $K^+ = 50$  mM (Supplementary Figure 2) and for VC-II-PGQ-2, a conversion from the parallel topology

at  $K^+ = 0$  mM to hybrid topology at  $K^+ = 50$  mM, were noticed (Figure 4B).

On analyzing the CD melting graph of the VC-PGQs, a clear picture of the stability of the GQ structures were obtained. In all the VC-PGQs, the presence of  $K^+$  predominantly revealed better stabilization and a higher melting temperature value ( $T_m$ ) as compared to  $Na^+$ ,  $Li^+$ , and  $Mg^{2+}$  (Figure 5 and Supplementary Figure 2). The  $T_m$  even rises higher when the potassium concentration is increased from 50 to 200 mM in all cases. A distinct difference in  $T_m$  value is observed when cation is absent and when cation is present; in the case of VC-I-PGQ-1, for example, the  $T_m$  value when no cation is present is 41.15 °C while it shoots up to as high as 78.31 °C when the  $K^+$  concentration is 200 mM.

## Molecularity of the VC-PGQs Evaluated Through Electrophoretic Mobility Shift Assay

Gel electrophoresis helps in resolving DNA or RNA oligonucleotides according to their molecular size, and therefore this technique is useful for assessing the molecularity of the GQ structures. G-quadruplex structures are predominantly formed by folding of the interconnection of Guanine residues in one single strand of DNA or RNA forming the intra-molecular GQ structures, but Guanine residues from more than one strand can assemble together to give rise to inter-molecular GQ structures



(Figures 1B,C). Thus, when these GQ sequences are evaluated through gel electrophoresis, their mobility is different from their linear counterparts. The intra-molecular GQ structures forming the compact structures run faster while the inter-molecular GQ structures run slower due to the association of multiple strands in comparison to the non-GQ forming sequences of the same length (Viglaský et al., 2010). Interestingly, all the VC-PGQs of both chromosomes show faster mobility and, therefore, are visibly ahead of their linear counterparts (Supplementary Figure 4). The similar band mobility shift was observed in the positive control, *cKit22*, as compared to its linear counterpart. Thus, all of them form the intra-molecular GQ structures, which are preferable for the stability and propensity of the GQ formation. In certain cases, like VC-I-PGQ-2, VC-I-PGQ-3, and VC-I-PGQ-7, we observed the appearance of two bands, which may be attributed to the formation of two populations of GQ with different conformations (Supplementary Figure 4).

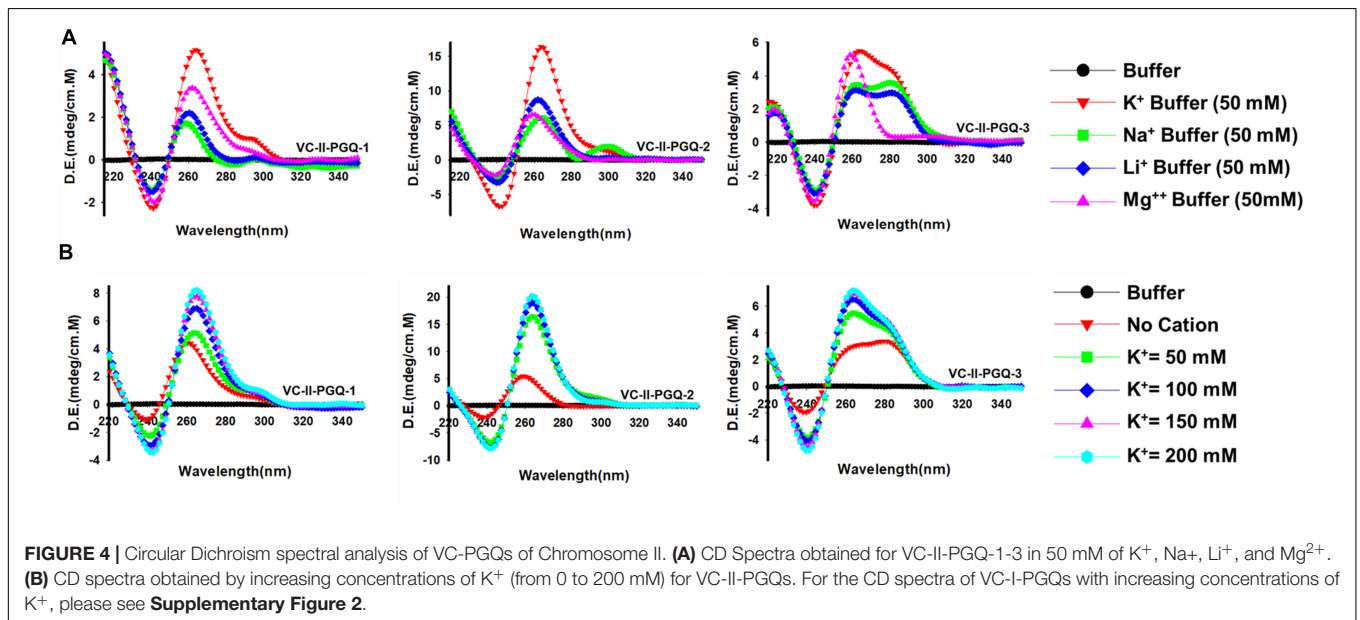
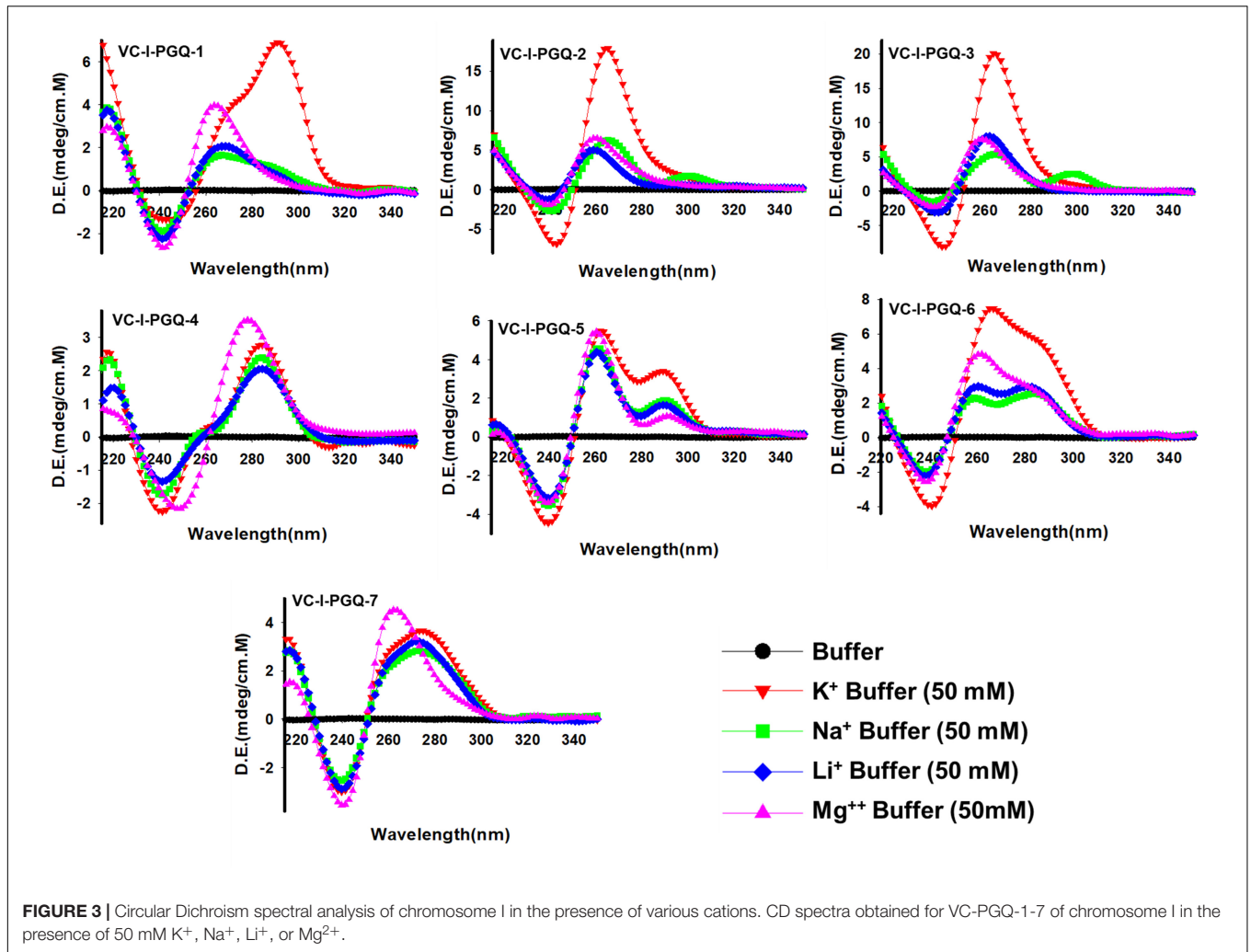
### Strong Binding Affinity of VC-PGQs With the Known GQ Binding Ligand, TMPyP4

GQ structures have been explored as potential drug targets in recent decades for a number of diseases including cancer and viral and bacterial diseases. Parallely, the development of GQ targeting ligands have also increased in pace, with molecules like Quarfloxin (targeting the telomeric GQ structures) reaching the

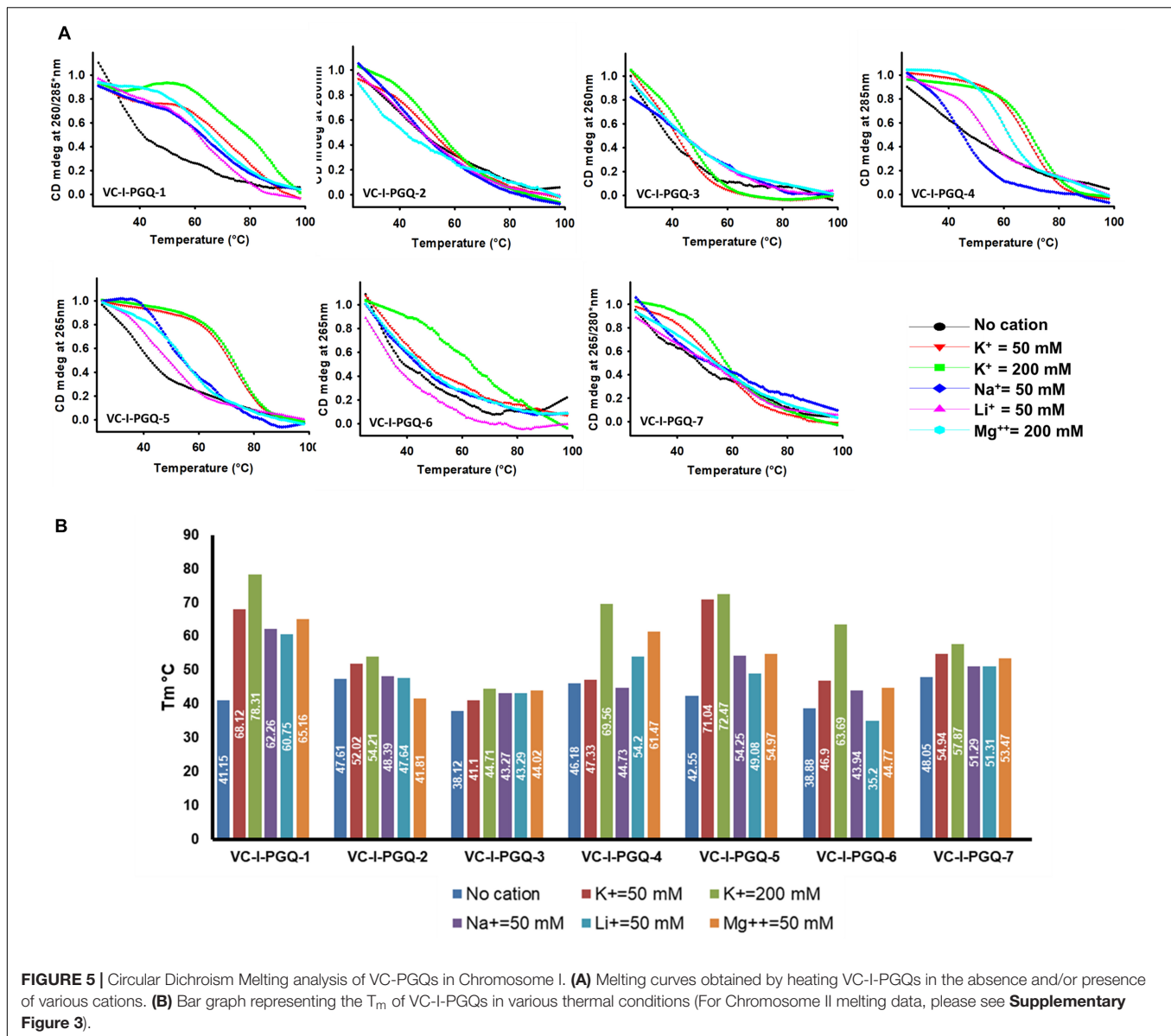
clinical trial phase (Yang and Okamoto, 2010). A number of GQ specific ligands, like TMPyP4, Braco-19, and Pyridostatin, specifically bind to GQ structures over the duplex DNA and therefore are extensively used to check the binding ability of the GQ structures (Ruggiero and Richter, 2018). In our study we have used TMPyP4 molecules to evaluate the propensity of the GQ structures to bind with the GQ binders. TMPyP4 is a cationic porphyrin molecule that can stack and stabilize the GQ structures due to its intrinsic properties (Parkinson et al., 2007). Isothermal Calorimetric Titration technique was utilized to study the binding affinity of the VC-PGQs with the TMPyP4 molecule as it is one of the most reliable, sensitive, and accurate techniques commonly used for studying biomolecular interaction and provides the complete thermodynamic profile in one go (Funke and Weisz, 2019). ITC studies revealed that all the VC-PGQs possessed a strong binding affinity towards TMPyP4 as the enthalpograms show the proper binding curves (Figures 6, 7). The thermodynamic reaction for all the PGQ sequences gave negative enthalpy values that indicate the energetically favorable interaction with the TMPyP4 molecule. VC-II-PGQ-1 had the highest binding affinity towards TMPyP4 with the  $K_d$  value being 0.22  $\mu$ M, followed by VC-I-PGQ-7 having a  $K_d$  value of 0.1  $\mu$ M. These two VC-PGQs were further evaluated by molecular dynamics studies. The other VC-PGQs also provided significant binding affinity values with negative enthalpic energy, as mentioned in Supplementary Table 3.

### Molecular Docking Simulation Revealed the Molecular Interaction of TMPyP4 With the VC-PGQs

Circular Dichroism analysis revealed the strong interaction of VC-PGQs with potassium cation that leads to their stabilization. Similarly, ITC analysis revealed the strong interaction of TMPyP4 with the GQs. To gain further insight into the molecular interaction of potassium and TMPyP4 with the *V. cholerae* G-quadruplexes, VC-I-PGQ-7 and VC-II-PGQ-1 were selected as they showed the highest binding affinity with TMPyP4 in the ITC experiments. First, their structures were modeled and then equilibrated and simulated for 10 ns (Figures 8, 9). The modeled structures of VC-I-PGQ-7 (Figures 8A–E) and VC-II-PGQ-1 (Figures 9A–E) showed the characteristic tetrad formation and tetrad stacking, further affirming the formation of GQ conformations by these PGQs. The global fluctuations analysis revealed the initial fluctuations in the modeled GQs till 200 ps, which thereafter remained constant throughout the production run of 10 ns (Figures 8G, 9G). Radius of gyration, another factor for analyzing the stability of the modeled structures that depicts the global compactness of the system, was also analyzed during the simulation analysis (Figures 8F, 9F). The minimal fluctuations in the  $R_g$  further affirms the stability of the modeled structures. On comparing the global fluctuations in the whole molecule as compared to that of the Guanine residues participating in G-tetrad and stacking formation, it was observed that the tetrad guanine showed lesser fluctuations as compared to the whole macromolecule (Figures 8G, 9G). This supports the strong interaction of







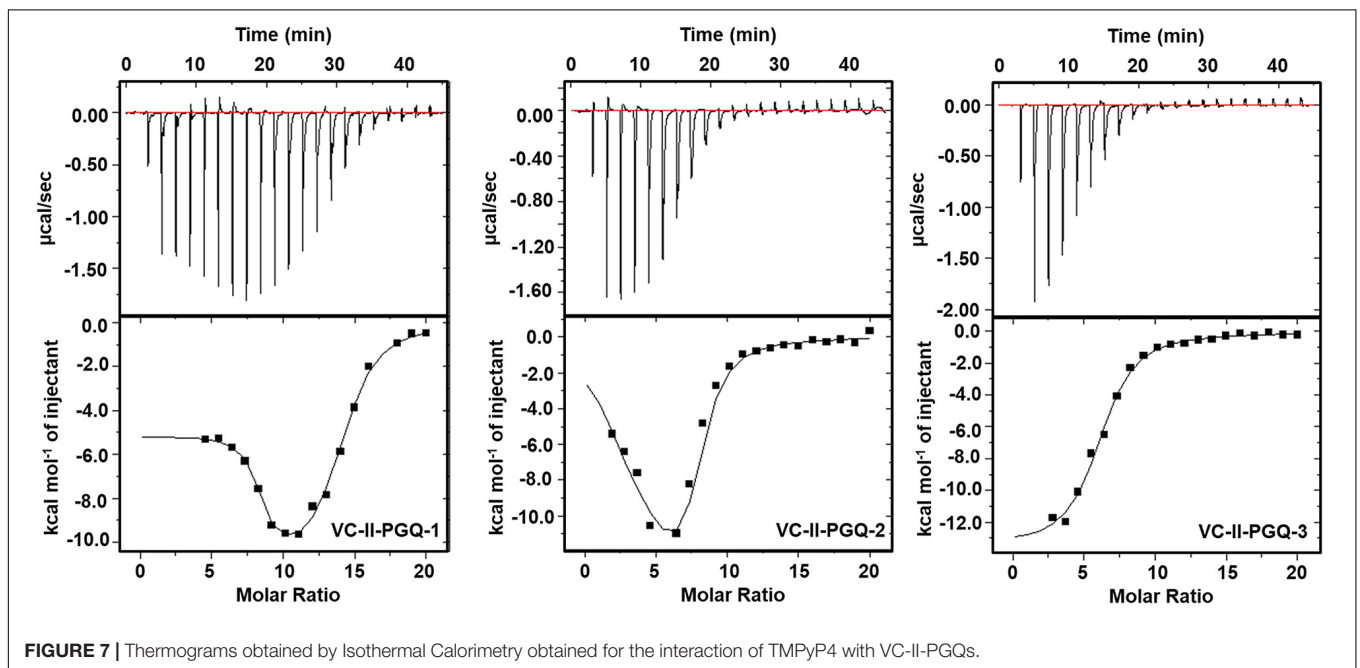
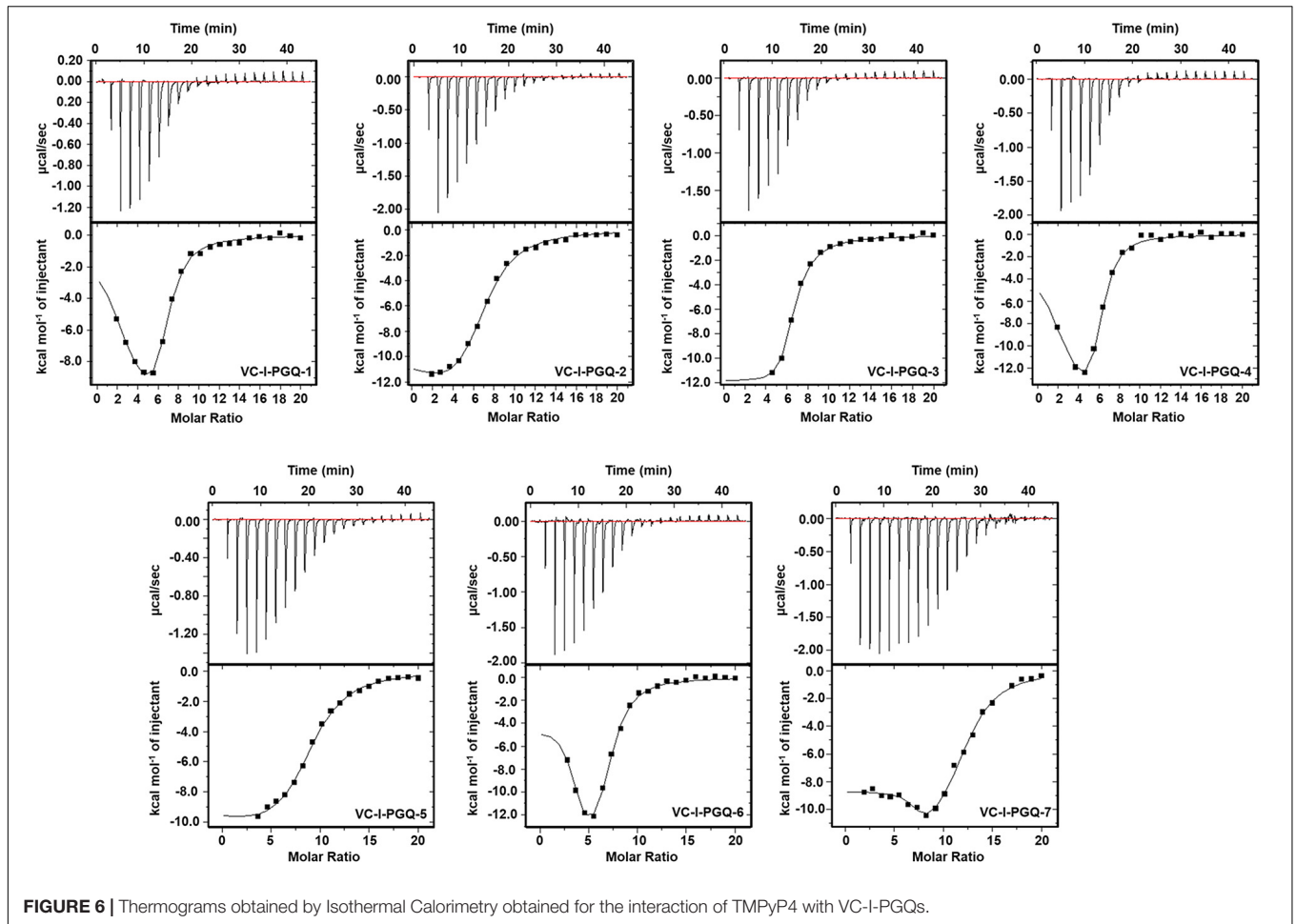
the Guanines of G-tetrad by classical Watson and crick and Hoogsteen hydrogen bonds.

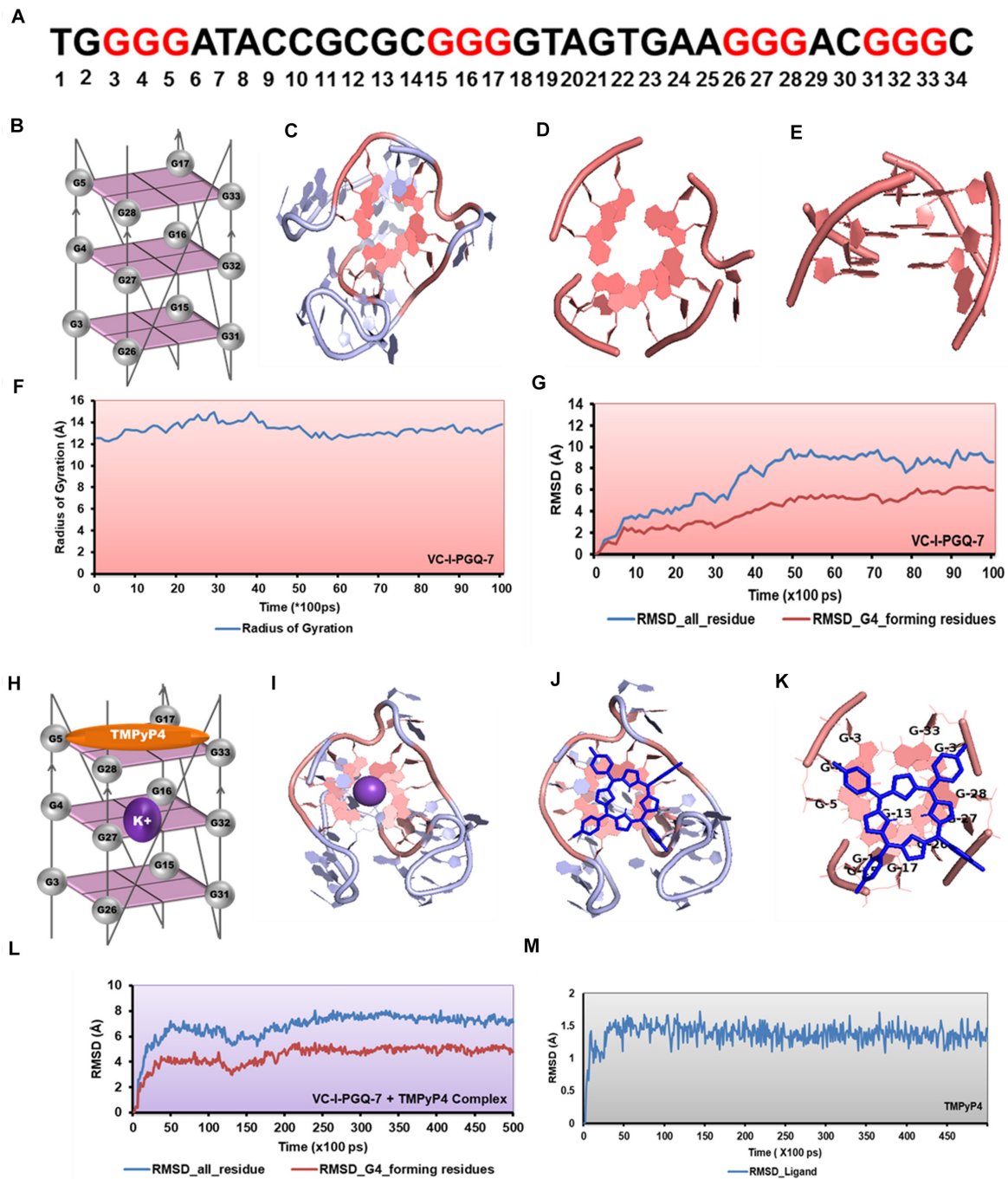
Interaction analysis of K<sup>+</sup> cation with the VC-I-PGQ-7 and VC-II-PGQ-1 revealed the binding of the cation at the middle of the GQ. It is fitted between the 1 and 2 tetrads in both VC-PGQs (**Figures 8H–I**). This affirms the hypothesis of stabilization of GQ conformations by various cations where the cations fit between the tetrad neutralizes the negative charges, thereby leading to stabilization of the PGQs. The docking analysis of VC-I-PGQ-7 with TMPyP4 revealed a favorable interaction with a binding energy of  $-7.43$  kcal/mol and predicted inhibitory constant ( $K_i$ ) of  $3.57$   $\mu$ M. Structural analysis of the docked complex revealed the stacking of TMPyP4 on the bottom tetrad. Molecular simulation analysis revealed the stable VC-I-PGQ-7 – TMPyP4 complex formation with minimal global fluctuations in the GQs and the ligand throughout the 50 ns simulation (**Figures 8L,M**).

Similarly, TMPyP4 binds with VC-II-PGQ-1 with a binding energy of  $-7.44$  kcal/mol and the predicted  $K_i$  of  $3.52$   $\mu$ M, and the ligand stacks upon the upper tetrad and forms a stable complex which was supported by the molecular dynamics simulation analysis (**Figure 9**). The constant RMSD of GQ and TMPyP4 through the trajectory depicted the stable and energy favorable complex formation between the GQs and ligands (**Figures 9L,M**).

## DISCUSSION

*Vibrio cholerae* is an autochthonous inhabitant of aquatic systems, often causing an acute life-threatening diarrheal condition termed cholera (Reidl and Klose, 2002). The bacteria have caused seven pandemics till date, with the first six pandemics

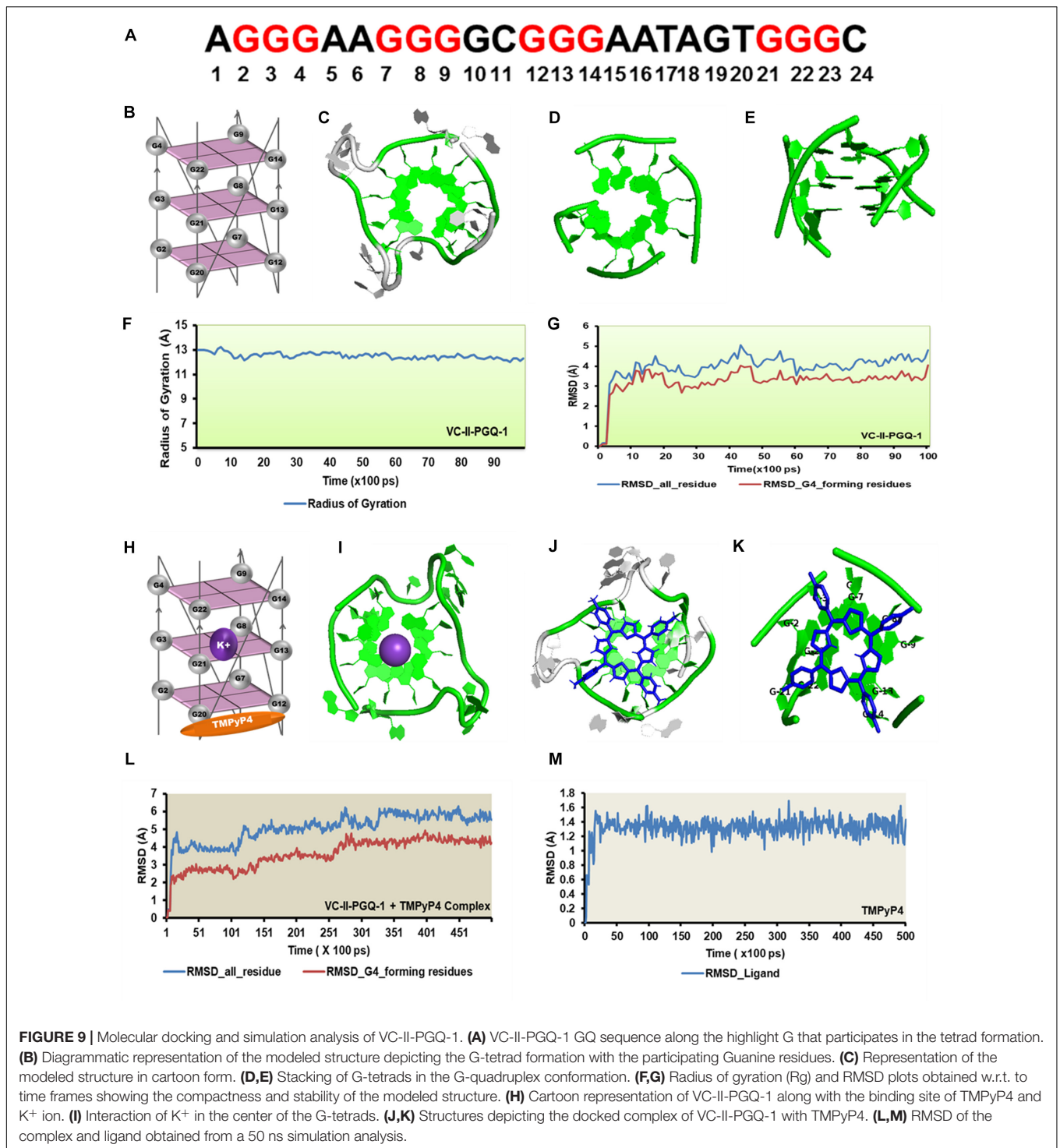




**FIGURE 8 |** Molecular docking and simulation analysis of VC-I-PGQ-7. **(A)** VC-I-PGQ-7 GQ sequence along with the highlight G that participates in the tetrad formation in the modeled structure. **(B)** Diagrammatic representation of the modeled structure depicting the G-tetrad formation with the participating Guanine residues. **(C)** Representation of the modeled structure in cartoon form. **(D,E)** Figures depicting the stacking of G-tetrads in the G-quadruplex conformation. **(F,G)** Radius of gyration (Rg) and RMSD plots obtained w.r.t. to time frames showing the compactness and stability of the modeled structure. **(H)** Cartoon representation of VC-I-PGQ-7 along with the binding site of TMPyP4 and  $K^+$  ion. **(I)** Interaction of  $K^+$  in the center of the G-tetrads. **(J,K)** Structures depicting the docked complex of VC-I-PGQ-7 with TMPyP4. **(L,M)** RMSD of the complex and ligand obtained from a 50 ns simulation analysis.

between 1899 and 1923 caused by the classical biotype of O1 strain, while the latest (which has been ongoing since 1961) is being attributed to the El Tor strain. *V. cholerae* possess a

unique bipartite genome, i.e., that whole genetic information is coded in two chromosomes (Escudero and Mazel, 2017). This distinctive feature probably provides the bacteria with the added



**FIGURE 9 |** Molecular docking and simulation analysis of VC-II-PGQ-1. **(A)** VC-II-PGQ-1 GQ sequence along the highlight G that participates in the tetrad formation. **(B)** Diagrammatic representation of the modeled structure depicting the G-tetrad formation with the participating Guanine residues. **(C)** Representation of the modeled structure in cartoon form. **(D,E)** Stacking of G-tetrads in the G-quadruplex conformation. **(F,G)** Radius of gyration (Rg) and RMSD plots obtained w.r.t. to time frames showing the compactness and stability of the modeled structure. **(H)** Cartoon representation of VC-II-PGQ-1 along with the binding site of TMPyP4 and  $K^+$  ion. **(I)** Interaction of  $K^+$  in the center of the G-tetrads. **(J,K)** Structures depicting the docked complex of VC-II-PGQ-1 with TMPyP4. **(L,M)** RMSD of the complex and ligand obtained from a 50 ns simulation analysis.

advantage of a shorter replication time and, thereby, shorter doubling time. Among the two chromosomes, chromosome I is the larger one being ~3 Mb long and coding for the characterized genes possessing essential roles in cellular functioning including DNA replication, transcription, translation, cellular metabolism, and biosynthesis of the cell wall and those involved the bacterial pathogenesis. In contrast, chromosome II is small

chromosome of ~1 Mb in length with a greater proportion of the uncharacterized genes. Chromosome II is believed to have been acquired from a large plasmid by the *Vibrio* family during evolution (Heidelberg et al., 2000). The virulence of *V. cholerae* is attributed to a number of factors which include cholera toxin, toxin-coregulated pilus (TCP) that helps in intestinal colonization, O-antigen of Lipopolysaccharides, the bacteria's

motility accommodated by the polar flagellum, porins, and some other unrecognized genes (Klose, 2001). The *V. cholerae* genome has been constantly evolving with the changes in the environmental conditions, and the *V. cholerae*'s mobilome is majorly responsible for its adaptation (Banerjee et al., 2014). Horizontal gene transfer, gene-capturing systems termed integrons, and other mobile genetic elements encompassing the transposons, bacteriophages, conjugative plasmids, ICEs, and SXT elements contribute to most the gene adaptation by the bacteria (Mooi and Bik, 1997; Cho et al., 2010). With changes in genetic constituent, the bacteria has also developed drug resistance, and even multi-drug resistant strains have emerged. Along with the mobile genetic elements, efflux pumps and spontaneous chromosomal mutations are a significant cause behind the antibiotic resistance mechanism (Kitaoka et al., 2011a). Therefore, to curb this bacterial infection, a conserved motif is essential to act as a potential drug target, and GQ structure forming sequences has proven to be one of the ideal motifs in this respect. G-quadruplex motifs, depending upon its location, regulates the expression of the genes. PGQ motifs at the upstream/promoter region of the gene are reported to regulate the transcription of the harboring genes and are widely used as therapeutic targets against various pathogenic viruses, bacteria, protozoa, and other human diseases including cancer and neurological disorders (Balasubramanian et al., 2011; Kota et al., 2015; Tosoni et al., 2015; Perrone et al., 2017a; Shankar et al., 2020). Stabilization of PGQ motifs in the open reading frame of a gene has been reported to suppress their expression (Endoh et al., 2013b). For example, GQ motif in the ORF region of human estrogen receptor alpha regulates its translation (Endoh et al., 2013a). Similarly, stabilization of GQ motif causes an unusual  $-1$  ribosomal frameshift in mammalian cells (Endoh and Sugimoto, 2013). G-quadruplex motifs in the C9orf72 are responsible for the blockade of various functional proteins responsible for neurodegenerative diseases and are one of the main causes of Amyotrophic lateral sclerosis and frontotemporal dementia (Wang et al., 2019). In bacterial species, the presence of GQ motifs in the open reading frame have been shown to have an impeding effect on the transcription of the PGQ harboring genes (Mishra et al., 2019a,b). The GQ formed in either strands reduces the translation efficiency by either impeding the transcriptional machinery or translational machinery. The difference in the transcriptional or translational effect varies with the position of the PGQ in the sense or anti-sense strand. The PGQ motifs in the sense strand inhibit the transcriptional machinery while in the anti-sense strand of the ORF region the motifs lead to modulation in translation (Agarwal et al., 2014; Holder and Hartig, 2014; Armas et al., 2017). As these GQs impede the formation of mRNA, various strategies have been used for blocking the formation of these GQ motifs in the essential genes of the cells (Armas et al., 2017). Hence, the subsequent stabilization of the PGQs at the upstream region or in the ORF region of sense or anti-sense strands by the G4 specific ligands can be used as a potential therapeutic approach for inhibiting the expression of the desired PGQ harboring genes. With this hypothesis, the genome of the *V. cholerae* was screened thoroughly for the presence of conserved GQ motifs.

Through our genome-wide search of the recent pandemic-causing *V. cholerae* strain, O1 biovar El Tor str. N16961, we observed ten of the most conserved VC-PGQ forming sequences across both the chromosomes of the bacteria. These PGQs were conserved in all the available strains of *V. Cholerae*, including MDR and XDR strains. The most extensively studied multi-drug resistant strain, O1 serotype str. 2010EL-1786, harbors nine out of ten conserved VC-PGQs (**Supplementary Table 4**). In chromosome I, two of the sequences, VC-I-PGQ-1 and VC-II-PGQ-3, were found to be conserved in the ORF of the methyl-accepting chemotaxis proteins (MCPs). Chemotaxis is an essential property of many bacteria, including *V. cholerae*, possibly contributing to the virulence and survival factor of the bacteria. This phenomenon of bacterial motility in response to environmental stimuli involves a cascade of signal transduction processes in which the MCPs are an integral part. The membrane scanning MCPs receives the signal for the chemoeffectors and thereafter undergo a conformational change due to binding of an attractant or leaving of a repellent. This change in the MCPs is detected by the CheA–CheW complex and, consequently, autophosphorylates the CheA protein kinase, and the subsequent phosphate is passed on to the response regulator, CheY. The phosphorylated CheY ultimately controls the flagellar motor that results in the reorientation of the bacterial cell movement (Boin et al., 2004). While VC-I-PGQ-4 is harbored in the ORF of the important gene, *rtxA* codes for another exotoxin with leukotoxic, hemolytic, and leukocyte-stimulating activities (Lin et al., 1999). ORF regions of two other important genes coding GGDEF family protein, acting as diguanylate cyclase involved in the nucleotide cyclization process and osmolality sensor protein, possess the VC-I-PGQ-6 and VC-I-PGQ-7 sequences, respectively. Certain VC-PGQs were also found in the promoter regions of certain genes coding for Orotate phosphoribosyltransferase, the essential enzyme for pyrimidine biosynthesis, OgbE GTPase, and cobyric acid synthase (**Table 1**). The VC-I-PGQ-3, VC-I-PGQ-5, VC-II-PGQ-1, and VC-II-PGQ-2 located in the gene promoters may have specialized roles to play in the bacterial gene modulation which need to be analyzed further. Additionally, VC-II-PGQ-2 is present at multiple locations in *V. cholerae* and is also conserved in various human pathogenic bacterial species including *E. coli*, *K. pneumoniae*, *Yersinia pestis*, *Shigella flexneri*, *Salmonella enterica*, *Pseudomonas aeruginosa*, and *M. tuberculosis* (Kaplan et al., 2016). The presence of this PGQ motif in human pathogens strengthens the plausibility of its essential role in the host pathogenesis.

All the biophysical techniques, including NMR, CD, and EMSA, portrayed that the predicted VC-PGQs have the potential to form stable GQ structures and even have the tendency to bind specifically to the GQ binding ligand, TMPyP4, thereby opening up new hopes for the treatment of this prolonged bacterial infection. The molecular dynamic studies helped us to get a better insight regarding the GQ orientation and their interaction pattern with TMPyP4, which also depicted the stability of the GQ-ligand complexes. The present study provides the preliminary base as of how non-canonical secondary nucleic acid structures can bear crucial gene modulatory roles

in *V. cholerae* and should be further validated in the real bacterial paradigm. Therefore, other than the conventional drug targets of *V. cholerae*, such as the signaling and transport proteins, toxins protein, ion channel regulators, etc., conserved GQ structure in the *V. cholerae* genome could rise as an alternate therapeutic solution.

## DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

AK designed the data conceptualization and methodology. US and NJ performed the *in silico* and *in vitro* experiments. AK, US, PM, and NJ collectively wrote the manuscript. PK, TKS, and AK did the critical review and editing. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Adrian, M., Heddi, B., and Phan, A. T. (2012). NMR spectroscopy of G-quadruplexes. *Methods* 57, 11–24. doi: 10.1016/j.ymeth.2012.05.003
- Agarwal, T., Roy, S., Kumar, S., Chakraborty, T. K., and Maiti, S. (2014). In the sense of transcription regulation by G-quadruplexes: asymmetric effects in sense and antisense strands. *Biochemistry* 53, 3711–3718. doi: 10.1021/bi401451q
- Armas, P., David, A., and Calcaterra, N. B. (2017). Transcriptional control by G-quadruplexes: In vivo roles and perspectives for specific intervention. *Transcription* 8, 21–25. doi: 10.1080/21541264.2016.1243505
- Asamitsu, S., Bando, T., and Sugiyama, H. (2018). Ligand design to acquire specificity to intended G-quadruplex structures. *Chem. A Eur. J.* 25, 417–430. doi: 10.1002/chem.201802691
- Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011). Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Dis.* 10, 261–275. doi: 10.1038/nrd3428
- Banerjee, R., Das, B., Balakrishna Nair, G., and Basak, S. (2014). Dynamics in genome evolution of *Vibrio cholerae*. *Infect. Genet. Evol.* 23, 32–41. doi: 10.1016/j.meegid.2014.01.006
- Bhartiya, D., Chawla, V., Ghosh, S., Shankar, R., and Kumar, N. (2016). Genome-wide regulatory dynamics of G-quadruplexes in human malaria parasite *Plasmodium falciparum*. *Genomics* 108, 224–231. doi: 10.1016/j.ygeno.2016.10.004
- Bhattacharyya, D., Mirihana Arachchilage, G., and Basu, S. (2016). Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* 4:38. doi: 10.3389/fchem.2016.00038
- Biswas, B., Kandpal, M., Jauhari, U. K., and Vivekanandan, P. (2016). Genome-wide analysis of G-quadruplexes in herpesvirus genomes. *BMC Genom.* 17:949. doi: 10.1186/s12864-016-3282-1
- Boin, M. A., Austin, M. J., and Häse, C. C. (2004). Chemotaxis in *Vibrio cholerae*. *FEMS Microbiol. Lett.* 239, 1–8. doi: 10.1016/j.femsle.2004.08.039
- Bugaut, A., and Balasubramanian, S. (2012). 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucl. Acid. Res.* 40, 4727–4741. doi: 10.1093/nar/gks068
- Cahoon, L. A., and Seifert, H. S. (2009). An Alternative DNA Structure is Necessary for Pilin Antigenic Variation in *Neisseria gonorrhoeae*. *Science* 325, 764–767. doi: 10.1126/science.1175653
- Cho, Y.-J., Yi, H., Lee, J. H., Kim, D. W., and Chun, J. (2010). Genomic evolution of *Vibrio cholerae*. *Curr. Opin. Microbiol.* 13, 646–651. doi: 10.1016/j.mib.2010.08.007
- Cogoi, S., and Xodo, L. E. (2006). G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucl. Acid. Res.* 34, 2536–2549. doi: 10.1093/nar/gkl286
- Dai, J., Chen, D., Jones, R. A., Hurley, L. H., and Yang, D. (2006a). NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucl. Acid. Res.* 34, 5133–5144. doi: 10.1093/nar/gkl610
- Dai, J., Dexheimer, T. S., Chen, D., Carver, M., Ambrus, A., Jones, R. A., et al. (2006b). An Intramolecular G-Quadruplex Structure with Mixed Parallel/Antiparallel G-Strands Formed in the Human BCL-2 Promoter Region in Solution. *J. Am. Chem. Soc.* 128, 1096–1098. doi: 10.1021/ja055636a
- Das, B., Verma, J., Kumar, P., Ghosh, A., and Ramamurthy, T. (2020). Antibiotic resistance in *Vibrio cholerae*: Understanding the ecology of resistance genes and mechanisms. *Vaccine* 38, A83–A92. doi: 10.1016/j.vaccine.2019.06.031
- Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013a). Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor  $\alpha$ . *Nucl. Acid. Res.* 41, 6222–6231. doi: 10.1093/nar/gkt286
- Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013b). Suppression of Gene Expression by G-Quadruplexes in Open Reading Frames Depends on G-Quadruplex Stability. *Angew. Chem. Int. Edit.* 52, 5522–5526. doi: 10.1002/anie.201300058
- Endoh, T., and Sugimoto, N. (2013). Unusual  $-1$  Ribosomal Frameshift Caused by Stable RNA G-Quadruplex in Open Reading Frame. *Analyt. Chem.* 85, 11435–11439. doi: 10.1021/ac402497x
- Escudero, J. A., and Mazel, D. (2017). Genomic Plasticity of *Vibrio cholerae*. *Int. Microbiol. Offic. J. Span. Soc. Microbiol.* 20, 138–148.
- Feigon, J., Koshlap, K. M., and Smith, F. W. (1995). 1H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol* 261, 225–255. doi: 10.1016/S0076-6879(95)61012-X
- Fleming, A. M., Ding, Y., Alenko, A., and Burrows, C. J. (2016). Zika Virus Genomic RNA Possesses Conserved G-Quadruplexes Characteristic of the Flaviviridae Family. *ACS Infect. Dis.* 2, 674–681. doi: 10.1021/acinfecdis.6b00109
- Funke, A., and Weisz, K. (2019). “Revealing the Energetics of Ligand-Quadruplex Interactions Using Isothermal Titration Calorimetry,” in *G-Quadruplex Nucleic Acids: Methods and Protocols*, eds Edn, eds D. Yang and C. Lin (New York, NY: Springer), 45–61. doi: 10.1007/978-1-4939-9666-7\_3
- Giacani, L., Brandt, S. L., Puray-Chavez, M., Reid, T. B., Godornes, C., Molini, B. J., et al. (2012). Comparative Investigation of the Genomic Regions Involved in Antigenic Variation of the TprK Antigen among Treponemal Species, Subspecies, and Strains. *J. Bacteriol.* 194, 4208–4225. doi: 10.1128/JB.00863-12

## FUNDING

The work was supported by Science and Engineering Research Board, Department of Science and Technology, Govt. of India (EMR/2016/3897 to AK), (EMR/2016/003547 to PK), and Department of Biotechnology, Govt. of India (BT/PR20319/BBE/117/189/2016 to PK).

## ACKNOWLEDGMENTS

We are thankful to the Sophisticated Instrumentation Facility at IIT Indore for NMR and CD experiments. NJ to CSIR, New Delhi and US and PM to MHRD, New Delhi for their individual fellowships.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00935/full#supplementary-material>

- Hall, R. H. (2018). Curbing cholera. *Sci. Transl. Med.* 10:eaat9483. doi: 10.1126/scitranslmed.aat9483
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477–483. doi: 10.1038/35020000
- Hema, M., Princy, S. A., Sridharan, V., Vinoth, P., Balamurugan, P., and Sumana, M. (2016). Synergistic activity of quorum sensing inhibitor, pyrazine-2-carboxylic acid and antibiotics against multi-drug resistant *V. cholerae*. *RSC Adv.* 6, 45938–45946. doi: 10.1039/C6RA04705J
- Hershman, S. G., Chen, Q., Lee, J. Y., Kozak, M. L., Yue, P., Wang, L.-S., et al. (2008). Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 36, 144–156. doi: 10.1093/nar/gkm986
- 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
- Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., De Groot, B. L., et al. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73. doi: 10.1038/nmeth.4067
- Hun Yoon, S., and Waters, C. M. (2019). *Vibrio cholerae*. *Trends Microbiol.* 27, 806–807. doi: 10.1016/j.tim.2019.03.005
- Huppert, J. L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucl. Acids Res.* 35, 406–413. doi: 10.1093/nar/gkl1057
- Islam, M. T., Alam, M., and Boucher, Y. (2017). Emergence, ecology and dispersal of the pandemic generating *Vibrio cholerae* lineage. *Int. Microbiol. Offic. J. Span. Soc. Microbiol.* 20, 106–115.
- Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29, 1859–1865. doi: 10.1002/jcc.20945
- Kaplan, O. I., Berber, B., Hekim, N., and Doluca, O. (2016). G-quadruplex prediction in *E. coli* genome reveals a conserved putative G-quadruplex-Hairpin-Duplex switch. *NAR* 44, 9083–9095. doi: 10.1093/nar/gkw769
- Kikin, O., D'antonio, L., and Bagga, P. S. (2006). QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucl. Acids Res.* 34, W676–W682. doi: 10.1093/nar/gkl253
- Kim, N. (2019). The Interplay between G-quadruplex and Transcription. *Curr. Med. Chem.* 26, 2898–2917. doi: 10.2174/0929867325666171229132619
- Kitaoka, M., Miyata, S. T., Unterweger, D., and Pukatzki, S. (2011a). Antibiotic resistance mechanisms of *Vibrio cholerae*. *J. Med. Microbiol.* 60, 397–407. doi: 10.1099/jmm.0.023051-0
- Klose, K. E. (2001). Regulation of virulence in *Vibrio cholerae*. *Int. J. Med. Microbiol.* 291, 81–88. doi: 10.1078/1438-4221-00104
- Kong, J.-N., Zhang, C., Zhu, Y.-C., Zhong, K., Wang, J., Chu, B.-B., et al. (2018). Identification and characterization of G-quadruplex formation within the EP0 promoter of pseudorabies virus. *Sci. Rep.* 8:14029. doi: 10.1038/s41598-018-32222-7
- Kota, S., Dhamodharan, V., Pradeepkumar, P., and Misra, H. S. (2015). G-quadruplex forming structural motifs in the genome of *Deinococcus radiodurans* and their regulatory roles in promoter functions. *Appl. Microbiol. Biotechnol.* 99, 9761–9769. doi: 10.1007/s00253-015-6808-6
- Kumar, P. A., Patterson, J., and Karpagam, P. (2009). Multiple antibiotic resistance profiles of *Vibrio cholerae* non-O1 and non-O139. *Jpn. J. Infect. Dis.* 62, 230–232.
- Kypr, J., Kejnovská, I., Renčíuk, D., and Vorlíčková, M. (2009). Circular dichroism and conformational polymorphism of DNA. *Nucl. Acids Res.* 37, 1713–1725. doi: 10.1093/nar/gkp026
- Lee, D. S. M., Ghanem, L. R., and Barash, Y. (2020). Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. *Nat. Commun.* 11:527. doi: 10.1038/s41467-020-14404-y
- Lin, W., Fuller, K. J., Clayton, R., Sexton, J. A., Rogers, M. B., Calia, K. E., et al. (1999). Identification of a *Vibrio cholerae* RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. *Proc. Nat. Acad. Sci.* 96:1071. doi: 10.1073/pnas.96.3.1071
- Lowrence, R. C., Subramaniapillai, S. G., Ulaganathan, V., and Nagarajan, S. (2019). Tackling drug resistance with efflux pump inhibitors: from bacteria to cancerous cells. *Critic. Rev. Microbiol.* 45, 334–353. doi: 10.1080/1040841X.2019.1607248
- Majee, P., Kumar Mishra, S., Pandya, N., Shankar, U., Pasadi, S., Muniyappa, K., et al. (2020a). Identification and characterization of two conserved G-quadruplex forming motifs in the Nipah virus genome and their interaction with G-quadruplex specific ligands. *Sci. Rep.* 10:1477. doi: 10.1038/s41598-020-58406-8
- Majee, P., Shankar, U., Pasadi, S., Muniyappa, K., Nayak, D., and Kumar, A. (2020b). Genome-wide analysis reveals a regulatory role for G-quadruplexes during Adenovirus multiplication. *Virus Res.* 283:197960. doi: 10.1016/j.virusres.2020.197960
- Melo, M. C. R., and Bernardi, R. C. (2018). NAMD goes quantum: an integrative suite for hybrid simulations. *Nat. Methods* 15, 351–354. doi: 10.1038/nmeth.4638
- Mishra, S. K., Jain, N., Shankar, U., Tawani, A., Sharma, T. K., and Kumar, A. (2019a). Characterization of highly conserved G-quadruplex motifs as potential drug targets in *Streptococcus pneumoniae*. *Scient. Rep.* 9:1791. doi: 10.1038/s41598-018-38400-x
- Mishra, S. K., Shankar, U., Jain, N., Sikri, K., Tyagi, J. S., Sharma, T. K., et al. (2019b). Characterization of G-Quadruplex Motifs in *espB*, *espK*, and *cyp51* Genes of *Mycobacterium tuberculosis* as Potential Drug Targets. *Mole. Ther. Nucl. Acid.* 16, 698–706. doi: 10.1016/j.omtn.2019.04.022
- Mishra, S. K., Tawani, A., Mishra, A., and Kumar, A. (2016). G4IPDB: A database for G-quadruplex structure forming nucleic acid interacting proteins. *Scient. Rep.* 6, 38144–38144. doi: 10.1038/srep38144
- Mooi, F. R., and Bik, E. M. (1997). The evolution of epidemic *Vibrio cholerae* strains. *Trends Microbiol.* 5, 161–165. doi: 10.1016/S0966-842X(96)10086-X
- Parkinson, G. N., Ghosh, R., and Neidle, S. J. B. (2007). Structural basis for binding of porphyrin to human telomeres. *Biochemistry* 46, 2390–2397. doi: 10.1021/bi062244n
- Patel, D. J., Phan, A. T., and Kuryavyi, V. (2007). Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucl. Acids Res.* 35, 7429–7455. doi: 10.1093/nar/gkm711
- Perrone, R., Lavezzo, E., Palù, G., and Richter, S. N. (2017a). Conserved presence of G-quadruplex forming sequences in the Long Terminal Repeat Promoter of Lentiviruses. *Scien. Rep.* 7, 1–11. doi: 10.1038/s41598-017-02291-1
- Perrone, R., Lavezzo, E., Riello, E., Manganello, R., Palù, G., Toppo, S., et al. (2017b). Mapping and characterization of G-quadruplexes in *Mycobacterium tuberculosis* gene promoter regions. *Scien. Rep.* 7:5743. doi: 10.1038/s41598-017-05867-z
- Reidl, J., and Klose, K. E. (2002). *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiol. Rev.* 26, 125–139. doi: 10.1111/j.1574-6976.2002.tb00605.x
- Rigo, R., Palumbo, M., and Sissi, C. (2017). G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochim. et Biophys. Acta* 1861, 1399–1413. doi: 10.1016/j.bbagen.2016.12.024
- Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theor. Comput.* 9, 3084–3095. doi: 10.1021/ct400341p
- Rother, M., Rother, K., Puton, T., and Bujnicki, J. M. (2011). ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucl. Acids Res.* 39, 4007–4022. doi: 10.1093/nar/gkq1320
- Rouleau, S., Glouzon, J.-P. S., Brumwell, A., Bisailon, M., and Perreault, J.-P. (2017). 3' UTR G-quadruplexes regulate miRNA binding. *RNA* 23, 1172–1179. doi: 10.1261/rna.060962.117
- Ruggiero, E., and Richter, S. N. (2018). G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucl. Acids Res.* 46, 3270–3283. doi: 10.1093/nar/gky187
- Saranathan, N., and Vivekanandan, P. (2018). *G-Quadruplexes: More Than Just a Kink in Microbial Genomes*. Cambridge: Cell Press. doi: 10.1016/j.tim.2018.08.011
- Shankar, U., Jain, N., Mishra, S. K., Sharma, T. K., and Kumar, A. (2020). Conserved G-Quadruplex Motifs in Gene Promoter Region Reveals a Novel Therapeutic Approach to Target Multi-Drug Resistance *Klebsiella pneumoniae*. *Front. Microbiol.* 11:1269. doi: 10.3389/fmicb.2020.01269
- Shimada, T., Arakawa, E., Itoh, K., Okitsu, T., Matsushima, A., Asai, Y., et al. (1994). Extended serotyping scheme for *Vibrio cholerae*. *Curr. Microbiol.* 28, 175–178. doi: 10.1007/BF01571061

- Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Nat. Acad. Sci. U S A* 99, 11593–11598. doi: 10.1073/pnas.182256799
- Tlučková, K., Marušič, M., Tóthová, P., Bauer, L., Šket, P., Plavec, J., et al. (2013). Human Papillomavirus G-Quadruplexes. *Biochemistry* 52, 7207–7216. doi: 10.1021/bi400897g
- Tosoni, E., Frasson, I., Scalabrin, M., Perrone, R., Butovskaya, E., Nadai, M., et al. (2015). Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucl. Acid. Res.* 43, 8884–8897. doi: 10.1093/nar/gkv897
- Verma, J., Bag, S., Saha, B., Kumar, P., Ghosh, T. S., Dayal, M., et al. (2019). Genomic plasticity associated with antimicrobial resistance in em *Vibrio cholerae* em. *Proc. Nat. Acad. Sci.* 116, 6226–6231. doi: 10.1073/pnas.1900141116
- Víglaský, V., Bauer, L., and Tlučková, K. (2010). Structural Features of Intra- and Intermolecular G-Quadruplexes Derived from Telomeric Repeats. *Biochemistry* 49, 2110–2120. doi: 10.1021/bi902099u
- Vorlíčková, M., Kejniovská, I., Sagí, J., Renčiuk, D., Bednářová, K., Motlová, J., et al. (2012). Circular dichroism and guanine quadruplexes. *Methods* 57, 64–75. doi: 10.1016/j.ymeth.2012.03.011
- Vy Thi Le, T., Han, S., Chae, J., and Park, H. J. (2012). G-quadruplex binding ligands: from naturally occurring to rationally designed molecules. *Curr. Pharm. Des.* 18, 1948–1972. doi: 10.2174/138161212799958431
- Walia, R., and Chaconas, G. (2013). Suggested Role for G4 DNA in Recombinational Switching at the Antigenic Variation Locus of the Lyme Disease Spirochete. *PLoS One* 8:e57792. doi: 10.1371/journal.pone.0057792
- Wang, S.-R., Min, Y.-Q., Wang, J.-Q., Liu, C.-X., Fu, B.-S., Wu, F., et al. (2016a). A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti-hepatitis C target. *Sci. Adv.* 2:e1501535. doi: 10.1126/sciadv.1501535
- Wang, S. R., Zhang, Q. Y., Wang, J. Q., Ge, X. Y., Song, Y. Y., Wang, Y. F., et al. (2016b). Chemical Targeting of a G-Quadruplex RNA in the Ebola Virus L Gene. *Cell Chem. Biol.* 23, 1113–1122. doi: 10.1016/j.chembiol.2016.07.019
- Wang, X., Goodrich, K. J., Conlon, E. G., Gao, J., Erbse, A. H., Manley, J. L., et al. (2019). C9orf72 and triplet repeat disorder RNAs: G-quadruplex formation, binding to PRC2 and implications for disease mechanisms. *RNA* 25, 935–947. doi: 10.1261/rna.071191.119
- Yang, D., and Hurley, L. H. (2006). Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucl. Nucleo. Nucl. Acid.* 25, 951–968. doi: 10.1080/15257770600809913
- Yang, D., and Okamoto, K. (2010). Structural insights into G-quadruplexes: towards new anticancer drugs. *Fut. Med. Chem.* 2, 619–646. doi: 10.4155/fmc.09.172
- Zhang, Z., He, X., and Yuan, G. (2011). Formation and recognition of G-quadruplex relevant for pilin antigenic variation in *Neisseria gonorrhoeae*. *Canad. J. Chem.* 90, 34–38. doi: 10.1139/v11-092

**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.

Copyright © 2020 Shankar, Jain, Majee, Kodgire, Sharma and Kumar. 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.