Chinese Journal of Natural Medicines

Structural basis of the major *TMPRSS2* promoter G-quadruplex and its complex with berberine

Zhiyu Tang, Yuting Bian, Shangran Li, Zhiyuan Chen, Yingying Wang, Yongqiang Zhang, Yipu Li, Yushuang Liu, Minghua Yang, Lingyi Kong, Kaibo Wang

Citation: Zhiyu Tang, Yuting Bian, Shangran Li, Zhiyuan Chen, Yingying Wang, Yongqiang Zhang, Yipu Li, Yushuang Liu, Minghua Yang, Lingyi Kong, Kaibo Wang, Structural basis of the major *TMPRSS2* promoter G-quadruplex and its complex with berberine, Chinese Journal of Natural Medicines, 2026, 24(2), 237-246. doi: 10.1016/S1875-5364(26)61094-1.

View online: https://doi.org/10.1016/S1875-5364(26)61094-1

Chinese Journal of Natural Medicines 24 (2026) 237-246

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Original article

Structural basis of the major *TMPRSS2* promoter G-quadruplex and its complex with berberine

Zhiyu Tang, Yuting Bian, Shangran Li, Zhiyuan Chen, Yingying Wang, Yongqiang Zhang, Yipu Li, Yushuang Liu, Minghua Yang, Lingyi Kong*, Kaibo Wang*

State Key Laboratory of Natural Medicines and Jiangsu Key Laboratory of Bioactive Natural Product Research, China Pharmaceutical University, Nanjing 210009, China

ARTICLE INFO

Article history:
Received 19 January 2025
Revised 27 March 2025
Accepted 4 June 2025
Available online 20 February 2026

Keywords:
DNA G-quadruplex
NMR solution structure
TMPRSS2
Berberine
Antivirus

ABSTRACT

TMPRSS2 plays a crucial role in facilitating the entry of both the influenza virus and the SARS-CoV-2 coronavirus into host cells. Recent studies have identified a guanine-rich sequence in the proximal promoter region of the *TMPRSS2* gene, which can form G-quadruplex structures (*TMPRSS2*-G4s) that are potential targets for small molecules to inhibit *TMPRSS2* expression. However, the structural details of the major *TMPRSS2*-G4 and its complex with small molecules remain unknown, hindering the development of antiviral drugs targeting *TMPRSS2*-G-quadruplexes (G4s). This study reports the first high-resolution nuclear magnetic resonance (NMR) solution structure of the major *TMPRSS2*-G4, which consists of a three-tetrad core parallel-stranded G4. Both 3′ and 5′ flanking regions form well-defined capping structures stabilized by multiple hydrogen bonds. Importantly, we found that berberine, an antiviral alkaloid, strongly binds to the major *TMPRSS2*-G4 and determined its binding complex structure with *TMPRSS2*-G4 at a 2∶1 binding stoichiometry. Each berberine molecule recruits an adjacent flanking residue, forming a coplanar structure superimposed on two outer G-tetrads. Moreover, we demonstrated that the major *TMPRSS2*-G4 can stably form within a longer deoxyribonucleic acid (DNA) context and be targeted by small molecules to inhibit DNA polymerase activity. Overall, this study provides structural insights into the recognition mechanism of small molecules by the major *TMPRSS2*-G4 and may facilitate the development of novel antiviral therapeutics targeting *TMPRSS2*-G4.

1. Introduction

Serine proteases, comprising approximately 75% of all human proteases, are crucial in regulating various physiological processes by processing proteins and peptides [1, 2]. Among them, type II transmembrane serine proteases (TTSPs) are trypsin-like membrane serine proteases that are involved in maintaining epithelial homeostasis, degrading extracellular matrix, activating growth factors, and initiating proteolytic cascades through cleavage of cell membrane proteins [3, 4]. TMPRSS2 is a TTSP belonging to the TTSP family [5]. Previous studies have primarily focused on the correlation between the *TMPRSS2:ERG* rearrangement and prostate cancer [6, 7]. However, it is important to note that TMPRSS2 facilitates the entry of influenza viruses and coronaviruses into host cells. By activating the spike (*S*) protein of coronaviruses or the hemagglutinin (HA) of influenza viruses, TMPRSS2 enables viral fusion and release of the viral genome into the cytoplasm [8, 9].

Both influenza viruses and coronaviruses can cause a range of illnesses in humans, varying from mild symptoms to severe disease [10, 11]. Disease severity depends on specific viral proteins expressed [12]. The HA protein and *S* glycoprotein mediate virus entry into host cells, while other proteins participate in packaging, assembly, and budding of new virions [13, 14]. Upon attachment to the host cell and interaction with entry receptors, viral proteins undergo proteolytic cleavage *via* TMPRSS2 or endosomal-lysosomal cysteine proteases such as cathepsin L or B [15, 16]. Activation by TMPRSS2 is associated with cell surface entry, representing the early entry pathway [9]. In contrast, activation *via* cathepsin L or B occurs within endosomes and requires acidification for membrane fusion, corresponding to the late entry pathway [9]. Notably, only TMPRSS2-mediated entry is essential for viral spread and pathogenesis within the host, whereas cathepsin L- or B-mediated entry is dispensable [16, 17]. For example, *TMPRSS2* is highly expressed in respiratory tract epithelial cells—primary targets of these viruses—while cathepsin L is expressed at low levels [9, 16]. Camostat mesylate, a clinically approved protease inhibitor, inhibits TMPRSS2 activity and reduces viral invasion into host cells [18]. However, this inhibitor has been associated with side effects such as cerebral hemorrhage, allergic reactions, cardiac arrest, and shock [19]. Therefore, exploring alternative mechanisms to suppress *TMPRSS2* expression is critical. Encouragingly, recent studies have revealed that the promoter region of the *TMPRSS2* gene contains several consecutive guanine-rich tracts capable of forming G-quadruplex (G4) structures [20]. These G4 structures can be stabilized by small molecules to downregulate *TM-*

* Corresponding author.
E-mail addresses: cpu_lykong@126.com (L. Kong); kbwang@cpu.edu.cn (K. Wang)

https://doi.org/10.1016/S1875-5364(26)61094-1

PRSS2 gene expression and thereby block viral entry into host cells [20, 21]. However, the detailed structure of the major *TMPRSS2* promoter G4 (*TMPRSS2*-G4) and its interaction with small molecules remain unknown, impeding the development of antiviral agents targeting *TMPRSS2*-G4.

G-quadruplexes (G4s) are noncanonical nucleic acid secondary structures [22, 23] formed through Hoogsteen hydrogen bonding between four guanines arranged in a square planar configuration (G-tetrads), which stack upon one another and are stabilized by monovalent cations such as K^+ and Na^+ [24, 25]. These structures are predominantly found in gene promoters and telomeres, where they regulate gene expression and maintain genomic stability [26, 27]. Additionally, G4s have been identified in viral genomes, including those of human immunodeficiency virus (HIV) [28], Epstein-Barr virus [29], and coronaviruses [22]. In these contexts, G4s play essential roles in various stages of the viral life cycle [21, 30].

Natural products have long served as valuable sources for developing antiviral and anticancer agents due to their structural diversity and potent pharmacological activities [31-34]. Among them, berberine and its derivatives have attracted considerable research interest owing to their broad spectrum of biological activities, including antiviral, antimicrobial, anti-inflammatory, and anticancer effects [35-38]. Notably, berberine exhibits antiviral activity [38, 39], yet the underlying molecular mechanisms remain poorly understood [40, 41].

In this study, using dimethyl sulfate (DMS) footprinting, gel electrophoretic mobility shift assay (EMSA), circular dichroism (CD) experiments, and nuclear magnetic resonance (NMR), we demonstrate that the proximal promoter region of the *TMPRSS2* gene predominantly forms a three-tetrad parallel-stranded G4 structure (*TMPRSS2*-G4). The NMR solution structure of this major *TMPRSS2*-G4 was determined in a K^+-containing solution. Furthermore, berberine was shown to bind strongly to and stabilize the major *TMPRSS2*-G4, and the complex structure was resolved. Additionally, deoxyribonucleic acid (DNA) polymerase stop assays confirmed that the major *TMPRSS2*-G4 can form stably in extended DNA sequences and be targeted by small molecules to inhibit DNA polymerase progression. This work provides a structural foundation for understanding the major *TMPRSS2*-G4 and its interactions with ligands, offering potential avenues for the design of effective antiviral therapeutics.

2. Material and methods

2.1. Sample preparation

The DNA oligonucleotides used in the experiments were purchased from Sangon Biotech (Shanghai) Co., Ltd. (Shanghai, China). These oligonucleotides were dissolved in a phosphate buffer containing KCl at pH 7.0, and DNA concentrations were measured using Nanodrop One/Onec (Thermo Fisher Scientific, USA) based on A260 absorbance. The DNA samples were annealed to form G4 structures by heating at 95 °C for 5 min, followed by slow or rapid cooling to room temperature. Berberine and coptisine were obtained from Shanghai Standard Technology Co., Ltd. (Shanghai, China), with purity > 98% as determined by HPLC analysis.

2.2. NMR spectroscopy

0.15–1.93 mmol·L^{-1} DNA was dissolved in 2–50 mmol·L^{-1} K^+-containing solution at pH 7.0, comprising 90% H_2O and 10% D_2O. Pu23m3 DNA was used for 2D NMR spectroscopy. NMR spectra were acquired using a 600 MHz Bruker spectrometer (Bruker Ascend 600, Germany). Topspin 4.1.4 (Bruker, Germany) was used for processing NMR data, and Sparky (UCSF) for analyzing 2D NMR spectra [42]. Spectral assignments were achieved through nuclear Overhauser effect spectroscopy (NOESY), double-quantum-filtered correlation spectroscopy (DQF-COSY), and heteronuclear single quantum correlation (HSQC) experiments, as described previously [43]. The DMSO content was 12% in 2D NMR samples and < 2% in 1D NMR samples.

2.3. NOE-derived structure determination

The structure calculations for free *TMPRSS2*-G4 and its complex with berberine followed established protocols [24, 42, 44]. Briefly, initial structures were generated and subjected to simulated annealing using Xplor-NIH 2.48. Subsequently, structures were refined using the Sander module of Amber 20 with implicit water. The ten lowest energy structures were selected for analysis [45].

2.4. DMS footprinting

DMS footprinting was performed according to a previously described protocol with minor modifications [46]. Briefly, 5′-FAM-labeled DNA oligonucleotides were slowly annealed by heating to 95 °C for 5 min and cooling to room temperature overnight. Methylation was carried out by exposing the DNA to 0.8% DMS for 4 min at 25 °C in a 200 μL volume. The reaction was quenched by adding 70 μL stop buffer (1.3 mol·L^{-1} NaOAc, 2.7 mol·L^{-1} β-mercaptoethanol, 3 mg·mL^{-1} sperm DNA). After phenol/chloroform extraction and ethanol precipitation, DNA was resuspended in 100 μL of 10% (*V/V*) piperidine in water and heated at 90 °C for 30 min. Precipitated DNA was dissolved in 80% deionized formamide, denatured at 95 °C for 5 min, and separated on a 20% denaturing polyacrylamide gel. DNA fragments were visualized using a GelGo imager (BIO-RAD, USA). Quantification was performed using ImageQuant 5.2 software.

2.5. CD and thermal denaturing experiments

CD spectra were recorded using a Jasco-1500 (Jasco Inc., Japan) at 25 °C across 230–330 nm, averaged over two scans [42]. Oligonucleotide samples (20 μmol·L^{-1}) were slowly annealed in a buffer containing 125 μmol·L^{-1} K_2HPO_4/KH_2PO_4 and 375 μmol·L^{-1} KCl at pH 7.0. Specific concentrations of berberine or coptisine were added, and mixtures were incubated for 2 h prior to measurement. CD spectra were collected using a quartz cuvette with a 1 mm path length and a response time of 1 sec. Melting analyses were conducted from 25 to 95 °C with a heating rate of 2 °C·min^{-1}.

2.6. Gel EMSA

EMSA was performed on a 1.5 mm thick 10 cm × 7 cm native gel containing 12.5 mmol·L^{-1} KCl and 16% acrylamide in 1 × TBE buffer at pH 8.0. DNA samples were dissolved in K^+-containing buffer (pH 7.0), denatured by heating to 95 °C for 5 min, and gradually cooled to room temperature overnight to allow proper annealing. Ligands were added at desired concentrations and incubated for 2 h. Gels were allowed to polymerize for about 40 min, and electrophoresis was conducted at 45 V for 2 h. DNA bands were visualized under UV light at 254 nm.

2.7. Fluorescence spectroscopy

Fluorescence measurements were carried out using an RF-6000 spectrofluorometer (Shimadzu Inc., Japan). Emission spectra were recorded in a 1 cm path-length quartz cell from 520 to 600 nm, with excitation fixed at 377 nm. Pu23m3 DNA was annealed at 95 °C for 5 min and rapidly cooled on ice. Titration experiments involved adding increasing concentrations of Pu23m3

DNA to a fixed solution of 0.2 μmol·L^{-1} berberine or coptisine in 2 mmol·L^{-1} K^+ solution (pH 7.0, total volume 3 mL). Fluorescence readings were taken after at least 2 min of incubation per concentration point. Dissociation constants (K_d) were calculated using GraphPad by fitting data to a model assuming a 2∶1 binding stoichiometry [47].

2.8. DNA polymerase stop assay

The assay was conducted as previously described [48]. The template DNA (5′-TCACCCTGGGCGGGCGGGGGCGGGGGCAAGATAGCTGCACGCAATTGCTATAGTGAGTCGTATTA-3′) was mixed with 5′-FAM-labeled primer (5′-FAM-TAATACGACTCACTATAGCAATTGC) in a 1∶1.2 molar ratio. Mixtures were heated to 95 °C for 5 min and slowly cooled to 25 °C. Berberine or coptisine was added at varying concentrations and incubated at 25 °C for 3 h. Primer extension was performed for 5 min in a 50 μL reaction mixture containing 0.12 μmol·L^{-1} DNA, 50 μmol·L^{-1} dNTPs, 0.4 U·μL^{-1} Bsu DNA polymerase (large fragment, NEB), 5 mmol·L^{-1} K^+ (pH 7.0), and 10 mmol·L^{-1} $MgCl_2$. Products were separated on a 12% denaturing polyacrylamide gel and visualized using a GelGo Imager (BIO-RAD, USA). Data were analyzed with Image Lab 6.1 software.

2.9. Quantification and statistical analysis

All experiments, including DNA polymerase stop assays and fluorescence spectroscopy, were repeated three times. Statistical analyses and graphing were performed using GraphPad Prism 9.5.0.

3. Results

3.1. The guanine-rich sequence in *TMPRSS2* proximal promoter forms G-quadruplex structures

The proximal promoter region of the *TMPRSS2* gene contains a G-rich sequence named Pu37 (Fig. 1A), consisting of seven G-tracts, each with two to five guanines, capable of forming multiple G4 structures [20]. However, the specific G4 conformations adopted by this sequence remain unclear, limiting the development of antiviral drugs targeting *TMPRSS2*-G4s. To identify the predominant G4 structure, we performed DMS footprinting on wild-type (wt) Pu37 DNA. Protection of guanine N7 positions from DMS methylation indicates Hoogsteen base pairing within G-tetrads, providing evidence of G4 formation under near-physiological conditions [46]. As shown in Fig. 1B, guanines in G-tracts I, II, III, and IV (G4–G6, G8–G10, G13–G16, and G18–G22) are protected, indicating formation of canonical three-tetrad G4 structures in K^+-containing solutions [28]. Notably, the GG dinucleotides G24–G25 and G27–G28 also show protection, suggesting possible formation of alternative species such as two-layered G4s or bulged three-layered G4s (Fig. 1B) [21, 24]. Given their prevalence in the human genome and high thermal stability [23], our subsequent studies focus on the canonical three-tetrad G4, likely the dominant species in the *TMPRSS2* proximal promoter.

G4 topologies are sensitive to environmental conditions, making them amenable to NMR analysis [23, 49]. A truncated wt Pu23 sequence (Fig. 1A and Table S1, Supporting information), containing all G-tracts necessary for canonical three-tetrad G4 formation, was subjected to NMR analysis. The ^{1}H NMR spectrum showed characteristic imino proton signals confirming G4 formation in wt Pu23 DNA (Fig. 1C), although broad and overlapping peaks suggested conformational heterogeneity. Since both G-tracts III and IV contain five guanines—sufficient to form distinct G3 tracts leading to polymorphic structures—we introduced the G13A mutation (Fig. 1A, Pu23m1 DNA) to limit isomer formation based on empirical rules [50, 51], supported by 1D ^{1}H NMR and DMS footprinting (Fig. 1B). Although Pu23m1 showed improved spectral quality, multiple imino proton peaks indicated residual polymorphism (Fig. 1C). To further reduce heterogeneity, we introduced G21A, yielding Pu23m2 DNA, which displayed 12 well-resolved imino proton peaks, indicative of a dominant G4 conformation (Figs. 1A and 1C). An additional C1T mutation produced Pu23m3 DNA, exhibiting similar but better-resolved peaks (Figs. 1A and 1C). Given that Pu23m3 contains two G_3NG_3 motifs typical of parallel G4s and that its 12 imino protons correspond to those observed in Pu23 and Pu23m1 (Fig. 1C), we concluded that the major G4 structure in the *TMPRSS2* proximal promoter is accurately represented by Pu23m3 DNA (*TMPRSS2*-G4). Thus, Pu23m3 was selected for structural characterization.

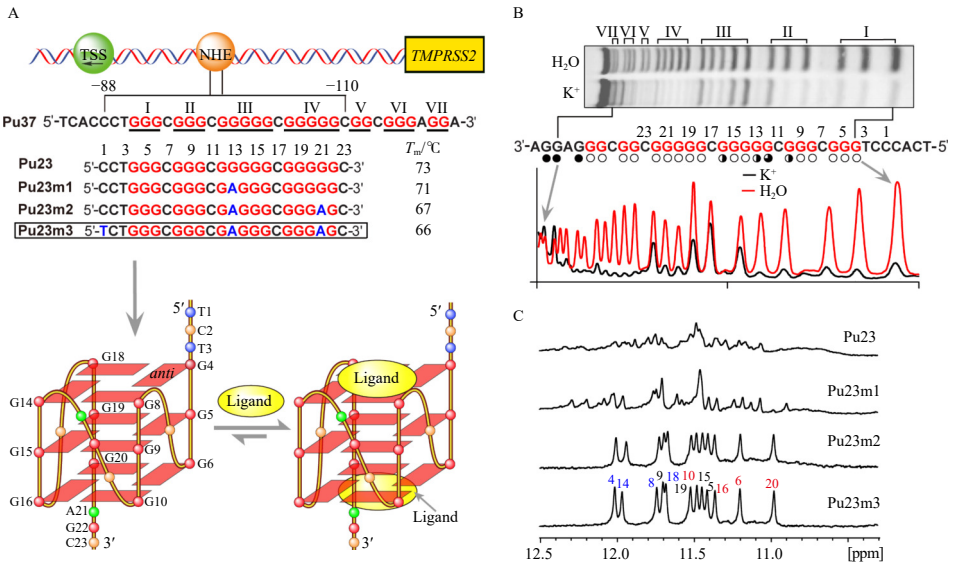


Fig. 1 The guanine-rich sequence in the *TMPRSS2* promoter forms G4s and its interaction with berberine. (A) Schematic of the human *TMPRSS2* gene promoter. G-rich sequence and its mutations in the proximal promoter region of the *TMPRSS2* gene are shown. Continuous guanines are numbered, underlined, and colored red. Mutated residues are colored blue. Melting temperatures of DNA sequences in 50 mmol·L^{-1} K^+ solution are indicated (above), along with a schematic representation of the topology of *TMPRSS2*-G4 (Pu23m3) and its ligand complex (below). (B) DMS footprinting data reveal guanines involved in major *TMPRSS2*-G4 formation in wild-type Pu37 DNA. (C) 1D ^{1}H NMR spectra of wild-type and mutant *TMPRSS2* gene promoter sequences. Conditions: 150 μmol·L^{-1} DNA, pH 7.0.

CD spectra of wild-type and mutant sequences exhibited negative bands at about 242 nm and positive bands at about 264 nm (Fig. 1A and Fig. S1, Supporting information), indicative of parallel G4 topologies [52]. High melting temperatures (T_m) from CD experiments suggest these G4s are stable under physiological conditions (Fig. S1, Supporting information).

3.2. NMR solution structure determination of the major TMPRSS2 gene, which can form G4s (TMPRSS2-G4s)

We have established that the major *TMPRSS2*-G4 in the proximal promoter region adopts a three-tetrad parallel-stranded conformation, effectively modeled by Pu23m3 DNA in K^+-containing solution. To obtain a comprehensive structural understanding, we aimed to determine its three-dimensional architecture. The 1D ^{1}H NMR spectrum of Pu23m3 revealed 12 sharp and well-resolved imino proton peaks (Fig. 1C), suitable for high-resolution structure determination. Accordingly, a suite of 1D and 2D NMR experiments, including NOESY, DQF-COSY, and ^{1}H–^{13}C HSQC, were acquired at various mixing times and temperatures in K^+-containing solution (Fig. 2 and Figs. S2–S5, Supporting information). Established protocols were applied to assign imino, aromatic, and sugar resonances consistent with a parallel G4 fold [53, 54]. Analysis of NMR data, particularly inter-guanine H1-H8 NOE cross-peaks (Fig. 2B), revealed a three-tiered G-tetrad core: G4–G8–G14–G18, G5–G9–G15–G19, and G6–G10–G16–G20 (Fig. 1A). The central tetrad G5–G9–G15–G19 is most protected from solvent exchange, as evidenced by slower imino proton exchange rates (Fig. S6, Supporting information). All flanking, loop, and tetrad residues adopt anti-glycosidic conformations, as indicated by moderate H1′-H6/H8 NOE intensities and characteristic carbon chemical shifts (Fig. 2C; Figs. S2 and S3, Supporting information). Collectively, the *TMPRSS2*-G4 is a three-tetrad-core, parallel-stranded G-quadruplex.

The high-resolution structure of *TMPRSS2*-G4 was determined using restrained molecular dynamics simulations incorporating distance restraints derived from NOESY data (Fig. 2; Fig. S4 and Tables S2–S5, Supporting information). A total of 535 NOE distance restraints, 48 hydrogen bond restraints, and 23 torsion angle restraints (Table 1) were employed. The final ensemble of ten lowest-energy structures showed excellent convergence, with heavy atom root-mean-square deviations (RMSD) of 0.51 ± 0.12 Å for the G-tetrad core and 0.69 ± 0.15 Å for all residues, respectively (Fig. 3A and Table 1).

The core of the *TMPRSS2*-G4 comprises three G-tetrads, with the T1·T3·A13 triad positioned at the 5′-end and the A21·C23 pair at the 3′-end (Fig. 4A; Figs. S7A and S8A, Supporting information). Specifically, the A13 residue is part of a 3-nucleotide loop (C11·G12·A13, Fig. S9A, Supporting information), which is stabilized by T1 and T3 from the 5′-end flanking residues (Fig. 4A; Fig. S7A, Supporting information). It is further stabilized through three potential hydrogen bonds: A13H61/T3O4, A13H62/T1O2, and A13N7/T1H3 (Fig. 4A). This 5′-end capping motif is supported by key NOE cross-peaks involving A13H8 to G14H8, G8H1, and T1Me; A13H2 to G14H1; T1H6 to G4H1, G8H1, and G8H8; and T3Me to G14H1, G18H1, G18H8, and G18H1′ (Fig. 4A; Fig. S9A, Tables S3 and S5, Supporting information). At the 3′-end, the AGC segment adopts a fold-back conformation, with an internal A21·C23 base pair stacking onto the outer G-tetrad. This conformation is stabilized by two potential hydrogen bonds between C23's -NH_2 and A21's sugar moiety (N3 and O4′, Fig. 4A; Fig. S8A and Table S4, Supporting information). Notably, several parallel-stranded G4s, including *MYC*-G4 [55], *PDGFR-β* vacancy G4 [24], and oxidized *BLM*-G4 [25], commonly exhibit this fold-back conformation in their 3′-end capping motifs. This structural feature appears essential for maintaining the overall architecture of

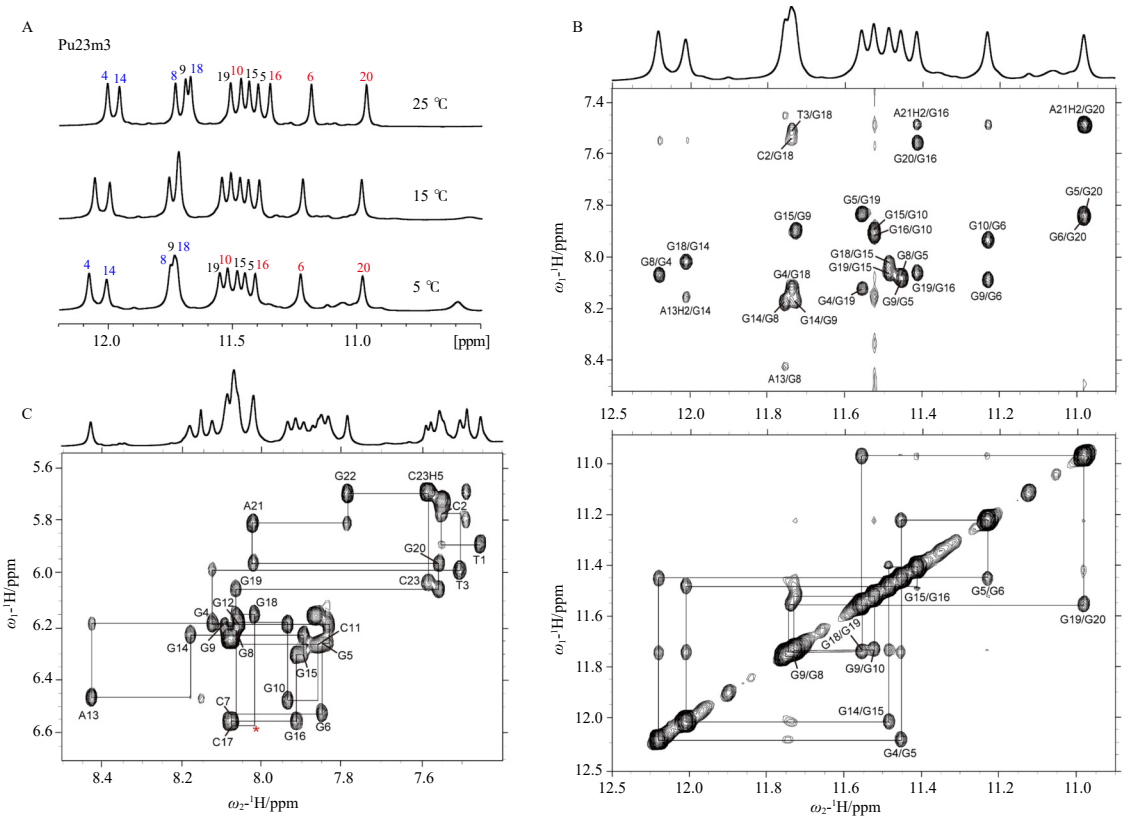


Fig. 2 NMR spectra of *TMPRSS2*-G4 (Pu23m3). (A) 1D ^{1}H NMR spectra of *TMPRSS2*-G4 at different temperatures, with imino proton signals from the 5′-end, middle, and 3′-end G-tetrads highlighted in blue, black, and red, respectively. (B) H1-H8 (top), H1-H1 (bottom), and (C) H1′-H6/H8 regions from NOESY spectra of free *TMPRSS2*-G4, showing sequential assignment pathways in 10 mmol·L^{-1} K^+ solution at 5 °C. Red asterisks denote missing connectivities. Conditions: 1.93 mmol·L^{-1} DNA, 350 ms mixing time, 10 mmol·L^{-1} K^+ solution, pH 7.0.

Table 1 NMR restraints and structural statistics of *TMPRSS2*-G4 and its complex with berberine.

	TMPRSS2-G4	BER-*TMPRSS2*-G4
NOE-based distance restraints		
Total	535	557
Intra-residue	319	319
Inter-residue	216	188
Sequential	158	152
Long range	58	36
Ligand	-	50
Other restraints		
Hydrogen bond restraints	48	48
Torsion angle restraints	23	23
G-tetrad planarity restraints	48	48
Structural statistics (mean ± SD, *n* = 10)		
Pairwise heavy atom RMSD (Å)		
All residues	0.69 ± 0.15	0.55 ± 0.14
G-tetrad core	0.51 ± 0.12	0.48 ± 0.14
Violations		
Max. NOE restraint violation (Å)	0.18	0.16
Mean NOE restraint violation (Å)	0.002 ± 0.012	0.001 ± 0.010

these parallel-stranded G4s. Furthermore, due to the recruitment of A13 by flanking residues, the C11-G12 segment forms a 2-nucleotide loop that resembles a vertical stacking conformation (Fig. S9A and Table S5, Supporting information).

3.3. Natural isoquinoline alkaloids berberine and coptisine strongly bind to and stabilize TMPRSS2-G4

Using the established *TMPRSS2*-G4 model system, our next objective was to identify small molecules capable of binding to and stabilizing this G4 structure. Berberine, abundant in *Coptis chinensis* Franch. (Fig. 5A), has demonstrated significant antiviral activity and strong affinity for parallel G4 structures [42, 56]. We therefore performed 1D ^{1}H NMR spectroscopy to assess the binding specificity of berberine and its derivative coptisine to *TMPRSS2*-G4. Upon titration, significant broadening and upfield chemical shifts in imino proton resonances were observed (Fig. 5B; Figs. S10A and S11A, Supporting information), indicating direct interaction with *TMPRSS2*-G4. Notably, spectra at ligand-to-*TMPRSS2*-G4 ratios of 3∶1 and 4∶1 displayed 12 well-resolved imino proton peaks, indicative of stable complex formation (Figs. 5B and S11A). Characteristic imino proton signals were also detected in complexes with Pu23m1 DNA (Figs. S10B and S11B, Supporting information), suggesting minimal impact of the G21A and C1T mutations on ligand binding to *TMPRSS2*-G4s. We further conducted CD, EMSA, and fluorescence-based titration experiments to characterize the binding of berberine and coptisine to Pu23m3 DNA. Both free Pu23m3 and its ligand-bound complexes exhibited CD spectra typical of parallel G4 topologies, confirming that the G4 fold is preserved upon ligand binding (Fig. 6A). At a 2∶1 ratio, berberine and coptisine increased the thermal stability of *TMPRSS2*-G4 by 23 and 25 °C, respectively (Fig. 6A). NMR titration revealed a 2∶1 binding stoichiometry, as evidenced by minor chemical shift changes upon addition of 2–3 equivalents of ligand (Fig. 5B). Ligands exhibit intermediate-to-fast exchange kinetics on the NMR timescale, often requiring higher ligand ratios than the stoichiometric binding ratio [49, 57]. This adjustment favors formation of a stable complex, as indicated by the presence of 12 well-defined imino proton peaks (Fig. 5B). K_d values were determined *via* fluorescence titration, where Pu23m3 DNA was titrated into solutions of fixed ligand concentration. Berberine and coptisine exhibited K_d values of 0.65 and 0.31 μmol·L^{-1}, respectively (Figs. 6B and 6C), comparable to previously reported affinities for *KRAS*-G4 and *MYC*-G4 [42, 49]. Additionally, native EMSA analysis confirmed that the *TMPRSS2*-G4–ligand complexes remain in monomeric G4 states (Fig. S12, Supporting information). Collectively, these results demonstrate that berberine and coptisine strongly bind to and stabilize *TMPRSS2*-G4.

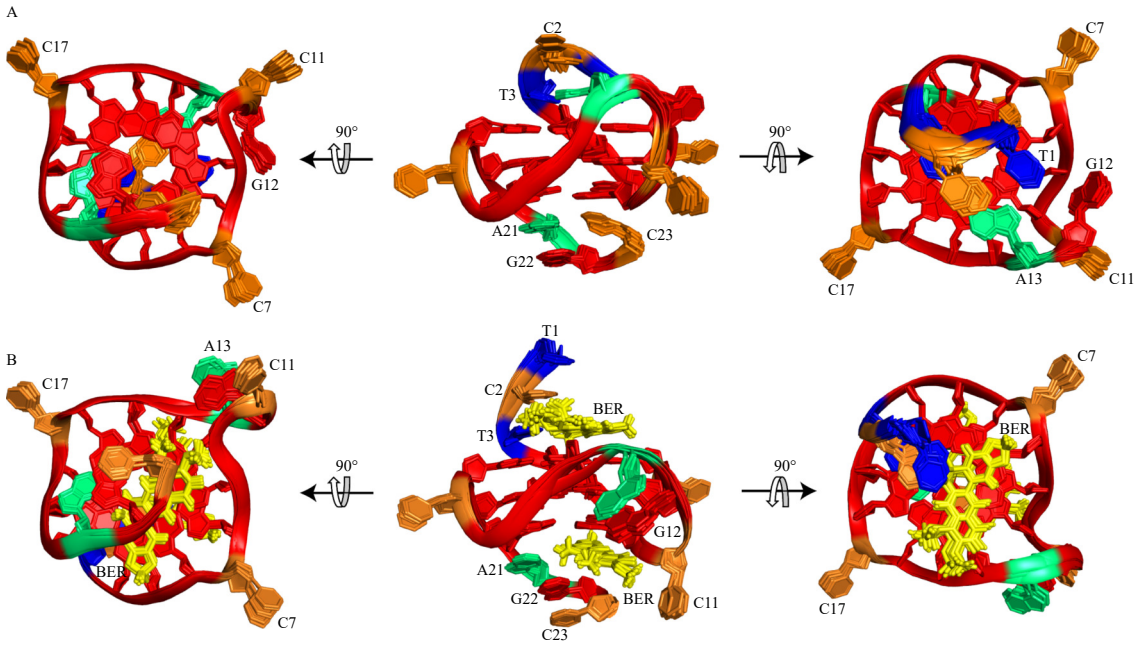


Fig. 3 Structure of *TMPRSS2*-G4 (Pu23m3) and its complex with berberine. The 10 lowest energy structures of the *TMPRSS2*-G4 (A) and BER-*TMPRSS2*-G4 complex (B) were individually superimposed following NOE-restrained structure calculations. Red, guanine; blue, thymine; orange, cytosine; green, adenine; yellow, berberine.

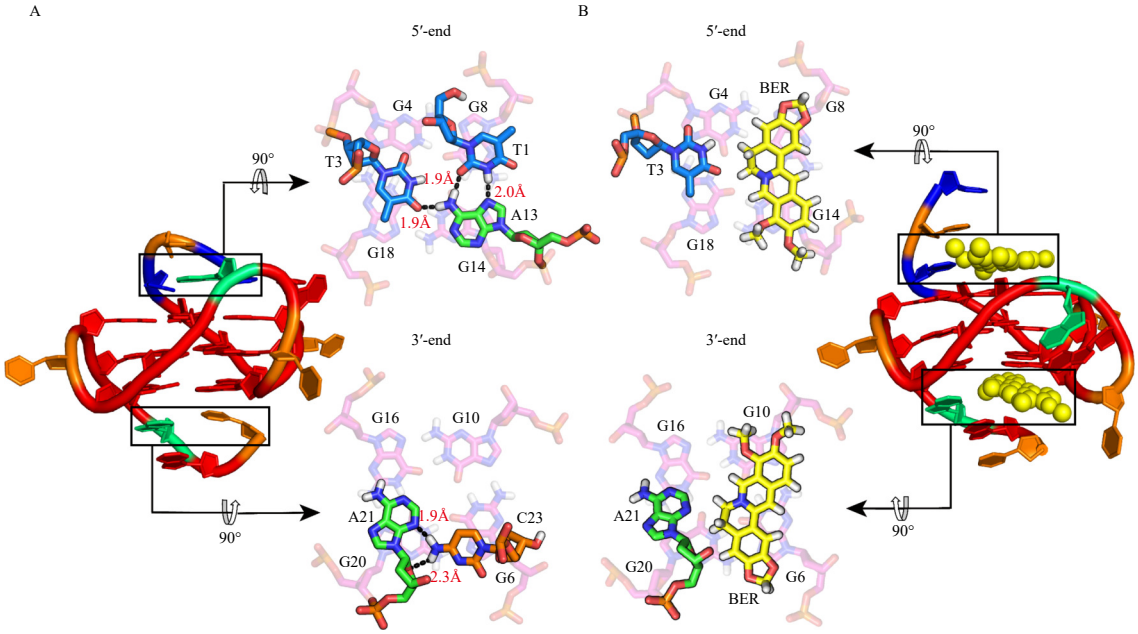


Fig. 4 NMR solution structure details of *TMPRSS2*-G4 (Pu23m3) and its complex with berberine. Cartoon representation of *TMPRSS2*-G4 and the specific top views of the 5′ and 3′-ends of *TMPRSS2*-G4 (A) and BER-*TMPRSS2*-G4 complex (B). Guanine is shown in red, adenine in green, thymine in blue, cytosine in orange, and berberine in yellow. Potential hydrogen bonds are denoted by black dashed lines.

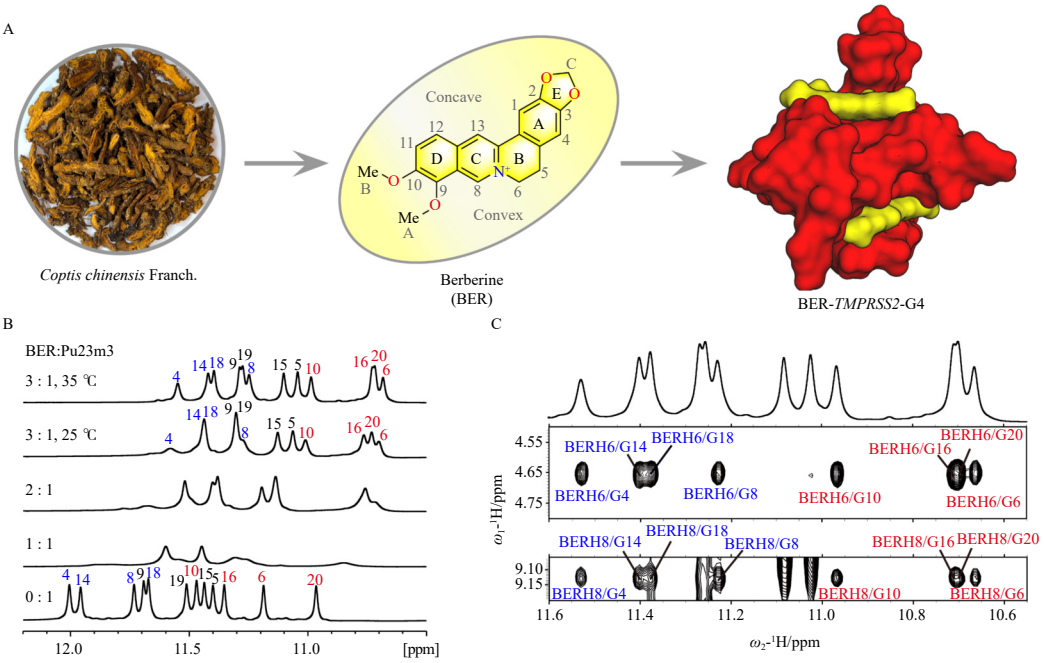


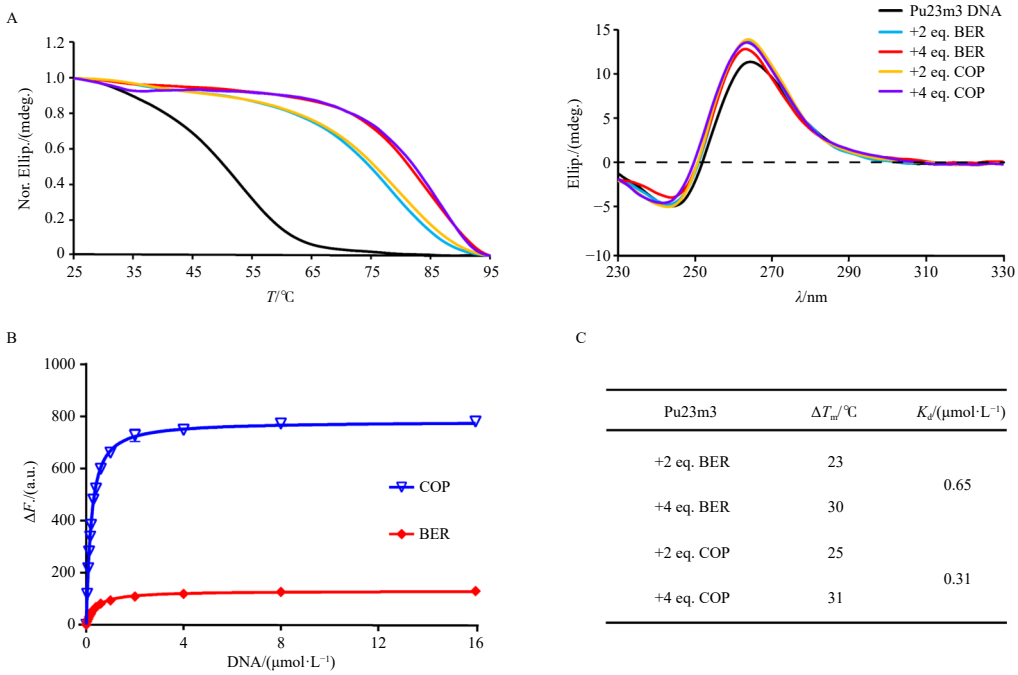
Fig. 5 *TMPRSS2*-G4 (Pu23m3) interaction with berberine. (A) *Coptis chinensis* Franch. is rich in berberine, and the surface view of the BER-*TMPRSS2*-G4 complex structure is shown. (B) 1D ^{1}H NMR spectra displaying imino region changes of *TMPRSS2*-G4 upon titration with berberine. The top spectrum corresponds to the 35 °C, 3∶1 BER-*TMPRSS2*-G4 complex. (C) Representative 2D NOESY spectra of the 3∶1 BER-*TMPRSS2*-G4 complex show intermolecular cross-peaks between berberine and *TMPRSS2*-G4 at 35 °C. Conditions: 1.93 mmol·L^{-1} DNA, 350 ms mixing time, 10 mmol·L^{-1} K^{+} solution, pH 7.0.

3.4. Berberine and coptisine impede DNA replication through targeting TMPRSS2-G4

The DNA polymerase stop assay enables detection of noncanonical DNA structures in template strands that hinder DNA polymerase progression during replication (Fig. 7A) [58]. This method is useful for investigating secondary G4 formation in extended DNA sequences resembling genomic contexts [59]. We employed this assay using a wild-type *TMPRSS2*-G4-containing template to evaluate whether *TMPRSS2*-G4 can form in a longer DNA context and potentially function biologically. Fig. 7B clearly shows a stalled product during early primer extension under potassium-dependent conditions (0–150 mmol·L^{-1}), confirming *TMPRSS2*-G4 formation in the extended template. Addition of berberine and coptisine led to a dose-dependent increase in stalled products (Fig. 7C), indicating that these ligands stabilize *TMPRSS2*-G4 and thereby inhibit DNA polymerase activity. These findings provide strong evidence for the potential formation of *TMPRSS2*-G4 in the human genome and suggest roles in regulating gene replication and transcription. Moreover, targeting *TMPRSS2*-G4 with small molecules may significantly modulate its biological functions.

3.5. NMR solution structure determination of TMPRSS2-G4 in complex with berberine

To elucidate the molecular recognition mechanism between



Pu23m3	ΔT_m/℃	K_d/(μmol·L^{-1})
+2 eq. BER	23	0.65
+4 eq. BER	30	
+2 eq. COP	25	0.31
+4 eq. COP	31	

Fig. 6 Biophysical characterizations of berberine and coptisine interacting with *TMPRSS2*-G4. (A) CD thermal melting curves and CD spectra of Pu23m3 DNA (20 μmol·L^{-1}) in the presence and absence of berberine and coptisine, measured in 0.5 mmol·L^{-1} K$^+$ solution at pH 7.0. (B) Fluorescence titration profiles of berberine and coptisine with Pu23m3 DNA (0.2 μmol·L^{-1} ligand in K$^+$-containing buffer, pH 7.0); K_d values were derived by fitting data to a 2∶1 binding model. Experiments were performed in triplicate. (C) Summary table of ΔT_m and K_d values for berberine and coptisine against Pu23m3 DNA.

small molecules and *TMPRSS2*-G4, we determined the structure of their complex (Fig. 5C). 1D ^{1}H NMR titration of berberine with *TMPRSS2*-G4 (BER-*TMPRSS2*-G4) revealed 12 distinct imino proton peaks (Fig. 5A), indicating formation of a stable, dominant complex suitable for structural analysis. Consequently, 2D NMR spectra (HSQC, NOESY, DQF-COSY) were acquired for the BER-*TMPRSS2*-G4 complex (Fig. 5C and Figs. S13-S16, Supporting information). Resonance assignments for berberine and DNA were made using standard procedures [42, 44] and referencing the free *TMPRSS2*-G4. In the complex, the parallel topology and G-tetrad organization were preserved (Figs. 1A, 5B, and 6A). The largest chemical shift perturbations (Δδ) occurred in H1/H8/H6 protons of outer tetrad guanines and 5′-/3′-flanking residues (Figs. S17 and S18, Supporting information), suggesting stacking of berberine onto the outer G-tetrads. NOE cross-peaks were observed between H6/H8 protons on berberine's convex side and imino protons of guanines in the two outer G-tetrads, indicating that the convex face of berberine faces the G-tetrad center (Figs. 5A and 5B). NOE-restrained MD simulations were used to determine the NMR solution structure of the BER-*TMPRSS2*-G4 complex, em-

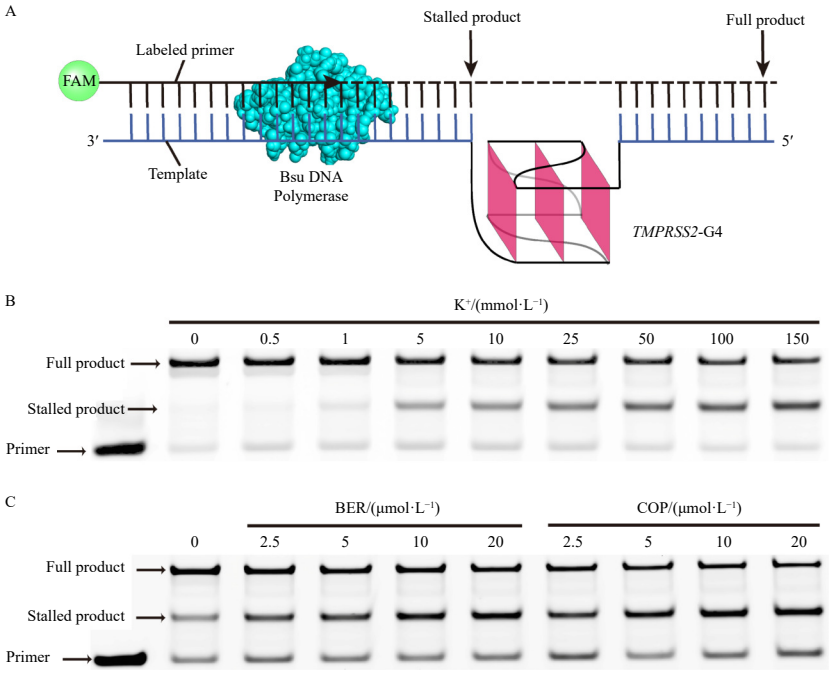


Fig. 7 DNA polymerase stop assay reveals the induction and stabilization of *TMPRSS2*-G4 by ligands and K$^+$. (A) Schematic of the assay: a FAM-labeled primer binds to a DNA template containing a wild-type *TMPRSS2*-G4-forming sequence, followed by extension catalyzed by Bsu DNA polymerase. (B and C) Both K$^+$ and berberine/coptisine stabilize *TMPRSS2*-G4, increasing stalled products and reducing full-length products.

ploying 557 NOE restraints (Table 1 and Tables S6–S11, Supporting information). Binding sites at both outer G-tetrads are well-defined, supported by 50 intermolecular NOE restraints (Fig. 5C; Figs. S15, S16 and Table S8, Supporting information). The final ten lowest-energy structures converged well, with RMSD values of 0.55 ± 0.14 Å for all residues and 0.48 ± 0.14 Å for the G-tetrad core (Fig. 3B and Table 1).

Berberine binding induces rearrangement of residues in both 5′- and 3′-flanking regions, resulting in continuous stacking of flanking sequences (Fig. 4B; Figs. S7B and S8B, Supporting information). This is supported by extensive sequential NOEs: from A21H8, A21H1′, A21H2′, A21H2″, A21H3′, and A21H4′ to G22H8; from G22H8, G22H1′, G22H2′, G22H2″, G22H3′, and G22H4′ to C23H6 at the 3′-end; and from C2H6 and C2H1′ to T3Me; from C2H1′, C2H2′, C2H2″, C2H3′, and C2H4′ to T3H6; and from C2H6, C2H5, C2H2′, C2H2″, C2H3′, C2H5′, and C2H5″ to T1H6 at the 5′-end (Figs. S7, S8 and Tables S8–S10, Supporting information). Intermolecular NOEs were also observed from G20H8, G20H2′, G20H2″, and G20H3′ to A21H8; from G6H8 to BERHC1/HC2; from G10H8 and G10H1′ to BERMeA/MeB; and from A21H2 to BERH5/H6 (Fig. 4B; Fig. S8 and Table S10, Supporting information), indicating that berberine recruits A21 into a coplanar arrangement stacking on the 3′-end G-tetrad. Similarly, T3 is recruited by berberine to form a coplanar stack on the 5′-end G-tetrad, as evidenced by NOEs from T3H6, T3H1′ to G4H8; from G18H1, BERH8, and BERMeA to T3Me; and from G8H8 to BERHC1/HC2 (Fig. 4B; Fig. S7 and Table S9, Supporting information). These observations align with prior studies on berberine complexes with *PDGFR-β* vG4 and *KRAS*-G4 [24, 42].

Furthermore, berberine disrupts hydrogen bonds within the T1·T3·A13 triad, causing A13 to shift away from the 5′-tetrad (Fig. 4 and Fig. S7, Supporting information). Consequently, A13 and G12 adopt a nearly parallel stacking conformation, supported by strong NOEs between G12H8, G12H1′, G12H2′, G12H2″, G12H3′, G12H4′, and A13H8 (Fig. S9 and Table S11, Supporting information). Similarly, hydrogen bonding in the A21·C23 pair at the 3′-end is disrupted, leading to loss of stacking between C23 and the 3′-tetrad (Fig. 4B; Fig. S8, and Tables S8, S10, Supporting information).

4. Discussion

G4s have been identified in telomeres, oncogene promoters, and G-rich satellite DNAs, and are associated with neurological disorders and cancers [60-62]. Recent studies have also revealed G4s in viral genomes [21, 63], where they play critical roles in viral life cycles and pathogenesis [63, 64]. Thus, targeting G4s may offer a novel strategy for antiviral therapy and facilitate the development of small-molecule antivirals [28].

Influenza and coronaviruses are highly transmissible, causing respiratory illnesses of varying severity, and posing substantial public health challenges due to prolonged incubation and seasonal recurrence [65-68]. In 2020, Hoffmann et al. demonstrated that *TMPRSS2* mediates SARS-CoV-2 *S* protein activation [18], showing that camostat, a serine protease inhibitor, blocks viral entry by suppressing *TMPRSS2*. That same year, Shen et al. found that G4-binding compounds stabilize *TMPRSS2*-G4, downregulating *TMPRSS2* expression and inhibiting influenza and coronavirus propagation [20]. In 2022, Liu et al. showed that stabilizing *TMPRSS2*-RNA-G4 reduces SARS-CoV-2 infection by lowering TMPRSS2 protein levels, and further demonstrated that pyridostatin (PDS), a G4-targeting ligand, stabilizes viral RNA G4s and inhibits viral RNA synthesis [41]. Collectively, these findings suggest that reducing *TMPRSS2* expression may be a viable strategy for antiviral intervention.

We aimed to target the major *TMPRSS2*-G4 for gene downregulation. However, the absence of a high-resolution structure has hindered the rational design of *TMPRSS2*-G4-targeting ligands. Here, we report the first high-resolution NMR structure of the major *TMPRSS2*-G4 formed by a G-rich sequence in the proximal promoter region (Figs. 1 and 2). Additionally, we determined the solution structure of its complex with berberine (Figs. 3B and 4B), providing a foundation for rational ligand design. For instance, in the 5′-flanking region, berberine recruits T3 into a coplanar arrangement, positioning T3O4 near BERH5/H6 (Fig. 4B and Fig. S7B, Supporting information). Modifying BERH5/H6 with a hydrogen bond donor could enhance binding specificity by addressing the multiple orientations of berberine observed in G4 pockets [42, 49]. Moreover, the positively charged nitrogen in berberine plays a key role in orienting the molecule toward the G-tetrad center *via* electrostatic interactions with carbonyl oxygens of guanines [42], ensuring the convex face facing inward—critical for complex stability. High-resolution NMR structures of G4-ligand complexes greatly aid in virtual screening and rational design of selective G4-targeting compounds [44, 49, 69].

5. Conclusion

In summary, targeting G4s expands therapeutic avenues for antiviral treatment [23, 28]. This study presents the first high-resolution NMR solution structures of the major *TMPRSS2* proximal promoter G4 and its complex with berberine, offering molecular insights into ligand recognition. These structures will facilitate screening and rational design of ligands with improved efficacy and selectivity for *TMPRSS2*-G4. Furthermore, we confirmed *TMPRSS2*-G4 formation in a genomic-like context and demonstrated that small molecules effectively bind and stabilize it, inhibiting DNA polymerase. The identification of the major *TMPRSS2*-G4 provides critical insight into its regulatory role and highlights its potential as a therapeutic target for viral infections.

Funding

This research was supported by the National Natural Science Foundation of China (Nos. 82322065, 82173707, and 82204241), the Innovation and Entrepreneurship (Shuangchuang) Program of Jiang-su Province (2024), the Natural Science Foundation of Jiangsu Province (No. BK20221039), the Project Program of State Key Laboratory of Natural Medicines (China Pharmaceutical University, No. SKLNMZZ2024JS12), the Fundamental Research Funds for the Central Universities (No. 2632025ZD06), and the Scientific Research Foundation for High-level Faculty, China Pharmaceutical University (No. 3150020065).

Supporting information

The supporting information is available free of charge at: DNA sequences, tables of proton chemical shifts and NOEs, EMSA, 2D-NOESY, $^{1}H-^{13}C$ HSQC, DQF-COSY, side view of structures, 5′- and 3′-end capping structures, fluorescence measurement, chemical shift differences, and CD spectra.

Declaration of competing interest

These authors have no conflict of interest to declare.

Notes

The data underlying this article are available in the Protein Data Bank (PDB) at https://www.rcsb.org/, and can be accessed with ID 8ZVE (free *TMPRSS2*-G4) and 8ZW3 (2∶1 berberine-*TMPRSS2*-G4 complexes). All other data are available in the article or supporting information. Raw data will be shared on request to the corresponding/first authors.

References

1 Kim S. TMPRSS4, a type II transmembrane serine protease, as a potential therapeutic target in cancer. *Exp Mol Med*. 2023;55(4):716-724. https://doi.org/10.1038/s12276-023-00975-5.
2 Vorstandlechner V, Laggner M, Copic D, et al. The serine proteases dipeptidyl-peptidase 4 and urokinase are key molecules in human and mouse scar formation. *Nat Commun*. 2021;12(1):6242. https://doi.org/10.1038/s41467-021-26495-2.
3 Tanabe LM, List K. The role of type II transmembrane serine protease-mediated signaling in cancer. *FEBS J*. 2017;284(10):1421-1436. https://doi.org/10.1111/febs.13971.
4 Barré O, Dufour A, Eckhard U, et al. Cleavage specificity analysis of six type II transmembrane serine proteases (TTSPs) using PICS with proteome-derived peptide libraries. *PLoS One*. 2014;9(9):e105984. https://doi.org/10.1371/journal.pone.0105984.
5 Shapira T, Monreal IA, Dion SP, et al. ATMPRSS2 inhibitor acts as a pan-SARS-CoV-2 prophylactic and therapeutic. *Nature*. 2022;605(7909):340-348. https://doi.org/10.1038/s41586-022-04661-w.
6 Bastus NC, Boyd LK, Mao X, et al. Androgen-induced TMPRSS2:ERG fusion in nonmalignant prostate epithelial cells. *Cancer Res*. 2010;70(23):9544-9548. https://doi.org/10.1158/0008-5472.CAN-10-1638.
7 Lu Q, Nunez E, Lin C, et al. A sensitive array-based assay for identifying multiple TMPRSS2:ERG fusion gene variants. *Nucleic Acids Res*. 2008;36(20):e130. https://doi.org/10.1093/nar/gkn585.
8 Saunders N, Fernandez I, Planchais C, et al. TMPRSS2 is a functional receptor for human coronavirus HKU1. *Nature*. 2023;624(7990):207-214. https://doi.org/10.1038/s41586-023-06761-7.
9 Steiner S, Kratzel A, Barut GT, et al. SARS-CoV-2 biology and host interactions. *Nat Rev Microbiol*. 2024;22(4):206-225. https://doi.org/10.1038/s41579-023-01003-z.
10 Gautam A, Boyd DF, Nikhar S, et al. Necroptosis blockade prevents lung injury in severe influenza. *Nature*. 2024;628(8009):835-843. https://doi.org/10.1038/s41586-024-07265-8.
11 Chu H, Hou Y, Yang D, et al. Coronaviruses exploit a host cysteine-aspartic protease for replication. *Nature*. 2022;609(7928):785-792. https://doi.org/10.1038/s41586-022-05148-4.
12 Wei X, Zhou Y, Shen X, et al. Ciclopirox inhibits SARS-CoV-2 replication by promoting the degradation of the nucleocapsid protein. *Acta Pharm Sin B*. 2024;14(6):2505-2519. https://doi.org/10.1016/j.apsb.2024.03.009.
13 Liu W, Zhao Y, Fan J, et al. Smoke and spike: benzo[a]pyrene enhances SARS-CoV-2 infection by boosting NR4A2-induced ACE2 and TMPRSS2 expression. *Adv Sci (Weinh)*. 2023;10(26):e2300834. https://doi.org/10.1002/advs.202300834.
14 Gowthaman R, Guest JD, Yin R, et al. CoV3D: a database of high resolution coronavirus protein structures. *Nucleic Acids Res*. 2021;49(D1):D282-D287. https://doi.org/10.1093/nar/gkaa731.
15 Zhao MM, Yang WL, Yang FY, et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct Target Ther*. 2021;6(1):134. https://doi.org/10.1038/s41392-021-00558-8.
16 Hoffmann M, Kleine-Weber H, Pöhlmann S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol Cell*. 2020;78(4):779-784.e5. https://doi.org/10.1016/j.molcel.2020.04.022.
17 Bian J, Li Z. Angiotensin-converting enzyme 2 (ACE2): SARS-CoV-2 receptor and RAS modulator. *Acta Pharm Sin B*. 2021;11(1):1-12. https://doi.org/10.1016/j.apsb.2020.10.006.
18 Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;181(2):271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052.
19 Hifumi T, Isokawa S, Otani N, et al. Adverse events associated with nafamostat mesylate and favipiravir treatment in COVID-19 patients. *Crit Care*. 2020;24(1):497. https://doi.org/10.1186/s13054-020-03227-4.
20 Shen LW, Qian MQ, Yu K, et al. Inhibition of influenza A virus propagation by benzoselenoxanthenes stabilizing *TMPRSS2* gene G-quadruplex and hence down-regulating *TMPRSS2* expression. *Sci Rep*. 2020;10(1):7635. https://doi.org/10.1038/s41598-020-64368-8.
21 Zhao C, Qin G, Niu J, et al. Targeting RNA G-quadruplex in SARS-CoV-2: a promising therapeutic target for COVID-19? *Angew Chem Int Ed Engl*. 2021;60(1):432. https://doi.org/10.1002/anie.202011419.
22 Hänsel-Hertsch R, Di Antonio M, Balasubramanian S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol*. 2017;18(5):279-284. https://doi.org/10.1038/nrm.2017.3.
23 Wang KB, Wang Y, Dickerhoff J, et al. DNA G-quadruplexes as targets for natural product drug discovery. *Engineering*. 2024;38:39-51. https://doi.org/10.1016/j.eng.2024.03.015.
24 Wang KB, Dickerhoff J, Wu G, et al. *PDGFR-β* promoter forms a vacancy G-quadruplex that can be filled in by dGMP: solution structure and molecular recognition of guanine metabolites and drugs. *J Am Chem Soc*. 2020;142(11):5204-5211. https://doi.org/10.1021/jacs.9b12770.
25 Wang KB, Liu Y, Li Y, et al. Oxidative damage induces a vacancy G-quadruplex that binds guanine metabolites: solution structure of a cGMP fill-in vacancy G-quadruplex in the oxidized *BLM* gene promoter. *J Am Chem Soc*. 2022;144(14):6361-6372. https://doi.org/10.1021/jacs.2c00435.
26 Robinson J, Raguseo F, Nuccio SP, et al. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res*. 2021;49(15):8419-8431. https://doi.org/10.1093/nar/gkab609.
27 Zheng KW, Zhang JY, He YD, et al. Detection of genomic G-quadruplexes in living cells using a small artificial protein. *Nucleic Acids Res*. 2020;48(20):11706-11720. https://doi.org/10.1093/nar/gkaa841.
28 de Nicola B, Lech CJ, Heddi B, et al. Structure and possible function of a G-quadruplex in the long terminal repeat of the proviral HIV-1 genome. *Nucleic Acids Res*. 2016;44(13):6442-6451. https://doi.org/10.1093/nar/gkw432.
29 Murat P, Zhong J, Lekieffre L, et al. G-quadruplexes regulate epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat Chem Biol*. 2014;10(5):358-364. https://doi.org/10.1038/nchembio.1479.
30 Métifiot M, Amrane S, Litvak S, et al. G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res*. 2014;42(20):12352-12366. https://doi.org/10.1093/nar/gku999.
31 Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod*. 2020;83(3):770-803. https://doi.org/10.1021/acs.jnatprod.9b01285.
32 Luo Z, Yin F, Wang X, et al. Progress in approved drugs from natural product resources. *Chin J Nat Med*. 2024;22(3):195-211. https://doi.org/10.1016/S1875-5364(24)60582-0.
33 Li T, Huang M, Lu J. Cancer statistics and trends in China: the potential of natural product application. *Chin J Nat Med*. 2024;22(8):673-675. https://doi.org/10.1016/S1875-5364(24)60649-7.
34 Hu Q, Zhang Y, Chen P, et al. Discovery and characterization of naturally occurring covalent inhibitors of SARS-CoV-2 M^{pro} from the antiviral herb Ephedra. *Chin J Nat Med*. 2024;22(9):797-807. https://doi.org/10.1016/S1875-5364(24)60577-7.
35 Kuo CL, Chi CW, Liu TY. The anti-inflammatory potential of berberine *in vitro* and *in vivo*. *Cancer Lett*. 2004;203(2):127-137. https://doi.org/10.1016/j.canlet.2003.09.002.
36 Tian Y, Tang G, Gao Y, et al. Carrier-free small molecular self-assembly based on berberine and curcumin incorporated in submicron particles for improving antimicrobial activity. *ACS Appl Mater Interfaces*. 2022;14(8):10055-10067. https://doi.org/10.1021/acsami.1c22900.
37 Zhang S, Xu P, Zhu Z, et al. Acetylation of $p65^{Lys310}$ by p300 in macrophages mediates anti-inflammatory property of berberine. *Redox Biol*. 2023;62:102704. https://doi.org/10.1016/j.redox.2023.102704.
38 Khan S, Hussain A, Attar F, et al. A review of the berberine natural polysaccharide nanostructures as potential anticancer and antibacterial agents. *Biomed Pharmacother*. 2022;146:112531. https://doi.org/10.1016/j.biopha.2021.112531.
39 Valipour M, Zakeri KZ, Abdollahi E, et al. Recent applications of protoberberines as privileged starting materials for the development of novel broad-spectrum antiviral agents: a concise review (2017–2023). *ACS Pharmacol Transl Sci*. 2024;7(1):48-71. https://doi.org/10.1021/acsptsci.3c00292.
40 Babalghith AO, Al-Kuraishy HM, Al-Gareeb AI, et al. The role of berberine in Covid-19: potential adjunct therapy. *Inflammopharmacology*. 2022;30(6):2003-2016. https://doi.org/10.1007/s10787-022-01080-1.
41 Liu G, Du W, Sang X, et al. RNA G-quadruplex in *TMPRSS2* reduces SARS-CoV-2 infection. *Nat Commun*. 2022;13(1):1444. https://doi.org/10.1038/s41467-022-29135-5.
42 Wang KB, Liu Y, Li J, et al. Structural insight into the bulge-containing *KRAS* oncogene promoter G-quadruplex bound to berberine and coptisine. *Nat Commun*. 2022;13(1):6016. https://doi.org/10.1038/s41467-022-33761-4.
43 Dickerhoff J, Jang J, Yang D. Best method to determine DNA G-quadruplex folding: the $^{1}H–^{13}C$ HSQC NMR experiment. *Methods*. 2024;221:35-41. https://doi.org/10.1016/j.ymeth.2023.11.013.
44 Wang KB, Dickerhoff J, Yang D. Solution structure of ternary complex of berberine bound to a dGMP–fill-in vacancy G-quadruplex formed in the *PDGFR-β* promoter. *J Am Chem Soc*. 2021;143(40):16549-16555. https://doi.org/10.1021/jacs.1c06200.
45 Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38. https://doi.org/10.1016/0263-7855(96)00018-5.
46 Li XM, Zheng KW, Zhang JY, et al. Guanine-vacancy–bearing G-quadruplexes responsive to guanine derivatives. *Proc Natl Acad Sci U S A*. 2015;112(47):14581-14586. https://doi.org/10.1073/pnas.1516925112.
47 Wang KB, Elsayed MSA, Wu G, et al. Indenoisoquinoline topoisomerase inhibitors strongly bind and stabilize the MYC promoter G-quadruplex and downregulate MYC. *J Am Chem Soc*. 2019;141(28):11059-11070. https://doi.org/10.1021/jacs.9b02679.
48 Liu Y, Li J, Zhang Y, et al. Structure of the major G-quadruplex in the human EGFR oncogene promoter adopts a unique folding topology with a distinctive Snap-Back loop. *J Am Chem Soc*. 2023;145(29):16228-16237. https://doi.org/10.1021/jacs.3c05214.
49 Dickerhoff J, Brundridge N, McLuckey SA, et al. Berberine molecular recognition of the parallel MYC G-quadruplex in solution. *J Med Chem*. 2021;64(21):16205-16212. https://doi.org/10.1021/acs.jmedchem.1c01508.
50 Wang ZF, Li MH, Chu IT, et al. Cytosine epigenetic modification modulates the formation of an unprecedented G4 structure in the WNT1 promoter. *Nucleic Acids Res*. 2020;48(3):1120-1130. https://doi.org/10.1093/nar/gkz1207.
51 Yang D, Okamoto K. Structural insights into G-quadruplexes: towards new anticancer drugs. *Future Med Chem*. 2010;2(4):619-646. https://doi.org/10.4155/fmc.09.172.
52 Del Villar-Guerra R, Trent JO, Chaires JB. G-quadruplex secondary structure obtained from circular dichroism spectroscopy. *Angew Chem Int Ed Engl*. 2018;57(24):7171-7175. https://doi.org/10.1002/anie.201709184.
53 Kerkour A, Marquevielle J, Ivashchenko S, et al. High-resolution three-dimensional NMR structure of the *KRAS* proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation. *J Biol Chem*. 2017;292(19):8082-8091. https://doi.org/10.1074/jbc.M117.781906.
54 Lin C, Dickerhoff J, Yang D. NMR studies of G-quadruplex structures and G-quadruplex-interactive compounds. *Methods Mol Biol*. 2019;2035:157-176. https://doi.org/10.1007/978-1-4939-9666-7_9.
55 Dickerhoff J, Dai J, Yang D. Structural recognition of the MYC promoter G-

quadruplex by a quinoline derivative: insights into molecular targeting of parallel G-quadruplexes. *Nucleic Acids Res.* 2021;49(10):5905-5915. https://doi.org/10.1093/nar/gkab330.
56 Zheng Z, Li X, Nie K, et al. Identification of berberine as a potential therapeutic strategy for kidney clear cell carcinoma and COVID-19 based on analysis of large-scale datasets. *Front Immunol.* 2023;14:1038651. https://doi.org/10.3389/fimmu.2023.1038651.
57 Kotar A, Wang B, Shivalingam A, et al. NMR structure of a triangulenium-based long-lived fluorescence probe bound to a G-quadruplex. *Angew Chem Int Ed Engl.* 2016;55(40):12508-12511. https://doi.org/10.1002/anie.201606877.
58 He YD, Zheng KW, Wen CJ, et al. Selective targeting of guanine-vacancy-bearing G-quadruplexes by G-quartet complementation and stabilization with a guanine-peptide conjugate. *J Am Chem Soc.* 2020;142(26):11394-11403. https://doi.org/10.1021/jacs.0c00774.
59 Takahashi S, Kotar A, Tateishi-Karimata H, et al. Chemical modulation of DNA replication along G-quadruplex based on topology-dependent ligand binding. *J Am Chem Soc.* 2021;143(40):16458-16469. https://doi.org/10.1021/jacs.1c05468.
60 Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov.* 2011;10(4):261-275. https://doi.org/10.1038/nrd3428.
61 Siddiqui-Jain A, Grand CL, Bearss DJ, et al. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A.* 2002;99(18):11593-11598. https://doi.org/10.1073/pnas.182256799.
62 Verma A, Yadav VK, Basundra R, et al. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res.* 2009;37(13):4194-4204. https://doi.org/10.1093/nar/gkn1076.
63 Figueroa GB, Souza SD, Pereira HS, et al. Development of a single-domain antibody to target a G-quadruplex located on the hepatitis B virus covalently closed circular DNA genome. *J Med Virol.* 2024;96(6):e29692. https://doi.org/10.1002/jmv.29692.
64 Ruggiero E, Richter SN. G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.* 2018;46(7):3270-3283. https://doi.org/10.1093/nar/gky187.
65 Liu D, Chen C, Chen D, et al. Mouse models susceptible to HCoV-229E and HCoV-NL63 and cross protection from challenge with SARS-CoV-2. *Proc Natl Acad Sci U S A.* 2023;120(4):e2202820120. https://doi.org/10.1073/pnas.2202820120.
66 Yao H, Song Y, Chen Y, et al. Molecular architecture of the SARS-CoV-2 virus. *Cell.* 2020;183(3):730-738.e13. https://doi.org/10.1016/j.cell.2020.09.018.
67 de Wit E, van Doremalen N, Falzarano D, et al. SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol.* 2016;14(8):523-534. https://doi.org/10.1038/nrmicro.2016.81.
68 Xiong Q, Cao L, Ma C, et al. Close relatives of MERS-CoV in bats use ACE2 as their functional receptors. *Nature.* 2022;612(7941):748-757. https://doi.org/10.1038/s41586-022-05513-3.
69 Farag M, Mouawad L. Comprehensive analysis of intramolecular G-quadruplex structures: furthering the understanding of their formalism. *Nucleic Acids Res.* 2024;52(7):3522-3546. https://doi.org/10.1093/nar/gkae182.

Natural medicine berberine suppresses host *TMPRSS2* expression through promoter G-quadruplex stabilization: a host-directed antiviral strategy

The host serine protease TMPRSS2 plays a pivotal role in viral entry by priming the spike proteins of numerous viruses, including influenza and corona-viruses. In this study, we unveil a novel host-directed antiviral mechanism of the natural compound berberine. We demonstrate that berberine directly binds to and stabilizes a G-quadruplex (G4) structure located in the promoter region of the *TMPRSS2* gene (*TMPRSS2*-G4). Using high-resolution NMR spectroscopy, we determined the solution structure of the *TMPRSS2*-G4–berberine complex. This interaction could lead to epigenetic suppression of *TMPRSS2* transcription, resulting in deregulation of TMPRSS2 protein on the surface of host cells. Consequently, berberine treatment could inhibit the proteolytic activation of viral spike proteins and block viral entry into host cells. Overall, our findings not only elucidate a potential epigenetic regulating mechanism of berberine but also establish host gene promoter G4s as promising targets for developing broad-spectrum, host-directed antiviral therapeutics.