9920–9937 *Nucleic Acids Research, 2023, Vol. 51, No. 18* *Published online 4 September 2023*
https://doi.org/10.1093/nar/gkad727

Novobiocin blocks nucleic acid binding to Polθ and inhibits stimulation of its ATPase activity

Aleem Syed [1,*], Frantisek Filandr[2], Jeffrey Patterson-Fortin[1], Albino Bacolla [3], Ramya Ravindranathan[1], Jia Zhou[1], Drew T. McDonald[4], Mohammed E. Albuhluli[5], Amy Verway-Cohen [4], Joseph A. Newman [6], Miaw-Sheue Tsai[4], Darin E. Jones[5], David C. Schriemer[2], Alan D. D'Andrea [1,7,*,†] and John A. Tainer [3,8,*,†]

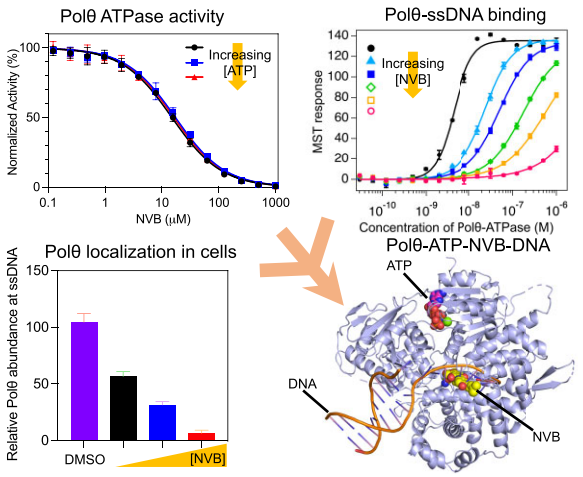
[1]Division of Radiation and Genome Instability, Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA, [2]Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB T2N 4N1, Canada, [3]Department of Molecular and Cellular Oncology, Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA, [4]Biological and System Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [5]Department of Pharmaceutical Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA, [6]Center for Medicines Discovery, University of Oxford, OX1 3QU, UK, [7]Center for DNA Damage and Repair, Dana-Farber Cancer Institute, Boston, MA 02215, USA and [8]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Received March 15, 2023; Revised August 10, 2023; Editorial Decision August 16, 2023; Accepted August 24, 2023

ABSTRACT

Polymerase theta (Polθ) acts in DNA replication and repair, and its inhibition is synthetic lethal in BRCA1 and BRCA2-deficient tumor cells. Novobiocin (NVB) is a first-in-class inhibitor of the Polθ ATPase activity, and it is currently being tested in clinical trials as an anti-cancer drug. Here, we investigated the molecular mechanism of NVB-mediated Polθ inhibition. Using hydrogen deuterium exchange-mass spectrometry (HX-MS), biophysical, biochemical, computational and cellular assays, we found NVB is a non-competitive inhibitor of ATP hydrolysis. NVB sugar group deletion resulted in decreased potency and reduced HX-MS interactions, supporting a specific NVB binding orientation. Collective results revealed that NVB binds to an allosteric site to block DNA binding, both *in vitro* and in cells. Comparisons of The Cancer Genome Atlas (TCGA) tumors and matched controls implied that *POLQ* upregulation in tumors stems from its role in replication stress responses to increased cell proliferation: this can now be tested in fifteen tumor types by NVB blocking ssDNA-stimulation of ATPase activity, required for Polθ function at replication forks and DNA damage sites. Structural and functional insights provided in this study suggest a path for developing NVB derivatives with improved potency for Polθ inhibition by targeting ssDNA binding with entropically constrained small molecules.

GRAPHICAL ABSTRACT



*To whom correspondence should be addressed. Tel: +1 713 563 7725; Fax: +1 713 794 3270; Email: jtainer@mdanderson.org
Correspondence may also be addressed to Alan D. D'Andrea. Email: alan_dandrea@dfci.harvard.edu
Correspondence may also be addressed to Aleem Syed. Email: aleem_syed@dfci.harvard.edu
†The authors wish it to be known that, in their opinion, the last two authors should be regarded as Joint Last Authors.
Present addresses:
Jia Zhou, Bristol Myers Squibb, Cambridge, MA 02142, USA.
Drew T. McDonald, UC Santa Barbara, Santa Barbara, CA 93106, USA.

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

INTRODUCTION

Killing tumors while leaving normal cells unharmed is the goal of precision cancer therapy, a challenging feat that is enabled by targeting tumor-specific vulnerabilities (1–8). The multifunctional polymerase theta (Polθ) with dual ATPase and polymerase activities represents one such vulnerability: loss of this protein causes only a mild phenotype in normal cells (9–12), but tumors, especially those with homology-directed repair (HDR) deficiency, are often addicted to Polθ function for their survival (12,13). Consequently, targeting Polθ potentiates PARP inhibitors (PARPi) in specifically killing HDR-deficient tumor cells, and re-sensitize PARPi-resistant tumor cells (14,15). The induced essentiality between Polθ and HDR-deficiency underscores the value of Polθ as a therapeutic target in the context of tumors with mutated HDR genes.

Polθ is an A-family polymerase with an unusually low fidelity when compared to other polymerases in this class (16,17). It contains an N-terminal helicase-like ATPase domain (HLD) and a C-terminal polymerase domain (PolD) connected by a mostly disordered central region. HLD is a super family 2 (SF2)-type helicase-like domain with DNA-dependent ATPase activity (18,19). PolD has templated-polymerase activity and an inactive proof-reading exonuclease domain (20,21). With these activities, Polθ plays a key role in an alternative DNA double-strand break (DSB) repair pathway called theta-mediated end-joining (TMEJ), also referred to as alternative end-joining (Alt-EJ), or microhomology-mediated end-joining (MMEJ) (22,23). TMEJ, along with single-strand annealing, are relatively enigmatic pathways to repair DNA DSBs compared to the better known pathways of homology directed repair (HDR) and non-homologous end-joining (NHEJ) (24). DSBs are the most severe form of DNA damage and are mostly repaired either by NHEJ with minimal DSB end processing or by HDR in an error-free fashion. HDR is initiated by resection of DSBs by nucleases generating 3′ single-stranded DNA (3′-ssDNA) overhangs (25). The ssDNA binding proteins RPA and RAD51 bind these long ssDNA followed by a homology search and then template-mediated repair. The DSBs with resected DNA are not ideal substrates for NHEJ. In the absence of a functioning HDR, TMEJ rescues these otherwise ill-fated breaks at the cost of genomic instability, due to the low fidelity of Polθ. Polθ ATP-dependent 3′-5′ translocation displaces RPA and RAD51 from the ssDNA and anneals both strands using a microhomology of 2–6 base pairs (bp), so that PolD can use microhomologies to template DNA synthesis for repair and replication fork rescue (12,22,26–28). The mutational signature for Polθ is a 2–6 bp microhomology plus insertions and deletions (indels) that scar repaired sites (29). Both the HLD and the PolD domains of Polθ are required for TMEJ; thus, both have been targeted with small-molecule inhibitors (14,15,30).

We previously discovered that the antibiotic NVB acts as an inhibitor of the ATPase activity of Polθ HLD (14). Consequently, a Phase I clinical trial of NVB is currently underway in patients with tumors that harbor aberrant DNA repair genes (NCT05687110). To inform enigmatic TMEJ and support ongoing preclinical and clinical studies, we therefore investigated the mechanism-of-action of NVB-mediated inhibition of Polθ ATPase activity. By combining hydrogen deuterium exchange-mass spectrometry (HX-MS), biochemical assays, microscale thermophoresis (MST), cellular assays and computational modelling, we determined that NVB is a non-ATP competitive inhibitor that binds to an allosteric site near the ssDNA binding channel in the ATPase core. This is contrary to the current view of NVB as an ATP competitive inhibitor of DNA gyrase. This NVB binding mode blocks ssDNA binding and inhibits ssDNA-mediated stimulation of Polθ ATPase activity. Importantly, we find that NVB blocks Polθ binding to ssDNA both *in vitro* and in cells. Using novobiocic acid (the aglycone of NVB), we investigated the orientation of NVB binding and established the contribution of the sugar group in enhancing the potency of NVB. Our study identifies the NVB binding pocket and provides a path for further optimization and investigation of the importance of ssDNA binding for Polθ biological functions at DSBs and stalled replication forks.

MATERIALS AND METHODS

Protein expression and purification

The Polθ ATPase domain (Polθ-ATPase, 67–894 aa) was expressed and purified from insect cells (19). Briefly, Sf9 cells were infected with Polθ-ATPase baculovirus for 52 h before harvesting. Cell pellets were resuspended in lysis buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 0.5 mM TCEP and 10 mM Imidazole) supplemented with Pierce EDTA-free protease inhibitor tablets (Thermo Fisher Scientific, cat# A32965). Lysates were incubated with TALON® metal affinity resin (Takara Bio, cat# 635502) for 90 min in a cold room for binding. Resin was subsequently washed, and bound protein was eluted in the elution buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 0.5 mM TCEP and 300 mM Imidazole). The His-tag from the eluted protein was cleaved by TEV protease and removed by repassing through the TALON® resin. Finally, protein was loaded on a pre-equilibrated size exclusion chromatography (SEC) column Superdex 200 Increase 10/300 GL or Superose 6 Increase 10/300 GL (Cytiva) in SEC buffer (10 mM HEPES (pH 7.5), 250 mM NaCl, 0.5 mM TCEP) and fractions with Polθ-ATPase were concentrated and flash-frozen. Polθ-ATPase (1–987 aa) was expressed and purified as reported previously (14).

Novobiocic acid preparation

Novobiocic acid was prepared as previously reported and spectroscopic data were consistent with literature reports (31,32). Briefly, a solution of 250 mg NVB sodium salt and 5 ml of 5 M NaOH was heated in a 20 ml reaction vial with stirring to 80°C for 45 min. The reaction mixture was allowed to cool to room temperature and neutralized with an equal volume of 5 M HCl. The aqueous solution was extracted with 3 × 10 ml of ethyl acetate. The combined organic solutions were washed with water (1 × 10 ml) and brine (2 × 10 ml). The organic solution was dried (Na_2SO_4), filtered and solvents removed under reduced pressure. The crude product was purified by medium pressure liquid chromatography (20% ethyl acetate in hexane to 60% ethyl acetate in hexane)

to yield N-(4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-4-hydroxy-3-(3-methylbut-2-en-1-yl)benzamide (novobiocic acid) (123.1 mg, 79% yield) as a light-yellow solid: ^{1}H NMR (300 MHz, DMSO-*d6*) δ 11.83 (br s, 1H), 10.41 (s, 1H), 10.04 (s, 1H), 9.17 (s, 1H), 7.74–7.70 (m, 2H), 7.57 (d, J = 8.6 Hz, 1H), 6.87 (t, J = 8.6 Hz, 2H), 5.31 (t, J = 7.4 Hz, 1 H), 3.27 (d, J = 7.4 Hz, 2H), 2.17 (s, 3H), 1.70 (s, 6H); LRMS for $C_{22}H_{21}NO_6$ m/z 394.19 $(M\text{-}H)^-$. Compound used for biological assays was greater than 95% pure based on NMR and HPLC with UV absorbance at 210 nm and 254 nm wavelengths.

ADP-glo ATPase assay

ADP-Glo assay (Promega) was used to study ATPase activity of the Polθ-ATPase (1–987 aa) and inhibition of this activity by NVB (Sigma, cat# 46531). First, serial dilutions of NVB were made in the assay buffer (40 mM Tris (pH 8.0), 20 mM $MgCl_2$, 50 mM KCl and 1 mM DTT). Then an equal volume Polθ-ATPase and ssDNA (5′-CCAGTGAATTGTTGCTCGGTACCTGCTAAC-3′ purchased from Integrate DNA technology with HPLC purification) mixture was added to each tube of NVB and incubated for 1 h at room temperature. The final concentrations were as follows: 50 nM Polθ-ATPase, 500 nM ssDNA (or 0 nM for reactions without ssDNA), and NVB concentration varying between 120 nM and 1 mM in a 30 μl reaction volume. The ATPase assay was initiated by adding 5 μl ATP (7x stock of the required final concentration was prepared in the assay buffer) and the reaction was further incubated for 2 h at room temperature. The ATPase activity was measured according to the manufacturer's instruction with minor modifications. The ATPase reaction was stopped by adding and incubating with a 30 μl of ADP-Glo reagent. Finally, a 60 μl kinase detection reagent was added to the reaction and incubated for an additional 1 h at room temperature before measuring the luminescence to quantify the ATPase activity on a CLARIOstar microplate reader (BMG Labtech). All data were background-subtracted using buffer-only wells and then normalized to enzyme-only wells (showing 100% activity). Luminescence data from three independent experiments were averaged. Computed mean and standard deviations were imported to Prism (GraphPad) and IC_{50} were estimated by fitting the normalized response vs. concentration of the inhibitor equation to the data in the Prism. To test if NVB is an ATP-competitive inhibitor, ATPase activity assays were performed at 10, 100 or 300 μM of final ATP concentrations and the resulting IC_{50}'s were compared. The concentrations of ATP were selected based on the previously measured K_m (40–50 μM) of ATP for Polθ using a pyruvate kinase, lactate-dehydrogenase enzyme-linked absorbance ATPase assay (19). We also tested the effect of different concentrations of ATP on IC_{50} of NVB using the same protocol except we omitted the 1 h pre-incubation step and instead added NVB and ATP together to initiate the ATPase reaction.

Microscale thermophoresis (MST)

The direct binding of purified Polθ ATPase domain and ssDNA was measured using MST. For the MST experiment, a 5′-Alexa Fluor® 647N-labelled ssDNA (5′-CCAGTGAATTGTTGCTCGGTACCTGCTAAC-3′) (Alexa647N-ssDNA) was purchased from Integrated DNA Technologies with HPLC purification. MST experiments were performed on a Monolith NT.115Pico system (NanoTemper). All MST experiments were performed in the following buffer: 10 mM HEPES (pH 7.5), 120 mM KCl (or 50 mM KCl for the low-salt binding), 0.5 mM TCEP and 0.05% Tween-20. Both ligand (unlabeled Polθ ATPase domain) and target (Alexa647N-ssDNA) dilution were prepared in the MST buffer. Equal volumes of 2 nM Alexa647N-ssDNA and serially two-fold diluted unlabeled Polθ ATPase domain were mixed to obtain a fixed concentration of Alexa647N–ssDNA (1 nM) and variable concentration of unlabeled Polθ ATPase domain (final concentrations ranging from 30 pM to 1 μM). The mixture was then incubated for 10 min at room temperature before loading into regular Monolith NT.115 capillaries for the MST measurements. After validating the assay conditions, the following instrument settings were used for the binding affinity experiments: 20% excitation power (or 15% for low-salt binding) in the Pico-RED channel and medium MST power with other default instrument settings, including the experiment temperature set to 25°C. Each binding experiment was performed in three independent runs and resulting raw MST data were loaded to the analysis software (MO.Affinity Analysis v2.3, NanoTemper) and fit with the Hill model. To study the effect of NVB, AMP-PNP or novobiocic acid on Polθ ATPase domain binding to ssDNA, MST experiments were performed as described above in the MST buffer with indicated concentration of the small molecules. For the MST experiments with AMP-PNP, both control and AMP-PNP runs were performed with MST buffer supplemented with 5 mM $MgCl_2$. MST data were presented as a change in normalized fluorescence (due to thermophoresis) as a function of unlabeled Polθ ATPase domain concentration. For presentation purposes, both raw and fit data were exported and plotted in IgorPro9 software (WaveMetrics).

Hydrogen deuterium exchange-mass spectrometry (HX-MS)

NVB or novobiocic acid were first dissolved in DMSO and then diluted in the assay buffer (25 mM HEPES pH/pD 7.4, 250 mM NaCl) at the indicated concentrations. ssDNA was reconstituted in water and then diluted in the assay buffer at the indicated concentrations.

For ligand-bound samples, Polθ was prepared to 4 μM then mixed with 4 mM NVB or novobiocic acid (both in 2% DMSO) or 8 μM ssDNA in 1:1 volume ratio, and pre-incubated for 30 min. For each sample, 5 μl of the pre-incubated protein was mixed with 5 μl of D_2O-based assay buffer to initiate the deuterium labelling at room temperature. After 2 min of labelling, the reactions were quenched by mixing each of the samples with 10 μl of quench/digestion buffer (500 mM glycine pH 2.3, 6 M urea) containing nepenthesin II digestion enzyme (NVB: 0.6 μg/μl, novobiocic acid: 0.4 μg/μl, ssDNA: 0.3 μg/μl). Samples were incubated for 2 min at 8°C and flash frozen. HX-MS control samples were prepared with matched DMSO concentrations and processed as above,

but with a quench/digestion buffer containing 0.2 μg/μl nepenthesin II. All samples and controls were prepared in triplicate.

HX-MS samples were analyzed on a Sciex TTOF 6600 instrument with an Optiflow Micro ESI Source, integrated with a Sciex Ekspert nanoLC 425 and a Trajan PAL HDX autosampler. Samples were manually injected into the cold compartment of the autosampler (set to 4°C) and desalted on a Luna C18(2) microtrap column (0.3 mm diameter, 5 μm particle size, 100 Å pore size, 20 mm length) for 90 sec at 50 μl/min mobile phase A (0.4% formic acid in MS-grade H_2O). Bound peptides were subsequently eluted and separated on Kinetex XB-C18 column (0.3 mm diameter, 2.6 μm particle size, 100 Å pore size, 50 mm length) connected directly to the ion source, using a linear 10-min gradient from 5%-35% mobile phase B (0.4% FA in MS-grade acetonitrile) at 8 μl/min. To obtain a list of peptides for deuterium analysis, sequence maps for Polθ were prepared from standard data-dependent acquisition analyses of control samples, using HX-PIPE (33). All HX-MS data were analyzed using the HX-DEAL functionality of the Mass Spec Studio (33). MSTools was used for workflow management (34). Associated HX-MS data are provided in standard exports format as supporting information (Supplementary HX data).

For higher sensitivity analysis of NVB-bound Polθ ATPase domain, a nanoHX-MS configuration was used. The same LC-MS system was used with the following modifications. The source was switched to an Optiflow Nano ESI Source and the cold chamber of the autosampler was outfitted with an Acclaim™ PepMap™ 100 C18 HPLC trap column for desalting (0.1 mm diameter, 5 μm particle size, 100 Å pore size, 20 mm length) at 10 μl/min mobile phase A for 3 min. Bound peptides were subsequently eluted and separated on nanoEase M/Z Peptide CSH C18 Column (75 μm diameter, 1.7 μm particle size, 130 Å pore size, 150 mm length), connected directly to the ion source, using a linear 10-min gradient from 5%-35% mobile phase B at 250 nl/min. Samples were prepared as above but at higher dilution and with reduced levels of digestion enzyme, and a separate data-dependent acquisition analysis conducted for sequence mapping.

Thermal stability assay

Protein thermal stability of Polθ-ATPase domain in the presence of either NVB, novobiocic acid or DMSO was measured using Prometheus nano Differential Scanning Fluorimetry (nanoDSF) (NanoTemper). Given that the intrinsic fluorescence emission from the protein coincides with the chromophore in the small molecule, we monitored scattering to measure the protein stability with temperature (14). Polθ ATPase domain (0.5 mg/ml in 25 mM HEPES (pH 8.0) and 200 mM NaCl) was incubated with the 500 μM of NVB, novobiocic acid or DMSO for 30 min at room temperature before samples were loaded into capillaries for the scattering measurement, in which temperature was increased at a rate of 1°C/min starting from 25°C till 90°C. The shift in the peak max of scattering was estimated from the first derivative of the scattering data. All data points were an average from three replicates and plotted in IgorPro9 software (WaveMetrics).

Mass photometry

We studied the oligomeric state of the Polθ ATPase domain in solution using mass photometry. The Polθ ATPase domain (250 nM) in the assay buffer (20 mM HEPES (pH 7.5), 150 mM NaCl, 1 mM TCEP) incubated with 500 μM of NVB for 1 h on ice before measurements. All experiments were performed at room temperature on a Refyn TwoMP mass photometry instrument (Refyn Ltd). Sample holder was prepared by attaching a sample well cassette (6-well silicone gasket) to a clean No. 1.5H high precision glass coverslip (24 × 50 mm, Thorlabs cat# CG15KH). Mass calibration mixture (consisting of β-amylase and thyroglobulin) was prepared in the assay buffer as well. We used a large image size setting for a 60 s data recording for each sample including mass calibration standards. Also, we used the 'droplet dilution' mode to find the focus. For each run, after finding the focus, the sample was diluted 10-fold by directly injecting sample into the droplet that was used to find the focus. Polθ ATPase domain in the assay buffer used as a control. For NVB treated Polθ ATPase domain, the dilution buffer contained 500 μM NVB. We also ran a mass calibration standard in the assay buffer containing 500 μM NVB and found no significant effect of NVB on the expected molecular weights of β-amylase or thyroglobulin. The data were analyzed in DiscoverMP software (Refyn Ltd). For each sample, all events along with amplitudes (%), mean (mass) and standard deviations (mass) for each peak were exported from the DiscoverMP. For the data presentation purpose, we generated histograms of exported events using Prism (GraphPad) and Gaussian fits were simulated using the exported mean and standard deviation for each peak from the DiscoverMP analysis software.

Molecular docking

For the Glide docking studies with the Schrödinger suite (www.schrodinger.com) we used a previously described protocol (14), except that a search grid was created around residue I191, given its proximity to a potential binding site for NVB, as identified by HX-MS data. For all the generated poses, binding free energy MM-GBSA calculations were performed using the Prime module and the top two models selected based on the resulting protein–ligand complex with the lowest MM-GBSA binding free energy value (35,36).

Cell lines

We previously used U2OS cells in studying the effect of NVB on Polθ-mediated MMEJ signaling (14). We continued to use U2OS cells in measuring the recruitment of Polθ to the damage sites as U2OS cells enable efficient transfection and expression of full-length N-terminus Myc-tagged Polθ in an endogenous *POLQ* knockout background. Further, the Myc-tag on Polθ enables analysis of Polθ localization as commercial antibodies available for Polθ are not validated for immunofluorescence. To generate a CRISPR-Cas9 *POLQ* knockout (KO) U2OS cell line, the multi-guide Gene Knockout Kit v2 (Synthego) was employed. Three guides (GGCUCAGCACGGACCCGGAG GAUUCGUUCUCGGGAAGCGG, AUGAAUCUUCUGCGUCG-

GAG) targeting human *POLQ* were electroporated using a 4D-Nucleofector X unit (Lonza) according to the manufacturer's instructions. At 48 h post-electroporation, genomic DNA was extracted and analyzed by DNA sequencing to confirm the disruption of the functional *POLQ* allele and Polθ levels by western blot analysis. The U2OS *POLQ* knockout cell line and RPE1 *TP53*$^{-/-}$*BRCA1*$^{-/-}$ cells (37) (used for DNA fiber assays) were maintained in F12 (Gibco) media supplemented with 10% FBS, and regularly tested for *mycoplasma* using MycoAlert Mycoplasma Detection Kit (Lonza) and identity confirmed at periodic intervals by the DFCI Molecular Diagnostics Core Human Cell Line Authentication service.

Transfections and plasmids

U2OS *POLQ* KO cells were seeded at 1×10^5 cells/ml into 6-well plates containing coverslips one day prior to planned transfection. Cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions and analyzed 48 h after transfection. The expression vector for full-length human Polθ with an N-terminus Myc-tag (pcDNA 3.1(Hygro) myc-hPolQ-Flag), was a gift from Agnel Sfeir (Addgene plasmid # 73132).

Laser micro-irradiation and immunofluorescence

At 24 h post transfection, U2OS cells growing on coverslips were pre-sensitized with BrdU as described (38) and treated with DMSO at the indicated concentrations of NVB (Selleckchem, cat# NSC2382), or novobiocic acid, for 24 h. The following day, 2 h prior to laser micro-irradiation, the media was refreshed. Laser micro-irradiation was performed using a two-photon laser (800 nm, 40% power) on an inverted laser scanning confocal microscope (LSM880NLO/Mai Tai Laser, Zeiss, Spectra-Physics). One hour after micro-irradiation, cells were washed in PBS once, pre-extracted for 5 min at 4°C and then fixed using 4% paraformaldehyde for 10 min at room temperature. Fixed cells were permeabilized for 5 min at 4°C. Following permeabilization, cells were incubated with a rat BrdU antibody (Abcam, cat# ab6326), and a mouse Myc-tag antibody (Cell Signaling Technology, cat# 2276) for 1 h at 37°C. After washing three times, cells were incubated with Alexa Flour™ 594 (Invitrogen, cat# A-21471), and Alexa Flour™ 488 (Invitrogen, cat# A-11029)-conjugated secondary antibodies for 1 h at 37°C. Following secondary antibody incubation, coverslips were washed four times and then mounted on glass slides using mounting media containing DAPI (Vectashield). Cells were imaged using a Zeiss AxioObserver microscope and quantified using Image J (NIH) where data was normalized to the DMSO-treated Polθ control, which was set to 100%.

DNA fiber assay with S1 nuclease digestion

The DNA fiber assay to quantify ssDNA gaps were performed as previously described (37) and adapted for FiberVision® (Genomic Vision) workflow. Briefly, RPE1 *TP53*$^{-/-}$*BRCA1*$^{-/-}$ cells were seeded for 24 h before the labelling experiment. On the day of the experiment, cells were labelled with 200 μM CldU (Sigma, cat# C6891) for 30 min, washed three times with warm PBS and then labelled with 100 μM IdU (Sigma, cat# I7125) for 2 h. DMSO or 200 μM NVB were added at the same time as the second label. Cells were washed three times with warm PBS and then permeabilized with CSK buffer (100 mM NaCl, 10 mM MOPS, 3 mM $MgCl_2$, 300 mM sucrose, 0.5% TritonX-100) for 10 min at room temperature. Permeabilized nuclei were washed once with PBS followed by S1 buffer (50 mM NaCl, 300 mM sodium acetate (pH 4.6), 10 mM zinc acetate, 5% glycerol) and then incubated with 10 U/ml of S1 nuclease (Sigma, cat# N5661) in S1 buffer for 30 min at 37°C. Nuclei were scrapped in 1 ml of PBS (with 0.1% BSA), centrifugated 5 min at 1500 rpm and resuspended in PBS. Agarose plug preparation and proteinase K treatment were performed as per FiberPrep DNA extraction protocol (Genomic Vision, cat# EXT001A). Agarose plugs were then digested with beta-agarase overnight. Samples were then poured into FiberComb wells and combed onto silanized coverslips (Genomic Vision, cat# COV-002) using the Molecular Combing System from Genomic Vision. Coverslips were denatured with 0.5 M NaOH/1 M NaCl for 8 min followed by dehydration in 70%, 90% and 100% ethanol for 3 min each. The coverslips were probed with rat BrdU antibody (clone BU1/75 (ICR1)) that recognizes CldU (Abcam, cat# ab6326), mouse BrdU antibody that recognizes IdU (BD Biosciences, cat #347580) and mouse anti-ssDNA (Developmental Studies Hybridoma Bank, University of Iowa). Further, coverslips were incubated with corresponding secondary antibodies: Cy5® goat anti-rat (Abcam, cat# ab6565), Cy3® goat anti-mouse (Abcam, cat# ab97035) and BV480 goat anti-mouse (BD Biosciences, cat# 564877). Fibers were scanned using FiberVision S. Fibers were analyzed with ImageJ. The lengths of IdU tracks in each fiber were measured by ImageJ and graphed using Prism (GraphPad). For each treatment, at least 200 fibers were analyzed.

Bioinformatic analyses

We used an in-house pipeline to assess the gene expression levels and to perform coexpression analyses and survival analyses in TCGA tumor types. Briefly, we downloaded the RNA-seq normalized rsem values for tumor samples and matched controls, and clinical data through the TCGA-assembler 2 software (39). We then used custom scripts to retrieve, process and plot the data in a parallel computing environment, as described (40). Of the 33 TCGA tumor types available, 15 contained at least 10 matched controls samples; these 15 tumor types were used for gene expression analyses.

TCGA tumor abbreviations are as follows: ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary

cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumors, THCA, thyroid carcinoma; THYM, thymoma, UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma.

For survival analyses, patient data for each type of tumors were separated into two groups, g_high, which contained samples with *POLQ* expression above the mean, and g_low, which contained samples with *POLQ* expression below the mean. Mutational signatures data were retrieved from the Catalogue Of Somatic Mutations In Cancer (COSMIC, https://cancer.sanger.ac.uk/signatures/) and processed using custom scripts. A master file was used in which each patient sample contained the percent of single base substitutions (SBSs) mapping to each of the 30 mutational signatures. The set of signatures were that of version 2, from 2015 (https://cancer.sanger.ac.uk/signatures/signatures_v2/). The master file was queried to compute the number of samples containing a given signature for each of the ∼20 500 available genes for which mutations had been catalogued, both for the g_high and the g_low samples. The difference in the number of g_high and g_low samples for any given gene was then assessed by Fisher's exact tests.

For gene set enrichment analyses (GSEA), we used the database for annotation, visualization and integrated discovery (DAVID, https://david.ncifcrf.gov); from there we selected the functional annotation clustering analysis using a medium classification stringency. The *P*-values were adjusted for multiple testing using the Benjamini-Hochberg correction. The top 100 genes most highly coexpressed with *POLQ* in the ARCHS4 web resource were obtained at https://maayanlab.cloud/archs4/gene/POLQ.

Statistical analyses

Statistical analyses were performed using Prism (GraphPad). All data are represented as mean ± SD, unless indicated otherwise. Significance was tested using one-way ANOVA for comparison of three or more sets, unless indicated otherwise.

RESULTS

NVB is a non-ATP competitive inhibitor of Polθ ATPase activity

Previously, we identified NVB as an inhibitor of Polθ ATPase activity and demonstrated that it binds and stabilizes the Polθ ATPase domain (14). Molecular docking studies centered on the ATP-binding pocket implied NVB binds to a tunnel near the ATP binding site instead to the ATP binding pocket in the ATPase domain. NVB was originally identified as a natural compound that inhibits the *E. coli* gyrase B (GyrB) ATPase domain (41,42). Structural analysis revealed that NVB inhibits GyrB ATPase activity by directly binding to the ATP-binding pocket (Supplementary Figure S1A) (43–45). Given the implied differences in NVB inhibition of Polθ ATPase compared to our assessment of the GyrB ATPase structural data, we sought to understand the underlying allosteric mechanism of NVB inhibition of Polθ. We initially measured NVB inhibition of Polθ ATPase activity in the presence of variable ATP concentrations using an ADP-Glo ATPase assay, where a 14-point NVB dose response curve was measured for three different concentrations (10, 100 or 300 μM) of ATP in the reaction (Figure 1A).

Our data show that there is no statistically significant change in the IC_{50} at three different tested concentrations of ATP (Figure 1A). Additionally, we also examined if adding NVB with the ATP, without pre-incubation with the enzyme, changes the IC_{50}. As with the pre-incubation, there was no significant effect of different concentrations of ATP on the measured IC_{50} (Supplementary Figure S1B). From these data, NVB is not directly binding to the ATP binding pocket, as seen for GyrB, but instead is a non-ATP competitive inhibitor of Polθ ATPase activity.

NVB binds to an allosteric site in the Polθ ATPase domain

To experimentally determine the NVB binding site in the Pol θ ATPase domain, we initially tested protein X-ray crystallography, as the Polθ ATPase domain has been crystallized previously (19). Unfortunately, our crystals of the Polθ ATPase domain in the presence of NVB were not suitable for useful diffraction experiments, consistent with potential ligand-induced conformational changes. Similarly, soaking experiments for pre-formed Polθ ATPase domain crystals with NVB were not fruitful. We therefore turned to hydrogen deuterium exchange-mass spectrometry (HX-MS) to map NVB binding site in Polθ ATPase domain.

HX-MS has the advantage of being a solution-based analytical technique that can determine the dynamic structural changes in the protein upon ligand/DNA binding (46–49). HX-MS measures the changes in mass associated with the exchange between amide hydrogens of the protein backbone and deuterium from the surrounding solvent (D_2O). The change in deuterium incorporation is compared between the ligand bound versus the ligand-free states of the protein to map potential binding sites and allosteric changes. HX-MS has been successfully used to map ligand binding sites and ligand-induced allosteric changes in a wide range of samples and states (49–54).

We measured changes in deuteration (Δ%D) in the Polθ ATPase domain treated with NVB, compared with the vehicle control. During method development, we noted a high resistance to digestion for the NVB-treated Polθ, which necessitated a more sensitive nanoHX-MS configuration. This mode allowed us to achieve 78% sequence coverage, whereas the conventional mode only generated 58% coverage. Based on the Δ%D in different regions of the protein using significance determinations as previously described (52), we identified three classes of perturbations (Figure 1B). First, strong protection from exchange was observed in the regions boxed in green. These regions represent the

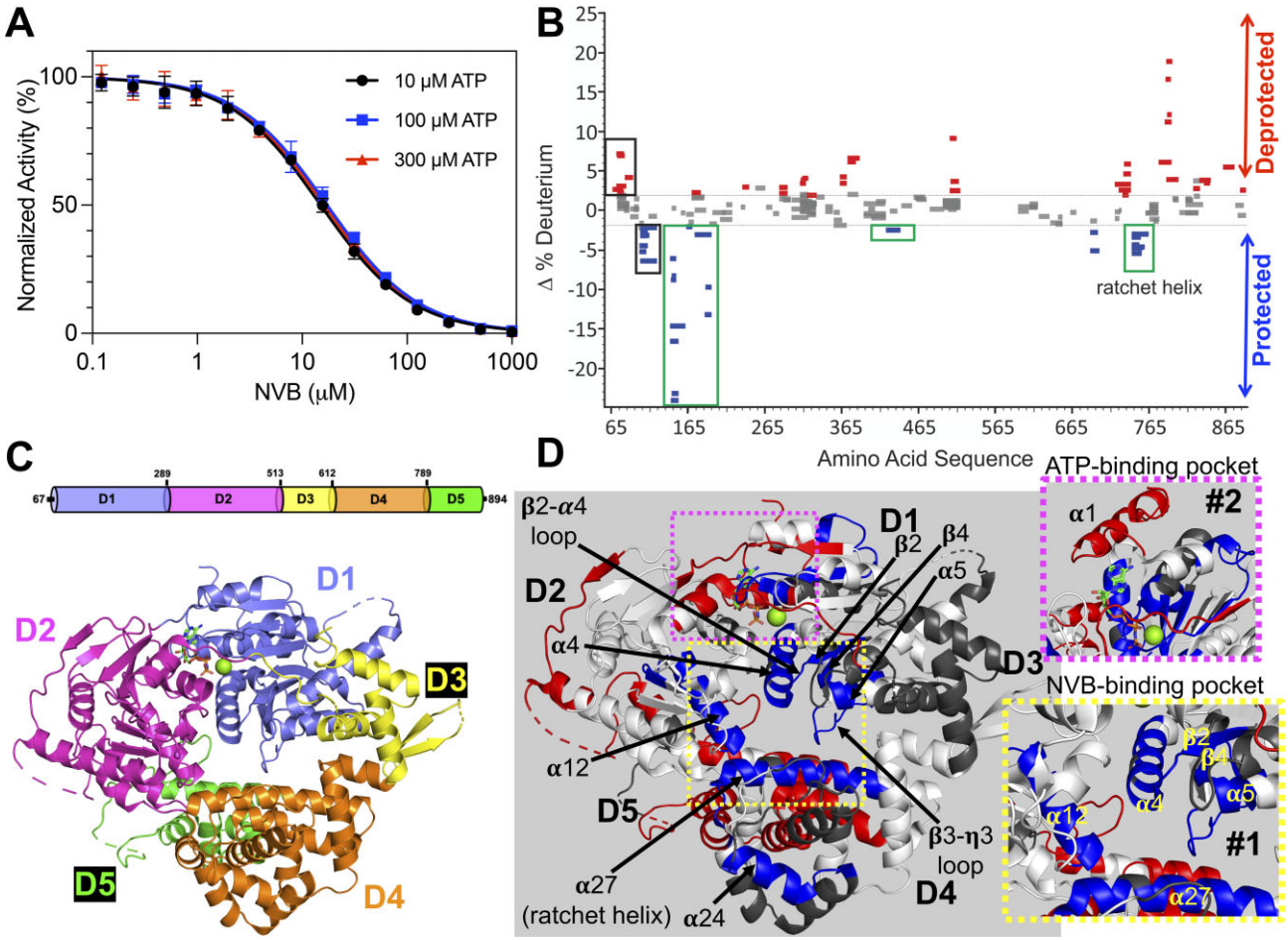


Figure 1. NVB mode of inhibition and binding site mapped onto the Polθ ATPase domain. (**A**) Inhibition of Polθ ATPase activity by NVB measured at the ATP concentrations of 10, 100 or 300 μM results in statistically similar IC_{50} values of 15 (95% CI 14–16), 17 (95% CI 16–18) or 16 (95% CI 15–18) μM, respectively. Normalized ATPase activity data represent the average response measured from three independent experiments and error bars show the standard deviation. (**B**) Woods plot showing the change in deuteration (Δ%D) in the Polθ ATPase domain in the presence of NVB compared to DMSO in hydrogen deuterium exchange-mass spectrometry (HX-MS) experiments ran in a nanoflow mode. Blue peptides show lower deuteration in the presence of NVB, whereas red peptides show higher deuteration. The threshold (dashed line) for significant Δ%D is defined as two standard deviations beyond the mean Δ%D values of all non-significant changes across all peptide replicates (unchanged peptides are shown in gray). Strong protection from exchange was observed in the regions boxed in green and both protection and deprotection were observed in the regions boxed in black. (**C**) Polθ ATPase domain sequence classified as five globular domains (D1–D5) and colored with corresponding domain boundaries. (**D**) Δ%D mapped onto the crystal structure of the Polθ ATPase domain (*blue*, lower; *red*, higher; *white*, unchanged; *black*, not recovered). Implicated NVB binding site in the Polθ ATPase core as guided by HX-MS data highlighted in **inset #1** (yellow dashed square) with the distinct nucleotide binding pocket highlighted in **inset #2** (pink dashed square). Key structural elements are highlighted with arrows (also see Supplementary Figure S2). D1–D5 represents different domains in the Polθ ATPase domain.

convergence of three domains (D1, D2 and D4, see domain definition in Figure 1C) and form a defined pocket within the interior of the protein (Figure 1D). Second, both protection and deprotection were observed in the regions boxed in black, which define the nucleotide binding site (Figure 1D). Third, the remaining perturbations formed a surprisingly extensive network of destabilizations that mapped to the periphery of the protein, except for small helix α24 on the face opposite the nucleotide binding site, which was stabilized. We previously observed that NVB generates a net stabilization of the Polθ ATPase domain (14). These findings are consistent with the increase in protease needed to achieve full digestion of the NVB-treated protein, relative to the drug-free control.

The pocket defined by the stabilizations most likely represents the binding site of NVB. Specifically, the site appeared to form at the junction of α4 and α5 helices from D1, α12 helix from D2 and α27 helix from D4 (**inset #1** in Figure 1D and see Supplementary Figure S2 for secondary structure element definition). Additionally, the β2 and β4 sheets in the D1 domain and β2-α4, β3-η3 loops surrounding this site also showed lower %ΔD. Interestingly, this site has been proposed previously to be the path for ssDNA binding in the Polθ ATPase domain, where the most extensive area of interface is formed by a 'ratchet helix' in the D4 domain (19) (here labeled as α27 in Figure 1D and highlighted in Figure 1B). The stabilization of the ratchet helix was accompanied by a particularly strong change in α4, which may be the axis along which nucleotide binding site is distorted (**inset #2** in Figure 1D). Notably, the previously proposed NVB binding site primarily overlaps with the destabilizations near the nucleotide binding site, suggesting it is not the

location for binding. Taken together, the HX-MS data indicate that NVB binds to an allosteric site in the Polθ ATPase domain, thereby impacting ATP hydrolysis and stabilizing the core of the domain.

ssDNA and NVB have overlapping binding sites

To confirm the DNA binding sites, we performed HX-MS analysis of the Polθ ATPase domain in the presence of ssDNA. First, we measured ssDNA substrate binding to the Polθ ATPase domain *in vitro* using microscale thermophoresis (MST). A 5′-Alexa647N-labelled ssDNA was used as a target in the MST experiment and an unlabeled Polθ ATPase domain was used as ligand, which was titrated into the target that was kept at a fixed-concentration. The Polθ ATPase domain bound to the ssDNA substrate with single digit nanomolar affinity in low-salt buffer conditions as measured by MST (Figure 2A). Additionally, we also observed that ssDNA stimulated ATPase activity of the Polθ ATPase domain by 6–8-fold at a given ATP concentration in the reaction (Supplementary Figure S3A).

Then we measured the changes in deuteration in the Polθ ATPase domain in the presence of the same ssDNA used in the MST and ATPase assays. We could achieve 77% protein coverage with a simpler and more conventional microflow HX-MS mode. We observed stabilizations in regions that strongly overlap with those induced by NVB. Yet, they were more extensive and mapped to a channel that extends to the core of the protein (Figure 2B and C). Stabilizations clustered around the contributions from the same three domains D1, D3 and D4. In addition to the 'ratchet helix' in domain D4, four other helices (α19, α20, α21 and α25) were also stabilized, supporting a much more extensive engagement of this domain compared to NVB. We also observed stabilization in the α4 and α5 helices from domain D1, and α16 from domain D3 (Figure 2C). Interestingly, these changes were not accompanied by the same strong perturbation of the nucleotide binding site, and the peripheral destabilizations of the protein were virtually nonexistent (Figure 2C). Furthermore, β2 through β5 sheets in domain D1 were also stabilized after ssDNA binding, indicating these changes may be sufficient to stimulate ATPase activity.

Next, we generated a Polθ ATPase domain-DNA binding model by structurally aligning the Polθ ATPase crystal structure to a closely-related DNA-bound Hel308 helicase (18). The DNA-bound model agreed with our HX-MS experimental data as most of the stabilized regions spanned around the path of the DNA in the model (Figure 2C). We then refined our previous molecular docking of NVB to the Polθ ATPase domain based on our nanoHX-MS data. A grid search centered around residue I191 from the N-terminus of the α5 helix in domain D1 resulted in two potential binding modes for NVB. Although the modeling appeared ambiguous as to the exact orientation (i.e. the sugar group pointing outside vs inside), the binding mode, in both orientations, was perpendicular to the path of incoming ssDNA (Figure 2D). Taken together, we conclude that NVB and ssDNA have overlapping binding sites in the Polθ ATPase domain.

NVB blocks polθ binding to ssDNA in cells and *in vitro*

To investigate whether NVB blocks ssDNA binding in cells, we tested the effect of NVB on the recruitment of Polθ to ssDNA in cells following DNA damage. First, we expressed a full-length N-terminus Myc-tagged Polθ in an endogenous *POLQ* knockout background. Transfected U2OS cells were pre-sensitized with 5-bromo-2′-deoxyuridine (BrdU) and treated with either DMSO or NVB at the indicated concentrations. Focused DNA damage was introduced in cells using laser micro-irradiation, and ssDNA at the laser stripes was imaged by immunofluorescence using a BrdU antibody. Simultaneously, Polθ abundance at these lesions was measured by staining for Myc-tag (Figure 3A). Immunofluorescence data showed that NVB reduced Polθ recruitment to ssDNA in cells in a dose-dependent manner (Figure 3B).

Recently, Polθ has been shown to seal post-replicative ssDNA gaps in HRD-deficient cells (55,56), whereas loss of Polθ resulted in the accumulation of unprocessed ssDNA gaps behind the replication forks. We reasoned that by blocking ssDNA binding, NVB would also affect the Polθ gap-sealing function. Therefore, we used a modified DNA fiber assay, which included S1 nuclease digestion of accumulated ssDNA. RPE1 *TP53*$^{-/-}$*BRCA1*$^{-/-}$ cells were sequentially labeled with the nucleoside analogs 5-chloro-2′-deoxyuridine (CldU) and 5-Iodo-2′-deoxyuridine (IdU) followed by S1 nuclease treatment (Figure 3C). Cells that were not treated with S1 nuclease were used as a control for ssDNA track length. Compared to vehicle treatment, NVB-treated cells showed a dramatic shortening in the IdU track lengths following S1 nuclease treatment (Figure 3D), suggesting accumulation of ssDNA gaps upon NVB treatment.

We also investigated the effect of NVB on Polθ ATPase domain binding directly to ssDNA *in vitro*. We inspected Polθ binding in the presence of NVB using MST, as described above for the low-salt buffer. Even at near-physiological conditions (120 mM KCl buffer), the purified Polθ ATPase domain bound to ssDNA with high affinity (4.4 ± 0.3 nM); however, addition of NVB in the MST buffer leads to a dose-dependent inhibition of Polθ binding to ssDNA (Figure 3E). In contrast, AMP-PNP, a non-hydrolysable analog of ATP, even at 2000 μM in the MST buffer was unable to block Polθ binding to ssDNA (Supplementary Figure S3B). Therefore, we concluded that NVB blocks Polθ binding to ssDNA and inhibits its key functions at DNA damage sites and ssDNA gaps in cells.

NVB sugar group defines its binding orientation

The HX-MS-guided molecular docking generated two very different orientations for NVB in the binding pocket, with the sugar group of NVB either pointing towards domains D1/D2 or towards domains D3/D4 (Figure 4A). To determine which of these orientations is correct, we compared NVB to novobiocic acid, an NVB derivative, that lacks the sugar group. We prepared novobiocic acid from NVB via base-catalyzed hydrolysis. First, we tested the effect of novobiocic acid on the thermal stability of Polθ ATPase domain with nanoDSF by measuring changes in the intrinsic fluorescence from tryptophan and tyrosine residues, as well as scattering, in the presence of the inhibitor or DMSO. As with NVB, novobiocic acid stabilized the Polθ ATPase

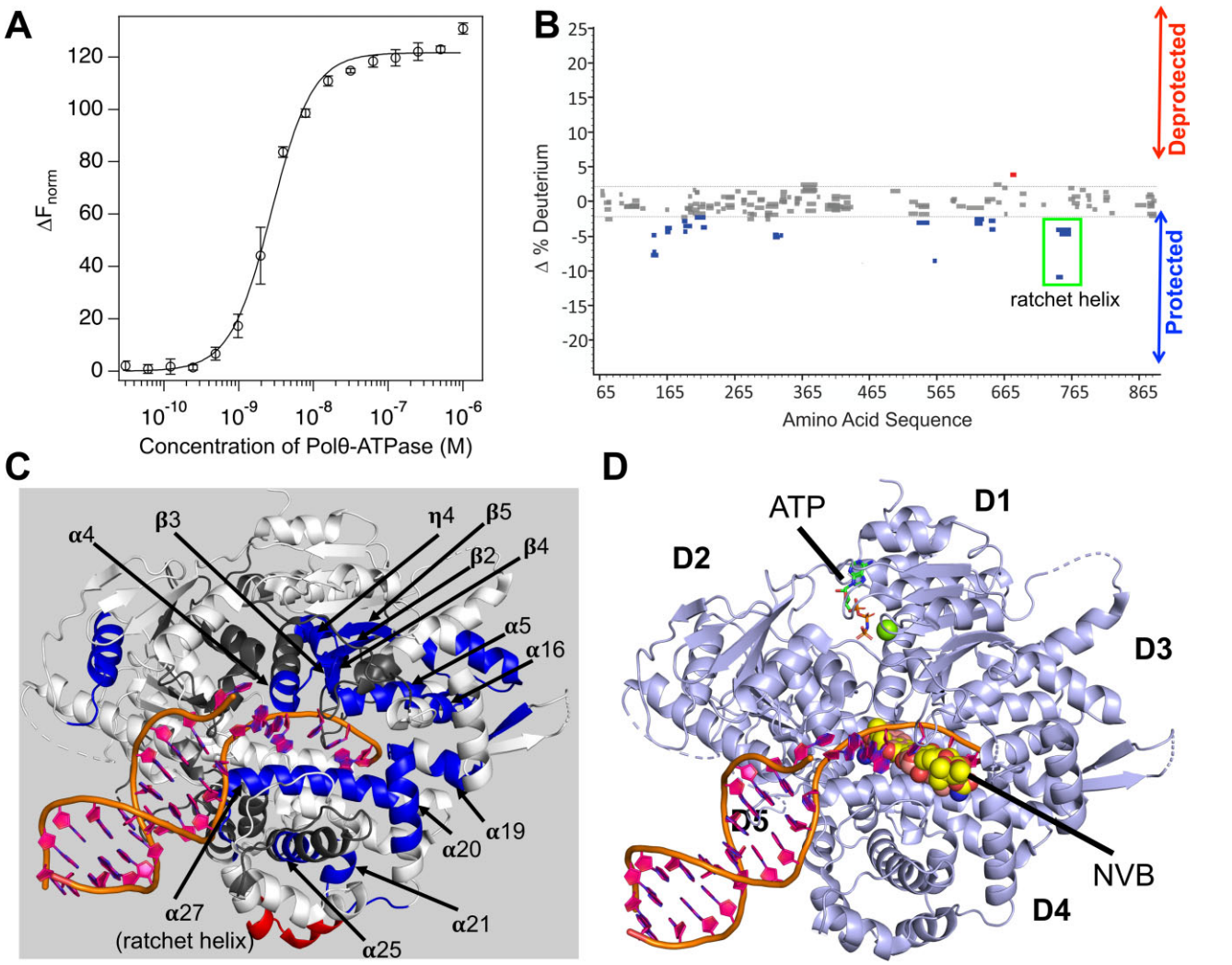


Figure 2. Mapping the ssDNA binding site in the Polθ ATPase domain**.** (**A**) Change in normalized fluorescence as result of thermophoresis in the MST experiment plotted as a function of concentration of unlabeled Polθ ATPase domain. The resulting curve represents ssDNA binding to Polθ ATPase domain in a low-salt buffer (50 mM KCl). The data points represent average of three independent measurements and error bars represent standard deviation. The measured EC_{50} value of Polθ ATPase domain binding to ssDNA is 2.7 ± 0.2 nM as estimated from the Hill fit of the averaged MST data. (**B**) Woods plot showing change in deuteration (Δ%D) in the Polθ ATPase domain in the presence of ssDNA compared to buffer alone in HX-MS experiment ran in a microflow mode. Blue peptides show lower deuteration in the presence of ssDNA, whereas red peptides show higher deuteration. The threshold (dashed line) for significant Δ%D is defined as two standard deviations beyond the mean Δ%D values of all non-significant changes across all peptide replicates (unchanged peptides are shown in gray). (**C**) Model of ssDNA-bound Polθ ATPase generated by structurally aligning Polθ ATPase domain with closely related DNA-bound Hel308 helicase (PDB ID:2P6R), Δ%D (*blue*, lower; *red*, higher; *white*, unchanged; *black*, not recovered) from HX-MS experiments and mapped onto the structure with key structural elements highlighted by arrows. For clarity, Hel308 is not shown in the model. (**D**) A model of the overlapping binding sites of NVB (HX-MS-guided docking) and ssDNA in the Polθ ATPase domain with D1–D5 labels for the different domains in the Polθ ATPase. NVB is shown as atom-colored spheres (center) away from the ATP (top left, sticks) in its binding site.

domain, indicating that novobiocic acid still binds the protein. However, the change in the scattering peak maximum was around +1°C, as opposed to +2°C in the presence of similar NVB concentration (Supplementary Figure S4A) (14). This increase in thermal stability could be partly due to the fact the NVB promotes tetramerization of the Polθ ATPase domain in solution, as measured by mass photometry (Supplementary Figure S4B). It is worth mentioning that mass photometry experiments were performed at a diluted concentration (25 nM), whereas the thermal stability experiments were performed at 5 μM concentration of the protein. Previous studies indicated that in solution the Polθ ATPase domain existed predominantly as tetramers in the concentration ranges 3.5–25 μM (19). Supporting our thermal stability measurements, we also observed a drop in the potency (as measured by IC_{50}) of novobiocic acid in the ATPase activity assays using ADP-Glo. We measured a four-fold weaker IC_{50} for novobiocic acid in assays performed at 100 μM ATP (Figure 4B). To test if the novobiocic acid can inhibit Polθ ATPase activity through the same mechanism as NVB, we measured Polθ binding to ssDNA in the presence of 1000 μM novobiocic acid. Using MST, we show that novobiocic acid can block ssDNA binding of Polθ *in vitro* (Supplementary Figure S4C). Additionally, we

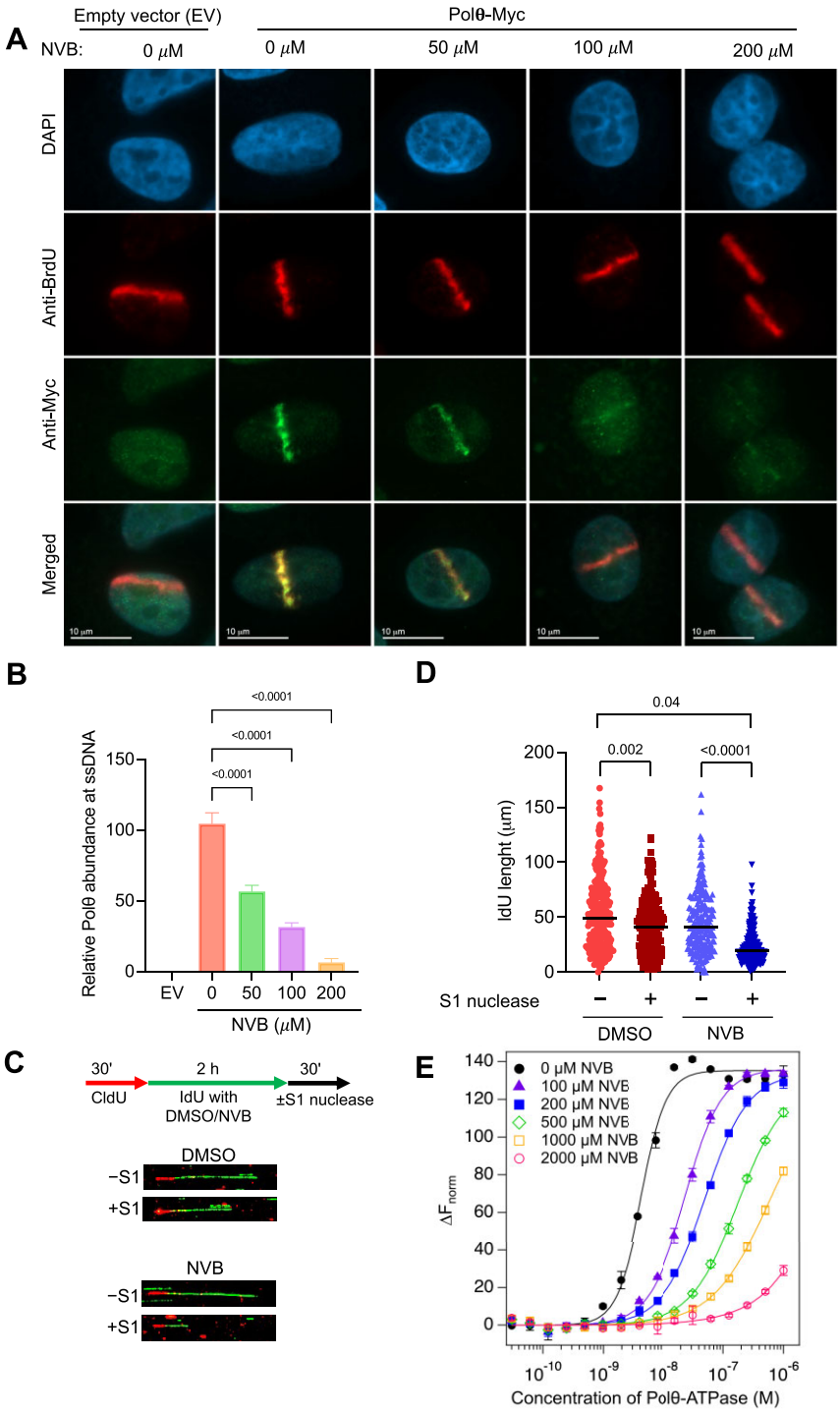


Figure 3. NVB blocks Polθ recruitment to ssDNA in cells and *in vitro*. (**A**) Representative images of immunofluorescence staining of ssDNA (labelled with anti-BrdU) and Polθ (labelled with anti-Myc) at laser micro-irradiated sites in U2OS cells in the presence of DMSO or NVB. Nuclei were stained with DAPI. Scale bar, 10 μm. (**B**) Quantification of relative Polθ abundance at ssDNA with variable concentrations of NVB. All data are presented as mean ± standard deviation from 100 cells for each treatment. Statistical significance was tested using one-way ANOVA for comparison of three or more sets. (**C**) *Top*, the schema for the CldU/IdU pulse-labeling protocol followed by S1 nuclease treatment; *bottom,* representative images of DNA fibers of RPE1 *TP53*$^{-/-}$*BRCA1*$^{-/-}$ cells exposed to DMSO or 200 μM NVB with or without S1 nuclease treatment. (**D**) Quantification of IdU track lengths in RPE1 *TP53*$^{-/-}$*BRCA1*$^{-/-}$ cells treated with DMSO or NVB with and without S1 nuclease treatment. Each dot represents 1 fiber; at least 200 fibers were quantified for each treatment and *P*-values were calculated using the Mann–Whitney test. (**E**) Change in normalized fluorescence as result of thermophoresis in the MST experiment, plotted as a function of concentration of unlabeled Polθ ATPase domain. The resulting curves represent ssDNA binding to Polθ ATPase domain in the absence (black solid circles) or presence (at the indicated concentrations) of NVB. The data points represent average of three independent measurements and error bars represent standard deviation. The measured EC_{50} values of Polθ ATPase domain binding to ssDNA are 4.4 ± 0.3, 23.3 ± 0.8 and 50 ± 2 nM for the NVB concentrations of 0, 100 and 200 μM, respectively, as estimated from the Hill fits of the averaged MST data. Binding constants were N.D. for the NVB concentrations of 500, 1000 and 2000 μM due to lack of saturation in the MST response curves.

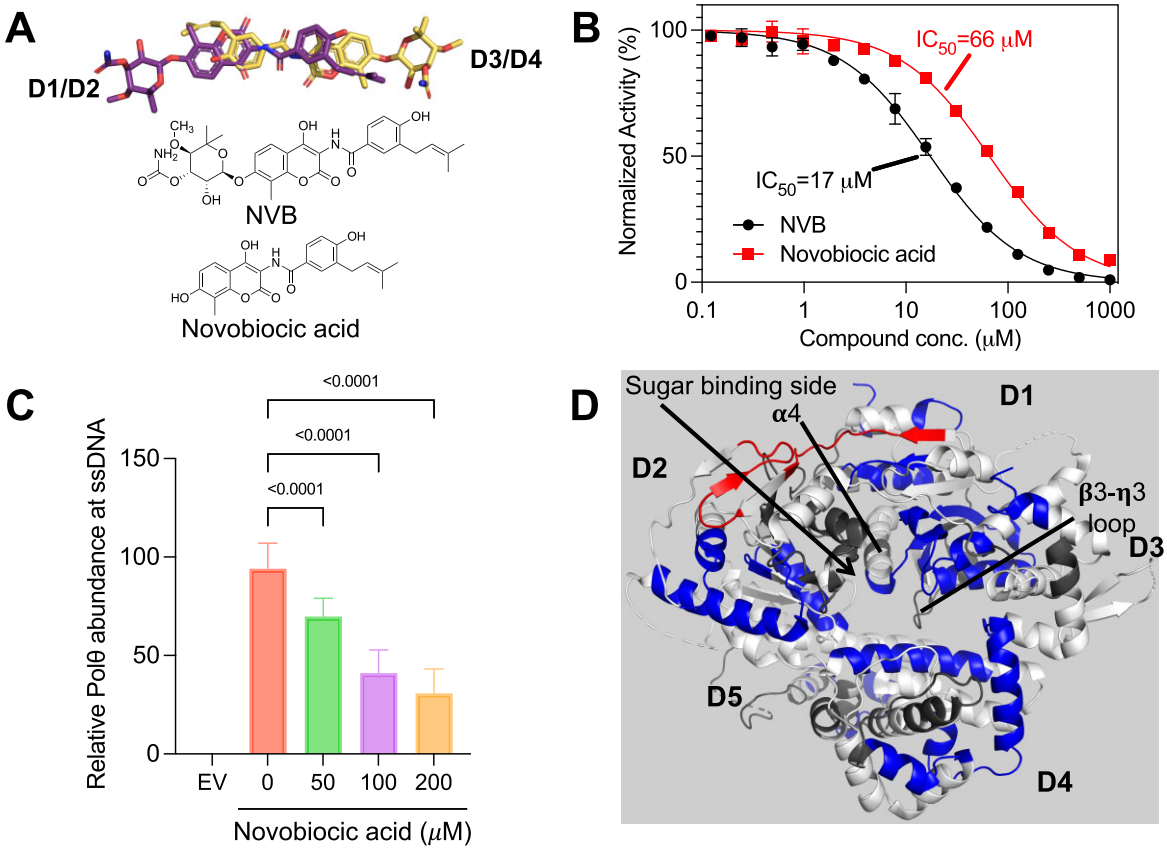


Figure 4. Sugar group of NVB impacts its potency. (**A**) Chemical structures of NVB and novobiocic acid. The two major potential orientations of NVB (shown in sticks) bound to Polθ ATPase domain predicted by HX-MS-guided molecular docking. The sugar group is pointing towards either D1/D2 or D3/D4. (**B**) Dose–response curves of NVB or novobiocic acid inhibition of ATPase activity at 100 μM ATP concentration with respective IC_{50}. Shown are mean ± standard deviation of three independent experiments. (**C**) Quantification of relative Polθ abundance at ssDNA with increasing concentrations of novobiocic acid from immunofluorescence micrographs of ssDNA and Polθ at laser micro-irradiated sites (see also Supplementary Figure S5). All data are presented as mean ± standard deviation from 100 cells. (**D**) HX-MS (operating in microflow mode) data on deuteration change (*blue*, lower; *red*, higher; *white*, unchanged; *black*, not recovered) in the presence of novobiocic acid mapped onto the Polθ ATPase structure. The key structural elements are highlighted with arrows (see also Supplementary Figure S6). D1–D5 represent different domains in the Polθ ATPase domain.

also measured novobiocic acid ability to block Polθ recruitment to DNA lesions created by laser micro-irradiation in U2OS cells, as performed for NVB. Like with NVB, Polθ abundance at the DNA damage site was impacted by novobiocic acid treatment in a dose-dependent manner (Figure 4C and Supplementary Figure S5). However, inhibition was not as effective as with NVB. At the highest dose of 200 μM, NVB blocked ∼96% of Polθ from being recruited to damage sites, whereas novobiocic acid only blocked ∼70%. Together, these assays show that novobiocic acid can still bind and inhibit Polθ through the same mechanism as NVB, but with an attenuated potency.

Finally, HX-MS was used to map the binding site of novobiocic acid. We could achieve 79% protein coverage with HX-MS in microflow mode, sufficient to compare results with the NVB data (Supplementary Figure S6). A shift in stabilization profile was generated, like that generated from ssDNA binding, indicating binding to the ATPase core. Domain D1 was markedly more stable when binding to novobiocic acid. In contrast to NVB binding, the influence on nucleotide binding site was muted upon novobiocic acid binding to the ATPase core and very little peripheral destabilization was observed. Notably, we only observed a partial stabilization of α4 helix in domain D1 compared to NVB-bound ATPase core. These combined measurements are consistent with computational docking, which placed sugar group binding towards D1/D2, and with the observation that loss of this key group affects NVB binding orientation and potency.

POLQ is overexpressed in cancer and associated with poor survival

Having defined its mechanism-of-action, we reasoned that the use of NVB for cell biology and medicine may be informed by the possible impact of *POLQ* overexpression on survival of patients with cancer. We therefore compared the levels of Polθ mRNA in TCGA tumors and matched controls. In all 15 types of tumors that were analyzed, *POLQ* mRNA levels were higher in tumor samples than in matched controls, with p-values as low as 2.9×10^{-51} in breast cancer (Figure 5A), showing that *POLQ* is consistently and strongly upregulated in cancer.

To evaluate the possible impact of *POLQ* overexpression on patient survival, for each of the 33 TCGA tumor types we applied the Kaplan–Meier estimator to two groups

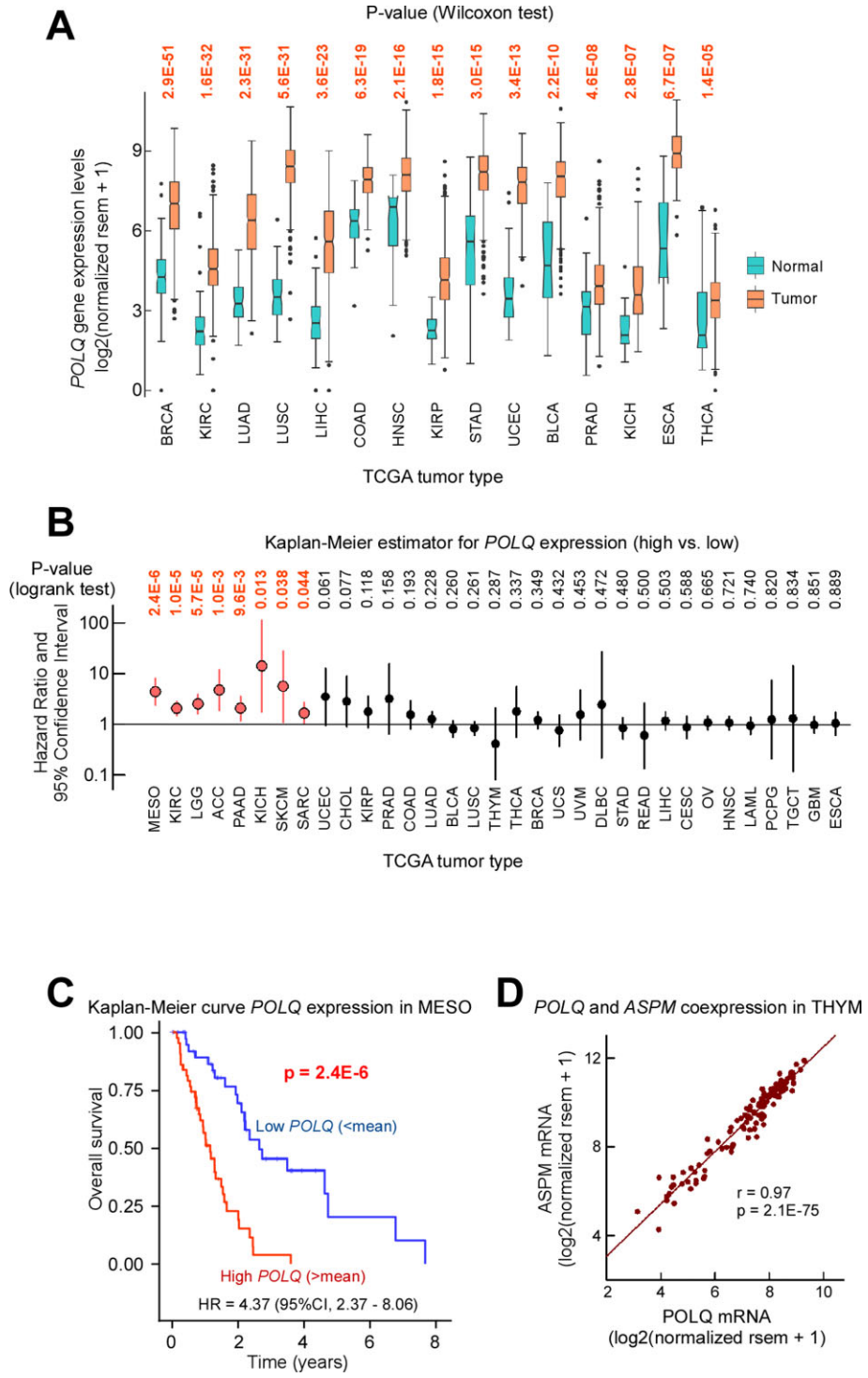


Figure 5. *POLQ* is overexpressed in cancer, and is associated to poor survival and high cell proliferation**.** (**A**) Box plot of *POLQ* mRNA levels in tumors and matched controls from TCGA ranked by *P*-value. Only tumor types with >10 control samples were analyzed. *P*-values from Wilcoxon rank sum tests. (**B**) Dot plot of hazard ratios (HR) and 95% confidence intervals (CI) from Kaplan–Meier survival curves for high (above the mean) versus low (below the mean) *POLQ* mRNA levels in TCGA tumor types, ranked by *P*-value. *P*-values from logrank tests. *Red*, increased risk for patients with high *POLQ* expression levels. (**C**) Example of Kaplan–Meier survival curve from the analyses in panel B. (**D**) Dot plot of correlation between mRNA levels of *POLQ* (X-axis) and *ASPM* (Y-axis), which regulates microtubule dynamics at spindle poles, in THYM. *MESO*, mesothelioma; *THYM*, thymoma; *ASPM*, assembly factor for spindle microtubules.

of patients, g_high, which contained tumor samples with *POLQ* expression above the mean, and g_low, which contained tumor samples with *POLQ* expression below the mean. A plot of the ensuing hazard ratios (HR) showed that in eight types of tumors (i.e. mesothelioma (MESO), kidney renal clear cell carcinoma (KIRC), brain lower grade glioma (LGG), adrenocortical carcinoma (ACC), pancreatic adenocarcinoma (PAAD), kidney chromophobe (KICH), skin cutaneous melanoma (SKCM) and sarcoma (SARC)), the g_high group incurred significantly higher risk of poor prognosis than the g_low group (Figure 5B), with a decrease in life expectancy of ∼1.5 years in half of patients with mesothelioma (Figure 5C). We were not able to obtain information on Polθ protein levels in TCGA tumors with matched controls since it was not targeted for reverse-phase protein array (RPPA) analyses in the TCGA project; however, a few experimental studies suggest that knocking down Polθ delays tumor growth (57–59). These data support the view that *POLQ* upregulation in cancer is part of a program aimed at sustaining tumor fitness.

POLQ overexpression is linked to defined mutational signatures and cell division

To learn more about the mechanisms by which *POLQ* overexpression might support tumor fitness, we conducted a comprehensive analysis of mutational signatures in TCGA tumors. The landscape of single base substitutions in a tumor originates from various mutational processes, each of which leaves its own signature (60). Therefore, we reasoned that a comparison between the g_high and g_low groups might inform on mechanisms of mutagenesis linked to *POLQ* overexpression. Our analysis showed that mutational signatures overall tended to occur more frequently in the g_high group than in the g_low group, and that g_high patients across various tumor types were likely to incur a surge in mutations arising from both deficient homologous recombination repair (signature 3) and error-prone repair following enzymatic deamination of cytosine (signature 13) (Supplementary Figure S7A).

As Polθ activity is unlikely to contribute to both signatures 3 and 13 solely on its own, we extended the mutational signature analysis to all available genes (∼20500) in the 33 TCGA tumor types and extracted all instances in which the g_high group for a given gene was strongly (p-value < 0.001) associated with a surge in both signatures 3 and 13 relative to the g_low group. The analysis returned a set of 1434 unique genes in a total of 7 tumor types (i.e. breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), brain lower grade glioma (LGG), lung adenocarcinoma (LUAD), stomach adenocarcinoma (STAD) and uterine corpus endometrial carcinoma (UCEC)). Gene Set Enrichment Analysis (GSEA) for this 1434 gene set revealed strong enrichment in components linked to cell cycle progression, mitosis and cell division, cellular structures, such as the kinetochore, which mediate cell division, ubiquitin conjugation, DNA replication, DNA repair and DNA damage responses (Supplementary Figure S7B, also see Supplementary Table S1). Signature 3 occurs frequently in breast cancer. Therefore, to assess whether the gene enrichment pattern observed was due to the BRCA samples alone, we repeated GSEA after removing the genes found in BRCA, which left a set of 501 genes. The result confirmed enrichment of the same category terms but with weaker p-values, as expected. We interpret these findings to imply that tumor fitness stems from a stimulation in DNA repair processes responding to an increase in cell proliferation, perhaps through the DNA replication stress response pathway at damaged DNA replication forks (61,62). Among others, this response demands Polθ activity, so its ability to inhibit DNA binding and stimulation of Polθ ATPase makes NVB an informative probe for Polθ ATPase functions.

To learn more about the link between *POLQ* expression and cell proliferation, we conducted a comprehensive co-expression analysis between *POLQ* and all available genes (∼20500) in the 33 TCGA tumor types, selected all results containing at least 100 samples, ranked the data by regression coefficient, selected the top 200 entries and identified from this list all unique genes, 119 total (Supplementary Table S1). GSEA of these 119 genes revealed the same profile as the mutational signature analysis, i.e. strong enrichment in genes involved in cell cycle and the kinetochore complex (Supplementary Figure S8), with stunning regression coefficients approaching unity (Figure 5D). Interestingly, a GSEA conducted on the top 100 *POLQ* coexpressed genes from ARCHS4 (Supplementary Table S1), a web-based resource compiling data from a wide set of RNA-seq projects (63), returned the same gene ontology category terms (Supplementary Figure S8). These composite meta-analyses raise the possibility that Polθ may be a component of the kinetochore structure, an exemplary prediction that may now be tested with Polθ inhibitors, including NVB.

DISCUSSION

Unrepaired DSBs can give rise to large-scale genome rearrangements, chromosomal translocations, mutagenesis and cell death, so DSB repair pathways have high biological and medical significance. Polθ plays a central role in the most enigmatic DSB repair pathway, variously termed TMEJ, MMEJ or Alt-EJ. We do know that TMEJ is mechanistically distinct from the major HDR and NHEJ DSB repair pathways and that critical TMEJ factors in mammalian cells are XRCC1, MRE11, PARP1, DNA ligase I or III, and Polθ (12,64). Although it is an inherently mutagenic DNA repair pathway, TMEJ is conserved in both archaea and eukaryotes, as is its associated DSB repair and replication fork nuclease MRE11, underscoring its fundamental biological importance (25,29,65). Besides its biological interest, Polθ has a role in shaping cancer genomes with characteristic indels that can promote PARP inhibitor therapeutic resistance (22,66,67). Moreover, tumors with mutations in cancer susceptibility genes BRCA1 and BRCA2 that cause HDR-deficiency are addicted to Polθ-mediated TMEJ (12,13). Given its potential for cancer-associated synthetic lethality and to overcome therapy resistance, Polθ is a high-priority target for precision cancer therapy (68), as evidenced by preclinical studies targeting its ATPase and the polymerase domains with small molecules (14,15,30). Our analysis furthermore supports and extends the observations that Polθ is overexpressed in cancer and associated

with poor survival, with links to defined mutational signatures and DNA repair processes responding to increased cell proliferation and replication stress.

Given these combined observations, TMEJ may also be a target for improving radiation therapy, as it is activated in irradiated human cells due to phosphorylation-dependent formation of the XRCC1 repair complex with the end-resection enzyme MRE11. Indeed, MRE11 endonuclease and exonuclease activities are required for TMEJ and for HDR (64,69). Furthermore, knockout of the GRB2 adaptor protein that efficiently targets MRE11 for HDR resulted in increased MRE11-XRCC1 complex and increased TMEJ (70). As XRCC1 forms an active repair complex with MRE11, Polθ and DNA ligase IIIα, which together support TMEJ *in vitro* (23,71), these combined observations suggest that XRCC1 adaptor complexes with MRE11 and Polθ efficiently initiate TMEJ. Furthermore, Polθ was shown to seal the post replicative gaps in HDR-deficient tumors (55,56). Notably, BRCA2 limits XRCC1 recruitment to suppress TMEJ at stalled forks. Without BRCA2 fork protection, XRCC1 complex with MRE11-Polθ-ligase IIIα may enable cells to complete DNA replication at the expense of increased genome instability (23). In fact, Polθ depletion or inhibition by NVB decreased radioresistance in lung adenocarcinoma while causing little toxicity to normal pulmonary epithelial cells (72).

We previously found that the ATPase-dead mutant of Polθ is still recruited to the damage sites in cells and the recruitment is only partially affected (14). Thus, inhibition of the ATPase activity cannot account for the complete loss of Polθ recruitment to the damage sites. Therefore, based on this previous observation and our current *in vitro* data showing the NVB inhibition of Polθ binding to ssDNA, we conclude that NVB inhibits ssDNA-mediated stimulation of Polθ ATPase activity as well as recruitment to DNA damage sites. However, persisting RPA and RAD51 filaments may also play a role in further blocking Polθ from being recruited to key DNA damage sites. Integrating our finding that NVB is a non-ATP competitive inhibitor that binds to the ssDNA-binding site in the Polθ ATPase core with existing data, we propose the following synthesis in the context of TMEJ. For its DNA damage repair functions, Polθ must be recruited to resected DSBs with ssDNA overhangs or ssDNA gaps formed by replication stress, to complete TMEJ or gap-sealing, respectively (Figure 6). In HDR-deficient tumors, Polθ can rescue ill-fated HDR commitments by displacing ssDNA binding proteins, such as RPA and RAD51, and channeling repair to TMEJ. DNA sequence microhomologies for Polθ binding are revealed by cellular nucleases such as MRE11, whose activity is essential for this pathway (64). Polθ ATPase anneals complementary sequences, producing a flapped substrate which may be removed by the flap endonuclease FEN1 (73,74). However, as shown by our data, NVB blocks Polθ binding to ssDNA, thereby both preventing Polθ from efficient recruitment to these crucial ssDNA sites and blocking its ATPase activation to prevent downstream DNA damage repair (Figure 6).

The crystal structure of Polθ ATPase core revealed several deep druggable pockets in the protein (19). The extended conformation of NVB seems ideal for targeting these deep pockets. However, in the DNA gyrase-NVB crystal structures, NVB directly binds the ATP pockets in a bent conformation (43–45). Here, our biochemical data indicate that NVB is a non-ATP competitive inhibitor of Polθ ATPase activity. Using a combination of HX-MS, MST, cellular assays and computational modelling, we determined that NVB binds to an allosteric site orthogonally to the ssDNA binding site of the ATPase core to inhibit ssDNA-stimulated Polθ ATPase activity. Importantly, we find that NVB blocks Polθ binding to ssDNA *in vitro* and in cells. Using the aglycone novobiocic acid, we investigated the correct orientation of NVB binding and established the sugar group's contribution to the overall inhibitor potency.

During our manuscript preparation, an independent study deposited to the bioRxiv portal reported a cryo-EM structural model of NVB-bound Polθ ATPase domain (75). These cryo-EM data support our findings and confirm that NVB is a non-ATP competitive inhibitor. Indeed, NVB binds to a distal non-canonical site (near the ssDNA binding site) in the ATPase core. In future studies, it will now be interesting to see if NVB binding involves any DNA mimicry or binding site remodeling or just steric blocking of ssDNA binding (76–78). Our findings here support and extend the cryo-EM complex results by independently identifying the binding pocket and by providing biophysical and biochemical data on NVB interactions and its competition with ssDNA *in vitro* and in cells.

Mechanistically, NVB is a well-characterized inhibitor of prokaryotic DNA gyrase and topoisomerase. Structural analyses of NVB with several prokaryotic gyrases and topoisomerases revealed an ATP-competitive nature of NVB mechanism-of-action (44,79). Further, NVB has been proposed to be a weak inhibitor of eukaryotic topoisomerase II and HSP90 (80–82). Although there is no structural information on these eukaryotic enzymes with NVB, based on the structural analogy, it has been proposed that NVB works through a similar ATP-competitive mechanism as DNA gyrase. However, NVB was also shown to target beyond the ATP-binding site in an *E coli* ATPase (LptB) involved in lipopolysaccharide (LPS) transport (83). The structural analyses of LptB-NVB-ADP ternary complex revealed that NVB binds to a distal interface between the ATPase and the transmembrane subunits of LPS transporter. Here, the NVB binding to LptB activates the ATPase. Recently, NVB was also shown to bind to the human autophagy Atg8 family adaptor protein LC3A, which anchors to the lipid bilayer of autophagosome (84). Interestingly, in contrast to Polθ inhibition, the removal of the sugar group in NVB improved the potency for LC3A. Together, these data show that NVB binding is not limited to ATP binding pockets.

Given that the Polθ belongs to the SF2-type family of ATPases with several key conserved motifs from this family, we checked other SF2-type ATPases as potential targets for NVB. Based on the HX-MS data, NVB targets an allosteric site (formed at the junction of D1, D2 and D4 domains in Polθ) near the ssDNA binding site. However, only one such motif (motif Ia: 140–146 aa from D1 in Polθ) is part of the NVB binding site from the HX-MS data. Notably, there is very little sequence conservation beyond these motifs between Polθ and other SF2-type ATPases (19,85). This

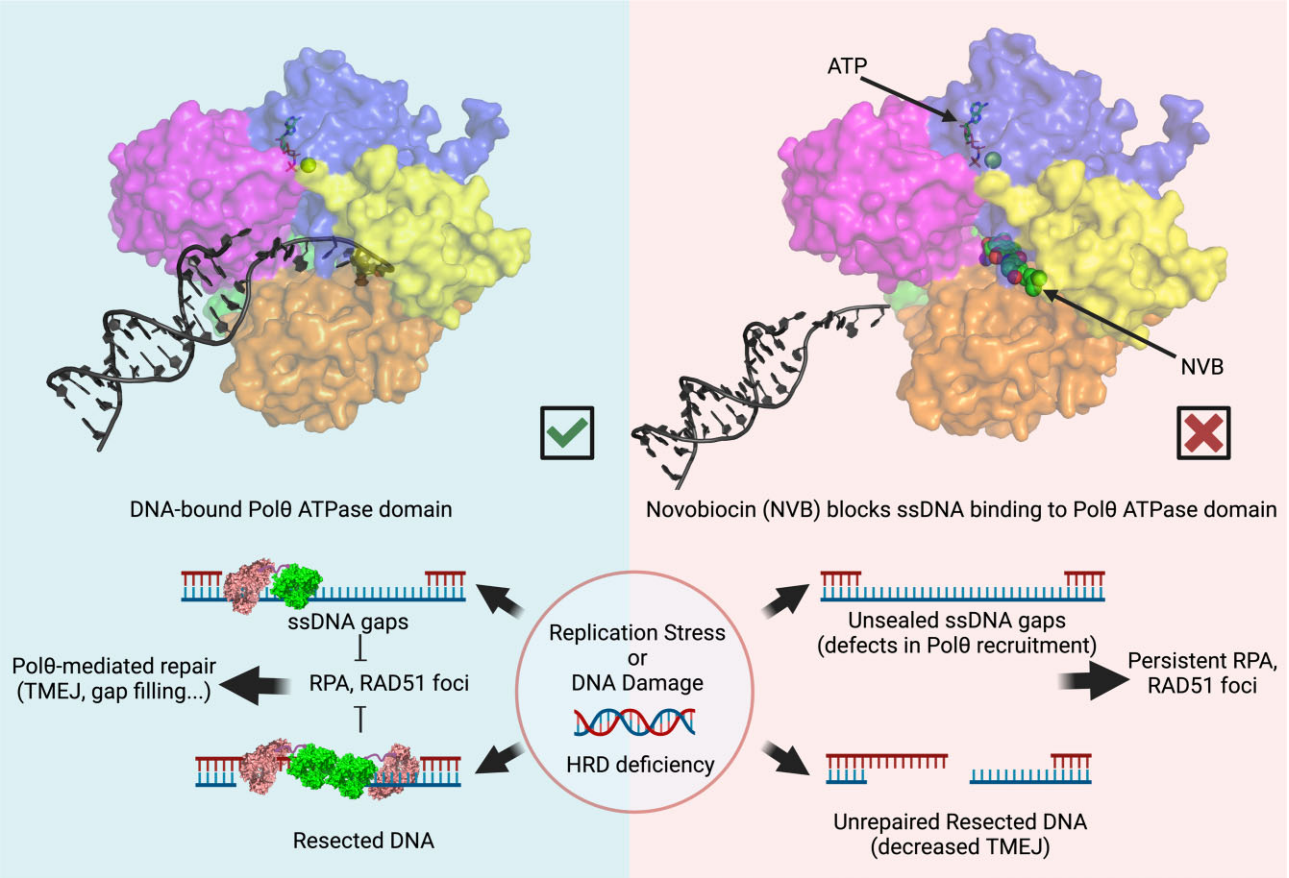


Figure 6. NVB inhibits Polθ by blocking both ssDNA binding and stimulation of its ATPase activity**.** In HDR-deficient tumors, Polθ plays an essential role in processing ssDNA gaps and resected DSBs and thus inhibiting the build-up of unresolved RPA and RAD51 foci. Both ATPase and polymerase domains are required for this activity. NVB inhibits Polθ recruitment to these key sites by blocking ssDNA binding and stimulation of its ATPase activity.

observation is further corroborated by our prior studies of NVB inhibition of SF2-type ATPases BLM, CHD1 and SMARCAL1 (14). We measured NVB inhibition of Polθ over these SF2-type ATPases. HELQ is the only other SF2-type ATPase that is structurally related to Polθ and maintains the sequence conservation beyond the canonical motifs (based on the AlphaFold2 predicted structure of HELQ for the Uniprot accession number Q8TDG4). Yet, HELQ lacks the key residues K151 and T175 that are shown to interact with the sugar of NVB (75). Thus, we reason that NVB likely inhibits Polθ more potently than other SF2-type ATPases. However, the potential of off-target binding beyond SF2-type ATPases underscores the value of using the NVB binding site of Polθ to improve drug potency.

NVB is effective in preclinical models in achieving cancer-specific killing (14,15,30); yet, it functions *in vitro* by competing with high-affinity nucleic acid binding by the Polθ ATPase core. Notably in the cell, there is strong regulation and competition among ssDNA binding proteins, so even modest modulation of Polθ DNA binding and ATPase activity may affect its DNA repair activity. In the cell, we think for example that Polθ function is balanced by RAD51C-XRCC3 stabilizing RAD51 filaments, that DNA blocking inhibitors can prevent proteins from attaining their high affinity DNA-binding conformation, and that inactive protein binding to DNA damage can block its canonical repair and enforce an alternative repair pathway (86–88).

The NVB binding site identified here may facilitate development of NVB derivatives with improved potency as the drug enters into early phase clinical trials (e.g. NCT05687110). Specifically, our HX-MS data point to a promising an allosteric binding channel that overlaps with the DNA binding site that could be targeted. Our findings, and the complementary cryo-EM structure, characterize NVB's mechanism-of-action and should aid in the design of NVB derivatives with improved potency. Given that we identified the sugar group as critical for binding orientation and potency, design efforts focused on modifying the coumarin tail with the sugar group intact (or replaced with equally potent bioisosteres) may be a good starting point (82). As an inhibitor with a defined mechanism-of-action, NVB provides a means to probe the underlying mechanism of TMEJ, to improve understanding of the source of oncogenic instability and to guide the clinical path for Polθ inhibitors beyond HDR-mutated tumors for added effective treatment strategies for cancer.

DATA AVAILABILITY

The TCGA RNA-seq normalized rsem data, TCGA clinical data and mutational signature files used for the

analyses are available at https://doi.org/10.5281/zenodo.7885656 (‘*tumortype*rnaSeq.tar’, ‘cliDat_tcga_18.tar’ and ‘signatureProfileSample.txt.zip’ files); source codes for TCGA analyses (tcgaAnalyses and sumbitMPI) are available at https://doi.org/10.5281/zenodo.7874703 and https://doi.org/10.5281/zenodo.7874772, respectively. Standard data export files for HX-MS analyses are included in the Supplementary HX-MS data folder. Any additional information required to reanalyze the data reported in this paper is available from John A. Tainer (jtainer@mdanderson.org) upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Center for Macromolecular Interaction (CMI) at Harvard Medical School for access to their Monolith NT.115Pico system (NanoTemper) and Refeyn TwoMP mass photometry (Refeyn, Ltd) instruments. We acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for HPC resources key to the reported computational research results (URL: http://www.tacc.utexas.edu). We are grateful to Ms. Kayleigh E. Fay (CMI) for aiding Mass Photometry data analysis, Dr. Zamal Ahmed for helpful advice on MST experiments, Dr. Davide Moiani for helpful advice on docking and Dr. Shrabasti Roychoudhury for helpful feedback on the manuscript. The illustration shown in the Figure 6 was created by Biorender (www.Biorender.com).

FUNDING

National Institute of Health [P01 CA092584 to J.A.T., D.C.S., M-S.T., D.E.J., J.A.N., in part]; National Cancer Institute [R35 CA220430 to J.A.T.]; Cancer Prevention and Research Institute of Texas [RP180813 to J.A.T.]; Robert A. Welch Chemistry Chair (G-0010 to J.A.T.); Lustgarten Foundation/Stand Up To Cancer Pancreatic Cancer Challenge grant (to A.D.D.); Breast Cancer Research Foundation (to A.D.D.); Ludwig Center at Harvard (to A.D.D.). Funding for open access charge: Robert A. Welch Chemistry Chair (G-0010).

Conflict of interest statement. The authors declare no competing interest. A.D.D. reports consulting for AstraZeneca, Bayer AG, Blacksmith/Lightstone Ventures, Bristol Myers Squibb, Cyteir Therapeutics, EMD Serono, Impact Therapeutics, PrimeFour Therapeutics, Pfizer, Tango Therapeutics and Zentalis Pharmaceuticals/Zeno Management; is an Advisory Board member for Cyteir and Impact Therapeutics; stockholder in Cyteir, Impact Therapeutics and PrimeFour Therapeutics; and reports receiving commercial research grants from Bristol Myers Squibb, EMD Serono, Moderna and Tango Therapeutics.

REFERENCES

1. Weinstein,I.B. (2002) Addiction to oncogenes–the Achilles heal of cancer. *Science*, **297**, 63–64.
2. Kantarjian,H., Sawyers,C., Hochhaus,A., Guilhot,F., Schiffer,C., Gambacorti-Passerini,C., Niederwieser,D., Resta,D., Capdeville,R., Zoellner,U. *et al.* (2002) Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *N. Engl. J. Med.*, **346**, 645–652.
3. Vogel,C.L., Cobleigh,M.A., Tripathy,D., Gutheil,J.C., Harris,L.N., Fehrenbacher,L., Slamon,D.J., Murphy,M., Novotny,W.F., Burchmore,M. *et al.* (2002) Efficacy and safety of Trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.*, **20**, 719–726.
4. Farmer,H., McCabe,N., Lord,C.J., Tutt,A.N.J., Johnson,D.A., Richardson,T.B., Santarosa,M., Dillon,K.J., Hickson,I., Knights,C. *et al.* (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, **434**, 917–921.
5. Bryant,H.E., Schultz,N., Thomas,H.D., Parker,K.M., Flower,D., Lopez,E., Kyle,S., Meuth,M., Curtin,N.J. and Helleday,T. (2005) Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*, **434**, 913–917.
6. Muller,F.L., Colla,S., Aquilanti,E., Manzo,V.E., Genovese,G., Lee,J., Eisenson,D., Narurkar,R., Deng,P., Nezi,L. *et al.* (2012) Passenger deletions generate therapeutic vulnerabilities in cancer. *Nature*, **488**, 337–342.
7. Zhao,D., Lu,X., Wang,G., Lan,Z., Liao,W., Li,J., Liang,X., Chen,J.R., Shah,S., Shang,X. *et al.* (2017) Synthetic essentiality of chromatin remodelling factor CHD1 in PTEN-deficient cancer. *Nature*, **542**, 484–488.
8. da Costa,A.A.B.A., Chowdhury,D., Shapiro,G.I., D’Andrea,A.D. and Konstantinopoulos,P.A. (2022) Targeting replication stress in cancer therapy. *Nat. Rev. Drug Discov.*, **22**, 38–58.
9. Harris,P.V., Mazina,O.M., Leonhardt,E.A., Case,R.B., Boyd,J.B. and Burtis,K.C. (1996) Molecular cloning of drosophila mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. *Mol. Cell. Biol.*, **16**, 5764–5771.
10. Muzzini,D.M., Plevani,P., Boulton,S.J., Cassata,G. and Marini,F. (2008) Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. *DNA Repair (Amst.)*, **7**, 941–950.
11. Shima,N., Munroe,R.J. and Schimenti,J.C. (2004) The mouse genomic instability mutation chaos1Is an allele of Polq that exhibits genetic interaction with atm. *Mol. Cell. Biol.*, **24**, 10381–10389.
12. Ceccaldi,R., Liu,J.C., Amunugama,R., Hajdu,I., Primack,B., Petalcorin,M.I.R., O’Connor,K.W., Konstantinopoulos,P.A., Elledge,S.J., Boulton,S.J *et al.* (2015) Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. *Nature*, **518**, 258–262.
13. Mateos-Gomez,P.A., Gong,F., Nair,N., Miller,K.M., Lazzerini-Denchi,E. and Sfeir,A. (2015) Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. *Nature*, **518**, 254–257.
14. Zhou,J., Gelot,C., Pantelidou,C., Li,A., Yücel,H., Davis,R.E., Färkkilä,A., Kochupurakkal,B., Syed,A., Shapiro,G.I. *et al.* (2021) A first-in-class polymerase theta inhibitor selectively targets homologous-recombination-deficient tumors. *Nature Cancer*, **2**, 598–610.
15. Zatreanu,D., Robinson,H.M.R., Alkhatib,O., Boursier,M., Finch,H., Geo,L., Grande,D., Grinkevich,V., Heald,R.A., Langdon,S. *et al.* (2021) Polθ inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. *Nat. Commun.*, **12**, 3636.
16. Seki,M. (2003) POLQ (Pol), a DNA polymerase and DNA-dependent ATPase in human cells. *Nucleic Acids Res.*, **31**, 6117–6126.
17. Arana,M.E., Seki,M., Wood,R.D., Rogozin,I.B. and Kunkel,T.A. (2008) Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.*, **36**, 3847–3856.
18. Büttner,K., Nehring,S. and Hopfner,K.-P. (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struct. Mol. Biol.*, **14**, 647–652.
19. Newman,J.A., Cooper,C.D.O., Aitkenhead,H. and Gileadi,O. (2015) Structure of the helicase domain of DNA polymerase theta reveals a possible role in the microhomology-mediated end-joining pathway. *Structure*, **23**, 2319–2330.
20. Zahn,K.E., Averill,A.M., Aller,P., Wood,R.D. and Doublié,S. (2015) Human DNA polymerase θ grasps the primer terminus to mediate DNA repair. *Nat. Struct. Mol. Biol.*, **22**, 304–311.

21. Seki,M., Masutani,C., Yang,L.W., Schuffert,A., Iwai,S., Bahar,I. and Wood,R.D. (2004) High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J.*, **23**, 4484–4494.
22. Ramsden,D.A., Carvajal-Garcia,J. and Gupta,G.P. (2021) Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. *Nat. Rev. Mol. Cell Biol.*, **23**, 125–140.
23. Eckelmann,B.J., Bacolla,A., Wang,H., Ye,Z., Guerrero,E.N., Jiang,W., El-Zein,R., Hegde,M.L., Tomkinson,A.E., Tainer,J.A. *et al.* (2020) XRCC1 promotes replication restart, nascent fork degradation and mutagenic DNA repair in BRCA2-deficient cells. *NAR Cancer*, **2**, zcaa013.
24. Ceccaldi,R., Rondinelli,B. and D'Andrea,A.D. (2016) Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.*, **26**, 52–64.
25. Syed,A. and Tainer,J.A. (2018) The MRE11–RAD50–NBS1 complex conducts the orchestration of damage signaling and outcomes to stress in DNA replication and repair. *Annu. Rev. Biochem.*, **87**, 263–294.
26. Mateos-Gomez,P.A., Kent,T., Deng,S.K., McDevitt,S., Kashkina,E., Hoang,T.M., Pomerantz,R.T. and Sfeir,A. (2017) The helicase domain of Polθ counteracts RPA to promote alt-NHEJ. *Nat. Struct. Mol. Biol.*, **24**, 1116–1123.
27. Schaub,J.M., Soniat,M.M. and Finkelstein,I.J. (2022) Polymerase theta-helicase promotes end joining by stripping single-stranded DNA-binding proteins and bridging DNA ends. *Nucleic Acids Res.*, **50**, 3911–3921.
28. Patterson-Fortin,J. and D'Andrea,A.D. (2020) Exploiting the microhomology-mediated end-joining pathway in cancer therapy. *Cancer Res.*, **80**, 4593–4600.
29. Sfeir,A. and Symington,L.S. (2015) Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway?*Trends Biochem. Sci*, **40**, 701–714.
30. Bubenik,M., Mader,P., Mochirian,P., Vallée,F., Clark,J., Truchon,J.-F., Perryman,A.L., Pau,V., Kurinov,I., Zahn,K.E. *et al.* (2022) Identification of RP-6685, an orally bioavailable compound that inhibits the DNA polymerase activity of Polθ. *J. Med. Chem.*, **65**, 13198–13215.
31. Mandler,M.D., Baidin,V., Lee,J., Pahil,K.S., Owens,T.W. and Kahne,D. (2018) Novobiocin Enhances Polymyxin Activity By Stimulating Lipopolysaccharide Transport. *J. Am. Chem. Soc.*, **140**, 6749–6753.
32. Le Bras,G., Radanyi,C., Peyrat,J.-F., Brion,J.-D., Alami,M., Marsaud,V., Stella,B. and Renoir,J.-M. (2007) New novobiocin analogues as antiproliferative agents in breast cancer cells and potential inhibitors of heat shock protein 90. *J. Med. Chem.*, **50**, 6189–6200.
33. Raval,S., Sarpe,V., Hepburn,M., Crowder,D.A., Zhang,T., Viner,R. and Schriemer,D.C. (2021) Improving spectral validation rates in hydrogen–Deuterium exchange data analysis. *Anal. Chem.*, **93**, 4246–4254.
34. Kavan,D. and Man,P. (2011) MSTools—Web based application for visualization and presentation of HXMS data. *Int. J. Mass spectrom.*, **302**, 53–58.
35. Borrelli,K.W., Cossins,B. and Guallar,V. (2010) Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility. *J. Comput. Chem.*, **31**, 1224–1235.
36. Knight,J.L., Krilov,G., Borrelli,K.W., Williams,J., Gunn,J.R., Clowes,A., Cheng,L., Friesner,R.A. and Abel,R. (2014) Leveraging data fusion strategies in multireceptor lead optimization MM/GBSA end-point methods. *J. Chem. Theory Comput.*, **10**, 3207–3220.
37. Lim,K.S., Li,H., Roberts,E.A., Gaudiano,E.F., Clairmont,C., Sambel,L.A., Ponnienselvan,K., Liu,J.C., Yang,C., Kozono,D. *et al.* (2018) USP1 Is required for replication fork protection in BRCA1-deficient tumors. *Mol. Cell*, **72**, 925–941.e924.
38. Kilgas,S., Kiltie,A.E. and Ramadan,K. (2021) Immunofluorescence microscopy-based detection of ssDNA foci by BrdU in mammalian cells. *STAR Protocols*, **2**, 100978.
39. Wei,L., Jin,Z., Yang,S., Xu,Y., Zhu,Y., Ji,Y. and Kelso,J. (2018) TCGA-assembler 2: software pipeline for retrieval and processing of TCGA/CPTAC data. *Bioinformatics*, **34**, 1615–1617.
40. Bacolla,A. and Tainer,J.A. (2022), Robust computational approaches to defining insights on the interface of DNA repair with replication and transcription in Cancer. In: Mosammaparast,N. (ed). *DNA Damage Responses*. Methods in Molecular Biology, Humana, NY, Vol. **2444**, pp. 1–13.
41. Hoeksema,H., Johnson,J.L. and Hinman,J.W. (1955) Structural studies on Streptonivicin,1 a new antibiotic. *J. Am. Chem. Soc.*, **77**, 6710–6711.
42. Gellert,M., O'Dea,M.H., Itoh,T. and Tomizawa,J. (1976) Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 4474–4478.
43. Lewis,R.J., Singh,O.M., Smith,C.V., Skarzynski,T., Maxwell,A., Wonacott,A.J. and Wigley,D.B. (1996) The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *EMBO J.*, **15**, 1412–1420.
44. Lamour,V., Hoermann,L., Jeltsch,J.-M., Oudet,P. and Moras,D. (2002) An open conformation of the thermus thermophilusGyrase B ATP-binding domain. *J. Biol. Chem.*, **277**, 18947–18953.
45. Zeth,K., Stanger,F.V., Dehio,C. and Schirmer,T. (2014) Structure of the N-terminal gyrase B fragment in complex with ADP·pi reveals rigid-body motion induced by ATP hydrolysis. *PLoS One*, **9**, e107289.
46. Chalmers,M.J., Busby,S.A., Pascal,B.D., West,G.M. and Griffin,P.R. (2014) Differential hydrogen/deuterium exchange mass spectrometry analysis of protein–ligand interactions. *Expert Rev. Proteomics*, **8**, 43–59.
47. Masson,G.R., Burke,J.E., Ahn,N.G., Anand,G.S., Borchers,C., Brier,S., Bou-Assaf,G.M., Engen,J.R., Englander,S.W., Faber,J. *et al.* (2019) Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods*, **16**, 595–602.
48. Filandrova,R., Kavan,D., Kadek,A., Novak,P. and Man,P. (2021) Studying protein–DNA interactions by hydrogen/deuterium exchange mass spectrometry. *Methods Mol. Biol.*, **2247**, 193–219.
49. Zandarashvili,L., Langelier,M.-F., Velagapudi,U.K., Hancock,M.A., Steffen,J.D., Billur,R., Hannan,Z.M., Wicks,A.J., Krastev,D.B., Pettitt,S.J. *et al.* (2020) Structural basis for allosteric PARP-1 retention on DNA breaks. *Science*, **368**, eaax6367.
50. Dawicki-McKenna,J.M., Langelier,M.-F., DeNizio,J.E., Riccio,A.A., Cao,C.D., Karch,K.R., McCauley,M., Steffen,J.D., Black,B.E. and Pascal,J.M. (2015) PARP-1 activation requires local unfolding of an autoinhibitory domain. *Mol. Cell*, **60**, 755–768.
51. Marciano,D.P., Dharmarajan,V. and Griffin,P.R. (2014) HDX-MS guided drug discovery: small molecules and biopharmaceuticals. *Curr. Opin. Struct. Biol.*, **28**, 105–111.
52. Bennett,M.J., Barakat,K., Huzil,J.T., Tuszynski,J. and Schriemer,D.C. (2010) Discovery and characterization of the laulimalide-microtubule binding mode by mass shift perturbation mapping. *Chem. Biol.*, **17**, 725–734.
53. Lim,X.-X., Shu,B., Zhang,S., Tan,A.W.K., Ng,T.-S., Lim,X.-N., Chew,V.S.Y., Shi,J., Screaton,G.R., Lok,S.-M. *et al.* (2021) Human antibody C10 neutralizes by diminishing Zika but enhancing dengue virus dynamics. *Cell*, **184**, 6067–6080.
54. Aceytuno,R.D., Piett,C.G., Havali-Shahriari,Z., Edwards,R.A., Rey,M., Ye,R., Javed,F., Fang,S., Mani,R., Weinfeld,M. *et al.* (2017) Structural and functional characterization of the PNKP–XRCC4–LigIV DNA repair complex. *Nucleic Acids Res.*, **45**, 6238–6251.
55. Belan,O., Sebald,M., Adamowicz,M., Anand,R., Vancevska,A., Neves,J., Grinkevich,V., Hewitt,G., Segura-Bayona,S., Bellelli,R. *et al.* (2022) POLQ seals post-replicative ssDNA gaps to maintain genome stability in BRCA-deficient cancer cells. *Mol. Cell*, **82**, 4664–4680.
56. Schrempf,A., Bernardo,S., Arasa Verge,E.A., Ramirez Otero,M.A., Wilson,J., Kirchhofer,D., Timelthaler,G., Ambros,A.M., Kaya,A., Wieder,M. *et al.* (2022) POLθ processes ssDNA gaps and promotes replication fork progression in BRCA1-deficient cells. *Cell Rep.*, **41**, 111716.
57. Li,J., Ko,J.M.-Y., Dai,W., Yu,V.Z., Ng,H.Y., Hoffmann,J.-S. and Lung,M.L. (2021) Depletion of DNA polymerase theta inhibits tumor growth and promotes genome instability through the cGAS-STING-ISG pathway in esophageal squamous cell carcinoma. *Cancers*, **13**, 3204.
58. Smolinska,A., Singer,K., Golchert,J., Smyczynska,U., Fendler,W., Sendler,M., van den Brandt,J., Singer,S., Homuth,G., Lerch,M.M. *et al.* (2022) DNA polymerase theta plays a critical role in pancreatic cancer development and metastasis. *Cancers*, **14**, 4077.
59. Oh,G., Wang,A., Wang,L., Li,J., Werba,G., Weissinger,D., Zhao,E., Dhara,S., Hernandez,R.E., Ackermann,A. *et al.* (2023) POLQ

inhibition elicits an immune response in homologous recombination–deficient pancreatic adenocarcinoma via cGAS/STING signaling. *J. Clin. Invest.*, **133**, e165934.
60. Alexandrov,L.B., Kim,J., Haradhvala,N.J., Huang,M.N., Tian Ng,A.W., Wu,Y., Boot,A., Covington,K.R., Gordenin,D.A., Bergstrom,E.N. *et al.* (2020) The repertoire of mutational signatures in human cancer. *Nature*, **578**, 94–101.
61. Osborn,A.J., Elledge,S.J. and Zou,L. (2002) Checking on the fork: the DNA-replication stress-response pathway. *Trends Cell Biol.*, **12**, 509–516.
62. Bacolla,A., Ye,Z., Ahmed,Z. and Tainer,J.A. (2019) Cancer mutational burden is shaped by G4 DNA, replication stress and mitochondrial dysfunction. *Prog. Biophys. Mol. Biol.*, **147**, 47–61.
63. Lachmann,A., Torre,D., Keenan,A.B., Jagodnik,K.M., Lee,H.J., Wang,L., Silverstein,M.C. and Ma'ayan,A. (2018) Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.*, **9**, 1366.
64. Dutta,A., Eckelmann,B., Adhikari,S., Ahmed,K.M., Sengupta,S., Pandey,A., Hegde,P.M., Tsai,M.-S., Tainer,J.A., Weinfeld,M. *et al.* (2016) Microhomology-mediated end joining is activated in irradiated human cells due to phosphorylation-dependent formation of the XRCC1 repair complex. *Nucleic Acids Res.*, **45**, 2585–2599.
65. Pérez-Arnaiz,P., Dattani,A., Smith,V. and Allers,T. (2020) Haloferax volcanii—A model archaeon for studying DNA replication and repair. *Open Biol.*, **10**, 200293.
66. Audebert,M., Salles,B. and Calsou,P. (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA Ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.*, **279**, 55117–55126.
67. Tobalina,L., Armenia,J., Irving,E., O'Connor,M.J. and Forment,J.V. (2021) A meta-analysis of reversion mutations in BRCA genes identifies signatures of DNA end-joining repair mechanisms driving therapy resistance. *Ann. Oncol.*, **32**, 103–112.
68. Higgins,G.S. and Boulton,S.J. (2018) Beyond PARP—POLθ as an anticancer target. *Science*, **359**, 1217–1218.
69. Shibata,A., Moiani,D., Arvai,A.S., Perry,J., Harding,S.M., Genois,M.-M., Maity,R., van Rossum-Fikkert,S., Kertokalio,A., Romoli,F. *et al.* (2014) DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol. Cell*, **53**, 7–18.
70. Ye,Z., Xu,S., Shi,Y., Bacolla,A., Syed,A., Moiani,D., Tsai,C.-L., Shen,Q., Peng,G., Leonard,P.G. *et al.* (2021) GRB2 enforces homology-directed repair initiation by MRE11. *Science Advances*, **7**, eabe9254.
71. Hammel,M., Rashid,I., Sverzhinsky,A., Pourfarjam,Y., Tsai,M.-S., Ellenberger,T., Pascal,J.M., Kim,I.-K., Tainer,J.A. and Tomkinson,A.E. (2021) An atypical BRCT–BRCT interaction with the XRCC1 scaffold protein compacts human DNA ligase IIIα within a flexible DNA repair complex. *Nucleic Acids Res.*, **49**, 306–321.
72. Rao,X., Xing,B., Wu,Z., Bin,Y., Chen,Y., Xu,Y., Zhou,D., Zhou,X., Wu,C., Ye,W. *et al.* (2023) Targeting polymerase θ impairs tumorigenesis and enhances radiosensitivity in lung adenocarcinoma. *Cancer Sci.*, **114**, 1943–1957.
73. Tsutakawa,S.E., Thompson,M.J., Arvai,A.S., Neil,A.J., Shaw,S.J., Algasaier,S.I., Kim,J.C., Finger,L.D., Jardine,E., Gotham,V.J.B. *et al.* (2017) Phosphate steering by Flap Endonuclease 1 promotes 5′-flap specificity and incision to prevent genome instability. *Nat. Commun.*, **8**, 15855.
74. Mengwasser,K.E., Adeyemi,R.O., Leng,Y., Choi,M.Y., Clairmont,C., D'Andrea,A.D. and Elledge,S.J. (2019) Genetic screens reveal FEN1 and APEX2 as BRCA2 synthetic lethal targets. *Mol. Cell*, **73**, 885–899.e886.
75. Guo,H., Wang,Y., Mao,J., Zhao,H., He,Y., Hu,Y., Li,J., Liu,Y., Guan,Z., Guo,A. *et al.* (2023) Cryo-EM structure of DNA polymerase theta helicase domain in complex with inhibitor novobiocin. bioRxiv doi: https://doi.org/10.1101/2023.01.20.524915, 21 January 2023, preprint: not peer reviewed.
76. Putnam,C.D., Shroyer,M.J.N., Lundquist,A.J., Mol,C.D., Arvai,A.S., Mosbaugh,D.W. and Tainer,J.A. (1999) Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with Escherichia coli uracil-DNA glycosylase. *J. Mol. Biol.*, **287**, 331–346.
77. Garcin,E.D., Arvai,A.S., Rosenfeld,R.J., Kroeger,M.D., Crane,B.R., Andersson,G., Andrews,G., Hamley,P.J., Mallinder,P.R., Nicholls,D.J. *et al.* (2008) Anchored plasticity opens doors for selective inhibitor design in nitric oxide synthase. *Nat. Chem. Biol.*, **4**, 700–707.
78. Houl,J.H., Ye,Z., Brosey,C.A., Balapiti-Modarage,L.P.F., Namjoshi,S., Bacolla,A., Laverty,D., Walker,B.L., Pourfarjam,Y., Warden,L.S. *et al.* (2019) Selective small molecule PARG inhibitor causes replication fork stalling and cancer cell death. *Nat. Commun.*, **10**, 5654.
79. Bellon,S., Parsons,J.D., Wei,Y., Hayakawa,K., Swenson,L.L., Charifson,P.S., Lippke,J.A., Aldape,R. and Gross,C.H. (2004) Crystal structures of Escherichia coli topoisomerase IV ParE subunit (24 and 43 Kilodaltons): a single residue dictates differences in novobiocin potency against topoisomerase IV and DNA gyrase. *Antimicrob. Agents Chemother.*, **48**, 1856–1864.
80. Yu,X.M., Shen,G., Neckers,L., Blake,H., Holzbeierlein,J., Cronk,B. and Blagg,B.S.J. (2005) Hsp90 Inhibitors identified from a library of Novobiocin analogues. *J. Am. Chem. Soc.*, **127**, 12778–12779.
81. Collins,A. (1990) What the papers say: topoisomerase II can relax; novobiocin is a mitochondrial poison after all. *Bioessays*, **12**, 493–494.
82. Burlison,J.A., Neckers,L., Smith,A.B., Maxwell,A. and Blagg,B.S.J. (2006) Novobiocin: Redesigning a DNA gyrase inhibitor for selective inhibition of Hsp90. *J. Am. Chem. Soc.*, **128**, 15529–15536.
83. May,J.M., Owens,T.W., Mandler,M.D., Simpson,B.W., Lazarus,M.B., Sherman,D.J., Davis,R.M., Okuda,S., Massefski,W., Ruiz,N. *et al.* (2017) The antibiotic novobiocin binds and activates the ATPase that powers lipopolysaccharide transport. *J. Am. Chem. Soc.*, **139**, 17221–17224.
84. Hartmann,M., Huber,J., Kramer,J.S., Heering,J., Pietsch,L., Stark,H., Odadzic,D., Bischoff,I., Fürst,R., Schröder,M. *et al.* (2021) Demonstrating ligandability of the LC3A and LC3B adapter interface. *J. Med. Chem.*, **64**, 3720–3746.
85. Fairman-Williams,M.E., Guenther,U.-P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.
86. Longo,M.A., Roy,S., Chen,Y., Tomaszowski,K.-H., Arvai,A.S., Pepper,J.T., Boisvert,R.A., Kunnimalaiyaan,S., Keshvani,C., Schild,D. *et al.* (2023) RAD51C-XRCC3 structure and cancer patient mutations define DNA replication roles. *Nat. Commun.*, **14**, 4445.
87. Nguyen,M.T., Moiani,D., Ahmed,Z., Arvai,A.S., Namjoshi,S., Shin,D.S., Fedorov,Y., Selvik,E.J., Jones,D.E., Pink,J. *et al.* (2021) An effective human uracil-DNA glycosylase inhibitor targets the open pre-catalytic active site conformation. *Prog. Biophys. Mol. Biol.*, **163**, 143–159.
88. Tubbs,J.L., Latypov,V., Kanugula,S., Butt,A., Melikishvili,M., Kraehenbuehl,R., Fleck,O., Marriott,A., Watson,A.J., Verbeek,B. *et al.* (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature*, **459**, 808–813.

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com