**Identification of a novel chemotherapeutic agent for tumors with DNA mismatch repair deficiencies**

Yongliang Zhang1,\*, Jennifer T. Fox1,\*, Young-Un Park2,3, Gene Elliott4, Ganesha Rai5, Mengli Cai6, Srilatha Sakamuru5, Ruili Huang5, Menghang Xia5, Kyeryoung Lee7, Min Ho Jeon8, Bijoy P. Mathew2,8, Hee Dong Park9, Winfried Edelmann7, Chan Young Park3, Sung You Hong2,8, David Maloney5, Kyungjae Myung1,2,3 †

Supplemental methods

1. Genome Instability Section, Genetics and Molecular Biology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA

2. Center for Genomic Integrity, Institute for Basic Science, Ulsan, 689-798, Korea

3. Department of Biological Sciences, School of Life Sciences, Ulsan National Institute of Science and Technology, Ulsan, 689-798, Korea

4. Transgenic Mouse Core, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA

5. National Center for Advancing Translational Sciences, National Institutes of Health. Rockville MD 20850, USA

6. Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

7. Department of Cell Biology, Albert Einstein College of Medicine, New York 10461, USA

8. School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology, 689-798, Ulsan, Korea

9. Institute of Drug Discovery, Research & Development, LG Life Sciences Ltd. 305-738, Daejeon, Korea

\* These authors contributed equally

† To whom correspondence should be addressed: E-mail kmyung@ibs.re.kr; Phone +82-52-217-5323; Fax +82-52-217-5323

**Supplemental Methods**

**Quantitative high-throughput screening (qHTS) using the ATAD5 luciferase reporter gene assay.** HEK293T cells stably expressing ATAD5-luc were dispensed at density of 2,000 cells/per well in 5 μL of culture medium into tissue culture-treated 1,536-well white/solid bottom assay plates (Greiner Bio-One) using a Multidrop Combi reagent dispenser (Thermo Scientific)(1). After the assay plates were incubated for 4-5 hours at 37°C to allow for cell adherence, 23 nL of compound was transferred via pin tool (Kalypsys) to columns 5-48 of the assay plates. A total of 344,385 compounds from the MLPCN library were screened at 6 concentrations ranging from 2.9 nM to 46 μM (PubChem AID: 493143). The controls included in columns 1-4 of each plate were: column 1, dose-response curve of 5-Fluorouridine (FUrd), a known stabilizer of ATAD5-luc, ranging from 2 nM to 92 μM; columns 2 and 4, 23 and 37 μM FUrd, respectively; and column 3, DMSO only. The treated assay plates were incubated for 16 hours at 37°C, followed by the addition of 5 μL/well ONE-Glo Luciferase Reagent (Promega) using a Flying Reagent Dispenser (FRD) (Aurora Discovery). After a 30 minute incubation at room temperature, the luminescence intensity was quantified using a ViewLux plate reader (Perkin Elmer). Raw plate reads for each titration point were normalized relative to the FUrd control (23 μM, 100%) and DMSO only wells (basal, 0%) and then corrected by applying a pattern correction algorithm using compound free control (DMSO) plates.

To eliminate false positives that act by stabilizing luciferase alone, the compounds were also tested in a luciferase counter screen. CMV-Luc2-Hygro HEK293 cells were dispensed at 2,000 cells/5 μL/well in assay medium containing 10% FBS. The cells were dispensed into tissue culture treated 1,536-well white/solid bottom assay plates (Greiner Bio-One) using a Flying Reagent Dispenser (FRD) (Aurora Discovery). After the assay plates were incubated for 5 hours at 37°C to allow for cell adherence, 23 nL compound was transferred via Pin Tool (Kalypsys) to rows 1-30 of the assay plates, resulting in the final concentrations ranging from 0.2 nM to 46 μM. DMSO only was transferred to row 31. In row 32, N6 Phenyl Adenosine at final concentrations of 46, 23, and 0 μM were transferred to columns 1-12, 13-24, and 25-48, respectively. The assay plates were incubated for 16 hours at 37°C, followed by the addition of ONE-Glo Luciferase reagent (Promega) at 5 μL/well using FRD. After a 30 min incubation at room temperature, the luminescence intensity was quantified using a ViewLux CCD-based plate reader (Perkin Elmer). Raw plate reads for each titration point were first normalized relative to N6 Phenyl Adenosine control (46 μM, 100%) and DMSO only wells (basal, 0%) and then corrected by applying a pattern correction algorithm using compound free control plates (DMSO) plates.

A total of 912 compounds were classified as potential ATAD5 activators, yielding a hit rate of 0.26%. The criteria for a positive hit were as follows: (1) a curve categorized as class 1.1, 1.2, 2.1, or 2.2 in the ATAD5-luc assay(2) (2) an efficacy >40% in the ATAD5-luc assay and (3) inactive or >10-fold less potent in the luciferase counter screen. From these 912 hits, 240 compounds were selected for testing in the MMR-deficient cell viability assays based on structure-activity relationship (SAR) analysis of the actives.

**Baicalein-Biotin synthesis procedure.** A solution of biotin (0.065 g, 0.268 mmol, 1 eq) in DMF (1 mL) was added to TEA (0.187 mL, 1.339 mmol, 5 eq) followed by BOP (0.154 g, 0.348 mmol, 1.3 eq) and stirred at room temperature for 15 min. 2-(4-(2-(2-Aminoethoxy)ethoxy)phenyl)-5,6,7-trihydroxy-4H-chromen-4-one (0.1 g, 0.268 mmol) was then added to the reaction mixture and stirred for an additional 45 minutes at room temperature. The crude product was purified by preparative HPLC to collect major peaks containing the product. The crude mixture obtained after liophilization was dissolved in methanol (4 mL) containing few drops of concentrated HCl and refluxed for 1 h. After removing the volatiles, the crude product was dissolved in 0.5 mL DMSO and purified on a preparative HPLC to obtain pure product 5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-N-(2-(2-(4-(5,6,7-trihydroxy-4-oxo-4H-chromen-2-yl)phenoxy)ethoxy)ethyl)-pentanamide. LC-MS retention time: (method 1; 8 min run) 3.966 min and (method 2; 4 min run) 2.604 min, HRMS (ESI): m/z (M + H)+ calcd for C29H34N3O9S, 600.2026; found, 600.2020.



**Xenograft assay.** The nude mice were housed under standard conditions. All animal studies were performed with an approved protocol from Animal Care and Use Committee of the National Human Genome Research Institute, National Institutes of Health. Six-week-old nude mice were subcutaneously injected at one flank with 2 x 106 colon cancer cells. When the diameter of tumor reached approximately 1 cm, two pellets containing baicalein or placebo were implanted besides the tumor. The pellet releases the same amount of active product per unit time and stops after the designated time is reached. Each pellet has 26.3mg baicalein or placebo (21-day release, Innovative Research of America). The control pellets (implanted in control group) are called C-111 placebo, which contained all the components of matrix except baicalein. This Matrix-Driven Delivery system has been used by many groups with different chemicals integrated(3-5). Baicalein was purchased from Sigma and was sent to the company for making it into their matrix. About two weeks after implantation, tumors were harvested from euthanized mice, and the tumor volume (mm3) was determined using water displacement method. In Fig. 5 and Fig. S7, all experiments were done at two different times. In HEC59 cell xenograft experiments, four mice were used in control group and another four were used in baicalein group at the first time and two mice were used in control group and three mice were used in baicalein group at the second time. Thus, total seven mice for baicalein and six mice for control were tested for HEC59 xenograft experiments. In HT29 cell xenograft experiments, three mice were used in each group for the first experiment and two mice were used in each group of second experiment. Thus, total five mice for baicalein and another five mice for control were tested for HT29 xenograft experiments. In Fig. S8 for time dependent experiment, four mice were used in each group. In Fig. S9 of LoVo xenograft experiments, three mice were used in each group.

**siRNA transfection.** Cells were grown to approximately 40% confluence. The cells were transfected with 5 nM negative control siRNA (Ambion), MSH2 siRNA (Qiagen, Hs\_MSH2\_6), XPF siRNA (Thermo Scientific Dharmacon®), CHK2 siRNA (Cell Signaling) or ATM siRNA (Cell Signaling) using the HiPerFect transfection reagent (Qiagen) according to the manufacturer's instructions. After incubation with siRNA for 48 hours, the cells were treated with baicalein. The cells were then either lysed and subjected to western blot analysis, or lysed to determine cell viability as described above.

**Colony formation assay.** 800 cells were seeded onto 10-cm tissue culture plates. The cells were incubated at 37 °C/5% CO2 for approximately 5 hours to allow for attachment, and then treated with 1:2 dilutions of baicalein (final baicalein concentration = 0-20 M). Twenty-four hours after treatment, the cells were washed with phosphate-buffered saline, allowed to recover in fresh medium for 7 days, and then stained with Giemsa stain to visualize viable colonies. Dose-response curves were generated using GraphPad Prism.

**AOM-DSS-induced colorectal tumor studies.** The *MSH2LoxP+/-VilCre+*micewere kindly provided by Dr. W. Edelmann (6) and bred under standard conditions. All animal studies were performed using an approved protocol from the Animal Care and Use Committee of the National Human Genome Research Institute, National Institutes of Health. Mice were divided into four groups, each containing 6-7 approximately 20-week-old male or female *MSH2LoxP/LoxPVilCre*mice. Groups 1 (*MSH2LoxP/LoxPVilCre*malemice) and 3 (*MSH2LoxP/LoxPVilCre*femalemice) were fed with the control diet for four weeks and then and treated with AOM/DSS. Groups 2 (*MSH2LoxP/LoxPVilCre*malemice) and 4 (*MSH2LoxP/LoxPVilCre*femalemice) were fed the baicalein-supplemented diet for four weeks and then treated with AOM/DSS. WT male mice that were approximately 20-weeks-old were also treated as described above. All mice continued to receive their respective diets until sacrifice. For the AOM/DSS treatment, mice in each group were given a single intraperitoneal (IP) injection of AOM (10mg/kg body weight). Seven days later, the mice were fed with 1.25% DSS in drinking water for the indicated cycles.

**Gel shift assay.** MutS gel shift assay was performed as previously reported(7). The sequences of DNA are as follows: Plus: CGGATCCGACTCATTCC*T*GCAGCGACTCCATGGGA

Minus: TCCCATGGAGTCGCTGC*G*GGAATGAGTCGGATCCG

**Cell cycle analysis.** Cells were treated with 1 mM hydroxyurea (Sigma) for 24 hours to synchronize the cells. The hydroxyurea was then washed out and replaced with media containing 60 ng/mL nocodazole (Sigma) and either DMSO or 100 μM baicalein. 24 hours later, the cells were harvested and fixed with 70% ethanol. Following treatment with RNase I (Roche) in phosphate-buffered saline, propidium iodide (Invitrogen) was added at a final concentration of 50 μg/mL. Cell cycle analysis was carried out by flow cytometry.

**EdU incorporation assay.** Cells were treated with DMSO or 100 μM baicalein for 24 hours. EdU incorporation was determined using the Click-iT EdU Flow Cytometry Assay Kit (AlexaFluor 488, Invitrogen) according to the manufacturer’s protocol.

**Double thymidine block.** When HeLa cell culture reached at 30% confluency, cells were washed with PBS twice and incubated with medium with 2 mM thymidine for 18 hours (first block). Then, thymidine was removed by washing cells with PBS and further incubated with fresh medium for 9 hours. After 9 hours incubation, cells were treated with 2mM thymidine for 17 hours (second block). Cells were then washed with PBS and cultured in fresh medium. Upon release from the thymidine block, G1/S cells are harvested at 0 hours, and G2 cells are harvested at 6 hours.

**NMR assay for DNA and baicalein binding.** All samples contained 0.1mM DNA molecules dissolved in 400L Sodium Phosphate Buffer (25mM, pH 6.3) and were incubated with DMSO or 0.4L 100mM baicalein overnight. One dimensional 1H NMR spectra was collected on a Bruker DMX600 spectrometer at 10°C. By comparing the spectra, the NMR resonance from the imino groups of the mismatch A/G bases can be assigned(8). The sequences of DNA are as follows: Plus: CATTCC*T*GCAGCGAC

Minus (Normal): GTCGCTGC*A*GGAATG

Minus (Mismatched): GTCGCTGC*G*GGAATG

**DNA and baicalein pull-down assay.** The 5’ digoxigenin-labeled DNA was synthesized by Integrated DNA Technologies. The sequences are as follows: Plus: CGGATCCGACTCATTCC*T*GCAGCGACTCCATGGGA

Minus (Matched): TCCCATGGAGTCGCTGC*A*GGAATGAGTCGGATCCG

Minus (Mismatched): TCCCATGGAGTCGCTGC*G*GGAATGAGTCGGATCCG

Digoxigenin-labeled DNA was incubated with biotin-labeled baicalein in binding buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.05% NP-40, 5 % Glycerol) for 6 hours at 4°C with gentle rocking. Streptavidin Sepharose beads (GE healthcare) were used to capture the baicalein-biotin via mixing for 2 hours at 4°C on a rotator. Beads were subsequently washed three times with the binding buffer. To elute bound DNA, the beads were incubated with elution buffer (10 mM EDTA pH 8.2 and 95% formamide) at 65°C for 5 minutes. The DNA was then transferred to a nitrocellulose membrane (Bio-Rad Laboratories), washed three times with 20×SSC buffer and crosslinked with UV. Co-precipitated DNA was detected using an antibody recognizing digoxigenin along with ECL Prime Western Blotting Detection Reagent (GE Healthcare).

**Biacore assay.** Surface Plasmon Resonance (SPR) experiments were performed using a Biacore T200 instrument at 25°C. The MutSα and SHPRH proteins diluted in 10 mM sodium acetate (pH 4.5) were immobilized on a series S CM5 chip (GE Healthcare) using amine coupling methods as described manufacturer instructions. The immobilization levels were ~1450 resonance unit (RU). For kinetic measurement, 1X PBS-P (PBS containing 0.005% surfactant P20) was used as running buffer. The varying concentrations of baicalein were flowed over the surface of a flow cell (fc 4) and a reference cell (fc 3) with 30 L/min for 60 sec with a dissociation time of 180 sec. The sensor chip surface was regenerated by running buffer containing 50% DMSO after each cycle. To test the dynamics of baicalein and matched or mismatched DNA interaction, the baicalein-biotin was immobilized on a series S SA chip (GE Healthcare) as described manufacturer instructions. The immobilization levels were ~300 resonance unit (RU). For kinetic measurement, HBS (25 mM HEPES pH 7.5, 150 mM NaCl containing 0.005% TWEEN-20) was used as running buffer. The matched or mismatched DNA flow was maintained with 30 L/min for 60 sec and then dissociated for 90 sec with HBS buffer. The sensor chip surface was regenerated by running buffer containing 10% DMSO plus 0.3% TWEEN-20 after each cycle. The DNA sequences are as follows: Plus: CGGATCCGACTCATTCC*T*GCAGCGACTCCATGGGACTTGCTTCGCTTCTGTC

Minus (Matched): GACAGAAGCGAAGCAAGTCCCATGGAGTCGCTGC*A*GGAATGAGTCGGATCCG

Minus (Mismatched): GACAGAAGCGAAGCAGGTCCCATGGGGTCGCTGC*G*GGAATGAGTCGGATCCG

To test the dynamics of baicalein, MutSα, and Mismatched DNA interaction, the Biotin-Mismatched DNA was immobilized on a series S SA chip (GE Healthcare) as described manufacturer instructions. The immobilization levels were ~200 resonance unit (RU). For kinetic measurement, HBS (25 mM Hepes (pH 7.5), 150 mM KCl, 5 mM MgCl2 containing 0.01% TWEEN-20) was used as running buffer. The Pre-incubated MutSα and baicalein flow was maintained with 30 L/min for 60 sec and then dissociated for

120 sec with HBS buffer. The sensor chip surface was regenerate with 50 mM NaOH after each cycle. MADSL-3’ Biotin-labeled Plus, 5-CGGATCCGACTCATTCC*T*GCAGCGACTCCATGGGATTTTTTTTTTTTTTT-3

MADSL-Minus, 5-TCCCATGGAGTCGCTGC*G*GGAATGAGTCGGATCCG-3

The data was analyzed by Biacore T200 Evaluation Software (GE Healthcare).

**Pulsed-field gel electrophoresis.** Cells were grown to approximately 90% confluence in 10-cm tissue culture plates and then treated with DMSO or 100 μM baicalein for 24 hours. ~6 x 105 cells were mixed in melted Pulsed-Field Certified Agarose (Bio-Rad) for a final agarose concentration of 0.75%. Agarose plugs were then digested in proteinase K reaction buffer (100 mM EDTA, pH 8.0, 0.2% sodium deoxycholate, 1% sodium lauryl sarcosine, and 1 mg/ml Proteinase K) at 50°C overnight and washed 4 times in wash buffer (20 mM Tris, pH 8.0, 50 mM EDTA). The plugs were loaded onto a 0.8% Pulsed-Field Certified Agarose gel (Bio-Rad). Separation was performed on a CHEF-DR III pulsed-field electrophoresis system (Bio-Rad; 100° field angle, 1200 sec switch time, 2 V/cm, 14°C) for 72 hours. The gel was stained with ethidium bromide, and double-strand breaks were quantified using ImageJ.

**Ethidium bromide displacement assay.** Ethidium bromide (BioRad) was diluted in 0.1M NaCl, 0.1M Tris, pH 8 for a final concentration of 6.43 μM. 70 μL of this ethidium bromide solution was then added to each well of a black, clear-bottom 96-well plate along with 10 μL of 0.1 μg/μL plasmid DNA (pcDNA3.1, Invitrogen) and 20 μL of the indicated compound. The mixture was incubated at room temperature for 30 minutes, and fluorescence was recorded using a Fluoroskan Ascent Fluorometer (Thermo Scientific, excitation = 485 nm, emission = 612 nm).

**Drug affinity responsive target stability (DARTS) assay.** 1 μg purified protein in 400 μL cold M-PER lysis buffer (Pierce) was incubated with protease and phosphatase inhibitors on ice for 10 minutes. 44.4 μL of 10 × TNC buffer (500 mM Tris-HCl pH 8.0, 500 mM NaCl, 100 mM CaCl2) was then added to the mixture. The mixture was split into two samples by transferring 210 μL into each of two tubes. DMSO was added to one tube and baicalein at a final concentration of 200 μM was added to the other. Following a 30 min incubation at room temperature pronase was diluted (Roche) to 25 μg/mL, 125 μg/mL or 1250 μg/mL with cold 1 × TNC buffer, and 2 μL was added to 50 μL aliquots of each protein sample. After 20 min, the digestion as terminated by boiling in SDS loading buffer for 5 min.

**References**

1. Fox JT, Sakamuru S, Huang R, Teneva N, Simmons SO, Xia M, et al. High-throughput genotoxicity assay identifies antioxidants as inducers of DNA damage response and cell death. Proc Natl Acad Sci U S A 2012;109(14):5423-8.

2. Inglese J, Auld DS, Jadhav A, Johnson RL, Simeonov A, Yasgar A, et al. Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. Proc Natl Acad Sci U S A 2006;103(31):11473-8.

3. Karayannis T, Au E, Patel JC, Kruglikov I, Markx S, Delorme R, et al. Cntnap4 differentially contributes to GABAergic and dopaminergic synaptic transmission. Nature 2014;511(7508):236-40.

4. Alkass K, Panula J, Westman M, Wu TD, Guerquin-Kern JL, Bergmann O. No Evidence for Cardiomyocyte Number Expansion in Preadolescent Mice. Cell 2015;163(4):1026-36.

5. Bhola NE, Jansen VM, Bafna S, Giltnane JM, Balko JM, Estrada MV, et al. Kinome-wide functional screen identifies role of PLK1 in hormone-independent, ER-positive breast cancer. Cancer Res 2015;75(2):405-14.

6. Kucherlapati MH, Lee K, Nguyen AA, Clark AB, Hou H, Jr., Rosulek A, et al. An Msh2 conditional knockout mouse for studying intestinal cancer and testing anticancer agents. Gastroenterology 2010;138(3):993-1002 e1.

7. Geng H, Du C, Chen S, Salerno V, Manfredi C, Hsieh P. In vitro studies of DNA mismatch repair proteins. Analytical biochemistry 2011;413(2):179-84.

8. Bhattacharya PK, Cha J, Barton JK. 1H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. Nucleic acids research 2002;30(21):4740-50.