Final report

„Complex enzyme activities supporting genome maintenance and cancer prevention in real-life crowded environments"

Project NKFIH K-116072

Oct 1, 2015 – Sep 30, 2020

PI: Mihály Kovács, Department of Biochemistry, Eötvös Loránd University

The genomes of living organisms is constantly damaged both by external factors and by the cells' intrinsic metabolic processes. In our project, we used novel approaches to explore the mechanisms and physiological roles of protein molecules that repair damaged DNA. Molecular functions were elucidated by biophysical techniques that we newly developed and optimized, while biological effects were unveiled by genetic and functional studies of the same protein system in bacteria and vertebrates (zebrafish). We have discovered a new molecular function by which DNA repair enzymes recognize and repair defective or potentially harmful genome segments, based on the geometry of multi-stranded DNA intermediates. In addition to the roles of DNA repair systems in somatic cells, our functional studies also revealed their functions in germ line development and fertility.

Specifically, our project was aimed at the elucidation of molecular mechanisms underlying the efficient propagation and control of homologous recombination (HR), a key mechanism enabling potentially error-free repair of DNA lesions. HR quality control mechanisms play key roles in avoiding the potentially deleterious or fatal consequences of imprecise HR, i.e. harmful genetic rearrangements, cell death and cancer.

Aim 1

(*i*) In Aim 1 we deciphered molecular activities of bacterial (*E. coli* RecQ) and human (Bloom's syndrome, BLM) helicases (central DNA recombination and repair enzymes) that lead to the disruption of DNA strand invasions contained in D-loops, key intermediates of HR. In ref. [1], reporting single-molecule and ensemble biophysical experiments, we published the discovery of a novel HR control mechanism whereby *E. coli* RecQ helicase uses its selective DNA invasion disruption and „shuttling" activities in order to selectively disrupt illegitimate (non-allelic, potentially harmful) HR invasions. Importantly, we also showed that these activities together suppress illegitimate recombination in vivo (cf. Aim 3) [1].

(*ii*) As a follow-up to the above publication, we have further developed the previously devised oligonucleotide-based D-loop processing assay with the inclusion of additional fluorescent sensors that enabled us to precisely dissect and quantify all possible D-loop disruption and stabilization (extension) pathways. Using the new assay we showed that, in contrast to *E. coli* RecQ that is biased toward D-loop disruption [1], human BLM helicase maintains a balance between D-loop disruption and stabilization. This mechanism enables BLM to exert its complex regulatory roles during eukaryotic HR (Harami GM, Seol Y, Pálinkás J, Kovács ZJ, Gyimesi M, Harami-Papp H, Neuman KC, Kovács M: Intrinsic balance between D-loop disruption and stabilization by human Bloom's syndrome helicase; manuscript). The manuscript describing this work was reviewed at Nature Communications, and the reviewers requested the inclusion of additional data on how key interaction partners of BLM, the topoisomerase III alpha-RMI1-RMI2 (TRR) complex, affect BLM's D-loop processing profile. Intriguingly, in our subsequent work we found that the TRR complex massively shifts BLM's D-loop processing balance toward D-loop disruption, thereby providing an efficient HR quality control mechanism for the BTRR (BLM-TRR) complex. These novel findings identified a role for BTRR in the early steps of HR, significantly extending previously prevailing views that BTRR is primarily involved in the late HR step of

double Holliday junction dissolution. Currently we are about to resubmit the manuscript that reports our combined findings on BLM and BTRR-catalyzed D-loop disruption.
(*iii*) We published a detailed monograph on the methods developed by us and others for the study of the enzymological-biophysical mechanisms of D-loop disruption and other DNA-restructuring enzymatic activities in Elsevier Methods [2].
(*iv*) Related to both Aims 1 and 2, we have extended our studies on the effect of molecular crowding agents (polyethylene glycol, dextran, bovine serum albumin) on the DNA-processing activities of *E. coli* RecQ and human BLM helicases. Our results delineated that, while the ATPase and ssDNA-binding properties of the enzymes are largely unaffected by molecular crowding, the D-loop processing reactions of BLM are affected by crowded conditions in a way conferring a preference toward D-loop stabilization (unpublished data).
(*v*) Besides molecular crowding, the formation of protein and nucleoprotein condensates by liquid-liquid phase separation (LLPS) is an emerging mechanism for intracellular compartmentalization, material distribution and biocatalysis. We recently published our discovery that *E. coli* single-stranded DNA binding (SSB) protein forms LLPS condensates in cellular-like conditions through multifaceted interactions involving all structural regions of the protein. SSB, ssDNA, and SSB-interacting molecules (including RecQ-family helicases) are highly concentrated within the condensates, whereas LLPS is overall regulated by interaction between SSB and ssDNA. Our results revealed a conserved mechanism by which bacterial cells store a pool of SSB and SSB-interacting proteins, and dynamic LLPS enables rapid mobilization of this pool to protect exposed ssDNA and repair genomic loci affected by DNA damage [3].
(*vi*) We published a further important mechanistic discovery on how RecQ helicases recognize and suppress illegitimate recombination events to serve genomic stability [4]. We employed an efficient combination of single-molecule magnetic tweezers and rapid kinetic approaches to establish that the helicase-and-RNase-D-C-terminal (HRDC) domain of RecQ helicase stabilizes intrinsic, DNA sequence-dependent, pauses of the helicase motor core in a DNA geometry-dependent manner. We found that, in the DNA unwinding mechanism of the helicase core, the DNA unwinding rate depends on the stability of the duplex DNA leading to transient sequence-dependent pauses. We detected non-linear amplification of these transient pauses by controlled DNA binding of the HRDC domain. The resulting DNA sequence- and geometry-dependent pausing is especially sensitive to mismatches and discontinuities in the DNA duplex, and thus it may well underlie a homology sensing mechanism that allows rapid disruption of unstable (illegitimate) and stabilization of stable (legitimate) DNA strand invasions, thereby providing an intrinsic mechanism for recombination quality control.
(*vii*) In collaboration with Peter Burkovics' lab (ELKH BRC), we assessed molecular mechanisms by which cells cope with the effects of formation of stable secondary structures (G-quadruplexes) that may lead to replication stalling. Several helicases have been implicated to regulate G4 unfolding to preserve genome integrity, but the processing of G-quadruplexes during DNA replication is thought to involve additional factors. We showed that the yeast Mgs1 protein robustly binds to G4 structures in vitro and preferentially acts at regions with a strong potential to form G4 structures in vivo, pointing to the role of Mgs1 in the preservation of genome integrity [5].
(*viii*) To quantitatively assess the effect of BLM helicase (and other genome maintenance factors) on the outcome of DNA damage tolerance and repair processes, we devised an SV40 large T antigen-based in vitro replication system using cellular extracts, in collaboration with Dávid Szüts' group (ELKH TTK). Our results based on Sanger and next generation sequencing, combined with DNA lesion-specific qPCR and replication efficiency measurements, indicated that BLM helicase interferes with T antigen-driven replication, in line with its proposed roles in the efficient disruption of HR intermediates requiring DNA strand extension . with the full length protein showing more activity than its helicase core region (Szeltner Z, Póti Á, Harami GM, Kovács M, Szüts D: Evaluation and modulation of DNA lesion bypass in an SV40 large T antigen-based *in vitro* replication system; manuscript under review at FEBS Open Bio).
(*ix*) Besides the roles of RecQ-family helicases in DNA replication and repair, we raised the hypothesis that these enzymes may play mechanistic roles in the disruption of R-loops (hybrid structures

comprising an RNA strand invading into a DNA duplex) that are formed during transcription and may halt transcription and/or DNA replication, and cause recombination events leading to harmful genomic rearrangements. Using the methodology described above (ii), we showed that both E. coli RecQ and human BLM helicases are indeed capable of disrupting R-loop structures, with RecQ being a more potent R-loop disruptor than BLM helicase (unpublished data).
(*x*) In collaboration with Keir Neuman's (National Heart, Lung and Blood Institute, USA) Jens Grundlach's lab (Univ. Washington, USA), we are deciphering the physical mechanism of DNA unwinding by RecQ-family helicases using single-molecule „nanopore tweezers". In this setup, the passing of the DNA strand (liberated by helicase-driven unwinding) through a lipid bilayer via the MspA a channel-forming protein is monitored at high spatiotemporal resolution to detect possible DNA sequence-dependent steps in the unwinding kinetics (unpublished data).
(*xi*) Besides the action of the RMI1/2 proteins in the helicase-containing BTRR complex (see point (*ii*) above), we hypothesize based on available indications that RMI1/2 may also influence the molecular „gating" mechanism of topoisomerase III alpha, thereby contributing to the DNA decatenation mechanism that is necessary for efficient dissolution of HR intermediate DNA structures. We are in process of performing single-molecule experiments to assess whether RMI1/2 promote the structural state of the topoisomerase that favors DNA decatenation (unpublished data).
(*xii*) We have utilized the rapid kinetic methodology used for deciphering the DNA-restructuring activities of DNA helicases also to elucidate the physicochemical basis of the activities of other metabolic proteins [6,7].

Aim 2
(*xiii*) In Aim 2 we deciphered the mechanism of assembly and remodeling of recombinase nucleoprotein filaments, key early HR intermediates. We published a structural and kinetic mechanism for the assembly of nucleoprotein filaments formed by human RAD51, yeast Rad51 and bacterial (*E. coli*) RecA proteins with ssDNA [8]. In this paper we reported a dynamic assembly mechanism prone for nucleoprotein structural rearrangements serving HR regulation. Knowledge of the nucleoprotein assembly mechanism served as a framework for analyzing RecQ-family helicase-catalyzed remodeling of nucleoprotein filaments.
(*xiv*) We have also devised plasmid-based D-loop disruption assays, which allow for the quantification of the helicase-catalyzed processing of native-like D-loops formed by prokaryotic (RecA) and eukaryotic (RAD51) recombinases. We use a modified PTBX3 plasmid into which a RecA-nucleoprotein filament (based on a fluorescently-labeled oligonucleotide) can invade to form a dynamic (mobile) D-loop. We used fluorescence polarization and electrophoretic mobility shift assays to analyze nucleoprotein and D-loop formation, respectively. We found that in this experimental setup RecA can form nucleoprotein filaments on ssDNA in a cooperative manner and can search and find the homologous region in the target plasmid to form a D-loop. This setup enables the analysis of nucleoprotein and D-loop processing mechanism of RecQ helicase constructs (unpublished data).
(*xv*) We have also investigated D-loop processing mechanisms at the single-molecule level. Using total internal reflection fluorescence (TIRF) microscopy it is possible to visualize individual, fluorescently-labeled D-loops. To generate fluorescent D-loops, Alexa488-labeled ssDNA strands of different length were produced to serve as the invading strand. The receiving dsDNA is biotinylated at each end, thus it binds to the appropriately treated glass surface of the flow cell. The dsDNA was visualized using SYTOX dye, thus enabling simultaneous detection of the dsDNA substrate and the D-loop position within it. Examined D-loops were generated using RecA protein. We could successfully detect single fluorophores in the applied TIRF microscope setup, reporting the helicase-mediated processing of D-loops in real time (unpublished data).
(*xvi*) We have further developed the magnetic tweezers-based single-molecule biophysical assays used to monitor helicase-catalyzed unwinding, applied originally in ref. [1] (see point (*i*) above). Using this setup in combination with solution kinetic methodologies, we elucidated a new molecular mechanism whereby RecQ helicase induces a change in the DNA binding mode of single-stranded DNA binding (SSB) protein *via* a specific protein-protein interaction [9]. This mechanism enables RecQ to liberate an

ssDNA segment on which its DNA-processing reactions can be efficiently initiated at sites appropriate for replication restart and DNA repair.

Aim 3

(*xvii*) In Aim 3, we elucidated the *in vivo* effect of molecular activities unveiled in Aims 1-2 in genome maintenance and HR quality control in bacterial (*E. coli*) and vertebrate (zebrafish, *Danio rerio*) model organisms. As a follow-up to the *in vivo* assessment of RecQ's HR quality control activities in ref. [1], in ref. [10] we reported that the specific functional alterations of bacterial HR initiation (RecBCD) and quality control (RecQ) systems exert cumulative effects on genomic stress survival of *E. coli* cells. Strikingly, however, these alterations can complement each other in terms of HR precision, reflecting the precise adaptive fine-tuning of both systems to their native (wild-type) functional environment.

(*xviii*) In the zebrafish subproject of Aim 3, we set out to generate a transgenic TrafficLight line that can be used for the quantification of DNA repair outcomes *via* differentiating between HR versus Non-Homologous End Joining (NHEJ) events. We successfully isolated two independent transgenic insertions. However, the specific zinc finger nuclease (ZFN), which was used in earlier cell culture studies with the TrafficLight construct, failed to cut its target site in our transgenic carriers. To overcome this setback, we used another construct that is applicable for any (pigmented) zebrafish line, as it primarily targets the well characterized golden locus. This construct enables the use of fluorescent reporters to differentiate and quantify HR and NHEJ events (unpublished data).

(*xix*) We successfully created a frameshift mutation in the zebrafish Bloom's syndrome helicase (*blm*) gene, resulting in a null allele. We in-crossed heterozygous parents and observed that *blm* null homozygotes are viable. As human patients with mutations in RecQ helicases often show the hallmarks of progeria, we measured the lifespan of homozygous mutants, compared them to heterozygous and-wild type siblings, and demonstrated a markedly reduced lifespan for *blm* -/- fish. Our results also show that compared to their wild-type counterparts even blm +/- heterozygous animals have compromised viability and significantly shorter lifespan. Surprisingly, our results show that, unlike a recently described rad51 mutant line, blm -/- homozygous fish do not show an elevated susceptibility to ionizing radiation. We also demonstrated that in a Blm-deficient background, the number of DNA lesions (as shown by anti-gammaH2AX antibody staining) is not altered compared to wild type counterparts, either in the absence or following gamma-irradiation. To expand these observations, we performed further genotoxicity assays using 1,2,3,4-diepoxybutane (DEB), an agent that crosslinks DNA strands. Our results show that Blm is not required for DNA repair after genotoxic treatment, as wild-type animals and blm mutants showed similar numbers of gammaH2AX-positive foci. Interestingly, we also observed that blm null mutant adult fish are always male. Such skewed sex-ratios have been observed before in zebrafish models of Fanconi anemia and were proposed to be linked to p53-dependent apoptosis in the early pool of primordial germ cells (PGCs). To test if the impairment in sex determination in blm mutants is also due to a similar mechanism, we have acquired the homozygous p53 mutant zebrafish strain and crossed them into blm mutant background to test if in the double mutants we can observe any females. We intercrossed blm mutants with tp53 mutants and we raised the progeny of the double-heterozygote incross. Interestingly, unlike in other mutants of DNA repair genes (e.g. of different Fanconi anemia genes) where a similar "all-male" phenotype could be rescued with the impairment of apoptosis, in blm;tp53 double mutants we still observed biased sex ratios (essentially, regardless of the tp53 genotype, homozygous blm mutants always develop into males). Also, as sex determination in zebrafish is dependent on the number of PGCs, we tested if the early expansion of the PGC pool at about 2 weeks post fertilization (wpf) that can be observed in wild-type female fish is observable in the blm null background. To do this, we performed primordial germ cell (PGC) counts in ~2-week old blm;Tg(ddx4:egfp) animals and showed that blm mutants indeed have smaller numbers of PGCs. As this observation is in line with the mechanism of sex determination in zebrafish (individuals that do not expand their PGCs during early stages of development will become males), we do believe that we have identified one important aspect of Blm function in this species. Similarly to the human phenotype, our *blm* null mutant males were sterile. We performed histological analysis of mutant testes and demonstrated that the mutants lack viable sperm, as spermatocytes are

arrested during meiosis. We were also able to show irregularities during spermatocyte meiosis in mutant fish using electron microscopy (EM). To test the basis of infertility, we explored the meiotic cell divisions in the testes blm-/- animals. We used antibodies against Sycp3, a marker of the meiotic synaptonemal complex. Our results showed that in blm mutant animals a meiotic arrest takes place, which explains why mutant testes lack mature spermatocytes. Taken together, our results show that zebrafish blm mutants recapitulate major hallmarks of the human disease. Moreover, some functions of zebrafish Blm bear additional importance in germ line development, and consequently in sex differentiation. Therefore, our model is a valuable tool for further understanding the developmental and molecular attributes of this rare disease, along with providing novel insights into the role of genome maintenance proteins in somatic DNA repair and fertility. We are about to submit the manuscript describing these findings to a high-end genetics journal (Annus T, Müller D, Jezsó B, Ullaga G, Harami GM, Orbán L, Kovács M*, Varga M*: Bloom's syndrome helicase prominently contributes to zebrafish longevity, germ line development and sex determination in a p53-mediated apoptosis-independent manner; manuscript).

(*xx*) Besides *blm*, we have also created novel mutations in other zebrafish RecQ ortholog genes. In humans, mutations of RECQL4 are the cause of the Rothmund-Thomson, RAPADILINO and Baller-Gerold syndromes, therefore we were especially pleased to identify a p.Ser271Glyfs*4 frameshift mutation in the *recql4* gene of zebrafish as the result of the CRISPR-mediated mutagenesis that we performed. We were able to isolate this mutation, and we have performed an additional outcross to ascertain that the mutation is in a wild-type background. After raising the progeny of *recql4* heterozygotes, we set out to characterize the phenotypes of recql4 homozygotes. We were also successful in identifying highly active sgRNAs against *recql*, another vertebrate RecQ paralog. Alongside with *wrn* and *recql5* genes, these strains provide highly useful tools to assess the genome maintenance activities of RecQ-family proteins in a vertebrate setting (unpublished data).

Dissemination of results, mentoring, awards and decorations

(*xxi*) We have acknowledged the K-116072 grant support in 10 papers published in highly-reputed international journals (2 of these papers appeared in PNAS, 2 in Nucleic Acids Research, and one in eLife, all with the corresponding authorship of the grant PI). 6 other manuscripts are in the pipeline and are about to be submitted. We disseminated our findings in the form of 16 domestic and 26 international conference presentations.

(*xxii*) During the grant period, one high-school student, 9 BSc students and 12 MSc students received research project mentorship in the PI's laboratory. A key member of the team, Gábor Harami successfully defended his PhD dissertation in 2016, and continued his work as a highly successful postdoctoral researcher in our laboratory (see below). During the project period, our students gave 12 presentations at TDK student conferences (4 of these were awarded with prizes), and won 2 ÚNKP (New National Excellence Program) fellowships. In 2020, Zoltán Kovács was awarded the Innovative PhD Prize of ELTE, and also won the prestigious Joseph Cours Fellowship.

(*xxiii*) In 2017, Gábor Harami was awarded the prestigious Premium Postdoctoral Fellowship of the Hungarian Academy of Sciences. In the next year, he won the prestigious Junior Prima Prize. In 2017, together with departmental colleagues András Málnási-Csizmadia and László Nyitray, the PI was awarded a joint Academy Prize of the Hungarian Academy of Sciences, for their work on motor enzyme mechanisms (see report in: Kovács Mihály, Málnási-Csizmadia András, Nyitray László: Motorfehérjékkel az Akadémiai Díjig. Biokémia (Journal of the Hungarian Biochemical Society), September 2017).

References – Publications that have arisen from the reported K-116072 grant support

* equal contributions

(1) Harami, G. M.*; Seol, Y.*; In, J.; Ferencziová, V.; Martina, M.; Gyimesi, M.; Sarlós, K.; Kovács, Z. J.; Nagy, N. T.; Sun, Y.; Vellai, T.; Neuman, K. C.*; Kovács, M.* Shuttling along DNA and Directed

Processing of D-Loops by RecQ Helicase Support Quality Control of Homologous Recombination. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (4), E466–E475. https://doi.org/10.1073/pnas.1615439114.

(2) Gyimesi, M.; Harami, G. M.; Kocsis, Z. S.; Kovács, M. Recent Adaptations of Fluorescence Techniques for the Determination of Mechanistic Parameters of Helicases and Translocases. *Methods San Diego Calif* **2016**, *108*, 24–39. https://doi.org/10.1016/j.ymeth.2016.04.028.

(3) Harami, G. M.*; Kovács, Z. J.*; Pancsa, R.; Pálinkás, J.; Baráth, V.; Tárnok, K.; Málnási-Csizmadia, A.; Kovács, M. Phase Separation by SsDNA Binding Protein Controlled via Protein-Protein and Protein-DNA Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (42), 26206–26217. https://doi.org/10.1073/pnas.2000761117.

(4) Seol, Y.*; Harami, G. M.*; Kovács, M.*; Neuman, K. C.* Homology Sensing via Non-Linear Amplification of Sequence-Dependent Pausing by RecQ Helicase. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.45909.

(5) Zacheja, T.; Toth, A.; Harami, G. M.; Yang, Q.; Schwindt, E.; Kovács, M.; Paeschke, K.; Burkovics, P. Mgs1 Protein Supports Genome Stability via Recognition of G-Quadruplex DNA Structures. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2020**. https://doi.org/10.1096/fj.202000886R.

(6) Horváth, G.; Egyed, O.; Tang, C.; Kovács, M.; Micsonai, A.; Kardos, J.; Toke, O. Ligand Entry in Human Ileal Bile Acid-Binding Protein Is Mediated by Histidine Protonation. *Sci. Rep.* **2019**, *9* (1), 4825. https://doi.org/10.1038/s41598-019-41180-7.

(7) Horváth, G.; Biczók, L.; Majer, Z.; Kovács, M.; Micsonai, A.; Kardos, J.; Toke, O. Structural Insight into a Partially Unfolded State Preceding Aggregation in an Intracellular Lipid-Binding Protein. *FEBS J.* **2017**, *284* (21), 3637–3661. https://doi.org/10.1111/febs.14264.

(8) Špírek, M.; Mlcoušková, J.; Belán, O.; Gyimesi, M.; Harami, G. M.; Molnár, E.; Novacek, J.; Kovács, M.*; Krejci, L.* Human RAD51 Rapidly Forms Intrinsically Dynamic Nucleoprotein Filaments Modulated by Nucleotide Binding State. *Nucleic Acids Res.* **2018**, *46* (8), 3967–3980. https://doi.org/10.1093/nar/gky111.

(9) Mills, M.*; Harami, G. M.*; Seol, Y.; Gyimesi, M.; Martina, M.; Kovács, Z. J.; Kovács, M.*; Neuman, K. C.* RecQ Helicase Triggers a Binding Mode Change in the SSB-DNA Complex to Efficiently Initiate DNA Unwinding. *Nucleic Acids Res.* **2017**, *45* (20), 11878–11890. https://doi.org/10.1093/nar/gkx939.

(10) Ferencziová, V.; Harami, G. M.; Németh, J. B.; Vellai, T.; Kovács, M. Functional Fine-Tuning between Bacterial DNA Recombination Initiation and Quality Control Systems. *PloS One* **2018**, *13* (2), e0192483. https://doi.org/10.1371/journal.pone.0192483.