E. coli DNA Repair Helicase Lhr is also a Uracil-DNA Glycosylase

Ed Bolt[1], Ryan J. Buckley[1], Karl M. Hanson[2], Nadia R. Ahmed[2], and Christopher Cooper[2]

[1]University of Nottingham School of Life Sciences
[2]University of Huddersfield School of Applied Sciences

January 18, 2023

Abstract

DNA glycosylases protect genetic fidelity during DNA replication by removing potentially mutagenic chemically damaged DNA bases. Bacterial Lhr proteins are well-characterized DNA repair helicases that are fused to additional 600-700 amino acids of unknown function, but with structural homology to SecB chaperones and AlkZ DNA glycosylases. Here we identify that *E. coli* Lhr is a uracil-DNA glycosylase that depends on an active site aspartic acid residue. We show that the Lhr DNA helicase activity is functionally independent of the uracil-DNA glycosylase activity, but that the helicase domains are required for fully active uracil DNA glycosylase activity. Consistent with uracil DNA glycosylase activity, deletion of *lhr* from the *E. coli* chromosome sensitized cells to oxidative stress that triggers cytosine deamination to uracil. The ability of Lhr to translocate single-stranded DNA and remove uracil bases suggests a surveillance role to seek and remove potentially mutagenic base changes during replication stress.

E. coli DNA Repair Helicase Lhr is also a Uracil-DNA Glycosylase

Ryan J. Buckley[1], Karl M. Hanson[2], Nadia R. Ahmed[2], Christopher D. O. Cooper[2 +] and Edward L. Bolt[1*]

[1]University of Nottingham, School of Life Sciences, UK

[2]School of Biological and Geographical Sciences, School of Applied Sciences, University of Huddersfield, UK.

*Correspondence to ed.bolt@nottingham.ac.uk

Summary

DNA glycosylases protect genetic fidelity during DNA replication by removing potentially mutagenic chemically damaged DNA bases. Bacterial Lhr proteins are well-characterized DNA repair helicases that are fused to additional 600-700 amino acids of unknown function, but with structural homology to SecB chaperones and AlkZ DNA glycosylases. Here we identify that *E. coli* Lhr is a uracil-DNA glycosylase that depends on an active site aspartic acid residue. We show that the Lhr DNA helicase activity is functionally independent of the uracil-DNA glycosylase activity, but that the helicase domains are required for fully active uracil DNA glycosylase activity. Consistent with uracil DNA glycosylase activity, deletion of *lhr* from the *E. coli*chromosome sensitized cells to oxidative stress that triggers cytosine deamination to uracil. The ability of Lhr to translocate single-stranded DNA and remove uracil bases suggests a surveillance role to seek and remove potentially mutagenic base changes during replication stress.

Keywords: Helicase; glycosylase; DNA repair; uracil; DNA replication

Introduction

Lhr (*L* arge *h* elicase-*r* elated) proteins are ATP-dependent 3' to 5' DNA translocases within the Superfamily 2 helicases [1]. The founder member of Lhr proteins was identified in bacteria [2], and subsequently Lhr was found to be widely distributed across all clades of archaea [3]. High amino acid sequence identity (typically

about 30%) between archaeal and bacterial Lhr proteins is limited to 800-900 amino acids that form helicase domains from the Lhr N-terminus ? called the 'Lhr-Core' [4]. Biochemical analyses of the Lhr-Core from the bacteria*Mycobacterium smegmatis* and *Pseudomonas putida* and from the archaeon *Methanothermobacter thermautotrophicus* have characterized Lhr translocation and helicase mechanism [5-7], and crystal structures of Lhr-Cores highlight similarities with translocation by the archaeal DNA repair helicase Hel308 [6, 8], especially in interactions between their winged helix and RecA-like domains [9, 10].

In addition to the Lhr-Core, bacterial Lhr proteins extend to 1400-1600 amino acids, in a C-terminal protein region of unknown function, called Lhr-CTD (Lhr-C-terminal domains). Structural modelling of the bacterial Lhr-CTD [11] and a subsequent cryo-EM structure [12], provided intriguing clues to Lhr-CTD function, including the presence of an array of tandem winged helix domains characteristic of the HTH_42 superfamily of proteins that have structural homology to the DNA glycosylase AlkZ [11]. Genetic analyses of the effects on bacterial and archaeal cells of deleting the *lhr* gene revealed mild sensitivities to agents that cause replication stress?UV irradiation [13] and azidothymidine (AZT) [14] ? and transcriptional up-regulation of*lhr* in response to mitomycin C [15]. In this work we report new insights about how Lhr contributes to DNA repair in bacteria. We demonstrate that the *E. coli* Lhr protein has uracil-DNA glycosylase activity, in addition to its well-characterized ATP-dependent DNA translocase functions, and that cells lacking Lhr are sensitive to oxidative stress.

Results

E. coli Lhr is an uracil DNA glycosylase requiring an active site aspartate

We investigated whether *E. coli* Lhr is capable of DNA glycosylase activity, as suggested from structural similarities between glycosylases and the uncharacterized C-terminal region of *E. coli* Lhr (Lhr-CTD, LHR amino acids 876-1538) [11, 12] (*Figure 1A*).*E. coli* Lhr-CTD protein fragment and full-length Lhr protein (1538 amino acid) were purified (*Figure 1B*) ? when Lhr-CTD (50 – 800 nM) mixed with a Cy5-end labelled 37-nt ssDNA molecule modified to contain a single uracil nucleotide located 18 nucleotides from the 5' ssDNA end (uracil-ssDNA) a single product was observed on alkaline treatment of reactions (*Figure 1C* , compare lanes 1-6 with 7-12), indicating DNA strand breakage at a DNA abasic site consistent with glycosylase activity. No product was observed from Lhr-CTD mixed with the same ssDNA lacking chemical modification (*Figure 1C* lanes 13-24).

To validate the uracil-DNA glycosylase activity of *E. coli*Lhr-CTD we sought to identify single amino acid substitutions that would inactivate it. Sequence alignment of *E. coli* Lhr-CTD and AlkZ, with which it has structural similarity [11, 12], were unproductive at identifying highly conserved residues because Lhr-CTD lacks the 'QxQ' motif characteristic of AlkZ protein active sites [16], therefore suggesting an alternate catalytic mechanism in LHR. Alternatively, potential active site amino acids were identified through visual scrutiny of the Phyre2 [17] predicted model of the *E. coli*Lhr-CTD, and in particular the positioning of side chains proximal to a proposed glycosylase active site (Figure 1D). Purified Lhr-CTD^{D1536A} (50 – 800 nM), gave no product when titrated into the uracil modified ssDNA after alkaline treatment, compared with Lhr-CTD (*Figure 1E*). We then tested whether substitution of the Lhr Asp-1536 residue inactivated uracil-DNA glycosylase active site chemistry or had some other effect on the protein that perturbed DNA binding. Unmutated Lhr-CTD was unable to form stable complex with DNA in EMSAs, when compared with full length Lhr (*Figure 2A*), therefore we purified and tested full Lhr^{D1536A}. Lhr was also active as an uracil-DNA glycosylase compared with Lhr-CTD (*Figure 2B*), but the Lhr^{D1536A} mutation inactivated glycosylase activity in agreement with inactive Lhr-CTD^{D1536A} fragment (*Figure 2C*). In EMSAs Lhr^{D1536A} formed stable complex with DNA similarly to Lhr (*Figure 2D*), therefore we conclude that Lhr is a uracil DNA glycosylase that requires an active site aspartic acid residue.

DNA glycosylase activity of *E. coli* Lhr is independent from its DNA helicase activity

Full length Lhr was substantially more active than Lhr-CTD as a uracil-DNA glycosylase when measured in assays as a function of time (Figure 3A) ? this may be explained by much more stable DNA binding by full length Lhr compared with Lhr-CTD that was observed in EMSAs (Figure 2A). We therefore continued to use

full-length Lhr to further investigate uracil-DNA glycosylase function against flayed duplex DNA molecules that are substrates for unwinding by the Lhr 3' to 5' DNA helicase activity [11]. For this work the duplex substrate was formed from annealing uracil-containing ssDNA with its unmodified complementary DNA strand, with uracil positioned 8nt from the fork branchpoint, 18 nt from the Cy5-DNA 5' end. Measured as a function of time, Lhr generated glycosylase product from the uracil duplex at least 5-fold more effectively than when incubated with uracil-ssDNA (*Figure 3B*), and Lhr was more active than Lhr-CTD on the uracil-fork DNA (*Figure 3C*). Neither Lhr or Lhr-CTD gave any glycosylase product when uracil was substituted for a single 8-oxoguanine residue at the same position in DNA (*Figure 3D*). The product of Lhr from uracil-DNA single strands or duplex migrated close to the 16 nt marker, indicating that the same glycosylase product was formed from both substrates. Glycosylase assay conditions in Figures 1-3 included Mg^{2+} in the reaction buffer, but Lhr was also active in buffers containing EDTA, Mn^{2+} and Ca^{2+} instead of Mg^{2+} (*Figure 3E* lanes 1-4) but was inactive as a glycosylase on DNA lacking a uracil residue (*Figure 3E* lanes 5-8). Lhr^{D1536} that is inactive as a uracil-DNA glycosylase was proficient at fork DNA unwinding (*Figure 3F*), and we therefore conclude that the uracil glycosylase activity of Lhr is functionally distinct from helicase activity, but we observe that glycosylase activity is enhanced when the helicase domains are present, by contributing to DNA binding (*Figure 2A*).

E. coli cells lacking LHR are sensitive to oxidative stress

Oxidative damage to DNA in *E. coli* cells includes deamination of cytosine to uracil and further oxidized uracil derivatives [18-21], triggering cytosine to adenine transversion mutations. We therefore assessed for a contribution from Lhr to DNA repair in *E. coli*cells that is consistent with *in vitro* uracil-DNA glycosylase activity. The *lhr* gene was deleted in *E. coli* MG1655 (Δ*lhr*) by recombineering, and we removed the inactivating antibiotic resistance marker, verified by sequencing across the deletion site. We first tested Δ*lhr* cells for sensitivity to azidothymidine (AZT), a previously reported phenotype for Lhr in*E. coli* cells [14]. In a viability plate assay after growing cells in broth (LB) containing a fixed 7.5 μg/mL AZT we observed 10-fold reduced viability of Δ*lhr* cells compared with wild type cells (*Figure 4A*), and similar moderate sensitivity of Δ*lhr*cells across AZT concentrations (*Figure 4B*), agreeing with the previous study [14]. We next measured survival of cells when grown in media containing hydrogen peroxide as a potent oxidizing agent. Hydrogen peroxide (12.5 mM) added to growth media after cells had reached OD_{600} of 0.3 resulted in significantly reduced growth of Δ*lhr* cells in exponential phase compared with wild type cells (*Figure 4C*). This agreed with 10-100-fold reduced cell viability compared with wild type cells when Δ*lhr* cells grown in the same way, but without hydrogen peroxide, were then spotted onto LB agar containing increasing concentrations of hydrogen peroxide to count their viability (*Figure 4D*).

Discussion

We provide biochemical evidence that *E. coli* Lhr is a uracil-DNA glycosylase (Lhr-UDG), a new function for bacterial Lhr proteins alongside their well-characterized 3' to 5' single DNA translocation activity that is stimulated by forked or flayed DNA substrates [5, 7, 11]. We show that the Lhr-UDG activity requires an active site aspartate residue (Asp-1536), similarly to the active site aspartate general base (Asp-62) that is essential for major groups of UNG/UDG proteins [22]. The Lhr UDG function is positioned in the previously uncharacterized Lhr-CTD ? though this fragment of Lhr was proficient as a 'stand-alone' uracil-DNA glycosylase, its activity was significantly increased by the presence of the Lhr helicase domains, probably by the helicase domains providing more stable DNA binding compared with Lhr-CTD, observed in EMSAs. Inactivating the Lhr-UDG activity did not inactivate DNA unwinding by Lhr, providing further support for the DNA binding functions of Lhr being concentrated in the helicase domains.

Loss of Lhr from bacterial cells (Δ*lhr*) causes mild sensitivity to AZT [14], a phenotype we also observed after generating Δ*lhr* cells and removing the inserted antibiotic resistance marker. We identified that Δ*lhr* cells were also significantly more sensitive than lhr^{+} cells to oxidative stress induced by hydrogen peroxide, which is one of several routes causing genetic damage by cytosine deamination in bacterial cells. This therefore supports our *in vitro* observations of Lhr-UDG function. UDGs are ubiquitous in nature, although this is the first report of a UDG fused to a DNA helicase. *E. coli* has a canonical UDG enzyme that functions in

global DNA repair coupled with stable DNA replication ? upregulation of Mycobacterial Lhr in response to mitomycin C treatment [15], and the sensitivity of *E. coli* cells to the polymerase inhibitor AZT when they lack *lhr* , may indicate that Lhr is activated as part of bacterial responses to specific forms of replication-stress. In this context removal of uracil from DNA by Lhr may protect genetic fidelity at sites that are overcoming blocked DNA replication. We cannot exclude that Lhr may be able to remove other lesions or chemical modifications from DNA, although we observed that it was inactive as a glycosylase against 8-oxoguanine, suggesting that is has at least specificity for recognizing pyrimidine damage over purines.

Experimental Procedures

Proteins

DNA sequences of primers and substrates, plasmids and *E. coli*strains are detailed in Supplementary data. *E. coli* MG1655 gene b1653, encoding Lhr, was PCR-amplified from genomic DNA, and cloned into pET14b using *Nde* I and *Hin* DIII restriction sites generating pRJB28 for expression of hexa-histidine tagged Lhr. Lhr-CTD (amino acids 876-1538) was amplified and cloned by ligation independent cloning (LIC) The resulting plasmid based on vector pNH-TrxT (GU269914.1 [23]). These plasmids were used to generate Lhr^{D1536A} using mutagenic primers in PCR by Q5 hot start polymerase, and resulting reactions were treated with *Dpn* I, T4 polynucleotide kinase, and DNA ligase. Plasmid DNA was extracted and sequenced from colonies after transforming reaction mixtures into *E. coli* .

Lhr and C-Lhr proteins were over-expressed in *E. coli* Rosetta 2 cells grown in MU broth with ampicillin and chloramphenicol selection. Cells were grown with shaking at 37°C to OD_{600} of 1.2 and transferred to an ice slurry for cooling before addition of IPTG (0.8 mM). Growth was continued for 10 hours at 18°C, cells were harvested and resuspended in 20 mM HEPES pH 8.0, 1.5 M ammonium sulfate, 20 mM imidazole, 10% (w/v) glycerol (Ni-NTA buffer A) containing 0.1 mM phenylmethylsulphonyl fluoride (PMSF). This process and purification was also followed for obtaining Lhr^{D1536A} protein. Cells were thawed and sonicated on ice, clarified by centrifugation, and soluble proteins were loaded into a 5 ml butyl sepharose column equilibrated with 20 mM HEPES pH 8.0, 1.5 M ammonium sulfate, 10% (w/v) glycerol (hydrophobic salt buffer A). The column was washed with 20 mM HEPES pH 8.0, 900 mM ammonium sulfate. Then a 5 ml Ni-NTA column pre-equilibrated with Ni-NTA buffer A was attached in tandem with the butyl sepharose column and columns washed with 20 mM HEPES pH 8.0 and 10% glycerol until no proteins were detectable by UV monitoring as eluting from the columns. The butyl sepharose column was removed and Lhr was eluted from the Ni-NTA column by increasing imidazole to 500 mM in 20 mM HEPEs and 10% glycerol. Lhr-containing fractions were pooled and dialyzed overnight into 20 mM HEPES pH 8.0, 150 mM NaCl, 10% (w/v) glycerol (low salt buffer A) and loaded into a 1 ml Q-sepharose column. Lhr eluted in an increasing gradient of NaCl to 1.5 M. Lhr fractions were pooled and dialyzed overnight for storing in 20 mM HEPES pH 8.0, 150 mM NaCl and 35% (w/v) glycerol for aliquoting, flash freezing, and storage at -80°C. *E. coli* uracil-DNA glycosylase and formamidopyrimidine DNA glycosylase (Fpg) control proteins (Figure 3) were purchased from New England Biolabs.

In vitro DNA binding, unwinding and glycosylase assays

DNA strands for substrate formation (supplemental data) were synthesized with Cy5 end-label. DNA binding was assessed using electrophoretic mobility shift assays (EMSAs). Reactions were incubated at 37°C for 20 minutes in helicase buffer (HB); 20 mM Tris pH 7.5, 10% (v/v) glycerol, 100 μg/ml BSA, using 12.5 nM Cy5-fluorescently labelled DNA substrate, 25 mM DTT and 5 mM EDTA, and then placed on to ice for 10 minutes. Orange G and 80% (v/v) glycerol (OG) was added to load reactions onto a 5% acrylamide TBE gel that was electrophoresed for 1 hour 30 minutes at 140 V. Gels were imaged using a Typhoon phosphor-imager (Amersham) at 633 nm using a R765 filter for Cy5 detection.

DNA unwinding assays were at 37°C in reactions containing buffer HB, 12.5 nM Cy5-fluorescently labelled DNA substrate, 25 mM DTT, 1.25 μM unlabeled 'trap' DNA, 5 mM ATP and 5 mM $CaCl_2$. Reactions were pre-incubated at 37°C for 5 minutes without the 'trap' or ATP before they were added together to start the reactions for 30 minutes at 37°C, stopped by addition of stock stop solution (4 μl per 20 μl reaction); 2

mg/mL proteinase K in 200 mM EDTA and 2.5% (w/v) SDS. OG dye was added for electrophoresis through a 10% acrylamide TBE gel for 45 minutes at 150 V. Gels were imaged using a Typhoon phosphor-imager (Amersham) at 633 nm using a R765 filter for Cy5 detection.

DNA glycosylase reactions were at 37°C in reaction mixtures containing buffer HB, 12.5 nM Cy5-fluorescently labelled DNA substrate, 25 mM DTT, 5 mM ATP, 4 mM $MnCl_2$ and 4 mM $CaCl_2$. Reactions were pre-incubated at 37°C before being initiated by addition of Lhr protein and (unless in a time course assay) allowed to continue for 30 minutes before addition of stock stop solution and 4 µl of 1 M NaOH. Reaction samples were boiled for 5 minutes and formamide added before loading into a 15% denaturing (8 M urea) acrylamide TBE gel for 4 hours at 5 watts per gel. Gels were imaged using a Typhoon phosphor-imager (Amersham) at 633 nm using a R765 filter for Cy5 detection, generating TIFFs that were measured using Gel Analyzer 19.1 (Lazar) software. Graphs of glycosylase activity were generated using Prism (GraphPad).

Generation of a chromosomal deletion of*E. coli lhr*

DNA constructs and strain genotypes are presented in the Supplementary material. *Lhr* deletion was by recombineering [24] and P1 transduction of an FRT (FLP recognition target) flanked Kan^r marker. To generate an effective P1 stock the overnight culture was used to inoculate 8 ml of Mu broth containing 6 mM $CaCl_2$. A sample of the cells (0.1 mL) grown at 37°C to OD_{600} 0.8-1.0 in a shaking water bath was added to four overlay tubes each containing 3 mL of 0.4% w/v Mu broth agar held at 42°C. P1 phage stock was diluted 10-100-fold in MC buffer (100 mM $MgSO_4$, 5 mM $CaCl_2$) and 0.05 mL, 0.1 mL or 0.2 mL of this diluted phage was added to the overlay tubes containing cells and molten agar and gently mixed. The remaining tube was left without phage as a control. The contents of each overlay tube was poured onto P1 agar plates and left to set for overnight growth at 37°C for 18 hours. Soft agar from phage-lysed plates was added to 1 ml of MC buffer (100 mM $MgSO_4$ and 5 mM $CaCl_2$) and 0.5 ml of chloroform for vigorous mixing before centrifugation at 5752 rcf for 20 minutes at 4°C. The supernatant was retrieved and mixed with chloroform (0.5 mL) for storage at 4°C as a P1 phage stock. MG1655 recipient strain was grown in a Mu Broth to OD_{600} 0.8 using a shaking water bath. Cells were pelleted, resuspended in 1 ml of MC buffer, and left at 25°C for 10 minutes. 0.2 ml of cells were added into 3 overlay tubes containing 0 ml, 0.05 ml and 0.2 ml of P1 lysate produced previously and incubated for 30 minutes at 37°C. Cell/P1 lysate mix was added to 2.5 ml of 0.6% agar, mixed gently and poured onto Mu Broth agar plates containing 30 µg/ml kanamycin and left to set. Plates were grown for 1-2 days, lid-up, at 37°C to allow colonies to develop. Colonies were then picked and purified by streaking onto Mu broth agar plates containing no antibiotic. This was repeated 3 times before plating again onto agar containing 30 µg/ml kanamycin for confirmation of gene knockout and Kan^r-FTR insertion.

Successful P1 treated MG1655 cells were transformed with pCP20. Transformants were picked and used to inoculate 8 ml of Mu broth containing no antibiotic. Culture was grown overnight at 45°C in a shaking water bath FLP recombinase expression and plasmid curation. Cells were then streaked onto Mu Broth agar plates to produce single colonies and grown at 37°C overnight. Colonies were re-streaked 3 times before replica plating onto Mu Broth agar plates containing 50 µg/ml ampicillin, 30 µg/ml kanamycin and then no antibiotic to confirm loss of the pCP20 plasmid. Isolates which only grew on the no antibiotic agar plates were grown overnight for glycerol stock production and streaked a further time for colony PCR diagnostic confirmation.

E. coli viability spot assays

Cell viabilities were measured from liquid cultures grown to OD_{600} 0.3-0.4 in a shaking water bath at 37°C monitored in the growth tubes by using a Spectronic 20+. Cells were then treated by addition to the growth media of hydrogen peroxide (H_2O_2) or AZT at concentrations stated in the results. Cells were grown for a further 30 minutes and then serially diluted into 1x M9 medium to arrest growth for spotting (10 ul) on to agar plates grown overnight in a 37°C incubator. For comparing growth curves cells were grown to OD_{600} 0.3-0.4 and then transferred into a 24-well flat-bottomed plate and H_2O_2 was added to appropriate wells to

the given concentration from a 0.98 M stock. Growth in the plates was monitored with orbital shaking in a FLUOstar microplate reader (BMG Labtech). OD_{600} readings were taken every 30 minutes in this time, and data was extracted and analyzed using Prism (GraphPad) software.

Author Contributions

ELB and CDOC designed the project and with RJB wrote the paper. RJB, KMH, NA and CDOC performed experiments, analyzed data, and generated images.

Acknowledgments

The work was supported by BBSRC grant BB/T006625-1 (ELB) and the BBSRC Doctoral Training Partnership (ELB/RJB).

References

1. Hajj, M., et al., *Phylogenetic Diversity of Lhr Proteins and Biochemical Activities of the Thermococcales aLhr2 DNA/RNA Helicase.*Biomolecules, 2021. **11** (7).

2. Reuven, N.B., et al., *The gene for the longest known Escherichia coli protein is a member of helicase superfamily II.* J Bacteriol, 1995. **177** (19): p. 5393-400.

3. Chamieh, H., H. Ibrahim, and J. Kozah, *Genome-wide identification of SF1 and SF2 helicases from archaea.* Gene, 2016.**576** (1 Pt 2): p. 214-28.

4. Ejaz, A., Y. Goldgur, and S. Shuman, *Activity and structure of Pseudomonas putida MPE, a manganese-dependent single-strand DNA endonuclease encoded in a nucleic acid repair gene cluster.* J Biol Chem, 2019. **294** (19): p. 7931-7941.

5. Ordonez, H. and S. Shuman, *Mycobacterium smegmatis Lhr Is a DNA-dependent ATPase and a 3'-to-5' DNA translocase and helicase that prefers to unwind 3'-tailed RNA:DNA hybrids.* J Biol Chem, 2013.**288** (20): p. 14125-34.

6. Ejaz, A., et al., *Structure of mycobacterial 3'-to-5' RNA:DNA helicase Lhr bound to a ssDNA tracking strand highlights distinctive features of a novel family of bacterial helicases.* Nucleic Acids Res, 2018. **46** (1): p. 442-455.

7. Ejaz, A. and S. Shuman, *Characterization of Lhr-Core DNA helicase and manganese- dependent DNA nuclease components of a bacterial gene cluster encoding nucleic acid repair enzymes.* J Biol Chem, 2018.**293** (45): p. 17491-17504.

8. Buttner, K., S. Nehring, and K.P. Hopfner, *Structural basis for DNA duplex separation by a superfamily-2 helicase.* Nat Struct Mol Biol, 2007. **14** (7): p. 647-52.

9. Northall, S.J., et al., *DNA binding and unwinding by Hel308 helicase requires dual functions of a winged helix domain.* DNA Repair (Amst), 2017. **57** : p. 125-132.

10. Johnson, S.J. and R.N. Jackson, *Ski2-like RNA helicase structures: common themes and complex assemblies.* RNA Biol, 2013.**10** (1): p. 33-43.

11. Buckley, R.J., et al., *Mechanistic insights into Lhr helicase function in DNA repair.* Biochem J, 2020. **477** (16): p. 2935-2947.

12. Warren, G.M., et al., *Oligomeric quaternary structure of Escherichia coli and Mycobacterium smegmatis Lhr helicases is nucleated by a novel C-terminal domain composed of five winged-helix modules.*Nucleic Acids Res, 2021. **49** (7): p. 3876-3887.

13. van Wolferen, M., X. Ma, and S.V. Albers, *DNA Processing Proteins Involved in the UV-Induced Stress Response of Sulfolobales.* J Bacteriol, 2015. **197** (18): p. 2941-51.

14. Cooper, D.L., D.C. Boyle, and S.T. Lovett, *Genetic analysis of Escherichia coli RadA: functional motifs and genetic interactions.* Mol Microbiol, 2015. **95** (5): p. 769-79.

15. Rand, L., et al., *The majority of inducible DNA repair genes in Mycobacterium tuberculosis are induced independently of RecA.* Mol Microbiol, 2003. **50** (3): p. 1031-42.

16. Mullins, E.A., et al., *Structure of a DNA glycosylase that unhooks interstrand cross-links.* Proc Natl Acad Sci U S A, 2017.**114** (17): p. 4400-4405.

17. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10** (6): p. 845-58.

18. Almatarneh, M.H., C.G. Flinn, and R.A. Poirier, *Mechanisms for the deamination reaction of cytosine with H2O/OH(-) and 2H2O/OH(-): a computational study.* J Chem Inf Model, 2008. **48** (4): p. 831-43.

19. Cadet, J. and J.R. Wagner, *DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation.* Cold Spring Harb Perspect Biol, 2013. **5** (2).

20. Kreutzer, D.A. and J.M. Essigmann, *Oxidized, deaminated cytosines are a source of C –> T transitions in vivo.*Proc Natl Acad Sci U S A, 1998. **95** (7): p. 3578-82.

21. Krokan, H.E., F. Drablos, and G. Slupphaug, *Uracil in DNA–occurrence, consequences and repair.* Oncogene, 2002.**21** (58): p. 8935-48.

22. Aravind, L. and E.V. Koonin, *The alpha/beta fold uracil DNA glycosylases: a common origin with diverse fates.* Genome Biol, 2000.**1** (4): p. RESEARCH0007.

23. Savitsky, P., et al., *High-throughput production of human proteins for crystallization: the SGC experience.* J Struct Biol, 2010.**172** (1): p. 3-13.

24. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97** (12): p. 6640-6645.

Figure Legends

Figure 1: The *E. coli* Lhr-CTD is a uracil DNA glycosylase requiring a catalytic aspartic acid

(A) AlphaFold 2 structural model of *E. coli* Lhr that is based on strong homology with the cryo-EM structure of Lhr helicase core and Lhr-CTD from *M. smegmatis* , respectively PDB: 5V9X and PDB:7LHL. The *E. coli* Lhr-core helicase (amino acids 1-897) contains RecA domains, a beta-sheet bundle (β) and a winged helix domain (WH) as indicated. Lhr-CTD (amino acids 898-1538) comprises folds with structural homology to SecB chaperones and AlkZ glycosylases, as indicted.

(B) Coomassie stained SDS-PAGE acrylamide gels showing purified Lhr and Lhr-CTD, with molecular mass ladder (M) values in kDa.

(C) Products from mixing Lhr-CTD (50, 100, 200, 400 and 800 nM) with 5' Cy5-ssDNA (12.5 nM) containing a d-Uracil base 18 located nucleotides from the fluorescent moiety as indicated (lanes 1-12), seen in a 15% denaturing acrylamide TBE gel. Addition of NaOH (lanes 8-12) causes β/δ elimination at the site of the abasic DNA product, resulting in DNA backbone cleavage. This confirms glycosylase protein activity. Marker (M) is made from known lengths of 5' Cy5 ssDNA.

(D) As for (C) in reactions containing unmodified 5' Cy5-ssDNA (12.5 nM).

(E) Phyre2 structural model of *E. coli* Lhr-CTD with predicted active site residues as labelled, including Lhr-CTD residue Asp-1536 that we mutated in this work.

(F) Products from mixing Lhr-CTD and Lhr-CTD^{D1536A}proteins with 12.5 mM d-uracil containing 5' Cy5-ssDNA substrate, viewed in a 18% acrylamide denaturing TBE gel. Product formation is shown every 5 minutes for 30 minutes, observing no glycosylase activity from Lhr-CTD^{D1536A}.

Figure 2: Lhr^{D1536A} is inactive as a glycosylase but binds to DNA

(**A**) EMSA assays showing Lhr (12.5, 25, 50, 100 and 200nM) complexes bound to DNA (12.5 nM) that are stable migrating through a 5% acrylamide TBE gel, compared with Lhr-CTD at the same concentrations.

(**B**) Products of Lhr glycosylase activity seen in an 18% acrylamide denaturing TBE gel were absent when reactions contained Lhr^{D1536A}. Proteins were used at 25, 50, 100 and 200 nM, with 12.5 nM of d-uracil containing 5' Cy5-ssDNA substrate.

(**C**) EMSA showing that Lhr^{D1536A} and Lhr (12.5, 25, 50, 100 and 200 nM) form stable complex with DNA in a 5% acrylamide TBE gel

Figure 3: Lhr is inactive against 8-oxoguanine, and its uracil DNA glycosylase activity on duplex DNA functions independently from Lhr helicase activity

(**A**) Time-dependent uracil DNA glycosylase activity of Lhr (50 nM) compared with Lhr^{D1536A}. The data shows means of glycosylase activity (n=3, with bars for standard error) alongside a representative gel used for quantification.

(**B**) Comparison of Lhr (50 nM) glycosylase activity on ss-, ds- and forked d-uracil containing DNA substrates (12.5 nM) as a function of time, with samples taken at time points indicated ? plots are means of two independent experiments showing standard error bars.

(**C**) Time-course assays (10, 20, 30 minutes) showing products from Lhr and Lhr-CTD (each 80 nM) mixed with the preferred flayed duplex uracil-DNA, seen in an 18% acrylamide denaturing TBE gel. Known length DNA strands are shown (M) and the positive control reaction (+ve) is product from 5 units of *E. coli* uracil DNA glycosylase.

(**D**) As for (C) except d-uracil DNA was replaced with otherwise identical 8-oxo-d-Guanine DNA, and the control reaction (+ve) shows product formed by 5 units of formamidopyrimidine DNA glycosylase (Fpg) protein.

(**E**) Lhr (80 nM) uracil-DNA glycosylase activity seen as products in 18% acrylamide denaturing TBE gels (lanes 1-4), after 30-minute reactions in either EDTA, manganese or calcium, each replacing magnesium as indicated, compared with unmodified DNA (lanes 5-8).

(**F**) DNA unwinding by Lhr and Lhr^{D1536A}proteins (12.5, 25, 50, 100 and 200 nM) on 12.5 nM of 5' Cy5 labelled flayed duplex DNA, seen in 10% acrylamide TBE gel.

Figure 4: *E. coli* cells lacking Lhr are sensitive to oxidative stress

(**A**) Viability spot tests showing moderately increased sensitivity of Δ*lhr* cells to AZT (7.5 μg/mL) compared with wild type (wt) cells, and (**B**) represented in viability curves when Δ*lhr* and wild type cells were grown independently in media containing AZT at 2.5, 5, 7.5 or 10 μg/mL. The plots show grow relative to wild type cells grown in media lacking AZT.

(**B**) Growth of Δ*lhr* and wild type cells monitored in 96 well plates in media containing 12.5 mM H_2O_2, and (**C**) viability spot tests comparing Δ*lhr* and wild type cells grown in H_2O_2 added to media at 1.5625, 3.75, 6.25 and 12.5 nM.

Hosted file

`Main Figures - Final.pptx` available at https://authorea.com/users/576754/articles/619423-e-coli-dna-repair-helicase-lhr-is-also-a-uracil-dna-glycosylase