Role of the nucleotide excision repair pathway proteins (UvrB and UvrD2) in recycling UdgB, a base excision repair enzyme in *Mycobacterium smegmatis*

Running title: UvrB and UvrD2 facilitate turnover of AP-site bound UdgB

Indu Kapoor[a], Abhirup Shaw[a], Arindam Naha[a], Elhassan Ali Fathi Emam[a] and Umesh Varshney[a, b, *]

[a] Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore, 560012, India

[b] Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, 560064, India

[*]*Correspondence to*: Umesh Varshney,

Phone: +918022932686, Fax: +918023602697,

E-mail: varshney@iisc.ac.in, uvarshney@gmail.com

Keywords: UdgB, UvrD2, UvrB, Nucleotide excision repair, Transcription Coupled Repair, mycobacteria

S1. Supplementary Methods

S1.1. Generation of constructs: pJAM2 and its derivatives, pJAM_FLAG and pJAMH_His, were used to overexpress and purify proteins. To derive pJAM_FLAG, a DNA sequence encoding FLAG-tag followed by a stop codon was introduced at XbaI restriction site, using oligomers UVL1 and UVL2. The plasmid pJAM_FLAG retains a single XbaI restriction site before the FLAG-tag and His_6-tag present after FLAG-tag is not translated. A hygromycin resistant version of pJAM2 plasmid, pJAMH_His, was made by deleting Kan^R cassette using NdeI restriction site. The vector fragment was end-filled by Klenow DNA polymerase, purified by phenol-chloroform extraction, and ligated with Hyg^R cassette excised from $pMV261^{hyg}$ using HpaI and DraI restriction sites.

The PCR carried out in 20 µl volume contained 250 ng genomic DNA of *Msm* mc^2155 strain, 0.5 units of Q5 DNA polymerase (NEB), 1X Q5 reaction buffer, 1X GC enhancer, 250 µM dNTPs, and 10 pmol each of the forward and reverse primers. Reaction conditions included an initial denaturation at 98°C for 2 min, followed by 30 cycles of 98 °C for 40 s, 50 – 60 °C for 35 s and 72 °C for 1 - 3 min and a final extension at 72 °C for 5 min.

For cloning in one or more plasmids, pJAM2, pJAM_FLAG and pJAMH_His, ORFs (*MSMEG_5031* (*Msm_udgB*), *MSMEG_5423* (*Msm_mfd*), *MSMEG_5534* (*Msm_uvrD1*), *MSMEG_1952* (*Msm_uvrD2*), and *MSMEG_2758* (*Msm_mysA*)) from *Msm* mc^2155 were PCR amplified using the respective primers with flanking BamHI or XbaI cloning sites (Table S1). Purified PCR products were ligated with the vector DNAs using BamHI and/or XbaI restriction sites, to obtain C-terminally His_6-tagged or FLAG-tagged proteins. In parallel, pJAM2 harboring a C-terminally truncated variant of *Msm_mfd*, encoding the first 1040 amino acids was also generated using *Msm_mfd*_BamHI_Fp and *Msm_mfd*_1040_XbaI_Rp primers.

To generate pMV261h_*udgB*, *Msm_udgB* encoding ORF, *MSMEG_5031,* was PCR amplified using the phosphorylated primers *Msm_udgB*_PFp (binds 187 bp upstream of ORF) and *Msm_udgB*_Rp (binds 123 bp downstream of ORF). PCR amplicon (1200 bp) was ligated into PvuII digested pMV261hyg vector.

Table S1. List of DNA oligomers used in the study to generate constructs and verify strains

DNA oligomer	**Sequence (5′ to 3′), restriction site is underlined**
*Msm_udgB*_PFp	GACACCGGTACCGGCGATCAGTG
*Msm_udgB*_Rp	CAGAAGCTTTGGTATCCGAGCCG
*Msm_udgB*_BamHI_Fp	CTGGTGGGATCCATGGGACCGATTTTC
*Msm_udgB*_XbaI_Rp	CGGTCATCTAGATTGCCCGTCACG
*Msm_uvrD1*_BamHI_Fp	TACGGATCCATGACTTCCCCC
*Msm_uvrD1*_XbaI_Rp	GGAGTCTAGAGAGCTTCTGCAG
*Msm_uvrD2*_BamHI_Fp	CGGTGTGGATCCTTGTCGGGCGGT
*Msm_uvrD2*_XbaI_Rp	GCAGCATCTAGACGAATCGTGGCGGTT
*Msm_mfd*_BamHI_Fp	ATCCAGGATCCATGACCGCACCG
*Msm_mfd*_XbaI_Rp	GCACCTTCTAGATCGCTTCGCCTC
*Msm_mfd*_1040_XbaI_Rp	GGACTTCTAGACGGTGTGGCAAC
*Msm_mysA*_XbaI_Fp	AGGCTCTAGAGTGGCAGCGACAAAG
*Msm_mysA*_XbaI_Rp	CATTTCTAGACTAGTCCAGGTAGTCGCG
UVL1	CTAGAGACTACAAGGACGACGACGACAAGTGAG
UVL2	CTAGCTCACTTGTCGTCGTCGTCCTTGTAGTCT
MsuvrB-Fp	CGCACCGGCAAACCCTTCG (binds 117 bps inside the ORF)
MsuvrB-Rp	CGCGAATTCAGTCACGACGAC (binds 50 bps downstream the ORF)

S1.2. Generation of *Msm uvrB*$^{-}$ *udgB*$^{-}$ strain: The plasmid, pPR-*Ms_uvrB-kan*^{R}, was electroporated in *Msm udgB*$^{-}$ for allelic exchange and the transformants were selected on LBT agar in the presence of Kan, Hyg and Gm at 30 °C. The transformants were inoculated in 2 ml of LBT Kan Hyg and incubated at 30 °C for 2 days. Bacterial cultures were serially diluted till 10^{-2}, and aliquots of 50 µl from each serial dilution were plated on LBT agar supplemented with Kan, Hyg and 10% sucrose. Following incubation at 39 °C for 3 days, isolated colonies obtained on the plate were patched on LBT agar containing either Kan, Hyg and 10% sucrose or Hyg and Gm. Colonies that grew only on LBT agar plates containing Kan, Hyg and 10% sucrose were further verified by PCR using ORF primers, *MsuvrB*-Fp and *MsuvrB*-Rp.

S1.3. Overexpression and purification of proteins

S1.3.1. *Msm*UvrA, *Msm*UvrB and *Msm*UvrC: N-terminally His_6-tagged NER proteins were overexpressed and purified from *Msm* [1].

S1.3.2. *Msm*UdgB_FLAG and *Msm*UdgB_His_6: C-terminally FLAG-tagged UdgB (used for *in vitro* pulldown assays) was purified from *Msm* mc^2155 strain harboring the pJAM_*udgB*_FLAG plasmid. A single colony from the transformation plate was inoculated in 2 ml LBT Kan and grown at 37 °C, 180 rpm for 2 days. LBT Kan (50 ml) was inoculated with 1% of saturated culture and incubated at 37 °C, 180 rpm for 12-14 h. Inoculation of 3 l LBT Kan was done with 1% primary culture and incubated at 37 °C, 160 rpm till OD_{600} reached 0.6. Protein was expressed using 0.5% acetamide for 4 h. The cell pellet was washed and resuspended in ice-cold 25 ml lysis buffer (50 mM Na_3PO_4 pH 7, 50 mM NaCl, 10% glycerol, and 2 mM β-mercaptoethanol). The following steps were performed at 4 °C. The cells were lysed by sonication and the lysate was subjected to centrifugation at 24K rpm, 4 °C for 1 h. The soluble cell-free extract (CFE) was loaded on heparin column (GE Healthcare), washed with 5 column

volumes of lysis buffer containing 200 mM NaCl and the protein was eluted with a gradient (0% to 100%) of NaCl in elution buffer [50 mM Na_3PO_4 pH 7, 1 M NaCl, 10 % glycerol (v/v) and 2 mM β-mercaptoethanol) at a flow rate of 1 ml/min for 30 min. The fractions eluted with 500 mM to 800 mM NaCl were enriched for the desired protein. The protein containing fractions were pooled and dialyzed to gradually obtain 50 mM NaCl concentration. The dialyzed fraction was loaded on Mono-S (GE Healthcare), washed with 3 column volumes of lysis buffer and the protein was eluted with a gradient of NaCl as explained above. Near homogenous protein fractions were pooled, dialyzed against buffer (50 mM Na_3PO_4 pH 7, 150 mM NaCl, 50% glycerol, 2 mM β-mercaptoethanol), quantified and stored at -20 ℃.

Similarly, C-terminally His_6-tagged UdgB (used for *in vitro* enzymatic reactions) was purified from *Msm ung*- strain harboring the pJAMH_*udgB*_His plasmid. Protein was overexpressed as described above and purified as discussed before [2].

S1.3.3. Translocases: *Msm*Mfd, *Msm*Mfd_1040 and *Msm*UvrD2 were purified using the construct pJAM_*mfd* or pJAM_*mfd*_1040 or pJAM_*uvrD2*_His, respectively in *Msm* mc^2155. Proteins were expressed using 0.5% acetamide followed by 6-8 h incubation at 37 °C and 180 rpm. Like UvrA, all three proteins were expressed in the soluble fraction and C-terminally His_6-tagged proteins were purified by affinity chromatography (Ni-NTA) followed by size-exclusion chromatography.

S1.3.4. *Msm*RNAP: RNAP was purified from *Msm* mc^2155/ pJAM_*mysA* strain by overexpressing the sigma factor, SigA (σ^A), to enrich the RNAP holoenzyme with σ^A factor [3].

Table S2. List of DNA oligomers used for biochemical assays

DNA Oligomer	Sequence (5′ to 3′), complementary sequence is underlined	Description
GU9	CTCAAGTGUAGGCATGCTTTTGCATGCCTGCACTTGA	37 nt long, makes a stem (16 bp) loop (4 nt) structure and contains G:U at 9th position in stem
GU26	GACTACGTACTGTCACGCTCAAGTGUAGGCATGCATCAGGCCAGATCTGCTTTTTAGCAGATCTGGCCTGATGCATGCCTGCACTTGAGCGTGACAGTACGTAGTC	106 nt long, makes a stem (51 bp) loop (4 nt) structure and contains G:U at 26th position in stem
UVL3	CTACTACGTACTGTCAGGGGTCCATUTTCACCGGAATCAGGCCAGATCTGCTAGTCTAGAGGATGCTAAGGTC	73 nt long ssDNA containing uracil at 25th position
UVL4	GCAGATCTGGCCTGATTCCGGTGAAGATGGACCCCTGACAGTACGTAGTAG	51 nt long and complementary to UVL3
5' Biotinylated SSU9	C(B)TCAAGTGUAGGCATGCAAGAGCT	24 nt long ssDNA harboring uracil at 9th position and biotin modification at 5' end
UVL9	AGCTCTTGCATGCCTGCACTTGAGTAGTCTAGAGGATGC	39 nt long and complementary to biotinylated SSU9
5' Biotinylated oligomer	GAC(B)GCTGCCGAATTCTGGCTTGCTAAAGGATAGTCGAATTTTCTCATTTT	51 nt long ssDNA with biotin modification at 5' end [4]

S2. Supplementary Results

S2.1. Raw data for growth curves

NOTE: Values starred (*) and highlighted in blue color were excluded while plotting the growth curves.

Time (hours)	*Msm* WT /pMV261h (U)				*Msm* WT /pMV261h (2.5 mM H_2O_2)			
0	0.106	0.103	0.102	0.105	0.11	0.114	0.107	0.108
3	0.148	0.157	0.156	0.161	0.166	0.176	0.156	0.16
6	0.148	0.153	0.147	0.152	0.135	0.143	0.132	0.132
9	0.197	0.201	0.199	0.206	0.152	0.169	0.161	0.161
12	*1.533**	0.302	0.287	0.29	0.196	0.224	0.199	0.203
15	0.432	0.474	0.456	0.458	0.299	0.347	0.317	0.315
18	0.645	0.717	0.691	0.686	0.463	0.542	0.485	0.495
21	0.863	0.915	0.896	0.89	0.665	0.778	0.766	0.759
24	1.047	1.099	1.083	1.083	0.882	0.979	0.964	0.956
27	1.184	1.237	1.216	1.232	1.078	1.163	1.144	1.157
30	1.263	1.302	1.287	1.304	1.241	1.28	1.268	1.293
33	1.312	1.34	1.313	1.323	1.318	1.336	1.342	1.363
36	1.317	1.35	1.321	1.346	1.349	1.353	1.36	1.379
39	1.327	1.359	1.317	1.342	1.369	1.374	1.372	1.397
42	1.32	1.339	1.318	1.333	1.377	1.372	1.382	1.411
45	1.336	1.345	1.317	1.33	1.359	1.361	1.369	1.387
48	1.321	1.339	1.312	1.328	1.363	1.364	1.364	1.38

Time (hours)	*Msm* WT /pMV261h_*udgB* (U)				*Msm* WT /pMV261h_*udgB* (2.5 mM H_2O_2)			
0	0.106	0.101	0.105	0.104	0.112	*0.103**	0.11	0.114
3	0.162	0.153	0.167	0.162	0.166	*0.143**	0.157	0.162
6	0.149	0.126	0.139	0.137	0.138	*0.095**	0.121	0.125
9	0.195	0.154	0.174	0.171	0.168	*0.094**	0.138	0.143
12	0.243	0.19	0.208	0.209	0.207	*0.088**	0.162	0.165
15	0.335	0.272	0.27	0.283	0.281	*0.094**	0.202	0.214
18	0.484	0.415	0.37	0.409	0.393	*0.086**	0.254	0.273
21	0.691	0.648	0.522	0.58	0.584	*0.092**	0.341	0.406
24	0.884	0.864	0.768	0.823	0.798	*0.091**	0.495	0.612
27	1.012	1.025	0.939	0.979	1.001	*0.090**	0.787	0.861
30	1.104	1.139	1.053	1.098	1.152	*0.086**	0.964	1.031
33	1.166	1.197	1.159	1.188	1.25	*0.094**	1.125	1.169
36	1.191	1.22	1.193	1.225	1.292	*0.088**	1.235	1.256
39	1.206	1.228	1.211	1.238	1.317	*0.089**	1.297	1.315
42	1.201	1.229	1.207	1.246	1.329	*0.092**	1.33	1.33
45	1.201	1.226	1.217	1.234	1.327	*0.083**	1.326	1.332
48	1.211	1.226	1.223	1.237	1.34	*0.088**	1.333	1.347

Time (hours)	*Msm uvrB*$^-$ /pMV261h (U)				*Msm uvrB*$^-$ /pMV261h (2.5 mM H_2O_2)			
0	0.106	0.109	0.111	0.111	0.112	0.11	0.11	0.115
3	0.162	0.156	0.169	0.175	0.165	0.169	0.169	0.171
6	0.14	0.143	0.149	0.154	0.133	0.137	0.129	0.143
9	0.184	0.18	0.185	0.185	0.158	0.163	0.146	0.168

12	0.245	0.234	0.256	0.245	0.19	0.195	0.169	0.195
15	0.38	0.351	0.404	0.382	0.258	0.262	0.222	0.279
18	0.581	0.528	0.612	0.573	0.359	0.361	0.292	0.363
21	0.751	0.712	0.769	0.735	0.546	0.54	0.427	0.555
24	0.894	0.833	0.909	0.844	0.732	0.741	0.625	0.703
27	1	0.969	1.022	0.963	0.849	0.9	0.766	0.834
30	1.089	1.067	1.09	1.067	0.998	1.047	0.91	0.98
33	1.13	1.112	1.137	1.122	1.125	1.149	1.054	1.09
36	1.152	1.142	1.159	1.143	1.167	1.181	1.115	1.125
39	1.149	1.141	1.148	1.127	1.199	1.219	1.149	1.141
42	1.14	1.135	1.136	1.127	1.206	1.224	1.159	1.159
45	1.123	1.121	1.112	1.089	1.196	1.217	1.154	1.149
48	1.106	1.123	1.107	1.068	1.19	1.2	1.16	1.157

Time (hours)	***Msm uvrB*⁻/pMV261h_*udgB* (U)**				***Msm uvrB*⁻/pMV261h_*udgB* (2.5 mM H_2O_2**			
0	0.109	0.107	0.109	0.105	0.107	0.109	0.109	0.106
3	0.158	0.168	0.179	0.167	0.158	0.156	0.16	0.159
6	0.143	0.144	0.153	0.136	0.11	0.122	0.135	0.11
9	0.151	0.154	0.168	0.147	0.116	0.13	0.149	*0.830**
12	0.177	0.178	0.198	0.175	0.117	0.134	0.176	0.121
15	0.234	0.236	0.267	0.229	0.141	0.166	0.205	0.133
18	0.331	0.327	0.36	0.316	0.141	0.167	0.239	0.132
21	0.472	0.464	0.481	0.463	0.169	0.208	0.306	0.147
24	0.633	0.644	0.647	0.653	0.176	0.273	0.457	0.167
27	0.749	0.754	0.716	0.723	0.218	0.386	0.619	0.161
30	0.847	0.859	0.834	0.816	0.284	0.544	0.689	0.188
33	0.961	0.988	0.945	0.901	0.451	0.681	0.803	0.263
36	1.055	1.062	1.033	1.005	0.602	0.768	0.898	0.341
39	1.058	1.067	1.05	1.011	0.67	0.851	0.967	0.487
42	1.064	1.078	1.055	1.026	0.756	0.955	0.994	0.588
45	1.068	1.058	1.045	0.997	0.794	0.983	1.011	0.651
48	1.064	1.074	1.039	1.003	0.861	0.982	1.003	0.735

Time (hours)	***Msm uvrB⁻ udgB⁻***				***Msm uvrB⁻ udgB⁻* (2.5 mM H_2O_2)**			
0	0.114	0.114	0.118	0.118	0.113	0.115	0.115	0.114
3	0.19	0.192	0.193	0.196	0.17	0.187	0.181	0.181
6	0.182	0.182	0.181	0.182	0.163	0.167	0.162	0.151
9	0.262	0.249	0.249	0.264	0.218	0.22	0.216	0.212
12	0.435	0.368	0.365	0.429	0.301	0.297	0.3	0.303
15	0.737	0.636	0.597	0.731	0.505	0.484	0.476	0.512
18	0.927	0.874	0.848	0.936	0.789	0.741	0.737	0.801
21	1.135	1.06	1.063	1.152	0.986	0.949	0.894	1.006
24	1.274	1.232	1.216	1.282	1.151	1.143	0.929	1.199
27	1.336	1.302	1.291	1.336	1.268	1.277	0.87	1.325
30	1.355	1.316	1.307	1.353	1.31	1.324	0.862	1.368
33	1.347	1.32	1.308	1.346	1.34	1.36	0.865	1.395
36	1.345	1.324	1.328	1.345	1.345	1.369	0.861	1.404
39	1.296	1.281	1.269	1.304	1.338	1.365	0.867	1.391
42	1.29	1.271	1.26	1.286	1.353	1.359	0.866	1.388
45	1.023	1.191	1.212	1.158	1.32	1.33	0.846	1.297

48	0.877	1.098	1.192	1.018	1.315	1.314	0.847	1.06

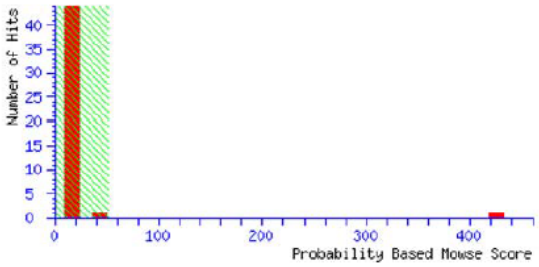
Time (hours)	**Media Control**			
0	0.064	0.064	0.064	0.064
3	0.063	0.063	0.063	0.063
6	0.064	0.064	0.064	0.064
9	0.066	0.066	0.066	0.066
12	0.066	0.066	0.066	0.066
15	0.068	0.068	0.068	0.068
18	0.068	0.068	0.068	0.068
21	0.068	0.068	0.068	0.068
24	0.068	0.068	0.068	0.068
27	0.068	0.068	0.068	0.068
30	0.068	0.068	0.068	0.068
33	0.068	0.068	0.068	0.068
36	0.068	0.068	0.068	0.068
39	0.067	0.067	0.067	0.067
42	0.068	0.068	0.068	0.068
45	0.068	0.068	0.068	0.068
48	0.068	0.068	0.068	0.068

S2.2. Raw MS analysis of interacting proteins (trypsin digested and subjected to MALDI) identified in pulldown assays

Protein band 1, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024413.dat)

Re-Search All | Search Unmatched

1. MSMEG_1368 **Mass:** 147106 **Score:** 425 **Expect:** 2.1e-039 **Peptides matched:** 56
DNA-directed RNA polymerase, beta' subunit (rpoC) [2.7.7.6] {Mycobacterium smegmatis MC2 155}

Sequence Coverage: 53%

Matched peptides shown in Bold Red

```
   1 MLDVNFFDEL RIGLATADDI RNWSYGEVKK PETINYRTLK PEKDGLFCEK
  51 IFGPTRDWEC YCGKYKRVRF KGIICERCGV EVTRAKVRRE RMGHIELAAP
 101 VTHIWYFKGV PSRLGYLLDL APKDLEKIIY FAAYVITSVD DEMRHNELST
 151 LEAEMAVEKK AVEDQRDADL EARAQKLEAD LAELEAEGAK SDVRRKVRDS
 201 GEREMRQLRD RAQRELDRLD EIWNTFTKLA PKQLIVDEVL YRELQDRYGE
 251 YFTGAMGAES IKKLIENFDI DAEAESLREV IRSGKGQKKL RALKRLKVVA
 301 AFQQSGNSPM GMVLDAVPVI PPELRPMVQL DGGRFATSDL NDLYRRVINR
 351 NNRLKRLIDL GAPEIIVNNE KRMLQESVDA LFDNGRRGRP VTGPGNRPLK
 401 SLSDLLKGKQ GRFRQNLLGK RVDYSGRSVI VVGPQLKLHQ CGLPKLMALE
 451 LFKPFVMKRL VDLNHAQNIK SAKRMVERQR PQVWDVLEEV IAEHPVLLNR
 501 APTLHRLGIQ AFEPQLVEGK AIQLHPLVCE AFNADFDGDQ MAVHLPLSAE
 551 AQAEARILML SSNNILSPAS GKPLAMPRLD MVTGLYYLTT LVEGATGEYQ
 601 AATKDAPEQG VYSSPAEAIM AMDRGALSVR AKIKVRLTEL RPPTDLEAQL
 651 FENGWKPGDA WTAETTLGRV MFNELLPKSY PFVNEQMHKK VQARIINDLA
 701 ERFPMIVVAQ TVDKLKDAGF YWATRSGVTV SMADVLVPPQ KQEILERHEA
 751 EADAIERKYQ RGALNHTERN ESLVKIWQDA TEEVGKALEE FYPADNPIIT
 801 IVKSGATGNL TQTRTLAGMK GLVTNPKGEF IPRPIKSSFR EGLTVLEYFI
 851 NTHGARKGLA DTALRTADSG YLTRRLVDVS QDVIVREHDC ETERGINVTL
 901 AERGPDGTLI RDAHVETSAF ARTLATDAVD ANGNVIIERG HDLGDPAIDA
 951 LLAAGITTVK VRSVLTCTSA TGVCAMCYGR SMATGKLVDI GEAVGIVAAQ
1001 SIGEPGTQLT MRTFHQGGVT GGADIVGGLP RVQELFEARV PRNKAPIADV
1051 AGRVRLEESD KFFKITIVPD DGGEEVVYDK LSKRQRLRVI THEDGTEGVL
1101 SDGDHVEVGD QLMEGAADPH EVLRVQGPRE VQIHLVKEVQ EVYRAQGVSI
1151 HDKHIEVIVR QMLRRVTIID SGSTEFLPGS LTERAEFEAE NRRVVAEGGE
1201 PAAGRPVLMG ITKASLATDS WLSAASFQET TRVLTDAAIN CRSDKLNGLK
1251 ENVIIGKLIP AGTGISRYRN IQVQPTEEAR AAAYTIPSYE DQYYSPDFGQ
1301 ATGAAVPLDD YGYSDYR
```

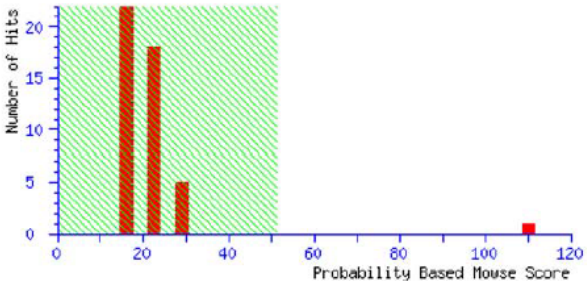
Search Parameters

```
Type of search         : Peptide Mass Fingerprint
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.5 Da
Peptide Charge State   : 1+
Max Missed Cleavages   : 0
```

Protein band 2, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024415.dat)

Re-Search All | Search Unmatched

1. MSMEG_6091 **Mass:** 93496 **Score:** 110 **Expect:** 6.7e-008 **Peptides matched:** 26
negative regulator of genetic competence ClpC-mecB {Mycobacterium smegmatis MC2 155}

Sequence Coverage: 36%

Matched peptides shown in **Bold Red**

```
     1 MFERFTDRAR RVVVLAQEEA RMLNHNYIGT EHILLGLIHE GEGVAAKSLE
    51 SLGISLEGVR SQVEEIIGQG QQAPSGHIPF TPRAKKVLEL SLREALQLGH
   101 NYIGTEHILL GLIREGEGVA AQVLVKLGAE LTRVRQQVIQ LLSGYQGKEA
   151 AEAGTGGRGG ESGNPSTSLV LDQFGRNLTA AAMEGKLDPV IGREKEIERV
   201 MQVLSRRTKN NPVLIGEPGV GKTAVVEGLA QAIVHGEVPE TLKDKQLYTL
   251 DLGSLVAGSR YRGDFEERLK KVLKEINTRG DIILFIDELH TLVGAGAAEG
   301 AIDAASILKP KLARGELQTI GATTLDEYRK YIEKDAALER RFQPVQVGEP
   351 TVEHTIEILK GLRDRYEAHH RVSITDSAMV AAATLADRYI NDRFLPDKAI
   401 DLIDEAGARM RIRRMTAPPD LREFDEKIAD ARREKESAID AQDFEKAAAL
   451 RDKEKQLVAQ RAEREKQWRS GDLDVVAEVD DEQIAEVLGN WTGIPVFKLT
   501 EEETTRLLRM EEELHKRIIG QEDAVKAVSK AIRRTRAGLK DPKRPSGSFI
   551 FAGPSGVGKT ELSKALANFL FGDDDALIQI DMGEFHDRFT ASRLFGAPPG
   601 YVGYEEGGQL TEKVRRKPFS VVLFDEIEKA HQEIYNSLLQ VLEDGRLTDG
   651 QGRTVDFKNT VLIFTSNLGT SDISKAVGLG FSQGGSENNY ERMKQKVHDE
   701 LKKHFRPEFL NRIDDIIVFH QLTQDEIIQM VDLMIGRVSN QLKTKDMALE
   751 LSDKAKALLA KRGFDPVLGA RPLRRTIQRE IEDQLSEKIL FEEIGPGQLV
   801 TVDVEGWDGE GQGEDAKFTF SGGPKRAETA EPDLAGAGAA GAPTAGTE
```

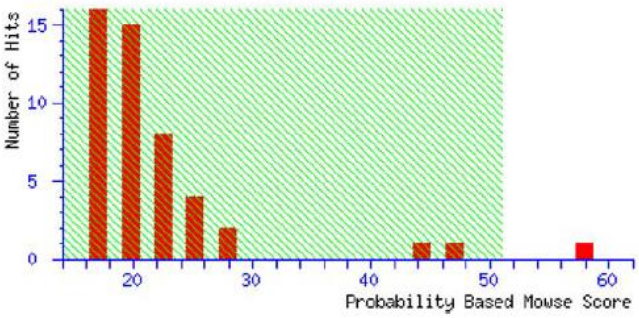
Search Parameters

```
Type of search         : Peptide Mass Fingerprint
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.5 Da
Peptide Charge State   : 1+
Max Missed Cleavages   : 1
```

Protein band 3, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024435.dat)

Re-Search All | Search Unmatched

1. MSMEG_0005 **Mass:** 74580 **Score:** 58 **Expect:** 0.011 **Peptides matched:** 11
DNA gyrase, B subunit (gyrB) [5.99.1.3] {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **23%**

Matched peptides shown in **Bold Red**

```
  1 MAAQKNNAPK EYGADSITIL EGLEAVRKRP GMYIGSTGER GLHHLIWEVV
 51 DNAVDEAMAG FATRVDVKIH ADGSVEVRDD GRGIPVEMHA TGMPTIDVVM
101 TQLHAGGKFD GETYAVSGGL HGVGVSVVNA LSTRLEATVL RDGYEWFQYY
151 DRSVPGKLKQ GGETKETGTT IRFWADPEIF ETTDYNFETV ARRLQEMAFL
201 NKGLTIELTD ERVTAEEVVD DVVKDTAEAP KTADEKAAEA TGPSKVKHRV
251 FHYPGGLVDY VKHINRTKTP IQQSIIDFDG KGPGHEVEIA MQWNAGYSES
301 VHTFANTINT HEGGTHEEGF RAALTSVVNR YAKDKKLLKD KDPNLTGDDI
351 REGLAAVISV KVAEPQFEGQ TKTKLGNTEV KSFVQKICNE QLQHWFEANP
401 AEAKTVVNKA VSSAQARIAA RKARELVRRK SATDIGGLPG KLADCRSTDP
451 SKSELYVVEG DSAGGSAKSG RDSMFQAILP LRGKIINVEK ARIDRVLKNT
501 EVQSIITALG TGIHDEFDIS KLRYHKIVLM ADADVDGQHI STLLLTLLFR
551 FMKPLVENGH IFLAQPPLYK LKWQRSEPEF AYSDRERDGL LEAGRAAGKK
601 INVDDGIQRY KGLGEMDAKE LWETTMDPSV RVLRQVTLDD AAAADELFSI
651 LMGEDVEARR SFITRNAKDV RFLDV
```

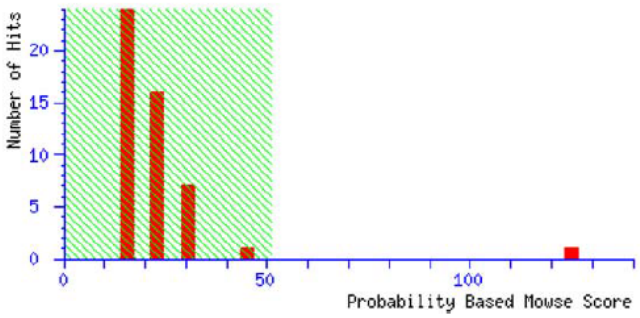
Search Parameters

```
Type of search          : Peptide Mass Fingerprint
Enzyme                  : Trypsin
Fixed modifications     : Carbamidomethyl (C)
Variable modifications  : Oxidation (M)
Mass values             : Monoisotopic
Protein Mass            : Unrestricted
Peptide Mass Tolerance  : ± 0.4 Da
Peptide Charge State    : 1+
Max Missed Cleavages    : 0
```

Protein band 4, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (./data/20210609/F024419.dat)

Re-Search All | Search Unmatched

1. MSMEG_0880 **Mass:** 60472 **Score:** 125 **Expect:** 2.1e-009 **Peptides matched:** 15
chaperonin GroL (groL) {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **42%**

Matched peptides shown in **Bold Red**

```
     1 MRPGPHLREH TPSRPSRALH LTTERAIPNP EEHFAMAKTI AYDEEARRGL
    51 ERGLNSLADA VKVTLGPKGR NVVLEKKWGA PTITNDGVSI AKEIELEDPY
   101 EKIGAELVKE VAKKTDDVAG DGTTTATVLA QALVREGLRN VAAGANPLGL
   151 KRGIEKAVEK VTETLLKSAK EVETKEQIAA TAGISAGDQS IGDLIAEAMD
   201 KVGNEGVITV EESNTFGLQL ELTEGMRFDK GYISGYFVTD AERQEAVLED
   251 PYILLVSSKV STVKDLLPLL EKVIQSGKPL LIIAEDVEGE ALSTLVVNKI
   301 RGTFKSVAVK APGFGDRRKA MLQDMAILTG GQVISEEVGL SLETADVSLL
   351 GKARKVVVTK DETTIVEGAG DAEAIQGRVA QIRAEIENSD SDYDREKLQE
   401 RLAKLAGGVA VIKAGAATEV ELKERKHRIE DAVRNAKAAV EEGIVAGGGV
   451 ALLQSAPSLE ELSLTGDEAT GANIVRVALS APLKQIALNG GLEPGVVAEK
   501 VSNLPAGHGL NAATGEYEDL LAAGVADPVK VTRSALQNAA SIAALFLTTE
   551 AVVADKPEKA AAPAGDPTGG MGGMDF
```

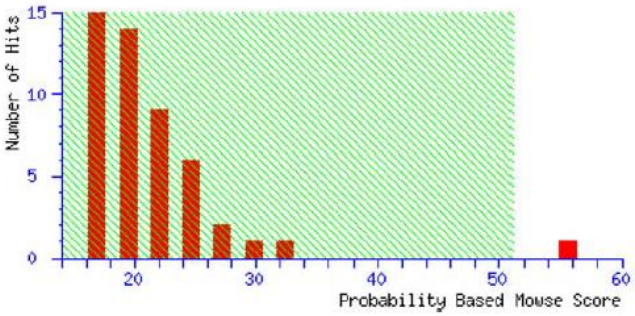
Search Parameters

```
Type of search         : Peptide Mass Fingerprint
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.29 Da
Peptide Charge State   : 1+
Max Missed Cleavages   : 0
```

Protein band 5, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024420.dat)

Re-Search All | Search Unmatched

1. MSMEG_2430 **Mass:** 54788 **Score:** 56 **Expect:** 0.018 **Peptides matched:** 9
signal recognition particle protein (ffh) {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **20%**

Matched peptides shown in **Bold Red**

```
     1 MFESLSDRLT GALQGLRGKG RLTDADIDAT TREIRLALLE ADVSLPVVRA
    51 FVARIKERAK GAEVSAALNP AQQVVKIVNE ELIGILGGET RQLAFAKTPP
   101 TVIMLAGLQG AGKTTLAGKL AKWLKDKGHS PLLVACDLQR PGAVNQLQIV
   151 GERAGVAVFA PHPGTAPGAH ETEGAGPGDP VAVAAAGLAE AKTKLYGVVI
   201 VDTAGRLGID DVLMAQAAGI RDAVQPDETL FVLDAMIGQD AVATAEAFRE
   251 GVGFTGVVLT KLDGDARGGA ALSVREVTGV PILFASAGEK LEDFDVFHPD
   301 RMASRILGMG DVLTLIEQAE QVFDAQKAEE AAAKIGTGEL TLEDFLEQML
   351 TIRKMGPIGN LLGMLPGAGQ MKDALAAVDD KQLDRVQAII RGMTPQERAD
   401 PKIINASRRL RIANGSGVTV SEVNQLVDRF FEARKMMSSM AGQMGMGFGR
   451 KSATRKAAKS KGKKGKNAKK GKGPTQPKVR NPLGAAGLPG GFPDLSNMPK
   501 GLDELPPGLA DFDLSKLKFP NK
```

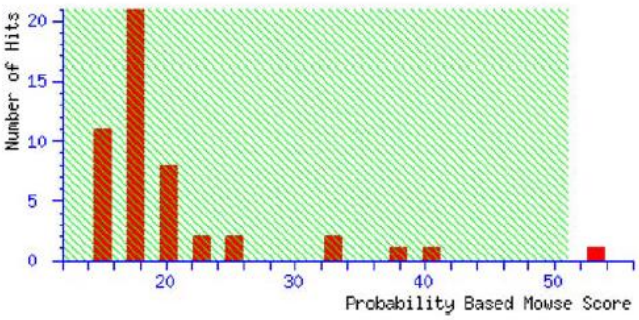
Search Parameters

```
Type of search          : Peptide Mass Fingerprint
Enzyme                  : Trypsin
Fixed modifications     : Carbamidomethyl (C)
Variable modifications  : Oxidation (M)
Mass values             : Monoisotopic
Protein Mass            : Unrestricted
Peptide Mass Tolerance  : ± 0.27 Da
Peptide Charge State    : 1+
Max Missed Cleavages    : 0
```

Protein band 6, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024421.dat)

Re-Search All | Search Unmatched

1. MSMEG_2389 **Mass:** 21217 **Score:** 53 **Expect:** 0.032 **Peptides matched:** 8
DNA-binding protein HU (hup) {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **42%**

Matched peptides shown in **Bold Red**

```
     1 MNKAELIDVL TTKMGTDRRQ ATAAVENVVD TIVRAVHKGD SVTITGFGVF
    51 EQRRRAARVA RNPRTGETVK VKPTSVPAFR PGAQFKAVIS GAQKLPADGP
   101 AVKRGVTAGP AKKAAKKAPA KKAAAKKTAT KAAAKKAPAK KAATKAPAKK
   151 AATKAPAKKA ATKAPAKKAA TKAPAKKAAA KAPAKKAATK APAKKAAAKK
   201 APAKKGRR
```

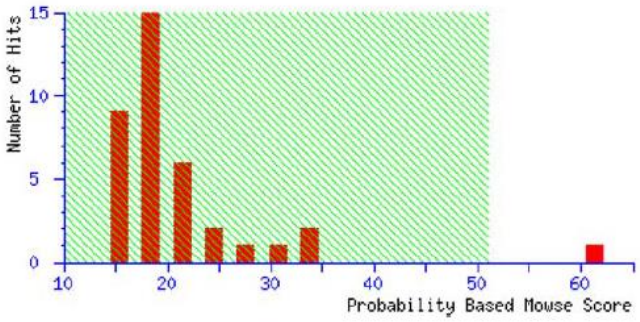
Search Parameters

```
Type of search         : Peptide Mass Fingerprint
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 200 ppm
Peptide Charge State   : 1+
Max Missed Cleavages   : 1
```

Protein band 7, Table 3

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).



Concise Protein Summary Report

Switch to full Protein Summary Report

To create a bookmark for this report, right click this link: Concise Summary Report (../data/20210609/F024412.dat)

Re-Search All | Search Unmatched

1. MSMEG_6092 **Mass:** 12446 **Score:** 61 **Expect:** 0.005 **Peptides matched:** 6
Lsr2 protein {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **33%**

Matched peptides shown in **Bold Red**

```
     1 MAKKVTVTLV DDFDGEATAD ETVEFGLDGV TYEIDLSAKN AAKLRNDLKQ
    51 WVEAGRRVGG RKRGRAATTT TRGRGAIDRE QSAAIREWAR RNGHNVSTRG
   101 RIPADVIDAF HAAT
```

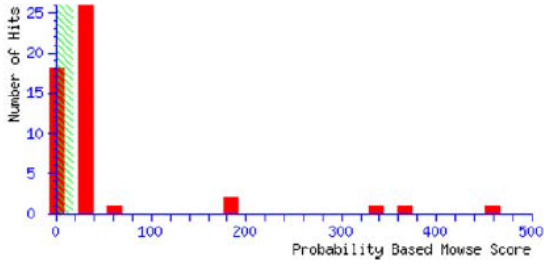
Search Parameters

```
Type of search         : Peptide Mass Fingerprint
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.5 Da
Peptide Charge State   : 1+
Max Missed Cleavages   : 1
```

Protein band 1, Table 4

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 17 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Peptide Summary Report

Switch to Protein Summary Report

To create a bookmark for this report, right click this link: Peptide Summary Report (../data/20210608/F024380.dat)

Select All | Select None | Search Selected | Error tolerant | Archive Report

1. MSMEG_3808 **Mass:** 109775 **Score:** 459 **Peptides matched:** 49
excinuclease ABC, A subunit (uvrA) {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **44%**

Matched peptides shown in **Bold Red**

```
     1 MGMNCVMRGT EHVRTACLDP RRPGRKPLAD RLVVKGAREH NLRGVDLDLP
    51 RDALIVFTGL SGSGKSSLAF DTIFAEGQRR YVESLSAYAR QFLGQMDKPD
   101 VDFIEGLSPA VSIDQKSTNR NPRSTVGTIT EVYDYLRLLY ARAGTPHCPV
   151 CGERIARQTP QQIVDQVLAM DEGLRFQVLA PVVRTRKGEF VDLFEKLNSQ
   201 GYSRVRVDGV VYPLTDPPKL KKQEKHDIEV VVDRLTVKAS AKQRLTDSIE
   251 TALNLADGIV VLEFVDREDD HPHREQRFSE KLACPNGHPL AVDDLEPRSF
   301 SFNSPYGACP ECTGLGIRKE VDPDLVVPDP ELTLAEGAVA PWSVGQSAEY
   351 FTRMLAGLGE EMGFDVNTPW KKLPAKARRA ILEGCDHQVH VRYKNRYGRT
   401 RSYYADFEGV MAFLQRRMEQ TDSEQMKERY EGFMRDIPCP ECNGTRLKPE
   451 ILAVTLSAGD FGAKSIAQVA ELSIADCADF LNSLTLGPRE QAIAGQVLKE
   501 IQSRLGFLLD VGLDYLSLSR AAATLSGGEA QRIRLATQIG SGLVGVLYVL
   551 DEPSIGLHQR DNRRLIDTLV RLRDLGNTLI VVEHDLDTIA HADWVVDIGP
   601 AAGEHGGQIV HSGTYDDLLR NPESLTGAYL SGKESIEVPA IRRPVDKKRQ
   651 ITVVGARENN LKEIDVAFPL GVLTSVTGVS GSGKSTMVND ILATVLANKL
   701 NGARLVPGRH TRVNGLDQLD KLVRVDQSPI GRTPRSNPAT YTGVFDKIRS
   751 LFAATTEAKV RGYQPGRFSF NVKGGRCEAC SGDGTIKIEM NFLPDVYVPC
   801 EVCHGARYNR ETLEVHYKGK TISEVLDMSI EEATEFFEPI SSIHRYLKTL
   851 VDVGLGYVRL GQPAPTLSGG EAQRVKLAAE LQKRSTGRTI YILDEPTTGL
   901 HFEDIRKLLK VINGLVDKGN TVIVIEHNLD VIKTSDWIID MGPEGGAGGG
   951 TVVAQGTPED VAANPDSYTG KFLAELLDVP TPKRKRRKVS A
```

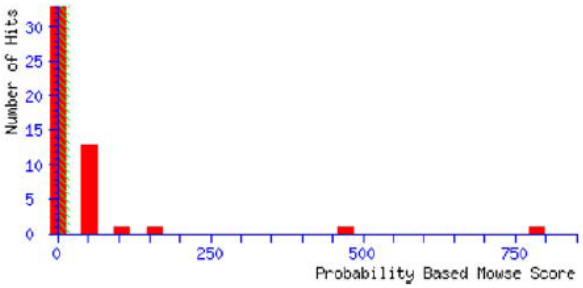
Search Parameters

```
Type of search         : MS/MS Ion Search
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.01 %
Fragment Mass Tolerance: ± 0.5 Da
Max Missed Cleavages   : 0
```

Protein band 1, Table 4

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 18 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Peptide Summary Report

Switch to Protein Summary Report

To create a bookmark for this report, right click this link: Peptide Summary Report (../data/20210608/F024376.dat)

Select All | Select None | Search Selected | Error tolerant | Archive Report

1. MSMEG_1439 **Mass:** 30396 **Score:** 786 **Peptides matched:** 64
ribosomal protein L2 (rplB) {Mycobacterium smegmatis MC2 155}

Sequence Coverage: **76%**

Matched peptides shown in **Bold Red**

```
     1 MGIRKYKPTT PGRRGASVSD FAEITRSTPE KSLVRPLHGK GGRNAHGRIT
    51 TRHKGGGHKR AYRVIDFRRH DKDGVNAKVA HIEYDPNRTA NIALLHYLDG
   101 EKRYIIAPQG LKQGDVIESG ANADIKPGNN LPLRNIPAGT VIHAVELRPG
   151 GGAKLARSAG VSIQLLGKEG TYAALRMPSG EIRRVDVRCR ATVGEVGNAE
   201 QSNINWGKAG RMRWKGKRPT VRGVVMNPVD HPHGGGEGKT SGGRHPVSPW
   251 GKPEGRTRKP NKPSDKLIVR RRRTGKKR
```

Search Parameters

```
Type of search         : MS/MS Ion Search
Enzyme                 : Trypsin
Fixed modifications    : Carbamidomethyl (C)
Variable modifications : Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.01 %
Fragment Mass Tolerance: ± 0.5 Da
Max Missed Cleavages   : 0
```

Supplementary References:

[1] U. Ray, S. Sharma, I. Kapoor, S. Kumari, V. Gopalakrishnan, S.V. Vartak, N. Kumari, U. Varshney, S.C. Raghavan, G4 DNA present at human telomeric DNA contributes toward reduced sensitivity to gamma-radiation induced oxidative damage, but not bulky adduct formation, Int J Radiat Biol, 97 (2021) 1166-1180.

[2] T. Srinath, S.K. Bharti, U. Varshney, Substrate specificities and functional characterization of a thermo-tolerant uracil DNA glycosylase (UdgB) from Mycobacterium tuberculosis, DNA Repair (Amst), 6 (2007) 1517-1528.

[3] A. China, V. Nagaraja, Purification of RNA polymerase from mycobacteria for optimized promoter-polymerase interactions, Protein Expr Purif, 69 (2010) 235-242.

[4] R.S. Thakur, A. Desingu, S. Basavaraju, S. Subramanya, D.N. Rao, G. Nagaraju, Mycobacterium tuberculosis DinG is a structure-specific helicase that unwinds G4 DNA: implications for targeting G4 DNA as a novel therapeutic approach, J Biol Chem, 289 (2014) 25112-25136.

Figure S1

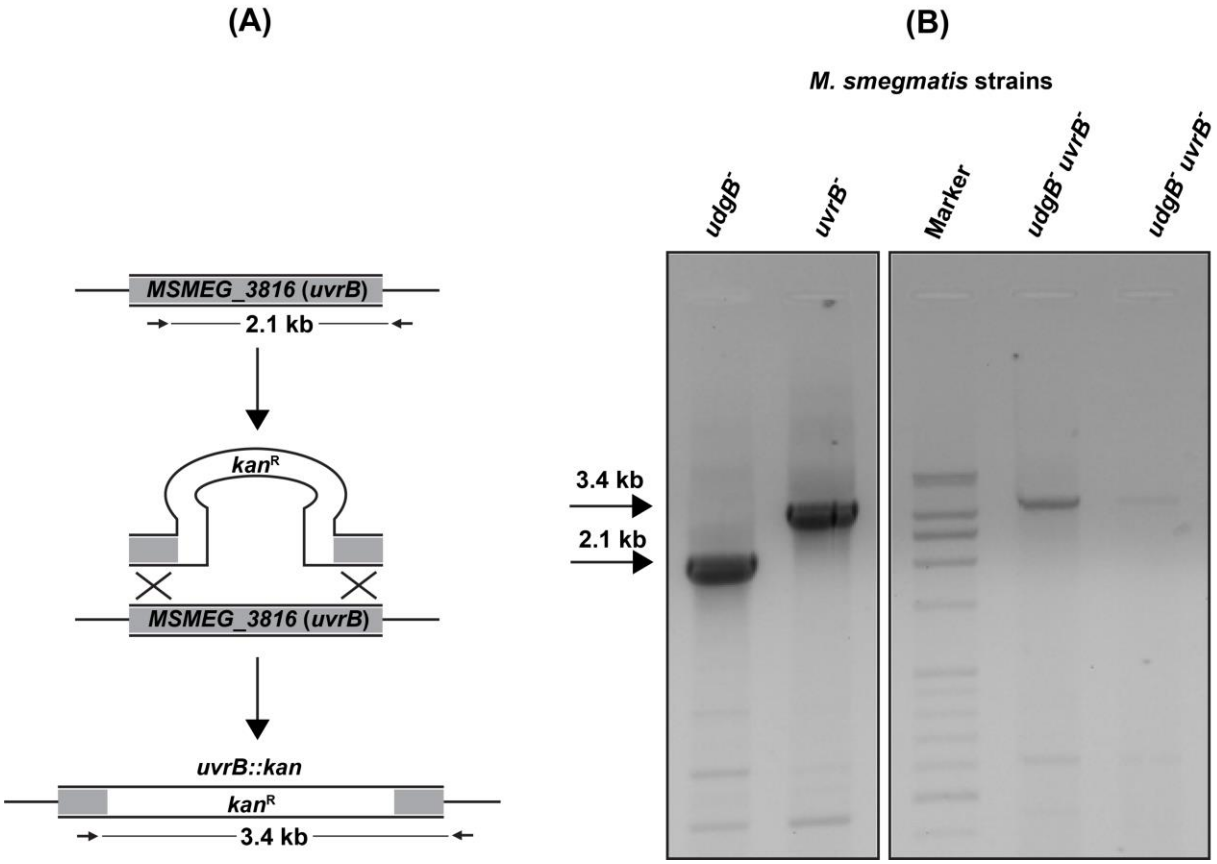


Fig. S1: Verification of *uvrB* deletion in *M. smegmatis udgB*$^{-}$ strain. (A) Schematic showing *MSMEG_3816* (*uvrB*) and Δ*uvrB::kan* loci with the binding sites of screening primers (shown with small horizontal arrows). **(B)** Representative agarose gel showing the amplification of *uvrB* locus, using the flanking primers, from the control (*udgB*$^{-}$, parent) and knockout strains, where the wild type and *uvrB::kan* alleles resulted in 801 bp and 1594 bp long amplicons, respectively.

Figure S2

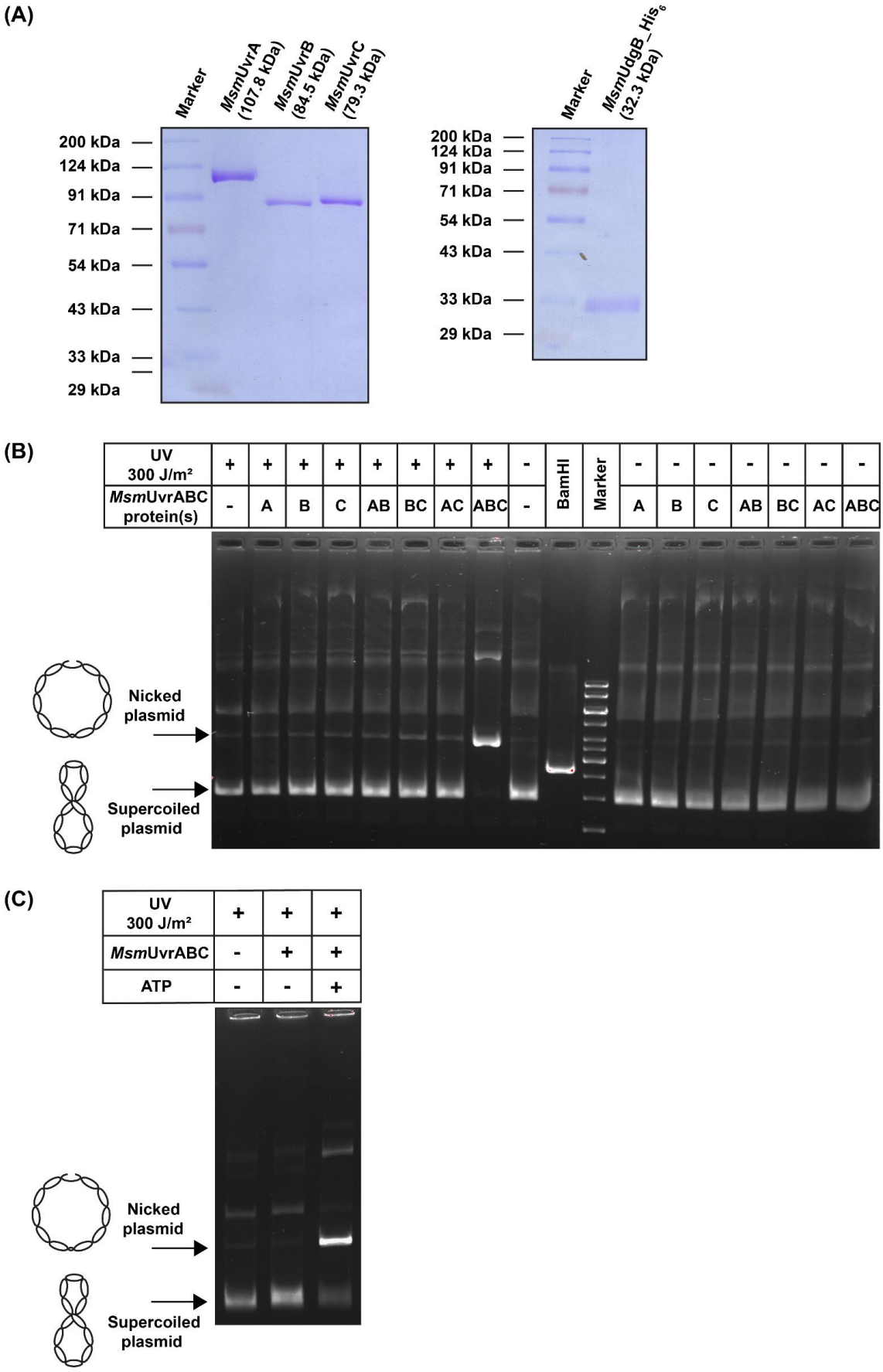


Fig. S2: Purification and activity analysis of *Msm*UvrABC proteins. (A) SDS-PAGE showing the purity and migration profile of UvrA, UvrB, UvrC and UdgB. **(B)** Analysis of the UV-irradiated supercoiled pUC18 (irradiated with 300 J/m^2 UV light) or unirradiated upon their treatments with UvrABC proteins (either one at a time or two at a time or all three proteins together). Resolution of reaction mixtures on 1% agarose gel displays the migration profile of irradiated supercoiled plasmid and its conversion into nicked circular plasmid after treatment with UvrABC complex. **(C)** Controls for panel (B) examining UvrABC activity in the absence and presence of ATP on the UV-irradiated supercoiled pUC18.

Figure S3

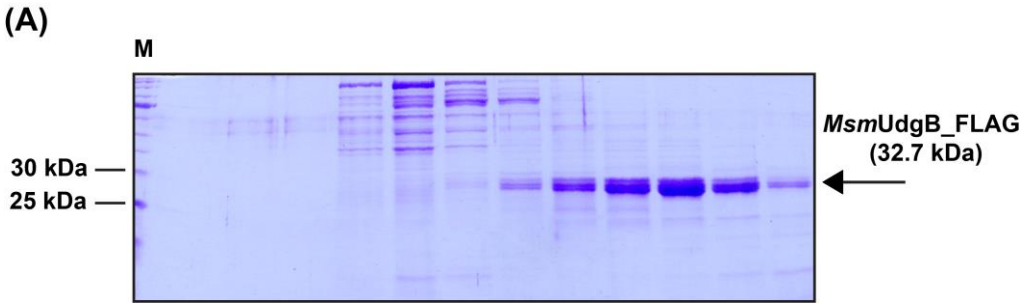


Fig. S3: Purification profile of *Msm*UdgB_FLAG. Representative SDS-PAGE showing the purification (from gel filtration column) and migration profile of UdgB_FLAG protein containing C-terminal FLAG-tag.

Figure S4

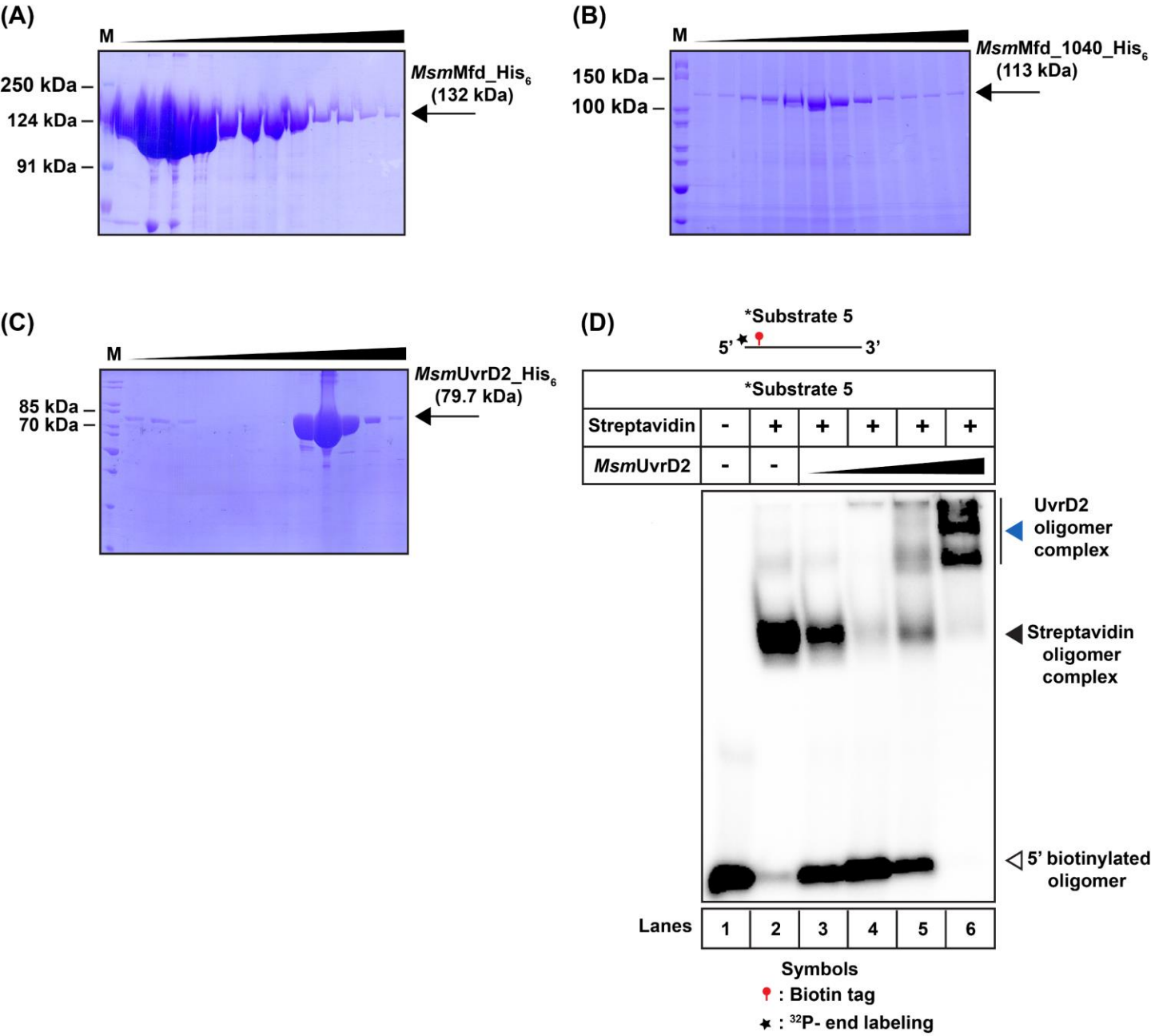


Fig. S4: Purification of *Msm*Mfd and *Msm*UvrD2 proteins, and activity analysis of UvrD2. (A), (B) and (C) SDS-PAGE demonstrating the purification profiles (from Ni-NTA column) and mobility pattern of Mfd, Mfd_1040 and UvrD2, respectively. (D) Streptavidin displacement assay. Native PAGE depicting the migration of 5' ^{32}P-labelled and biotinylated substrate 5 (lane 1) and streptavidin complex with the substrate (lane 2). Increasing amounts of the single strand DNA translocase activity of UvrD2 result in streptavidin displacement, migration of substrate at its original position and formation of UvrD2 complex with ssDNA (lanes 3-6). The asterisk denotes 5' ^{32}P-end labelling of the substrate.

Figure S5

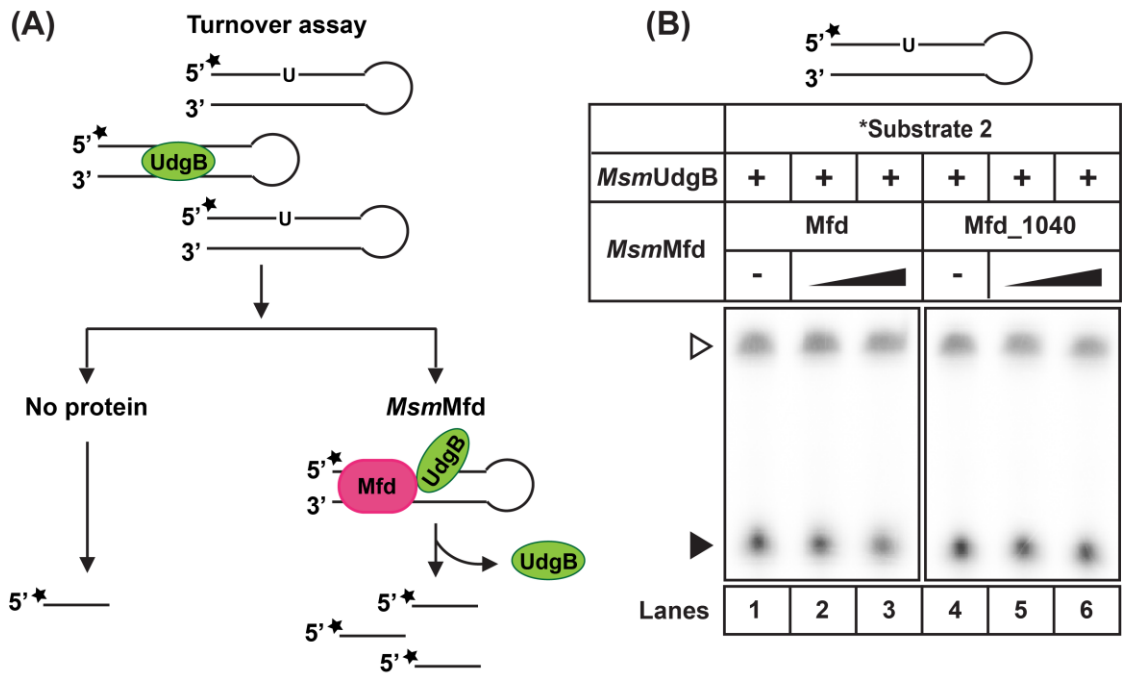


Fig. S5: Impact of *Msm* translocases on UdgB AP-DNA complex. (A) Schematic of the turnover assay showing a possible increase in the UdgB reaction product due to the possible translocase action of Mfd (Note: increased product will be seen only if Mfd dislodged UdgB from its complex with AP-DNA). (B) UdgB turnover assays in the presence of either Mfd (full length) or Mfd_1040 (176 amino acids deleted from C-terminal). ^{32}P-labelled substrate 2 was treated with the limiting amount of UdgB followed by incubation with the translocase (Mfd or Mfd_1040). Reaction mixtures from turnover assays were analysed on denaturing PAGE, where the empty black and filled black arrowheads denote ^{32}P-labelled substrate 2 and UDG reaction product, respectively. The asterisk indicates 5' ^{32}P-end labelling of the substrate.

Figure S6

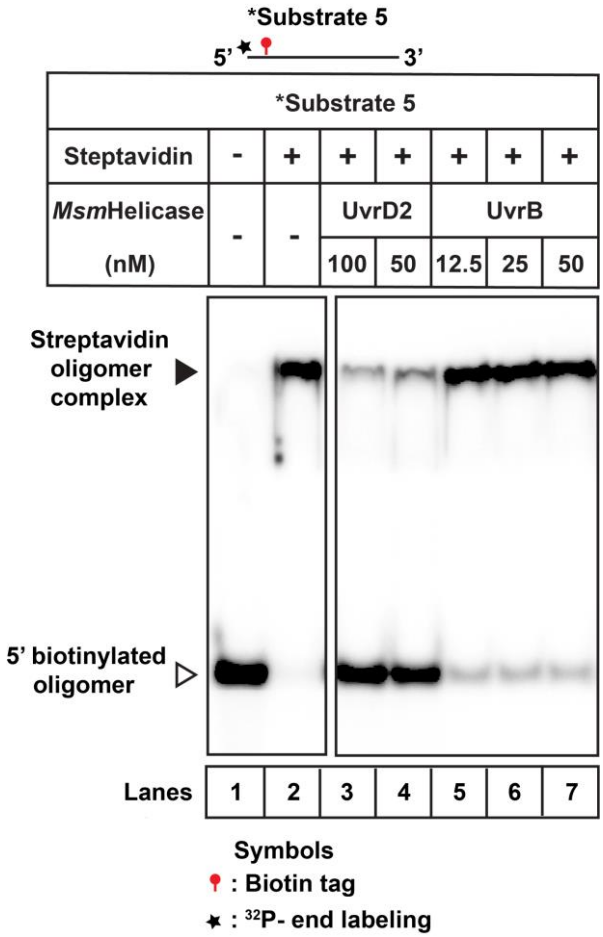


Fig. S6: Activity analysis of *Msm*UvrD2 and *Msm*UvrB. Streptavidin displacement assay. Native PAGE showing the migration of 5' ^{32}P-labelled and biotinylated substrate 5 (lane 1) and streptavidin complex with the substrate (lane 2). Single strand DNA translocase activity of UvrD2 (lanes 3 and 4) or UvrB (lanes 5-7) results in streptavidin displacement, migration of substrate at its original position. The asterisk denotes 5' ^{32}P-end labelling of the substrate.