Killen et al. BMC Molecular Biology 2012, 13:33
http://www.biomedcentral.com/1471-2199/13/33

BMC Molecular Biology

RESEARCH ARTICLE Open Access

Escherichia coli RecG functionally suppresses human Bloom syndrome phenotypes

Michael W Killen[1,2], Dawn M Stults[3], William A Wilson[4] and Andrew J Pierce[1,5*]

Abstract

Defects in the human *BLM* gene cause Bloom syndrome, notable for early development of tumors in a broad variety of tissues. On the basis of sequence similarity, *BLM* has been identified as one of the five human homologs of *RecQ* from *Escherichia coli*. Nevertheless, biochemical characterization of the BLM protein indicates far greater functional similarity to the *E. coli* RecG protein and there is no known RecG homolog in human cells. To explore the possibility that the shared biochemistries of BLM and RecG may represent an example of convergent evolution of cellular function where in humans BLM has evolved to fulfill the genomic stabilization role of RecG, we determined whether expression of RecG in human BLM-deficient cells could suppress established functional cellular Bloom syndrome phenotypes. We found that RecG can indeed largely suppress both the definitive elevated sister chromatid exchange phenotype and the more recently demonstrated gene cluster instability phenotype of BLM-deficient cells. In contrast, expression of RecG has no impact on either of these phenotypes in human cells with functional BLM protein. These results suggest that the combination of biochemical activities shared by RecG and BLM fill the same evolutionary niche in preserving genomic integrity without requiring exactly identical molecular mechanisms.

Background

Human cells possess five proteins with clear sequence homology to the *E. coli* RecQ protein: BLM, WRN, RECQL, RECQL4 and RECQL5. These proteins are all implicated in preserving genomic integrity (reviewed in [1,2]). Functionally, inherited homozygous defects in *BLM*, *WRN* or *RECQL4* cause human disease: Bloom syndrome, Werner syndrome and Rothmund-Thomson/RAPADILINO/Baller-Gerold syndromes respectively. Bloom syndrome is particularly striking for its predisposition to early-onset malignancy with a broad distribution of cancer types similar to that seen with sporadic tumors in the general population [3].

Sequence homology of BLM with RecQ notwithstanding, characterization of the *in vitro* activities of BLM demonstrates significant similarities to the biochemistry of the *E. coli* RecG DNA translocase protein. Both BLM [4,5] and RecG [6,7] can bind to and regress multi-stranded DNA structures that model stalled replication forks. Similarly, both BLM [8,9] and RecG [10,11] have the capacity to bind to and branch migrate Holliday junctions. Both BLM [12,13] and RecG [14] have also been shown to dismantle D-loops where a 3′-OH ssDNA has invaded a homologous DNA duplex, although the mechanism by which RecG carries out this reaction is less well established [15]. The manner by which RecG accomplishes these tasks is in large part made clear by its crystal structure [16]: a RecG monomer binds at a model replication fork by inserting a C-terminal protein wedge domain into the fork. The body of RecG then functions as a double stranded DNA translocase to pull on and reanneal DNA template strands through the body of the protein. At such time as the nascent DNA strands encounter the wedge domain, they are stripped off and allowed to anneal together resulting in the formation of a Holliday junction. As RecG continues to translocate on the dsDNA, the branch point of the Holliday junction is effectively migrated. In the absence of high-resolution structural information it remains unclear precisely how BLM carries out these activities.

The BLM protein also possesses activities it is not known to share with RecG. BLM can act in concert with EXO1 at double stranded DNA ends to cause a 5′-3′ single

* Correspondence: piercea@medimmune.com
[1]Department of Microbiology, Immunology and Molecular Genetics, Markey Cancer Center, University of Kentucky, Lexington, KY, USA
[5]Present address: MedImmune, LLC, Gaithersburg, MD, USA
Full list of author information is available at the end of the article

BioMed Central
© 2012 Killen et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

stranded resection that exposes a free ssDNA 3′ end suitable for loading with Rad51 [17], reminiscent of the combined activities of the *E. coli* RecQ helicase and RecJ 5′-3′ exonuclease [18,19]. Alternatively, BLM can also functionally interact with the DNA2 exonuclease to carry out a similar reaction [20]. BLM has strong unwinding activity on G-quadruplex DNA structures [21] as well as both ssDNA annealing [22,23] and/or strand exchange activities [24]. Notably, BLM has many well-characterized protein-protein interactions, including those with RMI1, C16orf75 (RMI2) and TOP3A [25-28] that collectively mediate double-Holliday junction dissolution, as well as direct interaction with the Rad51 recombinase [29] and with the multi-component Fanconi anemia protein containing BRAFT complex [30]. In contrast, RecG functions in a largely monomeric manner [31].

The mechanistic similarities between BLM and RecG have led us and others [32] to speculate that *E. coli* RecG and human BLM may be functional analogs. In order to test this hypothesis and to determine the extent to which the shared biochemical activities of BLM and RecG are responsible for suppressing the functional cellular phenotypes observed in human cells lacking BLM, we reasoned that suitable expression of RecG might suppress a BLM defect. The best characterized cellular phenotype of BLM deficiency is a 10-fold elevated frequency of sister chromatid exchanges [33], thought to represent a hyper-recombination phenotype indicative of elevated crossing-over and overall genomic instability. In addition, we have recently demonstrated that BLM deficiency causes a striking destabilization of the highly repetitive human ribosomal RNA gene clusters (the 'rDNA'), with recombination-mediated genomic restructuring of these clusters increased 100-fold over cells wild-type for BLM function [34]. Accordingly, we engineered several semi-humanized RecG protein expression systems and stably introduced these constructs into human cells either wild-type or defective for the BLM protein. We then assayed the resulting RecG transgene expressing cells for changes in these two phenotypes.

Methods

Protein expression constructs

The coding sequence for RecG was isolated from *E. coli* TOP10 (Invitrogen) genomic DNA with the addition of a consensus Kozak sequence [35] and the SV40 large T-antigen nuclear localization signal (PKKKRKV) by PCR using primers 5′-gggggggatccagccaccatggctccaaaaaaaaagcgcaaag tggcgatgaaaggtcgcctg and 5′-gggggggatatcgcggccgcttacgc attcgagta (RecG sequences underlined) followed by cloning into the pCAGGS mammalian constitutive expression vector. A cryptic polyadenylation sequence in the RecG coding sequence 5′-AATAAA (AsnLys) was removed via silent mutagenesis to 5′-AACAAG (AsnLys). Carboxy-terminal additions of the enhanced green fluorescent protein (EGFP) coding sequence were constructed either to produce a RecG-EGFP fusion protein with a GSG linker peptide (Figure 1A) "RecG fuse" (predicted molecular weight: 104.7 kDa), or to produce two proteins in a bi-cistronic manner: RecG fused to the Strep-tag II for potential affinity purification [36] followed by a GSG linker, the 2A polyprotein 'self-cleaving' sequence from *Thosea asigna* virus [37,38], a VAT peptide linker and EGFP (Figure 1B) "RecG 2a". In the bi-cistronic construct the predicted molecular weight of the RecG polypeptide is 80.8 kDa and the predicted molecular weight of the EGFP polypeptide is 27.3 kDa. All final construct sequences were verified by direct DNA sequencing. The finished constructs express essentially the entirety of the wild-type *E. coli* RecG polypeptide, with small amino- and carboxy-terminal extensions. The EGFP expression vector used for subcloning the EGFP coding sequence was pEGFP-N1 (Clontech). The RecG expression constructs with full annotation are available from the non-profit Addgene plasmid repository (http://www.addgene.org) as plasmids #31274 and #31275.

Cell lines

Cell lines used that express normal BLM include the SV40-transformed fibroblast line GM00637 (Coriell Cell Repositories) "*BLM*+" and the cervical carcinoma line HeLa S3 (ATCC: American Type Culture Collection) "HeLa". Cell lines used that are deficient in BLM protein include GM08505 (Coriell Cell Repositories) which are SV40 transformed fibroblasts derived from a Bloom syndrome patient homozygous for the Ashkenazi Jewish founder BLM mutation (6-bp del/7-bp ins) at nucleotide 2281 of the open reading frame, hereafter referred to as "*BLM*-". A second BLM deficient line used was GM16375 (Coriell Cell Repositories) which are EBV transformed lymphocytes from a French-Canadian Bloom syndrome patient homozygous for a C>A transversion resulting in a (S595X) termination mutation, hereafter referred to as "*BLM*- FC". The BLM-defective line stably suppressed by either BLM cDNA expression, or by a control empty vector are the lines PSNF5 "*BLM*-: cDNA" and PSNG13 "*BLM*-: empty vector" respectively, as described in [39] (kind gift from Ian Hickson) both derived originally from the GM08505 line (Coriell Cell Repositories). Lines were generally grown in MEM with 10% fetal bovine serum, with *L-gln* and antibiotic supplementation at 37C in a humidified 5% CO_2 incubator.

Stable cell lines were generated by electroporating either a RecG expression construct or a control EGFP expression construct into cells, followed by unselected cell population expansion, one round of flow-sorting enrichment for green fluorescent cells, further unselected expansion, and finally a second round of flow-sorting enrichment for green fluorescent cells. Clonal and subclonal derivatives of

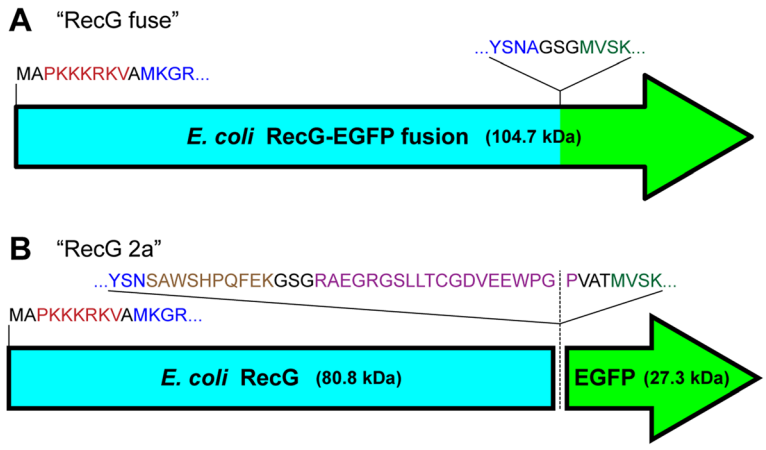

Figure 1 RecG expression constructs. 1**A**) RecG expressed as a fusion protein with EGFP. Lettering colors: black: linker amino acids, red: nuclear localization signal, blue: RecG coding sequence, green: EGFP coding sequence. 1**B**) RecG co-translated with EGFP as two separate polypeptides. Lettering colors: black: linker amino acids, red: nuclear localization signal, blue: RecG coding sequence, brown: strep-tag II, purple: 2a 'self-cleaving' peptide, green: EGFP coding sequence. Dotted line represents the division of the precursor polyprotein into two independent polypeptides. The full-length wild-type *E. coli* RecG polypeptide initiates with the indicated MKGR. . . peptide sequence and terminates with the indicated . . .YSNA peptide sequence.

these highly enriched fluorescent populations were subsequently derived by limiting dilution. All of the transgene expressing cell lines generated and used in this work are either clonal or subclonal, with the exception of the "*BLM*- FC: RecG fuse" line: *BLM*- FC cells were transduced by a high-titer lentivirus (Welgen, Inc.) containing an expression cassette for the RecG-EGFP fusion construct (Figure 1A) and separated into green fluorescent and non-fluorescent populations by flow sorting, and the *BLM*-: EGFP line which was isolated on the basis of drug resistance only. Stable cell lines were chosen for further experiments on the basis of the levels of EGFP expressed in these lines, as measured by flow cytometry.

Sister chromatid exchange assays

Sister chromatid exchange assays were performed largely according to [40] with minor modifications [34]. Cellular metaphase spreads were imaged and scored individually by counting the number of visible exchanges and the number of chromosomes in each unique high powered microscope field to calculate the number of SCEs per chromosome. In our hands, scoring the number of SCEs per chromosome is more robust than scoring the number of SCEs per metaphase and is relatively insensitive to the ploidy of the cell line under investigation [41]. The resulting SCEs/chromosome figures were binned and plotted. All statistical tests were performed using unbinned data. SCE experiments were generally performed by collecting fixed metaphase cells from one or two experiments, which were subsequently dropped on microscope slides to release the metaphase spreads, stained and counted together in a single session at the microscope.

Western blotting

Protein extracts were prepared using RIPA buffer as described previously [34]. All resolving SDS-PAGE gels used 9% acrylamide and were blotted onto Hybond-ECL nitrocellulose membrane (Amersham Biosciences, cat. #RPN68D). Primary antibodies used were: rabbit anti-GFP (Cell Signaling Technologies, cat. #2555), rabbit anti-beta-actin (Cell Signaling Technologies, cat. #4970), rabbit anti-beta-tubulin (NeoMarkers, cat. #RB-9249-PO). The secondary antibody was ImmunoPure Antibody donkey anti-rabbit IgG conjugated with horseradish peroxidase (Pierce, cat. #31458). Blots were developed using an Amersham™ ECL Plus western blotting detection system (GE Healthcare, cat. #RPN2132) and imaged with a Storm 860 PhosphorImager (Molecular Dynamics).

GCI analysis

Gene cluster instability analysis was carried out as described previously [34]. Briefly, genomic DNA was prepared in the solid phase by digesting single cell suspensions in agarose with proteinase K in the presence of sarkosyl and EDTA, rinsed thoroughly and equilibrated in 50% glycerol/10 mM Tris/1 mM EDTA pH 8.0 and stored at -20C. 10 μl agarose slices containing approximately 1 μg genomic DNA were equilibrated in suitable restriction digestion buffer and digested overnight with EcoRV (New England Biolabs). Digested DNA still in solid form was loaded into a 1% PFC agarose (Bio-Rad) gel and run in 0.5× TBE buffer (44.5 mM Tris base, 44.5 mM boric acid, 1 mM EDTA pH 8.0) using a CHEF-MAPPER system (Bio-Rad) at 14C. Pulsed-field electrophoretic conditions were a field strength of 6 V/cm with 120° separation between field vectors. Field switch times varied from 3

seconds to 90 seconds with a 'ramp factor' of 0.357. Gels were run for 24 hours, equilibrated to 0.5% glycerol in water, then dried, rehydrated, probed with a radiolabeled probe specific for the human rDNA [42] and imaged with a Storm 860 PhosphorImager (Molecular Dynamics).

EGFP quantitation by flow cytometry

For cell lines carrying EGFP transgenes, the fold increase in fluorescence relative to non-fluorescent cells was calculated by dividing the geometric mean value of green fluorescence emitted from the fluorescent cell sub-population by the geometric mean value of background green autofluorescence from the non-fluorescent sub-population of the same cell line.

Results

RecG expression suppresses the *BLM*- elevated SCE phenotype

We stably transfected *E. coli* RecG expressing transgenes into human cells with both BLM-defective and BLM-normal genetic backgrounds. Two different expression systems were used, a direct fusion protein between RecG and EGFP, and a bi-cistronic system where separate RecG and EGFP polypeptides are co-translated. RecG fusion proteins in *E. coli* largely retain functionality [43], and indeed the RecG-EGFP fusion protein plasmid used in this work increases the resistance of deletion recG *E. coli* to mitomycin C (Robert Lloyd, personal communication). We found a significant RecG dose-dependent decrease in SCEs per chromosome in *BLM*- cells (Table 1, Figure 2) with the highest level of RecG expression tested reducing the elevation of SCEs by 75%. In our hands, expression of the human BLM protein from a cDNA construct in *BLM*- cells reduces this characteristic elevated SCE phenotype by 80% (Figure 2A, Table 1 and [34]), so *E. coli* RecG expression is nearly as effective as the human protein in this system. RecG suppressed the elevated *BLM*- SCE phenotype either when expressed as an EGFP fusion protein or as a separate polypeptide co-translated with EGFP. In all cases where RecG is expressed in a *BLM*- line, mean SCEs/chromosome are reduced with high confidence (two-tailed Mann–Whitney test, P<0.001). Expression of EGFP alone without RecG in *BLM*- cells has no statistically significant effect on SCE levels (Figure 2A, Table 1). That RecG suppression of the *BLM*- elevated SCE phenotype is not dependent on the genetic background of the cells other than with respect to the loss of BLM is demonstrated by manner in which RecG suppresses sister chromatid phenotypes both in cells of Ashkenazi Jewish origin and separately in cells of French-Canadian origin (Figure 3A). Both of these lines are essentially BLM-null due to a 6-bp deletion/7-bp insertion, and due to an S595X mutation respectively yet presumably have otherwise heterogeneous genetic backgrounds. RecG is expressed at a low level in "*BLM*- FC: RecG fuse" cells (Figure 3B, Table 1), resulting in a less dramatic, 25% reduction in the elevated SCE phenotype, nevertheless in a statistically highly significant manner (P=0.0004, Mann–Whitney two-tailed test).

Table 1 SCE levels and EGFP expression

Cells	SCE/Chr	EGFP	% EGFP+	HPF	Xomes
HeLa	0.09	1.0	0	21	605
HeLa: RecG fuse	0.08	42.6	98	32	835
BLM+	0.10	1.0	0	22	863
BLM+: RecG 2a	0.10	49.3	100	30	999
BLM-: cDNA	0.28	1.0	0	26	1625
BLM-: RecG 2a high	0.33	148	93	33	1861
BLM-: RecG fuse high	0.34	29.8	87	26	1149
BLM-: RecG 2a low	0.68	12.4	92	36	2017
BLM-: EGFP	1.02	27.3	57	48	3013
BLM-: empty vector	1.05	1.0	0	25	1354
BLM-	0.98	1.0	0	16	883
BLM- FC: RecG fuse	0.70	6.4	100	55	1981
BLM- FC	0.91	1.0	0	49	2077
BLM-: RecG fuse med	0.43	30.6	87	29	1925
BLM-: lost RecG fuse	0.86	8.96	1	25	1316

Cells: The stable cell line being investigated.
SCE/Chr: Median number of sister chromatid exchanges observed per chromosome (sister chromatid pair).
EGFP: Relative levels of fluorescence for cells expressing RecG and/or EGFP, relative to cells not expressing EGFP.
% EGFP+: The fraction of cells in a population with above-background green fluorescence due to expression of EGFP.
HPF: The number of high powered microscope fields scored.
Xomes: The total number of sister chromatid pairs scored for exchanges.

RecG has no effect on SCE levels in cells expressing normal BLM protein

In order to establish the *BLM*- specificity of the RecG SCE reduction effect, we carefully examined SCE levels in two different *BLM*+ lines that stably express relatively high levels of either the RecG-EGFP fusion protein, or the bicistronic RecG/EGFP construct (Figure 4). We found that RecG expression had no statistically significant effect on either the median levels of SCEs per chromosome or in the distribution of SCEs seen in either *BLM*+ line (Table 1).

Loss of RecG expression restores the elevated SCE phenotype to *BLM*- cells

One of the clonal *BLM*- lines stably expressing medium levels of the RecG-EGFP fusion protein (Figure 5B, "*BLM*-: RecG fuse med") gradually lost EGFP expression during two months of continuous cell culture in the absence of selection, likely due to epigenetic silencing of the RecG-EGFP transgene. In the resulting "*BLM*-: lost RecG fuse" line, SCE levels were re-elevated to nearly those of either the parental *BLM*- line, or the control *BLM*- line containing a

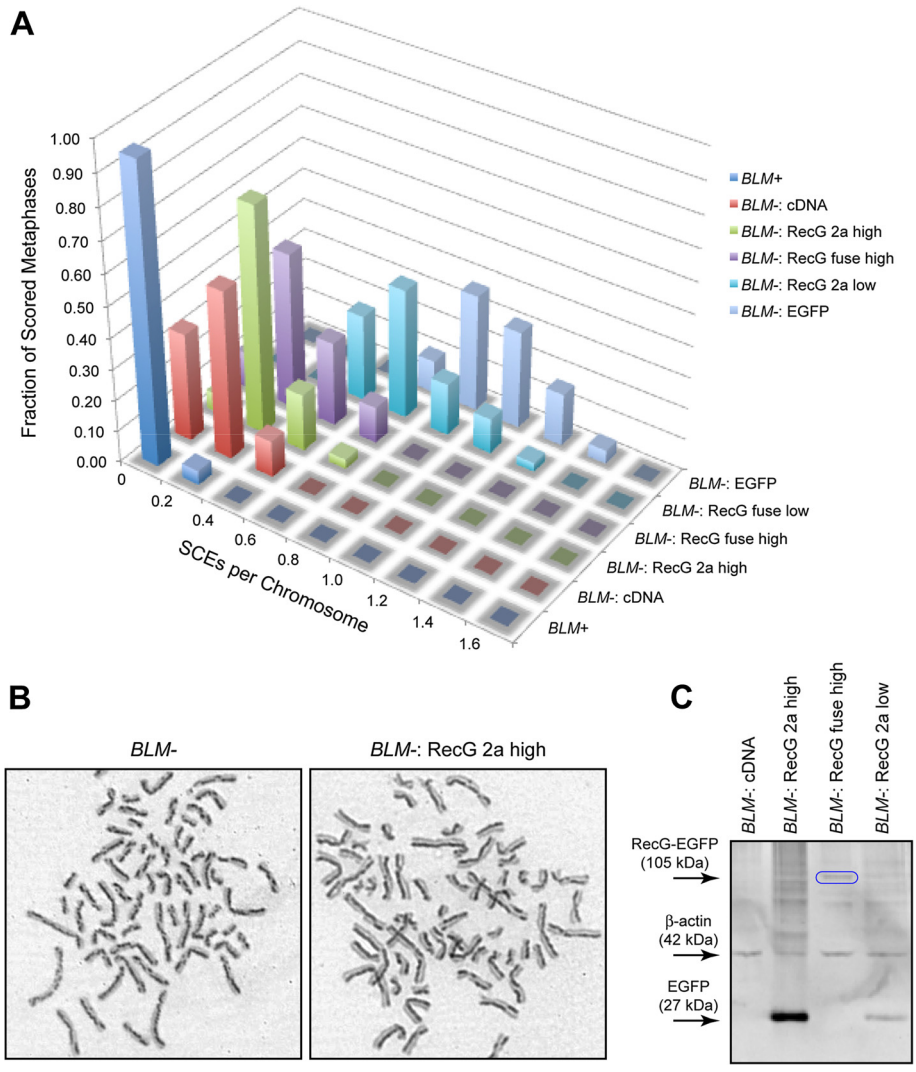

Figure 2 RecG expression suppresses the *BLM*- elevated SCE phenotype. 2**A**) Normalized sister chromatid exchange frequencies. "RecG 2a" denotes cells expressing RecG and EGFP as separate co-translated proteins. "RecG fuse" denotes cells expressing RecG and EGFP together as a fusion protein. 2**B**) Representative sister chromatid exchange metaphase spreads with sister chromatids differentially stained. The harlequin staining pattern of the *BLM*- cells is greatly reduced upon expression of RecG. 2**C**) RecG expression levels. Protein extracts from the indicated cell lines are blotted with an anti-GFP antibody and with anti-β-actin as a loading control. RecG-EGFP fusion proteins (105 kDa) are indicated with blue ovals. EGFP is detected as a proxy biomarker for RecG in the "RecG 2a" co-translated lines. See also Table 1 for EGFP quantitation by flow cytometry.

transgene for EGFP alone "*BLM*-: EGFP" (Figure 5A, Table 1), confirming that RecG expression is solely responsible for suppressing the elevated SCE phenotype in *BLM*- cells.

RecG expression reduces the elevated gene cluster instability (GCI) of *BLM*- cells

Loss of BLM very strongly increases the rate of spontaneous recombination-mediated genomic restructuring in the highly repetitive human rRNA gene clusters [34]. In *BLM*- cells, recombination alters gene cluster lengths on the order of every round of mitotic cell division, an approximate 100-fold increase over the rate in *BLM*+ cells. This rapid randomization of cluster lengths produces a diagnostic ladder-like electrophoretic karyotype with each rung on the ladder separated by the 48.5 kb unit length of the human rDNA repeat, particularly in the range of gene cluster lengths from 250 kb to 550 kb (Figure 6 – shaded in red, red dotted side bracket). The intensity of individual cluster bands is related to how early in the clonal expansion the mitotic recombination event took place, with more intense bands having arisen earlier in the expansion of a clonal population [34]. Expression of RecG suppresses this elevated GCI phenotype: the rate of spontaneous cluster restructuring is greatly reduced (Figure 6 – red carets) although as with

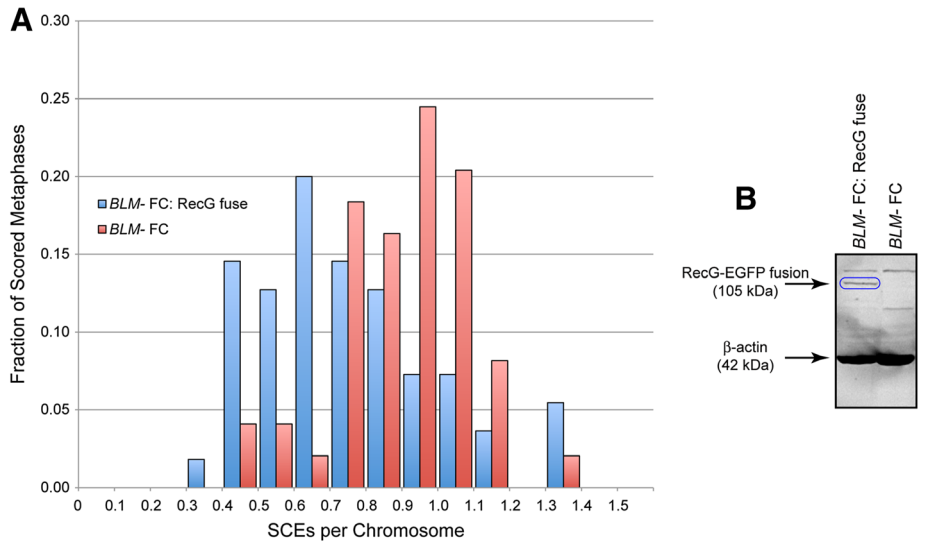

Figure 3 RecG expression reduces SCE in a French-Canadian *BLM*- line. 3**A**: Normalized sister chromatid exchange frequencies. "*BLM*- FC": EBV-immortalized lymphocytes homozygous for an inactivating (S595X) mutation. "*BLM*- FC: RecG fuse": the same cells expressing low levels of the RecG-EGFP fusion protein. 3**B**: Western blot showing relative protein levels. Blue oval: RecG-EGFP fusion protein. "*BLM*- FC: RecG fuse": protein extract from French-Canadian *BLM*- cells stably expressing low levels of the RecG-EGFP fusion protein. "*BLM*- FC": protein from French-Canadian *BLM*- cells lacking a RecG transgene. See also Table 1 for EGFP quantitation by flow cytometry.

suppression of the elevated SCE phenotype of *BLM*- cells, there is a residual low level of GCI even at the highest levels of RecG expression assayed. Notably, expression of RecG is particularly effective at restabilizing gene clusters in the previously highly unstable 250 kb to 550 kb size range, as seen by the significant reduction of gene cluster lengths detected in this range. When GCI is quantified by the average number of minor bands

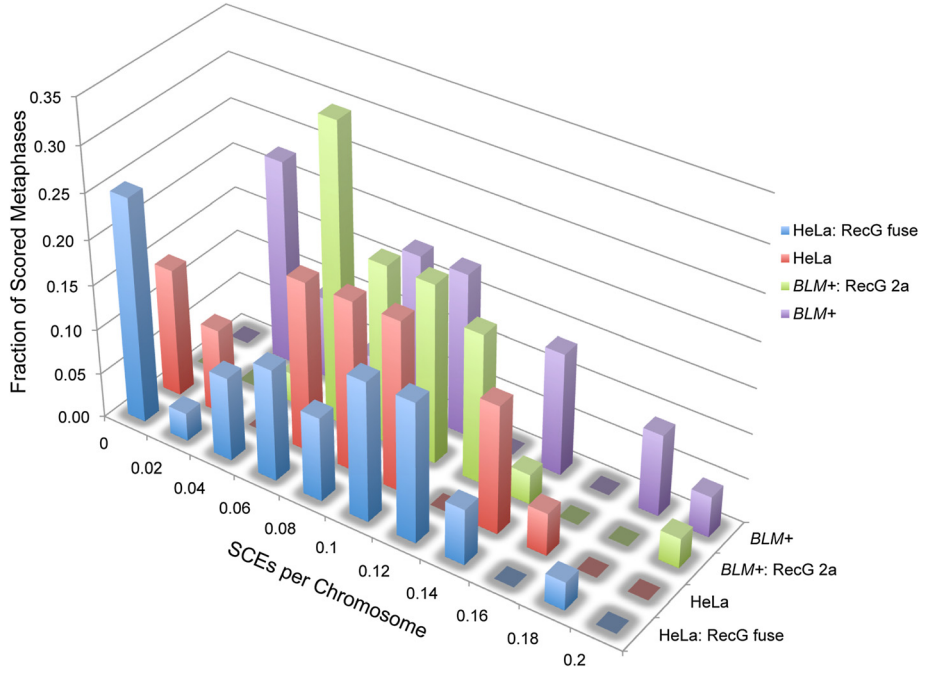

Figure 4 RecG expression has no effect on SCE levels in cells with normal BLM protein. Normalized sister chromatid exchange frequencies are shown. Note the scale of the x-axis in this figure ranges from 0 to 0.2 for the distribution of SCEs/chromosome in lines with normal BLM function, whereas the scale of the x-axis in Figure 2A ranges from 0 to 2.0 SCEs/chromosome, consistent with the known 10-fold elevation of sister chromatid exchanges in *BLM*- lines.

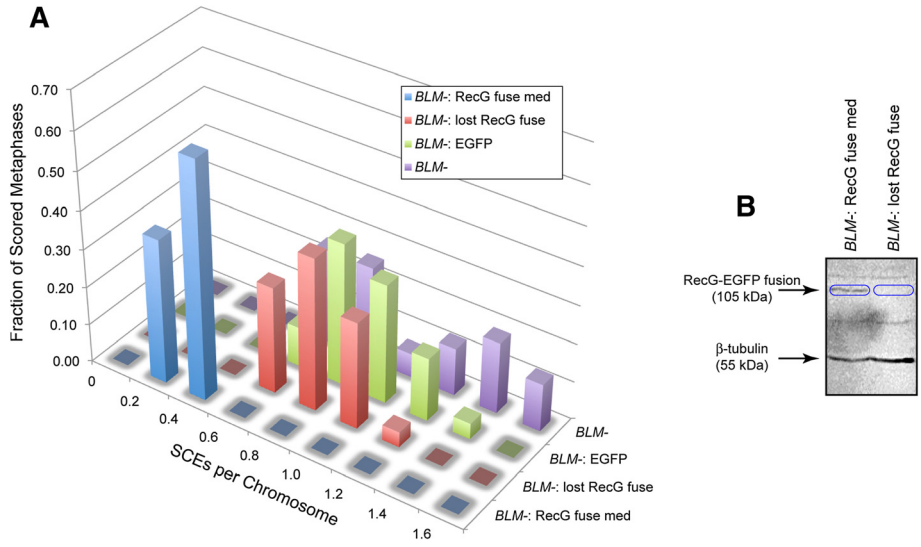

Figure 5 Loss of RecG expression in *BLM*- cells restores the elevated SCE phenotype. 5**A**) Normalized sister chromatid exchange frequencies are shown. 5**B**) Loss of expression of RecG-fuse is shown by western blotting (blue ovals). See also Table 1 for EGFP quantitation by flow cytometry.

observed per clone [34], expression of RecG suppresses the elevated BLM GCI phenotype by more than 3-fold overall, from 8.7 minor bands per clone in the *BLM*- cells to 2.6 minor bands per clone in the *BLM*- cells that express RecG.

RecG expression does not affect gene cluster instability in *BLM*+ cells

Expression of RecG as an independent polypeptide co-translated with EGFP did not have an obvious effect on gene cluster instability in a cell line wild-type for BLM

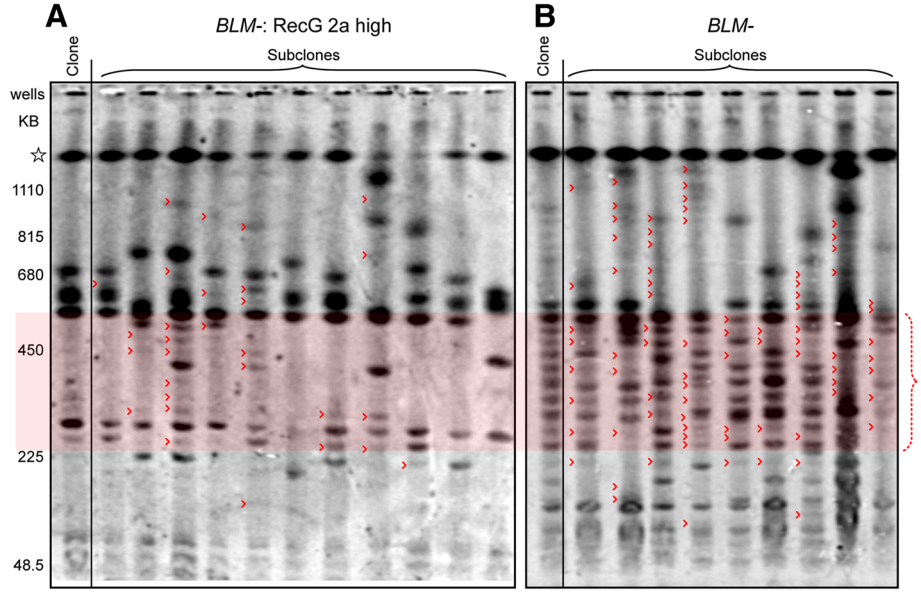

Figure 6 RecG stabilizes the gene cluster instability of *BLM*- cells. Southern blotting detection of the rDNA clusters with resolution up to 1 Mb is shown. Cluster lengths are calibrated by *S. cerevisiae* chromosome size markers. Open star is the 1 Mb resolution limit of the gel: all clusters larger than 1 Mb co-localize to this band. Clonal populations are shown on the left of each panel, with a collection of subclones derived from this clonal population to the right. Red carets: mitotic recombination indicated by minor intensity gene cluster bands. The zone of particularly high instability in the 250 kb to 550 kb size range is indicated by red shading and by the red dotted bracket at the right. 6**A**) "*BLM*-: RecG 2a high": *BLM*- cells stably expressing higher levels of cotranslated RecG/EGFP. 6**B**) The parental *BLM*- line from which the "*BLM*-: RecG 2a high" cells were derived.

protein (Figure 7) even though RecG is expressed at a relatively high level when measured by flow cytometry detection of the co-translated EGFP protein in this line (Table 1: *BLM*+: RecG 2a). The changes to the banding pattern in the subclones relative to that of the parental clonal population represent mitotic recombination events that occurred sufficiently late in the expansion of the parental clonal population as to be undetectable as visible minor-intensity banding. The relatively small number of changes to major intensity bands in the subclones and the lack of detectable minor bands in any of the cell populations is consistent with the low levels of spontaneous gene cluster restructuring seen in other *BLM*+ transformed human cell lines [34].

Discussion

We show here that expression of the *E. coli* RecG protein is able to largely, but not completely, suppress two functional human cellular phenotypes of BLM deficiency: elevated sister chromatid exchange, and elevated gene cluster instability. Although this result is in accord with the several shared biochemical activities of these two proteins, this functional suppression across different kingdoms of biology is interesting, particularly when considering that there appears to be no relationship of evolutionary sequence homology between RecG and BLM. Both RecG and BLM are members of the DEXDc superfamily of helicase proteins, and both contain both DEXDc and HELICc subdomains (Figure 8 and [44]), however, the *E. coli* RecQ protein is the clear evolutionary homolog of BLM on the basis of sequence similarity within these two conserved domains. The helicase family membership of RecG notwithstanding, RecG does not primarily fulfill the classical function of a helicase but rather seems to have co-opted the molecular engine of a helicase to translocate on double stranded DNA. We suggest that this manner of translocation may also be true of BLM particularly during the branch migration of Holliday junctions since the crystal structure of a Holliday junction shows minimal single-stranded DNA character [45]. Both RecG and BLM have extended amino acid sequences N-terminal to their DEXDc superfamily domains and it is possible that significant structure/function similarities to RecG may reside in this region. Alternatively, BLM may maintain a portion of the Holliday junction in transient single-stranded form as Holliday junctions branch migrate together during junction dissolution.

The *E. coli* RecG protein, with essentially no sequence similarity to human BLM, would seem unlikely to engage in any of the well-characterized protein-protein interactions that are important for function of the BLM protein in the wild-type human cellular context, although we have not experimentally ruled out these interactions. The capacity with which RecG can suppress *BLM*- cellular phenotypes suggests therefore that the physiological role of BLM in suppressing the two phenotype we investigated, sister chromatid exchanges and gene cluster instability, is to perform the same molecular reactions that can be performed in a human cell by the RecG monomer alone, namely direct manipulations of DNA structures. This is not to rule out a role for BLM and associated protein-protein interactions in suppression of other manifestations of genomic instability, or potentially BLM accessory proteins may help localize BLM to DNA structures upon which both BLM and RecG can act, thereby reducing the required level of BLM in the cell relative to the large amount of ectopic RecG expressed in the cell lines of this study. Activities of BLM not known to be shared by RecG, such as the unwinding of G-quadruplex DNA, must play a minor role in both sister chromatid exchange and gene cluster instability suppression assayed herein, or alternatively, a G-quadruplex activity of RecG may remain to be discovered.

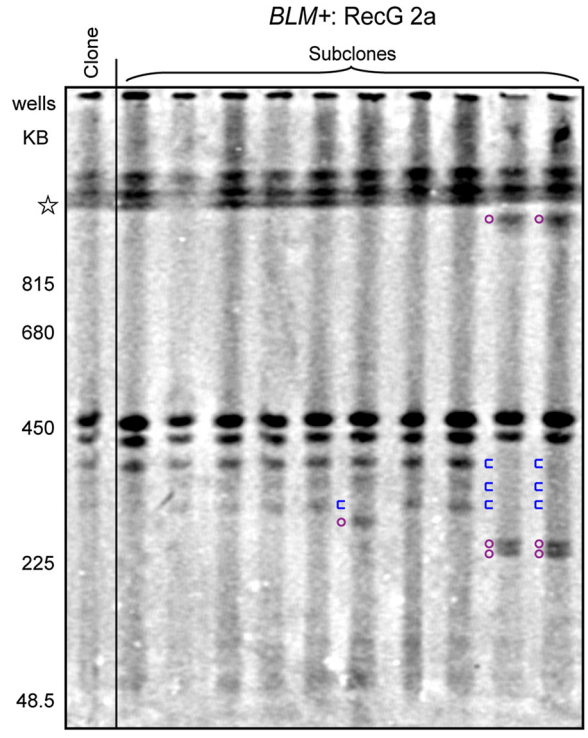

Figure 7 RecG has no effect on gene cluster instability in cells with normal BLM. Southern blotting detection of the rDNA clusters in *BLM*+ cells stably expressing higher levels of co-translated RecG/EGFP, "*BLM*+: RecG 2a". Cluster lengths resolved up to 1 Mb are shown. Cluster lengths are calibrated by *S. cerevisiae* chromosome size markers. Open star is the 1 Mb resolution limit of the gel: all clusters larger than 1 Mb co-localize in this region. A clonal cell population is shown in the far left lane of the gel. A collection of subclonal populations each derived from a single cell of this clonal population are shown to the right. Plum circles: relocated major band clusters. Blue braces: former location of major band clusters from the parental clone.

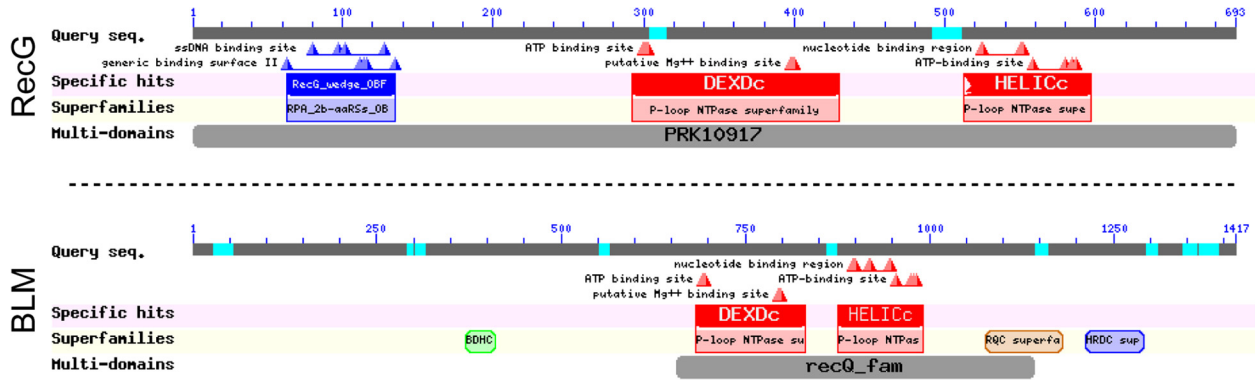

Figure 8 Domain-based alignment of *E. coli* RecG and human BLM proteins. All conserved domain analysis below is from [44], RecG accession GI:42669, BLM accession GI:4557365. RecG Conserved Domains: RecG_wedge_OBF: A subfamily of OB folds corresponding to the OB fold found in the N-terminal (wedge) domain of Escherichia coli RecG. DEXDc: DEAD-like helicases superfamily. A diverse family of proteins involved in ATP-dependent RNA or DNA unwinding. This domain contains the ATP-binding region. HELICc: Helicase superfamily c-terminal domain; associated with DEXDc-, DEAD-, and DEAH-box proteins, yeast initiation factor 4A, Ski2p, and Hepatitis C virus NS3 helicases. PRK10917: ATP-dependent DNA helicase RecG; Provisional BLM Conserved Domains: BDHCT: This is a C-terminal domain in Bloom's syndrome DEAD helicase subfamily. DEXDc: DEAD-like helicases superfamily. A diverse family of proteins involved in ATP-dependent RNA or DNA unwinding. This domain contains the ATP-binding region. HELICc: Helicase superfamily c-terminal domain; associated with DEXDc-, DEAD-, and DEAH-box proteins, yeast initiation factor 4A, Ski2p, and Hepatitis C virus NS3 helicases. RQC domain: This DNA-binding domain is found in the RecQ helicase among others and has a helix-turn-helix structure. The RQC domain, found only in RecQ family enzymes, is a high affinity G4 DNA binding domain. HRDC domain: The HRDC (Helicase and RNase D C-terminal) domain has a putative role in nucleic acid binding. recQ_fam: ATP-dependent DNA helicase, RecQ family. All proteins in this family for which functions are known are 3′-5′ DNA-DNA helicases. These proteins are used for recombination, recombinational repair, and possibly maintenance of chromosome stability. This family is based on the phylogenomic analysis of JA Eisen (1999, Ph.D. Thesis, Stanford University).

We suggest that RecG and BLM fulfill the same physiological function in *E. coli* and human cells respectively: they recognize stalled/regressed DNA replication forks and then utilize their common biochemical functions to mitigate this topological genomic damage. Given the well-established biochemical similarities between the two proteins and the capacity for RecG to functionally substitute for BLM shown in this work, we consider this the most parsimonious conclusion. Nevertheless, an alternative possibility consistent with our data is that the lesions recognized by BLM and RecG may be distinct from each other, but that may be converted though subsequent processing into a common form. For example, human cells, lacking RecG, may over-produce the classic substrate for RecG: a replication fork missing the nascent leading strand which may then be converted, by some as yet uncharacterized mechanism, into a bona fide distinct substrate for the BLM protein. The converse situation is also possible: that in the absence of BLM, the lesion recognized by BLM is converted into a different lesion that can be recognized and repaired by RecG. An additional possibility is that the lesion recognized by BLM and RecG may be the same, yet the two proteins may convert the lesion to different, but still less genomically destabilizing alternative structures. Mechanistic details aside, that the bacterial RecG protein can relieve human genomic instability phenotypes resulting from BLM protein deficiency demonstrates the highly conserved nature of DNA structural lesions and the type of mitigating enzymatic processing required to preserve genomic stability.

The role of BLM in human genomic stabilization is well-established. In contrast, a major mechanistic role of RecG in *E. coli* genomic stabilization has only recently been elucidated. In the absence of RecG, PriA initiates spurious replication forks throughout the bacterial genome, causing poorly-controlled genomic over-replication and compromised viability [46]. In the absence of RecG, cells have an absolute requirement for a ssDNA exonuclease activity, unless PriA is also eliminated [47]. It will be interesting to determine whether human *BLM*-cells share this over-replicated phenotype. Curiously, the *E. coli* RecQ protein rather than protecting the cell from aberrant recombination structures as the human BLM and *E. coli* RecG proteins do, seems to act instead to promote their formation [48]. Cells lacking the UvrD-mediated inhibition of recombination and also lacking RecG are killed by formation of intermolecular recombination intermediates (IRIs) that interfere with correct genome segregation to daughter cells. Formation of IRIs is caused by the action of RecQ and partner proteins and causes "death by recombination". Deletion of RecQ restores cellular viability [49]. An analogous phenotype in human cells might be the elevated formation of anaphase bridges in cells deficient for BLM [50]. One evolutionary interpretation consistent with our functional

data here would be that although the ancient common protein ancestor of both RecQ and BLM was preserved in evolutionary descent through both lineages to provide a core helicase domain, the functionality of these proteins diverged in opposite directions. Since there are no RecG homologs by sequence similarity in human cells, it would appear that at least one of the five human RecQ paralogs, BLM, has at least to some extent taken on the important recombination molecular transactions provided to *E. coli* by the RecG protein, and is an example of convergent evolution.

Conclusion

E. coli RecG and human BLM play similar, and to some extent interchangeable, roles in suppressing genomic instability phenotypes when expressed in human cells. Accordingly, given their lack of sequence homology, *E. coli* RecG and human BLM may be an example of convergent evolution.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MWK carried out cell line development and molecular analysis, DMS participated in gene cluster instability experiments, WAW subcloned the RecG coding sequence from *E. coli*. AJP directed the studies and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors very gratefully acknowledge the assistance of Dr. Bob Lloyd (University of Nottingham) for assessing the functionality of the RecG-EGFP fusion protein in *E. coli*. This work was supported by the Markey Cancer Research Center.

Author details

[1]Department of Microbiology, Immunology and Molecular Genetics, Markey Cancer Center, University of Kentucky, Lexington, KY, USA. [2]Present address: Western Kentucky University, Bowling Green, KY, USA. [3]Sarah Cannon Research Institute, Nashville, TN, USA. [4]Department of Radiation Medicine, College of Medicine, University of Kentucky, Lexington, KY, USA. [5]Present address: MedImmune, LLC, Gaithersburg, MD, USA.

Received: 21 January 2012 Accepted: 11 October 2012
Published: 30 October 2012

References

1. Monnat RJ Jr: **Human RECQ helicases: roles in DNA metabolism, mutagenesis and cancer biology.** *Semin Cancer Biol* 2010, **20**(5):329–339.
2. Bernstein KA, Gangloff S, Rothstein R: **The RecQ DNA helicases in DNA repair.** *Annu Rev Genet* 2010, **44**:393–417.
3. German J: **Bloom's syndrome. XX. The first 100 cancers.** *Cancer Genet Cytogenet* 1997, **93**(1):100–106.
4. Ralf C, Hickson ID, Wu L: **The Bloom's syndrome helicase can promote the regression of a model replication fork.** *J Biol Chem* 2006, **281**(32):22839–22846.
5. Machwe A, Xiao L, Groden J, Orren DK: **The Werner and Bloom syndrome proteins catalyze regression of a model replication fork.** *Biochemistry (Mosc)* 2006, **45**(47):13939–13946.
6. McGlynn P, Lloyd RG, Marians KJ: **Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled.** *Proc Natl Acad Sci USA* 2001, **98**(15):8235–8240.
7. Robu ME, Inman RB, Cox MM: **Situational repair of replication forks: roles of RecG and RecA proteins.** *J Biol Chem* 2004, **279**(12):10973–10981.
8. Karow JK, Constantinou A, Li JL, West SC, Hickson ID: **The Bloom's syndrome gene product promotes branch migration of Holliday junctions.** *Proc Natl Acad Sci USA* 2000, **97**(12):6504–6508.
9. Plank JL, Wu J, Hsieh TS: **Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration.** *Proc Natl Acad Sci USA* 2006, **103**(30):11118–11123.
10. Whitby MC, Vincent SD, Lloyd RG: **Branch migration of Holliday junctions: identification of RecG protein as a junction specific DNA helicase.** *EMBO J* 1994, **13**(21):5220–5228.
11. Grove JI, Harris L, Buckman C, Lloyd RG: **DNA double strand break repair and crossing over mediated by RuvABC resolvase and RecG translocase.** *DNA Repair (Amst)* 2008, **7**(9):1517–1530.
12. Bachrati CZ, Borts RH, Hickson ID: **Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase.** *Nucleic Acids Res* 2006, **34**(8):2269–2279.
13. van Brabant AJ, Ye T, Sanz M, German IJ, Ellis NA, Holloman WK: **Binding and melting of D-loops by the Bloom syndrome helicase.** *Biochemistry (Mosc)* 2000, **39**(47):14617–14625.
14. McGlynn P, Al-Deib AA, Liu J, Marians KJ, Lloyd RG: **The DNA replication protein PriA and the recombination protein RecG bind D-loops.** *J Mol Biol* 1997, **270**(2):212–221.
15. Briggs GS, Mahdi AA, Weller GR, Wen Q, Lloyd RG: **Interplay between DNA replication, recombination and repair based on the structure of RecG helicase.** *Philos Trans R Soc Lond B Biol Sci* 2004, **359**(1441):49–59.
16. Singleton MR, Scaife S, Wigley DB: **Structural analysis of DNA replication fork reversal by RecG.** *Cell* 2001, **107**(1):79–89.
17. Nimonkar AV, Ozsoy AZ, Genschel J, Modrich P, Kowalczykowski SC: **Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair.** *Proc Natl Acad Sci USA* 2008, **105**(44):16906–16911.
18. Courcelle J, Crowley DJ, Hanawalt PC: **Recovery of DNA replication in UV-irradiated Escherichia coli requires both excision repair and recF protein function.** *J Bacteriol* 1999, **181**(3):916–922.
19. Handa N, Morimatsu K, Lovett ST, Kowalczykowski SC: **Reconstitution of initial steps of dsDNA break repair by the RecF pathway of E. coli.** *Genes Dev* 2009, **23**(10):1234–1245.
20. Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, Kowalczykowski SC: **BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair.** *Genes Dev* 2011, **25**(4):350–362.
21. Sun H, Karow JK, Hickson ID, Maizels N: **The Bloom's syndrome helicase unwinds G4 DNA.** *J Biol Chem* 1998, **273**(42):27587–27592.
22. Cheok CF, Wu L, Garcia PL, Janscak P, Hickson ID: **The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA.** *Nucleic Acids Res* 2005, **33**(12):3932–3941.
23. Machwe A, Xiao L, Groden J, Matson SW, Orren DK: **RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange.** *J Biol Chem* 2005, **280**(24):23397–23407.
24. Chen CF, Brill SJ: **An essential DNA strand-exchange activity is conserved in the divergent N-termini of BLM orthologs.** *EMBO J* 2010, **29**(10):1713–1725.
25. Raynard S, Bussen W, Sung P: **A double Holliday junction dissolvasome comprising BLM, topoisomerase IIIalpha, and BLAP75.** *J Biol Chem* 2006, **281**(20):13861–13864.
26. Singh TR, Ali AM, Busygina V, Raynard S, Fan Q, Du CH, Andreassen PR, Sung P, Meetei AR: **BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome.** *Genes Dev* 2008, **22**(20):2856–2868.
27. Wu L, Hickson ID: **DNA helicases required for homologous recombination and repair of damaged replication forks.** *Annu Rev Genet* 2006, **40**:279–306.
28. Xu D, Guo R, Sobeck A, Bachrati CZ, Yang J, Enomoto T, Brown GW, Hoatlin ME, Hickson ID, Wang W: **RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability.** *Genes Dev* 2008, **22**(20):2843–2855.
29. Wu L, Davies SL, Levitt NC, Hickson ID: **Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51.** *J Biol Chem* 2001, **276**(22):19375–19381.
30. Meetei AR, Sechi S, Wallisch M, Yang D, Young MK, Joenje H, Hoatlin ME, Wang W: **A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome.** *Mol Cell Biol* 2003, **23**(10):3417–3426.
31. McGlynn P, Mahdi AA, Lloyd RG: **Characterisation of the catalytically active form of RecG helicase.** *Nucleic Acids Res* 2000, **28**(12):2324–2332.

32. Amor-Gueret M: **Bloom syndrome, genomic instability and cancer: the SOS-like hypothesis.** *Cancer Lett* 2006, **236**(1):1–12.
33. Chaganti RS, Schonberg S, German J: **A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes.** *Proc Natl Acad Sci USA* 1974, **71**(11):4508–4512.
34. Killen MW, Stults DM, Adachi N, Hanakahi L, Pierce AJ: **Loss of Bloom syndrome protein destabilizes human gene cluster architecture.** *Hum Mol Genet* 2009, **18**(18):3417–3428.
35. Kozak M: **Determinants of translational fidelity and efficiency in vertebrate mRNAs.** *Biochimie* 1994, **76**(9):815–821.
36. Maier T, Drapal N, Thanbichler M, Bock A: **Strep-tag II affinity purification: an approach to study intermediates of metalloenzyme biosynthesis.** *Anal Biochem* 1998, **259**(1):68–73.
37. Donnelly ML, Hughes LE, Luke G, Mendoza H, ten Dam E, Gani D, Ryan MD: **The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences.** *J Gen Virol* 2001, **82**(Pt 5):1027–1041.
38. Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, Vignali DA: **Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector.** *Nat Biotechnol* 2004, **22**(5):589–594.
39. Gaymes TJ, North PS, Brady N, Hickson ID, Mufti GJ, Rassool FV: **Increased error-prone non homologous DNA end-joining–a proposed mechanism of chromosomal instability in Bloom's syndrome.** *Oncogene* 2002, **21**(16):2525–2533.
40. Perry P, Wolff S: **New Giemsa method for the differential staining of sister chromatids.** *Nature* 1974, **251**(5471):156–158.
41. Stults DM, Killen MW, Shelton BJ, Pierce AJ: **Recombination phenotypes of the NCI-60 collection of human cancer cells.** *BMC Mol Biol* 2011, **12**:23.
42. Stults DM, Killen MW, Pierce HH, Pierce AJ: **Genomic architecture and inheritance of human ribosomal RNA gene clusters.** *Genome Res* 2008, **18**(1):13–18.
43. Mahdi AA, McGlynn P, Levett SD, Lloyd RG: **DNA binding and helicase domains of the Escherichia coli recombination protein RecG.** *Nucleic Acids Res* 1997, **25**(19):3875–3880.
44. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH: **CDD: a Conserved Domain Database for the functional annotation of proteins.** *Nucleic Acids Res* 2011, **39**(Database issue):D225–D229.
45. Ortiz-Lombardia M, Gonzalez A, Eritja R, Aymami J, Azorin F, Coll M: **Crystal structure of a DNA Holliday junction.** *Nat Struct Biol* 1999, **6**(10):913–917.
46. Rudolph CJ, Upton AL: *Harris L*. Lloyd RG: Pathological replication in cells lacking RecG DNA translocase. Mol Microbiol; 2009.
47. Rudolph CJ, Mahdi AA, Upton AL, Lloyd RG: **RecG protein and single-strand DNA exonucleases avoid cell lethality associated with PriA helicase activity in Escherichia coli.** *Genetics* 2010, **186**(2):473–492.
48. Magner DB, Blankschien MD, Lee JA, Pennington JM, Lupski JR, Rosenberg SM: **RecQ promotes toxic recombination in cells lacking recombination intermediate-removal proteins.** *Mol Cell* 2007, **26**(2):273–286.
49. Fonville NC, Blankschien MD, Magner DB, Rosenberg SM: **RecQ-dependent death-by-recombination in cells lacking RecG and UvrD.** *DNA Repair (Amst)* 2010, **9**(4):403–413.
50. Chan KL, North PS, Hickson ID: **BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges.** *EMBO J* 2007, **26**(14):3397–3409.

doi:10.1186/1471-2199-13-33
Cite this article as: Killen *et al.*: *Escherichia coli* RecG functionally suppresses human Bloom syndrome phenotypes. *BMC Molecular Biology* 2012 **13**:33.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central