

Non-Bloom syndrome-associated partial and total loss-of-function variants of BLM helicase

Hamed Mirzaei and Kristina H. Schmidt¹

Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, Tampa, FL 33620

Edited* by Richard D. Kolodner, Ludwig Institute for Cancer Research, La Jolla, CA, and approved October 9, 2012 (received for review June 15, 2012)

Bloom syndrome (BS) is an autosomal recessive disorder caused by mutations in the RecQ-like DNA helicase BLM, which functions in the maintenance of genome stability. Using a humanized model of *Saccharomyces cerevisiae* that expresses a chimera of the N terminus of yeast Sgs1 and the C terminus of human BLM from the chromosomal *SGS1* locus, we have functionally evaluated 27 BLM alleles that are not currently known to be associated with BS. We identified nine alleles with impaired function when assessed for hypersensitivity to the DNA-damaging agent hydroxyurea (HU). Six of these alleles (P690L, R717T, W803R, Y811C, F857L, G972V) caused sensitivity to HU that was comparable to known BS-associated or helicase-dead alleles, suggesting that they may cause BS and, in the heterozygous state, act as risk factors for cancerogenesis. We also identified three alleles (R791C, P868L, G1120R) that caused intermediate sensitivity to HU; although unlikely to cause BS, these partial loss-of-function alleles may increase risk for cancers or other BS-associated complications if a person is homozygous or compound heterozygous for these alleles or if they carry a known BS-associated allele.

Bloom syndrome (BS) is a rare genetic disorder with 264 affected individuals worldwide currently included in the BS registry (1). BS is characterized by proportional dwarfism, erythema on sun-exposed skin, hyper- or hypopigmented areas of skin, immunodeficiency, and subfertility (2, 3). Persons with BS have a striking predisposition to cancer development and increased risk for early-onset type-II diabetes (3). Life expectancy is short, ranging from 1 to 49 y, with a mean age at death of 27 (1, 4).

BS is caused by mutations in the *BLM* gene, whose 22 exons encode a 3'-5' DNA helicase that is related to RecQ of *Escherichia coli* and Sgs1 of *Saccharomyces cerevisiae* (5, 6). Nineteen recurring and 45 unique mutations in *BLM* have been identified as causative of BS (7). *Blm*^{Ash}, which occurs at a frequency of ~0.5% in the Ashkenazi Jewish population, is the most common BS-associated mutation, a frameshift caused by deletion of six nucleotides and insertion of seven different nucleotides (8–11). Most other BS-associated mutations are also frameshifts or point mutations that lead to premature termination of BLM upstream of or within the helicase domain, causing complete loss of function. However, 13 missense mutations—7 in the helicase domain and 6 in the RecQ C-terminal (RQC) domain—have also been identified in persons with BS (7). Five of these missense mutations have been characterized in vitro and were found to impair BLM's ability to bind and hydrolyze ATP or to bind DNA (12). Although a crystal structure of the BLM helicase core has not yet been reported, modeling of known BS-associated mutations onto the crystal structure of *E. coli* RecQ has revealed molecular and structural explanations for these defects (12). Crystal structures of other RecQ-like DNA helicases, most notably those of the helicase core of human RecQL1 and the winged-helix (WH) domain of Werner syndrome protein (WRN) in a complex with DNA, have revealed that even domains that show no significant sequence identity, such as the WH domains, are structurally and functionally conserved within the RecQ helicase family (13, 14). Normal BLM function also depends on its physical interaction with topoisomerase III α and Rmi1/Rmi2. Together they form the so-called BRT (BLM/Rmi1/Rmi2/topoisomerase III α) complex, which is involved in dissolving homologous recombination intermediates as non-cross-overs (reviewed

in ref. 15). Interactions with Rad51, Mlh1, p53, and ataxia telangiectasia mutated (ATM) protein kinase have also been reported (16, 17). Thus, BS cells are hypersensitive to DNA-damaging agents, show elevated rates of mitotic recombination, especially between sister chromatids but also between homologous chromosomes, and accumulate abnormal replication intermediates (18–21). Evidence that heterozygosity for BS-associated mutations is associated with increased cancer risk comes from a study that showed a greater than twofold increase in colorectal cancer risk in carriers of the *blm*^{Ash} mutation, although this association is somewhat uncertain (22, 23). Associations between noncoding single-nucleotide polymorphisms (SNPs) in *BLM* and breast cancer and to a lesser extent bladder cancer, malignant melanoma, acute myeloid leukemia, and myelodysplastic syndromes have also been reported (24, 25), and it has been suggested that these noncoding SNPs may be linked to functional SNPs in exons of *BLM*. However, it is unknown which, if any, of the 75 currently known coding *BLM* variants alter BLM function. Here we have used an *S. cerevisiae* system that expresses a chimera composed of the N terminus of Sgs1 and the C terminus of BLM under control of the chromosomal *SGS1* promoter to functionally evaluate 27 variants of the *BLM* gene that are not currently known to be associated with BS.

Results

Identification of Total Loss-of-Function BLM Alleles Not Currently Associated with BS. Like BS cells, yeast cells lacking Sgs1 are hypersensitive to DNA-damaging agents (26). Deletion of the C-terminal 800 residues of Sgs1 eliminates the HRDC (helicase and RNase D C-terminal), helicase, and RQC domains, the latter two of which make up the helicase core of Sgs1. Hydroxyurea (HU) sensitivity of cells expressing this Sgs1 truncation is suppressed by replacing the missing C-terminal 800 residues of Sgs1 with the C-terminal half of human BLM (residues 648–1417) (27) (Fig. 1A and B). Fig. 1B shows that the function of this chimera depends on enzymatic activity contributed by BLM because mutating the conserved lysine in the Walker A motif of BLM (K695R) or introducing mutations known to cause BS (C1036F, C1055S, D1064V) inactivated the chimera, causing severe sensitivity to HU. The region of BLM included in the chimera contains 41 of the 75 missense mutations from the National Center for Biotechnology Information database of Short Genetic Variations (dbSNP, <http://www.ncbi.nlm.nih.gov/projects/SNP/>) that are not currently known to be associated with BS and 2 additional, rare variants (c.2069C>T, P690L; c.3970C>T, H1324T) reported by German et al. (7). Using PolyPhen (28), we ranked these 43 variants according to their probability to impair BLM function (Table S1). The 11 variants predicted to be “probably damaging” (score >2.0) and

Author contributions: H.M. and K.H.S. designed research; H.M. performed research; H.M. contributed new reagents/analytic tools; H.M. and K.H.S. analyzed data; and H.M. and K.H.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: kschmidt@usf.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1210304109/-DCSupplemental.

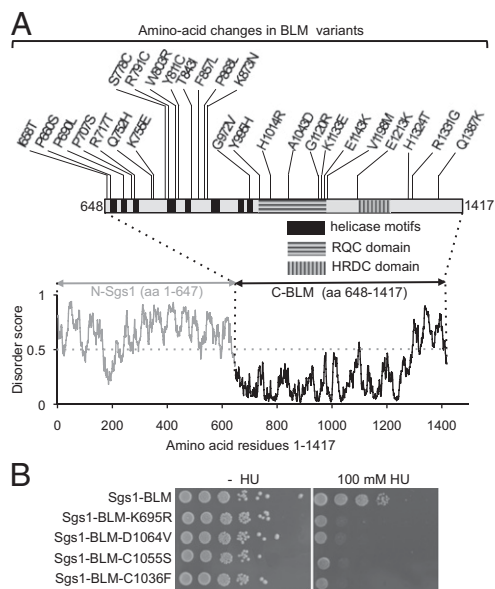


Fig. 1. Location of 27 BLM variants evaluated in this study. (A) The C-terminal 770 residues of BLM were fused to the 647 N-terminal residues of Sgs1 to construct a chimera (Sgs1-BLM) that confers resistance to hydroxyurea in cells lacking Sgs1. Conserved domains are indicated by black boxes (helicase motifs) and by horizontally (RQC) or vertically (HRDC) striped boxes (12). Disorder scores are from a server for the prediction of intrinsically unstructured regions of proteins (IUPred, <http://iupred.enzim.hu/>), with >0.5 predicting disorder and <0.5 predicting order. (B) HU sensitivity of yeast cells expressing the Sgs1-BLM chimera, the helicase-dead Sgs1-BLM-K695R chimera, and chimeras with mutations known to cause BS (C1036F, C1055S, D1064V).

the 11 variants predicted to be “possibly damaging” (score >1.5) were selected for in vivo analysis (Table 1). In addition, we chose the three highest-ranked variants with a score <1.5 (“benign”) and two lower-ranked variants for a total of 27 amino acid changes to be functionally evaluated (Fig. 1A and Table 1). These 27 variants were introduced into *BLM* cDNA and stably integrated at the chromosomal *SGS1* locus. Expression of the BLM variants from the endogenous *SGS1* promoter ensured cell cycle-regulated expression and copy number control.

By assessing the ability to grow in the presence of HU, we identified six total loss-of-function *BLM* alleles that are currently not associated with BS: P690L, R717T, W803R, Y811C, F857L, and G972V (Fig. 2). Expression of these BLM variants caused cells to be as sensitive to HU as cells expressing the helicase-dead K695R allele (Fig. 2) or known BS-associated alleles (Fig. 1B). Yeast cells expressing these alleles were also hypersensitive to the alkylating agent methane methylsulfonate (MMS) and 4-nitroquinoline 1-oxide (4NQO), which induces bulky DNA adducts (Fig. S1). Three of these six mutations affect residues in conserved helicase motifs (Fig. 3A): P690 maps to the Walker A motif, R717 to motif Ia, and G972 to motif VI. W803 is located in a highly conserved aromatic-rich (AR) loop that follows the Walker B motif. Y811 and F857 are located between motifs II and III and between motifs III and IV, respectively. Only W803R affects a residue that is conserved in all five human RecQ homologs (BLM, WRN, RecQL1, RecQL4, and RecQL5), whereas the others affect residues that differ in at least one of them. The arginine at position 717 is not conserved, with leucine or isoleucine being present at those positions in the other four human RecQ-like helicases, in Sgs1, and in RecQ. Although the structure of the helicase domain of BLM has not been solved, those of *E. coli* RecQ and human RecQL1 were recently reported (14, 29). Based on ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) alignments, P690, R717, W803, Y811, F857, and G972 of BLM correspond to P114, I141, W227, Y235,

F281, and N397 of RecQL1 with no significant gaps in the alignment (Fig. 3A). Modeling the BLM residues onto the RecQL1 structure revealed that P690, R717, W803, and Y811 are located in the first of the two RecA-like lobes of the helicase domain, whereas G972 is located in the second lobe, with F857 in the loop connecting the two lobes (Fig. 3C). Notably, all six affected residues are located near the surface of the deep cleft between the lobes. Residues exposed along this cleft, including those that make up the conserved helicase motifs, are critical for ATP binding and hydrolysis as well as single-stranded DNA binding (12, 14, 29–32).

Identification of Partial Loss-of-Function *BLM* Alleles. In addition to total loss-of-function alleles, we identified three alleles (R791C, P868L, G1120R) that caused slight growth inhibition on HU, suggesting a partial loss of function (Fig. 2 and Table 1). Cells expressing these alleles also showed slightly increased sensitivity to 4NQO, and to a lesser extent to MMS (Fig. S1). These partial loss-of-function mutations affect residues in the helicase core; however, they are not on the surface of the cleft between the two lobes. Instead, R791 is in an internal β -sheet near the invariant aspartic acid (D795) and glutamate (E796) residues of the Walker B motif (Fig. 3A and D), whereas P868 is farthest from the cleft in a loop that follows the first β -sheet of lobe 2 (Fig. 3E). The third affected residue, G1120, is located in the α 2- α 3 loop of

Table 1. Coding, nonsynonymous *BLM* variants functionally evaluated in this study

Mutation	Domain	PolyPhen score	HU
I658T	HD	1.860	–
H660Y	HD (motif 0)	2.211	–
P690L	HD (motif I)	2.724	+++
P707S	HD	1.951	–
R717T	HD (motif Ia)	2.008	+++
Q752H	HD	1.921	–
K755E	HD	1.493	–
S778C	HD	1.773	–
R791C	HD (motif II)	2.495	+
W803R	HD (AR loop)	3.757	+++
Y811C	HD	2.616	+++
T843I*	HD	1.796	+
F857L	HD	2.001	+++
P868L	HD	2.724	+
K873N	HD	1.649	–
G972V	HD (motif VI)	2.397	+++
Y995H	RQC (zinc)	1.716	–
H1014R	RQC (zinc)	2.360	–
A1043D	RQC (zinc)	1.440	–
G1120R	RQC (WH)	2.172	+
K1133E	RQC (WH)	1.550	–
E1143K	RQC (WH)	1.554	–
V1198M	–	0.375	–
E1213K	HRDC	1.470	–
H1324T	–	1.098	–
R1331G	–	1.787	–
Q1387K	–	1.569	–

Amino acid changes in BLM were identified in the dbSNP database and mapped to BLM. HD, helicase domain; RQC, RecQ C-terminal domain with zinc-binding and winged-helix subdomains. Polymorphism phenotyping (PolyPhen) scores (28) estimate the probability of nonsynonymous mutations to affect protein function (>2, probably damaging; <2, possibly damaging; <1.5, benign). HU refers to Fig. 2 (+++), sensitive as helicase-dead K695R; +, intermediate sensitivity between wild type and K695R; –, equivalent to wild type). *BLM* variants P690L and H1324T were reported by German et al. (7). *A variant originally reported in the dbSNP database as a missense mutation but later identified as the deletion of exons 11 and 12, resulting in a premature stop codon (3, 6) is indicated.

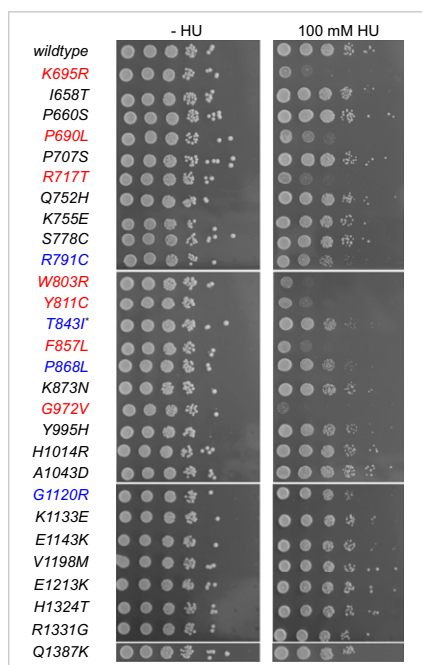


Fig. 2. HU sensitivity of yeast cells expressing BLM variants. Variants were chromosomally expressed from the endogenous *SGS1* promoter, and 10-fold dilutions of exponentially growing cultures were spotted on YPD with HU or without HU. Variants that cause sensitivity equivalent to that of the helicase-dead variant (*K695R*) are shown in red (total loss of function), those that grow as well as the wild type are in black, and those that cause HU sensitivity intermediate between the wild type and the helicase-dead variant are in blue (partial loss of function). The asterisk marks a BLM variant listed in the dbSNP database that was originally reported as a missense mutation (6) but was subsequently identified to be a deletion of exons 11 and 12, resulting in a premature stop codon (3).

the WH domain and appears to be the only BLM variant that exhibits impaired function due to a missense mutation in this domain (Fig. 3 *B, F*, and *G*). The WH domain of RecQ-like helicases has recently been implicated in DNA recognition and strand separation (13, 14). Although there is no significant sequence identity between the WH domains of RecQ helicases, ClustalW2 and structure-based alignments show that G1120 is conserved in Sgs1, RecQ, and all four human RecQ-like DNA helicases that possess an RQC domain (Fig. 3*B* and Fig. S2*B*) (13). When we changed G1120 to a different residue (G1120E) we observed the same partial defect, further supporting the importance of G1120 for BLM function (Fig. S3).

Although predicted by PolyPhen to be probably damaging, the *H1014R* and *H660Y* alleles did not cause hypersensitivity to HU. None of the *BLM* alleles predicted to be possibly damaging or benign impaired BLM function in the HU sensitivity assay (Fig. 2 and Table 1). Of the 17 amino acid changes that did not impair BLM function, 13 are located in domains for which crystal structures or solution structures have been determined in at least one member of the RecQ-like DNA helicase family (Fig. S4). Seven of these amino acid changes map to the helicase domain and affect residues along the periphery of lobe 1 or lobe 2 away from the cleft between the lobes where the conserved helicase motifs are located (Fig. S4*A*). Three other amino acid changes are located in the zinc-binding subdomain (Fig. S4*B*) and two are located in the α 3- α 4 loop of the WH subdomain, which, in contrast to the α 2- α 3 loop, where the partially defective G1120R mutation is located, is not known to be involved in DNA binding (Fig. S4*C*). Finally, one amino acid change that did not impair BLM function maps to the HRDC domain (Fig. S4*D*).

Corresponding Amino Acid Changes Impair Sgs1 Function. Of the five human RecQ-like DNA helicases, BLM is most closely related to Sgs1. Alignments between the two proteins show 50% identity across the helicase core, where the affected residues in all nine functionally impaired BLM variants are located. To determine whether the functional impairment was limited to BLM or whether the equivalent mutations also affected Sgs1 function, we constructed nine *sgs1* alleles that corresponded to the nine partially or fully impaired *BLM* alleles (Fig. 4*A* and Fig. S2). Identifying the equivalent residues in Sgs1 was unambiguous except for the proline residue at position 868, which in BLM is located in a lysine-rich loop (⁸⁶⁸PKKPKK⁸⁷³) recently suggested to interact with DNA (12). This proline is not present in Sgs1 and, in the absence of a crystal structure of Sgs1, it is unclear whether the loop is conserved. However, because the positive charge of the loop has been suggested to be functionally important for DNA binding, we decided to mutate the first of three lysines in this region (K881) to leucine. The nine mutations were introduced into the *SGS1* gene on a CEN/ARS plasmid expressed under control of the native *SGS1* promoter and then tested for their ability to complement the HU hypersensitivity of an *sgs1Δ* mutant (Fig. 4*B*). The *W816R* and *Y824C* alleles could not suppress the HU hypersensitivity of the *sgs1Δ* mutant whereas *P701L*, *F870L*, and *G1115R* partially suppressed it, indicating impaired function of these five *sgs1* alleles. The *I732T*, *R804C*, *K881L*, and *G983V* alleles appeared to fully suppress HU sensitivity of the *sgs1Δ* mutant, suggesting they retained activity that was similar to that of the *SGS1* wild-type allele. Taken together, four of the six amino acid changes that caused a total loss of BLM function also impaired Sgs1 function, whereas one of the three partial loss-of-function mutations did (Table 2). Notably, all *BLM* mutations that affected residues also conserved in Sgs1 and *E. coli* RecQ impaired Sgs1 function, whereas those that are not conserved in *E. coli* RecQ, Sgs1, and BLM did not (Fig. 4*A*, Table 2, and Fig. S2).

Genetic Interactions Between BLM Variants in Diploid Cells. To assess possible genetic interactions between *BLM* alleles, we constructed diploids with various combinations of wild-type, partial loss-of-function, and total loss-of-function alleles (Fig. 5 and Fig. S5). As expected, diploids homozygous or compound heterozygous for total loss-of-function alleles were most sensitive to HU. However, whereas we expected that diploids homozygous or heterozygous for wild-type alleles would not be hypersensitive to HU regardless of the status of the second allele, we observed that diploids heterozygous for total loss-of-function and wild-type alleles (Fig. 5, rows 4–7) were more sensitive to HU than diploids that were biallelic for wild-type alleles (Fig. 5, rows 1–3). We made similar observations when we included additional total loss-of-function alleles (*K695R*, *C1055S*) in the analysis (Fig. S5). HU sensitivity of diploids homozygous for the partial loss-of-function *P868L* allele or heterozygous for *P868L* and one of the total loss-of-function alleles fell between that of diploids biallelic for wild-type or total loss-of-function alleles (Fig. S5). These findings could be explained by insufficient wild-type BLM protein or a lower activity of multimeric BLM helicase complexes that consist of a mixture of wild-type and nonfunctional monomers.

Discussion

In this study, we have used *S. cerevisiae* to functionally evaluate 27 BLM variants that are not currently associated with BS. The mutant *BLM* alleles were chromosomally expressed from the endogenous *SGS1* promoter. The advantage over plasmid-based approaches or expression from inducible or constitutive promoters is that BLM expression in this system is cell cycle-regulated and copy number-controlled. Similar to Sgs1, BLM appears in S phase and is maintained at similar levels in G2/M, but has mostly disappeared by the time cells reenter G1 (33). Regulation of BLM expression by the chromosomal *SGS1* promoter was critical, because the goal of our study was to evaluate BLM variants that are not known to cause BS. Therefore, the

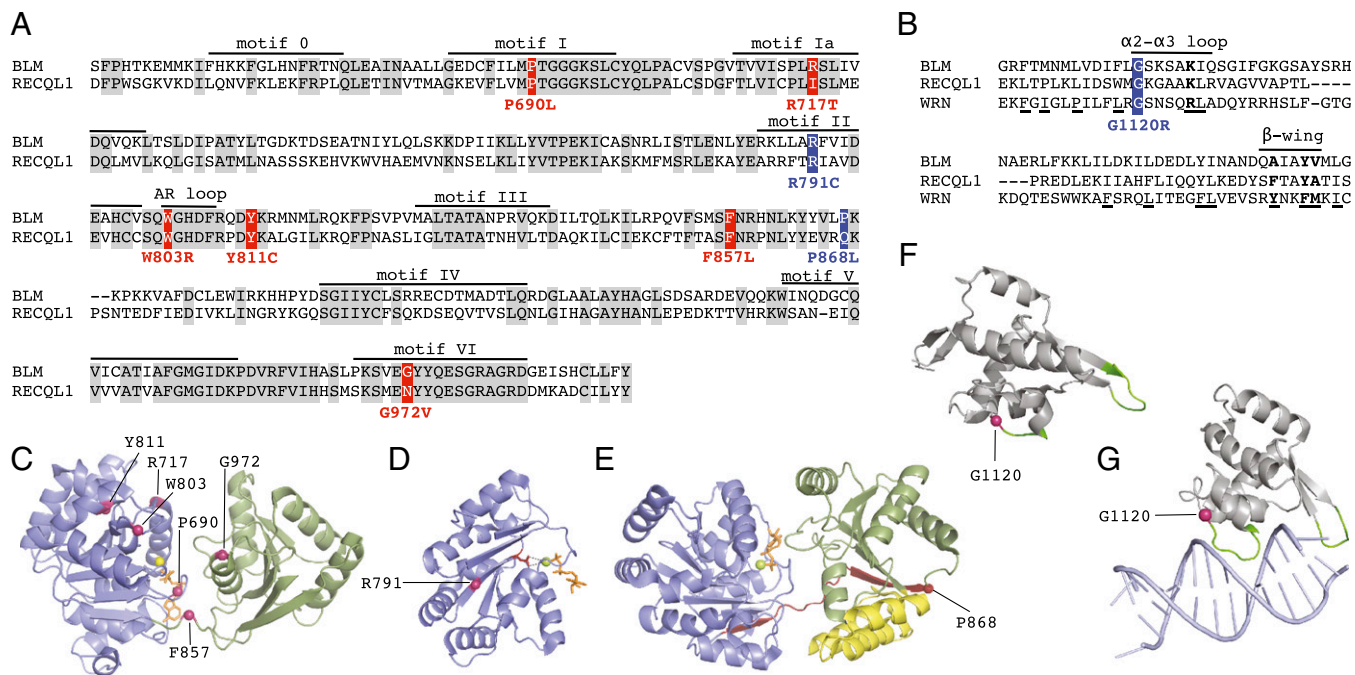


Fig. 3. Location of nine missense mutations found in this study to impair BLM function. (A) ClustalW2 alignment between the helicase domain of human BLM and RecQL1. Identical residues are shaded gray. Total and partial loss-of-function mutations are in red and blue, respectively. Helicase motifs and an aromatic-rich loop are indicated by horizontal lines. (B) Structure-based alignment of WH domains (partial) of BLM, RecQL1, and WRN, adapted from Kitano et al. (13). Structurally important residues are underlined and functionally important residues in the $\alpha 2$ - $\alpha 3$ loop and β -wing are shown in bold. The glycine in the G1120R mutation is shaded blue. (C–F) Residues affected in BLM variants are highlighted in the crystal structure of RecQL1 (14). (C) Location of the six total loss-of-function mutations (pink spheres) along the cleft between lobes 1 (blue) and 2 (green). ADP is shown as an orange stick model and Mg^{2+} is shown as a yellow sphere. (D) R791 is located in lobe 1, with the nearby, invariant residue D795 (Walker B) shown as a red stick model. (E) The residue corresponding to P868 is located at the base of lobe 2 where the zinc-binding domain (yellow) adjoins. The loop connecting lobes 1 and 2 and the β -sheets flanking this loop are indicated in red. (F) G1120 is located in the WH domain (gray). The $\alpha 2$ - $\alpha 3$ loop and β -wing are shown in green. (G) Location of the residue equivalent to G1120 in the WH domain of WRN crystallized with DNA (13). The $\alpha 2$ - $\alpha 3$ loop and β -wing are shown in green. Images were generated in PyMOL (<http://pymol.org>).

experimental system had to be designed such that besides wild-type and null alleles it could also detect partial loss-of-function alleles, which impair BLM function but at a level too weak to cause BS. The subtle defects exhibited by such weak alleles may be compensated for and thus go undetected in overexpression systems or in the absence of appropriate cell-cycle regulation.

Using this approach, we identified six total loss-of-function BLM alleles (P690L, R717T, W803R, Y811C, F857L, G972V) and three partial loss-of-function alleles (R791C, P868L, G1120R). Although biochemical analysis will uniquely identify the specific defect of each of these BLM variants, previous structure–function analyses of RecQ and other helicases as well as the in vitro characterization of known BS-associated mutations allow us to make some predictions. For example, although the P690L mutation does not affect one of the highly conserved GK(T/S) residues of motif I (Walker A), its location within this motif most likely leads to loss of function because of an inability to bind ATP and/or Mg^{2+} . The R717T mutation affects a variable residue of motif Ia. The effect of mutations in motif Ia has not yet been investigated, primarily because none of the missense mutations known to cause BS are located in this motif. However, structural analyses have revealed that residues from motif Ia engage in ssDNA binding (30, 31, 34, 35), and replacing the positively charged arginine with a neutral threonine may weaken ssDNA binding in this critical region. The third loss-of-function variant, W803, is located in an AR loop that was originally described in the PcrA helicase and shown to be critical for coupling ATP hydrolysis to DNA binding and unwinding (31, 36, 37). This AR loop is conserved in RecQ-like helicases, including BLM, Sgs1, and *E. coli* RecQ. The loss of function caused by mutating Y811, another aromatic residue just outside the AR loop, is likely due to

a similar inability to coordinate ATP hydrolysis with DNA binding/unwinding. F857 is located between motifs III and IV in the flexible loop that connects the two RecA-like lobes that make up the helicase domain. A mutation on this side of the interlobe cleft, which accommodates the ATP binding site, may interfere with ATP binding/hydrolysis. In SF-1 helicases, the “arginine-finger” motif VI in lobe 2 also reaches into this site to interact with ATP (31, 38), and it is plausible that the conserved arginine residues in motif VI of RecQ helicases (R978 and R982 of BLM), which also reach into the ATP binding site (39), serve the same function. Mutation of G972, which is located near this arginine finger, could therefore impair coordination between lobe 2 and ATP binding in lobe 1.

Besides total loss-of-function mutations, we identified three BLM alleles that retain partial activity (R791C, P868L, G1120R). Instead of being at or near the surface of the interlobe cleft where residues of the conserved motifs are exposed, the R791C and P868L mutations affect residues within or adjacent to internal β -sheets (Fig. 3 D and E). However, R791 is close to the invariant D795 of the DEXH motif II (Walker B) and, thus, the partial loss of function is most likely due to decreased ATP hydrolysis. In contrast, because the residue of RecQL1 equivalent to P868 of BLM is not near the interlobe cleft, its mutation is unlikely to affect the function of the conserved helicase motifs. Instead, its location near the bottom of lobe 2, where the zinc-binding domain adjoins the helicase domain (Fig. 3E), may weaken the interaction between these two domains, which is thought to be important for DNA binding and protein stability. In fact, P868 is the first residue in a lysine-rich loop (residues 869–873) that has been suggested to play a role in stable DNA binding by BLM (12). Unlike R791C and P868L, the G1120R

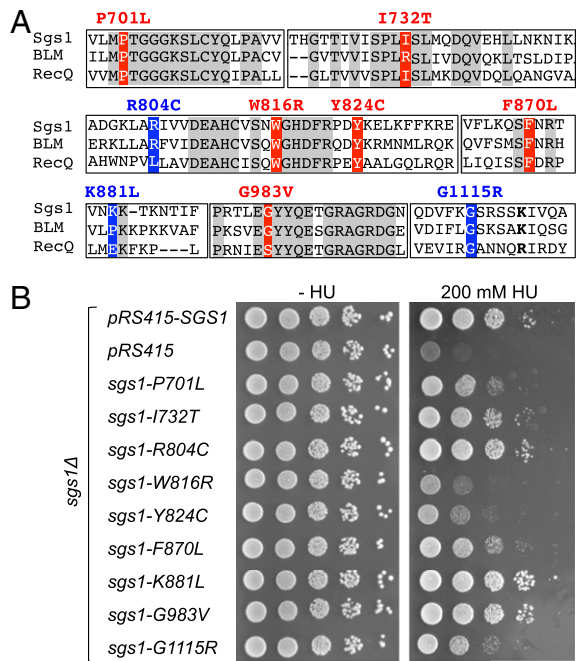


Fig. 4. HU hypersensitivity of *sgs1* mutants with amino acid changes corresponding to those that impair BLM function. (A) ClustalW2 alignment of Sgs1, human BLM, and *E. coli* RecQ. Amino acid changes in Sgs1 are indicated in red if the corresponding BLM mutation causes total loss of function or in blue if the corresponding BLM mutation causes partial loss of function. (B) *sgs1* mutant alleles were expressed from pRS415 and the native *SGS1* promoter to test for their ability to suppress HU hypersensitivity of *sgs1Δ* cells. Empty vector (pRS415) and vector expressing wild-type Sgs1 (pRS415-SGS1) served as negative and positive controls, respectively.

mutation is located in the RQC domain of the helicase core. It is the only known BLM variant whose defect is due to a missense mutation in the WH subdomain. A recent analysis of the crystal structure of the WH domain of WRN complexed with DNA indicates that in RecQ-like helicases two loops, one between the $\alpha 2$ and $\alpha 3$ helices and one between the $\beta 2$ and $\beta 3$ strands, are involved in sequence-independent DNA binding and strand separation, respectively, and undergo conformational changes upon DNA binding (13). The importance of the conserved glycine in the $\alpha 2$ - $\alpha 3$ loop is likely to stem from its low propensity (the lowest next to proline) for α -helix formation and, thus, its ability to maintain the loop structure and provide the flexibility needed for the conformational change (loop extension) and exposure of residues for DNA binding. Mutation analysis has

Table 2. HU hypersensitivity conferred by *sgs1* alleles coding for amino acid changes equivalent to those impairing BLM function

			HU hypersensitivity	
Amino acid change				
BLM	Sgs1	Domain	BLM	Sgs1
P690L	P701L	HD (motif I)	+++	++
R717T	I732T	HD (motif Ia)	+++	-
R791C	R804C	HD (motif II)	+	-
W803R	W816R	HD (AR loop)	+++	+++
Y811C	Y824C	HD	+++	+++
F857L	F870L	HD	+++	++
P868L	K881L	HD	+	-
G972V	G983V	HD (motif VI)	+++	-
G1120R	G1115R	RQC (WH)	+	++

confirmed the importance of the β -wing for DNA binding and helicase activity in WRN and RecQL1 (13, 14). The identification of the partially defective *G1120R* and *G1115R* alleles now also provides in vivo support for the functional importance of the $\alpha 2$ - $\alpha 3$ loop in BLM and Sgs1, respectively. Similarly, impairment of Sgs1 and BLM by the equivalent F870L and F857L mutations, respectively, suggests functional importance in both helicases of the $\beta 7$ - $\beta 8$ loop, where it reaches into the interlobe cleft. Equivalent mutations in or near the AR loop and the Walker A motif also impaired the function of both BLM and Sgs1. However, K881L (P868L in BLM) did not impair Sgs1 function, suggesting that the lysine-rich loop in lobe 2 of BLM and its role in helicase function are not conserved in Sgs1. Indeed, this finding provides in vivo support for the recent proposal that the lysine-rich loop and its function may be unique to BLM (12).

With the exception of the *blm^{Ash}* allele, which occurs at a frequency of $\sim 0.5\%$ in the Ashkenazi Jewish population, alleles known to cause BS are very rare ($< 0.1\%$) (7, 9, 10). Indeed, most BS mutations have only been found in a single individual with BS (7). It is therefore likely that additional BS alleles exist in the human population that have, due to their rarity and the autosomal recessive pattern of inheritance of BS, not yet been identified in a person with BS. We propose that the six total loss-of-function alleles may be candidates for new BS mutations. Based on the recent finding (40) that the frequency of *P868L* in the human population is $\sim 4.6\%$, we further propose that partially defective BLM variants (*R791C*, *P868L*, *G1120R*) are unlikely to be associated with full-scale BS, as the predicted homozygosity for the *P868L* allele in the human population far exceeds the incidence of BS.

Whether carriers for disease-causing *BLM* mutations also exhibit genome instability and are at increased risk for cancerogenesis is only beginning to be investigated. In one study, heterozygosity for the most common BS-associated mutation, *blm^{Ash}*, was associated with increased colorectal cancer risk, although this increase was not seen in another study (22, 23). Evidence of a gene dosage effect (haploinsufficiency) for *BLM* in colorectal tumorigenesis in mice also indicates that heterozygosity for *BLM* plays a role in cancer predisposition (41). Similarly, we observed that diploids heterozygous for loss-of-function *BLM* alleles were more sensitive to DNA-damaging agents than wild-type diploids. This hypersensitivity could be due to haploinsufficiency or decreased activity of helicase complexes that consist of mixtures of wild-type and mutant BLM monomers. The possibility that the defect may be specific to expression being controlled by the *SGS1* promoter and may not apply to protein levels expressed from the *BLM* promoter also cannot be excluded.

Besides the possibility of increased cancer risk for carriers of loss-of-function *BLM* mutations, hypomorphic *BLM* alleles—those that are partially functional and therefore do not cause all of the clinical features of BS—may be associated with increased cancer risk. With the identification of the *P868L* polymorphism as a hypomorphic *BLM* allele candidate, it will now be possible to determine whether such an association exists. A recent report,

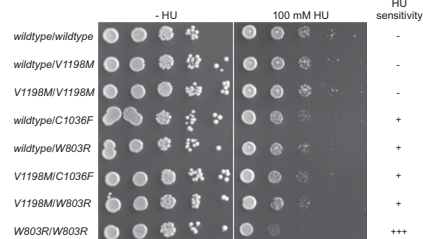


Fig. 5. Diploids heterozygous or homozygous for total loss-of-function alleles (*W803R*, *C1036F*) are more sensitive to HU than diploids homozygous for alleles exhibiting wild-type function (*wildtype*, *V1198M*).

which found an allele dosage-dependent effect of *P868L* on rectal cancer risk, does indeed point toward this possibility (42). With an estimated 1 in 10 individuals being carriers of *P868L* (40) and, thus, ~1 in 500 individuals being predicted to be homozygous, an association between *P868L* homozygosity and an elevated risk for cancer (or other risks associated with genome instability, such as hypersensitivity to chemotherapeutics or other genotoxic agents) would be significant. Despite the evidence from our yeast system, it will remain unclear whether *R791C* and *G1120R* are also hypomorphic alleles like *P868L* until the frequency of these alleles in the human population is determined. Notably, carriers of a known BS mutation (or one of the total loss-of-function alleles identified here) who also carry a hypomorphic allele such as *P868L*, that is, who have no fully functional *BLM* allele, might also be at higher risk for cancers commonly associated with BS, such as leukemia, lymphoma, and epithelial cancers (43), or they may exhibit a BS-like disorder, characterized by milder expression or later onset of a combination of BS-associated symptoms.

Methods

Yeast Strains, Plasmids, and Media. Strains were derived from RDKY3615 (*MATa*, *ura3-52*, *trp1Δ63*, *his3Δ200*, *leu2Δ1*, *lys2Bgl*, *hom3-10*, *ade2Δ1*, *ade8*, *hxt13::URA3*). Diploids were constructed by mating haploids with the desired *BLM* allele. *BLM* cDNA was purchased from Open Biosystems (clone

MGC286), and the *HIS3* cassette from pRS303 was inserted into a Spel site downstream of *BLM*. Variants were introduced into the cDNA by site-directed mutagenesis. A PCR product comprising nucleotides 1942–4251 of *BLM* and the *HIS3* marker cassette was used to replace nucleotides 1942–4341 of the chromosomal *SGS1* ORF in RDKY3615 by recombination-mediated integration (44), leading to expression of a fusion of the N terminus of Sgs1 and the C terminus of BLM (Sgs1-BLM). In-frame fusion at nucleotide 1941, the integrity of the integrated *BLM* cDNA sequence, and the presence of the desired variant were confirmed by sequencing. Mutations in plasmid-borne *SGS1* were introduced by site-directed mutagenesis and confirmed by sequencing of *SGS1*. Yeast strains and plasmids used in this study are listed in Tables S2 and S3. Yeasts were cultured in liquid yeast extract/peptone/dextrose (YPD) with shaking at 30 °C or grown on YPD agar at 30 °C.

Sensitivity to DNA-Damaging Agents. Cell cultures were grown overnight in YPD (or in synthetic media lacking leucine to maintain pRS415-derived plasmids), diluted to OD₆₀₀ 0.2, and incubated until they reached OD₆₀₀ 0.5–0.6. Tenfold dilutions were spotted on YPD and YPD supplemented with HU, MMS, or 4NQO as indicated.

ACKNOWLEDGMENTS. We thank Jessica Kennedy, Michael Damit, Christopher Prewitt, and Kyle Achors for excellent technical assistance; Dr. Sameer Varma (University of South Florida) for help with PyMOL; and Dr. Steven Brill (Rutgers University) for plasmid pRS415-SGS1. This work was supported by National Institutes of Health Grant R01GM081425 (to K.H.S.).

- Ellis NA, Sander M, Harris CC, Bohr VA (2008) Bloom's syndrome workshop focuses on the functional specificities of RecQ helicases. *Mech Ageing Dev* 129(11):681–691.
- Bloom D (1954) Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs; probably a syndrome entity. *AMA Am J Dis Child* 88(6):754–758.
- German J, Ellis NA (1998) Bloom syndrome. *The Genetic Basis of Human Cancer*, eds Vogelstein B, Kinzler KW (McGraw-Hill, New York), pp 301–315.
- German J, Passarge E (1989) Bloom's syndrome. XII. Report from the Registry for 1987. *Clin Genet* 35(1):57–69.
- Bennett RJ, Keck JL (2004) Structure and function of RecQ DNA helicases. *Crit Rev Biochem Mol Biol* 39(2):79–97.
- Ellis NA, et al. (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83(4):655–666.
- German J, Sanz MM, Ciocci S, Ye TZ, Ellis NA (2007) Syndrome-causing mutations of the *BLM* gene in persons in the Bloom's Syndrome Registry. *Hum Mutat* 28(8):743–753.
- Ellis NA, et al. (1998) The Ashkenazic Jewish Bloom syndrome mutation *blmAsh* is present in non-Jewish Americans of Spanish ancestry. *Am J Hum Genet* 63(6):1685–1693.
- Li L, Eng C, Desnick RJ, German J, Ellis NA (1998) Carrier frequency of the Bloom syndrome *blmAsh* mutation in the Ashkenazi Jewish population. *Mol Genet Metab* 64(4):286–290.
- Ellis NA, et al. (1994) Linkage disequilibrium between the FES, D15S127, and *BLM* loci in Ashkenazi Jews with Bloom syndrome. *Am J Hum Genet* 55(3):453–460.
- Straughen JE, et al. (1998) A rapid method for detecting the predominant Ashkenazi Jewish mutation in the Bloom's syndrome gene. *Hum Mutat* 11(2):175–178.
- Guo RB, et al. (2007) Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. *Nucleic Acids Res* 35(18):6297–6310.
- Kitano K, Kim SY, Hakoshima T (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* 18(2):177–187.
- Pike AC, et al. (2009) Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci USA* 106(4):1039–1044.
- Hickson ID (2003) RecQ helicases: Caretakers of the genome. *Nat Rev Cancer* 3(3):169–178.
- Wu L, Davies SL, Levitt NC, Hickson ID (2001) Potential role for the *BLM* helicase in recombinational repair via a conserved interaction with RAD51. *J Biol Chem* 276(22):19375–19381.
- Wang Y, et al. (2000) BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* 14(8):927–939.
- Chaganti RS, Schonberg S, German J (1974) A manifold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci USA* 71(11):4508–4512.
- Groden J, Nakamura Y, German J (1990) Molecular evidence that homologous recombination occurs in proliferating human somatic cells. *Proc Natl Acad Sci USA* 87(11):4315–4319.
- Groden J, German J (1992) Bloom's syndrome. XVIII. Hypermutability at a tandem-repeat locus. *Hum Genet* 90(4):360–367.
- Lönn U, Lönn S, Nylén U, Winblad G, German J (1990) An abnormal profile of DNA replication intermediates in Bloom's syndrome. *Cancer Res* 50(11):3141–3145.
- Gruber SB, et al. (2002) *BLM* heterozygosity and the risk of colorectal cancer. *Science* 297(5589):2013.
- Cleary SP, et al. (2003) Heterozygosity for the *BLM*(Ash) mutation and cancer risk. *Cancer Res* 63(8):1769–1771.
- Broberg K, et al. (2009) Association between polymorphisms in RMI1, TOP3A, and *BLM* and risk of cancer, a case-control study. *BMC Cancer* 9:140.
- Ding SL, et al. (2009) Genetic variants of *BLM* interact with *RAD51* to increase breast cancer susceptibility. *Carcinogenesis* 30(1):43–49.
- Mullen JR, Kaliraman V, Brill SJ (2000) Bipartite structure of the *SGS1* DNA helicase in *Saccharomyces cerevisiae*. *Genetics* 154(3):1101–1114.
- Mirzaei H, Syed S, Kennedy J, Schmidt KH (2011) Sgs1 truncations induce genome rearrangements but suppress detrimental effects of *BLM* overexpression in *Saccharomyces cerevisiae*. *J Mol Biol* 405(4):877–891.
- Adzhubei IA, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248–249.
- Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.
- Korolev S, Yao N, Lohman TM, Weber PC, Waksman G (1998) Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicases. *Protein Sci* 7(3):605–610.
- Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97(1):75–84.
- Zittel MC, Keck JL (2005) Coupling DNA-binding and ATP hydrolysis in *Escherichia coli* RecQ: Role of a highly conserved aromatic-rich sequence. *Nucleic Acids Res* 33(22):6982–6991.
- Dutertre S, et al. (2000) Cell cycle regulation of the endogenous wild type Bloom's syndrome DNA helicase. *Oncogene* 19(23):2731–2738.
- Kim JL, et al. (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: The crystal structure provides insights into the mode of unwinding. *Structure* 6(1):89–100.
- Korolev S, Hsieh J, Gauss GH, Lohman TM, Waksman G (1997) Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* 90(4):635–647.
- Dillingham MS, Soultanas P, Wigley DB (1999) Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res* 27(16):3310–3317.
- Dillingham MS, Soultanas P, Wiley P, Webb MR, Wigley DB (2001) Defining the roles of individual residues in the single-stranded DNA binding site of PcrA helicase. *Proc Natl Acad Sci USA* 98(15):8381–8387.
- Soultanas P, Dillingham MS, Velankar SS, Wigley DB (1999) DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J Mol Biol* 290(1):137–148.
- Killoran MP, Keck JL (2006) Sit down, relax and unwind: Structural insights into RecQ helicase mechanisms. *Nucleic Acids Res* 34(15):4098–4105.
- Wirtenberger M, et al. (2006) Interaction of Werner and Bloom syndrome genes with p53 in familial breast cancer. *Carcinogenesis* 27(8):1655–1660.
- Goss KH, et al. (2002) Enhanced tumor formation in mice heterozygous for *Blm* mutation. *Science* 297(5589):2051–2053.
- Frank B, et al. (2010) Colorectal cancer and polymorphisms in DNA repair genes WRN, RMI1 and *BLM*. *Carcinogenesis* 31(3):442–445.
- German J (1997) Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet* 93(1):100–106.
- Gietz RD, Woods RA (2006) Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol Biol* 313:107–120.