Saccharomyces cerevisiae Sen1 Helicase Domain Exhibits 5′- to 3′-Helicase Activity with a Preference for Translocation on DNA Rather than RNA*•

Stephen Martin-Tumasz and David A. Brow

From the Department of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin 53706

Received for publication, June 22, 2015, and in revised form, July 19, 2015 Published, JBC Papers in Press, July 20, 2015, DOI 10.1074/jbc.M115.674002

In the yeast Saccharomyces cerevisiae, the essential nuclear helicase Sen1 is required for efficient termination of transcription of short noncoding RNA genes by RNA polymerase II. However, the mechanism by which Sen1 promotes transcription termination is not known. Prior biochemical studies on the Sen1 homolog from Schizosaccharomyces pombe showed that it can bind and unwind both DNA and RNA, but the S. pombe protein is not essential and has not been demonstrated to function in transcription. Furthermore, Sen1 from either yeast has not previously been expressed as a recombinant protein, due to its large molecular mass (252 kDa in S. cerevisiae). Here, we report the purification and characterization of the 89-kDa S. cerevisiae Sen1 helicase domain (Sen1-HD) produced in Escherichia coli. Sen1-HD binds single-stranded RNA and DNA with similar affinity in the absence of ATP, but it binds RNA more stably than DNA in the presence of ATP, apparently due to a slower translocation rate on RNA. Translocation occurs in the 5′ to 3′ direction, as for the S. pombe protein. When purified from E. coli at a moderate salt concentration, Sen1-HD was associated with short RNAs that are enriched for the trinucleotide repeat (CAN)₃. We propose that Sen1 binds to RNAs and prevents their stable pairing with DNA, consistent with in vivo studies by others showing increased R-loop (RNA/DNA hybrid) formation when Sen1 activity is impaired by mutations. Our results are consistent with a model in which Sen1 promotes transcription termination by resolving R-loops.

The Saccharomyces cerevisiae SEN1 gene was first identified in a screen for mutations that inhibit pre-tRNA splicing; SEN stands for splicing endonuclease (1, 2). Later, the SEN1 and NRD1 genes were identified in a selection for mutations that cause read-through of an RNA polymerase II (pol II)² terminator in the antisense strand of the U6 RNA gene (3, 4). Sen1 function was subsequently found to be necessary for efficient termination of a variety of short pol II transcripts (5, 6). The pre-tRNA splicing defect of the sen1-1 mutant can be explained by decreased expression of the SEN2 gene due to read-through of an upstream small nucleolar RNA gene terminator, but Sen1 may participate in other RNA processing pathways (7–9) and in genome stability (9–11), in addition to pol II termination.

Sen1 belongs to the Upf1-like superfamily 1 helicases (12), which include the eukaryotic nonsense-mediated decay factor Upf1 (13) and human IGHMBP2, which appears to be involved in translation (14). These enzymes exhibit 5′- to 3′-helicase activity and act on both DNA and RNA duplexes with 5′ single-stranded tails. Mutations in the human SETX gene, which encodes the apparent ortholog of Sen1, called senataxin (Fig. 1A), are associated with the neurodegenerative disorders ataxia with oculomotor apraxia type 2 (AOA2) and ataxia and myotrophic lateral sclerosis type 4 (ALS4) (15, 16). Both AOA2 and ALS4 exhibit onset in the mid-teen years and slow progression of cerebellar ataxia. In contrast, mutations in the human IGHMBP2 gene are associated with distal spinal muscular atrophy type 1, which has an early childhood onset and results in rapid paralysis of the diaphragm and ensuing respiratory distress (17). The mechanisms by which defects in these two helicases cause degeneration of distinct populations of neurons are unknown.

Characterization of the biochemical properties of Sen1 will clarify its in vivo substrates and activities. Previously, a 5′- to 3′-DNA and RNA helicase activity purified from Schizosaccharomyces pombe cell extract was attributed to an ortholog of S. cerevisiae Sen1 (18). More recently, however, TAP-tagged Sen1 purified from S. cerevisiae exhibited no DNA/RNA duplex unwinding activity and did not stably bind RNA, despite having DNA- and RNA-dependent ATPase activity (19). Furthermore, S. cerevisiae Sen1 appears to be maintained at a low cellular

* This work was supported by National Institutes of Health Grant RO1 GM082956 (to D. A. B.). The authors declare that they have no conflicts of interest with the contents of this article.

1 To whom correspondence should be addressed: Dept. of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health, 420 Henry Mall, Madison, WI 53706. Tel.: 608-262-1475; Fax: 608-262-5253; E-mail: dabrow@wisc.edu.

2 The abbreviations used are: pol II, polymerase II; SUMO, small ubiquitin-like modifier; 6-FAM, 6-carboxyfluorescein; ss, single-stranded.
Substrate Binding and Activity of the Sen1 Helicase Domain

level by targeted proteolysis (20); thus purification of the native protein from yeast is difficult. To facilitate biochemical characterization of Sen1’s helicase activity, we sought to overexpress the functional Sen1 helicase domain in E. coli. Here, we report the purification and characterization of recombinant S. cerevisiae Sen1 helicase domain (Sen1-HD). Sen1-HD binds single-stranded RNA and DNA with similar affinity and, in the presence of ATP, translocates on both in a 5’ to 3’ direction. However, it translocates more efficiently on DNA than RNA. When overexpressed in E. coli, Sen1-HD binds heterogeneous short RNAs, a significant fraction of which contain the trinucleotide repeat (CAN)₄ and derive predominantly from the antisense strand of protein-coding genes. The in vitro activities of the Sen1-HD are consistent with Sen1’s proposed in vivo function on R-loops (RNA/DNA hybrids) (10, 11, 21), but the activity of the Sen1-HD may be modified in vivo by its flanking domains and by extrinsic factors.

Experimental Procedures

Plasmid Construction—DNA encoding Sen1 residues 1095–1876, here referred to as the helicase domain (Sen1-HD), was amplified by PCR from the plasmid YEp351SEN1C (2) using an upstream primer with an NheI restriction site and a downstream restriction site and cloned into pET28a-His6-Smt3 (22), creating pET28a-His₆-Smt3-Sen1-HD.

Protein Expression and Purification—pET28a-His₆-Smt3-Sen1-HD was transformed into Rosetta2 (DE3)-competent E. coli Rosetta strain, protein of the correct molecular mass was expressed but was insoluble. To improve protein solubility, the Sen1-HD was fused to yeast SUMO (Smt3) as follows. The Sen1-HD was amplified by PCR from pET21b-Sen1-HD, adding an upstream BglII restriction site and downstream by BsaI and SalI restriction site. This fragment was cloned into pET28a-His₆-Smt3 (22), creating pET28a-His₆-Smt3-Sen1-HD.

Production of Rabbit Antiserum against Sen1-HD—Antiserum was raised in two New Zealand White rabbits by Harlan Bioproducts for Science, Inc. (Madison, WI), using their standard 112-day protocol. Initial immunization was done with 0.75 mg of Sen1-HD per rabbit in PBS + 2 mM MgCl₂ and 10 mM imidazole. Booster immunization at 4, 8, and 12 weeks after initial immunization used 0.1 mg of protein per rabbit. See Ref. 24 for use of the antiserum for Western blotting.

Nucleic Acid Complexes—Oligonucleotides used were purchased from Integrated DNA Technologies (IDT; see Table 1), except for the 84-nucleotide RNA, which was prepared by in vitro transcription as follows. A “gBlock” (IDT) containing the 84-nucleotide sequence flanked upstream by the T7 promoter and a BamHI restriction site and downstream by Bsal and BamHI sites was cloned into the BamHI site of pUC19. The plasmid was amplified in E. coli DH5α cells and isolated using a GeneJet plasmid Maxiprep kit (Thermo Scientific). After digestion with Bsal and phenol/chloroform extraction, the plasmid was transcribed in vitro using T7 RNA polymerase (a kind gift from Samuel Butcher), and the RNA was purified on an 8% polyacrylamide gel with 8 M urea. RNA was identified by UV shadowing, soaked out of the gel slice into 0.3 M sodium acetate, and precipitated with ethanol. The RNA was further purified by phenol/chloroform extraction and stored at −20 °C in 10 mM Tris-Cl, pH 7, 0.1 mM EDTA (TE). RNA concentration was determined by Nanodrop, using the estimated extinction coefficient at 260 nm (773,700 M⁻¹ cm⁻¹, based on the Oligo-Analyzer tool from IDT).

The N-terminal His₆ tag and SUMO protein were removed by incubating His₆-SUMO-Sen1-HD with SUMO protease for 30–60 min at room temperature in Wash Buffer B. SUMO protease was expressed from a His₆-SUMO protease expression vector (22) and purified as described above, except it was dialyzed into Wash Buffer B + 50% (v/v) glycerol and stored at −20 °C. The cleaved His₆-SUMO fragment and His₆-SUMO protease were removed by passing the digestion reaction over a HisPur cobalt spin column and retaining the flow-through (containing Sen1-HD).

Final purification was achieved using an AKTA Pure FPLC system (GE Healthcare). All steps were performed in Sen1 Buffer (50 mM NaPO₄, pH 6, 10 mM imidazole, 1 mM dithiothreitol, 2 mM MgCl₂, 5% (v/v) glycerol) with varying concentrations of NaCl. The protein was fractionated first by size, using a Superdex 200 10/300 GL gel filtration column run in Sen1 buffer + 1 M NaCl. Fractions containing Sen1-HD were pooled and then further purified using a 5-ml HiTrap heparin HP column. Sen1-HD was loaded onto the column in Sen1 buffer + 50 mM NaCl and then eluted with a linear gradient of 50 mM to 2 M NaCl in Sen1 buffer. Fractions containing Sen1-HD were pooled and dialyzed overnight at 4 °C into 100 volumes of Sen1 buffer + 50 mM NaCl, using a 20,000 molecular weight cutoff Slide-A-Lyzer dialysis cassette.

Purity was judged using SDS-PAGE (4–15% acrylamide gradient) and Coomassie staining. Protein concentration was determined by Nanodrop (Thermo Scientific), using the estimated extinction coefficient at 280 nm (70710 M⁻¹ cm⁻¹, based on the ExPASy ProtParam tool). The best yields obtained were about 0.5 mg/liter of culture.

The purity of the purified protein was confirmed by Coomassie staining and mass spectrometry (ESI-MS) measurement.
Duplexes (see Table 2 and Fig. 3) were prepared by heat annealing 10-μl reactions containing 25 μM short oligonucleotide and 50 μM long oligonucleotide in either water or TE. Reactions were boiled for 2 min and allowed to cool slowly to 4 °C over ~3 h. Reactions were then diluted to a final duplex concentration of 1 μM by addition of cold TE and stored at −20 °C.

Trimeric complexes (see Table 2 and Fig. 3) were prepared by heat annealing 10-μl reactions as described above, except containing 20 μM long strand and 37.5 μM each short strand in TE supplemented with 25 mM EDTA. After annealing, glycerol was added to a final concentration of 10% (v/v), and the hybrids were resolved by 12% native PAGE in TBE (50 mM Tris base, 50 mM boric acid, 1.25 mM EDTA) at room temperature. The trimer band was excised from the gel and transferred to a tube containing 1 ml of TBE + 100 mM NaCl. Nucleic acid was soaked out of the gel slice overnight at 4 °C with rotation and then stored at ~20 °C.

Biochemical Assays—Binding assays were performed in Binding Assay Buffer (42 mM NaPO₄, pH 6, 42 mM NaCl, 4.7 mM MgCl₂, 8% (v/v) glycerol, 3.4 mM imidazole, 0.3 mM DTT) in a 25-μl volume. Reactions contained 1 nM (trimer) or 20 nM (monomer or dimer) nucleic acid and varying amounts of Sen1-HD (monomer or dimer) nucleic acid and 600 nM DNA trap (reverse complement to the labeled strand or strands). ATP and Sen1-HD concentrations were optimized experimentally. Reactions were preincubated for 5 min at 30 °C (or 100 °C for samples marked “Boil”) and then incubated with either 0 or 250 nM Sen1-HD for up to 30 min at 30 °C. Samples without protein were incubated for 30 min at 30 °C. All reactions were then incubated for 10 min at 30 °C following the addition of 12 μl of Stop Buffer (100 mM EDTA, 2.5% SDS, 1 mg/ml proteinase K).

Both assays were analyzed using native PAGE as described above. All helicase and most binding assays were analyzed on 12% gels, whereas binding assays performed on trimeric substrates were analyzed on 6% gels. The gels were imaged on a Typhoon FLA 9000 scanner (GE Healthcare).

Gel Analysis—Gel images were quantified with ImageJ (National Institutes of Health). After cropping, the built-in ImageJ background subtraction tool was used to subtract most of the background (light background, 50 pixel rolling ball radius), and the appropriate window and level was identified using ImageJ’s autosetting. The same size box was used to measure the intensity of all free and bound bands, along with a single representative background measurement. The integrated density values were transferred into Excel (Microsoft), and the fractional presence of a given species was calculated as shown in Equation 1,

$$F_n = \frac{I_n - B}{I_n + U_n - 2B - F_0}$$

where $F$ is the fraction of nucleic acid present in the species of interest (free or bound); $I_n$ is the integrated density from the band of interest; $U_n$ is the integrated density from the other band; $B$ is the background integrated density; and $n$ is the time point or titration point. This was calculated as fraction free for helicase assays and fraction bound for binding assays.

Apparent dissociation constants were calculated by averaging the observed fraction bound from 6-FAM and Cy5 channels for each trimer at each titration point. These data were fit to the standard one-site binding Equation 2 in GraphPad Prism.
Substrate Binding and Activity of the Sen1 Helicase Domain

Results

Expression of Recombinant Sen1-HD in E. coli—Previous in vivo deletion studies showed that the minimal essential region of Sen1 is between residues 1089 and 1907, which corresponds closely to the helicase domain (Fig. 1A) (24). Based on alignment of the sequence of S. cerevisiae Sen1 with senataxin and its paralogs Upf1 and IGHMBP2 (Fig. 2), we chose to express in E. coli a fragment of the SEN1 gene encoding residues 1095–1876. This fragment encodes about 50 residues N-terminal to the predicted beginning of the first RecA domain and about seven residues C-terminal to the end of the second RecA domain (Figs. 1A and 2). To obtain soluble protein, we found it necessary to fuse the Sen1 fragment to yeast SUMO, which was subsequently cleaved off, along with a His6 tag, during purification. A major fraction of the soluble Sen1-HD purified as an apparent monomer on a gel filtration column (Fig. 1, B and C).

As a negative control, we sought to express a mutant form of the Sen1-HD deficient in helicase activity. We attempted expression of four different mutant alleles in E. coli as follows: K1363R in motif 1; T1779P in motif 5; K1788E in motif 5; and T1819A in motif 6 (Fig. 2). The sen1-K1363R mutation is recessive lethal in yeast, whereas the remaining alleles result in heat-sensitive growth and defects in pol II termination (24). 3 T1779P and K1788E correspond to the T2373P and K2382E AOA2 disease mutations in senataxin (Fig. 1A). None of the mutant expression constructs yielded significant amounts of protein in E. coli. It seems unlikely that all four substitutions would result in unfolded proteins in E. coli. An alternative possibility is that helicase function is required to stabilize the Sen1-HD in E. coli (see “Discussion”).

Recombinant Sen1-HD Binds Strongly to Single-stranded DNA and RNA—Sen1-HD’s binding affinity for DNA/DNA, DNA/RNA, and RNA/RNA trimers with a 28-nucleotide single-stranded region (Fig. 3A and Table 2) was investigated using an electrophoretic mobility shift binding assay (Fig. 3B). In the absence of ATP, Sen1-HD bound strongly to all three trimers

\[
F_n = \frac{F_{\max} X_n}{K_d + X_n} \quad \text{(Eq. 2)}
\]

where \(F_n\) is the fraction bound at titration point \(n\); \(F_{\max}\) is the maximum fraction bound, and \(X_n\) is [Sen1-HD] at titration point \(n\). Results are reported as calculated \(K_d\) ± the reported standard error.

Microarray Analysis of Sen1-HD-associated RNA—Bound RNA was purified from Sen1-HD eluted from a sulfopropyl-Sepharose column by extraction with an equal volume of phenol/chloroform/isoamyl alcohol, 25:24:1, pH 6.6 (Ambion), and concentrated by ethanol precipitation. First-strand cDNA was synthesized from ~10 μg of Sen1-associated RNA using GoScript reverse transcriptase (Promega) and random hexamer primers in the presence of 4 mM amino-allyl dUTP, 8 μg/ml actinomycin D, and 0.5 units/μl SUPERase-In ribonuclease inhibitor (Life Technologies, Inc.). After hydrolysis with 0.25 M NaOH for 30 min at 65 °C and neutralization with HCl, the cDNA was purified with a MERmaid SPIN kit (MP Biomedicals). The cDNA was coupled to 4.5 μl of Cy3 mono-reactive dye (GE Healthcare) according to the manufacturer’s protocol and purified with a MERmaid kit. The cDNA was hybridized to a custom E. coli K12 MG1655 tiled genome microarray (Roche NimbleGen) and scanned as described previously (23).

3 X. Chen and D. A. Brow, unpublished data.
tested, with apparent dissociation constants between 50 and 75 nM (Fig. 3C). There was no apparent binding preference for DNA versus RNA under these conditions. At lower concentrations of Sen1-HD (100 nM), there was a single-shifted band, likely representing a 1:1 Sen1-HD trimer complex. However, as the concentration of Sen1-HD increased additional higher order bands appear, suggesting that Sen1-HD forms multimers on the nucleic acid substrate. At the highest concentrations (1 M), almost all of the nucleic acid appeared to be in or near the well. This result suggests that at high concentrations multiple Sen1-HD molecules bind to a trimer or that aggregation of the Sen1-HD prevents the nucleic acid from entering the gel.

FIGURE 2. Manual alignment of the amino acid sequences of the helicase domains from S. cerevisiae (S.c.) Upf1, human IGHMBP2, senataxin, and S. cerevisiae Sen1. Residues that are identical or similar between at least two proteins are highlighted. The two RecA domains (1A and 2A) and two insertion domains (1B and 1C) are indicated above the sequences, as are the canonical helicase motifs. Underlined residues in Sen1 are the sites of mutations tested for expression in E. coli (see text).

The helicase domain of the Sen1 paralog Upf1 bound ~10 nucleotides of single-stranded RNA (2S). To determine the optimal length of single-stranded nucleic acid for Sen1-HD binding, we used five different dimer DNA constructs with 5 single-stranded tails ranging from 0 to 28 nucleotides long (Fig. 3A and Table 2). Sen1-HD exhibited lower affinity for all five dimers tested than for the trimer constructs (Fig. 3D). However, binding was significantly enhanced by the presence of a 10-nucleotide ssDNA tail, and it was strongest with the 15- and 28-nucleotide tails (compare 360 nM lanes in Fig. 3D). Thus, it appears that the Sen1 and Upf1 helicase domains contact a similar length of single-stranded nucleic acid. In addition, Sen1
may have a preference for a single-stranded region flanked by duplexes.

**Sen1-HD Possesses 5' to 3' DNA and RNA Translocase Activity**—To identify the presence and directionality of Sen1-HD’s helicase activity, we used the trimeric nucleic acid substrates, which have distinct fluorescent labels on the two short strands. Based on its binding to tailed duplexes (Fig. 3D), Sen1-HD likely binds predominantly in the single-stranded region of the trimer. Thus, 3' to 5' translocase activity would result in the displacement of the 6-FAM-labeled short strand, whereas 5' to 3' activity would result in the displacement of the Cy5-labeled short strand (Fig. 3A). Incubation of Sen1-HD with all the DNA trimeric substrates results in a decrease in the trimer band both in the 6-FAM and in the Cy5 channels (Fig. 4A). In the 6-FAM channel, a new band appears immediately below the trimer, corresponding to the long strand bound only to the 6-FAM oligonucleotide (because the band does not appear in the Cy5 channel). In the Cy5 channel, the band corresponding to the Cy5-labeled oligonucleotide paired to the trap strand, which is predominant in the boiled sample, increases in intensity. These results demonstrate 5' to 3' translocation on the long strand, displacing the Cy5-labeled short strand.

Examining the results for all four combinations of DNA and RNA for the short and long strands, there is evidence for 5' to 3' translocation along both DNA and RNA and no evidence for any 3' to 5' translocation (Fig. 4B). Sen1-HD’s helicase activity appears strongest when translocating along DNA to displace DNA (blue line in Fig. 4B) and is robust but slightly weaker when translocating along DNA to displace RNA (red line in Fig. 4B). When translocating along RNA, Sen1-HD’s activity becomes both weaker and more variable. Sen1-HD can translocate along the RNA and displace DNA; the average activity is somewhat lower than along the DNA/RNA trimer, but the final fraction unwound ranges from as low as 0.11 to as high as 0.32 (green line in Fig. 4B). Finally, Sen1-HD displays a very low ability to translocate along RNA and displace RNA (purple line in Fig. 4B).

**Effect of ATP on Sen1-HD Binding to ssDNA and ssRNA Is Consistent with Slower Translocation on RNA**—Helicase activity requires both binding to and translocation along the substrate nucleic acid. Our binding experiments above showed that Sen1-HD has a very similar affinity for DNA and RNA (Fig. 3B), but its helicase activity is greater when translocating along DNA. This result suggests that Sen1-HD may be able to translocate more efficiently along DNA than RNA. To test this proposal, we used a modified version of the binding assay. We incubated Sen1-HD with either 28-nucleotide single-stranded DNA or RNA in the presence of ATP, which should allow Sen1-HD to translocate off the oligonucleotides, resulting in decreased occupancy.

In the absence of ATP, we observed nearly complete binding to both ssDNA and ssRNA at the lowest Sen1-HD concentration (Fig. 5). This result is consistent with what we observed in
binding assays using the trimeric complexes. In the presence of ATP, unbound DNA and RNA persist even at relatively high Sen1-HD concentrations, consistent with release of the Sen1-HD from the nucleic acid via translocation off the end. As predicted, the destabilizing effect of ATP is greater on ssDNA than ssRNA (Fig. 5), consistent with the proposal that Sen1-HD translocates more efficiently along ssDNA than ssRNA. We infer that the lower helicase activity of Sen1-HD when displacing a nucleic acid strand from RNA rather than DNA is due to a decreased efficiency of translocation.

Sen1-HD Purified from E. coli at Moderate Salt Concentration Is Bound to Short RNAs—Our initial purification scheme for Sen1-HD used 0.3M (rather than 1M) NaCl buffer for lysis and metal-affinity purification, followed by gel filtration chromatography in 50 mM (rather than 1M) NaCl and cation exchange chromatography using sulfopropyl-Sepharose rather than heparin-Sepharose. Under these conditions, the purified Sen1-HD had significant bound nucleic acid, as evidenced by broad peaks on the gel filtration and ion exchange columns, and high absorbance of 260 nm light relative to 280 nm light. The presence of nucleic acid was confirmed by phenol/chloroform extraction of the sample and detection by agarose gel electrophoresis and ethidium bromide staining (Fig. 6A, input). On the nondenaturing gel, the nucleic acid co-migrated with DNA fragments in the 100–300-bp size range. Incubation of the nucleic acid at 37 °C for 30 min with or without RNase-free DNase had no apparent effect on the nucleic acid (Fig. 6A, no add. and DNase). However, incubation with RNase T1 resulted in elimination of the staining material (Fig. 6A, RNase). Thus, under these conditions Sen1 co-purifies with short RNA species from E. coli cell lysates.

We reasoned that the identity of the Sen1-HD-associated RNAs might provide a clue to the sequence specificity, if any, of RNA binding by Sen1-HD. To determine the sequence of these small RNAs, we used them as template for synthesis of first-strand Cy3-labeled cDNA and hybridized the cDNA to a tiled microarray of the E. coli genome (23). After manual filtering of spurious signals, we used MEME to search for common sequence patterns (31). We analyzed three different groups of probes as follows: all probes at or above the 99.8th percentile for signal intensity; all pairs of adjacent or overlapping probes at or above the 98th percentile; and all triplets of adjacent or overlapping probes at or above the 95th percentile. In all cases, the most common motif (present in about a quarter of the probes) is the trinucleotide CAN repeated four times, where N is any nucleotide but most often C (Fig. 6B). Thus, Sen1 appears to bind RNA in vivo and exhibits some degree of selectivity for the sequence (CAN)₄.

Discussion

Activity of the Recombinant Sen1 Helicase Domain—Here, we present the first biochemical characterization of recombinant Sen1-HD. Consistent with the reported activity of native Sen1 purified from S. pombe (18), S. cerevisiae Sen1-HD has 5’-to 3’-helicase activity (Fig. 4). The S. pombe study reported 5’-to 3’-helicase activity from both a high molecular mass (181 kDa) form of Sen1 and a 95-kDa presumptive proteolytic cleavage product of the full-length enzyme (18), which likely corresponds closely to our 89-kDa Sen1-HD protein. This finding suggests that the protein domains flanking the Sen1-HD do not significantly modify its helicase activity, at least in the absence of other cellular factors. It is therefore surprising that full-
length Sen1 purified from *S. cerevisiae* had no detectable helicase activity (19). However, the helicase assay for native *S. cerevisiae* Sen1 used 1 mM ATP and an RNA/DNA duplex with a 5’ “tail” of single-stranded RNA. Using a similar substrate, we observed maximal helicase activity between 8 and 16 mM ATP (data not shown). Thus, the native *S. cerevisiae* Sen1 may have been well below its optimal ATP concentration. Indeed, *S. pombe* 181-kDa Sen1 was reported to “hardly support” unwinding of an all-RNA substrate at 0.5 mM ATP but to efficiently unwind an all-DNA substrate under the same conditions (18).

In addition, the concentration of native Sen1 used was apparently 50 nM, and we obtained optimal activity between 250 and 500 nM Sen1-HD (data not shown). Therefore, the apparent lack of helicase activity using tagged, full-length Sen1 purified from *S. cerevisiae* cells (19) may be a result of nonoptimal assay conditions.

We observed a marked preference of the Sen1-HD for translocation on DNA versus RNA (Fig. 4). This preference could be due to the absence of the flanking protein domains, because studies on *S. pombe* Sen1 showed about 3-fold lower activity of the 95-kDa HD-like form compared with the 181-kDa form on an RNA duplex substrate but similar activities on a DNA duplex substrate (18). The preference of Sen1-HD for translocation on DNA was also reflected in the greater destabilizing effect of ATP on binding of Sen1-HD to ssDNA than to ssRNA (Fig. 5). Interestingly, yeast Upf1, which has much shorter helicase-flanking domains than Sen1, exhibits a similar differential binding behavior in the presence of ATP (13).

Substrate Binding and Activity of the Sen1 Helicase Domain
HD. Although we have not examined the in vivo origin of the Sen1-HD-associated RNAs in detail, we note that of the top 21 matches to the (CAN)₄ consensus, 18 correspond to the anti-sense strand of annotated protein-coding genes and only three correspond to the sense strand. This skewed distribution is intriguing, because one function of Sen1 in S. cerevisiae is silencing of genes via cotranscriptional association with their antisense transcripts (6, 26, 27). Conceivably, engagement of Sen1-HD with antisense transcripts in E. coli protects it from degradation, whereas the mutant versions of Sen1-HD that expressed poorly are not functionally engaged and therefore are degraded. A study that mapped the sites of cross-linking of Sen1 to RNA in live yeast cells reported that no conserved motifs were identified (28), so it is as yet unknown whether the (CAN)₄ motif may contribute to Sen1 function in yeast.

Conclusions and Prospects—The results described here show that, despite the lack of helicase activity observed for affinity-purified native S. cerevisiae Sen1, purified recombinant S. cerevisiae Sen1-HD has helicase activities typical of other Upf1-like SF1 helicases, with the possible exception of a higher $K_m$ value for ATP. These activities include the ability to translocate on DNA as well as RNA. Although the natural targets of Upf1 and IGHMBP2 appear to be RNA, nuclear Sen1 and its human ortholog senataxin may function on DNA/RNA hybrids (R-loops) at sites of transcription (10, 11, 21). Our ability to obtain reasonable quantities of apparently homogeneous recombinant Sen1-HD from E. coli opens up the possibility of structural analysis by x-ray crystallography. Although Sen1 is a paralog of Upf1 and IGHMBP2 (29, 30), its 1B and 1C insertion domains within the first RecA domain are dissimilar to those in Upf1 and IGHMBP2 (Fig. 2). Because the 1B and 1C domains are typically involved in substrate recognition, it will be of particular interest to determine the structure and function of these domains in Sen1.

Author Contributions—S. M. T. and D. A. B. designed the study and wrote the paper. S. M. T. carried out all experiments and analyzed the data. S. M. T. and D. A. B. interpreted the data and approved the final version of the manuscript.

Acknowledgments—We thank Elaine Brow and Amanda Carney for constructing pET21b-Sen1-HD; Jon Audhya for pET28a and SUMO protease; Sam Butcher for T7 RNA polymerase; Xin Chen for assisting with helicase assays; Kevin Myers, Nicole Ellis, and Tricia Kiley for assistance with the microarray analysis; the Dahlberg and Fox laboratories for use of their equipment; Jim Keck and Basu Bhattacharaya for helpful discussions and control helicase; and Eric Montemayor and Xin Chen for comments on the manuscript.

References

1. Winey, M., and Culbertson, M. R. (1988) Mutations affecting the tRNA-splicing endonuclease activity of Saccharomyces cerevisiae. Genetics 118, 609–617
2. DeMarini, D. J., Winey, M., Ursic, D., Webb, F., and Culbertson, M. R. (1992) SEN1, a positive effector of tRNA-splicing endonuclease in Saccha-
romyces cerevisiae. Mol. Cell. Biol. 12, 2154–2164
3. Steinmetz, E. J., and Brow, D. A. (1996) Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol. Cell. Biol. 16, 6993–7003
4. Steinmetz, E. J., and Brow, D. A. (1998) Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. Proc. Natl. Acad. Sci. U.S.A. 95, 6699–6704
5. Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3’-end formation of RNA polymerase II transcripts. Nature 413, 327–331
6. Steinmetz, E. J., Warren, C. L., Kuehner, J. N., Panbehi, B., Ansari, A. Z., and Brow, D. A. (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol. Cell 24, 735–746
7. Ursic, D., Himmel, K. L., Gurley, K. A., Webb, F., and Culbertson, M. R. (1997) The yeast SEN1 gene is required for the processing of diverse RNA classes. Nucleic Acids Res. 25, 4778–4785
8. Rasmussen, T. P., and Culbertson, M. R. (1998) The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nuclear RNAs in Saccharomyces cerevisiae. Mol. Cell. Biol. 18, 6885–6895
9. Ursic, D., Chinchilla, K., Finkel, J. S., and Culbertson, M. R. (2004) Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. Nucleic Acids Res. 32, 2441–2452
10. Mischo, H. E., Gómez-González, B., Grzegorczyk, P., Rondón, A. G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N. J. (2011) Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol. Cell 41, 21–32
11. Alzu, A., Bermejo, R., Béginis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., and Liberis, G. (2008) Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell 151, 835–846
12. Jankovsky, E. (2011) RNA helicases at work: binding and rearranging. Trends Biochem. Sci. 36, 19–29
13. Czaplinski, K., Weng, Y., Hagan, K. W., and Peltz, S. W. (1995) Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. RNA 1, 610–623
14. Guenther, U.-P., Handoko, L., Laggerbauer, B., Jablonka, S., Chari, A., Alzheimer, M., Ohmer, J., Plötter, O., Gehring, N., Sickmann, A., von Au, K., Schuelke, M., and Fischer, U. (2009) IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). Hum. Mol. Genet. 18, 1288–1300
15. Moreira, M. C., Klur, S., Watanabe, M., Németh, A. H., Le Ber, I., Moniz, J. C., Tranchant, C., Aubourg, P., Tazir, M., Schulz, L., Delannoy, M., Schulz, J. B., Puguot, J., Calvas, P., Shizuka-Ikeda, M., et al. (2004) Senat-
taxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-oculapara-
axia 2. Nat. Genet. 36, 225–227
16. Chen, Y. Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., Dierick, I., Abel, A., Benkner, M. L., Rabin, B. A., Nicholson, G. A., Auer-Grumbach, M., Wagner, K., De Jonghe, P., Griffin, J. W., et al. (2004) DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). Am. J. Hum. Genet. 74, 1128–1135
17. Grohmann, K., Schuelke, M., Diers, A., Hoffmann, K., Lucke, B., Adams, C., Bertini, E., Leonhardt-Horti, H., Muntoni, F., Ouvrier, R., Pfeifer, A., Rossi, R., Van Maldergem, L., Wilmshurst, J. M., Wiener, T. F., et al. (2001) Mutations in the gene encoding immunoglobulin μ-bonding pro-
tein 2 cause spinal muscular atrophy with respiratory distress type 1. Nat. Genet. 29, 75–77
18. Kim, H.-D., Choe, J., and Seo, Y.-S. (1999) The sen1 + gene of Schizosac-
charomyces pombe, a homolog of budding yeast SEN1, encodes an RNA and DNA helicase. Biochemistry 38, 14697–14710
19. Porrua, O., and Libri, D. (2013) A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. Nat. Struct. Mol. Biol. 20, 884–891
20. DeMarini, D. J., Papa, F. R., Swaminathan, S., Ursic, D., Rasmussen, T. P., Culbertson, M. R., and Hochstrasser, M. (1995) The yeast SEN3 gene encodes a regulatory subunit of the 26S proteasome complex required for ubiquitin-dependent protein degradation in vivo. Mol. Cell. Biol. 15, 6311–6321
21. Chan, Y. A., Aristizabal, M. J., Lu, P. Y., Luo, Z., Hamza, A., Kobot, M. S.,
Substrate Binding and Activity of the Sen1 Helicase Domain

Stirling, P. C., and Hieter, P. (2014) Genome-wide profiling of yeast DNA: RNA hybrid prone sites with DRIP-Chip. *PLoS Genet.* 10, e1004288

Lee, C.-D., Sun, H.-C., Hu, S.-M., Chiu, C.-F., Homhuan, A., Liang, S.-M., Leng, C.-H., and Wang, T.-F. (2008) An improved SUMO fusion protein system for effective production of native proteins. *Protein Sci.* 17, 1241–1248

Myers, K. S., Yan, H., Ong, I. M., Chung, D., Liang, K., Tran, F., Keleş, S., Landick, R., and Kiley, P. J. (2013) Genome-scale analysis of *E. coli* FNR reveals complex features of transcription factor binding. *PLoS Genet.* 9, e1003565

Chen, X., Müller, U., Sundling, K. E., and Brow, D. A. (2014) *Saccharomyces cerevisiae* Sen1 as a model for the study of mutations in human senataxin that elicit cerebellar ataxia. *Genetics* 198, 577–590

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011) Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* 41, 693–703

Castelnuovo, M., Rahman, S., Guffanti, E., Infantino, V., Stutz, F., and Zenklusen, D. (2013) Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat. Struct. Mol. Biol.* 20, 851–858

Castelnuovo, M., Zaug, J. B., Guffanti, E., Maffioletti, A., Camblong, J., Xu, Z., Clauder-Münster, S., Steinmetz, L. M., Luscombe, N. M., and Stutz, F. (2014) Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. *Nucleic Acids Res.* 42, 4348–4362

Creamer, T. J., Darby, M. M., Jamonnak, N., Schauhency, P., Hao, H., Wheelan, S. J., and Corden, J. L. (2011) Transcriptome-wide binding sites for components of the *Saccharomyces cerevisiae* non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS Genet.* 7, e1002329

Cheng, Z., Muhlrad, D., Lim, M. K., Parker, R., and Song, H. (2007) Structural and functional insights into the human Upf1 helicase core. *EMBO J.* 26, 253–264

Lim, S. C., Bowler, M. W., Lai, T. F., and Song, H. (2012) The Ighmbp2 helicase structure reveals the molecular basis for disease-causing mutations in DMSA1. *Nucleic Acids Res.* 40, 11009–110022

Bailey, T. L., and Elkan, C. (1994) *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp. 28–36, AAAI Press, Menlo Park, CA