

Werner Syndrome Protein Contains Three Structure-specific DNA Binding Domains*

Cayetano von Kobbe‡, Nicolas H. Thomä§¶, Bryan K. Czyzewski§, Nikola P. Pavletich§, and Vilhelm A. Bohr‡¶

From the ‡Laboratory of Molecular Gerontology, National Institute on Aging, National Institutes of Health, Baltimore, Maryland 21224 and the §Structural Biology Program and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Werner syndrome (WS) is a premature aging syndrome caused by mutations in the WS gene (*WRN*) and a deficiency in the function of the Werner protein (WRN). WRN is a multifunctional nuclear protein that catalyzes three DNA-dependent reactions: a 3′-5′-exonuclease, an ATPase, and a 3′-5′-helicase. Deficiency in *WRN* results in a cellular phenotype of genomic instability. The biochemical characteristics of WRN and the cellular phenotype of *WRN* mutants suggest that WRN plays an important role in DNA metabolic pathways such as recombination, transcription, replication, and repair. The catalytic activities of WRN have been extensively studied and are fairly well understood. However, much less is known about the domain-specific interactions between WRN and its DNA substrates. This study identifies and characterizes three distinct WRN DNA binding domains using recombinant truncated fragments of WRN and five DNA substrates (long forked duplex, blunt-ended duplex, single-stranded DNA, 5′-overhang duplex, and Holliday junction). Substrate-specific DNA binding activity was detected in three domains, one N-terminal and two different C-terminal WRN fragments (RecQ conserved domain and helicase RNase D conserved domain-containing domains). The substrate specificity of each DNA binding domain may indicate that each protein domain has a distinct biological function. The importance of these results is discussed with respect to proposed roles for WRN in distinct DNA metabolic pathways.

Werner syndrome (WS)¹ is a rare autosomal recessive disorder characterized by the early onset of aging symptoms, such as graying of the hair, cataracts, osteoporosis, atherosclerosis, type II diabetes mellitus, and high incidence of malignant neoplasm (1). The gene mutated in WS (*WRN*), encodes a nuclear 1432-amino acid protein (WRN) (2). WRN belongs to the

RecQ helicase family of proteins, which is a conserved group of proteins implicated in several aspects of DNA metabolism (2, 3). Three of the RecQ helicases, WRN, Bloom (BLM), and Rothmund-Thomson, are associated with heritable human diseases (3). WRN is a multifunctional protein with three DNA-dependent catalytic activities: 3′-5′-exonuclease (4, 5), ATPase, and 3′-5′-helicase (6) (Fig. 1A). WRN is the only member of the human RecQ family known to include an exonuclease domain.

Cells from WS patients display replication defects, genomic instability, and altered telomere dynamics, suggesting an important role for WRN in DNA metabolic pathways (for review, see Ref. 7). This is consistent with the observation that many proteins that interact with WRN are involved in recombination, replication, transcription, and telomere structure or repair (8–15). *In vitro*, WRN binds and is catalytically active toward non-canonical DNA structures such as recombination intermediates (16), replication forks (17), repair intermediates (13, 18), and telomeric ends (19). For example, WRN interacts with Holliday junctions, forked duplexes, 5′-overhang duplexes, and D-loops.

Typically, WS cells carry a C-terminal truncated WRN that lacks a nuclear localization signal (20). The C-terminal region of WRN also plays a central role in WRN protein-protein interactions (for review, see Ref. 7) and nucleolar targeting (30). The conserved RQC domain (RecQ conserved) includes the nucleolar targeting sequence (nuclear localization signal-dependent) and participates in WRN protein-protein interactions (9, 30). The C-terminal region of WRN also contains the conserved HRDC (helicase, RNase D Conserved) domain (2), whose function is still unclear, but which may play a role in DNA binding in *Saccharomyces cerevisiae* Sgs1 and human RecQ homologues (21).

While the enzymatic activities of WRN have been studied extensively (7), very little is known about the DNA binding properties of WRN. A more complete understanding of WRN DNA binding properties is essential to define the mechanism by which WRN recognizes, binds, and processes DNA substrates. Earlier biochemical studies on full-length WRN suggest that WRN binds DNA with low efficiency (22, 23).

This study examines the DNA binding properties of a series of truncated recombinant WRN proteins and addresses whether distinct functional domains of WRN display differential DNA binding properties that might be required in different stages of helicase and exonuclease reactions. These WRN variants were challenged with several WRN DNA substrates including long forked duplex, Holliday junction, 5′-overhang duplex, ssDNA, and blunt-ended duplex (Table I). The results demonstrate that WRN has three distinct structure-specific DNA binding domains that co-localize with the exonuclease, RQC, and HRDC regions of WRN, respectively. The importance

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Supported by a Human Frontier Science Program Long Term Fellowship.

¶ To whom correspondence should be addressed: Laboratory of Molecular Gerontology, NIA, NIH, 5600 Nathan Shock Dr., Baltimore, MD 21224. Tel.: 410-558-8162; Fax: 410-558-8157; E-mail: vbohr@nih.gov.

¹ The abbreviations used are: WS, Werner syndrome; WRN, Werner syndrome protein; RQC, RecQ conserved domain; HRDC, helicase, RNase D conserved domain; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; BLM, Bloom; GST, glutathione S-transferase; aa, amino acids; HJ, Holliday junction; ATPγS, adenosine 5′-O-(thiotriphosphate).

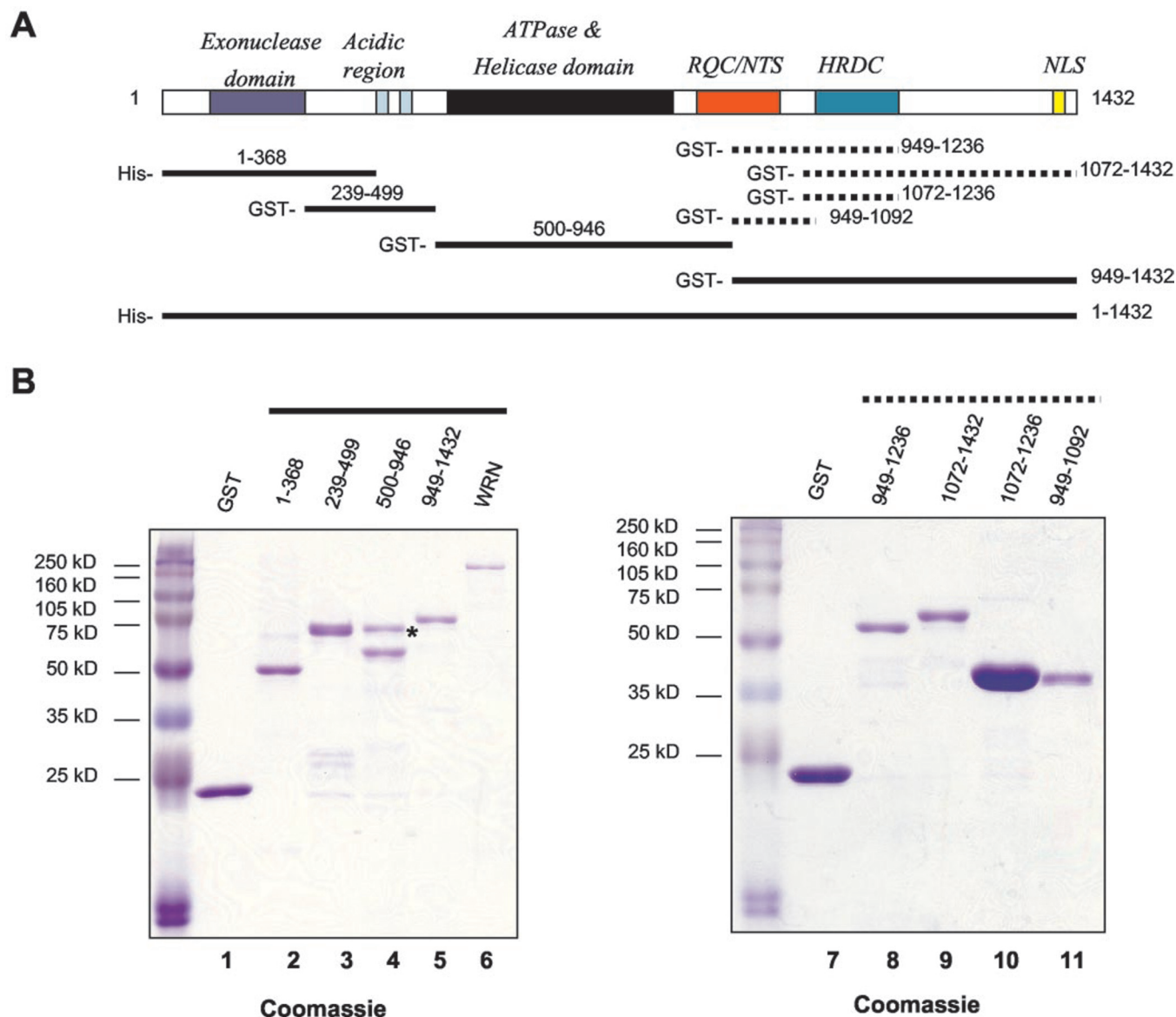


FIG. 1. *A*, functional domains of WRN and recombinant WRN fragments used in DNA binding studies. *Solid lines* are fragments that cover the entire WRN sequence. *Dashed lines* are C-terminal WRN fragments. *Numbers* indicate WRN amino acid coordinate. The abbreviations used are as follows: NTS, nucleolar targeting sequence. *B*, WRN fragments were purified and analyzed by SDS-PAGE and visualized by Coomassie Blue staining. Approximately 1 μ g of each recombinant fragment was loaded, except in lanes 7 and 10, where \sim 3 and \sim 6 μ g of each recombinant were loaded, respectively. *Numbers* indicate WRN amino acid coordinate. The asterisk indicates mobility of full-length GST-WRN helicase domain (aa 500–946). *C*, strategy to identify WRN DNA binding regions. *D*, limited proteolysis of WRN RQC and HRDC domains. Purified WRN RQC-HRDC (aa 950–1360) protein (lane 1) was treated with different concentrations of subtilisin (lanes 2–8) as described under “Materials and Methods.” The resulting fragments were separated by SDS-PAGE and stained with Coomassie Blue. GST alone is a leftover of the purification, and it serves as an internal control of protease resistant protein. *Asterisks* indicate intermediate products of the subtilisin digestion. *Black arrows* indicate the two protease-resistant fragments. The N-terminal sequence revealed that the 17- and 14-kDa fragments start at the amino acid 958 and either 1128 or 1148, respectively.

of this finding and its implications for WRN cellular functions are discussed.

MATERIALS AND METHODS

Recombinant Proteins—The cloning and expression of most of the GST-WRN fragments was described previously (9). Briefly, lysates from *Escherichia coli* BL21(DE3) pLysS expressing recombinant proteins were obtained by sonication of bacterial pellets in lysis buffer (2 \times phosphate-buffered saline and a mixture of protease inhibitors, Amersham Biosciences). After sonication, Triton X-100 was added to a final concentration of 0.1%. Lysates were then cleared by centrifugation at 35 K for 60 min at 4 $^{\circ}$ C in a Beckman Ti60 rotor. Glutathione-Sepharose beads were added to the lysates, and the mixture was rotated for 1 h 30 min at 4 $^{\circ}$ C. Beads were washed five times with lysis buffer plus Triton X-100 (0.1%), and the recombinant proteins were eluted with glutathione elution buffer (100 mM Tris-HCl, pH 8, 150 mM NaCl, 5% glycerol, 0.1% Triton X-100, and 10 mM glutathione). The purified

GST-WRN fragments were then analyzed by SDS-PAGE and visualized by Coomassie staining. Recombinant, hexahistidine-tagged WRN and the N-terminal domain of WRN (N-fragment, aa 1–368) were purified as described previously (24). GST-WRN RQC (aa 950–1360) was cloned into the NotI-KpnI restriction sites of pGEX-TEV vector and purified basically as described above.

DNA Substrates—The telomeric short and long forked, blunt-ended duplex, 5'-overhang duplex, ssDNA, and Holliday junction (HJ) substrates were described elsewhere (15, 16, 25).

Electrophoretic Mobility Shift Assay—Binding reactions (20 μ l) were conducted in standard binding buffer (40 mM Tris-HCl, pH 7.0, 1 mM EDTA, 20 mM NaCl, 8% glycerol, and 20 μ g/ml bovine serum albumin). Protein and DNA substrate concentrations are indicated in the figure legends. Reactions were incubated for 30 min on ice and then directly loaded on 5% nondenaturing polyacrylamide gels (37.5:1). Electrophoresis was carried out at a constant voltage of 12.8 V/cm at 4 $^{\circ}$ C in 1 \times TAE (40 mM Tris acetate, 1 mM EDTA, pH 8.0),

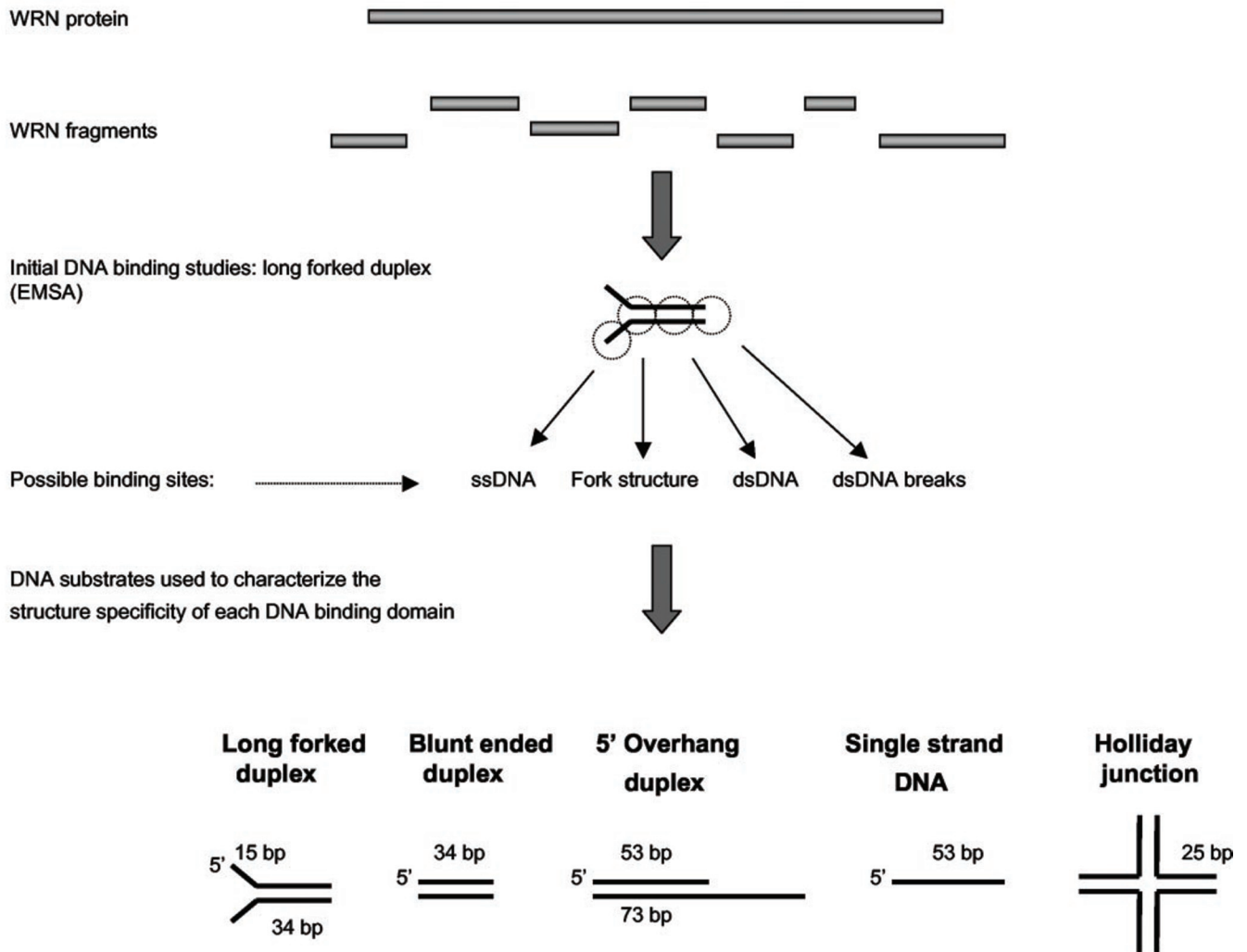
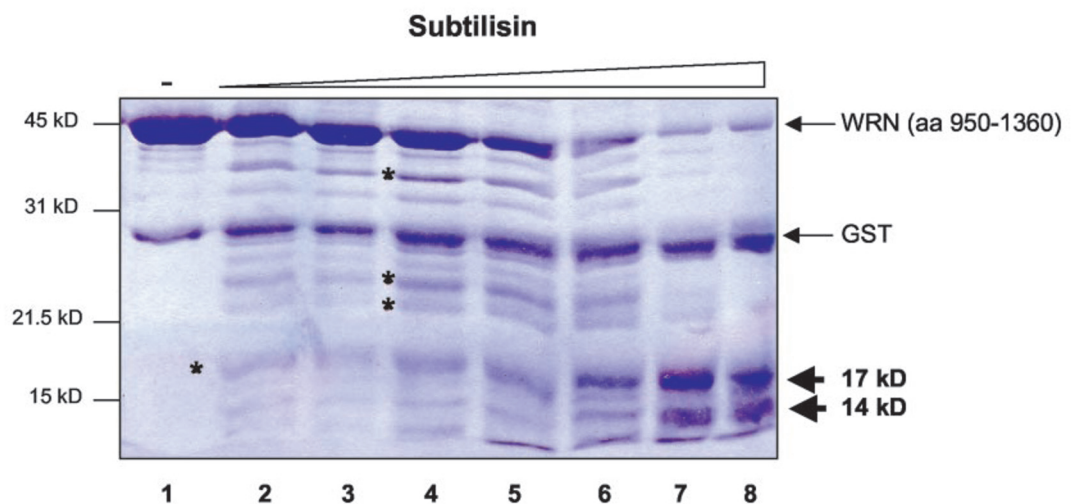
C**D**

FIG. 1—continued

for 2 h 20 min. Products were visualized using a PhosphorImager and ImageQuant software (Molecular Dynamics).

Helicase Reactions—Reactions were performed in binding buffer plus 4 mM MgCl₂ and 2 mM ATP, at 37 °C for 15 min. The DNA substrate (short forked duplex) and protein concentrations are indicated in the figure legend to Fig. 2B. Reactions were terminated by the addition of 3× stop dye (50 mM EDTA, 40% glycerol, 0.9% SDS,

0.1% bromophenol blue, and 0.1% xylene cyanol). Products were analyzed on 12% native polyacrylamide gels as described for electrophoretic mobility shift assay.

Limited Proteolysis of the WRN C Terminus and Identification of the C-terminal Domain Structure—A WRN protein fragment comprising amino acids 950–1360 was subjected to partial subtilisin proteolysis. In a typical experiment, 30 μg of protein was digested with increasing

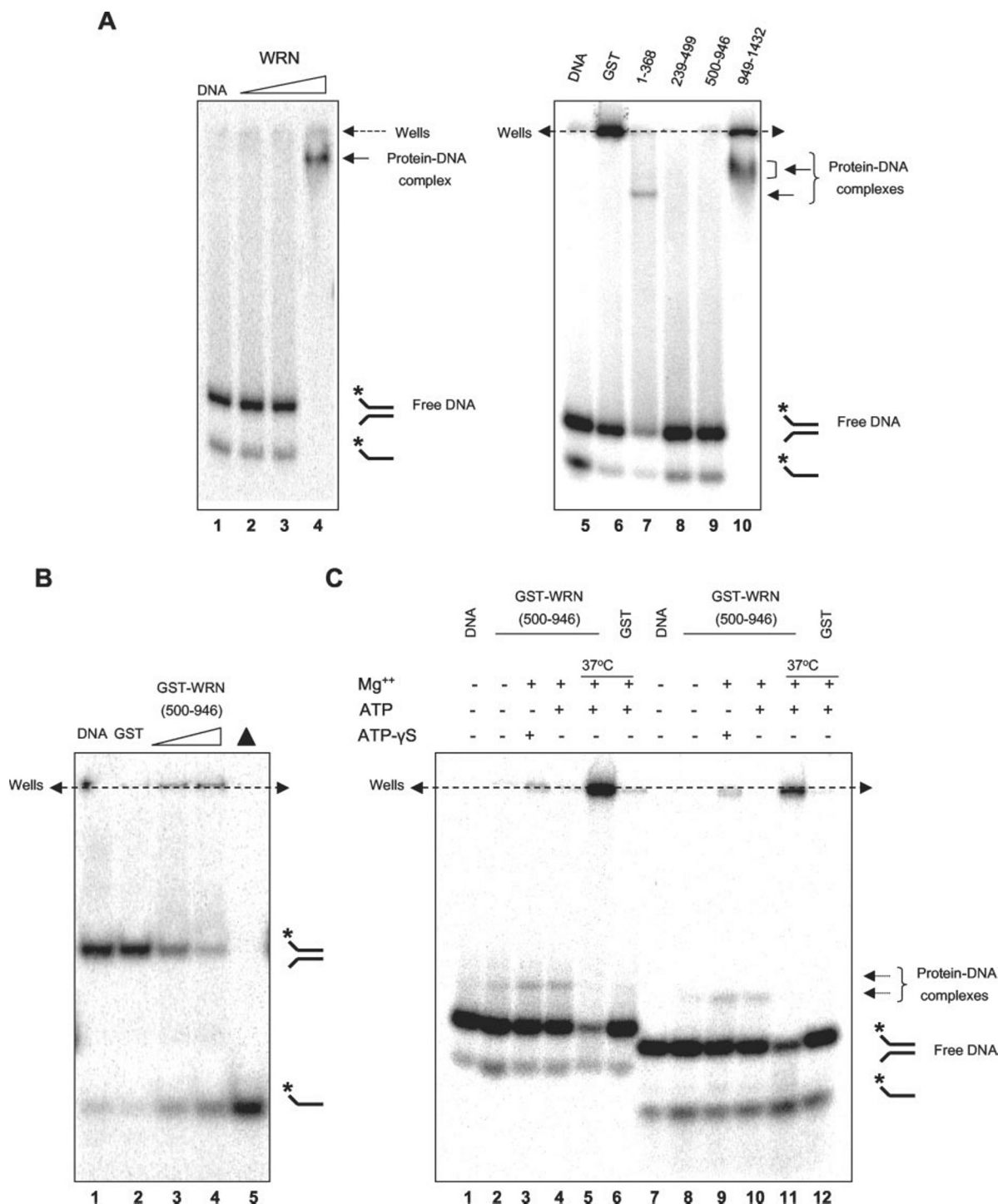


FIG. 2. DNA binding properties of WRN N- and C-terminal domains. A, gel mobility shift assay; full-length WRN (0.1, 0.2, and 1 pmol, lanes 2–4), recombinant WRN fragments (1 pmol, lanes 7–10), or GST (1 pmol, lane 6) were incubated with 30 fmol of radiolabeled long forked duplex DNA substrate for 30 min on ice. The protein-DNA complexes were then analyzed as described under “Materials and Methods.” Dashed lines indicate the origin of electrophoresis. B, helicase assay; GST (2 pmol, lane 2) or GST-WRN helicase domain (1 and 2 pmol, lanes 3 and 4, respectively) were incubated with 10 fmol of radiolabeled short forked duplex for 15 min at 37 °C. the filled triangle (lane 5) is control reaction in which substrate was denatured by boiling. C, gel mobility shift assay; GST (1 pmol, lanes 6 and 12) or GST-WRN helicase domain (1 pmol, lanes 2–5 and 8–11) were incubated with 30 fmol long (lanes 1–6) or short forked substrate (lanes 7–12) for 30 min. The reactions were carried out on ice (lanes 1–4 and 7–10) or at 37 °C (lanes 5, 6, 11, and 12) in the presence or absence of ATP, ATP γ S, and/or Mg²⁺ as indicated in the figure. D, top, functional domains of WRN and recombinant C-terminal fragments (dashed lines). Numbers indicate WRN amino acid coordinate. Bottom, gel mobility shift assay; recombinant WRN fragments (1 pmol, lanes 2–6) or GST (1 pmol, lane 2) were incubated with 30 fmol of radiolabeled long forked duplex for 30 min on ice. Dashed lines indicate the origin of electrophoresis.

amounts of subtilisin (Fluka) ranging from 0.065 to 2% (w/v), for 10 min on ice. The reaction was stopped by addition of phenylmethylsulfonyl fluoride to a final concentration of 10 mM. Samples were subsequently analyzed by SDS-PAGE and transferred to polyvinylidene difluoride membrane for N-terminal sequencing.

RESULTS

The DNA binding activity of WRN was localized to specific protein regions by performing DNA binding assays with a series of recombinant truncated WRN variants and various

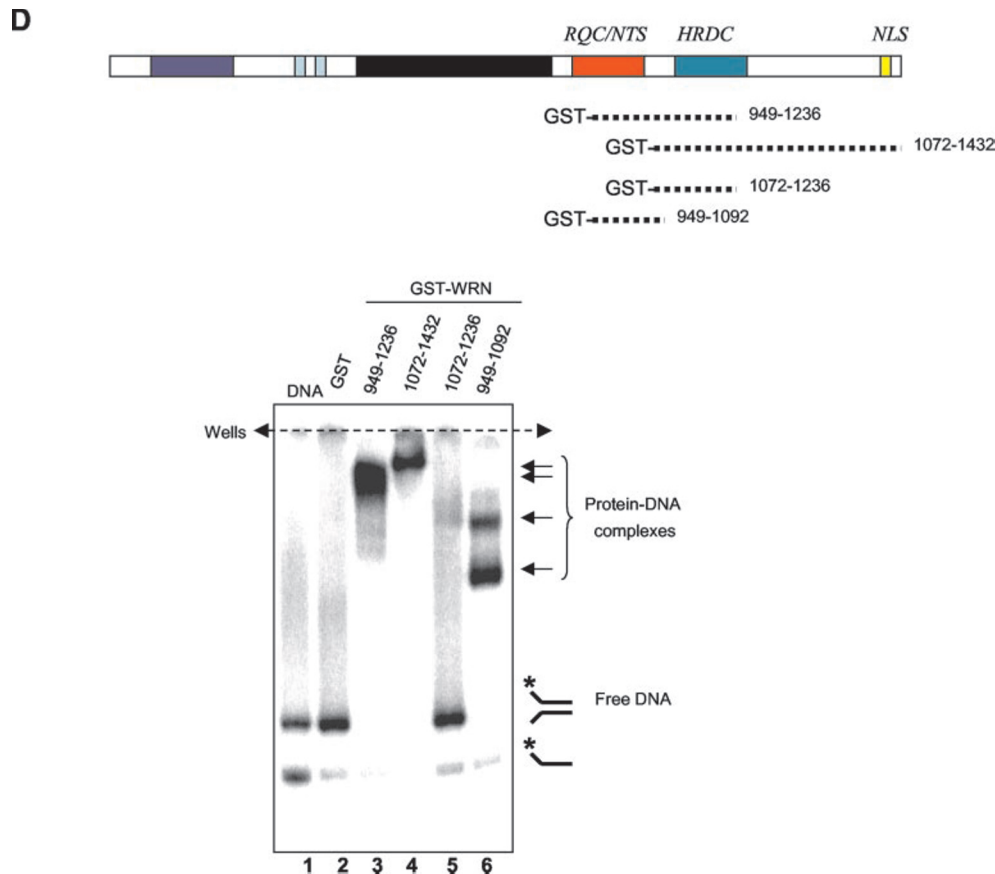


FIG. 2—continued

DNA substrates (Figs. 1, A and B). DNA binding was assayed on ice in the absence of ATP and divalent cations so that DNA substrates would not be degraded or altered by WRN catalytic activity. The long forked duplex DNA substrate has ssDNA and dsDNA regions and is a good substrate for WRN enzymatic activities (17). The long forked DNA substrate was used for initial assessment of WRN DNA binding capacity, because it can be recognized as ssDNA, dsDNA, a dsDNA break, or a forked structure (Fig. 1C). The forked DNA substrate is also recognized by replication protein A, a ssDNA binding protein, and Ku 70/80, a dsDNA break binding protein complex (data not shown). As shown in Fig. 2A (lanes 1–4), full-length recombinant WRN binds to the long forked substrate when the binding assay is carried out at a high enzyme concentration (lane 4). WRN binds to DNA more strongly in the presence of ATP than in its absence (22), which may explain the relatively weak binding of WRN to the long forked DNA substrate under the standard binding conditions used in these experiments.

On the primary amino acid sequence level, WRN can be subdivided into three characterized structural domains, the exonuclease, the RecQ like helicase, and the HRDC domain (Fig. 1A). We examined whether these domains expressed by themselves would constitute autonomously folded domains. The integrity of the catalytic exonuclease (aa 1–368) and helicase (aa 500–946) expression constructs were verified by virtue of their enzymatic activity. The exonuclease construct (aa 1–368) possesses 3′–5′-exonuclease activity identical to that of the full-length protein (24). The helicase domain expression construct (aa 500–946) was initially designed based on sequence alignment with other RecQ helicases and was subsequently shown to be catalytically active (Ref. 18 and this study; Fig. 2B). In the absence of catalytic activities, the domain organization of the C terminus was assessed on the basis of limited

proteolysis experiments (Fig. 1D). The rationale behind this approach is that unfolded proteins are generally less resistant to proteases and disappear without distinct intermediates. Folded proteins are more resistant, and proteolysis often reveals distinct intermediates corresponding to subdomains. Limited subtilisin digest of a purified C-terminal WRN fragment (aa 950–1360), which included the RQC and HRDC domains, produced two protease-resistant fragments (Fig. 1D). N-terminal sequencing showed that the 17-kDa fragment started at residue 958 (corresponding to the RQC domain), while the 14-kDa fragment started at either residue 1128 or 1148 (corresponding to the HRDC domain). This limited proteolysis pattern indicated that the C-terminal region of WRN contains two folded domains that encompass the RQC and HRDC domains, respectively. Based on the observed molecular weight in the SDS-PAGE, we estimated the C terminus of the structural domain corresponding to the RQC domain to be around residue 1092. A recombinant RQC fragment with similar boundaries is resistant to limited proteolysis by subtilisin (data not shown), providing further evidence that the RQC fragment is folded. In the studies of the HRDC domain, we also included the linker region connecting RQC and HRDC domains. This linker is predicted to be unstructured by secondary structure prediction programs (data not shown). Thus, the HRDC construct used starts at amino acid 1072 and ends at 1236.

Having established these boundaries, we designed and expressed various constructs containing either RQC (aa 949–1092), HRDC (aa 1072–1236), RQC and HRDC (aa 949–1236), HRDC containing C-terminal domain (aa 1072–1432), and the C-terminal half of WRN starting at 949 up to the very C terminus (aa 949–1432) (Fig. 1, A and B).

Four WRN fragments were initially used to localize WRN

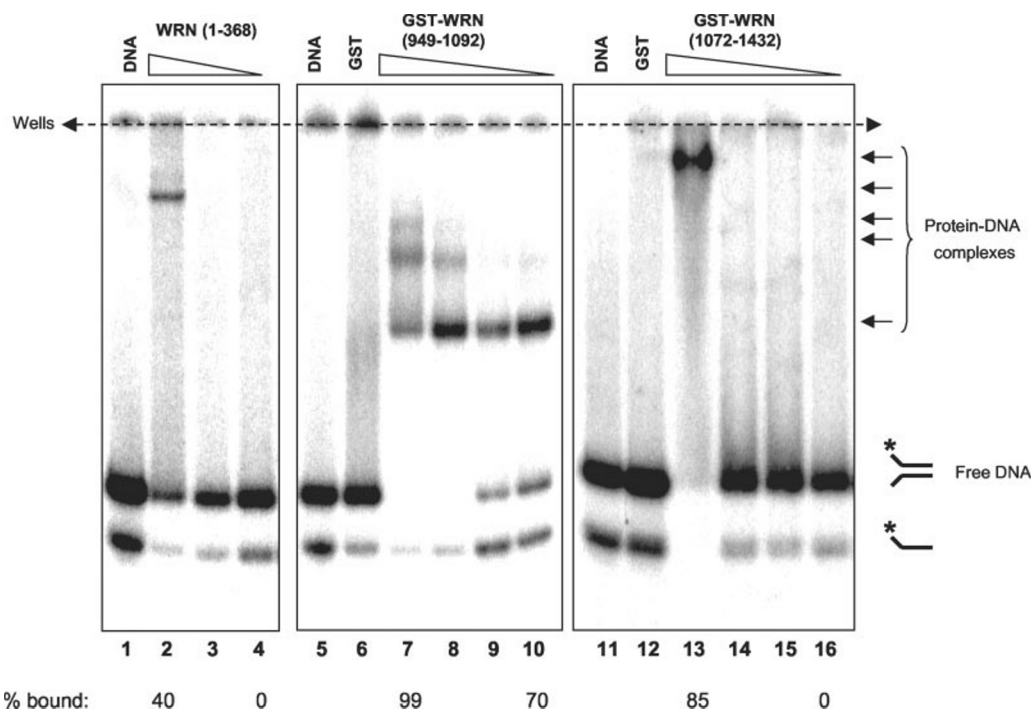


FIG. 3. **Interaction between WRN RQC-containing domain and long forked duplex.** Recombinant N-WRN (1, 0.5, and 0.2 pmol, lanes 2–4), GST-WRN RQC domain (1, 0.5, 0.2, and 0.1 pmol, lanes 7–10), GST-WRN HRDC-containing region (1, 0.5, 0.2, and 0.1, lanes 13–16), or GST (1 pmol, lanes 6 and 12) were incubated with 30 fmol of radiolabeled long forked substrate for 30 min on ice. Dashed lines indicate the origin of electrophoresis.

DNA binding domains. Binding activity was detected in the N-terminal exonuclease domain (aa 1–368; Fig. 2A, lane 7) and a C-terminal fragment (aa 949–1432; Fig. 2A, lane 10), but GST (Fig. 2A, lane 6), the acidic region (aa 239–499; Fig. 2A, lane 8), and the helicase domain (aa 500–946; Fig. 2A, lane 9) did not bind the long forked DNA substrate specifically. However, the GST-WRN helicase domain unwinds DNA in the presence of ATP and Mg^{2+} (18), suggesting that the interaction between this WRN fragment and the DNA substrate may have been too weak to detect by gel mobility shift assay in the absence of enzyme cofactors. This idea was tested by comparing binding activity in the presence and absence of ATP, ATP γ S, and Mg^{2+} at 37 °C and on ice. As shown in Fig. 2B, the helicase domain efficiently unwound the short forked DNA substrate in binding buffer plus ATP and Mg^{2+} (lanes 3 and 4), but GST did not (lane 2). In binding reactions on ice, plus Mg^{2+} and ATP (Fig. 2C), the WRN helicase domain formed a weak protein-DNA complex with short and long forked DNA substrates (lanes 3, 4, 9, and 10). Specific binding did not occur in reactions at 37 °C, although DNA was detected at the origin of electrophoresis (lanes 5 and 11). This supershifted DNA might be a high molecular weight DNA-helicase domain complex that only forms under these conditions. Under the same conditions, GST does not form a similar high molecular weight complex (lanes 6 and 12). These results indicate that WRN includes three distinct DNA binding regions: the N-terminal region, the C-terminal region, and the helicase domain. The DNA binding activity of the N- and C-terminal domains appears to be fairly robust, but the helicase domain binds DNA weakly and requires ATP/ Mg^{2+} for binding.

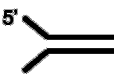
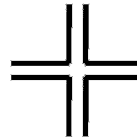
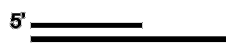


The WRN C-terminal DNA binding domain was localized more precisely with four purified GST-WRN fragments that lie within the aa 949–1432 region. The long forked DNA was used as the DNA binding substrate (Fig. 2D). Three of the C-terminal subfragments, aa 949–1236, 949–1092, and 1072–1432 formed strong protein-DNA complexes (lanes 3, 4, and 6), but the aa 1072–1236 fragment and GST did not bind the DNA

substrate (lanes 2 and 5). These results suggest that the WRN C-terminal region has two DNA binding subdomains, one associated with the RQC domain (aa 949–1092) and one associated with the HRDC containing C-terminal domain (aa 1072–1432).

Binding curves were carried out with a constant DNA concentration (long forked substrate) and variable concentrations of the WRN N-terminal, RQC, or HRDC regions of WRN (Fig. 3). The N-terminal region (aa 1–368) and the HRDC-containing domain (aa 1072–1432) only formed stable DNA complexes at high protein concentrations; at a ratio of 1 pmol of protein to 30 fmol of DNA, the N-terminal region bound 40% of the DNA substrate and the HRDC-containing domain bound 85% (Fig. 3, lanes 2 and 13). The RQC domain (aa 949–1092) bound the forked DNA substrate with much higher affinity, forming complexes with 70–99% of the DNA substrate at 0.1–1.0 pmol protein per assay (Fig. 3, lanes 7–10). In addition, three protein-DNA complexes formed when RQC was present at 1 pmol (50 nM; lane 7). At lower protein concentrations (6.2–25 nM), the protein-DNA complexes were faster migrating (lanes 8, 9, and 10, corresponding to 25, 12.5, and 6.2 nM, respectively). This result suggests that the RQC domain may form a higher order oligomer that assembles in a stepwise manner and that may involve distinct binding sites on the DNA substrate (ssDNA, dsDNA, or forked structures).

The structure specificity of the three DNA binding regions of WRN was tested with four DNA substrates known to form *in vivo* in DNA metabolic pathways that are linked to WRN function (Table I). The substrates used were a 34-bp blunt-ended duplex (sequence identical to the long forked substrate), a 5'-overhang duplex, ssDNA (same sequence as labeled strand of 5'-overhang substrate), and a Holliday junction. The N-terminal region and the HRDC-containing domain of WRN did not bind the blunt-ended DNA substrate (Fig. 4A, lanes 1–4 and 9–12), and the RQC-containing domain only bound this substrate at high enzyme concentration (50 nM, 17% binding efficiency; Fig. 3B, lane 7). This result suggests that the RQC

TABLE I
Biological pathways in which the substrates used in this study could be formed

DNA substrates	DNA metabolic pathways in which these substrates could potentially be formed
<p>Long forked duplex</p> 	<p><u>Transcription</u>, <u>replication</u> fork intermediate, <u>telomeric</u> ends (D-loop), others.</p>
<p>Holliday junction</p> 	<p><u>Recombination</u> intermediate, <u>replication</u> fork regression, others.</p>
<p>5' Overhang duplex</p> 	<p><u>Repair</u> intermediate, others.</p>
<p>Single strand DNA</p> 	<p><u>Transcription</u>, <u>repair</u> intermediate, <u>replication</u> fork intermediate, <u>telomeric</u> ends (D-loop), <u>HR</u> (ssDNA tail formation), others.</p>
<p>Blunt ended duplex</p> 	<p><u>HR</u> and <u>NHEJ</u> (dsDNA breaks formation), others.</p>

domain has lower affinity for blunt-ended than for forked DNA (50 nM, 99% binding efficiency; Fig. 3, lane 7), in agreement with previous reports (22, 23).

The 5'-overhang duplex is a DNA repair intermediate. The N-terminal region bound to this DNA substrate poorly (Fig. 4B, lane 3, 7% binding), whereas the HRDC-containing domain showed 27% binding to the DNA substrate at a high concentration (Fig. 4B, lane 7) and formed a faint smeared species during electrophoresis. In contrast, the RQC domain bound efficiently to this substrate, binding 71% of the DNA substrate at 50 nM and 10% of the DNA substrate at 12.5 nM (Fig. 4B, lanes 5 and 6). The 5'-overhang is a good substrate for the exonuclease activity of full-length WRN and also for the exonuclease domain (aa 1–368), despite the fact that this domain does not form a gel-stable protein-DNA complex under the conditions tested. A similar observation was noted above for the WRN helicase domain (Fig. 2C).

ssDNA was only a substrate for the RQC domain, which bound up to 60% of the input DNA substrate (Fig. 4B, lanes 9–12). This is in agreement with previous studies showing that WRN binds efficiently to ssDNA (23). The result in Fig. 4 strongly suggests that the RQC domain mediates the interaction between WRN and ssDNA.

HJs resemble recombination and replication fork arrest intermediates. Specific binding was not detected between the HJ substrate and the WRN N-terminal domain (Fig. 4C, lanes 2–4) or GST (Fig. 4C, lanes 6 and 10). In contrast, the RQC domain bound tightly to the HJ substrate at protein concentrations of 12.5–50 nM. The HRDC-containing domain also bound the HJ domain with high affinity forming detectable complexes at 25–50 nM (lanes 11 and 12). Thus, the HRDC-containing domain bound the HJ substrate with higher affinity than the long forked or the 5'-overhang duplex substrates. Although the RQC and HRDC-containing domains of WRN both bind efficiently to the HJ substrate, the complex between the HJ and RQC ap-

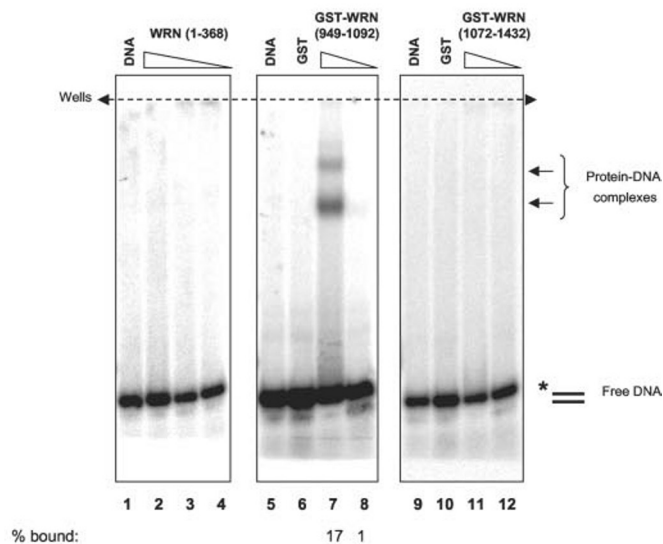
pears to be somewhat more stable (less smeared) than the complex between the HJ and the HRDC-containing domain (Fig. 4C, lanes 7, 8, and 11).

DISCUSSION

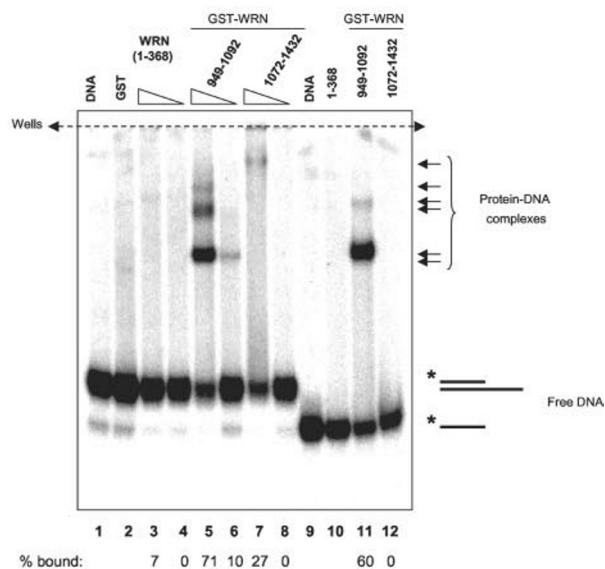
Many studies of WRN have focused on WRN catalytic activities or the potential biological role of WRN in DNA recombination, replication, repair, or telomere maintenance (7). In contrast, WRN binding to its DNA substrates is relatively poorly studied and poorly characterized. This study characterizes the DNA binding properties of a series of truncated recombinant WRN variants. The experiments were designed to analyze the DNA binding capacity of distinct WRN subdomains. Five DNA substrates that resemble typical *in vivo* recombination intermediates were used so that high affinity structure-specific DNA binding could be readily detected. The results of these binding studies indicate that WRN has three regions that form stable interactions with DNA. It is interesting to note that WRN binds DNA in a structure-specific rather than a sequence-specific manner and that each WRN DNA binding domain recognizes different DNA substrates (Fig. 5).

This study was conducted with recombinant WRN fragments. Individual exonuclease and helicase domains of WRN have previously been shown to constitute autonomously folded proteins exhibiting catalytic activity comparable with full-length WRN (Refs. 18 and 24 and this study). Other members of the family of HRDC domains have been expressed as individual recombinant proteins, and the structure of the Sgs1 HRDC has been solved demonstrating that the HRDC domain is independently folded (21, 26, 27). To address the domain organization of the WRN C-terminal region, we analyzed a fragment encompassing the RQC and HRDC domains using limited proteolysis (Fig. 1D). This analysis indicated that the RQC and HRDC domains are independently folded structural domains. Very recently, the structure of the catalytic core of

A



B



C

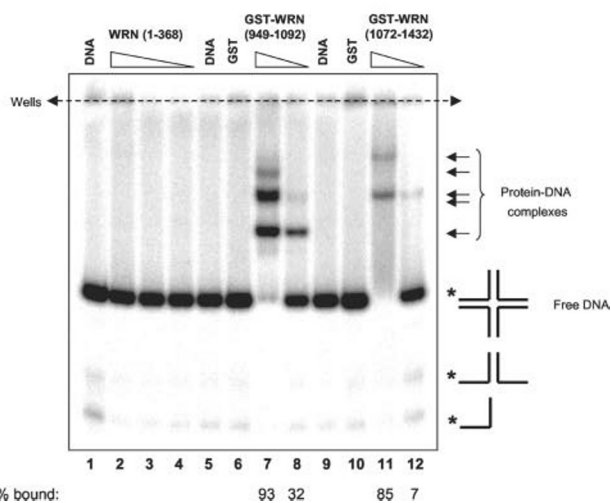


FIG. 4. DNA substrate specificity of WRN DNA binding domains. A, recombinant N-WRN (1, 0.5, and 0.2 pmol, lanes 2–4), GST-WRN RQC domain (1 and 0.1 pmol, lanes 7 and 8), GST-WRN HRDC-containing region (1 and 0.1 pmol, lanes 11 and 12), or GST (1 pmol,

E. coli RecQ has been described (28). In this study the authors identified the RecQ RQC domain as an independently folded subdomain, predicting that the RQC domain would have DNA binding and protein-protein interaction activities. Since RecQ is highly homologous to WRN, this study supports observations that the RQC is a functional domain and our results with the WRN RQC domain demonstrate their prediction about its function.

The experiments described here examine WRN domains binding to ssDNA, dsDNA, blunt, and overhanging dsDNA ends and forked DNA (Fig. 1C). WRN catalytic activities, which could degrade or modify these DNA substrates, were suppressed during DNA binding assays by omitting ATP and Mg^{2+} . However, the WRN helicase domain (aa 500–946) did not bind DNA under these conditions unless ATP and Mg^{2+} were included. This WRN domain demonstrates specific, but weak, binding to long and short forked DNA substrates in the presence of ATP and Mg^{2+} . The WRN helicase domain is catalytically active toward these substrates under the same conditions.

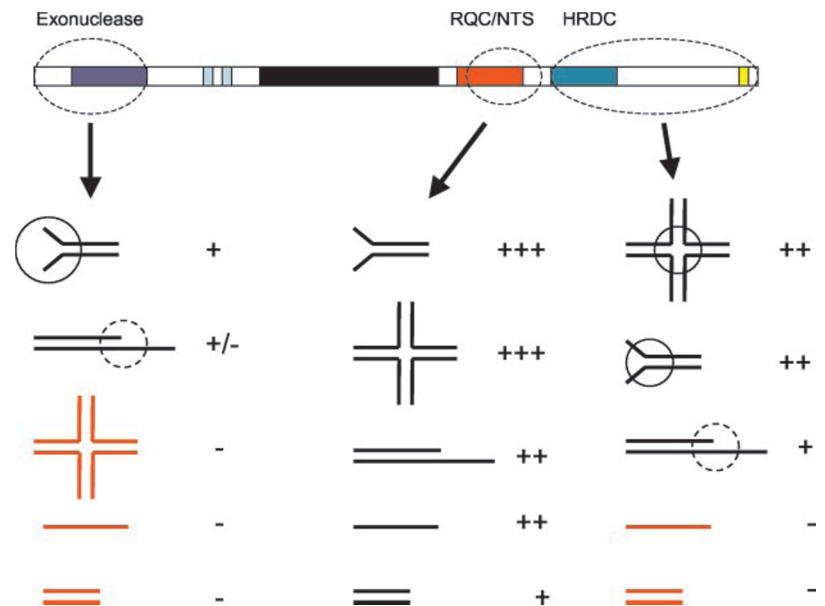
The WRN N-terminal exonuclease-containing domain (aa 1–368) binds tightly to forked DNA structures (17) but weakly to the 5'-overhang duplex (13, 14, 22). This result indicates that WRN catalytic activity and DNA binding do not always correlate with each other. In this study, the N-terminal region of WRN bound weakly to the long forked substrate, and the acidic region (aa 239–499) did not bind this DNA substrate. This result suggests the N-terminal 239 amino acids of WRN (aa 1–239) is required for specific binding to open forked DNA (*i.e.* forked duplex) and that this domain does not specify binding to forked DNA in a closed conformation such as an HJ DNA substrate. A previous study reported that the WRN exonuclease domain (aa 70–240) interacts with a 5'-overhang DNA substrate (29), which is consistent with our observation of a similar but weak interaction.

This study establishes the presence of two DNA binding domains in the WRN C-terminal region corresponding to the conserved RQC- and HRDC-containing domains. The RQC-containing domain (aa 949–1092) has broad specificity but binds with less affinity to blunt-ended duplex DNA than to other DNA substrates. Recent studies show that the WRN RQC domain is clearly multifunctional; it has been reported to stimulate FEN-1 incision (9), facilitate nuclear localization signal-dependent targeting to the nucleolus (30), and facilitate protein-protein interactions with FEN-1, BLM, TRF2, and PARP-1 (9, 14, 15, 31). Interestingly, the RQC domain is the only region that binds ssDNA (Figs. 4B and 5). *In vivo*, this substrate (ssDNA) is continuously generated during transcription, and most transcription occurs in the nucleoli. The fact that a RQC deleted WRN protein is excluded from the nucleoli (30) supports our conclusions that this domain is important for the biological functions of WRN.

The WRN HRDC-containing domain (aa 1072–1432) binds to the forked duplex and HJ with high affinity and to the 5'-overhang duplex with lower affinity. In contrast, this domain

lanes 6 and 10) were incubated with 30 fmol of radiolabeled blunt-ended duplex DNA for 30 min on ice. Dashed lines indicate origin of electrophoresis. B, recombinant N-WRN (1 and 0.5 pmol, lanes 3 and 4, and 1 pmol in lane 10), GST-WRN RQC domain (1 and 0.1 pmol, lanes 5 and 6, and 1 pmol in lane 11), GST-WRN HRDC-containing region (1 and 0.1 pmol, lanes 7 and 8, and 1 pmol in lane 12), or GST (1 pmol, lane 2) were incubated with 30 fmol of radiolabeled 5'-overhang (lanes 2–8) or 53-bp ssDNA (lanes 10–12) for 30 min on ice. C, recombinant N-WRN (aa 1–368) (1, 0.5, and 0.2 pmol, lanes 2–4), GST-WRN RQC domain (aa 949–1092), (1 and 0.1 pmol, lanes 7 and 8), GST-WRN HRDC-containing region (1 and 0.1 pmol, lanes 11 and 12), or GST (1 pmol, lanes 6 and 10) were incubated with 30 fmol of radiolabeled HJ for 30 min on ice.

FIG. 5. Subdomain localization and DNA structure specificity of WRN DNA binding activities. Substrates are indicated in order of decreasing binding efficiency (indicated by +, +/-, and -) from top of the figure to bottom. Red indicates no binding. Circles in solid lines indicate DNA binding activity. Circles in dashed lines indicate a low affinity binding region. +++, >90% binding efficiency; ++, 50–90% binding efficiency; +, <50% binding efficiency; +/-, 5–10% binding efficiency.



does not bind blunt-ended duplex or ssDNA substrates. This result suggests that the HRDC-containing C-terminal domain binds preferentially to forked DNA in an open (forked or 5'-overhang) or closed conformation (HJ). This observation is consistent with previous analysis of structural homology within HRDC domains (21) and other experimental data. For example, the HRDC domain of *S. cerevisiae* RecQ helicase Sgs1p binds DNA (21), as do the HRDC domains of RecQ and BLM (26, 27). It is interesting to note that WS has been associated with mutations in the HRDC domain (32). Importantly, when the HRDC domain is expressed alone (aa 1072–1236), we could not detect any specific binding to the long forked substrate (Fig. 2D, lane 5). This result suggests that the region of WRN between aa 1236 to 1432 either might help the HRDC domain for stable DNA binding or directly interact with the DNA. Future studies will be necessary to clarify the role of this C-terminal region of WRN, which thus far has no known function.

Our results demonstrate that in addition to the DNA-dependent catalytic activities, WRN contains stable DNA binding domains (Fig. 5). It is known that both activities of WRN, exonuclease and helicase, can cooperate in the same substrate to process it more efficiently (17). Thus, the WRN DNA binding domains may function as auxiliary regions mediating specific intermolecular contacts, not only with DNA but also with other proteins, generating multiprotein-DNA complexes involved in the processing of DNA intermediates generated in either replication, recombination, DNA repair, and/or telomere metabolism. It is possible that in the context of full-length WRN, these DNA binding regions could be modulated (turned on/off) via post-translational modification and/or protein binding, affecting its accessibility toward specific DNA substrates generated in different DNA metabolic pathways. In this way, the extensive post-translational modifications of WRN via acetylation, SUMOylation and phosphorylation (25, 33, 34), and the increasing number of WRN interacting proteins showing binding to both the N- and C-terminal parts of WRN support this conclusion (for review, see Ref. 7).

Some of these WRN binding proteins could in turn compete for the stable DNA binding of WRN and thus modulate the processivity of WRN toward the substrate. For example, the exonuclease activity WRN is dramatically stimulated by the interaction with the Ku complex (13). In this context it could be possible that Ku competes with the N-terminal region

of WRN for stable DNA binding, thus increasing the exonuclease activity (processivity) by decreasing the stable binding to the substrate.

WRN is a multifunctional and modular nuclear protein containing 3'-5'-exonuclease and helicase activities. These catalytic activities reside in two different regions of the protein, the N-terminal and middle part of the protein, aa 1–368 and 500–946, respectively. As described above, when expressed as recombinant fragments, both domains are active, indicating that the enzymatic activities do not require the function of other WRN regions. However, our results suggest that other WRN regions could modulate the helicase and/or exonuclease functions via auxiliary contacts to DNA. If this was the case, it could explain why different RecQ helicase members, conserving the central helicase domain, perform different cellular roles due to the presence of specific N- and C-terminal domains. In this way, by having more conserved regions, more overlapping biochemical and cellular functions would be expected. It is tempting to speculate WRN and BLM may demonstrate partial functional redundancy, because both enzymes include conserved helicase, RQC, and HRDC domains. *In vitro* both proteins can bind and process similar substrates (16, 35–37), and *in vivo* WS and Bloom's syndrome share some phenotypic and cellular characteristics, thus supporting the hypothesis that both proteins act in the same cellular pathways, as has been proposed previously (14, 38–40). However, the WRN N-terminal exonuclease domain is unique, indicating that BLM could not compensate for a loss of function mutation in this WRN domain.

This study shows that the WRN exonuclease region binds forked DNA structures, which may implicate WRN activity in DNA replication. This possibility is consistent with the WS cellular phenotype which is characterized by a prolonged S-phase (41) and hyper-recombination (40, 42). It was recently proposed that WRN plays a role in progression of the DNA replication fork (43).

In summary, this study and other recent findings (for review, see Ref. 7) are consistent with diverse biological functions for different WRN regions including intracellular trafficking, protein-protein interactions, and structure-specific DNA binding. Future studies are needed to better define the biological roles of WRN and to understand the mechanisms by which both the DNA-dependent catalytic activities and DNA binding domains of WRN are regulated *in vivo*.

Acknowledgments—We thank Lale Dawut for technical assistance, Dr. David Wilson III for scientific discussions, Dr. Tomoko Kawai, Dr. Patricia Opresko, Dr. Wen-Hsing Cheng, and Dr. Robert M. Brosh, Jr. for sharing radiolabeled substrates. We also thank Dr. Jeanine Harrigan, Dr. Kazunari Hashiguchi, and Dr. Sudha Sharma for critical reading of the manuscript. Finally we thank Dr. James L. Keck for sharing unpublished results.

REFERENCES

- Martin, G. M. (1978) *Birth Defects Orig. Artic. Ser.* **14**, 5–39
- Yu, C. E., Oshima, J., Fu, Y. H., Wijsman, E. M., Hisama, F., Alisch, R., Matthews, S., Nakura, J., Miki, T., Ouais, S., Martin, G. M., Mulligan, J., and Schellenberg, G. D. (1996) *Science* **272**, 258–262
- Karow, J. K., Wu, L., and Hickson, I. D. (2000) *Curr. Opin. Genet. Dev.* **10**, 32–38
- Huang, S., Li, B., Gray, M. D., Oshima, J., Mian, I. S., and Campisi, J. (1998) *Nat. Genet.* **20**, 114–116
- Kamath-Loeb, A. S., Shen, J. C., Loeb, L. A., and Fry, M. (1998) *J. Biol. Chem.* **273**, 34145–34150
- Gray, M. D., Shen, J. C., Kamath-Loeb, A. S., Blank, A., Sopher, B. L., Martin, G. M., Oshima, J., and Loeb, L. A. (1997) *Nat. Genet.* **17**, 100–103
- Opresko, P. L., Cheng, W. H., von Kobbe, C., Harrigan, J. A., and Bohr, V. A. (2003) *Carcinogenesis* **24**, 791–802
- Brosh, R. M., Jr., Orren, D. K., Nehlin, J. O., Ravn, P. H., Kenny, M. K., Machwe, A., and Bohr, V. A. (1999) *J. Biol. Chem.* **274**, 18341–18350
- Brosh, R. M., Jr., von Kobbe, C., Sommers, J. A., Karmakar, P., Opresko, P. L., Piotrowski, J., Dianova, I., Dianov, G. L., and Bohr, V. A. (2001) *EMBO J.* **20**, 5791–5801
- Spillare, E. A., Robles, A. I., Wang, X. W., Shen, J. C., Yu, C. E., Schellenberg, G. D., and Harris, C. C. (1999) *Genes Dev.* **13**, 1355–1360
- Szekely, A. M., Chen, Y. H., Zhang, C., Oshima, J., and Weissman, S. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11365–11370
- Lebel, M., Spillare, E. A., Harris, C. C., and Leder, P. (1999) *J. Biol. Chem.* **274**, 37795–37799
- Cooper, M. P., Machwe, A., Orren, D. K., Brosh, R. M., Ramsden, D., and Bohr, V. A. (2000) *Genes Dev.* **14**, 907–912
- von Kobbe, C., Karmakar, P., Dawut, L., Opresko, P., Zeng, X., Brosh, R. M., Jr., Hickson, I. D., and Bohr, V. A. (2002) *J. Biol. Chem.* **277**, 22035–22044
- Opresko, P. L., von Kobbe, C., Laine, J. P., Harrigan, J., Hickson, I. D., and Bohr, V. A. (2002) *J. Biol. Chem.* **277**, 41110–41119
- Mohaghegh, P., Karow, J. K., Brosh, J. R., Jr., Bohr, V. A., and Hickson, I. D. (2001) *Nucleic Acids Res.* **29**, 2843–2849
- Opresko, P. L., Laine, J. P., Brosh, R. M., Jr., Seidman, M. M., and Bohr, V. A. (2001) *J. Biol. Chem.* **276**, 44677–44687
- Harrigan, J. A., Opresko, P. L., von Kobbe, C., Kedar, P. S., Prasad, R., Wilson, S. H., and Bohr, V. A. (2003) *J. Biol. Chem.* **278**, 22686–22695
- Orren, D. K., Theodore, S., and Machwe, A. (2002) *Biochemistry* **41**, 13483–13488
- Matsumoto, T., Shimamoto, A., Goto, M., and Furuichi, Y. (1997) *Nat. Genet.* **16**, 335–336
- Liu, Z., Macias, M. J., Bottomley, M. J., Stier, G., Linge, J. P., Nilges, M., Bork, P., and Sattler, M. (1999) *Struct. Fold. Des.* **7**, 1557–1566
- Shen, J. C., and Loeb, L. A. (2000) *Nucleic Acids Res.* **28**, 3260–3268
- Orren, D. K., Brosh, R. M., Jr., Nehlin, J. O., Machwe, A., Gray, M. D., and Bohr, V. A. (1999) *Nucleic Acids Res.* **27**, 3557–3566
- Machwe, A., Ganunis, R., Bohr, V. A., and Orren, D. K. (2000) *Nucleic Acids Res.* **28**, 2762–2770
- Karmakar, P., Piotrowski, J., Brosh, R. M., Jr., Sommers, J. A., Miller, S. P., Cheng, W. H., Snowden, C. M., Ramsden, D. A., and Bohr, V. A. (2002) *J. Biol. Chem.* **277**, 18291–18302
- Janscak, P., Garcia, P. L., Hamburger, F., Makuta, Y., Shiraishi, K., Imai, Y., Ikeda, H., and Bickle, T. A. (2003) *J. Mol. Biol.* **330**, 29–42
- Bernstein, D. A., and Keck, J. L. (2003) *Nucleic Acids Res.* **31**, 2778–2785
- Bernstein, D. A., Zittel, M. C., and Keck, J. L. (2003) *EMBO J.* **22**, 4910–4921
- Xue, Y., Ratcliff, G. C., Wang, H., Davis-Searles, P. R., Gray, M. D., Erie, D. A., and Redinbo, M. R. (2002) *Biochemistry* **41**, 2901–2912
- von Kobbe, C., and Bohr, V. A. (2002) *J. Cell Sci.* **115**, 3901–3907
- von Kobbe, C., Harrigan, J. A., May, A., Opresko, P., Dawut, L., Cheng, W. H., and Bohr, V. A. (2003) *Mol. Cell. Biol.* **23**, 8601–8613
- Moser, M. J., Oshima, J., and Monnat, R. J., Jr. (1999) *Hum. Mutat.* **13**, 271–279
- Blander, G., Zalle, N., Daniely, Y., Taplick, J., Gray, M. D., and Oren, M. (2002) *J. Biol. Chem.* **277**, 50934–50940
- Kawabe, Y., Seki, M., Seki, T., Wang, W. S., Imamura, O., Furuichi, Y., Saitoh, H., and Enomoto, T. (2000) *J. Biol. Chem.* **275**, 20963–20966
- Brosh, R. M., Jr., Karow, J. K., White, E. J., Shaw, N. D., Hickson, I. D., and Bohr, V. A. (2000) *Nucleic Acids Res.* **28**, 2420–2430
- Li, J. L., Harrison, R. J., Reszka, A. P., Brosh, R. M., Jr., Bohr, V. A., Neidle, S., and Hickson, I. D. (2001) *Biochemistry* **40**, 15194–15202
- Yang, Q., Zhang, R., Wang, X. W., Spillare, E. A., Linke, S. P., Subramanian, D., Griffith, J. D., Li, J. L., Hickson, I. D., Shen, J. C., Loeb, L. A., Mazur, S. J., Appella, E., Brosh, R. M., Jr., Karmakar, P., Bohr, V. A., and Harris, C. C. (2002) *J. Biol. Chem.* **277**, 31980–31987
- Imamura, O., Fujita, K., Itoh, C., Takeda, S., Furuichi, Y., and Matsumoto, T. (2002) *Oncogene* **21**, 954–963
- Enomoto, T. (2001) *J. Biochem. (Tokyo)* **129**, 501–507
- Yamagata, K., Kato, J., Shimamoto, A., Goto, M., Furuichi, Y., and Ikeda, H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8733–8738
- Takeuchi, F., Hanaoka, F., Goto, M., Yamada, M., and Miyamoto, T. (1982) *Exp. Gerontol.* **17**, 473–480
- Cheng, R. H., Murano, S., Kurz, B., and Shmookler, R. (1990) *Mutat. Res.* **237**, 259–269
- Rodriguez-Lopez, A. M., Jackson, D. A., Iborra, F., and Cox, L. S. (2002) *Aging Cell* **1**, 30–39