Crystallographic and NMR Analyses of UvsW and UvsW.1 from Bacteriophage T4*

Received for publication, July 18, 2007, and in revised form, September 17, 2007 Published, JBC Papers in Press, September 17, 2007 DOI 10.1074/jbc.M705900200

Iain D. Kerr†1,2, Sivashankar Sivakolundu†1,3, Zhenmei Li‡, Jeffrey C. Buchsbaum§, Luke A. Knox¶, Richard Kriwacki‡1,4, and Stephen W. White‡1,5

From the †Department of Structural Biology, St. Jude Children’s Research Hospital, Memphis, Tennessee 38105, the ‡Department of Radiation Oncology, Milton S. Hershey Medical Center, Penn State University, Hershey, Pennsylvania 17033, and the ¶Department of Molecular Sciences, University of Tennessee Health Science Center, Memphis, Tennessee 38163

The uvsWXY system is implicated in the replication and repair of the bacteriophage T4 genome. Whereas the roles of the recombinase (UvsX) and the recombination mediator protein (UvsY) are known, the precise role of UvsW is unclear. Sequence analysis identifies UvsW as a member of the monomeric SF2 helicase superfamily that translocates nucleic acid substrates via the action of two RecA-like motor domains. Functional homologies to Escherichia coli RecG and biochemical analyses have shown that UvsW interacts with branched nucleic acid substrates, suggesting roles in recombination and the rescue of stalled replication forks. A sequencing error at the 3′-end of the uvsW gene has revealed a second, short open reading frame that encodes a protein of unknown function called UvsW.1. We have determined the crystal structure of UvsW to 2.7 Å and the NMR solution structure of UvsW.1. UvsW has a four-domain architecture with structural homology to the eukaryotic SF2 helicase, Rad54. A model of the UvsW-ssDNA complex identifies structural elements and conserved residues that may interact with nucleic acid substrates. The NMR solution structure of UvsW.1 reveals a dynamic four-helix bundle with homology to the structure-specific nucleic acid binding module of RecQ helicases.

Bacteriophage T4 remains an ideal model system in which to study the fundamental mechanisms of nucleic acid metabolism. DNA replication is particularly well suited for study in T4 because the viral genome encodes all the necessary replicative enzymes and accessory proteins that exist in higher organisms (1). During the early stage of the T4 infection cycle, DNA synthesis is initiated by a classic origin-dependent mechanism in which replication complexes are assembled onto persistent R-loops (RNA-DNA hybrids) (2). However, this mechanism is quickly abandoned, and during the latter stages of infection the switch is made to the more efficient recombination-dependent replication (RDR) to synthesize the bulk of the DNA (3). RDR produces a complex concatemeric network of T4 DNA that is ultimately resolved into the circularly permuted and terminally redundant genome that exists within the mature virus particles. The initiation stage of RDR is a recombination reaction in which the 3′-end of a single-stranded region of T4 DNA invades a homologous region of DNA to generate a displacement (D) loop intermediate. The T4 replication machinery is then assembled within the D loop to begin replication from the invading 3′-end (3).

Homologous recombination is a key process in DNA metabolism that remains poorly understood at the molecular level, and T4 is an ideal system in which to study the mechanism due to its central role in RDR and T4 replication. The T4 proteins UvsX and UvsY perform equivalent functions to the recombination proteins Rad51 and Rad52, respectively, in eukaryotic systems (4). UvsX and UvsY are both members of the uvsWXY system, which is associated with susceptibility to DNA damage because of UV radiation (hence the name). This is consistent with the roles that recombination and RDR are now known to play in the repair of DNA double-strand breaks (5). In addition, it has been known for some time that stalled and collapsed replication forks can lead to DNA double-strand breaks (6), and RDR has a pivotal role in repairing the lesions generated by these potentially lethal events (7). Thus, T4 has also emerged as a model system for studying the molecular basis of the DNA repair processes that serve to maintain genome stability.

UvsW is a T4-encoded protein that was first identified through mutations that have multiple and seemingly diverse effects on T4 DNA metabolism (8, 9). The uvsW gene is the third member of the uvsWXY system, and the multiple roles of UvsW were rationalized by in vitro and in vivo studies which

* This work was supported in part by National Institutes of Health Grant GM66934 (to S. W. W.), Cancer Center Core Grant CA21765, and the American Lebanese Syrian Associated Charities (ALSAC). Use of the Advanced Photon Source was supported by the United States Dept. of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. 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.

† Both authors made equal contributions to this work.
‡ Current address: Department of Cellular and Molecular Pharmacology, University of California, San Francisco, 1700 4th St., Byers Hall Room 509, San Francisco, CA 94158.
§ Current address: Bruker BioSpin Corp., 15 Fortune Dr., Billerica, MA 01821-3991.
¶ Current address: Department of Structural Biology, St. Jude Children’s Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105. Tel.: 901-495-3040; E-mail: Stephen.White@stjude.org.
† Both authors made equal contributions to this work.
‡ Current address: Department of Radiation Oncology, Milton S. Hershey Medical Center, Penn State University, Hershey, Pennsylvania 17033, and the ‡Department of Molecular Sciences, University of Tennessee Health Science Center, Memphis, Tennessee 38163.
3 This Current address: Bruker BioSpin Corp., 15 Fortune Dr., Billerica, MA 01821-3991.
4 To whom correspondence may be addressed: Dept. of Structural Biology, St. Jude Children’s Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105. Tel.: 901-495-3290; E-mail: Richard.Kriwacki@stjude.org.
5 To whom correspondence may be addressed: Dept. of Structural Biology, St. Jude Children’s Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105. Tel.: 901-495-3040; E-mail: Stephen.White@stjude.org.
revealed that it can catalyze a variety of nucleic acid processing functions as an ATP-dependent helicase (10). Primary sequence analyses (10) revealed the presence of seven conserved helicase motifs that are involved in substrate binding and hydrolysis (11, 12). The motifs specifically suggested that UvsW belongs to the SF2 family of helicases, and this was confirmed by the 2.3-Å crystal structure of a stable N-terminal fragment of the inactive UvsW point mutant K141R (UvsWNF) (13). The fragment comprises a small N-terminal domain that is structurally homologous to the DNA-binding domain of the T4 transcription factor MotA (14), and the first RecA-like motor domain (1A) that has been observed in all monomeric helicase structures to date (15). UvsW is also involved in the switch from early origin-dependent to late replication-dependent replication during the T4 infection cycle, and specifically catalyzes the disassembly of R-loops that act as primers at replication origins (16).

During our structural analysis of UvsWNF, we identified an error in the T4 genome sequence that results in a shorter UvsW open reading frame and the generation of a new open reading frame that encodes a 76-amino acid protein now referred to as UvsW.1. This open reading frame is highly conserved in T4-related viruses, and this suggests that the function of UvsW.1 is somehow related to that of UvsW. It was recently shown that there exists a direct and functional interaction between UvsW and UvsW.1 (17). Here we report the crystal structure of full-length UvsW and the NMR solution structure of UvsW.1. Unsurprisingly, the missing C-terminal region of our earlier truncated UvsWNF structure comprises the second (2A) RecA-like motor domain, and the closest structural homolog to full-length UvsW is Rad54. UvsW.1 has a simple four-helical bundle structure that has structural similarities to the HRDC motif found in RecQ helicases.

**MATERIALS AND METHODS**

**UvsW-K141R**

*Overexpression and Purification*—The full-length UvsW-K141R construct was provided by Dr. Kenneth Kreuzer (Duke University Medical Center). The gene was subcloned into the pET28a expression vector (Novagen) between the NcoI and H9252 restriction sites such that a His6 tag was cloned into the expression vector pET11d (Novagen) using appropriate primers from T4 genomic DNA (provided by Dr. Kenneth Kreuzer, Duke University Medical Center) and the first RecA-like motor domain (1A) that has been observed in all monomeric helicase structures to date (15). UvsW is also involved in the switch from early origin-dependent to late replication-dependent replication during the T4 infection cycle, and specifically catalyzes the disassembly of R-loops that act as primers at replication origins (16).

During our structural analysis of UvsWNF, we identified an error in the T4 genome sequence that results in a shorter UvsW open reading frame and the generation of a new open reading frame that encodes a 76-amino acid protein now referred to as UvsW.1. This open reading frame is highly conserved in T4-related viruses, and this suggests that the function of UvsW.1 is somehow related to that of UvsW. It was recently shown that there exists a direct and functional interaction between UvsW and UvsW.1 (17). Here we report the crystal structure of full-length UvsW and the NMR solution structure of UvsW.1. Unsurprisingly, the missing C-terminal region of our earlier truncated UvsWNF structure comprises the second (2A) RecA-like motor domain, and the closest structural homolog to full-length UvsW is Rad54. UvsW.1 has a simple four-helical bundle structure that has structural similarities to the HRDC motif found in RecQ helicases.

**UvsW**

*Sample Preparation*—The UvsW gene was obtained by PCR with appropriate primers from T4 genomic DNA (provided by Dr. Kenneth Kreuzer, Duke University Medical Center) and cloned into the expression vector pET11d (Novagen) using NdeI and BamH1 restriction sites such that a His6 tag was placed at the N terminus. Unlabeled and 13C/15N-labeled UvsW.1 was expressed in separate colonies of BL21 (DE3) cells.

**Crystallization, Data Collection, and Structure Solution**—Commercial screening kits were used to determine crystallization conditions at 18 °C. The hanging-drop vapor diffusion method was employed with a reservoir volume of 700 µl, and drops consisted of 1 µl of protein (4–11 mg ml⁻¹) and 1 µl of precipitant. The final, optimized conditions were 1.0 M sodium malonate pH 7.0, 11 mg ml⁻¹ protein, and these yielded crystals of dimensions 0.1 × 0.1 × 0.2 mm³. Cryoprotection was achieved by briefly soaking the crystals in 1.0 M sodium malonate pH 7.0, 100 mM NaCl, and 50% glycerol, and a 2.7-Å data set was collected from a single crystal at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory. Data were collected in 1° oscillations using a crystal-to-detector distance of 300 mm and 3-s exposures, and integrated, scaled, and merged using the HKL2000 package (18). UvsW-K141R crystals belong to space group C222₁, with unit cell dimensions a = 118.6 Å, b = 155.2 Å, c = 101.5 Å. The structure was determined by molecular replacement using the CCP4 program MOLREP (19), and the N-terminal fragment UvsWNF (PDB code 1RIF) as the search model. A clear solution was obtained with an R-factor of 51.5% and a correlation coefficient of 0.52, and the missing C-terminal domain was clearly visible in the 2Fo-Fc and Fo-Fc electron density maps and traced using COOT (20) and O (21). Iterative rounds of model building and refinement using CNS (22) yielded the final structure with R_work and R_free values of 21.7 and 25.0%, respectively. The final structure was then analyzed and validated using the program PROCHECK (23). Relevant data collection and refinement statistics are shown in Table 1.
Structures of UvsW and UvsW.1

Following induction with 1 mM isopropyl-1-thio-β-D-galactopyranoside and 3 h of additional growth, the cells were lysed using a sonicator, and UvsW.1 was purified from the clarified supernatant using a Ni2⁺ column with buffer E (50 mM sodium phosphate, pH 6.5, 50 mM NaCl, 1 mM dithiothreitol). Pure UvsW.1 was concentrated by ultrafiltration to 5 mg/ml and stored at −80 °C.

NMR Analysis—The backbone chemical shift assignments were obtained using a standard triple resonance assignment strategy through the analysis of two-dimensional 1H-15N HSQC and three-dimensional HNCA, HNCACB, and CBCA(CO)NH spectra recorded using a 600 MHz Varian INOVA NMR spectrometer at 25 °C and pulse sequences provided by the manufacturer. The side chain assignments were obtained through the analysis of three-dimensional C(CO)NH-TOCSY, H(CCO)NH-TOCSY, HCCH-COSY, and HCCH-TOCSY spectra. 1H-1H distance restraints were obtained through the analysis of 15N- and 13C-edited three-dimensional NOESY-HSQC spectra (mixing time, 120 ms), which were recorded using an 800 MHz Bruker AVANCE NMR spectrometer equipped with a TCI cryogenic probe. NMRPipe (24) and Felix (Accelrys) software were used for NMR data processing, display, and analysis.

Structure Calculation and Refinement—The NOE assignments and structure calculations were performed iteratively using the program ARIA 2.1 (25) as interfaced with the structure computation program CNS (22). A total of 1,976 unambiguous distance restraints and 98 dihedral restraints were used for the final structure calculation. Secondary 13Cα chemical shift values, corresponding to the difference of experimental and random coil values (Δδ 13Cα) (26), for UvsW.1 were used to generate dihedral restraints using the program TALOS (27). Final structure calculation was performed with 100 randomly generated starting structures. The fifty lowest energy structures were used for further refinement using the program AMBER 9 (28). The structural refinement was performed using ff02 Amber force field. The solvent was implicitly represented by Generalized-Born model to represent a NaCl concentration of 0.1 M. The structures generated by ARIA were first energy-minimized for 2 ps. This was followed by a simulated annealing step where the temperature was increased to 800 K and then cooled to 0 K. The 20 lowest energy structures were then analyzed and validated using the program PROCHECK-NMR (29). Relevant statistics are shown in Table 2.

Flexibility Studies—The flexibility of the UvsW.1 polypeptide backbone was assessed by measuring 1H-15N heteronuclear nuclear Overhauser effect (NOE) values and by monitoring the extent of amide proton exchange after dissolution in buffer comprised of D2O. The 1H-15N heteronuclear NOE experiments were performed with a 4 s period of amide 1H saturation using a train of 120° hard pulses separated by a 5-ms delay, or a 4-s delay period, before the first 1H pulse (30). The value of the 1H-15N heteronuclear NOE (hetNOE) for each amide was calculated as the ratio of peak intensities in the irradiated (Iirrad) and non-irradiated (Icont) spectra (hetNOE value = Iirrad/Icont). To perform the proportionate

---

### Table 1

| Structures of UvsW and UvsW.1 | Data collection | Native |
|-------------------------------|----------------|--------|
| Wavelength (Å)                | 1.05           |        |
| Beamline                      | SER-CAT 22-1D  |        |
| Resolution (Å)                | 50.2-2.7 (2.80-2.70) | |
| Space group                   | C2221          |        |
| Cell constants                | a = 118.64Å, b = 155.20Å, c = 101.52Å | |
| V <sub>c</sub>                | 3.96           |        |
| Total measurements            | 748125         |        |
| No. of unique reflections     | 25580          |        |
| Average redundancy            | 6.3 (4.1)      |        |
| R <sub>free</sub>             | 35.2 (3.2)     |        |
| Completeness (%)              | 97.4 (82.2)    |        |
| R <sub>merge</sub>            | 8.2 (41.8)     |        |
| Wilson B factor (Å<sup>2</sup>)| 45.6           |        |

### Table 2

| Statistics of 20 lowest energy UvsW.1 structures determined by NMR spectroscopy | All residues | Residues 10–70 |
|--------------------------------------------------------------------------------|--------------|----------------|
| Total number of unique NOEs                                                   | 1,976        |                |
| Intra                                                                       | 669          |                |
| Sequential                                                                  | 498          |                |
| Medium range                                                                | 578          |                |
| Long range                                                                  | 222          |                |
| Total number of dihedral restraints                                          | 98           |                |
| ϕ                                                                          | 49           |                |
| ψ                                                                          | 49           |                |
| R.m.s.d. from mean structure (Å)                                             |              |                |
| Backbone                                                                   | 2.21 ± 0.54  |                |
| All heavy atoms                                                             | 2.96 ± 0.59  |                |
| Structured region (residues 10–70)                                           |              |                |
| Backbone                                                                   | 0.58 ± 0.18  |                |
| All heavy atoms                                                             | 1.18 ± 0.21  |                |
| Distance restraint violations in 20 lowest-energy structures                 |              |                |
| Total no. of restraints violated by >0.4 Å                                   | 1            |                |
| Maximum distance restraint violation (Å)                                     | 0.41         |                |
| Torsion angle restraint violations in 20 low energy structures               |              |                |
| No. of restraints violated by >1                                            | 0            |                |
| Ramachandran torsion angle statistics                                        |              |                |
| Residues in most favored region                                             | 82.1%        | 93.8%          |
| Residues in allowed region                                                   | 13.4%        | 5.4%           |
| Residues in generously allowed region                                       | 3.2%         | 0.8%           |
| Residues in disallowed region                                                | 1.3%         | 0.0%           |
RESULTS

UvsW

Overall Structure—We previously reported that the Walker A inactive point mutant UvsW-K141R (generated for functional studies) is considerably more soluble than wild-type UvsW, and this mutant was used for the initial structural studies of the stable N-terminal fragment (residues 1–282) that we refer to as UvsWNF (13). UvsW-K141R was also used to generate the crystals of the full-length molecule (Met¹−Leu⁵⁰²), and an additional Leu⁵⁰³ was added at the C terminus from the cloning procedure. Crystals in space group C222₁ that diffracted to 2.7 Å were grown from sodium malonate pH 7.0, and the structure was determined by molecular replacement using the UvsWNF structure as a search model. The asymmetric unit contains one molecule of UvsW-K141R, and the refined model is missing residues 287–295 that were not visible in the electron density. These residues correspond to the linker region that connects the two RecA-like motor domains.

The structure of UvsW-K141R reveals that the full-length molecule comprises four domains, a small N-terminal domain (1B), two signature RecA-like domains 1A and 2A (15) and a small substructure (2B) built from extensions from domain 2A. The structures, mode of interaction, and relative orientation of domains 1B and 1A are virtually identical to those observed in the UvsWNF structure (13). A least squares superimposition of UvsW-K141R onto each molecule of the UvsWNF asymmetric unit matches 271 and 278 α-carbon positions with r.m.s.d. values of 0.81 Å and 0.56 Å, respectively. The largest positional differences occur in two flexible loop regions (residues 31–51 and 262–269), and there are minor backbone torsional differences in residues 194–205. Fig. 1 shows an α-carbon trace of the newly characterized domains 2A and 2B, and Fig. 2 shows the primary structure and the associated secondary structure assignment of the entire molecule. Note that the secondary structure nomenclature now reflects the revised domain nomenclature, which differs from that previously reported (13).

The structure of domain 2A is very similar to that of equivalent minimal RecA-like domains of other monomeric helicases as exemplified in the NS3 structure (31). The structure is organized around a 7-stranded parallel β-sheet in which the strand order is β3A₂−β4A₂−β2A₂−β5A₂−β6A₂−β1A₂−β7A₂, and six α-helices, α1A₂ to α6A₂, are packed around the β-sheet. The small 2B substructure comprises two α-helices, α1B₂ and α2B₂, that are inserted between β1A₂ and α1A₂, and a β-ribbon/loop connecting β6A₂ and α6A₂. This substructure can-
Structures of UvsW and UvsW.1

FIGURE 3. Structure of full-length UvsW. A, actual crystal structure showing the molecule in the open conformation. B, structure in the modeled closed conformation in which domains 2A and 2B have been rotated according to the structure of NS3-ssDNA. Each domain is labeled. Key conserved residues within the helicase motifs and the two flexible loops are shown, and these are colored blue and pink to reflect interactions with DNA and nucleotide, respectively. Note that in the closed conformation, the helicase motifs congregate at the domain interface and the two flexible loops are at the top of the molecule. The figures were produced using MOLSCRIPT (57) and rendered with RASTER3D (58).

FIGURE 4. Model of the UvsW-ssDNA complex. The figure represents a closer view of Fig. 3B in which ssDNA (pink) has been added by superimposition with the HCV NS3-ssDNA crystal structure. The DNA-binding motifs, key α-helices and conserved flexible loops are labeled. For clarity, the conserved nucleotide binding residues have been omitted. Note that the N termini of helices α4A1 and α2A2 point directly at the DNA phosphates (represented as spheres). Also, the path of the DNA ends at helix α2B2 that may represent the UvsW wedge or pin that serves to unwind the DNA duplex in RecG (36) and UvrD (35), respectively. The figure was produced using MOLSCRIPT (57) and rendered with RASTER3D (58).

not be considered an independently folded domain, but we label it as such to facilitate comparisons with other helical structures.

Fig. 3A shows a scheme of the entire UvsW-K141R crystal structure in which the locations of key residues within the signature SF2 helicase motifs as defined by Hall and Matson (12) are indicated. It can be seen that the motifs within the two domains are well separated, and an approximate 180° rotation of domains 2A and 2B is required to bring the motifs together to create the architecture that is typical of the monomeric helicases (15). We therefore believe that the observed crystal structure represents a non-functional conformation of the molecule. Fig. 3B shows the structure after this rotation has been performed relative to the NS3-ssDNA co-structure (1A1V), and it can be seen that the motif residues now adopt their typical positions in the vicinity of the 1A/2A domain interface. Therefore, the conformation shown in Fig. 3B will form the basis of subsequent discussions.

Functional Regions—The seven SF2 helicase motifs are arranged in the typical fashion observed in other monomeric helicase structures (15), and their locations within the UvsW primary structure are indicated in Fig. 2. Note that the assignment of motif IV differs from that described previously based on sequence analysis (10). Motifs I, II, III, and VI congregate at the 1A/2A domain interface and are involved in ATP binding and hydrolysis, and motifs IA, IV, and V have important roles in binding the translocating nucleic acid substrate (12, 15). Recently (32), an additional Q motif has been identified that serves to interact with the adenine ring of ATP, and this is present in UvsW as Q118 within helix α2A1 (Fig. 2). We previously used the structure of the NS3 RNA helicase bound to ssDNA to generate a putative model of a UvsWNF-ssDNA complex (13), and this model can now be expanded to include the full-length UvsW structure (Fig. 4). Four threonine residues, Thr166 and Thr167 in motif IA and the adjacent Thr211 and Thr214 are appropriately positioned to engage the DNA sugar phosphate backbone from domain 1A. Thr211 and Thr214 are within the TWQT sequence in UvsW that corresponds to the TXG motif found in other helicases (33). Lys355 and His356 within motif IV, and Ser404, Ser409, and Thr410 within motif V can all interact with DNA from domain 2A directly opposite from domain 1A. Sequence information is now available for a number of T4-like genomes, and they all contain UvsW homologs that show a high level of similarity. The functionally important motifs and residues described above are all highly conserved.

The model shown in Fig. 4 agrees very well with the majority of the SF1 and SF2 structures that contain bound DNA substrates and/or nucleotides, and is consistent with the so-called inchworm mechanism that has been suggested for the general translocating mechanism based on structural studies of the SF1 PcrA helicase (34). Recently, a more refined wrench and inchworm mechanism has been proposed based on structural studies of the SF1 UvrD helicase (35). Fig. 3 also reveals that helix dipoles may have a role in the interaction with DNA, and possibly in the translocating mechanism, which has not been previously noted. Helix α4A1 in domain 1A and helix α2A2 in domain 2A are both aligned such that their N termini, with...
positive partial charge due to the helix dipoles, point directly at the n and n + 3 phosphate groups, respectively, of the bound substrate. Significantly, these helices are equivalent in terms of the common fold of the RecA-like domain, and they appear to perform matching functions in binding the translocating substrate. The wrench and inchworm translocation mechanism invokes an ATP-driven relative movement of domains 1A and 2A that accompanies a cyclical binding and release of the DNA, and changes in the alignments of the dipoles with respect to the DNA phosphate groups may have an important role in this cycle.

The sequence conservation also reveals two other regions of UvsW that are highly conserved among T4-like sequences and clearly important for function. Both correspond to positively charged loop regions. The first is a basic/aromatic loop that has been described previously and is an extension of domain 1B (13). The second (residues 458–475) is part of domain 2B and contains five lysine residues, Lys641, Lys643, Lys648, Lys469, and Lys470. This loop is surprisingly well resolved in the electron density map and comprises a short proximal β-ribbon (residues 458–461 and 472–475) that is anchored by the side chain of Tyr471 and a salt bridge cluster involving Arg305, Asp308, and Asp457. These two loops, together with their associated domains 1B and 2B, create a superstructure above the motor domains that presumably recognize the UvsW DNA substrates (Fig. 3B). Although UvsW can functionally replace RecG, it is much smaller than the E. coli helicase and lacks the relatively large additional domains that RecG uses to recognize, bind and unwind its branched DNA substrates. Apparently, the UvsW superstructure can perform equivalent functions, but perhaps with less specificity. Finally, UvrD contains a so-called separation pin structure that mediates the unwinding of duplex DNA upstream of the ssDNA (35) and has a similar role to the wedge that has been identified in RecG (36). Although UvsW has neither of these structures, Fig. 4 shows that the modeled ssDNA points directly at the face of helix α2B2 in domain 2B, which may perform an equivalent function.

UvsW and Rad54—Using the Dali search algorithm (37), it was found that the newly characterized UvsW domains 2A and 2B most closely resemble the equivalent domains of Rad54. Two structures of Rad54 have been determined, one in the so-called open conformation (38) that corresponds to our crystal structure of UvsW (Fig. 3A), and a second in a closed conformation (39) that matches the model shown in Fig. 3B. Comparing the two closed conformations, the overall architectures of UvsW and Rad54 are very similar, and domains 1B and 2B of UvsW topologically correspond to the HD1 and HD2 domains, respectively, of Rad54 (39). Notably, Rad54 contains an α-helix in an equivalent position, both spatially and with respect to the RecA fold, to helix α2B2 in UvsW that we suggest may have a role in DNA duplex separation. Rad54 also contains an extended linker between domains 1A and 2A that forms a short α-helix, and the linker region of UvsW that is not visible in our electron density map may form an equivalent α-helix in the closed conformation.

**UvsW.1**

**Structure and Dynamics**—T4 UvsW.1 comprises 76 residues, and our construct contained an additional three residues at the N terminus after removal of the His6 tag (Gly, Ser, and His). The primary structure and secondary structure assignment are shown in Fig. 5. The superposition of the 20 lowest energy solution structures of UvsW.1 is shown in Fig. 6A, and a scheme of the lowest energy structure is shown in Fig. 6B. UvsW.1 is comprised of 4 α-helices, helix α1 (residues 5–12), helix α2 (residues 14–24), helix α3 (residues 26–44), and helix α4 (residues 47–69). Helices α2, α3, and α4 pack together into a rigid three-helix bundle with well-defined connecting loops, and helix α1 is somewhat separated from this core. Helix α1 exhibits only partial helicity and heightened dynamics relative to the core, and two lines of evidence suggest that it is inherently flexible. First, secondary 13Cα chemical shift values (Δδ 13Cα) are less than 2 ppm and values between 3–4 ppm are typically observed in fully populated α-helices. Second, the 1H–15N heteronuclear NOE (hetNOE) values for residues in helix α1 (and part of helix α2) are generally reduced (between 0.3–0.6) relative to those...
Structures of UvsW and UvsW.1

in the helical core. This indicates that helix α1 exhibits heightened dynamics on the high ps-low ns time scale in contrast to the core that appears to be rigid on this time scale. The seven residues at the C terminus following helix α4 are unstructured based on near zero Δ13Cα values and negative hetNOE values.

The amide protons of UvsW.1 are subject to rapid exchange with protons of water. For example, the majority of amide protons of UvsW.1 exchanged after resuspension of a lyophilized protein sample in 2H2O (pH 6.5) within 15 min. A small number of residues exhibited un-exchanged amide protons at this initial time point, including Ile21, Leu30, Glu31, Leu33, Tyr37, Lys38, Glu45, Leu69, Arg62, and Leu75, but these somewhat protected amides exhibited complete exchange by 20 min. These results show that, while UvsW.1 is rigid on the ns-ps timescale and possesses well formed secondary structure, the globular core must experience breathing motions on a longer time scale, which break the hydrogen bonds that stabilize secondary structure. Thus, the hydrophobic core of the 3-helix bundle of UvsW.1 may not be as well packed as has been observed for other proteins with helical bundle structures.

Functional Regions—UvsW.1 is highly conserved in T4-related viruses, and the only major differences are insertions and deletions that are consistent with the solution structure and which correspond to the linker between helices α1 and α2, and at the flexible C terminus. It has been shown that UvsW.1 forms a functional complex with UvsW (17), and such protein-protein interactions are typically mediated by surface hydrophobic patches. One such conserved patch is evident on the surface of the helical bundle core and is centered on residues Ile14, Met18, and Ile52 (Fig. 6C). Thus far, we have not been able to generate a stable UvsW/UvsW.1 complex for structural analyses using either x-ray crystallography or NMR, and cannot confirm whether this patch mediates the interaction between the two proteins. Another prominent feature of the protein is an extreme surface electronegativity, especially at the flexible C terminus, and this is evident from Fig. 6D.

DISCUSSION

We have confirmed that UvsW is an ATP-dependent monomeric helicase that belongs to the SF2 class. Genetic and biochemical experiments have suggested that UvsW can act on a variety of substrates and elicit a number of functions during the T4 infection cycle. UvsW was first identified as a T4 genetic locus that controls genome stability, and mutations within the UvsW gene rendered the virus particularly susceptible to DNA damage due to UV radiation and treatment with hydroxyurea (8, 10). These observations suggested that UvsW has prominent roles in DNA double-strand break repair (DSBR), and the rescue of collapsed replication forks, respectively. It is now understood that recombination plays a central role in DSBR, and this is consistent with the close association of the uvsW, uvsX, and uvsY genes within the T4 uvsWX system (5). The role of UvsW in rescuing damaged replication forks has been less obvious, but insights were gained from the ability of the helicase to rescue RecG-deficient E. coli strains (10). RecG is known to rescue stalled forks by a process of fork regression whereby daughter strands are reannealed to facilitate fork rescue using a recombination-dependent mechanism via a so-called chicken-foot intermediate (36, 40). Thus, UvsW may also be capable of catalyzing fork regression in a similar manner to RecG. Finally, it is known that UvsW is a key player in the switch from early to late replication during the T4 infection cycle by disassembling R-loops that initiate early origin-dependent replication (16).

Data presented in the accompanying article from the Kreuzer laboratory (Webb et al., 60) reveal how UvsW can perform these multiple roles. They demonstrated that UvsW is inactive on blunt or overlapping linear duplexes, and also on so-called Y structures with one duplex region and two single-stranded regions in the arms. Instead, the preferred substrate is branched DNA that contains at least two duplex regions, and any larger DNA structure that contains this minimal motif would be predicted to be a substrate for UvsW. Notably, these include R-loops and D-loops, and Holliday junctions (HJs) that undergo efficient branch migration (BM) in the presence of UvsW. Webb et al. (60) also demonstrated that UvsW can process T4-generated replication intermediates including those that are expected to mediate fork regression. Although linear duplexes with 3'-ssDNA overhangs have been shown by others to be a substrate for UvsW (17), this has not been confirmed by our colleagues. Our structure of full-length UvsW does not directly show how the helicase specifically recognizes this motif, and this can only be revealed by a UvsW-DNA costructure.

However, we have previously noted that domain 1B resembles the dsDNA-binding double-wing motif of the T4 transcription factor MotA (13), and the two basic, conserved loops in the closed form of the molecule (Fig. 3B) are well positioned to participate in substrate recognition.

SF2 helicases are generally regarded as DNA (and RNA) remodeling enzymes that comprise a pair of RecA-like motor domains, and additional domains that recognize the specific nucleic acid substrate. They have key roles in nucleosome remodeling, DNA repair and transcription, and frequently displace proteins from the substrate as part of the remodeling process (41) (42). Known crystal structures from SF2 family members include RecG (36), UvrB (43), eIF4A (44), hepatitis C virus NS3 (31), RecQ (45), Vasa (46), and Rad54 (38, 39). Many SF2 helicases are known to translocate dsDNA, but the actual mechanism has not been resolved. A so-called inchworm mechanism has been proposed for how translocation is catalyzed by the monomeric helicases (15), and recent studies on UvrD have largely confirmed these proposals and provided further structural and functional insights (35). We have made the additional observation that two helix dipoles appear to have important conserved roles in the translocation mechanism by providing electrostatic interactions with the backbone phosphate groups across the 1A/2A domain interface. It is not clear how a DNA duplex is accommodated in this translocation mechanism, but the conserved motor domain architecture and associated motifs suggest that the fundamental process is identical. The SF1 helicases PcrA and UvrD both provide aromatic side chains that interact with the bases of ssDNA and that translocates through the molecule (34, 35). UvsW does not contain equivalent aromatic residues, and this suggests a different mode of interaction with dsDNA. An alternative proposal for the mechanism of SF2 helicases is based on an open conformation of Rad54 (38) that also exists in our crystal structure. How-
ever, we suggest that the open conformation of UvsW results from the combined effects of (1) the extended linker that connects domains 1A and 2A, (2) the repulsion of the basic surfaces that accommodate the DNA substrate, and (3) the lack of large interacting domains above the motor domains in Rad54 and UvsW that can lock the molecules in their functional conformation.

RecG is known to regress stalled replication forks (40) and to branch migrate HJs (47), and the structure of the large RecG-DNA fork complex has provided the best structural insights to date into how these enzymes specifically operate on their target substrates (36). Apart from the paired RecA domains, UvsW has no structural homology to RecG, and the much smaller T4 enzyme does not appear to be specifically designed for this role. Rather, it appears to be a less specialized SF2 helicase that can be adapted to a variety of remodeling processes that are required during the T4 infection cycle. Its penchant for HJs and ability to efficiently catalyze HJ branch migration is consistent with the key roles that HJs play in T4 RDR. T4 has long been recognized as an ideal model system for studying fundamental biological processes, not only in prokaryotes but also in eukaryotes where there are considerable sequence homologies between functionally equivalent proteins (48). Therefore, it comes as no surprise that the closest known structural homolog of UvsW is the eukaryotic SF2 helicase Rad54. This is intriguing because the two proteins also share a number of functional similarities. Rad54 has emerged as a fundamental player in eukaryotic homologous recombination (49) and was described as a eukaryotic HR Swiss army knife in a recent review of the enzyme (50). In addition, RAD54 was first identified as a yeast gene in which mutations have a similar sensitivity to DSBs as rad51 and rad52 mutants (51). This is reminiscent of the uvsWXY system in T4 and may reflect the parallel importance of Rad51/Rad52/Rad54 and UvsX/UvsY/UvsW in eukaryotic and T4 DSB repair, respectively, via homologous recombination. Thus, our continuing studies of the T4 proteins and their interactions in HR should provide valuable information on a key repair system in all eukaryotic cells.

Finally, our structural analysis of UvsW1 may provide functional insights into another important class of SF2 helicases, the RecQ enzymes. This class includes the Bloom’s (BLM) and Werner’s (WRN) helicases, and yeast Sgs1, and they are known to have roles in genome stability and recombination (52, 53). The HRDC (helicase-and-ribonuclease D/C-terminal) domain is found in a number of RecQ helicases and displays significant sequence variability. The three known structures of HRDC, from E. coli RecQ (54), yeast Sgs1 (55), and WRN (56), all reveal a helical bundle that, superficially at least, resembles that of UvsW1. The role of the HRDC is currently ill-defined but one suggestion is that it provides substrate recognition elements to enhance that of the parent SF2 helicase. In a recent study (17), it was shown that UvsW and UvsW1 form a functional complex, and UvsW1 appears to function as a non-covalent domain of the helicase. In addition, DNA annealing has been observed in members of the RecQ family, and it was also reported that UvsW has DNA annealing properties that are partially inhibited by UvsW1 (17). We have identified a conserved hydrophobic patch on the highly electronegative surface of UvsW1, and mutagenesis should provide insights into its interacting partners when a functional phenotype is identified. Predictive algorithms and NMR solution studies both reveal that UvsW1 is somewhat dynamic and disordered, and this suggests that a binding partner (UvsW, DNA substrate, or both) is necessary to achieve a more fully folded, functional conformation. In terms of the putative homology to the HRDC domain, it may be significant that the uvsW gene is directly at the 3’-end of the uvsW gene. As suggested by the name, the HRDC domain is typically located at the C terminus of RecQ molecules, and although UvsW1 is encoded by a separate gene in T4-related viruses, it may originally have been covalently linked to UvsW. Indeed, an artificial UvsW–UvsW1 linked construct displayed the same biochemical properties as the UvsW/UvsW1 complex (17).

Acknowledgments—We thank Dr. Kenneth N. Kreuzer and co-workers at Duke University for the original UvsW constructs, helpful comments, and advice. We also thank Mary-Ann Bjornsti for helpful comments.

References

1. Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Rüger, W. (2003) Microbiol. Mol. Biol. Rev. 67, 86–156
2. Nossal, N. G., Dudas, K. C., and Kreuzer, K. N. (2001) Mol. Cell 7, 31–41
3. Kreuzer, K. N. (2005) Annu. Rev. Microbiol. 59, 43–67
4. Beernink, H. T., and Morrical, S. W. (1999) Trends Biochem. Sci. 24, 385–389
5. Klein, H. L., and Kreuzer, K. N. (2002) Mol. Cell 9, 471–480
6. Cox, M. M., Goodman, M. F., Kreuzer, K. N., Serratt, D. J., Sandler, S. J., and Mianis, K. J. (2000) Nature 404, 37–41
7. McGlynn, P., and Lloyd, R. G. (2002) Nat. Rev. Mol. Cell Biol. 3, 859–870
8. Derr, L. K., and Drake, J. W. (1990) Mol. Gen. Genet. 227, 257–264
9. Derr, L. K., and Kreuzer, K. N. (1990) J. Mol. Biol. 214, 643–656
10. Carles-Kinch, K., George, J. W., and Kreuzer, K. N. (1997) EMBO J. 16, 4142–4151
11. Gorbalenya, A. E., and Koonin, E. V. (1993) Curr. Opin. Struct. Biol. 3, 419–429
12. Hall, M. C., and Matson, S. W. (1999) Mol. Microbiol. 34, 867–877
13. Sickmier, E. A., Kreuzer, K. N., and White, S. W. (2004) Structure 12, 583–592
14. Li, N., Sickmier, E. A., Zhang, R., Joachimiak, A., and White, S. W. (2002) Mol. Microbiol. 43, 1079–1088
15. Singleton, M. R., Dillingham, M. S., and Wagley, D. B. (2007) Annu. Rev. Biochem. 76, 23–50
16. Dudas, K. C., and Kreuzer, K. N. (2001) Mol. Cell 21, 2706–2715
17. Nelson, S. W., and Benkovic, S. J. (2007) J. Biol. Chem. 282, 407–416
18. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 267, 307–326
19. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
20. Emmsley, P., and Cowtan, K. (2004) Acta Crystallogr. D. Biol. Crystallogr. 60, 2126–2132
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W., Jiang, S. J., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D. Biol. Crystallogr. 54, 905–921
23. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
24. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
25. Nilges, M., Macias, M. J., O’Donoghue, S. I., and Oschkinat, H. (1997) J. Mol. Biol. 269, 408–422
26. Schwarzinger, S., Kroon, G. I., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) J. Am. Chem. Soc. 123, 2970–2978
Structures of UvsW and UvsW.1

27. Cornilescu, G., Delaglio, F., and Bax, A. (1999) J. Biomol. NMR 13, 289–302
28. Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) J. Comput. Chem. 26, 1668–1688
29. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
30. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Biochemistry 33, 5984–6003
31. Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C., and Caron, P. R. (1998) Biochemistry 33, 5984–6003
32. Tanner, N. K., Cordin, O., Banroques, J., Doere, M., and Linder, P. (2003) Mol. Cell 11, 127–138
33. Pause, A., and Sonenberg, N. (1992) EMBO J. 11, 2643–2654
34. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S., and Wigley, D. B. (1999) Cell 97, 75–84
35. Lee, J. Y., and Yang, W. (2006) Cell 127, 1349–1360
36. Singleton, M. R., Scaife, S., and Wigley, D. B. (2001) Cell 107, 79–89
37. Holm, L., and Sanders, C. (1993) J. Mol. Biol. 233, 123–138
38. Durr, H., Korner, C., Muller, M., Hickmann, V., and Hopfner, K. P. (2005) Cell 121, 363–373
39. Thomas, N. H., Czyzewski, B. K., Alexeev, A. A., Mazin, A. V., Kowalczykowski, S. C., and Pavletich, N. P. (2005) Nat. Struct. Mol. Biol. 12, 350–356
40. McGlynn, P., and Lloyd, R. G. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 8227–8234
41. Korber, P., and Horz, W. (2004) Cell 117, 5–7
42. Pazin, M. J., and Kadonaga, J. T. (1997) Cell 88, 737–740
43. Theis, K., Chen, P. J., Skovranga, M., Van, H. B., and Kisker, C. (1999) EMBO J. 18, 6899–6907
44. Caruthers, J. M., Johnson, E. R., and McKay, D. B. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 13080–13085
45. Bernstein, D. A., Zittel, M. C., and Keck, J. L. (2003) EMBO J. 22, 4910–4921
46. Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006) Cell 125, 287–300
47. Whitby, M. C., and Lloyd, R. G. (1998) J. Biol. Chem. 273, 19729–19739
48. Bernstein, H., and Bernstein, C. (1989) J. Bacteriol. 171, 2265–2270
49. Myung, K., Datta, A., Chen, C., and Kolodner, R. D. (2001) EMBO J. 20, 113–116
50. Bernstein, D. A., and Keck, J. L. (2005) Structure 13, 1173–1182
51. Liu, Z., Macias, M. I., Bottomley, M. J., Stier, G., Linge, J. P., Nilges, M., Bork, P., and Sattler, M. (1999) Structure 7, 1557–1566
52. Kitano, K., Yoshihara, N., and Hakoshima, T. (2007) WRN, J. Biol. Chem. 282, 2717–2728
53. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
54. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
55. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Proteins 11, 281–296
56. Webb, M. R., Plank, J. L., Long, D. T., Hsieh, T.-S., and Kreuzer, K. N. (2007) J. Biol. Chem. 282, 34401–34411