The DNA Binding Site(s) of the *Escherichia coli* RecA Protein*

William M. Rehrauer and Stephen C. Kowalczykowski‡

*From the Department of Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611
and the Division of Biological Sciences, Sections of Microbiology and of Molecular and Cell Biology,
University of California, Davis, California 95616-8665*

Photochemical cross-linking has been used to identify residues in the *Escherichia coli* RecA protein that are proximal to and may directly mediate binding of DNA. Ultraviolet irradiation promotes specific and efficient cross-linking of the RecA protein to poly(deoxythymidylic) acid. Cross-linked peptides remaining covalently attached to the polynucleotide following proteolytic digestion with trypsin correspond to amino acids 61–72, 178–183, and 233–243 of the RecA protein primary sequence. Their location and surface accessibility in the crystal structure, along with the behavior of various *recA* mutants, support the assignment of the cross-linked regions to the DNA binding site(s) of the RecA protein. Functional overlap of amino acids 61–72 with an element of the ATP binding site suggests a structural mechanism by which nucleotide cofactors allosterically affect the RecA nucleoprotein filament.

Genetic recombination is a universal occurrence that serves to generate genetic diversity, to preserve genomic integrity, and to ensure proper partitioning of chromosomes. Because of its basic role in DNA metabolism, elucidating the mechanism of this cellular process is of considerable importance. The product of the *RecA* gene is virtually indispensable to genetic recombination in *Escherichia coli* (1, 2), and extensive analysis has revealed a unique and critical enzymatic activity: the RecA protein promotes the recognition and exchange of strands between homologous DNA molecules (3–6). Consequently, by providing a model system for investigating biochemical steps central to genetic recombination, characterization of the RecA protein has laid the foundation for the discovery of both structural (7, 8) and functional analogues not only from eubacteria but also from organisms ranging from bacteriophage to higher eukaryotes (for review see Ref. 9 and references therein).

Interaction of the RecA protein with DNA and a nucleotide triphosphate cofactor, such as ATP, is fundamental to its ability to catalyze homologous pairing and subsequent transfer of strands between a variety of DNA substrates *in vitro* (10, 11). The functional species of the RecA protein is a helical nucleoprotein filament assembled through cooperative polymerization on either single-stranded DNA or duplex DNA containing a single-stranded gap or tail. Nucleoprotein filament formation requires stoichiometric amounts of the RecA protein and the binding of ATP to induce an active conformation in which the DNA assumes a highly extended and unwound state (12–14).

This unusual DNA conformation appears to be universal to the mechanism of genetic recombination because homologs such as the uvsX and Rad51 proteins from T4 bacteriophage and *Saccharomyces cerevisiae*, respectively, form complexes with DNA that are structurally (15, 16) and functionally analogous (17, 18).

The manner by which the RecA nucleoprotein filament recognizes sequence homology within a duplex DNA target remains unknown. Understanding of this mechanism would be greatly facilitated by definition of those regions within the RecA protein polymer that interact with DNA and are consequently central to the recognition process. Despite the resolution of a crystal structure (19, 20), the sites responsible for DNA binding within the RecA protein remain undefined. Structural and mutational analyses substantiate that the production of covalent linkages between nucleotide bases and amino acids through the action of ultraviolet irradiation is a general approach to probe the molecular interactions between species at the immediate interface of protein-nucleic acid complexes (for review see Refs. 21 and 22 and references therein). Covalent bond formation is thought to occur by a free radical mechanism in which a hydrogen atom from a favorably positioned amino acid is abstracted by a photoexcited nucleic acid base (21). Although model studies demonstrate differences in photo reactivity (23), it is presumed that all amino acids can in principle be cross-linked to nucleic acid bases through the absorption of ultraviolet light provided that the participating functional groups are in close proximity.

In this study, we employ photochemical cross-linking to discern those amino acid residues within the RecA nucleoprotein complex that are intimately associated with and may potentially mediate binding of single-stranded DNA. We show that specific cross-linking of the RecA protein to polydeoxythymidylic acid is efficient and saturates at an apparent stoichiometry consistent with direct single-stranded DNA binding studies. These studies both complement and extend the recently reported cross-linking of the RecA protein to oligonucleotides (24). The peptides within the primary structure of the RecA protein that are covalently linked to poly(dT) have been identified, and their possible involvement in the binding of DNA is discussed. The cross-linking of a peptide that coincides with a portion of the nucleotide binding site may provide further insight as to how conformational changes in the RecA nucleoprotein filament elicited by the binding and hydrolysis of ATP are directly transduced to regions implicated in DNA binding.

MATERIALS AND METHODS

RecA protein was purified from *E. coli* strain JC12772 (25) using a modified preparative protocol[1] based on spermidine precipitation (26); its concentration was determined using an extinction coefficient of 2.7×10^4 M^{-1} cm^{-1} at 280 nm. Single-stranded polynucleotides, poly(dT) and poly(dA) acids, were purchased from Pharmacia Biotech Inc. The

‡ To whom correspondence should be addressed: Section of Microbiology, University of California, Davis, California 95616-8665. Tel.: 916-752-5938; Fax: 916-752-5939; E-mail: sckowalczykowski@ucdavis.edu.

[1] S. C. Kowalczykowski, manuscript in preparation.

FIG. 1. *A*, photochemical cross-linking of the RecA protein to poly(dT). Samples containing RecA protein and poly(dT) were prepared under conditions that either support (100 mM NaCl binding, *lanes 1–6*) or prohibit complex formation (1000 mM NaCl nonbinding, *lanes 7–12*) and were irradiated with shortwave UV light for the times indicated. *B*, proteolytic sensitivity of cross-linked RecA protein-poly(dT) complexes. After ultraviolet irradiation for 120 s, cross-linked complexes were incubated in the absence (*lane 2*) or the presence (*lane 3*) of trypsin. As a control, poly(dT) was also subjected to trypsin (*lane 1*).
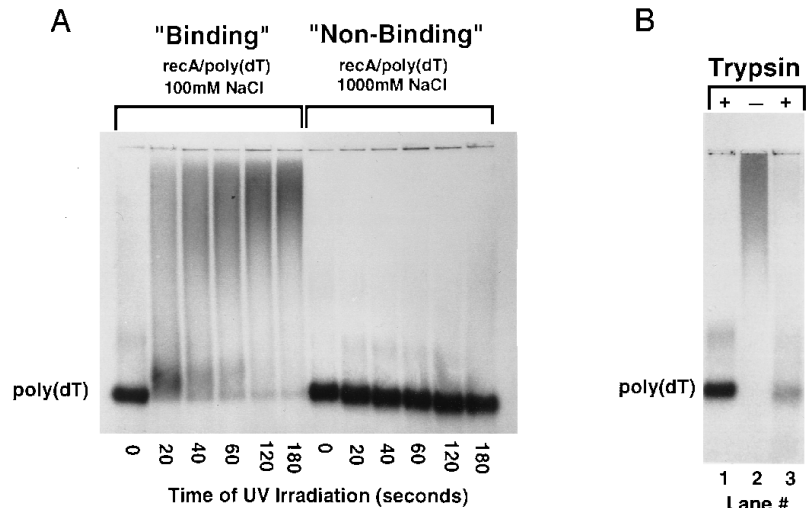
concentration of poly(dT) (average length, 300 nucleotides) and poly(dA) (average length, 390 nucleotides) was determined using an extinction coefficients of 8520 M^{-1} (nucleotide) cm^{-1} at 260 nm and 8600 M^{-1} (nucleotide) cm^{-1} at 257 nm, respectively. ATPγS[2] was purchased from Boehringer Mannheim and dissolved as a concentrated stock at pH 7.5 in TE buffer (10 mM Tris-HCl, 1 mM EDTA); its concentration was determined using an extinction coefficient of $1.54 \times 10^4\ M^{-1}\ cm^{-1}$ at 260 nm.

Photochemical Cross-linking of RecA Protein

Gel Assay—Reactions contained 10 μM RecA protein, 70 μM ^{32}P 5′-end labeled poly(dT), 25 mM Tris-HCl (pH 7.5), 6 mM MgCl$_2$, 0.1 mM dithiothreitol, and either 100 mM or 1000 mM NaCl. Samples (50 μl total) were incubated at 37 °C for 10 min. Irradiation of 25-μl aliquots was done on parafilm at a 7-cm height using a Spectroline XX-15F lamp (λ_{max} 254 nm; average intensity, 1.05×10^{-3} J/sec·cm^2). Samples were prepared for proteolysis by the addition of urea to a final concentration of 8 M, followed by microdialysis against 0.4 M NH$_4$HCO$_3$ (pH 8.0) and 8 M urea. To ensure complete denaturation, samples were reduced and carboxymethylated with excess dithiothreitol and iodoacetic acid, respectively. Prior to the addition of trypsin (Boehringer Mannheim, sequencing grade) at a 1:25 (w/w) ratio relative to RecA protein, samples were diluted with water to yield a final urea concentration of 2 M. Incubations were conducted in the presence of 0.1 mM CaCl$_2$ for 8 h at 37 °C. Samples were adjusted to 250 mM NaCl, denatured with 0.5% SDS, 25 mM EDTA, and made 0.025% bromphenol blue and xylene cyanol before electrophoresis on a 1.5% agarose gel and autoradiography.

Single-stranded DNA-dependent ATP Hydrolysis Assay—Samples (50 μl) containing RecA protein alone or in the presence of polynucleotide were irradiated as described for the gel assay. Following irradiation, samples were immediately placed on ice. The ATP hydrolysis activity of the RecA protein (μM ATP hydrolyzed/minute) was measured spectrophotometrically (27) using 30-μl aliquots of irradiated samples. Assays were done in a 300-μl reaction volume at 37 °C in a final buffer consisting of 25 mM Tris-HCl (pH 7.5), 6 mM MgCl$_2$, 0.1 mM dithiothreitol, 100 mM NaCl, 0.5 mM ATP (Pharmacia), 0.2 mg/ml NADH, 20 units/ml of both pyruvate kinase and lactate dehydrogenase, and 2 mM phosphoenolpyruvate (all from Sigma). The RecA protein ATP hydrolysis activity stimulated by saturating amounts of polynucleotide is not affected by addition of irradiated polynucleotide (data not shown). Thus, to ensure that any inhibition of single-stranded DNA-dependent ATP hydrolysis caused by ultraviolet irradiation was intrinsic to the RecA protein all assays included an excess of appropriate polynucleotide (7 μM). The percentage of ATP hydrolysis activity was determined by normalizing the rates of ATP hydrolysis measured for irradiated samples to those of nonirradiated samples. Because irradiation of the RecA protein either alone (100 mM or 1000 mM NaCl) or in the presence of polynucleotide under conditions that prohibited binding (1000 mM NaCl) resulted in equivalent amounts of inhibition, it was defined as background. Therefore, the percentage (differential) of inhibition in

ATP hydrolysis activity specifically due to the extent of RecA protein cross-linking was calculated by subtracting the amount of inhibition caused by irradiation of samples in the presence of 1000 mM NaCl ("nonbinding" conditions) from that observed in the presence of 100 mM NaCl ("binding" conditions).

Isolation and Identification of Cross-linked Peptides of the RecA Protein

Preparative amounts (reaction volume, 500 μl) of RecA protein-poly(dT) complexes were formed and exposed to shortwave ultraviolet light for 2 min, as described previously, with the only difference being that for samples containing 500 μM ATPγS irradiation was conducted for 3 min. Samples were prepared for proteolysis as detailed for Fig. 1 with the exception that after being denatured with 8 M urea the buffer of irradiated samples was exchanged to 0.4 M NH$_4$HCO$_3$ (pH 8.0), 2 M urea using a G-25 Sephadex fast desalting column (Pharmacia). Prior to reduction and carboxymethylation steps, samples were concentrated 4-fold by vacuum drying. Proteolytic digestions were conducted as described for the gel assay for a 24-h period with a second equivalent of trypsin being added after 12 h. Complexes of poly(dT)-cross-linked peptides were isolated in the void elution of a Superose12 column (Pharmacia) using 0.1 M NH$_4$HCO$_3$ (pH 8.0), 2 M urea, and 1 M NaCl and then desalted into 10 mM Hepes (pH 8.0), 10 mM NaCl. Amino acid sequences of cross-linked peptides were determined using an automated HPG1000A liquid phase protein sequencer. In the absence or the presence of ATPγS, control samples were irradiated in the presence of 1000 mM NaCl or 8 M urea, respectively. The controls were otherwise processed in parallel to cross-linked samples and yielded no amino acid sequence.

RESULTS

Cross-linking of the RecA Protein—Shortwave ultraviolet light promotes cross-linking of the RecA protein to polydeoxythymidylic acid as demonstrated directly by a decrease in electrophoretic mobility of poly(dT) under denaturing conditions (Fig. 1*A*). Consistent with product yield being contingent on the extent of cross-linking, the shift in polynucleotide mobility is dependent on the ultraviolet dosage and saturates over a range of 60 to 120 s of exposure (*lanes 1–6*). Sensitivity of the retarded species to proteolytic digestion by trypsin confirms the photochemical cross-linking of the RecA protein to poly(dT) (Fig. 1*B*). Although the active conformation of the RecA protein-ssDNA complex requires the binding of nucleotide cofactor, such as ATP, the RecA nucleoprotein filament is too dynamic for examination by cross-linking due to the potential for multiple ssDNA binding states that occur during the ssDNA-dependent hydrolysis of ATP (10, 28). However, use of the relatively nonhydrolyzable ATP analogue, ATPγS, permits separation of nucleotide binding and hydrolysis, thereby allowing for analysis of a stable form of the active RecA protein polymer assembled on ssDNA. Photochemical cross-linking between the

---

[2] The abbreviations used are: ATPγS, adenosine 5′-[γ-thio] triphosphate; ssDNA, single-stranded DNA.
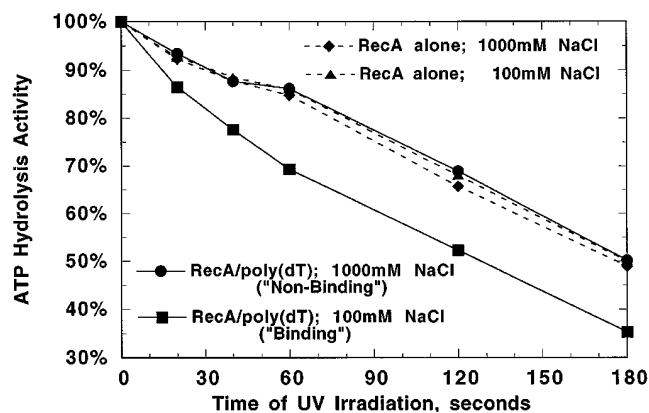
FIG. 2. **Effects of ultraviolet irradiation on the single-stranded DNA-dependent ATP hydrolysis activity of RecA protein.** The ATP hydrolysis activity was measured for 30 μl-aliquots withdrawn from irradiation reactions that contained 10 μM RecA protein, either in the absence or the presence of 70 μM poly(dT), and either 100 mM or 1000 mM NaCl (see legend in figure). The percentage of ATP hydrolysis activity was determined by normalizing the observed rates of ATP hydrolysis measured for irradiated samples to those of nonirradiated samples.

RecA protein and poly(dT) in the presence of ATPγS (500 μM) also yields complexes with both reduced mobility and protease sensitivity but requires longer irradiation times (~3 min) to reach a maximal extent (data not shown). This is not surprising because the effective intensity of incident light is most likely reduced due to absorption by the added nucleotide cofactor.

It is conceivable that photoactivated species that are generated in solution may collide to yield stable, yet spurious, covalent complexes; however, cross-linking to poly(dT) in either the absence or the presence of ATPγS is not observed when the RecA protein is denatured prior to irradiation, indicating that only the native RecA protein can be cross-linked (data not shown). To further address whether the observed cross-linking is due to a specific interaction between the RecA protein and the single-stranded polynucleotide, several control experiments were conducted. Similar to most protein-nucleic acid complexes (29), the stability of RecA protein-ssDNA complexes decreases with increasing salt concentrations (28). In agreement, photoinduced cross-linking of the RecA protein to poly(dT) is sensitive to the salt concentration as ultraviolet irradiation under conditions that prohibit complex formation (*i.e.* 1 M NaCl) fails to produce a species with reduced electrophoretic mobility (Fig. 1*A; lanes 7–12*). Conversely, cross-linking to poly(dT) is restored at 1 M NaCl in the presence of ATPγS (data not shown) due to the induction of the high affinity single-stranded DNA binding state of the RecA protein (28). Finally, photoadducts between the RecA protein and poly(dT) are not formed when either or both of the macromolecules are irradiated separately before mixing (data not shown).

Ultraviolet irradiation inhibits the single-stranded DNA-dependent ATP hydrolysis activity of the RecA protein in a dosage-dependent manner. As depicted in Fig. 2, equivalent amounts of "background" photoinactivation are observed when the RecA protein is irradiated either alone or in the presence of poly(dT) under conditions (1000 mM NaCl) that prohibit binding; however, inhibition is increased under conditions that permit formation and thus cross-linking of RecA protein-poly(dT) complexes. Therefore, photochemical cross-linking to polynucleotide results in a differential inhibition of the ATP hydrolysis activity of the RecA protein.

As shown in Fig. 3*A*, irradiation of RecA protein complexes formed with either poly(dT) or poly(dA) causes a differential inhibition of its ATP hydrolysis activity (defined as the inhibi-
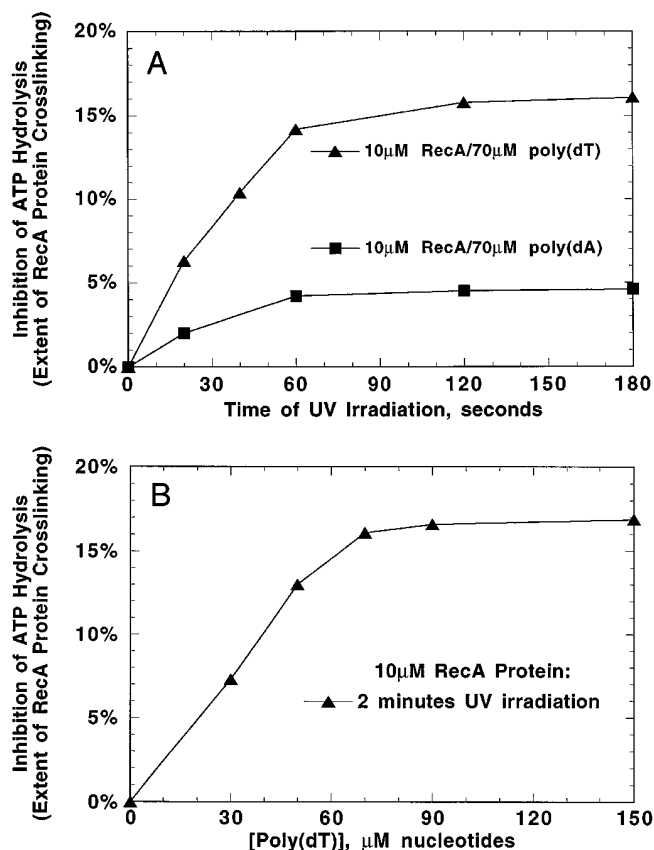

FIG. 3. **Differential inhibition of ATP hydrolysis activity due to photochemically induced cross-linking between RecA protein and polynucleotides.** ATP hydrolysis activity was measured, and differential inhibition was calculated as described under "Materials and Methods." *A*, the type of polynucleotide influences the extent of RecA protein cross-linking. Samples containing 10 μM RecA protein and 70 μM nucleotides of either poly(dT) (*triangles*) or poly(dA) (*squares*) were exposed to ultraviolet irradiation for the times indicated. *B*, the extent of RecA protein cross-linking determined as a function of the poly(dT) concentration. Samples comprised of 10 μM RecA protein and various amounts of poly(dT) (0, 30, 50, 70, 90, and 150 μM nucleotides) were irradiated with ultraviolet light for 2 min.

tion caused by ultraviolet light under conditions permitting complex formation (100 mM NaCl) minus that observed under conditions prohibiting complex formation (1000 mM NaCl)). In each case, the amount of ultraviolet exposure required to achieve maximal extents of differential inhibition (cross-linking) is approximately 70 s and is consistent with that estimated for poly(dT) using gel electrophoresis (Fig. 1*A*). Furthermore, the greater extent of RecA protein cross-linking to poly(dT) (~16%) relative to poly(dA) (~4%) agrees with previous studies that reveal thymine to be the most photoreactive nucleotide base (21). The quantitative relationship between photochemical cross-linking of the RecA protein and inhibition of its ATP hydrolysis activity is supported by chromatography studies that demonstrate that ~20% of the total RecA protein elutes from a strong anion exchange column associated with poly(dT) (data not shown). Therefore, these results indicate that formation of photoadducts between RecA protein and polynucleotides correlates to a loss in enzymatic activity. When studied as a function of poly(dT) concentration (Fig. 3*B*), the differential inhibition of ATP hydrolysis saturates at an apparent stoichiometry (7 ± 1 nucleotides per RecA protein monomer), which parallels that obtained using an assay that directly measures the binding of the RecA protein to single-stranded DNA (28, 30). Because ATPγS inhibits the ATP hydrolysis activity of the RecA protein and induces a high affinity ssDNA binding state

FIG. 4. **Regions of RecA protein photochemically cross-linked to poly(dT).** The three-dimensional structure of the RecA protein monomer is represented as a ribbon drawn along the peptide backbone (19). β-Strands within the major central domain are numbered *1–8* according to their order of occurrence in the primary sequence. Disordered loops between amino acids 157–164 (L1) and 195–209 (L2) are indicated by *dashed lines*. The position of ADP (*white*) is defined by diffusion into the crystal (20). Modeling was done using RasMol Molecular Visualization software (version 2.5). Tryptic peptides of the RecA protein covalently cross-linked to poly(dT), either in the absence or the presence of ATPγS, are labeled as follows: *green*, **Ile$_{61}$-Val-Glu-Ile-Tyr-Gly-Pro-Glu-Ser-Ser-Gly-Lys$_{72}$**; *red*, Leu$_{178}$-Ala-Gly-Asn-Leu-Lys$_{183}$; *blue*, **Glu$_{233}$-Gly-Glu-Asn-Val**-Val-Gly-Ser-Glu-Thr-Arg$_{243}$. The amino acid residues in bold type above were identified through amino acid sequencing.
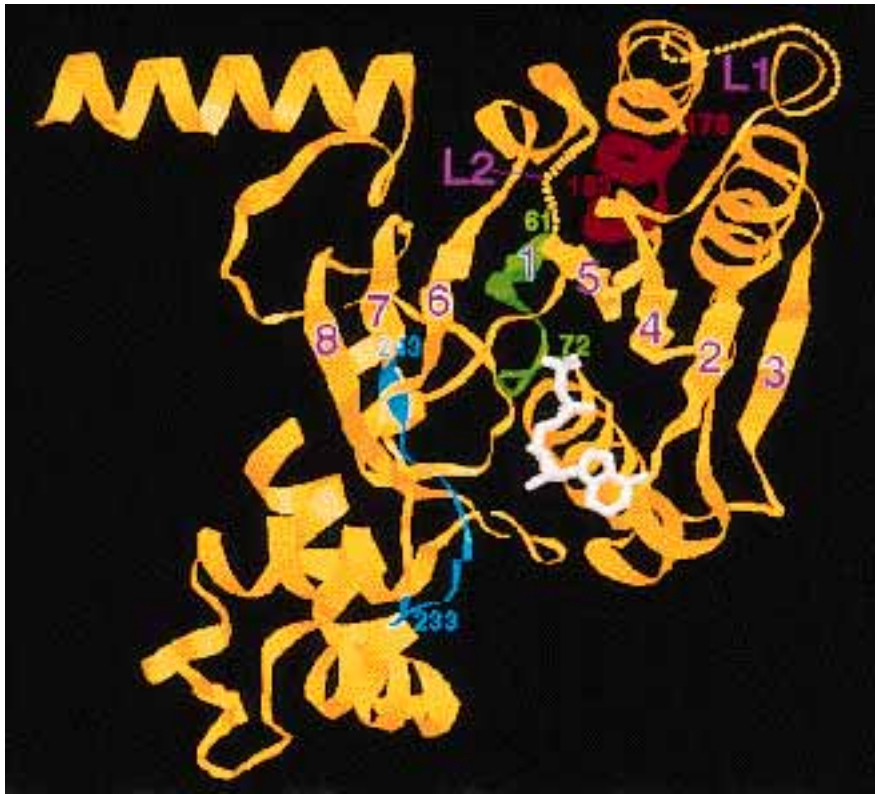
of the RecA protein stable out to concentrations of NaCl greater than 2 M, neither ATP hydrolysis nor ion exchange chromatography were used to characterize cross-linking of the RecA protein to poly(dT) in the presence of the relatively nonhydrolyzable cofactor.

Identification of Cross-linked RecA Peptides—Identification of regions of the RecA protein closely associated with single-stranded DNA was accomplished following extensive proteolytic digestion of cross-linked complexes with trypsin. Based on their covalent linkage to poly(dT), cross-linked peptides were isolated by size exclusion chromatography. The mixed amino acid sequences obtained were in approximately similar proportions and did not vary in the presence of ATPγS; they correspond to three unique tryptic peptides that spanned amino acids 61–72, 178–183, and 233–243 of the RecA protein (Fig. 4 legend). The cross-linking of lysine 183 is inferred by a dramatic reduction of the expected derivative in the sixth sequencing cycle; the ability to sequence beyond lysine 183 suggests that cross-linking not only inhibits cleavage by trypsin but involves the α-carbon or side chain constituents. Despite a distinct signal originating from amino acids 61–72, a specific site of cross-linking is uncertain and may result from multiple interactions within this region. The balance of sequence data matches the initial four to five residues of a tryptic peptide comprised by amino acids 233–243 of the RecA protein. The abrupt diminishment in repetitive yield of this sequence suggests that the peptide backbone is involved in cross-linking.

DISCUSSION

Comparison to the RecA Protein Crystal Structure—The crystal structure of the RecA protein has been solved at atomic resolution (19). The repeating unit cell is comprised of a single helical turn of six protein monomers (19) that forms a polymer structurally similar to images derived from electron microscopic examination of RecA protein-DNA filaments (31, 32). Flanked by two smaller subdomains at the amino and carboxyl termini, the major central domain of the RecA protein contains

the nucleotide binding domain identified by diffusion of ADP into crystals (20) and two regions of weak electron density, loop1 (L1) and loop2 (L2) (Fig. 4), that were proposed to be involved in DNA binding (19). Although not encompassed in either L1 or L2, the three regions identified through photochemical cross-linking to poly(dT) vertically traverse the major domain between these disordered loops (Fig. 4), suggesting one possible path for bound DNA. Though the disposition of DNA within the RecA protein filament is unknown, the location of amino acids 61–72 and 233–243 on the lower monomer surface within the interior of the polymer (Fig. 5, *A* and *C*) is consistent with both DNA residing near the axis of the filament (33, 34) and a most likely helical path for DNA binding (19, 35). In contrast, lysine 183 is distant from the polymer axis because it protrudes from the upper surface of the monomer that borders the deep groove formed between successive turns of the helical polymer (Fig. 5, *B* and *C*). Assuming that each monomer subunit equivalently contacts DNA, the path length defined along lysine 183 residues through one complete helical turn of the RecA protein polymer (Fig. 5*B*) cannot be easily accommodated by the internucleotide spacing of DNA within the nucleoprotein filament. One possible reservation is that because the x-ray structure is of an inactive, compact form of the RecA protein that was crystallized in the absence of both DNA and ATP (36), lysine 183 may be positioned differently within the active conformation of the protein. However, this explanation is not fully applicable, because specific cross-linking of lysine 183 to poly(dT), as well as the other peptides, is observed both in the absence and the presence of ATPγS, arguing that these same regions are proximal to and may interact with ssDNA despite the structural changes elicited throughout the RecA protein polymer by nucleotide cofactor binding. In agreement, lysine 183 is contained within a peptide that cross-links to oligo(dT)$_{14}$ (24). Thus, either the RecA protein monomer or polymer undergoes substantial structural alterations upon DNA binding (33) or a region proximal to bound DNA is located away from

FIG. 5. **Location of cross-linked regions within the RecA protein polymer.** A space-filling representation of a single turn of the helical polymer formed by six RecA protein monomers in the crystal structure (19, 20) is shown as viewed parallel to (both directions, *A* and *B*) and perpendicular to (*C*) the helical axis (RasMol, version 2.5). Alternating monomers of the RecA protein are colored *yellow* and *orange*, whereas ADP is *white*. Cross-linked amino acids 61–72, lysine 183, and 233–243 of the RecA protein are highlighted in *green*, *red*, and *blue*, respectively.
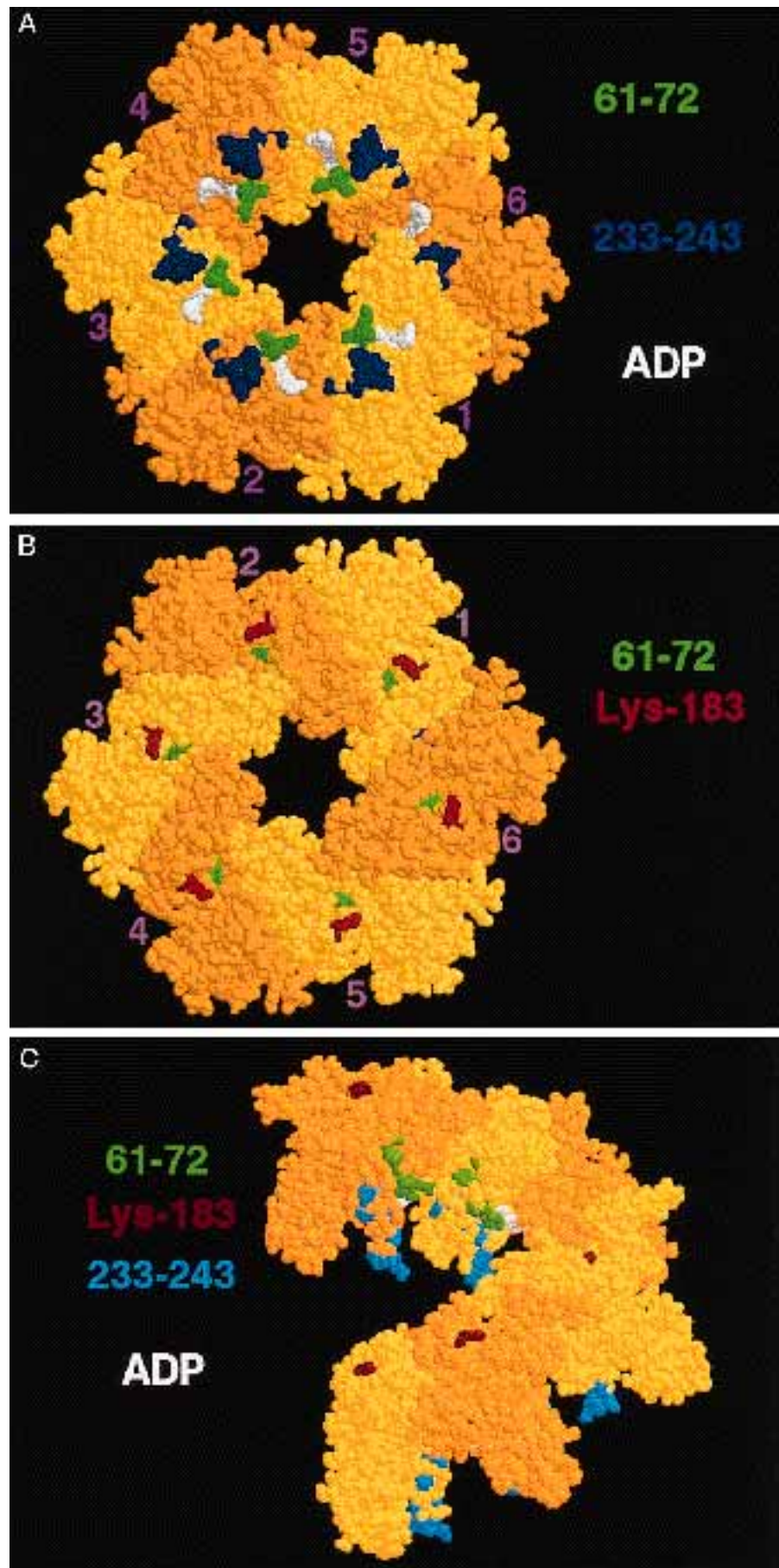
the helical axis; if the latter possibility is correct then not every monomer within the RecA nucleoprotein filament would necessarily interact with the same DNA molecule. Alternatively, the specific cross-linking of lysine 183 may represent the location of a secondary DNA binding site within the protein filament. Consequently, recognition between two DNA molecules may be mediated through binding sites comprised by the opposed surfaces that border the helical groove of the RecA protein polymer (Fig. 5, *A*, *B*, and *C*). The accessible nature of the groove within the RecA nucleoprotein filament may facilitate the transient and necessarily random contacts between DNA molecules that are central to the homology search process. Regardless of the specific binding mode for ssDNA, identification of indistinguishable patterns of cross-linking in both the absence or the presence of ATPγS indicates that these same regions of the RecA protein represent, at least in part, the binding site(s) for ssDNA whether the RecA nucleoprotein filament is in an inactive collapsed or an active extended conformation.

Comparison to Mutant RecA Proteins—In support of their assignment as regions that are in close proximity to and may mediate binding of DNA, mutations within or adjacent to cross-linked peptides affect both the specificity and affinity of the RecA protein for nucleic acids. Within the first peptide, the putative involvement of amino acids proline 67, glutamic acid 68, and serine 69 of the RecA protein in the interaction with DNA is consistent with these residues being highly conserved (7, 8) and being oriented toward the axis of the RecA protein filament (19) (Fig. 5, *A* and *C*); furthermore, extensive mutagenesis reveals that these positions have a unique and rather strict set of constraints needed to preserve functionality *in vivo*, leading to the proposal that these side chains are bound to DNA (37). Indeed, the RecAP67W (proline 67 to tryptophan) protein can be activated by DNA oligomers that normally do not support the formation of an active wild-type nucleoprotein filament (38). Preliminary analysis indicates that the enhanced activities displayed by the RecAP67W protein can be attributed to an increased affinity for single-stranded DNA.[3]

In the second peptide, a potential role for lysine 183 in the interaction with DNA is implicated by our photochemical cross-linking data and that of Morimatsu and Horii (24). A previous sequence comparison among a limited set of eubacterial RecA proteins implied that a positively charged residue at position 183 may be essential to function because substitutions for lysine are accompanied by the simultaneous change of the next amino acid to arginine or lysine (see Ref. 7 for compilation). Currently, a more extensive alignment of 62 RecA protein sequences from a diverse range of bacterial sources indicates that although lysine at residue 183 is not conserved, the most common amino acid at position 184 is either lysine or arginine (39); because exceptions to this conservation of amino acid 184 are most common within enterobacteria, it is conceivable that both structural and functional demands are preserved in the *E. coli* RecA protein by the presence of a positive charge at the preceding amino acid. Although such primary sequence relationships are suggestive, the fact that single substitutions for residues at positions 183 or 184 have dramatic effects on RecA protein-promoted activities substantiates the idea that this region is functionally significant. Substitution of lysine 183 by methionine confers a *recA⁻* phenotype and results in a protein that is diminished in both DNA binding and DNA-dependent ATP hydrolysis, is sensitive to inhibition by the E. coli single-stranded DNA binding protein, and is incapable of promoting

DNA strand exchange.[4] Moreover, mutation of the adjacent residue, glutamine 184, to lysine in the RecA1202 protein produces a variant that can be activated by either atypical nucleic acid substrates, such as rRNA and tRNA, or alternate nucleotide cofactors (40, 41). Located on an exterior surface that borders the helical groove of the RecA protein filament (Fig. 5*C*), amino acid residues 183 and 184 contribute to the intermolecular association between adjacent polymers within the crystal (19). This observation lead to the proposal that bundling of polymers may serve to regulate RecA protein binding to DNA (*i.e.* dissociation of bundles activates the RecA protein) (19). Although many aggregate forms of the RecA protein have been documented (42–44), the species fundamental to its interaction with DNA is unclear; nevertheless, specific cross-linking of lysine 183 within the interfilament contact region implies that binding to DNA is competitive with the bundling of a DNA-free form of the RecA protein.

Whereas initial residues (233–238) of the third identified region exhibit little amino acid conservation, glycine 239 is invariant among eubacterial RecA proteins (7, 8) and thus may allow for favorable interactions between DNA phosphates and polypeptide backbone amide NH groups. The terminal arginine residue (amino acid 243) is also conserved among eubacterial RecA proteins (7, 8), and its mutation to alanine results in a RecA protein with a lower apparent affinity for single-stranded DNA but that retains self-assembly properties indistinguishable from the wild-type protein (45). Furthermore, arginine 243 is included in a region (residues 243–310) of the RecA protein originally predicted to be important for interaction with DNA based on a primary sequence alignment among several proteins that bind cooperatively and with high affinity to single-stranded DNA (46). Based on structural analysis of filamentous bacteriophage gene 5 proteins (47, 48), a portion of this region within the RecA protein (residues 243–257) is proposed to be comprised of a set of antiparallel β-sheets that constitute a binding surface for single-stranded DNA, referred to as a "DNA binding wing." Whereas arginine 243 is located within one of the primarily parallel arranged β-sheets of the major central domain of the RecA protein (Fig. 4), the balance of this region (residues 247–257) is an interconnecting loop segment that contributes to the extensive interface formed between monomers in the RecA protein polymer (19). Thus, although it may include a portion of the DNA binding site of the RecA protein, any similarity of this region (residues 243–257) to the DNA binding wing of other single-stranded DNA binding proteins appears coincidental and is not manifest in conservation of structure.

Allosteric Mechanism—Nucleotide cofactors allosterically modulate both the stability (28, 49) and structure (12–14) of RecA protein-DNA complexes. ATP increases the affinity of the RecA protein for DNA while inducing an active extended nucleoprotein filament; conversely, ADP produces a low affinity binding state of the RecA protein and a compact inactive filament (10). Consequently, the DNA-dependent ATP hydrolysis activity of the RecA protein regulates its binding to and dissociation from DNA. Moreover, the utilization of ATP by the RecA protein bears a mechanistic resemblance to that of both translation elongation factor Tu and the signal-transducing ras p21 protein of the G-protein family in that NTP hydrolysis serves as a molecular switch that allows conversion between active (ATP-bound) and inactive (ADP-bound) conformations. Structural determination reveals that each of these diverse proteins contains a nucleotide binding fold comprised in part of a structurally conserved phosphate binding loop (20) or "A" site motif

---

[3] W. M. Rehrauer and S. C. Kowalczykowski, unpublished observations.

[4] E. Zaitsev and S. C. Kowalczykowski, personal communication.

DNA Binding Site(s) of the E. coli RecA Protein

having the consensus sequence (G/A)X_4GK(T/S) (50). In the case of the ras p21 (51–53) and EF-Tu proteins (54–56), the nature of the conformational change induced by the exchange of GTP for GDP was established by comparing the crystal structures of both active and inactive conformations, respectively. In general, structural changes, elicited by direct contacts with the terminal phosphate of bound GTP, propagate to other effector regions within the molecule. Story and Steitz (20) have advanced an analogous allosteric mechanism for the RecA protein in which glutamine 194 (terminal residue in β-sheet 5; Fig. 4) interacts with the γ phosphate of bound ATP to stabilize a conformation of the disordered loop L2 (residues 195–209) that has a high affinity for DNA. Although no peptides corresponding to the L2 region of RecA protein are detected in cross-linking to polydT, our results are consistent with and support the proposal based on the crystal structure. One of the peptides of the RecA protein identified by photochemical cross-linking (amino acids 61–72) coincides with the A site motif ($_{66}$GPESSGKT$_{73}$) and is situated very close in the monomer structure to both the disordered L2 region and the immediately preceding residue, glutamine 194. Consequently, the structural proximity of regions functioning in the interaction with the pyrophosphate moiety of a bound nucleotide cofactor and DNA may not only account for the correlation between the extent of RecA protein cross-linking and inhibition of ATP hydrolysis activity but also provides mechanistic insight into the allosteric properties exhibited by the RecA nucleoprotein filament. It is readily apparent that changes in interactions between either the phosphate binding loop (amino acids 66–73, specifically lysine 72 (57)) and glutamine 194 (20) with bound nucleotide cofactor that are elicited through the binding and hydrolysis of ATP could mediate a set of different interactions between regions of the RecA protein (amino acids 61–72 and the disordered L2 region) and DNA. Such a direct mechanism for allosteric regulation of DNA binding affinity may have general applicability to many other NTP-dependent nucleic acid binding proteins (58).

REFERENCES

1. Clark, A. J., and Margulies, A. D. (1965) *Proc. Natl. Acad. Sci. U. S. A.* **53**, 451–459
2. Clark, A. J. (1973) *Annu. Rev. Genet.* **7**, 67–86
3. Shibata, T., Cunningham, R. P., DasGupta, C., and Radding, C. M. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5100–5104
4. McEntee, K., Weinstock, G. M., and Lehman, I. R. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 2615–2619
5. Cassuto, E., West, S. C., Mursalim, J., Conlon, S., and Howard-Flanders, P. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3962–3966
6. Cox, M. M., and Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6018–6022
7. Roca, A. I., and Cox, M. M. (1990) *Crit. Rev. Biochem. Mol. Biol.* **25**, 415–456
8. Story, R. M., Bishop, D. K., Kleckner, N., and Steitz, T. A. (1993) *Science* **259**, 1892–1896
9. Kowalczykowski, S. C., and Eggleston, A. K. (1994) *Annu. Rev. Biochem.* **63**, 991-1043
10. Kowalczykowski, S. C. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20**, 539–575
11. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D., and Rehrauer, W. M. (1994) *Microbiol. Rev.* **58**, 401–465
12. Stasiak, A., and Egelman, E. H. (1988) in *Genetic Recombination* (Kucherlapati, R., and Smith, G. R., eds) pp. 265–308, American Society for Microbiology, Washington, D. C.
13. Heuser, J., and Griffith, J. (1989) *J. Mol. Biol.* **210**, 473–484
14. DiCapua, E., Schnarr, M., Ruigrok, R. W., Lindner, P., and Timmins, P. A. (1990) *J. Mol. Biol.* **214**, 557–570
15. Yu, X., and Egelman, E. H. (1993) *J. Mol. Biol.* **232**, 1–4
16. Ogawa, T., Yu, X., Shinohara, A., and Egelman, E. H. (1993) *Science* **259**, 1896–1899
17. Yonesaki, T., and Minagawa, T. (1985) *EMBO J.* **4**, 3321–3327
18. Sung, P. (1994) *Science* **265**, 1241–1243
19. Story, R. M., Weber, I. T., and Steitz, T. A. (1992) *Nature* **355**, 318–325
20. Story, R. M., and Steitz, T. A. (1992) *Nature* **355**, 374–376
21. Shetlar, M. D. (1980) *Photochem. Photobiol. Rev.* **5**, 105–197
22. Williams, K. R., and Konigsberg, W. H. (1991) *Methods Enzymol.* **208**, 516–539
23. Shetlar, M. D., Christensen, J., and Hom, K. (1984) *Photochem. Photobiol.* **39**, 125–133
24. Morimatsu, K., and Horii, T. (1995) *Eur. J. Biochem.* **228**, 772–778
25. Uhlin, B. E., and Clark, A. J. (1981) *J. Bacteriol.* **148**, 386–390
26. Griffith, J., and Shores, C. G. (1985) *Biochemistry* **24**, 158–162
27. Kowalczykowski, S. C., and Krupp, R. A. (1987) *J. Mol. Biol.* **193**, 97–113
28. Menetski, J. P., and Kowalczykowski, S. C. (1985) *J. Mol. Biol.* **181**, 281–295
29. Record, M. T., Jr., Anderson, C. F., and Lohman, T. M. (1978) *Q. Rev. Biophys.* **11**, 103–178
30. Lauder, S. D., and Kowalczykowski, S. C. (1991) *J. Biol. Chem.* **266**, 5450–5458
31. Egelman, E. H., and Stasiak, A. (1986) *J. Mol. Biol.* **191**, 677–697
32. Yu, X., and Egelman, E. H. (1990) *Biophys. J.* **57**, 555–566
33. DiCapua, E., Schnarr, M., and Timmins, P. A. (1989) *Biochemistry* **28**, 3287–3292
34. Egelman, E. H., and Yu, X. (1989) *Science* **245**, 404–407
35. Stasiak, A., and Di Capua, E. (1982) *Nature* **299**, 185–186
36. Yu, X., and Egelman, E. H. (1992) *J. Mol. Biol.* **227**, 334–346
37. Konola, J. T., Logan, K. M., and Knight, K. L. (1994) *J. Mol. Biol.* **237**, 20–34
38. Konola, J. T., Nastri, H. G., Logan, K. M., and Knight, K. L. (1995) *J. Biol. Chem.* **270**, 8411–8419
39. Karlin, S., and Brocchieri, L. (1996) *J. Bacteriol.*, in press
40. Wang, W.-B., and Tessman, E. S. (1986) *J. Bacteriol.* **168**, 901–910
41. Wang, W.-B., Tessman, E. S., and Tessman, I. (1988) *J. Bacteriol.* **170**, 4823–4827
42. Brenner, S. L., Zlotnick, A., and Griffith, J. D. (1988) *J. Mol. Biol.* **204**, 959–972
43. Wilson, D. H., and Benight, A. S. (1990) *J. Biol. Chem.* **265**, 7351–7359
44. Benight, A. S., Wilson, D. H., Budzynski, D. M., and Goldstein, R. F. (1991) *Biochimie (Paris)* **73**, 143–155
45. Morimatsu, K., and Horii, T. (1994) *Adv. Biophys.* **31**, 23–48
46. Prasad, B. V. V., and Chiu, W. (1987) *J. Mol. Biol.* **193**, 579–584
47. Jong, E. A. M. D., Duynhoven, J. P. M. V., Harmsen, B. J. M., Tesser, G. I., Konings, R. N. H., and Hilbers, C. W. (1989) *J. Mol. Biol.* **206**, 133–152
48. Skinner, M. M., Zhang, H., Leschnitzer, D. H., Guan, Y., Bellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N., Wang, A. H., and Terwilliger, T. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2071–2075
49. Menetski, J. P., Varghese, A., and Kowalczykowski, S. C. (1988) *Biochemistry* **27**, 1205–1212
50. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J.* **1**, 945–951
51. Pai, E. F., Kabsch, W., Krengel, U., Holmes, K. C., John, J., and Wittinghofer, A. (1989) *Nature* **341**, 209–214
52. Brunger, A. T., Milburn, M. V., Tong, L., deVos, A. M., Jancarik, J., Yamaizumi, Z., Nishimura, S., Ohtsuka, E., and Kim, S. H. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4849–4853
53. Tong, L. A., de Vos, A. M., Milburn, M. V., and Kim, S. H. (1991) *J. Mol. Biol.* **217**, 503–516
54. Jurnak, F. (1985) *Science* **230**, 32–36
55. Berchtold, H., Reshetnikova, L., Reiser, C. O. A., Schirmer, N. K., Sprinl, M., and Hilgenfeld, R. (1993) *Nature* **365**, 126–132
56. Kjeldgaard, M., and Nyborg, J. (1992) *J. Mol. Biol.* **223**, 721–742
57. Rehrauer, W. M., and Kowalczykowski, S. C. (1993) *J. Biol. Chem.* **268**, 1292–1297
58. Kowalczykowski, S. C., and Krupp, R. A. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3478–3482