Differential Proximity Probing of Two DNA Binding Sites in the *Escherichia coli* RecA Protein Using Photo-Cross-Linking Methods

Yi Wang* and Kenji Adzuma

The Rockefeller University, Box 256, New York, New York 10021

Received October 12, 1995; Revised Manuscript Received January 17, 1996[⊗]

ABSTRACT: The DNA strand-exchange reaction catalyzed by the *Escherichia coli* RecA protein occurs between the two DNA binding sites that are functionally distinct. Site I is the site to which a DNA molecule (normally single-stranded DNA) binds first; this first binding makes site II available for additional DNA-binding (normally double-stranded DNA). Photo-cross-linking was employed to identify the amino acid residues located close to the bound DNA molecule(s). A ssDNA oligo containing multiple 5-iodouracil residues (IdU) was cross-linked to RecA by irradiation with a XeCl pulse laser (308 nm), and the cross-linked peptides were purified and sequenced. To differentiate the two DNA binding sites, we used two protocols for making RecA−ssDNA complexes: (1) IdU-containing oligo was mixed with a stoichiometric excess of RecA, a condition which favors the binding of the oligo to site I, and (2) RecA was first allowed to bind to a nonphotoreactive oligo and then chased with the IdU-containing oligo, a condition which favors the binding of the IdU-oligo to site II. We observed that when RecA was in excess (site I probing), cross-linking occurred to Met-164 which is located in the disordered loop 1 of the RecA crystal structure [Story, R. M., Weber, I. T., & Steitz, T. A. (1992) *Nature 355,* 318−325]. When site II was probed, the majority of cross-linking occurred to Met-202 or Phe-203, located in loop 2. These results support the idea that, as predicted by Story and co-workers (1992), the disordered loops are involved in DNA binding. The results also suggest that the two sites are not only functionally but also physically distinct.

An early step in homologous recombination involves the pairing of two DNA molecules that share a similar sequence and the exchange of DNA strands in between. The recA protein of *Escherichia coli* (RecA) catalyzes this DNA strand-exchange reaction, normally between a single-stranded DNA (ssDNA) and a double-stranded DNA (dsDNA). This reaction occurs in three steps [for reviews, see Cox and Lehman (1987), Kowalczykowski (1991), and Kowalczykowski and Eggleston (1994)]. RecA first polymerizes onto the ssDNA in the presence of ATP, producing a nucleoprotein filament called a presynaptic complex. This complex then captures a dsDNA whose sequence is nearly identical, or "homologous", to the resident ssDNA. The resulting complex is called a synaptic complex, in which all three DNA strands are held close to each other. The final step is the release of a pair of strand-exchange products, the displaced strand and a heteroduplex. Product release requires the hydrolysis of ATP, but the two processes are not necessarily coupled (Menetski & Kowalczykowski, 1987; Rosselli & Stasiak, 1990).

RecA is thought to have two kinds of DNA binding sites, site I and site II (Takahashi et al., 1989; Müller et al., 1990; Zlotnick et al., 1990). Site I is defined as the site to which a DNA molecule binds first. The binding stoichiometry is three nucleotides per strand per RecA monomer (i.e., three bases/RecA monomer for ssDNA binding, three base pairs/RecA monomer for dsDNA binding). When an increasing amount of ssDNA is added to a fixed amount of RecA, the complex holding only one molecule of ssDNA predominates until the stoichiometry of DNA nucleotides versus RecA reaches 3:1 (Takahashi et al., 1989; Zlotnick et al., 1993); beyond this stoichiometry, RecA starts to bind the second ssDNA molecule. The maximal stoichiometry is 6:1. When RecA is added to a mixture of ssDNA and dsDNA, site I becomes occupied by the ssDNA because binding of RecA to ssDNA is much faster than to dsDNA (Pugh & Cox, 1987); this is the condition where the strand-exchange reaction normally occurs. However, each site can in principle accommodate either ssDNA or dsDNA, depending on the order of addition. Note that the definition of these two DNA-binding sites is a functional one, referring merely to their order of occupancy, and does not necessarily imply that the two sites are physically distinct.

The crystal structure of RecA polymer, without DNA or ATP, was solved by Story et al. (1992). The structure is a helical filament with a 6_1 symmetry and shows a good resemblance to the low-resolution structure of RecA−DNA−ATPγS complex which was derived from reconstruction of the electron-microscopic images (Egelman & Stasiak, 1986). The crystal structure contains two disordered loops, L1 (residues 156−165) and L2 (195−209), that are located close to the helical axis of the RecA polymer. Because DNA molecules within a RecA filament must reside close to the axis of the filament (Egelman & Stasiak, 1986), these disordered loops were proposed to be involved in DNA binding (Story et al., 1992).

Owing to their intrinsic heterogeneity, structural characterization of the complex formed between nonspecific DNA binding proteins, such as RecA, and their target has been a challenge to X-ray crystallography and NMR techniques. Photo-cross-linking can be used for probing the contact points between DNA and protein and may offer important informa-

* Corresponding author. Phone: (212) 327-7471. Fax: (212) 327-7183. E-mail: wangy@rockvax.rockefeller.edu.

[⊗] Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0006-2960/96/0435-3563$12.00/0

tion on the structures of macromolecular complexes for which direct structural analyses are difficult (Williams & Konigsberg, 1991). As a cross-linker, halogenated bases such as 5-iodouracil (IdU) have many advantages. They are isosteric to thymine (5-methyluracil); incorporation of a halogenated base in place of thymine is expected to cause little structural perturbation. Also, halogenated bases can be photoactivated by long-wavelength (>300 nm) UV light. Since proteins and natural DNA bases absorb little at wavelengths longer than 300 nm, photoactivation specific to the halogenated base can be attained without causing significant photodamage to proteins and DNA. In combination with laser, which delivers high-density monochromic photons, highly efficient cross-linking specific to the halogenated base can be accomplished within a very short time period (Willis et al., 1993).

In this study, we analyzed the structure of RecA−ssDNA complexes by a photo-cross-linking method employing laser irradiation and a halogenated base IdU as the cross-linker. By placing the IdU-containing ssDNA at either site I or site II, we differentially probed the locations of the first and second DNA binding sites in RecA.

MATERIALS AND METHODS

DNA Oligos. DNA oligos used in this study are 40 nucleotides in length and were chemically synthesized using standard phosphoramidite chemistry on an ABI (Applied Biosystems Inc) 392 DNA synthesizer. The regular, non-photoreactive oligo (called dT-oligo hereafter) has an arbitrary sequence: ATAACTTGTGAGGTCATCGGGCG-CACTAGTGTACATGCGC. The photoreactive oligo (called IdU-oligo hereafter) carries 5-iodouracil (IdU) in place of thymine residues (total of 10) in the same sequence and was synthesized using 5-iododeoxyuridine cyanoethyl phosphoramidite (Glen Research). To reduce possible deiodination of the IdU, the final deprotection step was carried out at room temperature for 24 h as suggested by the manufacturer. All the oligos were purified through sequencing gel electrophoresis. Some oligo preparations were labeled with ^{32}P at the 5′-termini using T4 polynucleotide kinase and [γ-^{32}P]-ATP (3000 Ci/mmol), followed by removal of unincorporated ATP using a spin column (Quick Spin G-25, Boehringer Mannheim). DNA concentrations (nucleotide molar, M) was calculated using an extinction coefficient of 1.0×10^4/M/cm at 260 nm.

Enzymes. RecA was purified from the overproduction system (strain JC12772) established by Uhlin and Clark (1981), according to Cotterill et al. (1982) with minor modifications. The final fractions were dialyzed against the RecA dilution buffer: 20 mM Tris-acetate (pH 7.5 at 1 M at room temperature), 0.1 mM EDTA,[1] 1 mM dithiothreitol, and 10% glycerol. The RecA concentration was calculated using an extinction coefficient of 2.16×10^4 M^{-1} cm^{-1} at 278 nm in the RecA dilution buffer.

Sources of other enzymes used in this study are as follows: restriction enzymes and T4 polynucleotide kinase, New England Biolabs; sequencing-grade trypsin (cleaves the C-termini of Lys and Arg), Boehringer Mannheim; sequencing-grade endoprotease Lys-C (preferentially cuts Lys but also Arg), Promega; endoproteinase Glu-C (Glu and Asp), Boehringer Mannheim.

Formation of RecA−ssDNA Complex. The standard reaction mixture contained 40 mM Tris-acetate (pH 7.5), 100 mM potassium acetate, 9 mM magnesium acetate, 1 mM dithiothreitol, 250 μM ATPγS, and 100 μg of bovine serum albumin (BSA) per mL. Incubation was carried out at 37 °C for 40 min unless otherwise noted. For site I probing, the preparative scale sample was made by adding 160 nmol of RecA (final concentration of 20 μM) to 160 nmol (in nucleotide) of IdU-containing oligo (20 μM). For site II probing, 100 nmol of RecA (10 μM) was first preincubated with 400 nmol of regular, thymine-containing oligo (40 μM) for 20 min; the mixture was then added with 400 nmol of IdU-oligo (40 μM), and the incubation was continued for an additional 20 min. Concentrations of RecA and ssDNA used in other experiments are described in the figure legends or text.

Photo-Cross-Linking. Photo-cross-linking was carried out with a XeCl excimer laser (308 nm, Lambda Physik EMG-50). The laser beam of 10 mm × 26 mm was directed unfocused through a quartz prism to the sample placed on an ice bath. The laser was operated in the power density range of 150−250 mW/cm^2 with frequencies of 5−10 Hz. For analytical scale, 10−20 μL samples were placed on a glass plate covered with parafilm. For preparative scale, 1-mL aliquots were placed in polypropylene tubes (1.5 cm i.d. × 5 cm height) with a gentle magnetic stirring. The total irradiation time for preparative cross-linking was about 2 min, which corresponds to 180 mW/cm^2. Note that the laser is a pulse laser with a pulse duration of only about 10 ns; the nominal irradiation time of 2 min means that the actual exposure time was only about 12 μs at 10 Hz (pulses/s).

Isolation of the Cross-Linked Peptides. The cross-linked samples were denatured at 60 °C for 20 min in the protease digestion buffer prior to digestion. Trypsin digestion was conducted in 0.1 M Tris-HCl (pH 8.5) and 0.01% SDS. Trypsin was added to a final of 1/10 (w/w) of the total RecA amount, followed by incubation at 37 °C for 10 h. The Lys-C digestion was carried out with 1/10 (w/w) of the total RecA, in 25 mM Tris-HCl (pH 8.5), 1 mM EDTA, and 0.01% SDS at 37 °C for 12 h. The Glu-C (protease V8) digestion was carried out with 1/10 (w/w) of the total RecA, in 50 mM potassium phosphate (pH 7.8) and 0.01% SDS at 37 °C for 20 h.

All the subsequent steps were carried out at 4 °C. The digested sample in two aliquots was injected into a Mono Q column (1 mL bed, Pharmacia LKB) equilibrated with 20 mM Tris-HCl (pH 7.5) and washed with the same buffer. The bound species, which contained free ssDNA and the cross-linked peptides, were eluted with 0.8 M KCl. The eluted fractions were combined (~4 mL total) and concentrated using Centricon-3 (Amicon) units. The concentrated sample was washed twice with aliquots of HPLC grade water on the Centricon units and further concentrated to a final volume of 500 μL. The sample was then lyophilized, dissolved in sequencing dye, and loaded on a 10% urea−polyacrylamide (sequencing) gel. The bands corresponding to each of the cross-linked peptides were identified by a brief exposure to an X-ray film, and the peptides were separately eluted into 3 mL of 0.1 M ammonium bicarbonate (pH 8.0)

[1] Abbreviations: ATPγS, adenosine 5′-*O*-(3-thiotriphosphate); PVDF, polyvinylidene difluoride; HPLC, high-performance liquid chromatography; EDTA, ethylenediaminetetraacetate; SDS, sodium dodecyl sulfate.

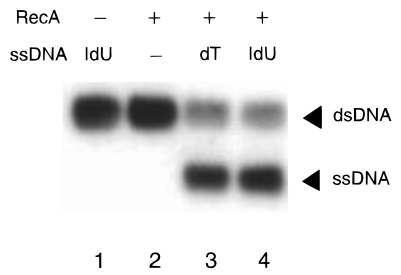

FIGURE 1: Comparison between normal oligo and IdU-substituted oligo in DNA synapsis. The synapsis reaction was carried out as described in Materials and Methods, in the buffer containing 250 μM ATPγS, 1 μM RecA, 1 μM (in nucleotide) 40-mer ssDNA oligo, 1 μM 40-base pair dsDNA, and 20 μM sonicated *E. coli* genomic DNA. The ssDNA oligo was either a regular oligo (labeled as dT) or IdU-substituted oligo (IdU). One of the strands of the 40-base pair dsDNA is identical to the dT-oligo and was labeled with ^{32}P. After a 30-min incubation at 37 °C, a dye mixture containing SDS was added to the final concentration of 0.5%, and the samples were separated by electrophoresis on an 8% polyacrylamide gel. The resulting autoradiogram is shown: lane 1, without RecA; lane 2, without ssDNA oligo; lane 3, reaction with regular dT-oligo; lane 4, reaction with IdU-substituted oligo.

at room temperature for 24 h. Each eluted peptide was directly injected to a reverse-phase column (PartiSphere C_{18}, Whatman) equilibrated with 50 mM potassium phosphate (pH 7.0) and eluted with a linear gradient of acetonitrile. The absorbance (214 nm) peak that was also radioactive was pooled and lyophilized. The peptide samples were then dissolved in water and immobilized onto a positively charged PVDF membrane (Immobilon N, Millipore) using a Bio-Rad microfiltration apparatus.

Throughout the purification, quantities of peptide−DNA cross-linked samples were estimated from the radioactivity. The amounts of final peptide samples varied from 20 to 250 pmol in molecules. Total recoveries of the cross-linked peptides in site-I and site-II probing experiments were about 10% and 20%, respectively, with some loss attributable to spontaneous degradation of the cross-linked peptides. There were no particular steps at which a large loss (>50%) occurred. The amino acid sequences of the cross-linked peptides were analyzed by an ABI 470 gas-phase peptide sequencer.

RESULTS

IdU-Containing Oligo Is as Efficient a Substrate as Normal Oligo in the DNA Synapsis Reaction. The photoreactive oligo used in this study was a 40-mer DNA oligo with an arbitrary sequence carrying 10 IdU residues. To examine whether the IdU substitution has any deleterious effect, we first compared the IdU-oligo with the normal oligo in the DNA synapsis reaction.

Each of IdU oligo and normal oligo was mixed with a homologous dsDNA in the presence of ATPγS and RecA. In this dsDNA, one of the strands was identical to the ssDNA oligo and was ^{32}P-labeled. The DNA synapsis reaction would eventually lead, after removal of RecA, to the displacement of the labeled strand which migrates faster than the original dsDNA in a polyacrylamide gel. As shown in Figure 1 (compare lane 3 with lane 4), the efficiencies of DNA synapsis as measured by the displacement of the labeled strand were virtually identical between the IdU oligo and normal oligo.

Efficient Photo-Cross-Linking Requires the Presence of RecA, IdU, and UV Irradiation. We compared the time dependence of the cross-linking reaction between IdU-containing oligo and normal oligo. The RecA−ssDNA complex was formed by incubating each oligo, labeled with ^{32}P, with RecA. The mixing ratio was 1:1 (DNA bases/RecA molecule), which is a stoichiometric excess of RecA over ssDNA. This ratio is expected to favor the formation of the RecA complex in which only one molecule of ssDNA is bound. The mixture was irradiated with a XeCl laser (308 nm) and the products were analyzed by 10% SDS−PAGE.

The irradiation produced a highly efficient cross-linking that was specific to the IdU oligo. After 15-s irradiation, we observed one prominent and several faint bands, each of which showed a slower mobility than those of the original oligo and of RecA (lane 2, Figure 2a). As the irradiation time increased, the total intensities of these new bands showed a steady increase while the intensity of free oligo decreased until almost none was left at 10 min irradiation. Formation of these new bands required the presence of RecA (lane 10); all of these new bands disappeared after treatment with protease K (data not shown). These results show that the bands represent RecA−DNA cross-linking rather than DNA−DNA cross-linking. The oligo that carries thymine residues instead of IdU produced only a very small amount of crosslinked products even after 10-min irradiation (lanes 7−9).

Quantification (data not shown) of the intensities of individual bands revealed two types of cross-linked products. The intensity of the one type reached a maximum at 2-min and then gradually decreased; this species is probably the one in which a single molecule of RecA was cross-linked to the oligo (indicated by "single XL" in Figure 2a). The other set of bands, which migrated slower than the singly-cross-linked product, displayed a constant increase in their intensities over time; these species probably reflect the cross-linking between a single DNA molecule and multiple RecA molecules ("multiple XL" in Figure 2a). Multiple cross-linking can happen because the 40-mer oligo contains 10 IdU residues and the RecA−oligo complex may contain up to 13 molecules of RecA per oligo.

The same samples as in Figure 2a were digested with trypsin and the products were analyzed by sequencing gel electrophoresis. The trypsin digestion of the IdU-cross-linked samples produced, over a smeary background, several bands that migrated slower than free DNA (Figure 2b). Comparison of the kinetics and intensities of these bands with the bands shown in Figure 2a suggests that the singly-cross-linked RecA−DNA species migrates as a cluster of bands in the sequencing gel (indicated by a bracket in Figure 2b). Other bands which migrated slower than the singly-cross-linked species (indicated by triangles) are probably the multiple cross-linked products.

Site I and Site II Can Be Probed Differentially. The experiments described above used a stoichiometric excess of RecA over the IdU oligo. Under these conditions, RecA−ssDNA complex in which DNA is bound only at site I is expected to predominate over the one in which both site I and site II are occupied. This expectation, however, must be treated with a caution. The main reason is kinetic; the second ssDNA molecule bound at site II appears to be kinetically inert in the presence of ATPγS (Takahashi et al., 1989; Zlotnick et al., 1993). Thus, once the doubly-bound

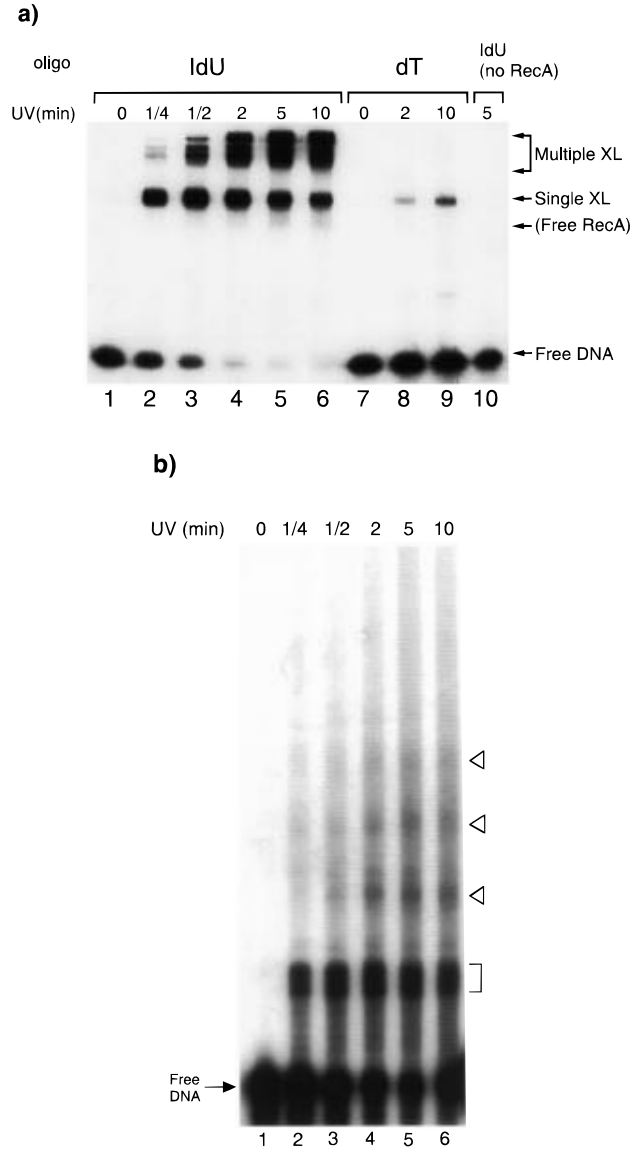

FIGURE 2: Time course of photo-cross-linking reaction. RecA−ssDNA complex was formed by incubating 5 μM RecA with 5 μM ^{32}P-labeled oligo containing either IdU or thymine residues. The mixtures were irradiated with a XeCl laser at 308 nm for the indicated time periods. (a) Autoradiogram of the irradiated samples analyzed by 10% SDS−PAGE. Positions of cross-linked species (XL), free DNA, and un-cross-linked RecA are indicated (RecA is not visible in the autoradiogram). Lanes 1−6, samples with IdU-oligo; lanes 7−9, with dT-oligo; lane 10, with IdU-oligo but without RecA. (b) Autoradiogram of the irradiated samples which were subsequently digested by trypsin and analyzed by 10% sequencing gel electrophoresis. Samples in each lane were derived from the samples shown in corresponding lanes of panel a. Bracket indicates the position of singly cross-linked product(s); triangles, positions of multiply cross-linked products.

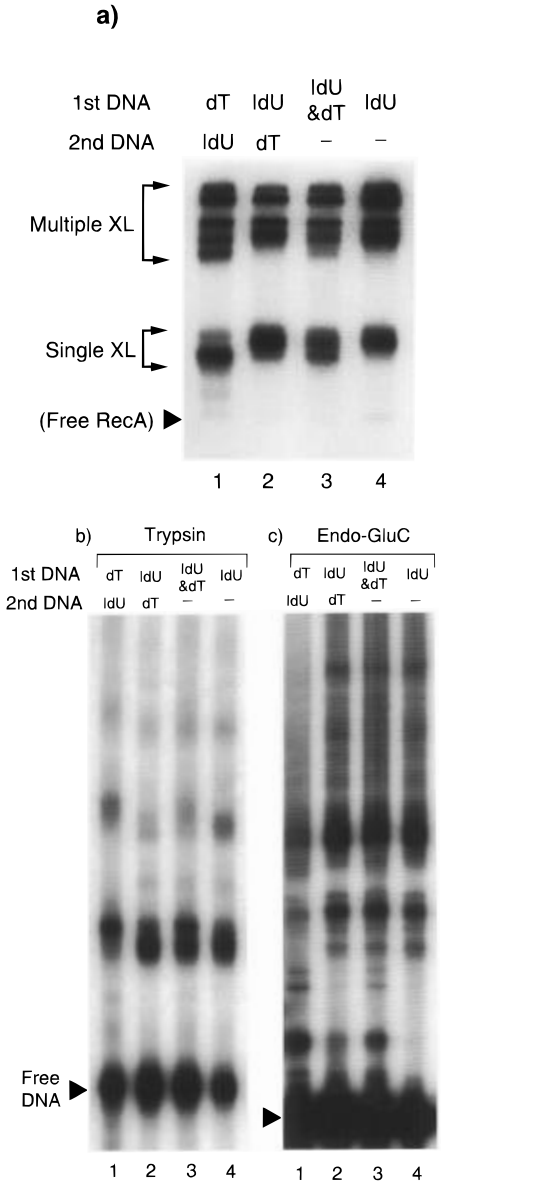

FIGURE 3: Differential cross-linking. All RecA-DNA complexes were prepared with ssDNA oligos. Cross-linking to site I: 5 μM (final concentration) of RecA was incubated with 5 μM of ^{32}P-labeled IdU-containing oligo in the presence of ATPγS for 20 min. A portion of the sample was then irradiated by UV (lane 4 in all three panels). Unlabeled dT-oligo (35 μM) was added to the remaining portion; the mixture was incubated for an additional 20 min and then irradiated (lane 2). Cross-linking to site II: 5 μM of RecA was first incubated with 20 μM unlabeled dT-oligo in the presence of ATPγS for 20 min. ^{32}P-labeled IdU-oligo (20 μM) was then added, and the mixture was incubated for an additional 20 min and then irradiated by UV (lane 1). Cross-linking to both site I and site II: 5 μM of RecA was incubated with a mixture of 20 μM each ^{32}P-labeled IdU-oligo and unlabeled dT-oligo in the presence of ATPγS for 20 min and then irradiated (lane 3). Analyses: a portion of the irradiated samples were directly loaded to 10% SDS−PAGE (a). The remaining portions were digested with either trypsin (b) or Glu-C (c); the digested samples were analyzed by 10% sequencing gel electrophoresis. All three panels show the autoradiograms.

complex was made, it would likely remain as doubly-bound even if there are free RecA molecules. Under these nonequilibrium conditions, there is no a priori guarantee that a singly-bound complex should predominate just because there are excess of free RecA. We therefore carried out the following experiments to directly analyze the differential binding of the two sites.

We first examined the kinetic stability of ssDNA molecule bound at site I. A ssDNA−RecA complex was formed by incubating ^{32}P-labeled IdU oligo with an excess of RecA (5 μM each, 1:1 stoichiometry) in the presence of ATPγS. This complex was then chased with an unlabeled, regular oligo, the concentration of which (35 μM) should be sufficient to saturate all the remaining DNA-binding sites of RecA. After further incubation, the samples were irradiated by UV laser and the products were analyzed by SDS−PAGE. As shown in Figure 3a (compare lane 2 with lane 4), adding an un-

cross-linkable oligo to the complex pre-formed with the IdU-oligo did not reduce the cross-linking efficiency at all; the pattern of cross-linking was also identical to the one without chase. These results show that the RecA−ssDNA complex formed during the first incubation is kinetically inert and the DNA molecules bound at site I are practically immobile under these conditions. Importantly, the results also show that binding of a second ssDNA does not interfere with the cross-linking of the ssDNA molecule bound at site I.

We next reversed the order of addition of the two oligos: unlabeled, regular ssDNA oligo first, followed by ^{32}P-labeled IdU oligo. The conditions were such that the first un-cross-linkable oligo should have saturated site I and site II should have been the only site available for the IdU oligo to bind (see Figure 3 legend for details). The resulting cross-linking pattern (lane 1 in Figure 3a) was clearly different from the one in which IdU oligo should have been bound predominantly at site I (lane 2 or lane 4). When RecA was added to a 1:1 mixture of IdU and regular oligos, the cross-linking pattern (lane 3) was essentially a 1:1 composite of the site I pattern (such as lane 2 or lane 4) and the site II pattern (lane 1).

A finer comparison was made by analyzing the protease-digestion patterns of these cross-linked products. When the IdU oligo was added first, the digestion patterns did not change even if the complex was chased with the un-cross-linkable oligo (compare lane 2 with lane 4 in Figures 3b,c). In the reverse order, a different digestion pattern was observed (lane 1).

These results demonstrate that the two DNA binding sites in RecA are nonequivalent in binding; if they were equivalent, we should not have observed any effects of the order of additions. Furthermore, the experiments described in this section provided us a means by which two DNA binding sites of RecA can be probed differentially.

DNA Bound at Site I Lies Close to Met-164 in L1 Disordered Loop. We prepared cross-linked samples in large scale for the peptide-sequence analysis as described in Materials and Methods. The site I sample was made in the presence of excess RecA. After UV irradiation, the sample was digested extensively with trypsin. Urea−polyacrylamide (sequencing) gel electrophoresis was used to separate individual cross-linked products. We purified five cross-linked samples (Figure 4a). The results of sequencing were identical: A-E-I-E-G-E-I-G-D-S-H-X-G-L-A-A-R, which corresponds to amino acids 153−169 of RecA. No standard PTH-amino acid was detected at the twelfth position where Met-164 is expected (Figure 5a). Beyond this cycle, normal yield resumed, indicating that Met-164 is the amino acid cross-linked to DNA. The fact that sequencing was able to proceed beyond Met-164 suggests that the cross-linking occurred to the Met side chain or the C_α of the peptide backbone. Met-164 is located within a disordered loop L1 (residues 157−164) in the RecA crystal structure.

We observed spontaneous degradation of purified cross-linked products, presumably as a result of the breakage of the cross-linking bond (Figure 4a). Degradation of I-2,3,4 produced I-1, whereas degradation of I-5 gave I-2,3,(4) as well as I-1. The I-1 comigrates with the free DNA. The initial yield of I-1 sequencing was very low, despite the fact that this sample contained the highest radioactivity. We think the I-1 was essentially an un-cross-linked oligo which had been contaminated with a degraded cross-linked product

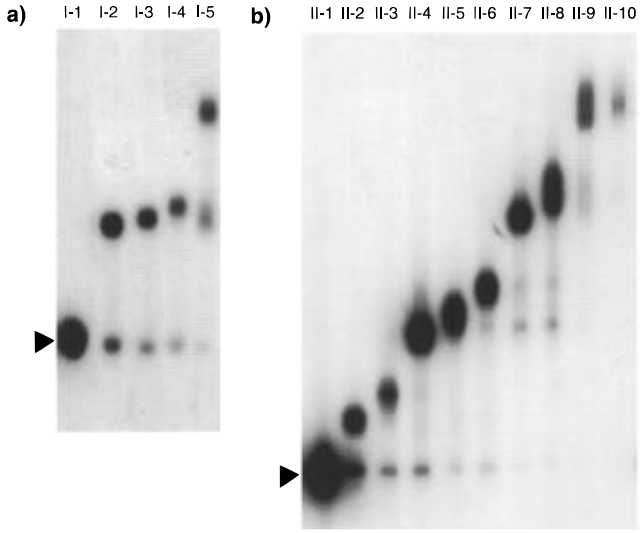

FIGURE 4: Purified cross-linked peptides. Autoradiograms of the samples separated on 10% sequencing gels. (a) Samples of site I-probing experiments. (b) Samples of site II-probing experiments. Triangle indicates the position of free oligo. Also see Table 1.

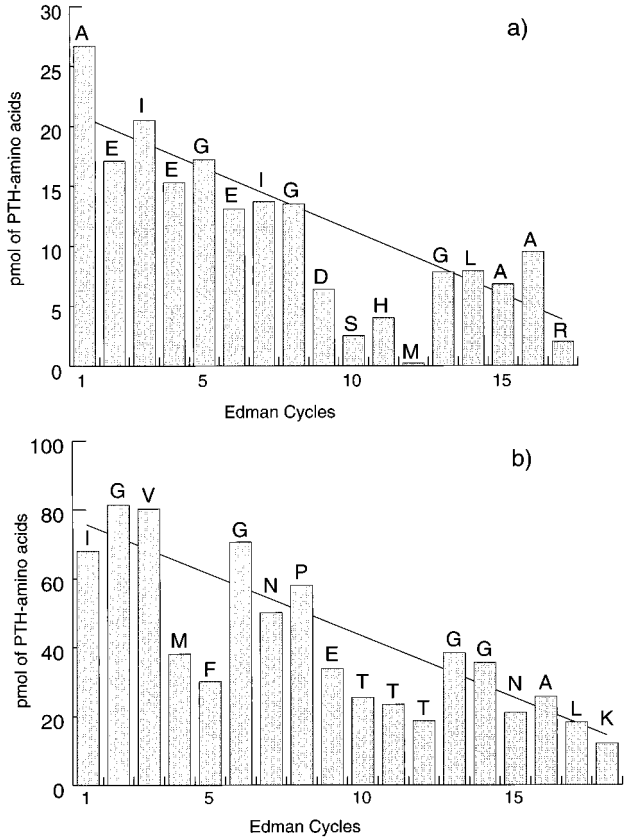

FIGURE 5: Results of Edman sequencing. (a) Analysis of site I peptide; an example from the sample I-2 is shown. Each column corresponds to, from left to right, each cycle of Edman degradation; the names of particular PTH amino acids detected at each cycles are shown by using single-letter codes. (b) Analysis of site II peptide 1 (sample II-5). Note: lower than average recoveries of certain PTH amino acids such as D (Asp), S (Ser), T (Thr), and H (His) are normal in this type of analysis (Allen, 1989).

which had lost a part of DNA and therefore lacked the ^{32}P label. The I-5 was probably the species in which two peptides were attached to a single DNA molecule, and I-2,3,4 were the singly-cross-linked molecules. Subtle differences in mobility among I-2,3,4 might be due to the differences in cross-linking positions on the DNA.

Table 1: Peptide Compositions of Site II Cross-Linked Samples

band in Figure 4b	peptide 1[a] (198−216)	peptide 2[a] (153−169)	peptide 3[a] (287−294)	peptide 4[a] (251−256)	total no. of mol submitted[b]	tentative band assignment[c]
II-10	20	6		10	20	1/4*, 2/4*
II-9	45	9			20	1/1/2, 1/1/1, etc.
II-8	91	16			49	1/1, 2/2, 1/2
II-7	82	15			55	1/1, 2/2, 1/2
II-6	36	15	7	8	34	1/3, 1/4, 2/3, 2/4
II-5	103	10			64	1, 2
II-4	102	13			113	1, 2
II-3	19	3	9	4	24	3, 4, (1), (2)
II-2	9	2	8	7	34	3, 4, (1), (2)
II-1					69	free DNA
Total	507	89	24	29	482	

[a] Numbers in these columns refer to the amount, expressed in picomoles, of each peptide detected from the Edman sequencing. [b] Amounts of cross-linked samples, in picomoles, estimated from the ^{32}P counts. [c] Numbers refer to the name of peptides (1−4). Presumptive multiple cross-linked products are indicated by using a slash (/); e.g., 1/3 means a doubly cross-linked species with both peptides 1 and 3 attached to the same DNA molecule. The assignment is as follows. II-1: free DNA. II-2,3: singly cross-linked species containing either peptide 3 or 4 (sizes of the two peptides are similar); each sample was contaminated with nonradioactive samples containing either peptide 1 or 2 (in parentheses), whose peptide part was intact but a portion of DNA was lost. II-4,5: singly cross-linked species containing either peptide 1 or 2. II-6: doubly cross-linked species containing either one of the peptide 1 or 2 together with either one of the peptide 3 or 4. II-7 and II-8: doubly cross-linked species containing two molecules of peptide 1 or 2. II-9: various kinds of triply cross-linked species. II-10: doubly cross-linked species similar to II-7,8, but with a peptide 4 derivative (asterisk) which was a result of incomplete digestion of Lys-C (data not shown).

DNA Bound at Site II Lies Close to Met-202/Phe-203 in L2 Disordered Loop. The site-II sample was made by first blocking site I with the un-cross-linkable oligo and then chasing with the IdU oligo. We used endoproteinase Lys-C to digest the cross-linked samples, purified 10 samples (Figure 4b, II-1 to II-10), and obtained one major and three minor peptide sequences (peptides 1−4 in Table 1). The sample II-1 was essentially the free DNA; each of the other nine samples produced a mixture of at least two peptide sequences.

Peptide 1 is the major species with a sequence: I-G-V-M-F-G-N-P-E-T-T-T-G-G-N-A-L-K, corresponding to amino acids 198−216. Although there was no particular cycle that failed to produce corresponding PTH-amino acids, a significant reduction of the step yield was detected at Met-202 and Phe-203 positions (Figure 5b; note that reduction of PTH-threonines is normal in Edman sequencing). These results indicate that cross-linking had occurred most probably at Met-202 and Phe-203 with a roughly equal efficiency. These two amino acids are located in the disordered loop L2 (residues 195−209).

Peptide 2 is a 24-residue peptide (153−176) of which the first 17 amino acids are identical to the site I peptide (153−169); this reflects the fact that, in the site-II analyses, we used Lys-C instead of trypsin. The step yield of each cycle, calculated by subtraction of the overlapping peptides, revealed that Met-164 is the position of cross-linking. The peptide 2 contained three more Met residues at positions 170, 171, and 175; these Met residues all gave a higher yield than Met-164 (data not shown). This peptide is therefore almost certainly a reflection of the residual binding to site I.

Two more peptides were identified in small quantities. One corresponds to residues 287−294 (peptide 3) and the other to 251−256 (peptide 4; in some samples, this peptide spanned a region 251−280 presumably due to incomplete Lys-C digestion). Each of these two peptides represented about 5% of the total cross-linked products in molecule. Elucidation of the cross-linked amino acid(s) was difficult because of the interference from other peptides of much higher abundance.

Why did each sample produce more than one peptide sequences? One of the reasons is because of multiple cross-linking. It is most evident in cases for II-6,7,8, the spontaneous degradation of which produced a faint band corresponding to II-4 (or possibly II-5; see Figure 4b). Also, even though each sample was identified as a single band on the autoradiogram, the areas corresponding to these bands might also have contained nonradioactive species whose peptide parts were larger but which lacked a portion of DNA. Table 1 contains our tentative interpretation of the cross-linked species in each samples, II-1 ∼ II-10.

DISCUSSION

Principles of Photo-Cross-Linking and Implied Physical Proximity. Photo-cross-linking is a reaction elicited by photon absorption. It normally proceeds very rapidly (in an order of microsecond), the end result of which is to freeze the noncovalent interactions such as protein−DNA binding [for reviews, see Hockensmith et al. (1991) and Pashev et al. (1991)]. The reaction outcome depends not only on time/space-averaged physical proximity between the two reactants but also on the chemical reactivity of potential target groups. Consequently, a general premise in interpreting photo-cross-linking experiments is that, while the two molecules that are close to each other may not necessarily cross-link, those that did cross-link must have been close to each other. But how close should they be?

The cross-linker used in this study, IdU, is a so-called "zero-length" cross-linker. That is, at the moment when cross-linking occurred, IdU and the target molecule must have essentially been in van der Waals contact. The actual time/space-averaged distance between these two molecules may be greater than this distance, depending on other variables such as flexibility of each molecule. In a case where cross-linking results and the co-crystal structure can be compared, a cross-linking distance of 4∼5 Å has been documented with 5-bromouracil (Dong et al., 1994). We also note here that RecA most likely binds to the phosphodiester backbone of DNA, rather than to the base (Leahy & Radding, 1986). In B-form DNA for instance, the 5-position

of pyrimidine, to which iodine of IdU is attached, is approximately 5 Å away from the phosphate.

DNA Molecules in the Active RecA−DNA−ATPγS Filament Are Located Close to Two Disordered Loops in the RecA Crystal Structure. In this study, we probed two functionally distinct DNA binding sites of RecA by photo-cross-linking. Site I, to which a DNA molecule binds first, was probed under the condition in which the photoreactive oligo was mixed with an excess of RecA (DNA base/RecA = 1:1). This mixing ratio should favor the formation of the RecA complex that holds only one molecule of ssDNA (Takahashi et al., 1989; Zlotnick et al., 1993). Cross-linking under this condition occurred at Met-164 which is located within a disordered loop L1 of the RecA crystal structure (Story et al., 1992). The complex for probing site II, the binding site for a second DNA molecule, was formed by first "blocking" site I with an un-cross-linkable DNA and then chasing this complex with IdU-containing oligo. This protocol should produce a doubly-bound ssDNA−RecA complex with the photoreactive oligo bound preferentially at site II. Cross-linking under this condition occurred mostly at Met-202 and Phe-203, both of which are located within the disordered loop L2.

The crystal structure of RecA polymer was solved in the absence of DNA or ATPγS. It is known that this structure is more similar to so-called inactive form of RecA filaments made in the presence of ADP or in the absence of nucleotide cofactors, rather than to the active filaments formed with ATP or ATPγS (Yu & Egelman, 1992). Consistent with this idea, Story and Steitz (1992) were able to diffuse ADP into the RecA crystal, but not ATPγS. To what extent might the crystal structure differ from the active RecA−DNA−ATPγS complexes?

Our cross-linking results indicate that regions close to the helical axis in the crystal structure are also close to the axis of the active RecA−DNA−ATPγS filament. The regions to which our cross-linking occurred, disordered loops L1 and L2, are precisely the ones which had been predicted to be involved in DNA binding by the crystallographic study (Story et al., 1992). These loops extend toward the helical axis of RecA polymer crystal, consistent with the idea that DNA molecules bound to RecA filaments must be located within a small distance (<17 Å) from the axis of filament (Egelman & Stasiak, 1986; Egelman & Yu, 1989). Certain amino acids in these loops are known to be highly conserved (Roca & Cox, 1990), and mutations of some of these amino acids cause various defects in primary RecA functions, including DNA binding and/or ATP-induced conformational change (Muench & Bryant, 1990; Muench & Bryant, 1991; Larminat et al., 1992; Cazaux et al., 1994; Nastri & Knight, 1994).

Recently, two other groups have also probed the DNA binding site(s) by photo-cross-linking methods (Morimatsu & Horii, 1995; W. M. Rehrauer and S. C. Kowalczykowski, personal communication). Both studies employed thymidine polymers as the DNA substrate, and the photoactivation was done by a conventional short-wavelength UV lamp (254 nm). No special attempt was made to differentiate the two DNA binding sites in either study. With a RecA−ssDNA complex made in the presence of ATPγS, Morimatsu and Horii (1995) identified two tryptic peptides corresponding to residues 178−183 and 89−106 (Tyr-103). Rehrauer and Kowalczykowski (personal communication) also found the peptide 178−183 (Lys-183) in addition to two other peptides, 61−72 and 233−243. In the crystal structure, all these peptides appear to be far (>17 Å) from the helical axis, except for the peptide 61−72, which is close to the axis and can also be very close to the L2 loop. It is not immediately clear why the results of these two groups differ considerably from ours (except, possibly, for the peptide 61−72). We note, however, that laser irradiation at 308 nm (i.e., our system) is believed to induce an excited state(s) that are different from the ones induced by irradiation using short-wavelength UV lamps (Dietz et al., 1987).

Two Functionally Distinct DNA Binding Sites of RecA May also Be Physically Distinct. The two DNA binding sites of RecA, site I and site II, are defined functionally in terms of the order of occupancy. The existence of the two physically distinct sites, however, has been an inference of but not necessarily a prerequisite for the two-site model. For example, site I and site II might simply be two different conformations of the same physical site. In this case, however, site I and site II cannot exist simultaneously in the same RecA molecule. Therefore to explain the functional duality of the two sites, one must invoke some kind of RecA dimer consisting of nonequivalent monomers, one of which provides the function of site I and the other monomer provides the function of site II [for example, see Lauder and Kowalczykowski (1991)].

In the crystal structure of RecA polymer (Story et al., 1992) and also in the electron-microscopic images of RecA−DNA filaments (Egelman & Stasiak, 1986), RecA dimers as a unit have not been observed. Individual RecA molecules in these polymers are related to one another with a simple helical symmetry. These observations imply, but not prove, the functional equivalence of all the individual RecA monomers in the polymer. If all the RecA monomers are truly functionally equivalent, then the two physically distinct sites must exist within a single RecA monomer in order for the RecA polymer to bind two DNA molecules simultaneously.

Our results show that the first DNA molecule bound to RecA is located close to Met-164 and the second DNA is located at a different place, near Met-202 and Phe-203, providing an independent evidence that site I and site II may be physically distinct. A caveat to this interpretation is that Met-164 (site I cross-link) and Met-202/Phe-203 (site II cross-links) can be very close to each other in three dimensions. Therefore, our results alone do not exclude the possibility of one site exhibiting two distinct conformations. Nonetheless, we can conclude that the observed difference between the site-I and site-II cross-linking does not result from a conformational difference between singly-bound and doubly-bound ssDNA−RecA complexes. When RecA was preincubated with the photoreactive oligo first and then chased with a regular oligo (thus forming a doubly-bound complex but with the cross-linkable oligo this time at site I), the resulting cross-linking pattern was indistinguishable from the one obtained with a singly-bound complex (Figure 3). That is, the first DNA molecule always cross-links to Met-164 under our conditions, regardless of the presence or absence of the second DNA molecule.

A Possible Arrangement of the Two DNA Binding Sites. To which region of RecA might the first and second DNA molecules bind? The resolution of our method is too low to locate the sites precisely. However, it is still possible to make a specific prediction as to the locations of DNA binding

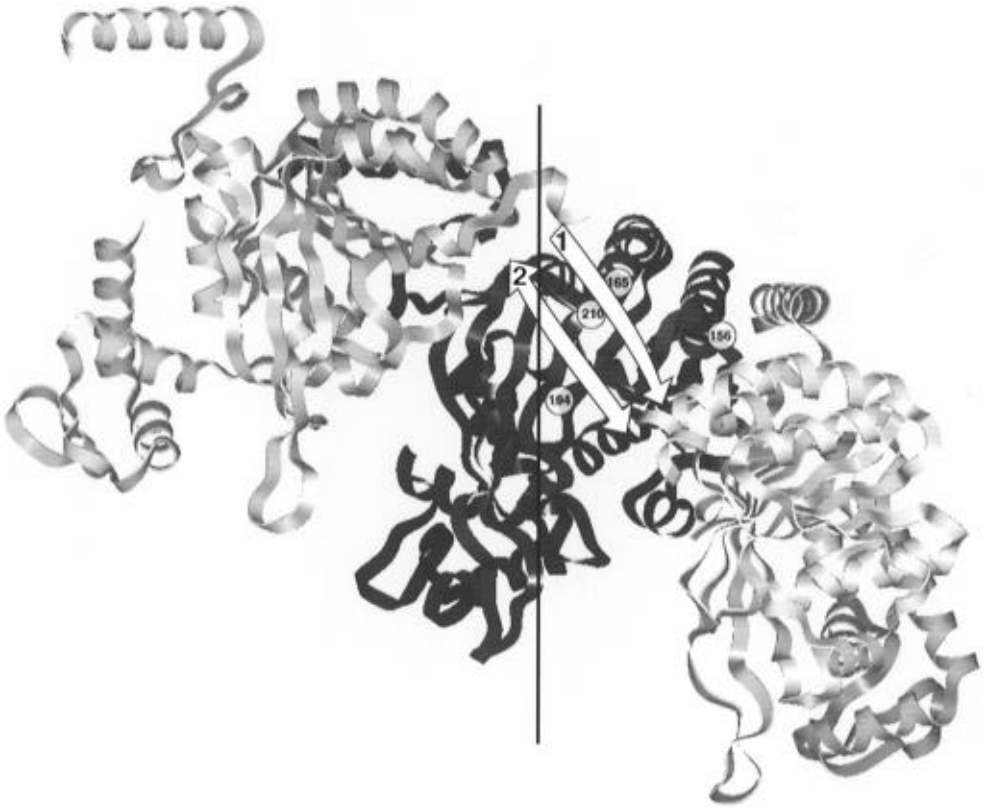

FIGURE 6: Possible arrangement of two DNA binding sites in RecA filament. Shown is a ribbon drawing of three consecutive RecA monomers (one-half turn of the polymer) viewed from inside of the filament, based on the atomic coordinates of RecA obtained from the Protein Data Bank, Brookhaven (Story et al., 1992). The vertical line represents the helical axis. Positions of two pairs of the amino acids that flank two disordered loops are indicated by circles with numbers in them (shown only for the middle monomer). E-156 and G-165 flank L1-loop; Q-194 and T-210 flank L2-loop. Possible locations of the first and second DNA molecules are shown by arrows. The direction of arrow indicates the 5′−3′ polarities of the DNA strands bound; the polarity of the first ssDNA is known (Stasiak et al., 1988) as indicated, while the polarity of the second ssDNA is not known.

sites, provided that there is no drastic structural difference between the crystal structure and active RecA−DNA complexes. Figure 6 schematically illustrates the crystal structure of RecA polymer containing three monomers and possible locations of the two ssDNA with respect to the center monomer viewed from the "inside" of RecA filament (Egelman & Stasiak, 1986; Story et al., 1992). Also shown in the drawing are the edges of two disordered loops L1 (157−164) and L2 (195−209).

Our cross-linking results indicate that the first DNA is bound close to Met-164. The exact atomic coordinates of Met-164 are not known, but the next residue (Gly-165) is visible in the crystal structure. An attractive region to place the first DNA molecule would be therefore near this Gly-165. Assuming an equivalent contact of all the RecA monomers with DNA, we predict that site I may be located somewhere within the "upper-right" area of the center RecA monomer in this drawing, the area which includes the L1 and possibly L2 loops (Figure 6).

Predicting the location of site II is complicated by the fact that the positions of Met-202 and Phe-203, the targets of site-II cross-linking, are ambiguous; these amino acids are in the middle of disordered loop L2, which is 15 amino acids long. It should be recalled here, however, that the presence of the second DNA molecule did not prevent the first DNA molecule from cross-linking to Met-164. It is therefore unlikely that the site II is located in between site I and Met-164. Placing the second DNA molecule close to the edges of L2 loop, thus to the "left" of the first DNA (see Figure 6), would be most consistent with our cross-linking results, although this is not the only possibility. We do not know whether this site binds the recipient strand or the outgoing strand during strand exchange.

Third DNA-Binding Site? In addition to the two major peptides identified as the site-I and site-II peptides, two other peptides, 3 (287−294) and 4 (251−256), were detected in the site-II probing experiments. The small proportions of these peptides (∼5% each of total) do not negate their importance; however, we do not think they represent any stable DNA binding sites. As mentioned earlier, stereochemical constraints dictate that the DNA strands bound to RecA filaments should not be more than 17 Å away from the axis of RecA filament (Egelman & Stasiak, 1986). The two minor peptides 3 and 4 are ∼50 and ∼35 Å away from the helical axis in the crystal structure, respectively.

We do not exclude the possibility that the UV irradiation might have damaged RecA before cross-linking occurred. Nonetheless, there might be some meaning for cross-linking to the peptide 3 (287−294). This peptide corresponds to one of the two C-terminal "lobes" protruding from the body of RecA (Story et al., 1992). Interestingly, one of the cross-linked peptides detected by Rehrauer and Kowalczykowski (personal communication) included the peptide 233−243 which corresponds to the other lobe. These lobes together point towards the deep helical groove of RecA filaments, and are probably involved in interaction with LexA repressor (Yu & Egelman, 1993). But another possible function of the region in and around these lobes might be to provide

some weak interaction with DNA, which must go in and out through this groove of RecA filament during DNA strand exchange. Kubista et al. (1990) in fact proposed the existence of the third, weak DNA binding site which is more exposed than the other two sites.

While this paper was being reviewed, a similar photo-cross-linking study of RecA−DNA interaction using IdU was published (Malkov & Camerini-Otero, 1995). Under a single binding condition, the authors identified two cross-linked amino acids, Met-164 and Phe-203, which are identical to the site-I and site-II cross-links in our study, respectively.

ACKNOWLEDGMENT

We thank Dr. David Mauzerall for providing access to the laser. Protein sequencing data were obtained at the Rockefeller University Protein/DNA Technology Center. We also thank Drs. Robert Craigie and Kiyoshi Mizuuchi for critical reading of the manuscript and Dr. Steve Kowalczykowski for sharing his unpublished results.

REFERENCES

Allen, G. (1989) *Sequencing of Protein and Peptides*, Elsevier, Amsterdam.
Cazaux, C., Larminat, F., Villani, G., Johnson, N., Schnarr, M., & Defais, M. (1994) *J. Biol. Chem. 269*, 8246−8254.
Cotterill, S. M., Satterthwait, A. C., & Fersht, A. R. (1982) *Biochemistry 21*, 4332−4337.
Cox, M. M., & Lehman, I. R. (1987) *Annu. Rev. Biochem. 56*, 229−262.
Dietz, T. M., von Trebra, R. J., Swanson, B. J., & Koch, T. H. (1987) *J. Am. Chem. Soc. 109*, 1793−1797.
Dong, Q., Blatter, E. E., Ebright, Y. W., Bister, K., & Ebright, R. H. (1994) *EMBO J. 13*, 200−204.
Egelman, E. H., & Stasiak, A. (1986) *J. Mol. Biol. 191*, 677−697.
Egelman, E. H., & Yu, X. (1989) *Science 245*, 404−407.
Hockensmith, J. W., Kubasek, W. L., Vorachek, W. R., Evertsz, E. M., & von Hippel, P. H. (1991) *Methods Enzymol. 208*, 211−236.
Kowalczykowski, S. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 539−575.
Kowalczykowski, S., & Eggleston, A. (1994) *Annu. Rev. Biochem. 63*, 991−1043.
Kubista, M., Takahashi, M., & Norden, B. (1990) *J. Biol. Chem. 265*, 18891−18897.
Larminat, F., Cazaux, C., Germanier, M., & Defais, M. (1992) *J. Bacteriol. 174*, 6264−6269.
Lauder, S., & Kowalczykowski, S. (1991) *J. Biol. Chem. 266*, 5450−5458.
Leahy, M. C., & Radding, C. M. (1986) *J. Biol. Chem. 261*, 6954−6960.
Malkov, V. A., & Camerini-Otero, R. D. (1995) *J. Biol. Chem. 270*, 30230−30233.
Menetski, J. P., & Kowalczykowski, S. C. (1987) *J. Biol. Chem. 262*, 2085−2092.
Morimatsu, K., & Horii, T. (1995) *Eur. J. Biochem. 228*, 772−778.
Muench, K. A., & Bryant, F. R. (1990) *J. Biol. Chem. 265*, 11560−11566.
Muench, K., & Bryant, F. (1991) *J. Biol. Chem. 266*, 844−850.
Müller, B., Koller, T., & Stasiak, A. (1990) *J. Mol. Biol. 212*, 97−112.
Nastri, H., & Knight, K. (1994) *J. Biol. Chem. 269*, 26311−22.
Pashev, I. G., Dimitrov, S. I., & Angelov, D. (1991) *Trends Biochem. Sci. 16*, 323−326.
Pugh, B. F., & Cox, M. M. (1987) *J. Biol. Chem. 262*, 1326−1336
Roca, A. I., & Cox, M. M. (1990) *Crit. Rev. Biochem. Mol. Biol. 25*, 415−456.
Rosselli, W., & Stasiak, A. (1990) *J. Mol. Biol. 216*, 335−352.
Stasiak, A., Egelman, E. H., & Howard-Flanders, P. (1988) *J. Mol. Biol. 202*, 659−662.
Story, R. M., & Steitz, T. A. (1992) *Nature 355*, 374−376.
Story, R. M., Weber, I. T., & Steitz, T. A. (1992) *Nature 355*, 318−325.
Takahashi, M., Kubista, M., & Norden, B. (1989) *J. Mol. Biol. 205*, 137−147.
Williams, K. R., & Konigsberg, W. H. (1991) *Methods Enzymol. 208*, 516−539.
Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R., & Koch, T. H. (1993) *Science 262*, 1255−1257.
Yu, X., & Egelman, E. (1992) *J. Mol. Biol. 227*, 334−346.
Yu, X., & Egelman, E. (1993) *J. Mol. Biol. 231*, 29−40.
Zlotnick, A., Mitchell, R. S., & Brenner, S. L. (1990) *J. Biol. Chem. 265*, 17050−17054
Zlotnick, A., Mitchell, R. S., Steed, R. K., & Brenner, S. L. (1993) *J. Biol. Chem. 268*, 22525−22530.

BI952438V