Photocross-links between Single-stranded DNA and Escherichia coli RecA Protein Map to Loops L1 (Amino Acid Residues 157-164) and L2 (Amino Acid Residues 195-209)*

(Received for publication, August 22, 1995, and in revised form, October 11, 1995)

Vladislav A. Malkov and R. Daniel Camerini-Otero‡
From the Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-1810

To function as a repair and recombination protein, RecA has to be assembled as an active filament on single-stranded DNA in the presence of ATP or its analogs. We have identified amino acids in the primary DNA binding site of RecA that interact with single-stranded DNA by photocross-linking. A nucleoprotein complex consisting of RecA protein bound to a monosubstituted oligonucleotide bearing a 5-iododeoxyuracil cross-linking moiety was irradiated with long wavelength ultraviolet radiation to effect cross-linking with RecA protein. Subsequent trypsin digestion, followed by purification and peptide sequencing, revealed the cross-linking of two independent peptides, amino acid residues 153-169 and 190-216. Met164 from loop L1 and Phe203 from loop L2 were determined to be the exact points of cross-linking. Thus, our data confirm and extend predictions about the DNA binding domain of RecA protein based on the molecular structure of RecA (Story, R. M., Weber, I. T., and Steitz, T. A. (1992) Nature 355, 318-325).

RecA protein is the central enzyme of bacterial homologous recombination. It also plays an important role in DNA repair and the SOS response (reviewed in Refs. 1-4). The discovery of eukaryotic RecA analogs Rad51 and DMC1, that have sequence homology to RecA and, possibly, similar biochemical activities, eukaryotic RecA analogs Rad51 and DMC1, that have sequence and the SOS response (reviewed in Refs. 1-4). The discovery of two independent peptides, amino acid residues 153-169 and 190-216. Met164 from loop L1 and Phe203 from loop L2 were determined to be the exact points of cross-linking. Thus, our data confirm and extend predictions about the DNA binding domain of RecA protein based on the molecular structure of RecA (Story, R. M., Weber, I. T., and Steitz, T. A. (1992) Nature 355, 318-325).

Photocross-linking has been used previously to determine the exact amino acids contacting DNA in a number of DNA (or RNA)-protein complexes (see Refs. 6-10 for mechanism and references). Willis et al. (11) demonstrated that 5-iodouracil-substituted DNA or RNA can be cross-linked with an extremely high yield by monochromatic 325-nm irradiation. From analogy with 5-bromodeoxyuridine, we may expect that UV irradiation of 5IdU with an emission peak at 312 nm can result in high yields of photocross-linking with tryptophans, tyrosines, and histidines and photocross-linking with lower yields with most of the other amino acids.

In this report we show that two loops, L1 and L2, of RecA protein are in close contact with single-stranded DNA. We have determined Met164 of loop L1 and Phe203 of loop L2 to be the exact positions of photocross-linking.

MATERIALS AND METHODS

RecA protein was isolated and purified as previously described (12). Oligonucleotides were synthesized by automated beta-cyanoethyl phosphoramidite DNA synthesis on a 380B DNA synthesizer (Applied Biosystems).

Preparative UV Irradiation, Protease Digestion, and Purification of Cross-linked Peptides—162 µg of the single-stranded oligonucleotide were preincubated with 10 µg of RecA protein for 1 h at 37°C in 25 ml in buffer containing 20 mM Tris-Cl, pH 7.7, 12.5 mM MgCl2, 0.4 mM DTT, and 0.1 mM ATP-S. The major portion of the oligonucleotide was synthesized with a phosphate at the 5' end, a small fraction was 32P-labeled by T4 polynucleotide kinase.

Irradiation in a UV Stratalinker 1800 (Stratagene), loaded with 312-nm bulbs (40 watts total) was carried out for 8 h. After irradiation the DNA-protein complex was dissociated by adding 15 mM EDTA, 5 mM ADP, 0.5 mM NaCl, and 0.3% SDS. A small oligo(dT) cellulose column was poured. The column was equilibrated with binding buffer (10 mM Tris-Cl, pH 8.0, 0.1 mM EDTA, 0.4 mM DTT, 0.1% SDS, 0.5 mM NaCl) prior to sample loading. Then it was washed with 10 volumes of binding buffer, and the sample was eluted in elution buffer 1 (10 mM Tris-Cl, pH 8.0, 0.05 mM EDTA, 0.4 mM DTT, 0.01% SDS). DNA-containing fractions were identified by following the 32P label. SDS and DTT were adjusted to 0.3% and 3 mM, correspondingly, in a reaction volume of 0.5 ml. The tube was heated at 85°C for 20 min to denature the protein. After denaturation, the reaction was allowed to cool, and 1 µl of dextranase buffer (20 mM Tris-HCl, pH 7.5, 1 mM CaCl2, 2 mM DTT) was added. After 4 h of dextranase digestion with 40 µg of enzyme at 37°C, an extra 40 µg of dextranase were added, and the cleavage was continued overnight.

The sample was applied to oligo(dt) cellulose as described above. The oligonucleotide was eluted with buffer 2 (10 mM Tris-HCl, pH 8.0, 0.05 mM EDTA, 0.4 mM DTT, 1 mM urea) and concentrated six times in a Speedvac. A small fraction was saved for a gel. The probe was heated at 85°C for 20 min. After denaturation the reaction was allowed to cool, and it was diluted 6-fold with trypsin buffer (50 mM Tris-HCl, pH 7.5, 1 mM CaCl2). After 4 h of modified trypsin digestion with 40 µg of enzyme at 37°C, an extra 40 µg of trypsin were added, and the cleavage was continued overnight.

The sample was again purified by oligo(dt) cellulose chromatography and then purified by gel electrophoresis. The probe was loaded onto a 10% denaturing (urea) PAGE, and the bands were cut from the gel, eluted, and sent to the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University for peptide sequencing.

Analytical UV Irradiation, Protease Digestion, and Purification of Cross-linked Peptides—Analytical complex formation, UV irradiation, and purification were essentially as described for preparative samples. Typically, 0.2-0.5 µg of oligonucleotide were preincubated for 1 h with

* 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.
‡ To whom correspondence should be addressed: Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, MD 20892-1810

1 The abbreviations used are: ss, single-stranded; 5IdU, 5-iododeoxyuridine; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; ATP-S, adenosine 5'-O-(thiotriphosphate); Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

30230

Downloaded from http://www.jbc.org/ by guest on July 27, 2018
an excess of RecA protein in 300 μl. The probe was then placed in a thin 1-mm quartz cuvette and irradiated for 4 h. Scaled-down oligo(dT) cellulose purifications were carried out as described above. The protease digestions were performed overnight with 10–20 μg of the appropriate protease. Finally, the cross-linked material was analyzed on 10% denaturing PAGE.

RESULTS

Analytical Digestion of Cross-linked RecA Protein—It was previously reported that RecA accomplishes its strand exchange activity by binding DNA at a minimum of two distinct sites (13, 14). In this study we were primarily interested in the localization of the major (highest affinity) single-strand DNA binding site, also called site I or the primary binding site. In order to exclude the possibility of the multiple occupancy of DNA in the RecA protein, we used excess protein, less than two nucleotides per one protein protomer. Since it is well established that the binding of ssDNA at the primary binding site is at a stoichiometry of three nucleotides per one protein molecule (15), we had a 1.5-fold molar excess of RecA.

The time course of the photocross-linking reaction showed that the reaction reached saturation with 19% of oligonucleotides cross-linked after 4 h, detected by gel-shift assay (Fig. 1, filled circles). Most of the photocross-links occurred at the 5′-iododeoxyuridine position, as the cross-linking yield was more than 10 times less for a control oligonucleotide containing the 5′-end labeled oligonucleotide for the same sequence. The specificity of V8 protease depends on the buffer contained 1.1 mM ADP; lane 3, the complex was formed in the presence of 0.3 mM ATPγS and 1.1 mM ADP; lanes 4–8, 0.1 mM ATPγS was used as a cofactor. Lanes 1–4, no competitive DNA was added; lanes 5–8, the molar excess of unsubstituted DNA added to the reaction mixtures is indicated at the top.

We utilized a two-step endoprotease digestion to obtain a complete trypsin digestion and to minimize sample losses. The first step was cleavage of cross-linked protein in the presence of 0.1% SDS by clostripain, which cleaves at the carboxylic end of all arginines. The second step was redigestion of cross-linked peptides in the presence of 1 M urea by trypsin, which cleaves after both arginines and lysines. The products of these two rounds of protease digestion were resolved by electrophoresis on a 10% polyacrylamide urea gel. The oligonucleotide and cross-linked samples were detected by autoradiography.

We resolved three different cross-linked peptides (Fig. 3, lane 2). To check whether we obtained the shortest completely cleaved peptides we combined different endoprotease digestions. A part of the sample after trypsin cleavage was saved for gel analysis (Fig. 3, lane 2), while the other part was repurified by oligo(dT) chromatography and then subjected to either clostripain (lane 3), trypsin (lane 4), or Staphylococcus aureus V8 protease (lanes 5–16) treatment. Clostripain did not change the cross-linked peptide pattern obtained (compare lanes 2 and 3). Additional trypsin treatment did not give rise to any new shorter products, but band C disappeared. Thus bands A and B contain the two shortest distinct peptides, while band C was an underdigested tryptic cross-linked peptide, that was converted into either A or B during the second round of trypsin cleavage. We expect the resulting peptides in samples A and B to be short because they migrated faster than samples cross-linked to a peptide 20 amino acids long (data not shown).

We also performed a time course of V8 protease cleavage of tryptic peptides in order to reveal the relationship between the three products (bands A, B, and C) and to determine whether these peptides have aspartic or glutamic acids in their sequence. The specificity of V8 protease depends on the buffer...
used for digestion. It cleaves specifically at the carboxylic side of glutamic and aspartic acid in 50 mM sodium phosphate, pH 7.8. In 50 mM ammonium bicarbonate, pH 7.8, this protease cleaves only after glutamic acid. Comparing the intensities of all the bands (lanes 5–16), it is clear that sample B was converted into B’ and, presumably, has one glutamic acid. Band A was relatively resistant to V8 digestion, but had at least two sites of cleavage (Glu or Asp) leading to A’. Band C behaved as B, so it was reasonable to expect that sample C represents the underdigestion of band B (see Fig. 4 for the peptide map). Amino Acid Sequencing of the Tryptic Peptides Cross-linked to ssDNA—The identity of the cross-linked peptides was unambiguously determined by amino acid sequencing. RecA-d3YDNA complexes were formed maintaining the same stoichiometry but under scaled up conditions. We used the same two-step endoprotease digestion. Small fractions after clostripain and trypsin rounds of digestion were loaded onto an analytical 10% denaturing PAGE (Fig. 5, lanes 2 and 3). Meanwhile, the main sample after the last digestion was loaded onto a 10% preparative urea PAGE, all of the bands were cut from the gel and eluted, and the peptides were sequenced.

The digestion was less complete than for small quantities. So, in addition to the typical triple band pattern observed in the lane 2 of Fig. 3, there was an extra band whose mobility coincided with the band before the trypsin treatment (Fig. 5, lane 2). Presumably, this top band represented the partially digested product. However, materials from all four gel shifted bands as well as the band corresponding to free oligonucleotide were sequenced. There was no peptide material in the major band corresponding to unbound oligonucleotide. All other gel-shifted bands gave amino acid sequences.

The gel-shifted band A contains amino acid residues 153–169 (Fig. 4), corresponding to loop L1 with small flanking regions in the molecular structure of RecA (17). The recovery of Met was only 5.8% that of the expected value; therefore, it is the probable site of cross-linking.

The peptide component of band B matched residues 199–217 of the RecA protein. It contains most of loop L2 and terminates at the nearest lysine on the carboxyl side. The yield of Phe in cycle 5 was only 21% of the anticipated level; thus, it appears to be a potential site of cross-linking. The remaining 20% of this peptide may be cross-linked at other positions along the peptide spanning residues 199–217, although the sensitivity of the sequencing was not sufficient to unambiguously determine these minor species. As we used long wavelength UV radiation resulting in minimum damage to biological substrates, it is unlikely that the decreased yield of phenylalanine results from photodestruction rather than photocross-linking. Also, as a control, we subjected a synthetic peptide derived from amino acids residues 193–212 to the same dose of UV radiation and did not observe any decrease of Phe level in the sequence (data not shown). Because the sequences of the cross-linked peptides A and B do not contain internal sites for cleavage by trypsin, they are the shortest possible tryptic peptides. This is consistent with the fact that the cross-linked peptides A and B represent limit digest segments.

Sequencing confirmed our preliminary conclusions about the nature of band C. The product C had peptide extending from residues 199–222 so C was indeed an underdigested band B, and lysine 216 was the position of undercleavage (see Fig. 4). As in the case of B, Phe203 was determined as the primary photocross-linking position.

The longest peptide, D, started upstream from loops L1 and L2 at residue 135. We sequenced only the first 22 amino acids from the amino end and detected no obvious position of cross-linking among these 22 amino acids. Thus, the simplest explanation is that lane D represents a partial digest of products A and B.

We can explain the electrophoretic mobilities of all peptide-DNA species in Fig. 3 on the basis of their peptide lengths and electrical charges (see Fig. 4 for peptide map). Sample A migrates faster than B, as it has one less amino acid and a –3 electrical charge of the peptide versus the uncharged peptide component of B. A presence of six extra amino acids slows down the mobility of C in comparison with B. Digestion with V8 protease cleaves product B into B’. B’ has the shortest cross-linked peptide, which is nine amino acids long, and exhibits the fastest mobility. The lengths of the digestion products of A’ based on the cleavage specificity of V8 protease and the known peptide sequence are predicted to be 15, 13, 11, and 9 amino acids long. As a result, the mobility of A’ is somewhat faster than A, but slower than that of B’. The inability to obtain an amino acid sequence for the products of B and B’. The inability to obtain all four predicted cleavage products of A’ and the relative resistance of the V8 cleavage of the other peptides could be the result of steric protection by the cross-linked oligonucleotide in the vicinity of these potential protease cleavage sites.

We have not found any evidence for any additional points of cross-linking and all peptides have been accounted for. Thus, we can conclude that: (i) there are only two independent points of photocross-linking, (ii) peptides A and B are the shortest products of tryptic digestion, and (iii) cross-links occurred at Met214 and Phe203.

**DISCUSSION**

Our cross-linking data confirm a prediction deduced from crystallographic studies of RecA protein (17, 18). Within the crystal, RecA protein is organized in a B3.7 A pitch right-handed helix with six protomers per turn (17). Assuming that this helix reflects the geometry of the DNA-RecA filament, and bearing in mind the fact that the DNA lies near the center (19), Story et al. (17) proposed that loops L1 and L2 lining the cavity down the filament axis are involved in DNA binding. This prediction, based on the x-ray data obtained in the absence of DNA and triphosphate cofactors, is confirmed and extended by our finding that these loops constitute the ssDNA (primary) binding site in the active DNA-RecA filament.

The identity of amino acids in the loop L2 is important for proper RecA function. In vitro, the recA430 mutation (Gly204→Ser) led to a decrease in ssDNA binding affinity (20). Two other mutations (Glu207→Gln and Gly211→Ala) exhibited a recA- phenotype in vivo (21). In vitro, these two mutants were deficient in promoting both the self-cleavage of LexA repressor and the DNA strand-exchange reaction (22). The DNA strand-ex-
change defects were correlated with an inability of mutant proteins to displace SSB from DNA (22). All of these data are consistent with the participation of loop L2 in ssDNA binding.

Gardner et al. (23) unambiguously demonstrated that a peptide (amino acids 193–212) spanning loop L2, binds to ssDNA (23). This observation combined with our cross-linking findings strongly argues for a direct participation of loop L2 in ssDNA binding.

While this manuscript was in preparation a study was published (24) that also used photocross-linking to identify the ssDNA binding domain of RecA. Surprisingly, the authors claimed two other regions to be the DNA binding sites of RecA in the presence of ATPγS, the regions of residues 89–106 and 178–183. As these peptides are located on the outside surface of the filament and far from the central axis, these observations are not in accord with the x-ray (17) and electron microscopy data (19), which place the DNA near that axis. Although the difference between our and their data is puzzling, it could be a consequence of different experimental designs. A lack of cross-links in loops L1 and L2 could be the result of a less powerful cross-linking strategy. Oligo(dT) was photocross-linked to RecA protein by short UV (254 nm) irradiation (24). Halogen-substituted uridines are more reactive than thymines. Thus, it is conceivable that, even though the oligo(dT) might have close contacts with loops L1 and L2, it might not be cross-linkable in these regions. In addition, the two regions of residues 89–106 and 178–183 could be part of the second DNA binding site of RecA. As we used a significant excess of RecA protein compared to the work of Morimatsu and Horii (24), only the primary DNA binding site was occupied, and this may explain why we did not have any cross-links with these secondary regions.

Acknowledgments—We thank members of the W. M. Keck Foundation biotechnology resource laboratory, especially Drs. Ken Williams and Kathy Stone, for peptide sequencing and helpful discussions. We are grateful to Dr. Carol Camerini-Otero and Linda Robinson for help and to George Poy for oligonucleotide synthesis. We also acknowledge Drs. Peggy Hsieh, Sergei Kozyavkin, and Oleg Voloshin for useful and stimulating conversations.

REFERENCES

1. Kowalczykowski, S. C. (1991) Biochimie (Paris) 73, 289–304
2. West, S. C. (1992) Annu. Rev. Biochem. 61, 603–640
3. Cox, M. M. (1993) Bioessays 15, 617–623
4. Kowalczykowski, S. C., Dixon, D. A., Egglieston, A. K., Lauder, S. D., and Rehrauer, W. M. (1994) Microbiol. Rev. 58, 538–602
5. Sung, P. (1994) Science 265, 1241–1243
6. Benson, F. E., Stasiak, A., and West, S. C. (1994) EMBO J. 13, 5764–5771
7. Egelman, E. H. (1993) Curr. Opin. Struct. Biol. 3, 189–197
8. Williams, K. R., and Koningsberg, W. H. (1993) Methods Enzymol. 208, 516–539
9. Willis, M. C., LeCuyer, K. A., Meisenheimer, K. M., Uhlenebeck, O. C., and Koch, T. H. (1994) Nucleic Acids Res. 22, 4947–4952
10. Hicke, B. J., Willis, M. C., Koch, T. H., and Cech, T. R. (1994) Biochemistry 33, 3364–3373
11. Willis, M. C., Hicke, B. J., Uhlenebeck, O. C., Cech, T. R., and Koch, T. H. (1993) Science 262, 1253–1257
12. Yancey-Wrona, J. E., and Camerini-Otero, R. D. (1995) Curr. Biol. 5, 1189–1198
13. Howard-Flanders, P., West, S. C., and Stasiak, A. (1984) Nature 310, 215–220
14. Kurumizaka, H., Ikawa, S., Ikeya, T., Ogawa, T., and Shibata, T. (1994) J. Biol. Chem. 269, 3068–3075
15. Koller, T., DiCapua, E., and Stasiak, A. (1983) Mechanisms of DNA Replication and Recombination (Cazzarelli, N., ed) pp. 723–729, Alan R. Liss, New York
16. Leahy, M., and Radding, C. M. (1986) J. Biol. Chem. 261, 6954–6960
17. Story, R. M., Weber, I. T., and Steltz, T. A. (1992) Nature 355, 318–325
18. Story, R. M., and Steltz, T. A. (1992) Nature 355, 374–376
19. Egelman, E. H., and Stasiak, A. (1986) J. Mol. Biol. 191, 677–697
20. Menetski, J. P., and Kowalczykowski, S. C. (1990) J. Mol. Biol. 211, 845–855
21. Cazaux, C., Larminat, F., and Defais, M. (1991) Biochimie (Paris) 73, 281–284
22. Cazaux, C., Larminat, F., Villani, G., Johnson, N. P., Schnarr, M., and Defais, M. (1994) J. Biol. Chem. 269, 8246–8254
23. Gardner, R. V., Voloshin, O. N., and Camerini-Otero, R. D. (1995) Eur. J. Biochem. 233, 419–425
24. Morimatsu, K., and Horii, T. (1995) Eur. J. Biochem. 228, 772–778

Fig. 4. Map of cross-linked peptides.Outlined R or K are potential sites of trypsin cleavage (arginine or lysine). Italicized E or D are potential sites of V8 protease cleavage (glutamic or aspartic acid). Shadowed M and F are the exact positions of photocross-linking (methionine 164 and phenylalanine 203).

Fig. 5. Two-step endoprotease digestion prior to peptide sequencing. On the left, a cartoon of the strategy. On the right, an autoradiogram of the digestion steps. Lane 1, oligonucleotide; lane 2, cross-linked RecA protein-DNA complexes treated with dostripain (R); lane 3, cross-linked complexes after additional treatment with modified trypsin (R + K).
Photocross-links between Single-stranded DNA and *Escherichia coli* RecA Protein
Map to Loops L1 (Amino Acid Residues 157-164) and L2 (Amino Acid Residues 195-209)

Vladislav A. Malkov and R. Daniel Camerini-Otero

*J. Biol. Chem.* 1995, 270:30230-30233.
doi: 10.1074/jbc.270.50.30230

Access the most updated version of this article at http://www.jbc.org/content/270/50/30230

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 23 references, 5 of which can be accessed free at http://www.jbc.org/content/270/50/30230.full.html#ref-list-1