The EMBO Journal vol.6 no.8 pp.2493–2498, 1987

The accessibility of DNA to dimethylsulfate in complexes with recA protein

Elisabeth Di Capua and Berndt Müller

Institute of Cell Biology, ETH Hönggerberg, CH-8093 Zürich, Switzerland

Communicated by Th.Koller

recA protein coats DNA co-operatively to form filaments ~100 Å thick, which in the presence of ATP, and more stably so in the presence of the non-hydrolyzable analog ATPγS, have a helical appearance with a deep cleft in the protein coat. This protein helix follows the DNA helix, to which it imparts a new helicity of 18.5 bp per turn of 97 Å pitch. Here we test the accessibilty of the DNA in the complex to modification by dimethylsulfate, and find that the complexed DNA is ~2-fold more reactive on the major groove side than it was in B-DNA (methylation of guanine N7), while it is protected ~2-fold on the minor groove side (methylation of adenine N3), suggesting that the protein coats the DNA along the minor groove. Furthermore, N3 of cytosine, a residue involved in base pairing, is found exposed in complexes with single strands as it is in naked single-stranded DNA, while it remains inaccessible in complexes with double strands, suggesting that the latter is not melted at this stage of the strand exchange reaction.

Key words: DNA–protein interaction/dimethylsulfate/recA

Introduction

recA of *Escherichia coli* is necessary for recombination. *In vitro*, the purified protein (mol. wt 38 000) catalyzes strand exchange (heteroduplex formation, the first step in the Holliday model for recombination) and strand assimilation (renaturation of complementary single strands) (for review, see Howard-Flanders *et al.*, 1984). To study the mechanism of these reactions, our approach is to characterize the interaction of recA protein with DNA, single-stranded and double-stranded. Stable complexes of recA with DNA are formed in the presence of ATPγS, a non-hydrolyzable analog of ATP (West *et al.*, 1980). For both single- and double-stranded DNA, the complexes are helical filaments of ~100 Å outer diameter and 97 ± 5 Å pitch (Figure 1, top). 6.17 recA units per turn of this protein helix (Di Capua *et al.*, 1982) interact with one turn of DNA of 18.5 bp [whereby the double helix is partially unwound by an average of 15°/bp (Stasiak and Di Capua, 1982)], or with 18.5 nucleotides of the single-strand (Koller *et al.*, 1983). Figure 1, bottom, summarizes these features in a sketch for clarity. The structure of these helical complexes led to the model (Howard-Flanders *et al.*, 1984) where one partner for strand exchange is complexed as described, ready to accept the second partner, which is proposed to invade the deep groove of the filamentous complex. This mechanistic model predicts that the DNA in the complex must be accessible in some way for homology recognition.

In the complexes with ATPγS, DNA is fully protected against DNases (Stasiak *et al.*, 1983; Bryant *et al.*, 1985). This is not surprising in view of the thick coat of protein apparent on electron micrographs, which is likely to prevent the access of large molecules. This paper describes studies performed with the small reagent dimethylsulfate. This reagent has been used by Leahy and Radding (1986) to study the interaction of recA with a single-stranded oligonucleotide, where they observed accessibility of both adenines and guanines. Dimethylsulfate is a refined probe for structural analysis of DNA sequences involved in the interaction with proteins (e.g. Siebenlist and Gilbert, 1980). The elegance of the method lies in that it can distinguish the grooves of double-stranded DNA, as it reacts with N3 of adenine in the minor groove, and N7 of guanine in the major groove. We found that in complexes with recA protein, guanine modification is increased ~2-fold, while the reaction with adenines is decreased ~2-fold. In addition, dimethylsulfate is a probe for unpaired cytosines (Kirkegaard *et al.*, 1983); it revealed that double-stranded DNA in complexes remains paired, while single-stranded DNA remains accessible.

Results

Accessibility of N3 of cytosine

N3 of cytosine is a residue involved in base-pairing. It is thus accessible to chemical modification in single-stranded DNA and inaccessible in duplex DNA (Kirkegaard *et al.*, 1983). Figure 2A and C shows that this is indeed the case: DNA in the absence of recA protein becomes methylated at N3 of cytosine only when melted prior to the reaction; in contrast, double-stranded DNA is protected from methylation. We are asking what is the influence of binding recA protein on the accessibility. Figure 2B (complexed single-stranded DNA) shows that, in the complex, N3 of cytosines remains accessible to dimethylsulfate to the same extent as in naked DNA; Figure 2D (complexed double-stranded DNA) shows that binding to recA does not make N3 of cytosines more accessible than in naked DNA.

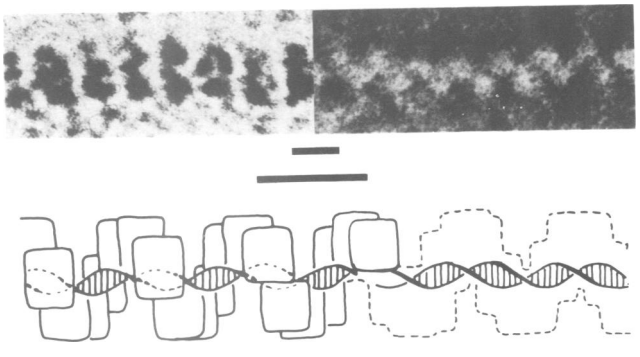

Fig. 1. Complexes of recA protein with double-stranded DNA. **Top**, electron micrographs of complexes adsorbed onto glow-discharged carbon support from a solution of 5 mM Mg acetate (Di Capua *et al.*, 1982). **Left**, shadowing with platinum; **right**, negative stain with uranylacetate. **Bottom**, summary sketch. Of the structure of the DNA in the complex only the helical parameters are known: 18.5 bp per turn of 97 Å pitch. The bases are drawn in a simplified way (bars: 100 Å). Complexes with single-stranded DNA have a similar appearance by electron microscopy. The structure of the DNA inside is not characterized.

© IRL Press Limited, Oxford, England

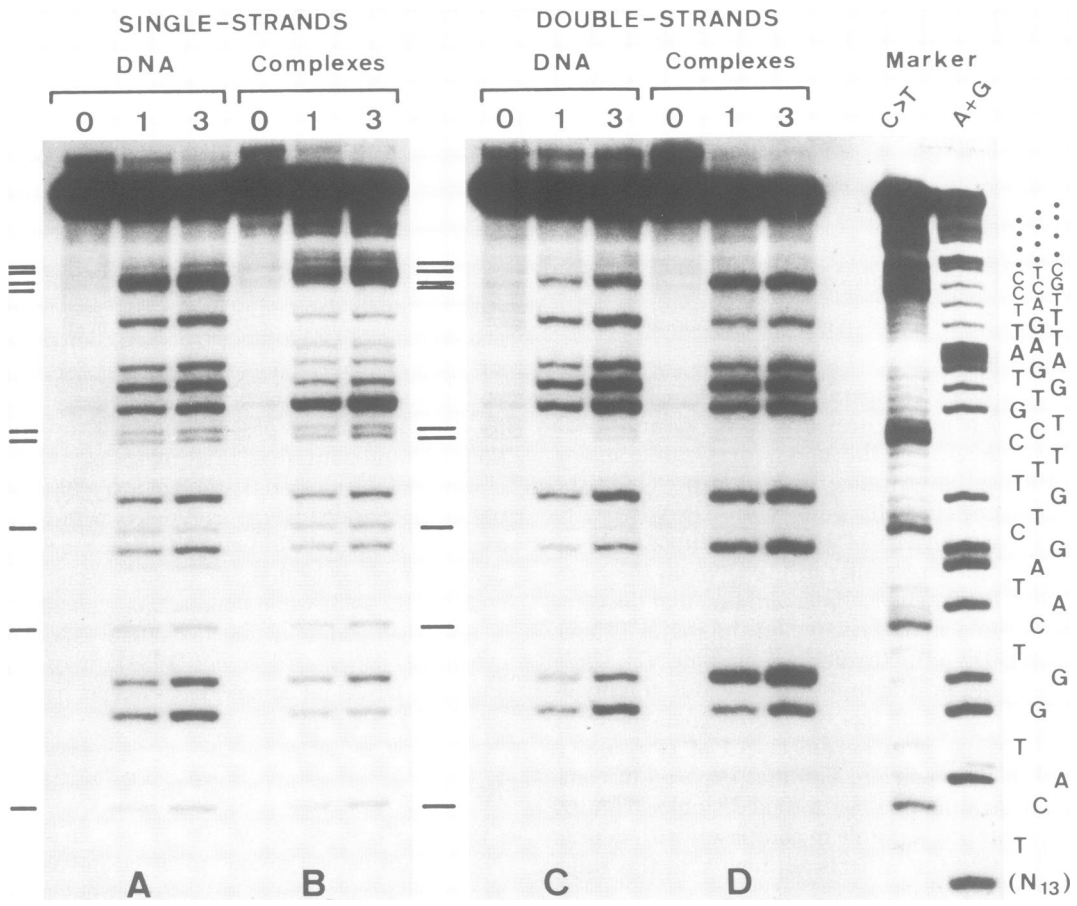

Fig. 2. Methylation of cytosine N3 in recA complexes. Acrylamide sequencing gel (20%) showing the bands resulting from specific cleavage of cytosines methylated at position N3. Because of the piperidine cleavage step, the methylated guanines appear also in this gel (strong bands); the positions of the cytosines are indicated by lines across the figure. The numbers at the top of the lanes indicate the time of dimethylsulfate treatment (0, 1, 3 min at 20°C). The single-stranded DNA was found not to renature under the conditions of complex formation and methylation (not shown).

Methylation of purines

In right-handed DNA helices, N7 of guanine lies in the major groove and is methylated very efficiently by dimethylsulfate. N3 of adenines is in the minor groove and is methylated at a slower rate, so that in B-DNA ~18% of the methylation is found on adenine N3, 78% on guanine N7, and very little on other residues (Lawley and Thatcher, 1970). This ratio was found also in our experiments. Figure 3C displays the DNA fragments obtained after chemical cleavage specific for methylated purines in double-stranded DNA. At any one time of the reaction with dimethylsulfate (lanes 1, 2, 4) we observe strong bands at the positions of guanines, and weaker bands at the positions of adenines. This observation was quantitized by scanning the lanes: the areas of the peaks were determined, the average of all adenine bands calculated and compared with the average of guanine bands; the result is displayed in Table Ia: in all six sequences that were used in this study, ~18% of the methylation in purines was found on adenines and 82% on guanines. A similar ratio was found for single-stranded DNA (Figure 3A, Table Ib).

In contrast, methylation of complexed DNA led to a different relationship: bands of methylated guanines are stronger, bands of methylated adenines are weaker or disappear altogether (Figure 3D). Quantitation of the relative methylation yields leads to ~94% of methylation on guanines, while ~6% is found on adenines in complexes with double-stranded as well as with single-stranded DNA (Table I).

This effect could be due to either of two mechanisms: (i) guanines are more exposed and adenines are protected, or (ii) only guanines are much more accessible, while adenines remain unaffected. On the gel of Figure 3, comparison of the lanes of equivalent times of methylation in naked versus complexed DNA suggests that the former case is true. This result could be biased in several ways. We always took care (i) to load carefully the same amount of DNA in each slot, (ii) to keep reaction conditions with dimethylsulfate very reproducible and (iii) we checked whether the mere presence of protein had an effect on the rate of methylation of naked DNA. The latter control can be done easily by mixing double-stranded DNA and recA protein in the absence of $ATP_{\gamma}S$; under these conditions, DNA is not bound as assayed by electron microscopy and by protection from DNases. We found no difference in methylation rate of DNA whether in the presence or in the absence of unbound protein. To the best of our ability, we thus made identical time points of methylation comparable to allow quantitative estimation of the relative rates of methylation in complexes versus naked DNA.

Figure 4 shows the quantitative results obtained from the analysis of six double-stranded fragments. Each bar represents the ratio of the area of a particular residue at a given time point of dimethylsulfate treatment in complexes versus naked DNA (see Materials and methods); guanine residues are shown as full lines, adenines as dotted lines. We see that most of the guanines have enhanced, while most of the adenines have reduced methyla-

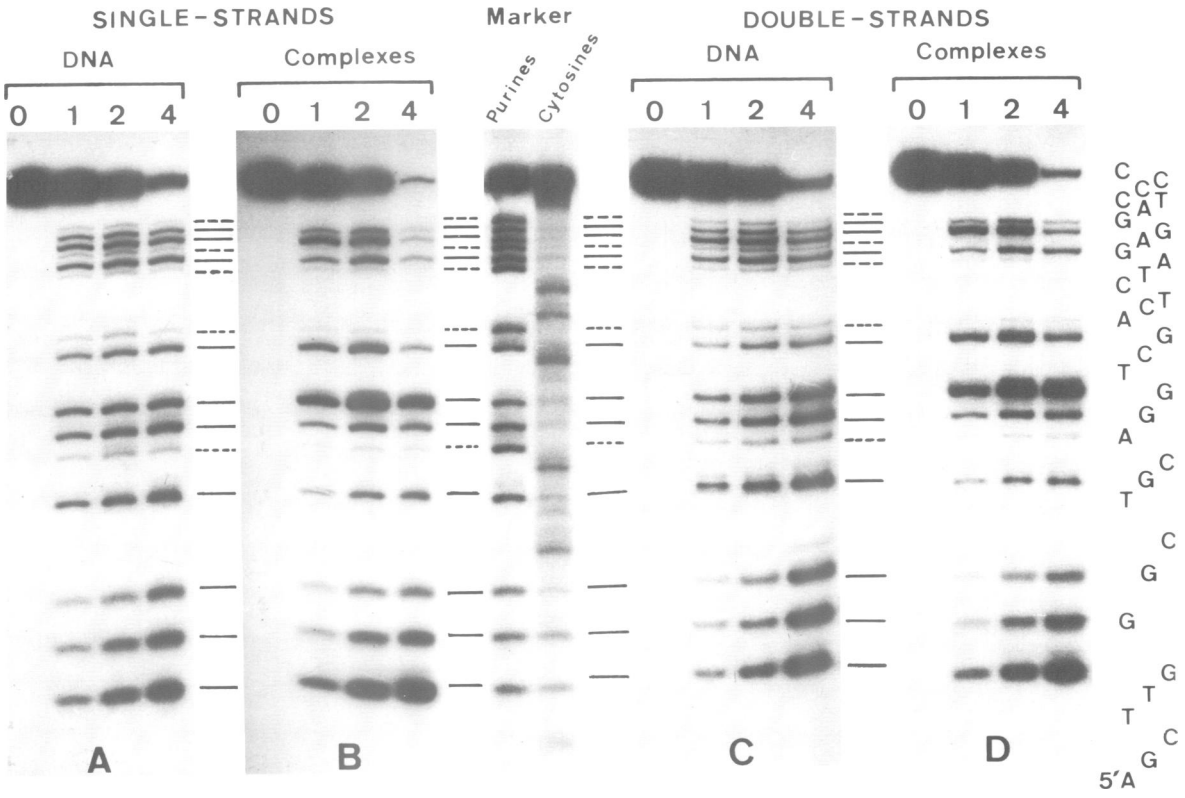

Fig. 3. Methylation of purines in recA complexes. The 20% acrylamide sequencing gel shows the bands resulting from specific cleavage at methylated guanines and adenines. The numbers indicate the time points (0, 1, 2, 4 min at 37°C) of dimethylsulfate treatment for naked or complexed DNA. Guanines are shown across the figure as full lines, adenines as dotted lines. On the right, the DNA sequence is shown.

tion. Averaging the results from all the fragments, we observe a 1.72-fold increase in guanine methylation (from 66 residues) and a 2.15-fold decrease in adenine methylation (from 40 residues). The fact that six different experiments (six different fragments) lead to the same result supports the validity of the approach of comparing equivalent lanes with each other: if methylation were uncontrolled, each experiment would lead to a different distribution of ratios.

The same analysis performed on the single-strands revealed a 1.47-fold increase of guanine methylation with 2/3 of the ratios lying between 0.61 and 3.42 (from 54 residues) and a 5-fold decrease of adenine methylation with 2/3 of the ratios lying between 0.07 and 0.52 (from 23 residues). The ratios with the single-stranded DNA scatter much more than the values for double-stranded DNA. Indeed, ~1/3 of the guanines were less methylated in complexes than in naked DNA, as illustrated in the fragment of Figure 2, an extreme case where seven out of nine guanines were methylated more weakly in the complex than in naked DNA. This irregular behaviour may be a feature of the structurally badly defined binding of recA to single-stranded DNA, reflected in the varying binding stoichiometries and contour lengths described in the literature (e.g. Flory *et al.*, 1984, Bryant *et al.*, 1985; discussed in Di Capua and Koller, 1987).

A side observation on the gels (e.g. Figure 3) is that, within a fragment, some G-residues get methylated much more strongly than others; this is true for naked DNA as well as DNA in complexes. More data would be necessary to analyze the rules governing these patterns, which are expected to be due to sequence-directed DNA structure at least in the case of naked DNA. In the case of complexes with double-stranded DNA, we found that 11 of the 13 guanines that are especially strongly

Table I. Relative amount of methylation found on adenines (expressed as % of total amount in purines) as determined from scans of gels

(a) Double-stranded DNA

Fragment (cf. Figure 4)	% Methylation found on adenines		Number of residues considered	
	DNA	Complexes	Adenines	Guanines
○	22	8[a]	5	14
●	18	11	3	7
△	16	7	8	8
▼	15	6	12	12
□	22	5	5	8
■	18	7	7	8
Average	18.5	7.3	40	57

(b) Single-stranded DNA

Fragment (cf. Figure 4)	% Methylation found on adenines		Number of residues considered	
	DNA	Complexes	Adenines	Guanines
○	20	9	5	12
●	14	5	3	9
△	21	2	8	8
▼	12	3	9	12
□	20	3	4	7
■	12	8	7	8
Average	16.5	5	36	56

[a]One A residue was reproducibly aberrantly methylated in this fragment.

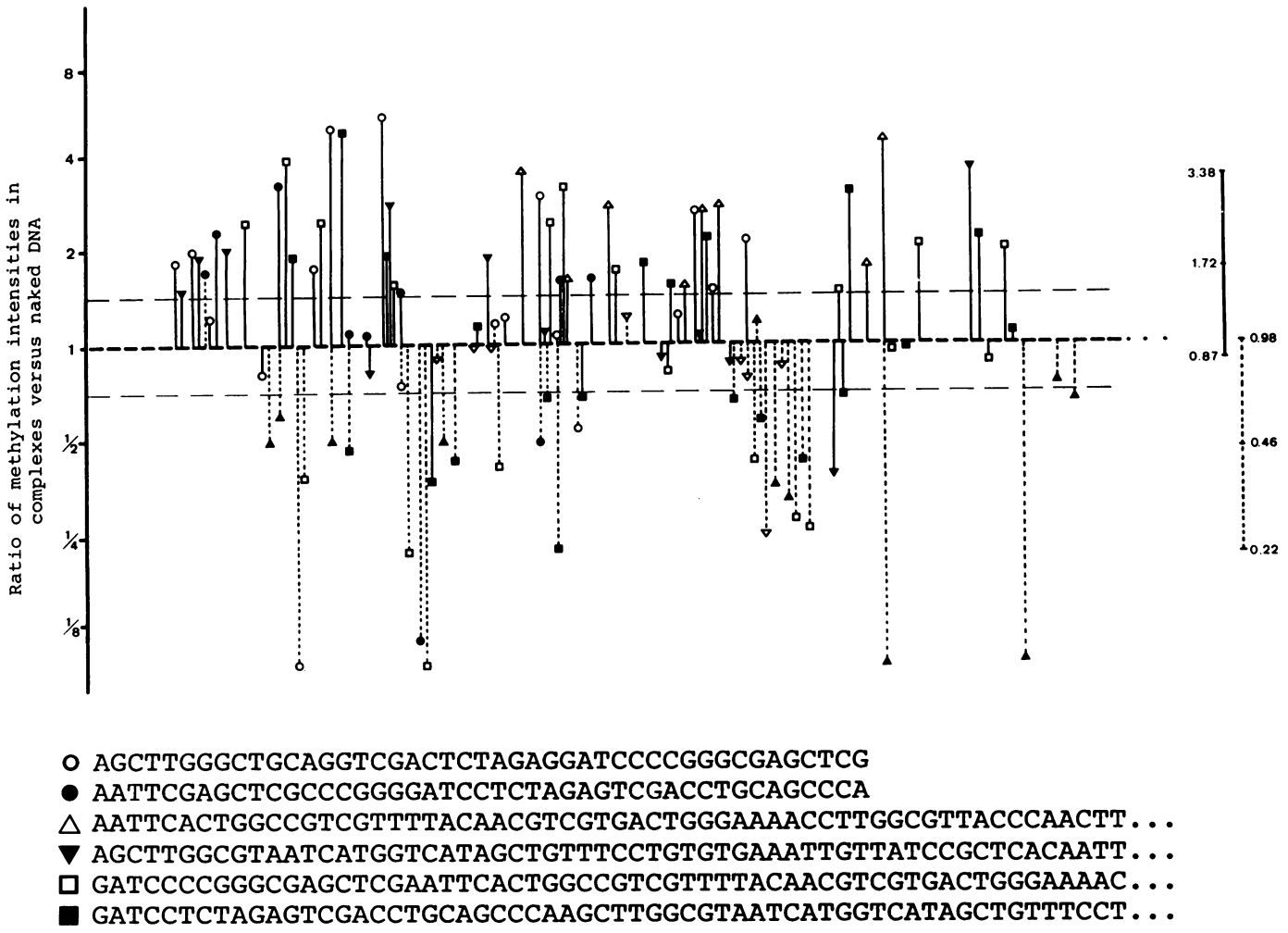

Fig. 4. Quantitation of methylation intensity on guanines and adenines in complexes versus DNA. The lower part of the figure shows the sequences used for the analysis. The symbol at their left is used at the tips of the bars in the graph to indicate from which fragment each arises. In the graph, data from all the fragments are drawn together. The bars express the magnitude of the difference in methylation in complexes versus DNA as the ratio of the intensities of the bands in the gels in a logarithmic scale. Full lines are guanines, dotted lines adenines. The bars on the right display the mean and SD for guanines (full line) and adenines (dotted line). Error estimation: the thin dashed lines at a ratio of 1.43 and 1/1.43 represent our estimate of the systematic error due to the imprecision of the method, as obtained independently of the methylation analysis (see Materials and methods). The amplitude of a bar within these boundaries can be considered as non-significant. It is, however, worth noting that even in this range most guanines are enhanced and most adenines decreased.

methylated in complexes (from all the fragments of Figure 4), are at position nx3 when counting from the 5′ end of the fragment. This frame of three is reminiscent of the binding stoichiometry of one recA per 3 bp in complexes. It suggests that the fragments get covered processively from precisely the 5′ end (Register and Griffith, 1985; Cassuto and Howard-Flanders, 1986) and that the three positions of the DNA triplet in contact with one recA subunit are not equivalent.

Discussion

The results of this study of accessibility of DNA to dimethylsulfate in complexes with recA protein show that the protein coat does not keep the solvent from the DNA, in contrast to the steric hindrance it represents for DNases (Stasiak *et al.*, 1983; Bryant *et al.*, 1985).

N3 of cytosines remains accessible in complexes with single-stranded DNA. The DNA in the complexes appears to expose the bases to base-pairing with the incoming partner. On the other hand, N3 of cytosines does not become accessible in complexes with double-stranded DNA. This speaks against a mechanism in which the DNA is prepared to homologous recognition of the bases by melting of the base-pairs.

The relative methylation rates of guanine and adenine change significantly. Leahy and Radding (1986) have performed a similar study on complexes with a single-stranded oligomer of 20 nucleotides. They concluded from their results that DNA is accessible in the complex. Our results show considerably more details: the enhanced methylation of guanines and the decreased methylation of adenines in complexes suggest that there are two 'faces' of the DNA ribbon, and that through the interaction with recA one face is made more reactive, while the other is relatively protected. The faces of double-stranded B-DNA are its major groove with guanine N7 and its minor groove with adenine N3. In the complex with recA, the DNA becomes stretched by a factor of 1.5, and unwound by 15°/bp, so we cannot speak of major and minor grooves anymore; however, the faces appear to be maintained.

The stretching and the unwinding are features typical for intercalating drugs; these have been found to intercalate from the minor groove (reviewed in Sobell *et al.*, 1983). The relative pro-

tection of the minor groove side of the double helix suggests that recA may act upon the DNA in a way similar to intercalation, pushing from the minor groove side. There exists, however, no indication as to whether an amino acid residue or the ATP moiety become inserted.

On the other face, guanine N7 reactivity is increased. This could be a property of the unwound conformation of the DNA. Indeed, N7 of guanine is increasedly methylated in the presence of ethidium bromide (E. Di Capua and B.Müller, unpublished). Another interpretation is that the complex establishes in its deep groove an environment that makes this residue more susceptible to nucleophilic attack. This may be relevant in the context of search for homology: N7 of guanine is proposed to allow base recognition via the major groove (reviewed in Morgan, 1979).

In the complex with single-stranded DNA, the relative reactivity of guanines and adenines are changed in the same sense as in complexes with double-strands, albeit less reproducibly so. This suggests that despite the absence of grooves in the single-strand, the DNA is 'reared' in a similar way, with one side of the ribbon covered by the protein, and the other side exposed in the cleft of the complex.

In summary, our results support the model that was proposed on the basis of electron microscopy (Figure 1, Stasiak *et al.*, 1983), with a DNA largely accessible in the groove of the recA complex. They also suggest a high potential for sequence homology recognition through N3 of cytosine and N7 of guanine. Finally, the relative protection of N3 of adenine supports a model in which recA binds along the minor groove of the DNA.

Materials and methods

DNA

Double-stranded DNA fragments were obtained by restriction digests of pUC13 plasmid with *Eco*RI, *Bam*HI or *Hind*II which all leave four protruding nucleotides at the 5′ end. This end was labelled with kinase and the DNA resected with a second enzyme. The sequences analyzed are shown on Figure 4. The size of the fragments was 45, 77, 93, 145 and 230 bp. recA protein was purified from strain KM 4104/pDR 1453 by the procedure of Cox *et al.* (1981) to fraction III. The concentration was determined using a E_{277}, 1% of 6.33 (Tsang *et al.*, 1985).

Complexes and methylation

Complexes were formed with DNA fragments end-labelled at one of the 5′ ends. Reaction conditions were (in 100 μl): 20 μM DNA in nucleotides, 20 μM recA protein, 2 mM $MgCl_2$, 0.5 mM $ATP_\gamma S$, 25 mM sodium cacodylate at pH 6.3. Complexes with double-stranded DNA were incubated for 20 min at 37°C; complexes with single strands (obtained by boiling the DNA just before mixing into the reaction) were incubated for 10 min at 37°C. In control experiments (not shown), no renaturation was found to occur under these conditions, either in the presence or in the absence of recA. After incubation, an aliquot was tested by treatment with excess DNases (DNase I to 0.1 mg/ml, phosphodiesterase I to 0.02 U/ml, incubation 10 min at 37°C). More than 95% of the DNA was protected.

Methylation with dimethylsulfate was with 1 μl per 100 μl solution. Less than one hit per fragment (>80% of the DNA intact) was achieved by incubation for 1 min at 20°C. Stronger methylation (almost all DNA degraded) was obtained with incubations up to 4 min at 37°C. The reaction was stopped with mercaptoethanol to 0.2 M, carrier tRNA to 20 μg/100 μl and ethanol precipitation. The DNA was then purified by phenol extraction in the presence of 0.1% SDS and repeated ethanol precipitation. Methylation and processing of naked DNA was exactly the same as that of complexed DNA, in either the absence of protein, or in the absence of $ATP_\gamma S$ but in the presence of protein.

Specific cleavage reactions

The cleavage reaction specific for methylated purines was according to Sienbenlist and Gilbert (1980): heating for 10 min at 95°C in 10 mM potassium phosphate (pH 7.0) followed by 10 min at 95°C after addition of sodium hydroxide to 0.1 M.

The reaction specific for methylated cytosines (Kirkegaard *et al.*, 1983) was 1 volume of DNA in H_2O, 1 volume dioxane and 2 volumes hydrazine, incubated at 0°C for 1 min. The hydrazine was removed by repeated EtOH precipitations and the DNA cleaved by the standard piperidine reaction (Maniatis *et al.*, 1982).

Analysis of the cleavage products was by electrophoresis on standard sequencing gels (Maniatis *et al.*, 1982). Care was taken to load the same amount of radioactivity in all slots of an experiment in order to allow comparison between tracks. The resulting autoradiographs were analyzed quantitatively by scanning on a Shimadzu CS 930 t.l.c. scanner. The peak area for each band was determined and used for calculations. The resolution power of scanning is low. Gels with 20% acrylamide were used for the separation of bands of 5–30 nucleotides, and 12% was used for the range 10–50 nucleotides. Longer fragments could easily be evaluated by eye, but the area measurements became imprecise. In several cases more than one pair of equivalent tracks was compared. The result was found not be dependent on the extent of methylation, as long as tracks with the same time of methylation were compared. Early timepoints (1 min at 20°C) were used for the calculation of Figure 4.

The ratios in Figure 4 were obtained by dividing the intensity of a band in complexes by the intensity in naked DNA. Figure 4 represents our data in a logarithmic scale. We chose the arithmetic mean and the standard deviation of the natural logarithms, because the natural logarithms of the ratios were found to be normally distributed. Expressed in rational numbers, these values yield the geometric mean of the ratios and an asymmetric interval around this mean, which spans ~2/3 of the ratios used for the calculation.

The ratios for adenines are inherently inaccurate due to the low intensities of the bands. Some of the adenines in complexes became indetectable.

Estimation of the systematic error

The systematic error arising from the imprecision of gels and scanning of bands was estimated independently from methylation in the following way. The same sequence was studied in four experiments. Eighteen well-resolved bands from this sequence were considered; the intensity of a band was determined by scanning, and the area was normalized to the average area of all bands in the lane, in order to allow comparision between experiments. For each residue, a ratio was calculated between the largest and the smallest value found for its normalized intensity among the four experiments. The mean of the natural logarithms from these ratios was 0.36 ± 0.14, which corresponds in rational numbers to a mean of 1.43 and an interval ranging from 1.25 to 1.63. It is shown in Figure 4 as dashed lines at ratios 1.43 and 1/1.43.

Acknowledgements

We are grateful to Ch.Radding and co-workers for communicating results prior to publication, which enabled us to finish our experiments in a shape that allowed comparision with theirs. We thank Th.Koller, A.Stasiak and H.Weber for suggestions and support, and H.Mayer-Rosa for help in the final version of the figures. This work was supported by Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (grant to Th.Koller).

References

Bryant,F.R., Taylor,A.R. and Lehman,I.R. (1985) *J. Biol. Chem.*, **260**, 1196–1202.
Cassuto,E. and Howard-Flanders,P. (1986) *Nucleic Acids Res.*, **14**, 1149–1157.
Cox,M.M., McEntee,K. and Lehman,I.R. (1981) *J. Biol. Chem.*, **256**, 4676–4678.
Di Capua,E., Engel,A., Stasiak,A. and Koller,Th. (1982) *J. Mol. Biol.*, **157**, 87–103.
Di Capua,E. and Koller,Th. (1987) In Lilley,D. and Eckstein,F. (eds), *Nucleic Acids and Molecular Biology*. Springer, Berlin, Vol. 1, pp. 174–185.
Flory,J., Tsang,S.S. and Muniyappa,K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7026–7030.
Howard-Flanders,P., West,S.C. and Stasiak,A. (1984) *Nature*, **309**, 215–220.
Kirkegaard,K., Buc,H., Spassky,A. and Wang,J.C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2544–2548.
Koller,Th., Di Capua,E. and Stasiak,A. (1983) In Cozzarelli,N.R. (ed.), *Mechanisms of DNA Replication and Recombination*. Alan R.Liss, New York, pp. 723–729.
Lawley,P.D. and Thatcher,C.J. (1970) *Biochem. J.*, **116**, 693.
Leahy,M.C. and Radding,C.M. (1986) *J. Biol. Chem.*, **261**, 6954–6960.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
Morgan,A.R. (1979) *Trends Biol. Sci.*, **October**, 244–247.
Register,J.C.,III and Griffith,J. (1985) *J. Biol. Chem.*, **260**, 12308–12312.
Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 122–126.
Sobell,H.M., Sakore,T.D., Jain,S.C., Banerjee,A., Bhandary,K.K., Reddy,B.S. and Lozansky,E.D. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 293–314.
Stasiak,A. and Di Capua,E. (1982) *Nature*, **299**, 185–186.
Stasiak,A., Di Capua,E. and Koller,Th. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 811–819.

Tsang,S.S., Muniyappa,K., Azhderian,E., Gonda,D.K., Radding,C.M., Flory,J. and Chase,J. (1985) *J. Mol. Biol.*, **185**, 295–309.
West,S.C., Cassuto,E., Mursalim,J. and Howard-Flanders,P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2569–2573.

Received on February 11, 1987