Dimerization of *Escherichia coli* UvrA and Its Binding to Undamaged and Ultraviolet Light Damaged DNA†

Sharlyn J. Mazur* and Lawrence Grossman‡

Department of Biochemistry, The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland 21205

Received October 15, 1990; Revised Manuscript Received January 30, 1991

ABSTRACT: The initial stages in the repair of damaged DNA by the *Escherichia coli uvr* system involve the recognition of damage by UvrA. We have examined in detail the binding of UvrA to DNA randomly damaged by ultraviolet light, undamaged DNA, and single-stranded DNA using nitrocellulose filter binding and gel mobility shift assays to arrive at the following model: UvrA dimers bind specifically to damaged DNA both in the presence and in the absence of ATP. The dimerization of UvrA is promoted by UvrA concentrations greater than 1 nM, the presence of ATP, or physiological temperatures, and the dimerization step dominates the temperature dependence of UvrA binding to DNA damaged by ultraviolet light. The apparent association constant for specific binding is dependent on the concentration of UvrA due to coupled dimerization, aggregation, and nonspecific binding reactions. At 1 nM UvrA, either with or without ATP, $K_{uv} \approx 10^9$ M^{-1}. The binding of UvrA to undamaged DNA is 10^3–10^4-fold weaker than the damage-specific binding. Both the strength of damage-specific binding and the discrimination between damaged and undamaged sites are affected by the salt concentration. The kinetics of association and dissociation reactions indicate that the primary effects of ATP are on the extent of UvrA dimerization rather than on the properties of the UvrA–uvDNA complex. The complexity of the interaction of UvrA, ATP, and DNA is indicated by the opposing effects of ATP binding and hydrolysis on UvrA dimerization.

In *Escherichia coli*, the UvrABC endonuclease initiates the repair of "bulky adducts" formed by compounds such as photoactivated psoralens, activated polycyclic aromatic hydrocarbons, *N*-acetoxy-2-(acetylamino)fluorene, mitomycin C, and *cis*-platin, in addition to the repair of photoproducts produced by ultraviolet light (Sancar & Sancar, 1988; Van Houten, 1990). Repair systems with a similar range of specificity are found in yeast and in humans (Friedberg, 1985).

A general understanding of the incision mechanism has developed (Grossman & Yeung, 1990; Selby & Sancar, 1990; Grossman et al., 1988), although fundamental details remain obscure. Since the primary chemical structures of the damage repaired by the *uvr* system have no structural elements in common, it is generally assumed that the enzyme recognizes the damage-induced distortion in the DNA helix (Pearlman et al., 1985). However, variations in the structure of undamaged DNA (Dickerson, 1983) and in the structure of the damaged site will limit the ability of the enzyme to discriminate between substrate and nonsubstrate DNA. UvrA, the only subunit that by itself binds specifically to damaged DNA, displays a limited preference for damaged over undamaged DNA (Seeberg & Steinum, 1982; Van Houten et al., 1987). For other systems in which the biological requirement is greater than the equilibrium discrimination, a kinetic element provides an additional level of discrimination (Hopfield, 1974; Ninio, 1975). In the *uvr* system, the requirement for ATP hydrolysis for overall incision in vitro suggests that the role of ATP hydrolysis is to increase specificity (Seeberg & Steinum, 1982; Yeung et al., 1983; Van Houten et al., 1988). However, the recent demonstration of a helicase activity of the UvrAB complex (Oh & Grossman, 1987, 1989) and the involvement of the cryptic ATPase activity of UvrB (Caron & Grossman, 1988) in the UvrAB helicase activity (Seeley & Grossman, 1989, 1990) suggest that the role of ATP in the incision mechanism is both complex and subtle.

To understand the mechanics of the incision reaction, we have examined in detail the binding reactions of UvrA with DNA damaged by ultraviolet light, undamaged DNA, and single-stranded DNA. These reactions are affected by the concentration of UvrA and by solution conditions such as temperature, the salt concentration, and the presence of ATP. The results indicate that UvrA binds specifically to UV-damaged sites as a dimer. ATP influences the dimerization but is not required for dimerization or specific binding. The specificity of UvrA exhibited in vitro is insufficient to explain the functioning of the UvrABC system in vivo, providing further evidence for the involvement of the UvrAB helicase in the mechanism of damage recognition in vivo.

Materials and Methods

UvrA. UvrA protein was purified as described previously (Yeung et al., 1986b) and further purified by spin dialysis to remove ATP with an Aminco Centricon 30.

DNA. The plasmids pHE6 (Milman, 1987) and pBR322 were prepared by chloramphenicol amplification. The runaway replication plasmid pPYC3 containing the 285 base pair (bp)[1] lac P–O insert (Yeung et al., 1983) was prepared from an overnight culture. The plasmids were purified by standard

† This work was supported by grants from the National Institutes of Health (GM 31110 and GM 22846) and from the U.S. Department of Energy (DE FG02-86ER60396) to L.G. An NSRA Training Grant (5 132 ES 07141-05 0201) from the National Institutes of Health provided support for S.J.M.

* Correspondence should be addressed to this author at the Department of Chemistry, The American University, Washington, DC 20016.

‡ Author to whom reprint requests should be addressed.

[1] Abbreviations: ATPγS, adenosine 5′-*O*-(3-thiotriphosphate); BB, binding buffer; bp, base pair(s); BSA, bovine serum albumin; dsDNA, double-stranded DNA; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FB, filter buffer, MOPS, 3-(*N*-morpholino)propanesulfonic acid; ssDNA, single-stranded DNA; Tris, tris(hydroxymethyl)-aminomethane; uvDNA, ultraviolet light irradiated DNA.

0006-2960/91/0430-4432$02.50/0

methods (Tanaka & Weisblum, 1975; Maniatis et al., 1982) including two equilibrium cesium chloride/ethidium bromide centrifugation steps. To minimize ethidium-mediated damage to the DNA, faint, nonfluoresent illumination was used for the recovery of the supercoiled plasmid band and the removal of ethidium bromide. RNA was removed by gel filtration. The concentrations of stock solutions were determined by the ultraviolet absorbance at 260 nm using the extinction coefficient 1.32×10^4 M^{-1} cm^{-1} (base pairs). Plasmid pBR322 prepared without chloramphenicol amplification and without cesium chloride/ethidium bromide centrifugation was obtained from Applied Genetics. Calf thymus DNA (Sigma) was used as a competitor. For competition experiments, plasmid DNA was cleaved with the restriction enzymes *Bam*HI or *Alu*I and purified by phenol extraction and ethanol precipitation.

The 285 bp *lac* P–O fragment contains a tandem repeat of the *lac* operator region of *E. coli* (Bogosian & Kane, 1987) and was cloned into the *Eco*RI site of the runaway replication plasmid pRLM24 (Yeung et al., 1983). The plasmid was digested overnight with *Eco*RI, and the fragment was separated from the 10.1 kbp vector by sucrose gradient centrifugation (Hardies & Wells, 1979). The fragment was labeled at the 3′ ends with [α-^{32}P]dATP using the filling-in function of T4 DNA polymerase. Labeled fragment was purified by extraction with phenol and phenol/chloroform, followed by gel filtration to remove unincorporated nucleotides (Maniatis et al., 1982). The specific activity of the fragment was determined by comparing the ethidium bromide fluorescence of the labeled fragment and closely spaced standards in a 2% agarose gel stained with ethidium bromide. Photographic negatives (Polaroid type 55) of the ethidium bromide fluorescence were analyzed with a Helena Laboratories quick scan densitometer. Since the response of photographic film to fluorescent signals is quite nonlinear (Pulleyblank et al., 1977), the standards were closely matched to the labeled sample.

Unlabeled and tritium-labeled single-stranded fd DNAs were prepared as described (Yeung et al., 1986a). The phage were precipitated by sodium chloride and poly(ethylene glycol) (Messing, 1983). The DNA was purified by extraction with phenol and phenol/chloroform (1:1) and dialyzed against 10 mM NaCl, 10 mM Tris·HCl, and 1 mM EDTA (pH 8.0). The concentration was determined spectrophotometrically using the value $\epsilon_{260} = 1.8 \times 10^4$ M^{-1} cm^{-1} (bases). The specific activity was 2×10^4 cpm/μg.

UV Irradiation. DNA samples were irradiated by ultraviolet light from a germicidal lamp at 1.9 J s^{-1} m^{-2}. Irradiation of the 285 bp fragment to 1.25 kJ/m^2 produces an average of two damaged sites per fragment (Haseltine et al., 1980).

Nucleoprotein Complex Formation. The standard binding buffer (BB) is 0.1 M KCl, 10 mM $MgCl_2$, 10 mM K-MOPS (pH 7.6), 1 mM EDTA, 5% glycerol, 1 mM DTT, and 50 μg/mL BSA. Filter buffer (FB) is BB without glycerol, DTT, or BSA. Changes from standard binding buffer are indicated in the figure legends. A low-salt binding buffer used for the gel shift experiments, 0.5 × BB, contains 50 mM KCl and 5 mM $MgCl_2$, and the remaining components are the same as in BB.

Generally, experiments were done in the absence of nucleotide or in the presence of 1 mM ATP. The ATPase activity of UvrA is affected by the type and concentration of DNA (Oh et al., 1989). The fraction of ATP hydrolyzed during incubation is less than 10% for the highest UvrA concentrations used and is much less than that for most of the experiments.

UvrA was added to labeled DNA in BB previously equilibrated to the incubation temperature under conditions appropriate to the specific experiment (see below and figure legends). After 10–15 min of incubation, unlabeled, unirradiated calf thymus DNA was added to a final concentration of 100 μM bp for a 10-s challenge immediately prior to filtration or loading on a gel. If the competing DNA is added before the UvrA, no significant binding to the labeled DNA is observed. However, in the brief challenge, only the rapidly dissociating complexes are removed.

Nitrocellulose Filter Binding. Nitrocellulose membrane filters (Schleicher & Schuell BA-85, 0.45 μm) were soaked in FB prior to use. For experiments with single-stranded DNA, the filters were soaked in 0.4 M KOH for 20 min, then rinsed extensively with distilled water, and stored at 4 °C in FB until used. Aliquots ranging from 5 to 500 μL were drawn through nitrocellulose filters by a negative pressure difference of 50–60 torr, and the filters were washed with 500 μL of FB. The filters were dried under heat lamps, and the radioactivity was determined by scintillation counting using toluene-based scintillation fluid.

Three factors must be evaluated to interpret the nitrocellulose filter binding data quantitatively. First, DNA–protein complexes are often not retained quantitatively by nitrocellulose filter binding (Riggs et al., 1970; Woodbury & von Hippel, 1983). Second, some small fraction of the DNA is retained even in the absence of specific binding protein. Finally, the damage produced in DNA by UV irradiation is located approximately at random, resulting in a Poisson distribution in the number of lesions per fragment. The number of proteins specifically bound to these sites is also described by a Poisson distribution. For a Poisson distribution in the number of complexes per fragment, the fraction of DNA molecules with n proteins bound is given by

$$f_n = e^{-\nu}\nu^n/n! \quad (1)$$

where ν is the average number of proteins bound per fragment. If each bound protein independently and identically imparts the probability (ϵ) that the fragment is retained by the filter (Woodbury & von Hippel, 1983), then the fraction passing through the filter is

$$f_0 = e^{-\epsilon\nu} \quad (2)$$

Correcting for the fragments retained in the absence of UvrA (*blank*) yields the expression for the average number of proteins bound per fragment:

$$\epsilon\nu = -\ln [(total - sample)/(total - blank)] \quad (3)$$

The *total* is determined by delivering an aliquot to a dry filter. The value of the *blank* is obtained from samples with no UvrA added and generally was around 2% of the *total*.

The filter efficiency can vary from one lot of nitrocellulose to another. Although the filter efficiency remained around 0.6 for most of the experiments, the last two figures contain results obtained with an aberrant lot of nitrocellulose. The filter efficiency was not determined precisely over the range of temperatures investigated, but the results require the filter efficiency to be considerably higher than 0.6. For simplicity, the filter efficiency with this batch of nitrocellulose was assumed to be 1.

Gel Mobility Shift Assay. Complexes were typically resolved on 2% agaorse gels using TA buffer (40 mM Tris–acetate/1 mM EDTA, pH 8.0) at about 7 V/cm in a Hoeffer Mighty Small vertical gel apparatus. Complexes can also be resolved in Tris–glycine electrophoresis buffer, in 4% acryl-

amide gels, and at a variety of field strengths.

Three methods for determining the average number of proteins bound per fragment are given here. Without any assumptions, the average number of proteins bound per fragment, ν, can be calculated from the distribution of DNA in the various bands as a direct average:

$$\nu = \sum n f_n / \sum f_n \quad (4)$$

where f_n is the fraction of DNA fragments that has n proteins bound. The other two methods assume the bound proteins are described by a Poisson distribution. The average number of proteins bound per fragment can be calculated from the ratio of the first complex band to the free DNA band:

$$f_1/f_0 = \nu e^{-\nu}/e^{-\nu} = \nu \quad (5)$$

Alternatively, the fraction free can be used to determine ν:

$$\nu = -\ln f_0 \quad (6)$$

The total radioactivity in the sample can be determined from the total counts in an aliquot or from the sum of the activity of all bands in a lane.

Stoichiometry Determination. The specific activity of the DNA fragment was determined by isotopic dilution. A small quantity of ^{32}P-labeled 285 bp fragment of high specific activity was mixed with a large amount of unlabeled fragment. The concentration of the DNA stock solution was determined spectrophotometrically. UvrA at 100–300 nM was mixed with UV-irradiated, ^{32}P-"spiked" 285 bp fragment and incubated at 37 °C for 15 min in BB without added BSA. Competing DNA was added to 1 mM (bp), and the sample was dialyzed by floating filter (Millipore type VS) against 0 °C TA buffer for 30 min. Omission of the dialysis step resulted in gels that had curved bands, but were otherwise the same. Addition of 5% glycerol without tracking dye facilitated loading the samples onto a 2-mm-thick vertical gel of 2% agarose. Complexes were resolved at 8 V/cm for about an hour. Duplicate lanes were used for protein and DNA analysis. The distribution of DNA was determined by slicing the gel into 2-mm pieces, dissolving each piece in 3 mL of scintillation fluid, and determining the radioactivity by liquid scintillation counting. The protein in the complex band was quantified by staining the gel for 90 min with Bio-Rad protein assay mix (Bio-Rad Laboratories) in the proportions 1 volume of gel/1 volume of H_2O/0.5 volume of dye mix. The profile of the absorbance at λ = 595 nm was determined by densitometry (Helena Laboratories), and compared to the absorbance of a series of BSA standards present in the gel. The dye reponse of UvrA compared to that of BSA was based on spectrophotometrically determined concentrations of UvrA and BSA. Control experiments show that the presence of DNA has no effect on the binding of Coomassie Blue to UvrA under the conditions of the assay. The extinction coefficient for UvrA, $\epsilon_{280} = 5.6 \times 10^4$ M^{-1} cm^{-1}, was calculated from its sequence (Husain et al., 1986) using component values for neutral 6 M guanidine hydrochloride (Edelhoch, 1967). This value is 20% higher than the value reported by Navaratnam et al. (1989).

Binding Experiments. The association constant for UvrA binding to UV-damaged DNA can be determined by measuring the extent of complex formation as a function of the protein concentration. The data are conveniently analyzed as a Scatchard plot (Scatchard, 1949) to describe the binding of a ligand (UvrA dimer, A_2) to a macromolecule with n sites (UV-irradiated DNA fragment):

$$\nu/[A_2] = Kn - K\nu \quad (7)$$

where ν is the average number of proteins bound per fragment and $[A_2]$ is the concentration of free UvrA dimer.

Kinetic Experiments. The rate of association of UvrA to UV-damaged DNA was measured by adding an aliquot of UvrA to a 37 °C solution of ^{32}P-labeled fragment DNA in binding buffer with or without ATP, to a final concentration of 100 pM UvrA (monomers) and 20 pM DNA fragment. After rapid mixing, the reaction solution is returned to a 37 °C water bath. Aliquots withdrawn at various times were challenged by the addition of 100 μM (bp) calf thymus DNA for 10 s before being filtered through nitrocellulose filters. The challenge removes UvrA bound to undamaged DNA. In standard binding buffer, the association reaction is generally observable when the UvrA concentration is in the range 0.02–0.10 nM and the concentration of DNA fragments is somewhat lower. The rate of complex formation is given by

$$d[A_2S]/dt = k_1[A_2][S] - k_{-1}[A_2S] \quad (8)$$

where $[A_2S] = \nu M_0$ is the concentration of specifically bound UvrA dimers, $[A_2]$ is the concentration of free UvrA dimer, and [S] is the concentration of free damaged sites. The reactant concentrations can be expressed in terms of the equilibrium concentrations and a variable, x, defining the distance from the equilibrium: $x = [A_2S]_{eq} - [A_2S] = [A_2] - [A_2]_{eq} = [S] - [S]_{eq}$. The rate law is then expressed:

$$\frac{dx}{x(c + x)} = k_1 t \quad (9)$$

where $c = [A_2]_{eq} + [S]_{eq} + K^{-1}$, and $K = k_1/k_{-1}$ is the equilibrium association constant. The integrated rate law can be expressed as a linear function of time (Weast, 1971):

$$\ln\left(\frac{c + x}{x}\right) - \ln\left(\frac{c + [A_2S]_{eq}}{[A_2S]_{eq}}\right) = k_1 c t \quad (10)$$

Association kinetic data were analyzed by linear regression of $\ln(c + x/x)$ as a function of time, assuming all of the UvrA was in the dimer form and assuming that the filter efficiency $\epsilon = 0.6$.

The rate of dissociation of UvrA–uvDNA complexes was measured by addition of competitor. In particular, 1 nM UvrA and 0.1 nM ^{32}P-labeled, UV-irradiated 285 bp fragment containing an average of two sites per fragment were incubated in BB at 37 °C for 15 min. The dissociation was initiated by adding UV-irradiated calf thymus DNA, at 37 °C and in BB, to a final concentration of 100 μM bp. The amount of UvrA bound to the labeled DNA as a function of time was determined by the nitrocellulose filter binding assay. Dissociation kinetic data were analyzed according to

$$\ln(\nu/\nu_0) = -k_{-1}t \quad (11)$$

where ν_0 is the number of specific complexes per fragment at the time the competitor was added.

The dissociation rate for UvrA–single-stranded DNA complexes was determined by the combined effects of dilution and addition of a competitor. Complexes formed in 37 °C BB, with or without 1 mM ATP, from 50 nM UvrA and 12 μM bases (2 nM molecules) single-stranded ^{3}H-labeled fd DNA were diluted 10-fold into 37 °C BB, with and without 1 mM ATP, containing 200 μM (base pairs) UV-irradiated calf thymus DNA. Aliquots taken at various times were filtered through KOH-treated nitrocellulose filters.

Competition Experiments. The extent of binding of the protein to a labeled fragment containing specific sites in the presence of an excess of unlabeled undamaged DNA is indicative of the binding constant to undamaged DNA. The

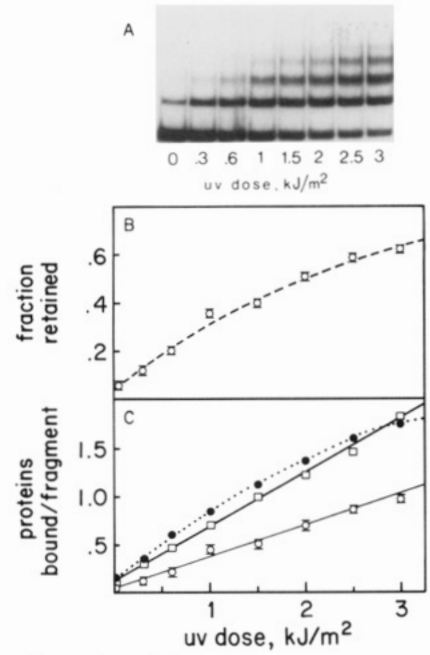

FIGURE 1: Comparison of the gel electrophoresis assay and the nitrocellulose filter binding assay for detection of UvrA binding to UV-damaged DNA. Solutions containing 10 nM UvrA and 1 nM ^{32}P-labeled UV-irradiated 285 bp fragment in BB were incubated at 37 °C for 15 min. Unirradiated calf thymus DNA was added as a competitor, and the samples were mixed for 10 s. Two volumes of 0 °C TA buffer with 5% glycerol were added, and 12-μL aliquots were quickly loaded onto a 2% agarose gel or filtered through nitrocellulose filters. (A) Autoradiogram resulting from gel shift assay. (B) Retention of UvrA–DNA complexes by nitrocellulose filters. Aliquots were filtered through nitrocellulose filters in triplicate. The total amount of radioactivity in an aliquot was 3.7×10^3 cpm. (C) Analysis of gel electrophoresis and nitrocellulose filter binding data. Nitrocellulose filter binding data (○) were transformed according to eq 3 to determine the product $\epsilon\nu$. Linear regression gives $\epsilon\nu = (0.32\ \text{kJ}^{-1}\ \text{m}^2)X + 0.06$ complexes per fragment ($r = 0.993$), where X is the dose in kilojoules per square meter. For the gel shift assay, the DNA-containing bands were cut from the dried gel, and the radioactivity was determined by liquid scintillation counting. The average number of proteins bound per fragment is calculated by a direct average (eq 4) (●) or from the ratio f_1/f_0 (eq 5) (□). The points obtained from the eq 4 are connected by a smooth dotted curve. Linear regression of the data analyzed using eq 5 gives $\nu = (0.55\ \text{kJ}^{-1}\ \text{m}^2)X + 0.14$ ($r = 0.999$).

inverse of the apparent specific equilibrium constant, K_{app}^{-1}, is a linear function of the concentration of competitor:

$$K_{app}^{-1} = \frac{([S]_0 - [A_2S])(\frac{1}{2}[A]_0 - [A_2S])}{[A_2S]} = \frac{1}{K_s} + \frac{K_N[N]}{K_s} \quad (12)$$

where $[S]_0$ and $[A]_0$ are the total concentrations of damaged sites and UvrA monomers, respectively.

If the competitor contains two classes of sites, a large fraction of nonspecific sites that bind the protein weakly and a small fraction of sites that bind the protein tightly, the competition plot can be linear or curved, depending on the relative strengths of specific, nonspecific, and tight binding modes (S. Mazur, unpublished calculations).

Results

Comparison of the Gel Mobility Shift Assay and the Nitrocellulose Filter Binding Assay. Specific complexes of proteins and DNA fragments containing the recognition sequence are often detected by nitrocellulose filter binding (Riggs et al., 1970), gel mobility shift assay (Garner & Revzin, 1981; Fried & Crothers, 1981), or DNase I footprinting methods (Brenowitz et al., 1986). The first two of these methods also are useful for detecting complexes of a DNA repair protein binding to randomly located sites of damage. The formation of damage-specific UvrA–DNA complexes as a function of the dose of ultraviolet light is depicted in Figure 1.

The distribution in the number of proteins bound per fragment can be seen directly from the gel shift experiment shown in Figure 1A. The fastest moving band in each lane corresponds to free DNA. The successive bands lying above the free DNA band contain one, two, and so on protein complexes bound per fragment. The number of proteins bound per fragment increases as the dose of ultraviolet light to the DNA increases, indicating that UvrA is bound specifically to sites of UV damage. Specific complexes also can be detected by nitrocellulose filter binding. DNA with no bound protein passes through the filter while DNA with one or more proteins bound is retained by the filter. In Figure 1B, the fraction of DNA retained by the filter increases with increased dose, again indicating the detection of UvrA bound specifically to damaged sites.

A quantitative analysis of the binding data is shown in Figure 1C. The average number of proteins bound per fragment can be determined directly from the distribution of DNA among the bands in the gel shift assay according to eq 4. Although these data are noticeably curved, linear regression gives $\nu = [(0.54\ \text{kJ}^{-1}\ \text{m}^2)X + 0.24]$ complexes/fragment with the correlation coefficient $r = 0.993$, where X is the UV dose in kilojoules per square meter. Assuming a Poisson distribution and using the ratio of the fraction of the DNA in the first DNA–protein band to the fraction in the free band according to eq 5 gives $\nu = (0.55\ \text{kJ}^{-1}\ \text{m}^2)X + 0.14$ with $r = 0.999$. Assuming a Poisson distribution and determining ν from the fraction free according to eq 6 gives $\nu = (0.55\ \text{kJ}^{-1}\ \text{m}^2)X + 0.44$ ($r = 0.97$) when the aliquot total is used to calculate the fraction free, and $\nu = (0.52\ \text{kJ}^{-1}\ \text{m}^2)X + 0.20$ ($r = 0.998$) when the sum of the bands total is used to calculate the fraction free. The agreement between the dose dependence of complex formation by the four methods is excellent, with only a 4% variation in the slope. However, the extrapolated extent of complex formation in the absence of UV irradiation varies by 50% among the four methods.

The nitrocellulose filter binding data are analyzed by assuming a Poisson distribution of bound protein. The results of transforming filter binding data using eq 3 are also shown in Figure 1C. Linear regression gives $\epsilon\nu = (0.32\ \text{kJ}^{-1}\ \text{m}^2)X + 0.06$ with $r = 0.993$. The difference in the dose dependence of complex formation determined by the two methods indicates that complexes are not retained quantitatively by the filters and that the filter efficiency under these conditions is $\epsilon = 0.58$. The filter efficiency depends on the composition of the binding buffer. For example, although UvrA–UV-damaged DNA complexes are quite stable in Tris–acetate electrophoresis buffer, they bind to nitrocellulose filters inefficiently in that buffer. Many protein–nucleic acid complexes are not retained quantitatively by nitrocellulose filters (Riggs et al., 1970; Yarus & Berg, 1970), even though the protein itself is retained quantitatively (Strauss et al., 1981). The presence of more than one binding site complicates the analysis, especially if the bound proteins interact (Senear et al., 1986; Bailey, 1979). We have assumed that each bound protein has an independent and identical effect on the retention of the DNA fragment (Woodbury & von Hippel, 1983) as applied to randomly located sites.

Up to a dose of 3 kJ/m^2, the Poisson distribution accurately describes the average number of proteins bound per fragment. Even so, some departure from the Poisson distribution is ev-

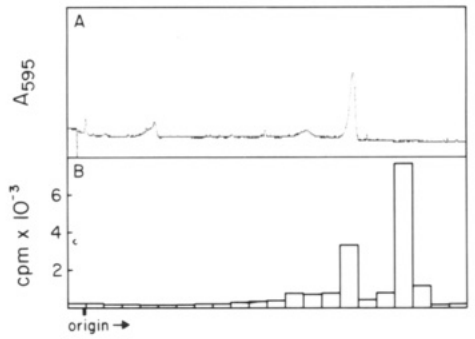

FIGURE 2: Determination of the stoichiometry of the UvrA–uvDNA complex by the gel shift assay. (A) Distribution of UvrA protein. The amount of UvrA in the nucleoprotein band is determined from the extent of dye binding to UvrA compared to BSA standards in the same gel. The dye binding affinities of UvrA and BSA were determined by using spectrophotometrically measured concentrations of the proteins. (B) Distribution of DNA. The peak of radioactivity with the greatest mobility corresponds to free 285 bp DNA. The second peak corresponds to DNA fragments with one protein bound. The amount of DNA in the nucleoprotein complex band is determined from the specific activity of the DNA. The ratio of UvrA molecules to DNA fragments in the complex is 2.24 ± 0.18, from 10 determinations, showing that UvrA binds to UV-damaged DNA as a dimer.

ident. Analysis of the distribution of complexes at higher binding densities indicates that the observed fraction of fragments with more than three proteins bound is lower than that predicted by the Poisson distribution. UvrA protects a 33 base pair region when it is bound to a damaged site (Van Houten et al., 1987), and consequently, the number of proteins that can bind to the same fragment is limited (McGhee & von Hippel, 1974).

The persistence of a linear increase in binding with the dose increasing to 3 kJ/m^2 is surprising. Most potential pyrimidine dimer sites are nearly in a photodynamic steady state with a dose of 4 kJ/m^2 (Haseltine et al., 1980). However, the 6-4 photoproducts continue to accumulate with increased dose (Lippke et al., 1981).

Although it has been suggested that UvrA induces a bend in the helix backbone when it binds to a damaged site (Pearlman et al., 1985), the width of the complex bands is essentially the same as the width of the free DNA band, and consequently, the variable location of bound UvrA does not affect the mobility of the complex in this gel system (Wu & Crothers, 1984).

All complexes observed in these assays are relatively stable since they survive a 10-s challenge by a high concentration of unlabeled, undamaged DNA. In the absence of the challenge, nonspecifically bound protein dissociates during electrophoresis, producing smearing between the bands. The stable complexes observed with unirradiated DNA may be located at fragment ends, sites of damage introduced unintentionally during purification and storage, or unusual conformations present in undamaged DNA. Substantial binding of UvrA to undamaged DNA has been observed repeatedly (Seeberg & Steinum, 1982; Yeung et al., 1986a; Van Houten et al., 1987). The fraction of unirradiated DNA which forms stable complexes is somewhat variable from one preparation to another, and increases with the age of ^{32}P-labeled DNA samples.

Determination of UvrA–DNA Complex Stoichiometry. The stoichiometry of the UvrA–DNA complex can be determined directly from gel mobility shift experiments. As illustrated in Figure 2, a separate determination of the amount of protein and DNA in a DNA–protein complex band shows that UvrA binds to UV-damaged sites as a dimer. The amount of DNA in the DNA–protein band is determined accurately from the measured radioactivity and the specific activity of the DNA.

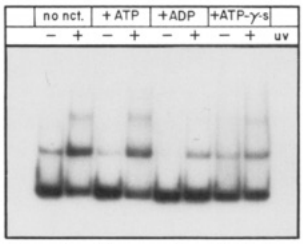

FIGURE 3: Effect of nucleotides on UvrA binding to irradiated and unirradiated DNA. Solutions containing 1 nM ^{32}P-labeled 285 bp DNA fragment (with and without 1 kJ/m^2 irradiation) and 5 nM UvrA in 0.5 × BB containing one of the following: no added nucleotide, 1 mM ATP, 1 mM ADP, or 0.1 mM ATPγS, were incubated at 37 °C for 10 min, then diluted with 0.5 volume of 0 °C TA buffer, and loaded onto a 2% agarose gel. The autoradiogram shows the image from the gel shift assay.

For stoichiometry experiments, the specific activity of the DNA fragment was determined by isotopic dilution. The amount of UvrA in the band is determined accurately by an indirect method. The binding capacities of spectrophotometrically determined amounts of UvrA and BSA for Coomassie Blue G250 were compared. In the gel shift experiments, the extent of dye binding to UvrA in the DNA–protein complex band and to BSA standards in the same gel was measured by a densitometer. From 10 determinations, the molar ratio of UvrA to the 285 base pair fragment is UvrA/DNA fragment = 2.24 ± 0.18. Thus, in the absence of ATP, UvrA binds to UV-damaged sites as a dimer.

The gel mobility shift experiment depicted in Figure 3 demonstrates that UvrA binds to DNA as a dimer in the presence of 1 mM ATP, 1 mM ADP, or 0.1 mM ATPγS and in the absence of nucleotide. All of the complexes detected in the gel assay are stable, having been challenged with a high concentration of unlabeled DNA prior to being loaded on the gel. ATP has little effect on the extent of complex formation, although fewer complexes are formed with undamaged DNA. ADP diminishes but does not abolish specific binding. In the presence of ATPγS, stable complexes are formed between UvrA and undamaged DNA or with UV-irradiated DNA. The qualitative effects of these nucleotides are in agreement with those found by using the DNase I footprint method (van Houten et al., 1988).

In agreement with earlier results from nitrocellulose filter binding experiments (Seeberg & Steinum, 1982; Yeung et al., 1986a) and DNaseI footprint experiments (Van Houten et al., 1988), UvrA binds to damaged DNA in the absence of ATP. In these reports, specific binding in the absence of ATP amounted to approximately half of the specific binding observed in the presence of ATP. Previous interpretations have emphasized the increased binding and specificity in the presence of ATP and have suggested that the function of the ATPase activity of UvrA is to promote specific binding to damaged sites. Here it is more interesting to recognize that for conditions in vitro, UvrA binds specifically to damaged sites even in the absence of ATP.

Equilibrium Properties of Specific Complexes. The binding of UvrA to UV-irradiated DNA was investigated quantitatively by nitrocellulose filter binding. An analysis of the titration of 10 pM UV-irradiated 285 bp fragment with UvrA, in the presence and absence of ATP, according to eq 7 is shown in Figure 4. The plots are fairly linear, giving an apparent specific binding constant of 4×10^{10} M^{-1} for the absence of ATP and 1.4×10^{10} M^{-1} in the presence of 1 mM ATP, assuming that UvrA is binding as a dimer and that the filter efficiency is $\epsilon = 0.6$. Both curves show trailing at higher protein concentrations, possibly due to non-UV-specific binding

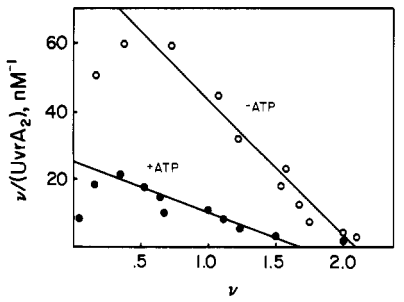

FIGURE 4: Specific binding of UvrA to UV-damaged sites. Nitrocellulose filter binding data are plotted as a Scatchard plot. The binding of from 0.02 to 1.5 nM UvrA to 0.01 nM 285 bp fragment, irradiated to 1.25 kJ/m^2, is greater in the absence of ATP than in the presence of 1 mM ATP. The apparent specific binding constants for UvrA dimers binding to UV-damaged sites is 4×10^{10} M^{-1} in the absence of ATP and 1.4×10^{10} M^{-1} in the presence of 1 mM ATP, assuming UvrA binds as a dimer and $\epsilon = 0.6$.

at fragment ends or sites of spurious damage. The deviations from linearity observed at low protein concentrations suggest dimerization of the protein or positive cooperativity. Similar results are obtained by using irradiated linearized pBR322 (data not shown). UvrA titrations at a 10-fold higher concentration of UV-irradiated fragment show weaker binding and little effect of ATP: $K_s = 3 \times 10^9$ M^{-1} in the absence of ATP and $K_s = 2 \times 10^9$ M^{-1} in the presence of ATP (data not shown).

These binding constants can be compared to the value of $K_s = 1 \times 10^8$ M^{-1} obtained by using the DNase I protection method to measure binding to a 37 bp fragment containing a uniquely located psoralen adduct (Van Houten et al., 1987). Although the most obvious differences between the two experiments are the methods and the nature of the damaged sites, the more significant differences are the UvrA concentration and the buffer composition. These factors are further discussed below.

The observed decrease in binding upon addition of ATP is unexpected. Previous studies at higher concentrations of DNA and UvrA have demonstrated an increase in binding with the addition of ATP (Seeberg & Steinum, 1982; Yeung et al., 1986a; Van Houten et al., 1988). In addition, since the K_m for ATP decreases from 200 to 30 μM upon addition of UV-irradiated DNA (Oh et al., 1989), a simple, coupled-equilibria model predicts that binding to UV-irradiated DNA should increase 5-fold in the presence of ATP. At the low concentrations of fragment, competitive effects from undamaged sites or DNA ends are minimized; thus, the increased binding in the absence of ATP cannot be explained by diminished binding to ends or undamaged sites. In addition, experiments with ^{35}S-labeled UvrA indicate that the nonspecific loss of UvrA by surface adhesion is decreased by addition of ATP (Oh et al., 1989). Thus, at low protein concentrations, the decrease in binding upon the addition of ATP reflects an inherent change in the properties of UvrA and is not the indirect result of increased nonspecific binding to undamaged DNA or fragment ends or nonspecific sticking to surfaces.

Association and Dissociation Kinetics. The decrease in binding seen at low protein concentrations in the presence of ATP is the result of a slower rate of association. The rate of dissociation of specific complexes is not significantly affected by the presence of ATP. The results of association and dissociation experiments are given in Table I. At low concentrations, the rate of association of UvrA to specific sites is slow enough to be measured by the nitrocellulose filter binding assay. The apparent rate constant for association to UV-specific sites is 4×10^7 M^{-1} s^{-1} in the presence of ATP and 3×10^8 M^{-1} s^{-1} in the absence of ATP, when the total UvrA concentration is 0.1 nM. In contrast, the rate of dissociation is little affected by ATP. The measured dissociation rate constant is 3.7×10^{-3} s^{-1} in the presence of ATP and 3.3×10^{-3} s^{-1} in the absence of ATP. The predominant effect of ATP on the specific binding of UvrA is on properties of the protein itself rather than on the DNA–protein complex.

Table I: Rate Constants for UvrA–DNA Complexes

rate constant	−ATP	1 mM ATP	DNA
k_a (M^{-1} s^{-1})	3×10^8	4×10^7	UV-irradiated 285 bp[a]
k_d (s^{-1})	3.3×10^{-3}	3.7×10^{-3}	UV-irradiated 285 bp[b]
$k_{d,ss}$ (s^{-1})	1.4×10^{-4}	7.5×10^{-5}	single-stranded ^{3}H-fd[c]

[a] Association rate constants were obtained for the reaction of 0.1 nM UvrA and 20 pM UV-irradiated 285 bp fragment in BB at 37 °C. [b] Dissociation rate constant for complexes formed from 1 nM UvrA and 0.1 nM UV-irradiated 285 bp fragment in BB at 37 °C. [c] Dissociation rate constant for complexes formed from 50 nM UvrA and 12 μM (bases) single-stranded fd DNA.

Although the dissociation kinetics of UvrA bound to UV-irradiated 285 bp fragment are well described by a single rate constant, the dissociation kinetics of complexes formed on some fragments indicate heterogeneity in complex stability. The nature of the damage or the sequence in which the damage is imbedded may influence the stability of the nucleoprotein complex. Sequence context effects have been observed in the rate of pyrimidine dimer formation (Haseltine et al., 1980), in the rate of cleavage by restriction enzymes (Alvers et al., 1984), and in the efficiency of excision by UvrABC (Myles et al., 1987; Seeberg & Fuchs, 1990).

The dissociation kinetics of UvrA bound to unirradiated DNA show two components (data not shown). Most UvrA bound to unirradiated DNA dissociates too quickly to be measured. However, a small fraction of the complexes dissociates with a rate similar to the rate of dissociation from irradiated DNA. The rapidly dissociating fraction arises from UvrA which is bound at truly undamaged sites. The slowly dissociating complexes may arise from UvrA bound at fragment ends, at unintentionally introduced sites of damage, or from UvrA molecules found in a tight binding conformation.

Dissociation of Single-Stranded DNA–UvrA Complexes. In contrast to the behavior of UvrA–uvDNA complexes, ATP does affect the rate of dissociation of UvrA bound to single-stranded DNA. Analysis of the dissociation of UvrA from ^{3}H-labeled fd DNA yields $k_{d,ss} = 1.4 \times 10^{-4}$ s^{-1} in the absence of ATP, but $k_{d,ss} = 7.5 \times 10^{-5}$ s^{-1} in the presence of 1 mM ATP. The dissociation from single-stranded DNA is 20 times slower than the dissociation from UV-damaged sites in the absence of ATP and 40 times slower in the presence of ATP. The slow rate of dissociation of UvrA from single-stranded DNA may reflect its role in DNA scanning during damage recognition (Oh & Grossman, 1987, 1989).

The rate of association of UvrA to single-stranded DNA is quite slow. The abysmally low filter efficiency for UvrA–single-stranded DNA complexes (data not shown) prevents a quantitative determination of the association rate constant. Assuming that each UvrA bound affects the filter retention independently and identically (Woodbury & von Hippel, 1983) allows the filter efficiency to be estimated as only about 1%. By combining the dissociation rate constant with the binding constant from competition experiments (see below), the association rate constant for the formation of single-stranded DNA–UvrA complexes can be estimated to be 10^4 times slower than the association rate constant for UV-damaged sites. The filter efficiency does not similarly affect the determination of dissociation rate constants because only the ratio

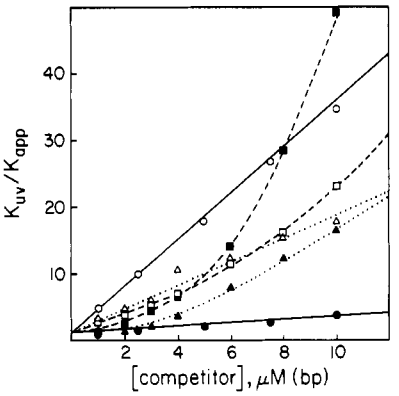

FIGURE 5: Competition between UV-damaged sites and undamaged DNA for binding of UvrA. Solutions containing 1 nM UvrA, 0.1 nM ^{32}P-labeled 285 bp fragment irradiated to 1.9 kJ/m^2, and varying amounts of unlabeled, undamaged double-stranded DNA were incubated in BB at 37 °C for 15 min. Damage-specific complex formation was measured by nitrocellulose filter binding. The data were analyzed by using eq 12 and assuming that UvrA binds as a dimer and $\epsilon = 0.6$. Open symbols indicate BB, and closed symbols indicate BB + 1 mM ATP. The competitor was supercoiled pHE6 (○, ●), *Bam*HI-cut pHE6 (□, ■), or *Alu*I-cut pHE6 (△, ▲).

of complex concentration enters the calculation.

Binding of UvrA to Undamaged DNA. The strength of UvrA binding to undamaged DNA or single-stranded DNA is determined most accurately by competition experiments. The complexes detected in these experiments are the same stable complexes formed at UV sites detected in experiments discussed previously. Questions of the efficiency of detecting unstable complexes are avoided by using competition to determine nonspecific binding constants. The results of representative competition experiments are shown in Figure 5.

Under these conditions, binding to undamaged DNA sites is only 10^3–10^4 times weaker than binding to a UV-damage site. In the presence of ATP, the binding to undamaged sites in supercoiled DNA is about 10^4-fold weaker than the binding to UV sites on a linear DNA fragment, but in the absence of ATP, the binding of undamaged sites is only about 10^3-fold weaker. The binding of UvrA to undamaged sites in linear DNA is about 10^3-fold weaker than the binding to UV sites in either the presence or absence of ATP.

Competition by linear DNA in the absence of ATP produces nearly linear competition plots, although some curvature is evident at high competitor concentrations. Competition by linear DNA in the presence of ATP produces curved competition plots. Curvature in competition plots indicates heterogeneity in the binding of the protein to the competitor. Increasing the concentration of blunt ends in the competing DNA by a factor of 20 by using *Alu*I-cut pHE6 DNA does not significantly alter the extent of competition although some differences are apparent at high concentrations. From these data, it is not possible to determine the association constant for UvrA binding to fragment ends.

DNA prepared by chloramphenicol amplification and cesium chloride/ethidium bromide centrifugation and DNA obtained from Applied Genetics, which was purified by a proprietary method not using those steps, are indistinguishable in a competition experiment (data not shown). Like the dissociation kinetics experiments, which indicate some slowly dissociating complexes are formed with all unirradiated DNA samples examined, these competition experiments indicate the formation of some tightly bound complexes with all DNA samples examined.

It is not surprising that a protein as versatile as UvrA exhibits a limited discrimination. The prevalence of tightly bound complexes in all DNA samples examined, despite care in the preparation of the DNA, suggests that these complexes are a natural feature of the interaction of UvrA and undamaged DNA and may reflect some aspect of the mechanism of damage recognition by the Uvr system.

Table II: UvrA–DNA Association Constants[a]

constant (M^{-1})	−ATP	1 mM ATP	DNA
K_{uv}	3×10^9	2×10^9	UV-irradiated 285 bp[b]
$K_{ds,sc}$	4×10^6	3×10^5	supercoiled pHE6, pBR322[c]
K_{ds}	2×10^6	2×10^6[d]	*Bam*HI-cut pHE6, pBR322[c]
K_{ss}	7×10^6	2×10^7	single-stranded fd DNA[c]

[a] Equilibrium association constants were obtained for binding of 1 nM UvrA and 0.1–0.2 nM UV-irradiated 285 bp fragment in BB at 37 °C. [b] From direct binding measurements. [c] From competition experiments. [d] Value from the low concentration region of the curved competition plot.

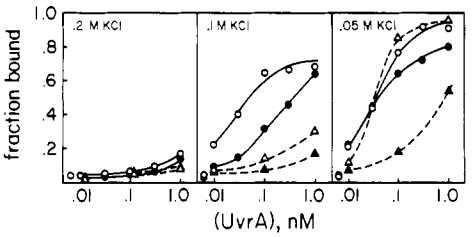

FIGURE 6: Effect of monovalent salt concentration on the specific and nonspecific binding of UvrA. Solutions containing 10 pM irradiated (1.25 kJ/m^2) or unirradiated 285 bp fragment were mixed with varying concentrations of UvrA in binding buffer of the standard composition except with 0.2, 0.1, or 0.05 M KCl, and incubated at 37 °C for 15 min. The fraction of DNA retained by the nitrocellulose filter is shown as a function of the UvrA concentration. Note that the UvrA concentration is plotted on a log scale. Open symbols denote the absence of ATP, and closed symbols denote the presence of 1 mM ATP. The retention of irradiated DNA is shown by circles and unirradiated DNA by triangles.

Equilibrium constants for the binding of UvrA at 1 nM to UV-damaged sites, double-stranded DNA, and single-stranded DNA are given in Table II. Under these conditions, ATP significantly decreases the binding of UvrA to undamaged supercoiled DNA and significantly increases the binding to single-stranded DNA, but has little effect on the binding to UV-damaged sites or to linear undamaged sites. The high affinity of UvrA for single-stranded DNA has been noted previously (Seeberg & Steinum, 1982).

Single-stranded DNA decreases the K_M of the UvrA ATPase by a factor of 30, from 150 μM in the absence of effector to 5 μM in the presence of 30 μM bases of single-stranded DNA (Oh et al., 1989). From the simple, linked-equilibria model of effector binding, UvrA could be expected to bind to single-stranded DNA more tightly in the presence of ATP by a factor of 30. Instead, the stability increases only by a factor of 3.

Effects of Solution Conditions: Salt Concentration, Protein Concentration, and Temperature. Like most DNA–protein interactions (Record et al., 1976), the binding of UvrA to DNA is strongly affected by the concentration of monovalent and divalent salts. The qualitative effects of varying the KCl concentration by a factor of 4 on the strength and specificity of binding can be seen in Figure 6. At 0.1 M KCl, UvrA binds specifically to UV-irradiated DNA in the presence or absence of ATP. At low protein concentrations, the specific binding is decreased by the presence of 1 mM ATP, as discussed previously. At the higher protein concentrations, even the unirradiated fragment is partially retained, especially in the absence of ATP. At the low salt concentration of 0.05 M

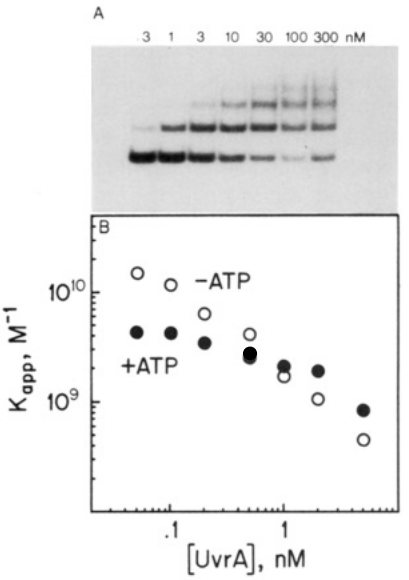

FIGURE 7: (A) Gel shift assay showing the concentration dependence of UvrA binding to uvDNA. UvrA and ^{32}P-labeled, UV-irradiated (1.5 kJ/m^2) 285 bp fragment were combined in the constant ratio of 2.5 UvrA monomers per DNA fragment at the indicated fragment concentrations. The specific activity of the DNA was varied so that the amount of radioactivity per lane was constant. The reactants were incubated in 0.5 × BB for 10 min at 37 °C and then analyzed by the gel shift assay. (B) Concentration dependence of the apparent specific binding constant. UvrA and ^{32}P-labeled, UV-irradiated linearized pBR322 were combined in BB, with or without 1 mM ATP, and incubated at 37 °C for 10 min. The UV irradiation produces an average of 10 damaged sites per plasmid. Complex formation was measured by nitrocellulose filter binding. Data were analyzed assuming UvrA binds as a dimer and $\epsilon = 0.6$.

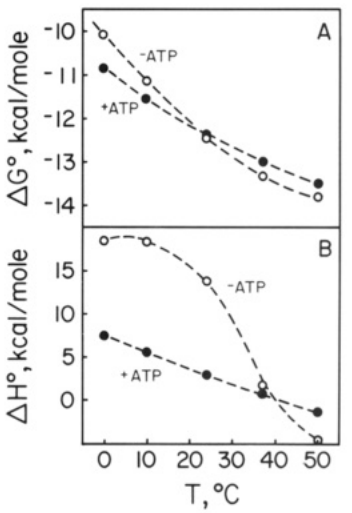

FIGURE 8: Temperature dependence of UvrA binding to UV-damaged DNA. Reactants were combined in BB equilibrated to various temperatures to give 0.5 nM UvrA and 0.1 nM ^{32}P-labeled UV-irradiated (1 kJ/m^2) 285 bp fragment, in the absence or presence of 1 mM ATP. Reaction mixtures were incubated for 15 min at the various temperatures and analyzed by nitrocellulose filter binding. The specific binding constant is calculated by assuming UvrA binds as a dimer and that $\epsilon = 1$. (A) The apparent standard free energy of binding is calculated from $\Delta G° = -RT \ln K_{uv}$. (B) The apparent standard enthalpy of binding is determined from the slope of the apparent specific binding constant plotted against inverse temperature.

KCl, binding is much tighter, but the discrimination between damaged and undamaged DNA is lost in the absence of ATP. Even in the presence of ATP, UvrA binds extensively to undamaged DNA. Conversely, at the high salt concentration of 0.2 M KCl, the binding of UvrA to irradiated DNA is very weak both with and without ATP. Variation in the concentration of monovalent salt thus affects both the strength of UvrA binding to DNA and the specificity for damaged DNA.

As mentioned above, the effect of ATP on the damage-specific binding of UvrA depends on the protein concentration. In Figure 7A, a gel mobility shift experiment shows that the increase in binding constant at low concentrations is not due to the formation of a different type of complex. The complex formed at the lowest concentration has the same mobility as the first complex band formed at high concentration, that is, a UvrA dimer bound to the fragment. In Figure 7B, the results of nitrocellulose filter binding experiments show a strong concentration dependence of the apparent binding constant. At UvrA concentrations around 10^{-10} M, the binding of UvrA is actually tighter in the absence of ATP than in its presence. The apparent association constant for the absence of ATP is more strongly dependent on concentration, and as a consequence, the curves cross at around 1 nM protein. At higher protein concentrations, the apparent association constant is greater in the presence of ATP than in its absence.

The concentration dependence displayed in Figure 7 is the result of several processes. At low concentrations, DNA binding in the presence of ATP is reduced because ATP hydrolysis results in dissociation of the UvrA dimers, whereas in the absence of ATP, dimers are stable at low concentrations. In the absence of additional mechanisms, this would lead to a concentration-independent equilibrium constant for the absence of ATP and an equilibrium constant that increased with concentration in the presence of ATP. An additional element that can produce a downward trend in the apparent binding constant is the aggregation of UvrA.

The temperature dependence of the thermodynamic quantities derived from the apparent specific binding constant is displayed in Figure 8. The free energy of binding decreases with increasing temperature both in the presence and in the absence of ATP. The two curves cross around 25 °C. At lower temperatures, ATP increases binding, but at higher temperatures, the converse is true. The entropy of binding is obtained from the dependence of $\Delta G°$ on temperature. The overall binding reaction is driven by the increase in entropy. The entropy of binding decreases with increasing temperature.

The standard enthalpy is determined from the slope of the apparent specific binding constant plotted against the inverse temperature. In both the presence and absence of ATP, the enthalpy of the reaction is temperature dependent. At 0 °C, the enthalpy contributes unfavorably by 18 kcal/mol in the absence of ATP and by 8 kcal/mol in the presence of ATP. This unfavorable enthalpy decreases to zero near physiological temperatures for both the presence and absence of ATP.

In addition, although changes in the heat capacity can be determined less accurately than changes in the standard enthalpy or entropy, the temperature dependence of $\Delta H°$ suggests that in both the presence and absence of ATP, the binding of UvrA to UV-damaged sites results in a decrease in the heat capacity. In the absence of ATP, ΔC_p is -8×10^2 cal/(mol·K), whereas in the presence of ATP, this is reduced to -2×10^2 cal/(mol·K).

The change in heat capacity in the absence of ATP is very large. Large negative heat capacity changes have been observed to accompany the folding of small globular proteins (Creighton, 1983) and in general result from a reaction in which the products are more compact or have fewer solvent and small ion molecules bound compared to the reactants.

The thermodynamic data suggest that the binding of UvrA to UV-damaged sites is a hydrophobically driven process. The general characteristics of hydrophobically driven reactions are (1) the enthalpy is unfavorable to indifferent, (2) the entropy

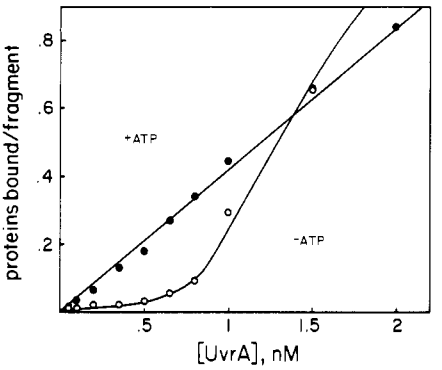

FIGURE 9: Binding of UvrA to uvDNA at 0 °C. UvrA protein, from 0.05 to 2 nM, was mixed with 0.1 nM ^{32}P-labeled 285 bp fragment, irradiated to 1 kJ/m^2, in BB with or without 1 mM ATP, and incubated for 15 min at 0 °C. Nucleoprotein complex formation was measured by nitrocellulose filter binding. The data were analyzed assuming UvrA is a dimer and that $\epsilon = 1$.

is favorable, and (3) the heat capacity is negative. The simplest interpretation of these data, in accord with data presented earlier, is that dimerization of UvrA, a process which is likely to be hydrophobically driven, is required for specific binding and that this process dominates the temperature dependence of the reaction. The dimerization of UvrA can be expected to involve the burying of a hydrophobic face of the monomer.

A prediction of this interpretation is that at low temperatures, the binding of UvrA to UV sites will reflect the requirement for dimerization of the protein. This, in fact, is observed. In Figure 9, the binding of UvrA to UV-damaged DNA at 0 °C is sigmoidal in the absence of ATP, but linear in the presence of ATP. At low temperatures and low protein concentrations, UvrA exists largely as a monomer in the absence of ATP, and consequently little specific binding to UV-damaged sites is observed. At higher concentrations, UvrA dimerizes, and specific binding increases. ATP promotes the dimerization of UvrA at low protein concentrations and increases specific binding. Sedimentation velocity experiments at 4 °C show that in the absence of ATP, UvrA has the mobility of a monomer, but in the presence of ATP, UvrA has a mobility intermediate between that of a monomer and a dimer (Oh et al., 1989).

Discussion

Dimerization of UvrA. The effects of ATP on the state of aggregation of UvrA and on the specific and nonspecific binding of UvrA to DNA have been obscured by the number of interrelated binding reactions and the sensitivity of those reactions to conditions of UvrA concentration, DNA concentration, temperature, and salt concentration. The results presented above combined with previously published results support the following model:

UvrA dimerizes in solution. Dimerization is promoted by physiological temperatures, high (greater than 1 nM) concentrations of UvrA, or the presence of ATP. However, some UvrA exists in the dimeric form even at low temperatures, lower concentrations of UvrA, or in the absence of ATP. The hydrolysis of ATP returns UvrA dimers to the monomer form. UvrA dimers bind specifically to sites of UV damage. The presence of ATP has little effect on the properties of the UvrA–uvDNA complex. The binding of UvrA to supercoiled, undamaged DNA is decreased by the presence of ATP whereas its binding to single-stranded DNA is increased. Both the strength of damage-specific binding and the discrimination between damaged and undamaged sites are affected by salt concentration.

In solution, the state of aggregation of UvrA is influenced by the presence of nucleotides and the protein concentration (Oh et al., 1989; Orren & Sancar, 1989). Sedimentation velocity experiments and equilibrium centrifugation experiments conducted at 4 °C showed that in the absence of nucleotides UvrA is a monomer while in the presence of ADP or ATPγS UvrA is a dimer and in the presence of ATP UvrA has an intermediate mobility, suggesting an equilibrium between monomers and dimers (Oh et al., 1989). Similarly, gel filtration and sedimentation velocity experiments were combined to show that, in the presence of ATP, UvrA monomers and dimers are in equilibrium and that the apparent sedimentation coefficient of UvrA is concentration dependent (Orren & Sancar, 1989). The temperature dependence of the binding of UvrA to UV-damaged sites shown in Figures 8 and 9 suggests that dimerization is a hydrophobic process that is promoted by physiological temperatures. From the gel mobility shift data presented here, only dimeric UvrA specifically bound to UV-damaged sites is detected. Thus, although nucleotides promote the dimerization of UvrA, neither nucleotide binding nor hydrolysis is required for dimer formation.

The surprising decrease in UvrA binding to UV-damaged sites in the presence of ATP shown in Figure 4 and the effect of ATP on the rate of association of UvrA to uvDNA given in Table I illuminate the opposing effects of ATP on the properties of UvrA. These observations can be explained by ATP hydrolysis causing dissociation of the UvrA dimer, thereby reducing the levels of reactive dimeric species. At higher protein concentrations, mass action allows the dimer to re-form quickly, but at low protein concentrations, the monomers remain dissociated. This explanation is supported by sedimentation velocity experiments in which the mobility of UvrA is intermediate between that of a monomer and a dimer (Oh et al., 1989).

Although ATP has little direct effect on the properties of the UvrA–uvDNA complex, the effects of ATP can be manifested indirectly through decreasing the level of UvrA dimer at low protein concentration or through decreasing the extent of competition from nonspecific sites. The increased levels of UvrA binding in the presence of ATP at high UvrA concentrations and high undamaged DNA concentrations measured previously (Seeberg & Steinum, 1982; Yeung et al., 1986a) are partly due to decreased levels of competition from undamaged sites. However, the illusion that the interaction of ATP, UvrA, and DNA can be described by simple coupled-equilibrium models is removed by the repeated failure of the simple model to describe the data. The recent demonstration that the activities of the two ATP binding sites in UvrA are not independent (Thiagalingam & Grossman, 1991) provides further support for the complexity of the interactions.

The apparent association constant for UvrA binding to a DNA fragment containing a psoralen adduct was determined by the DNaseI footprint method to be about 1×10^8 M^{-1} for a fragment concentration of 2.2 nM and for protein concentrations in the range of 10 nM (Van Houten et al., 1987). This is a factor of 30 lower than the low concentration value obtained from the data shown in Figure 3 for binding of UvrA to UV-irradiated 285 bp DNA in the presence of ATP, and approximately a factor of 10 lower than the high concentration values indicated in Figure 7. This difference may be attributed to the difference in the nature of the damage or the method of measuring binding, but a more plausible explanation is the difference in the buffer conditions. In the DNaseI footprint experiments, the concentration of KCl was 50 mM. In low-salt buffers, the binding to undamaged DNA is increased, and the

apparent specific binding constant may be decreased by competition from undamaged sites.

It is noteworthy that high salt concentrations (300 mM KCl) were used by Orren and Sancar (1989) in analyzing the protein composition of UvrAB–DNA complexes. Finding very little UvrA bound to the DNA after passing UvrAB–DNA complexes through a column with a high-salt buffer, they postulated that the role of UvrA was to transport UvrB to a damaged site, at which point it was released. The results shown in Figure 6 show that the binding of UvrA to DNA is very weak in solutions containing high concentrations of salt. Similarly, although UvrB has not been shown to bind to DNA directly, the high salt concentration may increase the strength of a hydrophobic interaction of UvrB with DNA, leading to a stable complex. While the manipulation of conditions provides insight, the exquisite sensitivity of DNA–protein interactions to solution conditions makes generalization from one condition to another perilous.

Specificity of UvrA for Damaged Sites. The specificity of UvrA found for UV-damaged sites is much lower than the specificity of 10^7 for DNA binding proteins such as the *lac* repressor (von Hippel et al. 1974) or specificities of 10^5 found for DNA modifying enzymes such as the restriction endonuclease *Eco*RI (Terry et al., 1983) or the DNA repair enzyme photolyase (Husain & Sancar, 1987). The low equilibrium specificity of this system is an expected consequence of its versatility in recognizing diverse types of damage and supports the inclusion of dynamic recognition in the mechanism (Oh & Grossman, 1989; Pu et al., 1989).

In the presence of DNA polymerase I and helicase II, which are required for the Uvr system to function enzymatically (Caron et al., 1985; Husain et al., 1985b), the in vitro turnover rate is 0.08 min^{-1} (ABC complex)$^{-1}$ (Husain et al., 1985a). This number is in good agreement with the reported in vivo rate of 0.12–0.3 min^{-1} (ABC complex)$^{-1}$ (Tang & Patrick, 1977). Considering the differences in conditions in vitro and in vivo, the close agreement is itself informative. One of the greatest differences is that the concentration of undamaged base pairs is quite high in vivo.

For an enzyme, E, acting on a specific substrate, S, in the presence of competing nonsubstrate, N, the simplest model describing the rate is given by

$$E + N \underset{}{\overset{K_N}{\rightleftharpoons}} EN \qquad (13)$$

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Assuming that both specific sites and nonspecific sites are in excess over enzyme, then the steady-state turnover rate is given by

$$v_{obs} = V/[E]_0 = k_2k_1[S]/[(k_{-1} + k_2)(1 + K_N[N]) + k_1[S]] \quad (14)$$

The mechanism of the UvrABC endonuclease is obviously more complex, but the qualitative effects of competition from undamaged DNA will be similar. For the present discussion, we assume that the equilibrium constant for nonspecific binding is $K_N = 10^6$ M^{-1} and that the specific complex association and dissociation rate constants are $k_1 = 4 \times 10^7$ M^{-1} s^{-1} and $k_{-1} = 4 \times 10^{-3}$ s^{-1}. This corresponds to a static discrimination factor of 10^4. In addition, we assume that the catalytic rate constant $k_2 = 1.4 \times 10^{-3}$ s^{-1}.

The in vitro turnover rate of 0.08 min^{-1} (UvrABC complex)$^{-1}$ was measured by both an incision and an excision assay (Husain et al., 1985a). In the incision assay, reaction mixtures contained 2.2 nM (molecules) pBR322 with 10 photoproducts per plasmid, 2.9 nM UvrA, and an excess of UvrB, UvrC, DNA polymerase I, and helicase II. The competition from nonspecific sites is expressed as $K_N[N] = 9.5$, and the turnover rate, calculated from eq 14, is $v_{obs} = 1.32 \times 10^{-3}$ s^{-1}. In the excision assay, the concentration of pBR322 was 12.6 nM, the concentration of *cis*-[^{3}H]platin adducts was 177 nM, and the concentration of UvrA, again limiting, was 32 nM. The competition from nonspecific sites is $K_N[N] = 50.4$. The calculated turnover rate is $v_{obs} = 1.35 \times 10^{-3}$ s^{-1}. Although competition from nonspecific sites is greater in the excision assay by a factor of 5, the turnover rate is hardly affected by it.

The situation in vivo is distinctly different. Tang and Patrick (1977) measured the rate of removal of pyrimidine dimers as 100 h^{-1} $cell^{-1}$ under conditions of liquid holding recovery. Assuming the number of ABC complexes per cell is around 10 (Sancar & Sancar, 1988), the turnover rate is 0.12–0.24 min^{-1} (UvrABC)$^{-1}$, or $v_{obs} = (2–4) \times 10^{-3}$ s^{-1}. Assume, for the moment, that the rate constants and equilibrium constants measured in vitro also apply in vivo. In an *Escherichia coli* cell, with a volume of 10^{-15} L and with 2.5 genome equiv per cell (von Hippel et al., 1974), the concentration of DNA base pairs is [N] = 17 mM, and the competition product is $K_N[N] = 1.7 \times 10^4$. With 100 damaged sites and 10 UvrABC complexes per cell, the turnover rate calculated from eq 14 is $v_{obs} = 9.4 \times 10^{-5}$ s^{-1}, lower than the in vitro rate by a factor of 15. The agreement of the in vivo and in vitro rates indicates that competition from undamaged sites has an unexpectedly small effect on the rate of incision in vivo. The small effect of competition from undamaged sites could be the result of differences in conditions in vivo and in vitro. An alternative explanation, which we favor, is that the lack of an effect may indicate the inclusion of a scanning step in the mechanism of damage recognition (Oh & Grossman, 1989; Dowd & Lloyd, 1989).

Determination of Protein–Nucleic Acid Complex Stoichiometry. The gel mobility shift assay provides a direct means for determining the stoichiometry of protein–nucleic acid complexes (Garner & Revzin, 1982), providing the amounts of protein and DNA can be determined accurately. Most methods for determining the amount of protein exhibit variability in response from one protein to another. The binding of Coomassie Blue G250 (Bradford, 1976) is quite sensitive, but can vary by up to a factor of 4 in response (Bio-Rad Laboratories, 1987). Determinations of the stoichiometry of protein complexes based solely on the extent of dye binding and the molecular weight are unreliable (Orren & Sancar, 1989). Similarly, methods based on binding studies or reaction rates cannot be used to determine stoichiometries unless the absolute protein concentrations are known. Ultraviolet absorbance can be used to determine the amount of protein to within 20%, providing the sequence of the protein is known (Butler et al., 1977), but requires relatively large amounts of protein. The small variation results from the effects of the local environment on the absorbance of chromophoric amino acids and can be removed by measuring the absorbance under denaturing conditions (Edelhoch, 1967).

When these two methods are combined, ultraviolet absorbance is used to calibrate the dye binding response of UvrA in situ. The distribution of dye in the gel is then an accurate and precise measure of the amount of UvrA in a band. Control experiments establish that buffer components and DNA do not interfere with the binding of Coomassie Blue G250 under these conditions.

Once calibrated for a particular protein, the gel-based Bradford assay for determining the stoichiometry of DNA–protein complexes is about as sensitive as in vivo labeling with ^{35}S and is more convenient. Both methods require an absolute determination of protein concentration. Since radioactively labeled proteins are generally not prepared in sufficient quantity to measure the ultraviolet absorbance directly, the specific activity of the protein is often determined indirectly through the extent of dye binding compared to known amounts of unlabeled protein. The relative sensitivity of the two methods depends on the number of sulfur atoms per protein molecule and the dye binding capacity of the protein. For example, in vivo labeling of UvrA resulted in a specific activity of 7×10^4 cpm/μg (Oh et al., 1989). The amount of UvrA that can be detected easily (3500 cpm) corresponds to 50 ng of protein. The gel-based dye binding assay is linear to at least 200 ng of UvrA, and a good signal is obtained from 100 ng of UvrA. Thus, the dye binding method ranges from being as sensitive to 5 times less sensitive than in vivo labeling.

ACKNOWLEDGMENTS

The efforts of E. Y. Oh in purifying the UvrA, S. Thiagalingam in removing ATP from the UvrA, and Lark Claassen in verifying the dye binding capacity of UvrA are gratefully acknowledged.

Registry No. ATP, 56-65-5; UvrABC endonuclease, 81611-73-6; KCl, 7447-40-7.

REFERENCES

Alvers, J., Pingoud, A., Haupt, A., Langowski, J., Peters, F., Maass, G., & Wolff, C. (1984) *Eur. J. Biochem. 140*, 83–92.

Bailey, J. M. (1979) *Anal. Biochem. 93*, 204–206.

Bio-Rad Laboratories (1987) *Bio-Rad Protein Assay*, Bio-Rad Technical Bulletin 1069 EG Ed., Bio-Rad Laboratories, Richmond, CA.

Bogosian, G., & Kane, J. F. (1987) *Nucleic Acids Res. 15*, 7185.

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.

Brenowitz, M., Senear, D. F., Shea, M., & Ackers, G. K. (1986) *Methods Enzymol. 130*, 132–181.

Butler, A. P., Revzin, A., & von Hippel, P. H. (1977) *Biochemistry 16*, 4757–4768.

Caron, P. R., & Grossman, L. (1988) *Nucleic Acids Res. 16*, 10891–10902.

Caron, P. R., Kushner, S. R., & Grossman, L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4925–4929.

Creighton, T. E. 1983 *Proteins*, W. H. Freeman, New York.

Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.

Dowd, D. R., & Lloyd, R. S. (1989) *J. Mol. Biol. 208*, 701–707.

Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.

Fried, M. G., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505–6525.

Friedberg, E. C. (1985) *DNA Repair*, W. H. Freeman, New York.

Garner, M. M., & Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047–3060.

Garner, M. M., & Revzin, A. (1982) *Biochemistry 21*, 6032–6036.

Grossman, L., & Yeung, A. T. (1990) *Mutat. Res. 236*, 213–221.

Grossman, L., Caron, P. R., Mazur, S. J., & Oh, E. Y. (1988) *FASEB J. 2*, 2696–2701.

Hardies, S., & Wells, R. D. (1979) *Gene 7*, 1–14.

Haseltine, W. A., Gordon, L. K., Lindan, C. P., Grafstrom, R. H., Shaper, N. L., & Grossman, L. (1980) *Nature 285*, 634–641.

Hopfield, J. J. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4135–4139.

Husain, I., & Sancar, A. (1987) *Nucleic Acids Res. 15*, 1109–1120.

Husain, I., Chaney, S. G., & Sancar, A. (1985a) *J. Bacteriol. 163*, 817–823.

Husain, I., Van Houten, B., Thomas, D. C., Abdel-Monem, M., & Sancar, A. (1985b) *Proc. Natl. Acad. Sci. U.S.A 82*, 6774–6778.

Husain, I., Van Houten, B., Thomas, D. C., & Sancar, A. (1986) *J. Biol. Chem. 261*, 4895–4901.

Lippke, J. A., Gordon, L. K., Brash, D. E., & Haseltine, W. A. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3388.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469–489; *103*, 679 (correction).

Messing, J. C. (1983) *Methods Enzymol. 101*, 20–78.

Milman, G. (1987) *Methods Enzymol. 153*, 482–491.

Myles, G. M., Van Houten, B., & Sancar, A. (1987) *Nucleic Acids Res. 15*, 1227–1243.

Navaratnam, S., Myles, G. M., Strange, R. W., & Sancar, A. (1989) *J. Biol. Chem. 264*, 16067–16071.

Ninio, J. (1975) *Biochimie 57*, 587–595.

Oh, E. Y., & Grossman, L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3638–3642.

Oh, E. Y., & Grossman, L. (1989) *J. Biol. Chem. 264*, 1336–1343.

Oh, E. Y., Claassen, L., Thiagalingam, S., Mazur, S. J., & Grossman, L. (1989) *Nucleic Acids Res. 17*, 4145–4159.

Orren, D. K., & Sancar, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5237–5241.

Pearlman, D. A., Holbrook, S. R., Pirkle, D. H., & Kim, S. H. (1985) *Science 227*, 1304–1308.

Pu, W. T., Kahn, R., Munn, M. M., & Rupp, W. D. (1989) *J. Biol. Chem. 264*, 20697–20704.

Pulleyblank, D. E., Shure, M., & Vinograd, J. (1977) *Nucleic Acids Res. 4*, 1409–1418.

Record, M. T., Jr., Lohman, T. M., & de Haseth, P. (1976) *J. Mol. Biol. 107*, 145–158.

Riggs, A. D., Suzuki, H., & Bourgeois, S. (1970) *J. Mol. Biol. 48*, 67–83.

Sancar, A., & Sancar, G. B. (1988) *Annu. Rev. Biochem. 57*, 29–67.

Scatchard, G. (1949) *Ann. N.Y. Acad. Sci. 51*, 660.

Seeberg, E., & Stenum, A.-L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 988–992.

Seeberg, E., & Fuchs, R. R. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 191–194.

Seeley, T. W., & Grossman, L. (1989) *Proc. Natl. Acad. Scid. U.S.A. 86*, 6577–6581.

Seeley, T. W., & Grossman, L. (1990) *J. Biol. Chem. 265*, 7158–7165.

Selby, C. P., & Sancar, A. (1990) *Mutat. Res. 236*, 203–211.

Senear, D. F., Brenowitz, M., Shea, M., & Ackers, G. K. (1986) *Biochemistry 25*, 7344–7354.

Strauss, H. S., Boston, R. S., Record, M. T., Jr., & Burgess, R. R. (1981) *Gene 13*, 75–87.

Tanaka, T., & Weisblum, B. (1975) *J. Bacteriol. 121*, 354–362.

Tang, M. S., & Patrick, M. H. (1977) *Photochem. Photobiol. 26*, 247–255.
Terry, B. J., Jack, W. E., Rubin, R. A., & Modrich, P. (1983) *J. Biol. Chem. 258*, 9820–9825.
Thiagalingam, S., & Grossman, L. (1991) *J. Biol. Chem.* (in press).
Van Houten, B. (1990) *Microbiol. Rev. 54*, 18–51.
Van Houten, B., Gamper, H., Sancar, A., & Hearst, J. E. (1987) *J. Biol. Chem. 262*, 13180–13187.
Van Houten, B., Gamper, H., Hearst, J. E., & Sancar, A. (1988) *J. Biol. Chem. 263*, 16533–16560.
von Hippel, P. H., Revzin, A., Gross, C. A., & Wang, A. C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4808–4812.
Weast, R. C., Ed. (1971) *Handbook of Chemistry and Physics*, 51st ed., Chemical Rubber Publishing Company, Cleveland, OH.
Woodbury, C. P., Jr., & von Hippel, P. H. (1983) *Biochemistry 22*, 4730–4737.
Wu, H.-M., & Crothers, D. M. (1984) *Nature 308*, 509–513.
Yarus, M., & Berg, P. (1970) *Anal. Biochem. 35*, 450–465.
Yeung, A. T., Mattes, W. B., Oh, E. Y., & Grossman, L. (1983) *Proc. Natl. Acad. Sci. U.S.A 80*, 6157–6161.
Yeung, A. T., Mattes, W. B., & Grossman, L. (1986a) *Nucleic Acids Res. 14*, 2567–2582.
Yeung, A. T., Mattes, W. B., Oh, E. Y., Yoakum, G. H., & Grossman, L. (1986b) *Nucleic Acids Res. 14*, 8535–8556.

Crystal Structure Analysis of the B-DNA Dodecamer CGTGAATTCACG[†,‡]

Teresa A. Larsen, Mary L. Kopka, and Richard E. Dickerson*
Molecular Biology Institute, Department of Chemistry and Biochemistry, and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, California 90024-1570
Received September 11, 1990; Revised Manuscript Received January 3, 1991

ABSTRACT: The crystal structure of the DNA dodecamer C-G-T-G-A-A-T-T-C-A-C-G has been determined at a resolution of 2.5 Å, with a final *R* factor of 15.8% for 1475 nonzero reflections measured at 0 °C. The structure is isomorphous with that of the Drew dodecamer, with the space group $P2_12_12_1$ and cell dimensions of $a = 24.94$ Å, $b = 40.78$ Å, and $c = 66.13$ Å. The asymmetric unit contains all 12 base pairs of the B-DNA double helix and 36 water molecules. The structure of C-G-T-G-A-A-T-T-C-A-C-G is very similar to that of C-G-C-G-A-A-T-T-C-G-C-G, with no major alterations in helix parameters. Water peaks in the refined structure appear to represent a selection of peaks that were observed in the Drew dodecamer. The minor-groove spine of hydration at 2.5 Å is fragmentary, but as Narendra et al. (1991) [*Biochemistry* (following paper in this issue)] have observed, lowering the temperature leads to a more complete representation of the spine.

This paper reports the crystal structure analysis of the B-DNA dodecamer with the sequence C-G-T-G-A-A-T-T-C-A-C-G. The sequence was chosen as a variant of the "Drew" dodecamer, C-G-C-G-A-A-T-T-C-G-C-G, as part of an investigation of lexitropsins, synthetic analogues of netropsin that are designed to "read" a particular base sequence along the floor of the minor groove of B-DNA. Netropsin (Figure 1a) is a planar drug molecule that binds within the minor groove and requires a binding region of four successive A·T base pairs (Zimmer, 1975; Kopka et al., 1985a,b). It can be regarded as a polymer whose monomer is a planar amide coupled to a five-membered pyrrole ring. Longer analogues of the natural netropsin molecule in Figure 1a with eight or nine monomer units have been synthesized by Lown, Dervan, and others (Lown & Krowicki 1985; Wade & Dervan 1987). As expected, the longer analogues demand longer stretches of A·T base pairs as binding sites. The A·T specificity arises from a combination of steric hindrance from the N2 amines of the guanines along the floor of the minor groove, the intrinsic narrowness of A·T zones of the minor groove in B-DNA (Yoon et al., 1988), and the deeper electrostatic potential in A·T regions of the minor groove than in G·C regions (Pullman, 1983).

Kopka et al. (1985a,b) and Lown et al. (1986a,b) proposed independently that netropsin be modified to make it G·C specific at a given position by replacing a pyrrole by furan or imidazole (Figure 1b). This replacement of a ring C-H by O or N would simultaneously make room for the N2 amine group of guanine and provide a new hydrogen-bond acceptor that could bond to that amine. Hence in a long synthetic -(amide-ring)$_n$- analogue of netropsin, each pyrrole-ring site would demand an A·T base pair, and each imidazole-ring site might require a G·C pair. In this manner one might be able to engineer a drug analogue capable of "reading" a specific sequence of the order of 10 base pairs in length. Such a sequence, chosen because it was known to be present in a tumor or foreign cell, would occur at random in the host genome only once in $2^{10} = 1024$ times, assuming that a lexitropsin site could distinguish between an A·T and a G·C pair but could not detect the orientation of the pair (Dickerson et al., 1986). By directing or vectoring the groove-binding drug specifically to the aberrant cell, one might hope for a substantial reduction in toxicity of the drug and of unwanted chemotherapeutic side effects.

Among the first lexitropsins to be synthesized by Lown were analogues of netropsin with one or the other of the pyrroles replace by imidazole. Such a molecule should recognize a four base pair site in which one of the two inner base pairs was G·C

[†] This research was supported by NSF Grant DNB85-01682 and NIH Program Project Grant GM-31299.
* To whom correspondence should be addressed, at the Molecular Biology Institute.
[‡] Crystallographic coordinates and structure factors have been submitted to the Brookhaven Protein Data Bank.

0006-2960/91/0430-4443$02.50/0