Use of a Combined Enzymatic Digestion/ESI Mass Spectrometry Assay To Study the Effect of TATA-Binding Protein on Photoproduct Formation in a TATA Box†

Yinsheng Wang, Michael L. Gross, and John-Stephen Taylor*

Department of Chemistry, Washington University, One Brookings Drive, St. Louis, Missouri 63130

Received June 4, 2001; Revised Manuscript Received August 6, 2001

ABSTRACT: Recently, it was reported that TATA-binding protein (TBP) enhances (6-4) photoproduct formation in a TATA box under UVC irradiation [Aboussekhra and Thoma (1999) *EMBO J. 18*, 433−443]. The conclusions of that study were based on an indirect enzymatic assay that was not specific for (6-4) photoproducts. Herein we report the use of a recently developed coupled enzymatic digestion/mass spectrometry assay [Wang et al. (1999) *Chem. Res. Toxicol. 12*, 1077−1082] to identify unambiguously and quantify the photoproducts formed in a TATA box-containing dodecamer duplex sequence in the presence or absence of TBP binding. Exposure of the adenovirus major late promoter TATA box to a high dose of UVC irradiation in the absence of the C-terminal domain of yeast TBP leads to predominant formation of the cis-syn dimer within the T_4 tract, whereas exposure in the presence of TBP leads to almost exclusive formation of the (6-4) photoproduct. In contrast, the (6-4) product is not detected at high doses of UVB irradiation in the absence of TBP but is detected in the presence of TBP, although the cis-syn product predominates. When the products of UVB irradiation were subsequently exposed to a high dose of UVC irradiation in the presence of TBP, the (6-4) photoproduct again becomes nearly the exclusive photoproduct, indicating that the cis-syn dimer is being reversed to TT by UVC light. Both cis-syn and (6-4) photoproducts are formed in approximately equal amounts upon irradiation with small doses of UVC in the presence of TBP, but the fraction of (6-4) photoproduct increases with dose. Through the use of a TATA box containing a site-specifically deuterated thymine, it was found that (6-4) photoproducts formed most selectively at the second and third positions of the T_4 tract upon either UVB or UVC irradiation in the presence of TBP. By using the same substrate, it was found that UVC-induced TA* formation was inhibited by TBP binding and that TA* formation was greatest at the 5′ end of the TATA sequence.

Skin cancers are the most prevalent type of cancer, the majority of which are thought to arise from mutations caused by DNA synthesis past cis-syn and (6-4)[1] photoproducts induced by the UVB component of sunlight (Figure 1) (*1*−*3*). Determining the origin of UV-induced mutations is made complex by the fact that DNA in vivo is intimately associated with proteins that can affect photoproduct formation and repair. The ability of proteins to affect DNA photoproduct formation has been known for some time and is most striking in bacterial spores that are very resistant to UV light (*4*). In *Bacillus subtilis* spores, α/β-type small, acid-soluble protein (SASP) causes the DNA to adopt an A conformation which inhibits formation of the cis-syn and (6-4) products and enhances formation of the spore photoproduct (*5*). The ability of proteins to modulate photoproduct formation ultimately became the basis of a technique known as photofootprinting, which can be used to monitor protein binding to DNA both in vitro and in vivo (*6*−*9*).

Research over the course of the past decade has shown that a variety of proteins can modulate DNA photoproduct formation. For example, it was found that nucleosomes lead to preferential formation of cis-syn dimers in heterogeneous sequences at sites where the DNA backbone is furthest from the histone surface (*10*). On the other hand, a different pattern of cis-syn dimer formation was found in long T tracks in nucleosomes (*11*). Whereas cis-syn dimers form within the nucleosomes, (6-4) photoproducts form preferentially in the linker regions between nucleosome core particles (*12*−*14*). In T7 phage, cis-syn and (6-4) products are strongly enhanced by the presence phage proteins, and the ratio of cis-syn to (6-4) product formation increases with increasing wavelength (*15*). Both enhancement and inhibition of photoproducts relative to naked DNA were also observed in vivo at sites of promoters of genes such as c-jun, c-fos, and PCNA (*9*). More recently, TFIIIA binding was found to modulate formation of photoproducts in the *Xenopus borealis* 5S rRNA gene (*16*).

We first became interested in exploring the effect of TATA-binding protein (TBP) on the photochemistry of TATA sites when we were studying the structure and mutagenic properties of TA* photoproducts (*17*−*19*). We were especially curious to know whether TBP would enhance or inhibit TA* formation and perhaps cause the formation of new types of photoproducts between T and A. TBP is a

† This research was supported by NIH Grant CA40463 and Alchemy International. The National Center for Research Resources of the NIH supports the mass spectrometry research resource (Grant P41RR00954).

* To whom correspondence should be addressed. Phone: (314) 935-6721. Fax: (314) 935-4481. E-mail: taylor@wuchem.wustl.edu.

[1] Abbreviations: (6-4) photoproduct, pyrimidine−pyrimidone photoproduct; AdMLP, adenovirus major late promoter; Da, Dalton; ESI, electrospray ionization; MS, mass spectrometry; TBP, TATA-binding protein; UVA, 320−400 nm; UVB, 280−320 nm; UVC, 200−280 nm.

10.1021/bi0111552 CCC: $20.00
Published on Web 09/08/2001

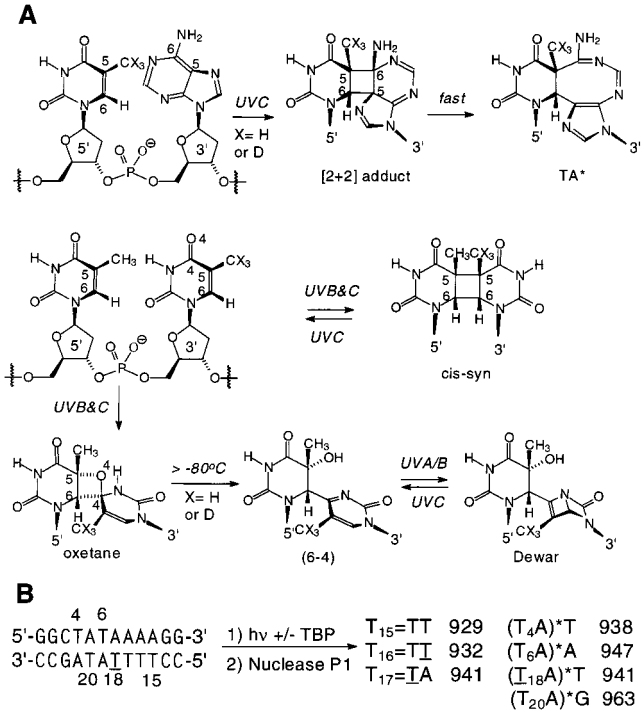

FIGURE 1: Photochemistry of TT and TA sites in DNA. Panel A: Photochemistry of TA and TT sites showing the pathways leading to the TA* photoproduct and the cis-syn cyclobutane dimer, (6-4) photoproduct, and its Dewar valence isomer. UVA = 320−400 nm, UVB = 280−320 nm, and UVC = 200−260 nm. Panel B: Expected photoproduct-containing trinucleotides and their masses that result from nuclease P1 degradation of the TATA box sequence shown with the site of a trideuterio-labeled thymidine underlined.

central protein in transcription as it is required by all three eukaryotic RNA polymerases to initiate the transcription of ribosomal, messenger, small nuclear, and transfer RNA at TATA sequences (*20*). TBP associates with about 10 other transcription-associated factors (TAFs) to form the primary promoter recognition factor TFIID (*21*, *22*).

X-ray crystallography studies of TBP/TATA box complexes have revealed that the saddle-shaped TBP core wraps around DNA in the minor groove (*23*−*27*). The DNA is dramatically bent toward the major groove, and the distortion is induced by the insertion of phenylalanine side chains into the first and last steps of the TATA box and is favored by the intrinsic bendability of the TA steps. Most interestingly, examination of the base pair overlaps of the various TA steps in the structure suggested that TA* photoproduct formation might be inhibited at certain sites and that a different class of TA photoproducts might be enhanced.

A number of years ago, it was discovered that UVC-induced photoproduct formation at a TT site within a TATA box in the yeast *GAL1* and *GAL10* promoters was enhanced in its transcriptionally active state (*7*). More recently, Thoma and co-workers concluded that the product being enhanced in the *GAL10* promoter, as well as the *SNR6* promoters, was the (6-4) product and that the cis-syn dimer was being suppressed (*28*). Their conclusion was based on the results of differential cleavage assays with the cis-syn dimer-specific T4 endonuclease V enzyme and the *Neurospora crassa* UV damage endonuclease (UVDE), which had been shown to cut specifically at sites of cis-syn dimers and (6-4) photoproducts (*29*). It is likely, however, that the *N. crassa* UVDE enzyme also cleaves at sites of other photoproducts, because the homologous UVDE from *Schizosaccharomyces pombe* is able to recognize an array of UV-induced DNA photoproducts including cis-syn, trans-syn-I, trans-syn-II, (6-4), and Dewar photoproducts (*30*). As a result, such an assay can be used to distinguish only between cis-syn dimers and some other type of photodamage but not to establish rigorously the type of damage.

To identify unambiguously and quantify isomeric DNA photoproducts formed in DNA, we developed a combined enzymatic digestion/tandem mass spectrometric assay (*31*). In this assay, a photoproduct-containing oligodeoxynucleotide is digested with nuclease P1 to give photoproduct-containing trinucleotides, which can be distinguished by distinctive fragmentations of their $[M - H]^-$ ions. The relative amounts of the isomeric photoproducts can then be quantified directly from the intensity ratios of the characteristic fragment ions (*32*). Herein, we make use of this method, along with site-specific isotopic labeling, to investigate of the effect of TBP binding on the photochemistry of the adenovirus major late promoter (AdMLP) TATA box DNA sequence. We intend to establish rigorously that the photoproduct preferentially induced by TBP binding in UVC light is indeed the (6-4) product and conversely that TA* photoproduct formation is reduced by TBP binding. We will also show that TBP binding increases the preference for forming the (6-4) product relative to the cis-syn dimer with increasing doses of UVC light. In addition, we will show that TBP binding also enhances (6-4) photoproduct formation by the more biologically relevant UVB light present in sunlight at sea level, and we will propose mechanisms to account for all of the observed effects.

EXPERIMENTAL PROCEDURES

Materials and Methods. The yTBPc expression vector was provided by late Prof. Paul Sigler (Yale University). BL21-(DE3) *E. coli* competent cells were from Novagen (Madison, WI). Ni−NTA resin was from Qiagen (Valencia, CA). Thymidine was obtained from Peninsula Laboratories, Inc. PtO_2 was from Aldrich. Automated DNA synthesis was carried out on an Expedite 8909 synthesizer (Applied Biosystems, Foster City, CA). Commercially prepared oligodeoxynucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, IA) and used without further purification. Nuclease P1 (200 units/mg) was from Boehringer-Mannheim (Mannheim, Germany) and was used without further purification. Photoproduct-containing oligonucleotides used as standards were prepared and purified as previously described (*19*, *31*, *33*, *34*). UV irradiation was carried out with a Spectroline Model XX-15F lamp housing fitted with a Longlife filter. UVC irradiation was carried out with two 15 W BLE-1T155 bulbs (254 nm), and irradiation with UVB light was carried out with two 15 W BLE-1T158 lamps (broad band from 270 to 400 nm peaking at 310 nm with spikes at 313 and 365 nm). ESI mass spectrometry was carried out on a Classic LCQ ion trap mass spectrometer (Finnigan, San Jose, CA).

Preparation of Yeast TATA-Binding Protein. A pET-21d vector containing the C-terminal conserved region of *Saccharomyces cerevisiae* TBP (yTBPc) with a His7 tag at the N-terminus (MGHHHHHHHGSLVPRGSRGS**S61**...**M240**) was used to transform competent BL21(DE3) *Escherichia*

coli cells. The cells were grown in Terrific broth medium and were induced by 0.1 mM IPTG when the absorbance at 600 nm reached 1.0. After 6 h of induction, cells were harvested by centrifugation, the cell pellet was resuspended in lysis buffer (250 mM ammonium acetate, 10 mM PMSF, 10% glycerol, 25 mM Tris, pH 8.0), and the cells were ruptured by a French press or by sonication. After centrifugation at 13000 *g* for 30 min, the supernatant was collected and passed through a Ni−NTA column. The column was washed sequentially with lysis buffer and lysis buffer with 20, 50, 80, 110, 150, and 200 mM imidazole. The fractions were monitored by SDS−PAGE, and those containing yTBPc were combined and stored at −80 °C. The imidazole-containing protein solution was diluted to approximately 0.1 mg/mL immediately before use and dialyzed three times for 4 h against 1 L of 250 mM ammonium acetate, 100 mM potassium chloride, and 10% glycerol.

Preparation of Thymidine-d_3-Containing Oligodeoxynucleotides. Thymidine perdeuterated at the methyl group was prepared by a published procedure (*35*) with some modification. Three grams of thymidine in 20 mL of D_2O was placed in 100 mL round-bottom flask, to which was added 0.167 g of PtO_2. The flask was evacuated and filled with H_2 twice. The solution was then heated at 60 °C and held under a positive pressure H_2 atmosphere by way of a hydrogen-filled balloon. After 7 days, the solution was filtered and returned to the reaction flask, a second portion of PtO_2 (0.167 g) was added, and the mixture was heated for another 14 days under a H_2 atmosphere. The reaction products were separated by flash column chromatography with 80/20 dichloromethane/2-propanol as the solvent system to give 1.0 g of the desired product. Proton NMR and product-ion mass spectra (ESI ion trap in the MS/MS mode) of the thymine fragment ion showed that >95% of the C5-methyl hydrogens were exchanged for deuterium. The β-cyanoethyl phosphoramidite of thymidine-d_3 was synthesized by a standard procedure (*36*) and used to prepare d(CCTTT**T**ATAGCC) (**T** = dT-d_3) by standard automated DNA synthesis.

Preparation and Irradiation of the TATA Box DNA. The AdMLP TATA box DNA was prepared by annealing d(GGCTATAAAAGG) and d(CCTTTTATAGCC) (1 mM each) by heating in 250 mM ammonium acetate and 100 mM potassium chloride for 3 min in a 93 °C water bath, followed by cooling to room temperature over a 2 h period. The annealed duplex was then diluted to 3.75 μM in the same buffer containing 10% glycerol, degassed with N_2 for 5 min, transferred to a 3.2 cm i.d. centrifuge cap, put into a N_2-filled zip-lock bag, and irradiated on ice for the times indicated in the figures. UVB and UVC irradiation was carried out at a measured intensity of approximately 2 mW/cm^2. Irradiation of the duplex in the presence of the TBP protein was carried out in a similar fashion, except that the indicated concentration of protein was allowed to equilibrate with the duplex at room temperature for 30 min. Following irradiation, the solutions were dialyzed three times for 4 h against 1 L of doubly distilled (dd) H_2O, phenol extracted, dried by centrifugal evaporation under vacuum in a Savant Speedvac, and then desalted by dialysis against dd H_2O. Initially, high doses of UVC light (100 000 J/m^2) were used to increase the yield of photoproducts so that product analysis could be achieved with good signal-to-noise ratio in the mass spectra. There was no evidence under these conditions that sufficient amounts of other products were being formed that were interfering with the subsequent enzyme-coupled mass spectrometric analysis.

Combined Enzymatic/Mass Spectrometric Assay of Photoproducts. Enzymatic digestion of the irradiated samples was carried out with 1 μL of 1 unit/μL nuclease P1 in 25 μL of 10 pmol/μL DNA for 5 min at 37 °C. ESI-MS experiments were carried out as described earlier (*34*). Briefly, 50/50/0.5 (v/v) $CH_3CN/H_2O/NH_3$ was used as the carrier solvent for electrospray, and the flow rate was held at 5 μL/min by using a syringe injector. A 5 μL aliquot of the digestion solution was injected in each run. The spray voltage was 4.6 kV, and the capillary temperature was maintained at 200 °C. The mass width for precursor selection was set at m/z = 2 for determination of the isotope-labeled samples and at m/z = 3 for determination of the other samples. The center masses selected for zoom-scan experiments of synthetic dT-d_3 and the nuclease P1 digestion of irradiation products were 243 and 931, respectively.

Quantification of Photoproduct Formation. To quantify the relative amount of cis-syn and (6-4) isomers formed, calibration curves were obtained using authentic mixtures of cis-syn and (6-4) product-containing ODNs d(GTAT[]-TAT) and d(GT[]TTGA) (*32*, *34*) (10 μM total) as previously described (*32*). To determine the relative yields of TA* products, calibration curves were obtained for known mixtures of purified TA* product-containing ODNs d(GGC-TATA*A), d(GTATTA*T) (*34*), d(GAGTA*CG), and d(GAGTA*GG) with different bases flanking the 3′ side of the TA* product, and MS/MS was carried out by *alternately* scanning two different channels. For example, the spectrometer was set to acquire the product-ion spectrum of the [M − H]$^-$ ion of d(pTA*A) in channel A and d(pTA*C) in channel B by scanning each channel for 3 s. The spectra obtained in the two channels were then signal averaged, and the peak area of the fragment ion of m/z 616 that is characteristic of TA* photoproducts was used for calibration. To determine the absolute amount of TA* product, 5 μM purified d(GGCTATA*A) was added to 25 μM irradiated TATA box DNA. The product-ion spectrum of the [M − H]$^-$ ion of d(pTA*A) was then acquired in one channel and that of d(pTA*T) or d(pTA*G) was obtained in the other. The absolute amounts of the TA* formed at p(TA)*T and p(TA)*G sites were then calculated on the basis of the intensity ratios of peaks corresponding to the product ion of m/z 616 in the two channels.

RESULTS AND DISCUSSION

We chose to investigate photoproduct formation in the TATA box of the adenovirus major late promoter (AdMLP) instead of the *GAL1*, *GAL10*, or *SNR6* promoters for two principal reasons. The first is that there are a number of crystal structures containing the AdMLP core sequence complexed to TATA-binding proteins. The second is that the AdMLP sequence, unlike the others, contains a T_4 tract that can be used to probe photoproduct formation at multiple sites within the complex. We chose to study the effect of yeast TBP on photoproduct formation because the initial observations of photoproduct enhancement were made with the yeast protein. In addition, the interaction between yeast

TBP and AdMLP TATA box has been thoroughly characterized in solution. The protein binds bidirectionally as a mixture of axial isomers, and only in the presence of the transcription factors TFIIA and TFIIB will it bind with a highly specific orientation and axial position (*37*, *38*). To lay a foundation for a thorough investigation of this system, we first chose to examine the effect of the C-terminal DNA binding domain of yeast TBP on the photochemistry of the AdMLP TATA box.

Combined Enzymatic Digestion/Mass Spectrometry Assay. To identify and quantify photoproducts formed in TATA box DNA, we made use of a combined enzymatic digestion/mass spectrometry assay that we recently developed to study photoproduct formation in oligodeoxynucleotides (*31*). This assay is based on the ability of nuclease P1 to degrade DNA containing internal dinucleotide photoproducts to trinucleotides of the form d(pN[]NN), where [] indicates a photomodification. Photoproduct-containing trinucleotides result because the active site of nuclease P1 has a binding pocket for a base 5′ to the phosphodiester bond to be cleaved (*39*). Because the bases in dinucleotide photoproducts, such as the cis-syn dimer and (6-4) product, are fused together, they cannot be bound by the enzyme, and hence cleavage cannot occur at the 3′-phosphodiester bonds adjacent to either base of the photoproduct. Photoproducts arising in the same sequence context can be distinguished from each other because they have distinctive fragmentation patterns (*31*).

The most abundant product ion formed in the fragmentation of the $[M - H]^-$ ion of d(pT[c,s]TA) occurs at *m*/*z* 803 and corresponds to the loss of adenine. In contrast, the most abundant product ion of the isomeric $[M - H]^-$ ion of d(pT-[6-4]TA) occurs as an ion of *m*/*z* 825, which is produced by the loss of a neutral fragment ($C_4H_3NO_3$), although a significant amount of the ion of *m*/*z* 803 is also produced, and this comes from the loss of adenine. The most abundant product ion of the $[M - H]^-$ ion of the isomeric d(pTA*T) photoproduct is an ion at *m*/*z* 616 formed by the loss of pdT. The cis-syn and (6-4) photoproducts in d(pT[]TG) (data not shown) and d(pT[]TT) (*32*) can be distinguished in a similar manner. The only difference between the (6-4) and Dewar photoproducts of d(pT[]TA) is that the fragment ion at *m*/*z* 825 of the Dewar product is only approximately one-third as abundant as of that of the (6-4) product (*31*). The fragmentation behavior of the (6-4) and Dewar products of d(pT[]TT) were not previously investigated.

Photoproduct Formation in the TATA Box Sequence in the Presence or Absence of yTBPc. Comparison of the product-ion mass spectra of the nuclease P1 digestion products of the UVC-irradiated TATA DNA, in the presence or absence of yTBPc (Figure 2), shows that TBP binding greatly affects the distribution of these photoproducts. Whereas the fragment ion at *m*/*z* 616 corresponding to the TA* product is very prominent in the absence of TBP, it is greatly diminished in the presence of TBP. Although the fragment ion corresponding to (6-4) and/or Dewar products is of very low abundance in the absence of TBP, there is a prominent ion of *m*/*z* 803 which would indicate that the cis-syn dimer of TT is the predominant photoproduct under these conditions. In the presence of TBP, however, the ion of *m*/*z* 825, which is characteristic of the (6-4) and/or Dewar products, becomes the most prominent fragment. A similar trend can be seen for the two d(pT[]TT) sites.

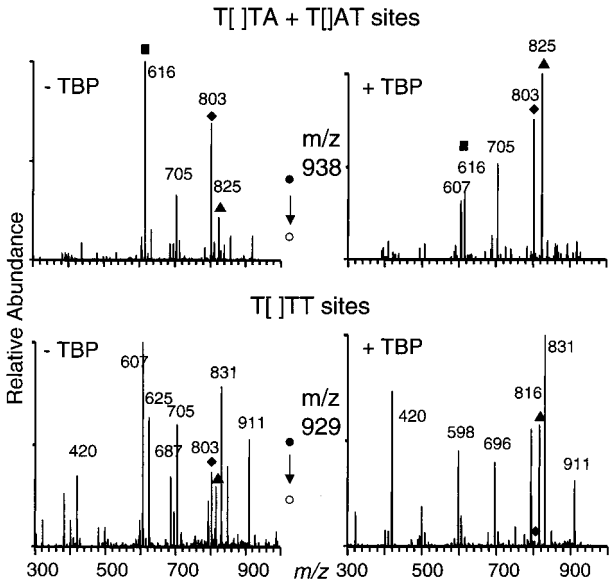

FIGURE 2: Product-ion spectra of ESI-produced $[M - H]^-$ ions of *m*/*z* 929 (pT[ ]TT) and 938 (pT[ ]TA and pT[ ]AT) from the nuclease P1 digestion products of the 3.75 μM TATA duplex that had been subjected to 1.5 h of UVC irradiation (~100 000 J/m^2) at 0 °C in the presence of 9.4 μM TBP. The product ions characteristic of the various products are labeled as ▲ for (6-4) or Dewar, ◆ for cis-syn, and ■ for TA*.

Distinguishing (6-4) and Dewar Photoproducts. Because the product-ion spectra of $[M - H]^-$ ions of d(pT[6-4]TT) and d(pT[Dewar]TT) show no obvious difference (data not shown), and those of d(pT[6-4]TA) and d(pT[Dewar]TA) show only a difference in relative abundances of the fragment ion formed by a loss of 113 Da (*31*), we needed an alternative method for distinguishing these two products. After some experimentation, we found that the Dewar product of d(pT[]-TA) gives a unique fragment ion at *m*/*z* 749 $[M - 113 - H_3PO_4 + Na - 2H]^-$, as seen in the product-ion spectra of $[M + Na - 2H]^-$ ions (Figure 3A,B). Although the abundance of this unique fragment ion is not high, the spectra are reproducible. The distinction also holds true for the (6-4) and Dewar products of d(pT[]TT) and d(pT[]TG) (spectra not shown). Compared with the two reference spectra, the product-ion spectrum of the nuclease P1 digestion products of the UVC-irradiated TATA box-containing ODN in the presence of TBP shows that the preferentially formed photoproduct at the d(pT[]TA) site is the (6-4) product, not its Dewar valence isomer (Figure 3C). A similar result was found for the d(pT[]TT) site (data not shown).

Quantifying the Effects of TBP Binding on the Selectivity of TT Photoproduct Formation. To quantify the selectivity of cis-syn and (6-4) photoproduct formation by mass spectrometry, we needed to obtain calibration curves for the two photoproducts (*32*). To do so, we mixed authentic cis-syn and (6-4) product-containing ODNs in different molar ratios, digested them with nuclease P1, and measured the peak intensities of the [M − base] and the [M − 113] ions, which are characteristic of the cis-syn and (6-4) photoproducts, respectively. Calibration curves made by plotting $I_{M-133}/(I_{M-113} + I_{M-base})$ versus fraction of (6-4) product were linear for mixtures of d(pT[c,s]TN) and d(pT[6-4]TN) when N = A or G. In contrast, the calibration curve for a mixture of d(pT[c,s]TT) and d(pT[6-4]TT) is not linear, which can be explained by assuming that the ionization and fragmentation

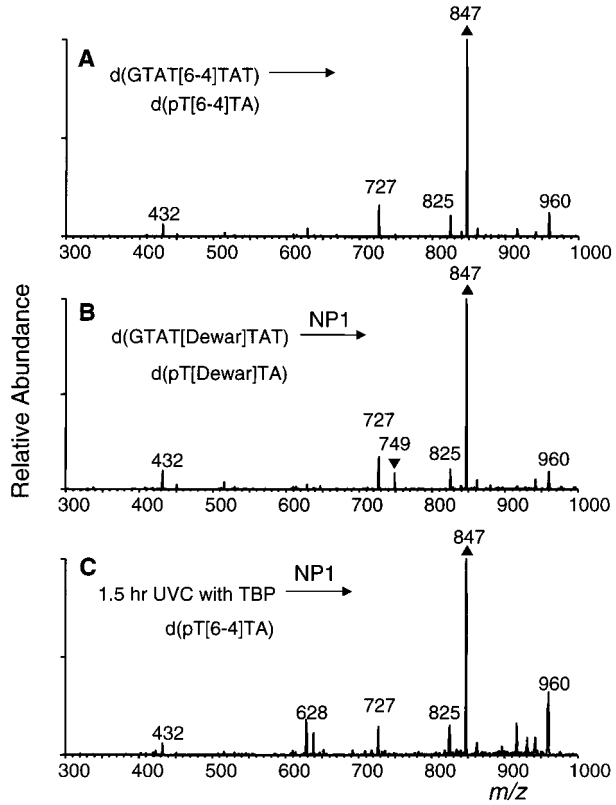

FIGURE 3: Product-ion spectra of ESI-produced $[M + Na - H]^-$ ions of pT[ ]TA obtained from nuclease P1 digestion of authentic d(GTAT[6-4]TAT), d(GTAT[Dewar]TAT), and the photoproduct mixture resulting from 1.5 h UVC irradiation (~100 000 J/m^2) of the TATA box duplex in the presence of TBP (TBP/DNA molar ratio 2.5, vide infra). The product ions characteristic of the various products are labeled as ▲ for (6-4) or Dewar and ▼ for Dewar alone.

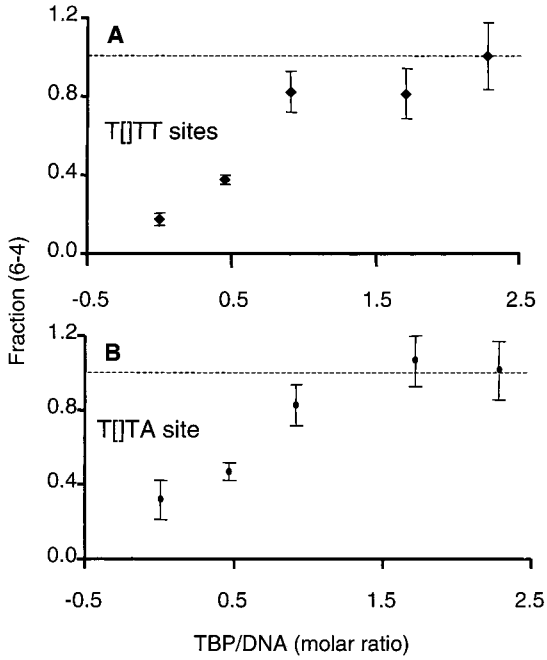

FIGURE 4: Fraction of (6-4) product formed in 3.75 μM TATA box DNA upon 1.5 h of UVC irradiation as a function of the number of equivalents of yTBPc: (a) T[]TT sites; (b) T[]TA site. Error bars were standard deviations based on four independent measurements.

efficiencies of the two isomers are different (*32*). To quantify the effect of TBP binding on the selectivity of photoproduct formation, the relative yields of cis-syn and (6-4) photoproducts were determined as a function of protein concentration in the presence of 3.75 μM DNA (Figure 4). With increasing ratios of TBP to TATA DNA, the fraction of (6-4) photoproducts increased linearly from approximately 0.3 with no TBP to nearly 1.0 with 1 equiv of TBP. The linear behavior points to the formation of a tightly bound complex, which is consistent with the reported K_d value of approximately 2×10^{-9} M (*40*).

Possible Origins of Photoproduct Formation Selectivity. There are a number of conceivable factors that could explain the almost exclusive formation of (6-4) photoproducts in the T track of TATA box upon intensive irradiation with UVC in the presence of TBP. One possibility is that the DNA in the bound complex adopts a conformation that kinetically favors the formation of (6-4) photoproducts over cis-syn dimers. Indeed, the average distances between atoms of adjacent T's in the T_4 tracts of two crystal structures of TATA sequences bound by TBP (*23*, *24*) are consistently shorter for (6-4) photoproduct formation than for cis-syn dimer formation (3.83 ± 0.36 vs 4.26 ± 0.27 Å). Distances alone, however, are not likely to be a good indicator of photoproduct formation efficiency, as the distances are also shorter in a reported crystal structure of an unbound DNA duplex containing a T_4 tract (*41*) (4.13 ± 0.19 vs 4.42 ± 0.2 Å). More significantly, the distances appear to decrease more upon TBP binding for (6-4) product formation than for cis-syn dimer formation (0.3 vs 0.16 Å). The most important feature is likely to be whether the structure has the flexibility to achieve the transition state geometry (*42*), which will depend on the dynamics of the bases involved. This is not so easy to ascertain from a crystal structure alone.

Another possible explanation relates to the fact that cis-syn photoproducts are well-known to revert to the parental DNA upon UVC irradiation, whereas the (6-4) photoproducts do not (*43*, *44*). It is, therefore, possible that both cis-syn and (6-4) products are being formed initially, but upon prolonged irradiation with UVC light, the cis-syn dimers photorevert to their parental forms, which irreversibly convert to the (6-4) products at a rate that is enhanced by TBP binding (Figure 1). To test this idea, we examined the fraction of (6-4) photoproduct formed as a function of UVC irradiation time in the range of 0.5−10 min and found that the fraction of (6-4) product does indeed increase with increasing irradiation time (Figure 5). It is also possible that TBP enhances the reversal of cis-syn dimers, although we do not have enough data to address this issue at this time.

Selectivity of TT Photoproduct Formation by UVB Light. To test further the idea that photoreversal might play a role in determining the observed selectivity of (6-4) photoproduct formation by UVC light, we irradiated the TATA-containing ODNs with UVB light, which does not cause significant photoreversal of cis-syn dimers. In the absence of TBP, no (6-4) photoproduct formation could be detected at either TTT or TTA sites (Figure 6), but the (6-4) photoproduct did form at both sites in the presence of TBP (Figures 6 and 7), although with a lower selectivity than that observed with low doses of UVC light (Figure 5). This would suggest that TBP binding does indeed enhance (6-4) photoproduct formation. When the UVB-irradiated product was further irradiated with UVC light, the (6-4) product was again formed

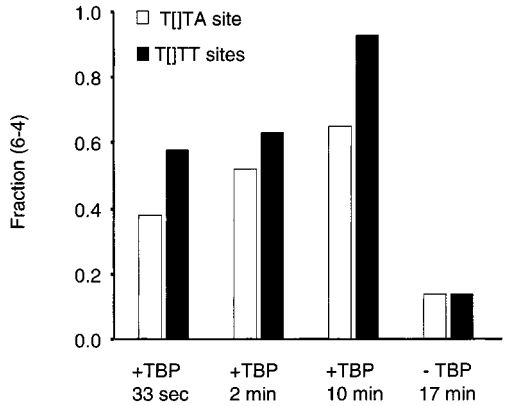

FIGURE 5: Time dependence of the fraction of (6-4) product formed in the 3.75 μM TATA box sequence by UVC irradiation in the presence or absence of 7.5 μM TBP. The doses used were approximately 100 J/m^2 (33 s), 360 J/m^2 (2 min), 900 J/m^2 (10 min), and 1500 J/m^2 (17 min).

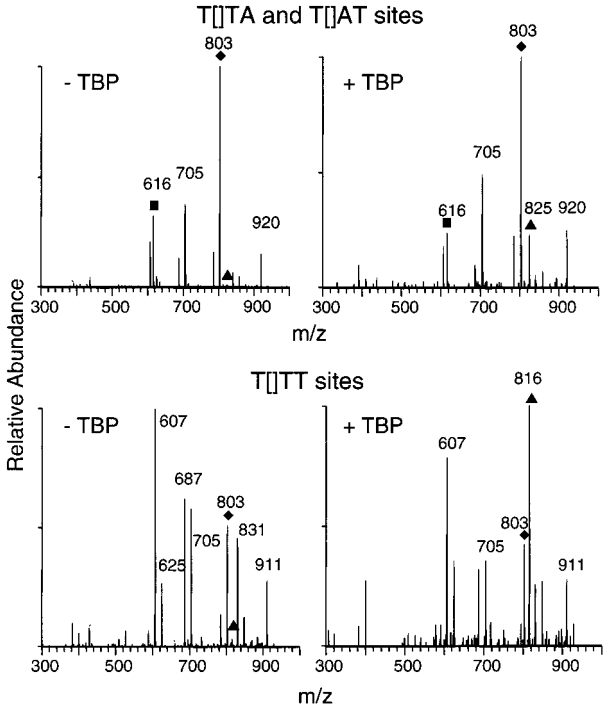

FIGURE 6: Product-ion spectra of ESI-produced negative ions at *m*/*z* 929 (pT[ ]TT) and 938 (pT[ ]TA and pT[ ]AT) from the nuclease P1 digestion products of 90 min of UVB (~100 000 J/m^2) irradiation products of 3.75 μM TATA box DNA in the presence or absence of 7.5 μM TBP. The product ions characteristic of the various products are labeled as ▲ for (6-4), ◆ for cis-syn, and ■ for TA*.

preferentially (Figure 7), demonstrating that the selectivity of (6-4) photoproduct formation is also enhanced by the photoreversal of the cis-syn dimers.

Quantifying TT Photoproduct Formation Selectivity at Otherwise Identical Sites. In the experiments described above with native DNA, we were unable to distinguish between photoproduct formation at $T_{15}T_{16}$ and $T_{16}T_{17}$ because nuclease P1 digestion of photoproducts formed at these sites both give d(pT[]TT). To distinguish these two sites, we site-specifically labeled the methyl group of T_{18} (Figure 1). By doing so, photoproduct formation at $T_{15}T_{16}$ could be distinguished from photoproduct formation at $T_{16}T_{17}$ because the trinucleotide d(pT_{16}[]$T_{17}T_{18}$) produced by nuclease P1 digestion will be 3 Da greater than d(pT_{15}[]$T_{16}T_{17}$), which lacks the deuterated

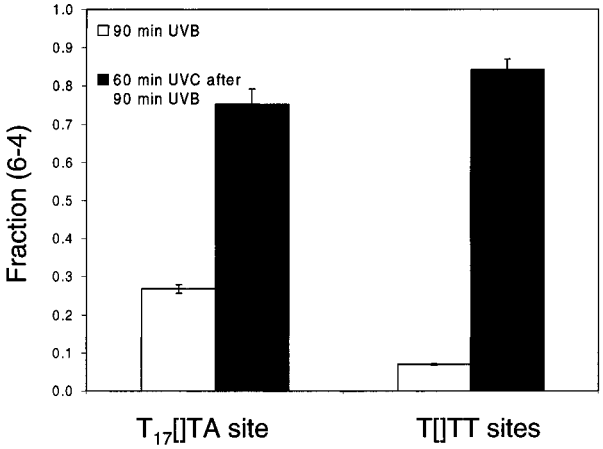

FIGURE 7: Fraction of (6-4) photoproduct formed in 3.75 μM TATA box DNA in the presence of 7.5 μM TBP by 90 min of UVB irradiation or by 90 min of UVB (~100 000 J/m^2) followed by 60 min of UVC irradiation. Error bars are standard deviations based on three independent measurements.

nucleotide. The T_{18} site was also chosen for labeling because it could also be used to distinguish between TA* formation at the two otherwise identical TAT sites in the TATA box DNA (see next section). The site-specifically labeled dodecamer d(CCTTT**T**ATAGCC), where **T** is dT-d_3, was prepared by standard automated solid-phase synthesis with thymidine-d_3 phosphoramidite which had been prepared by standard methodology (*36*) from thymidine-d_3 (*35*). In the absence of TBP, the selectivity of (6-4) product formation by UVC light is least at the $T_{16}T_{17}$ site and greatest at the $T_{15}T_{16}$ and $T_{17}T_{18}$ sites, whereas in the presence of TBP, it is greatest at the $T_{16}T_{17}$ site (Figures 8 and 9). These results are consistent with the observation that the shortest distances for the formation of (6-4) products are at the $T_{16}T_{17}$ site in the X-ray crystal structures of TBP/TATA box complexes (3.65 ± 0.34 Å vs an average of 3.83 ± 0.36 Å) (*23*, *24*). The observed effect at the $T_{16}T_{17}$ site is likely to represent the effect of averaging the many different bound species that Schepartz and co-workers showed to exist in solution (*37*, *38*). Upon UVB irradiation, (6-4) product formation in the absence of TBP binding is barely detectable (data not shown) but increases dramatically in the presence of TBP and parallels what was observed for UVC irradiation.

Effect of TBP on TA Photoproduct Formation.* To quantify the effect of TBP on TA* product formation, we extended our previously described methodology for quantifying the selectivity of cis-syn and (6-4) photoproduct formation for the determination of the absolute concentration of the TA* photoproduct in an isomeric mixture. To do so, we obtained calibration curves for TA* products flanked on the 3′ side by different bases by using known mixtures of purified TA*-containing ODNs. We found that the yields of the characteristic fragment ion of *m*/*z* 616, which results from loss of pdN from d(pTA*N), are very similar for TA* products with different 3′ flanking bases. By using the TATA box duplex labeled at T_{18}, we could distinguish TA* formation at the $T_4A_5T_6$ and $T_{18}A_{19}T_{20}$ sites because the $[M - H]^-$ ions of d(p***T***$_{18}$A*T) and d(pT_4A*T) differ by three mass units. To determine the absolute amount of TA* photoproducts produced at d(pTAG) and d(pTAT) sites, a known amount of a pure TA*A-containing ODN d(GGCTATA*A) (*34*) was added to the T_{18}-d_3 TATA box duplex following irradiation but prior to nuclease P1 digestion.

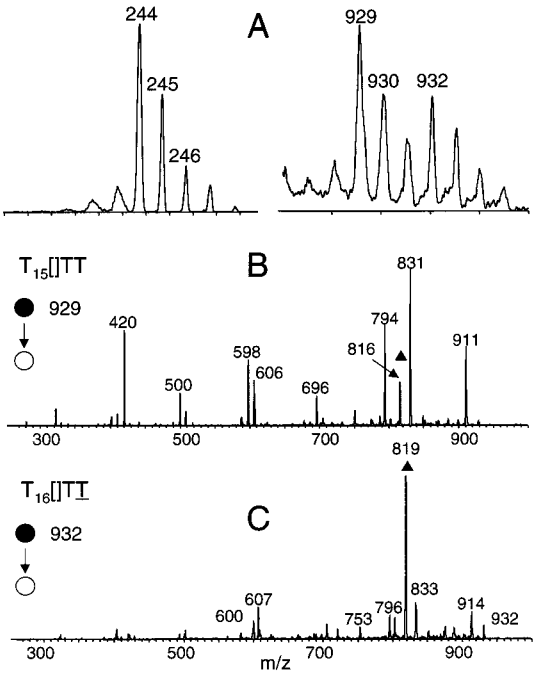

FIGURE 8: (A) Left panel: Negative ion zoom-scan ESI-MS of prepared deuterated thymidine. Right panel: Negative ion zoom-scan ESI-MS of the nuclease P1 digestion product of T_{18}-d_3-labeled TATA box DNA (3.75 μM) irradiated with 1.5 h UVC in the presence of 7.5 μM TBP. (B) Product-ion spectrum of the $[M - H]^-$ ion of *m/z* 929 (pT=TT) in (A). (C) Product-ion spectrum of the $[M - H]^-$ ion of *m/z* 932 (pT=T**T**) in (A). The product ion established to be characteristic of the (6-4) product is labeled as ▲. The additional peaks seen in spectrum B compared to spectrum C are likely to be due to contamination of the *m/z* 929 parent ion with an impurity ion.

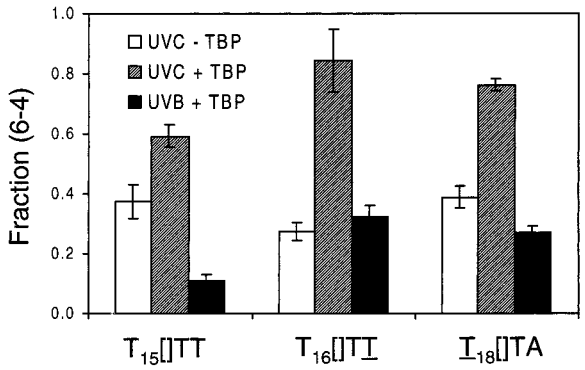

FIGURE 9: Specificity of (6-4) photoproduct formation in the TATA box DNA (3.75 μM) resulting from 1.5 h of UVC irradiation in the absence or the presence of 7.5 μM TBP as determined by isotope labeling and tandem mass spectrometry. UVB irradiation of the duplex DNA in the absence of TBP produces very little (6-4) photoproduct and was not shown in this figure.

Because some d(pTA*A) product also comes from irradiation of the TATA-containing substrate, the actual yields of the other TA* photoproducts are likely to be somewhat higher.

Figure 10 shows that formation of TA* by UVC irradiation is suppressed at the two TAT sites upon TBP binding, whereas TBP binding apparently does not affect TA* formation at the TAG site. No significant amount of TA* photoproduct was detected under UVB irradiation. The moderate effect of TBP binding on TA* formation by UVC cannot be readily explained by consideration of the relevant interatomic distances or by other structural features of the TBP/TATA box complex. From the available crystal structures of human TBP with the AdMLP core TATA box sequence, it would appear that the phenyl side chain of a phenylalanine should be wedged between T_4A_5·$T_{20}A_{21}$; yet surprisingly this interaction does not appear to interfere greatly with TA* formation at these sites. Again, the dynamics of the bases may be the most important determinant of photoproduct yield. On the other hand, the lack of a highly site-specific effect of TBP binding on TA* formation is consistent with the solution binding studies of Schepartz and co-workers, which indicate that yeast TBP binds bidirectionally and in multiple registers (*37*, *38*). Such equilibrating structures would be expected to average the effects of any one structure on the photochemistry of the TATA box.

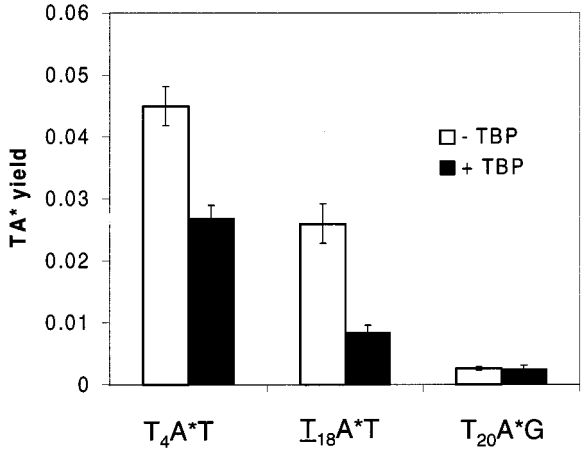

FIGURE 10: Absolute yields of TA* photoproduct formed in 25 μM TATA box DNA isolated from 1.5 h of UVC irradiation of 3.75 μM TATA box DNA in the presence or the absence of 7.5 μM TBP.

Biological Implications. Sunlight at sea level consists of wavelengths above approximately 280 nm, which includes UVB but not UVC light. The ability of TBP to enhance the formation of (6-4) photoproducts in TATA boxes by UVB light may therefore have important biological consequences. If the (6-4) product enhances binding of TBP to a TATA box, it may cause an increase in transcription of the damaged gene and at the same time decrease the transcription of undamaged genes by selectively recruiting transcription factors, as occurs in transcription factor hijacking (*45*, *46*). A further consequence of enhanced binding of the photodamaged DNA to a transcription factor would be that repair of the photoproduct would be inhibited, increasing the probability of a mutagenic event that could permanently affect transcription of the affected gene. TBP is known to bind to UVC-irradiated heterogeneous DNA (*47*), although the photoproducts to which the TBP was binding were never established. It is known, however, that photorepair of a cis-syn dimer or excision repair of a (6-4) product within the T_3 tract of the *SNR6* promoter, d(TATAAATA)·d(TATTTATA), is not inhibited by the presence of TBP (*28*), suggesting that both products inhibit rather than enhance TBP binding to the TATA box. In contrast, the authors of that study argued that TBP should be expected to bind tightly to the (6-4) product because TBP binding enhances its formation, and they proposed other explanations for the failure of TBP to inhibit its excision repair.

It does not necessarily follow, however, that TBP should bind to the (6-4) product in the TATA box simply because TBP enhances (6-4) product formation. Formation of the

(6-4) product is a two-step process (Figure 1) that involves initial photochemical formation of an oxetane intermediate, followed by a thermal rearrangement to the (6-4) product, which has a dramatically different structure. So while TBP might enhance formation of the oxetane intermediate by stabilizing the transition state leading to this product, there is no reason to expect that it should bind well to the (6-4) product that results from rearrangement of the initially formed oxetane intermediate. If the (6-4) product decreases the affinity of TBP for the promoter, which is more likely to be the case, it could reduce transcription of the damaged gene, which in turn could have transient mutagenic consequences if the gene affected plays an important role in DNA repair. Photoproduct formation within a TATA box might also inhibit transcription by changing the orientation of TBP. The only way to address these issues will be to determine the binding affinity and orientation of TBP for TATA box sequences containing site-specific photoproducts.

Given the rather substantial effect of TBP binding on the photochemistry of TATA boxes, it is likely that other proteins that change the DNA conformation may also significantly affect the photochemistry of the bound DNA. For example, the transcription activator SRY is able to interact with A/T base pairs (*48*), and the 3-D structure of SRY and its natural substrates reveal a common motif of side-chain intercalation driving the deformation of the DNA helix (*49*). High-mobility group 1 and 2 (HMG-1 and -2) proteins are abundant chromosomal proteins in a variety of eukaryotic cells (*50*), and they have a general feature of being able to induce DNA bending independent of sequence and to bind to bent DNA (*51*). HMG-1 protein binds preferentially to distorted sites such as the DNA lesions formed by the anticancer drug cisplatin (*52*, *53*) and UVC-irradiated ODNs (*54*).

CONCLUSIONS

By combining nuclease P1 digestion and tandem mass spectrometry, we demonstrated unambiguously that the UVC photoproduct enhanced by yTBPc binding is the (6-4) photoproduct, which is presumably the same product observed to be enhanced in studies of yeast promoters (*28*). We determined that TBP binding induces the highly selective formation of the (6-4) product by high doses of UVC light, in part because TBP binding enhances (6-4) product formation and in part because cis-syn dimer formation is photoreversible under these conditions. We also showed that TBP binding enhances formation of the (6-4) product by the more biologically relevant UVB light found in sunlight at sea level. In addition, we demonstrated that both quantitative and site-specific information can be obtained from the combined nuclease P1 digestion and tandem mass spectrometry method in conjunction with site-specifically isotopically labeled DNA. We believe that nuclease P1 digestion coupled with tandem mass spectrometry and isotope labeling is a powerful quantitative method for studying DNA damage formation and processing at specific sites in DNA duplex substrates containing multiple and redundant sites. It will now be interesting to see how the selectivity of photoproduct formation in the AdMLP TATA box is affected by TBP binding in the presence of transcription factors known to increase the selectivity of binding.

ACKNOWLEDGMENT

The authors thank the late Prof. Paul B. Sigler and Dr. Sean Juo (Yale University) for providing the yTBPc vector and protocol for purification of the yTBPc.

REFERENCES

1. Cadet, J., and Vigny, P. (1990) in *Bioorganic Photochemistry* (Morrison, H., Ed.) pp 1−272, John Wiley and Sons, New York.
2. Sarasin, A. (1999) *Mutat. Res. 428*, 5−10.
3. Lehmann, A. R. (2000) *Gene 253*, 1−12.
4. Donnellan, J. E., Jr., and Setlow, R. B. (1965) *Science 149*, 308−310.
5. Nicholson, W. L., Setlow, B., and Setlow, P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8288−8292.
6. Becker, M. M., and Wang, J. C. (1984) *Nature 309*, 682−687.
7. Selleck, S. B., and Majors, J. (1987) *Nature 325*, 173−177.
8. Becker, M. M., Lesser, D., Kurpiewski, M., Baranger, A., and Jen-Jacobson, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6247−6251.
9. Tornaletti, S., and Pfeifer, G. P. (1995) *J. Mol. Biol. 249*, 714−728.
10. Gale, J. M., Nissen, K. A., and Smerdon, M. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6644−6648.
11. Schieferstein, U., and Thoma, F. (1996) *Biochemistry 35*, 7705−7714.
12. Mitchell, D. L., Nguyen, T. D., and Cleaver, J. E. (1990) *J. Biol. Chem. 265*, 5353−5356.
13. Gale, J. M., and Smerdon, M. J. (1990) *Photochem. Photobiol. 51*, 411−417.
14. Suquet, C., Mitchell, D. L., and Smerdon, M. J. (1995) *J. Biol. Chem. 270*, 16507−16509.
15. Fekete, A., Vink, A. A., Gaspar, S., Modos, K., Berces, A., Ronto, G., and Roza, L. (1999) *Photochem. Photobiol. 69*, 545−552.
16. Liu, X., Conconi, A., and Smerdon, M. J. (1997) *Biochemistry 36*, 13710−13717.
17. Zhao, X., and Taylor, J.-S. (1996) *Nucleic Acids Res. 24*, 1561−1565.
18. Zhao, X., Nadji, S., Kao, J. L.-F., and Taylor, J.-S. (1996) *Nucleic Acids Res. 24*, 1554−1560.
19. Zhao, X., Kao, J. L. F., and Taylor, J.-S. (1995) *Biochemistry 34*, 1386−1392.
20. Burley, S. K. (1996) *Curr. Opin. Struct. Biol. 6*, 69−75.
21. Struhl, K., and Moqtaderi, Z. (1998) *Cell 94*, 1−4.
22. Burley, S. K., and Roeder, R. G. (1996) *Annu. Rev. Biochem. 65*, 769−799.
23. Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995) *Nature 377*, 119−128.
24. Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., and Burley, S. K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4862−4867.
25. Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature 365*, 512−527.
26. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature 365*, 520−527.
27. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol. 261*, 239−254.
28. Aboussekhra, A., and Thoma, F. (1999) *EMBO J. 18*, 433−443.
29. Yajima, H., Takao, M., Yasuhira, S., Zhao, J. H., Ishii, C., Inoue, H., and Yasui, A. (1995) *EMBO J. 14*, 2393−2399.
30. Avery, A. M., Kaur, B., Taylor, J. S., Mello, J. A., Essigmann, J. M., and Doetsch, P. W. (1999) *Nucleic Acids Res. 27*, 2256−2264.
31. Wang, Y., Taylor, J. S., and Gross, M. L. (1999) *Chem. Res. Toxicol. 12*, 1077−1082.

32. Wang, Y., Rempel, D. L., Taylor, J. S., and Gross, M. L. (2001) *Anal. Chem. 73*, 185−191.
33. Smith, C. A., Baeten, J., and Taylor, J. S. (1998) *J. Biol. Chem. 273*, 21933−21940.
34. Wang, Y., Taylor, J. S., and Gross, M. L. (1999) *J. Am. Soc. Mass Spectrom. 10*, 329−338.
35. Brush, C. K., Stone, M. P., and Harris, T. M. (1988) *J. Am. Chem. Soc. 110*, 4405−4408.
36. Gait, M. J. (1984) *Oligonucleotide synthesis. A practical approach*, IRL Press, Washington, DC.
37. Kays, A. R., and Schepartz, A. (2000) *Chem. Biol. 7*, 601−610.
38. Cox, J. M., Hayward, M. M., Sanchez, J. F., Gegnas, L. D., van der Zee, S., Dennis, J. H., Sigler, P. B., and Schepartz, A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 13475−13480.
39. Romier, C., Dominguez, R., Lahm, A., Dahl, O., and Suck, D. (1998) *Proteins 32*, 414−424.
40. Hahn, S., Buratowski, S., Sharp, P. A., and Guarente, L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5718−5722.
41. Han, G. W., Kopka, M. L., Cascio, D., Grzeskowiak, K., and Dickerson, R. E. (1997) *J. Mol. Biol. 269*, 811−826.
42. Becker, M. M., and Wang, Z. (1989) *J. Mol. Biol. 210*, 429−438.
43. Wang, S. Y. (1960) *Nature 188*, 844−846.
44. Haseltine, W. A., Gordon, L. K., Lindan, C. P., Grafstrom, R. H., Shaper, N. L., and Grossman, L. (1980) *Nature 285*, 634−641.
45. Zhai, X., Beckmann, H., Jantzen, H. M., and Essigmann, J. M. (1998) *Biochemistry 37*, 16307−16315.
46. Treiber, D. K., Zhai, X., Jantzen, H. M., and Essigmann, J. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5672−5676.
47. Vichi, P., Coin, F., Renaud, J. P., Vermeulen, W., Hoeijmakers, J. H., Moras, D., and Egly, J. M. (1997) *EMBO J. 16*, 7444−7456.
48. Harley, V. R., Jackson, D. I., Hextall, P. J., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P. N. (1992) *Science 255*, 453−456.
49. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−714.
50. Johns, E. W. (1982) *The HMG Chromosomal Proteins*, Academic Press, London.
51. Lilley, D. M. (1992) *Nature 357*, 282−283.
52. Hughes, E. N., Engelsberg, B. N., and Billings, P. C. (1992) *J. Biol. Chem. 267*, 13520−13527.
53. Pil, P. M., and Lippard, S. J. (1992) *Science 256*, 234−237.
54. Pasheva, E. A., Pashev, I. G., and Favre, A. (1998) *J. Biol. Chem. 273*, 24730−24736.

BI0111552