Fluorescent and Photochemical Properties of a Single Zinc Finger Conjugated to a Fluorescent DNA-Binding Probe†

Martin Thompson and Neal W. Woodbury*

Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604

Received August 16, 1999; Revised Manuscript Received November 17, 1999

ABSTRACT: A single zinc finger derived from the DNA-binding domain of the glucocorticoid receptor (GR) has been tethered to the intercalating fluorophore thiazole orange, and the DNA recognition characteristics of the conjugate have been examined. DNA sequence specificity for the peptide−dye conjugate, determined by steady-state fluorescence measurements and photoactivated DNA cleavage experiments, reproduce the binding features of response element recognition found in the native GR. The thiazole orange is able to intercalate and fluoresce when the conjugate binds, at concentrations where little fluorescence is observed from either the conjugate alone or the conjugate mixed with DNA lacking the zinc finger target sequence. The conjugate preferentially targets a 5′-TGTTCT-3′ sequence (the native glucocorticoid receptor element) with a dissociation constant of about 25 nM. Lower binding affinities (up to 10-fold) are observed for single site variants of this sequence, and much lower affinity (40−50-fold) is observed for binding to the estrogen response element (which differs from the glucocorticoid receptor element at two positions) as well as to nonspecific DNA. Footprinting reactions show a 4−6 base pair region that is protected by the zinc finger moiety. Photocleavage assays reveal a several base pair region flanking the recognition sequence where the tethered thiazole orange moiety is able to intercalate and subsequently cleave DNA upon visible light exposure. Thiazole orange is also shown to oxidize the 5′-G of remote GG sequences, depending on the details of the intervening DNA sequence. Small synthetic protein−dye conjugates such as this one are potentially useful for a variety of purposes including sequence-specific probes that work under physiological conditions (without melting and hybridization of DNA), sequence-specific photocleavage agents, and self-assembling components in electron and energy transfer systems that utilize DNA as a scaffold and/or photochemical medium.

Steroid hormone receptors play an important regulatory role in activating transcription by binding DNA (*1*, *2*). There is growing interest in the design of molecular mimics of the DNA-binding regions of such regulatory proteins (*3*, *4*), which, by binding responsive loci, might affect a specific genetic function. Both the study and the utility of such specific DNA-binding proteins would be enhanced if the binding protein were able to report its binding state and some information about its environment. In addition, it would be useful if controlled activities, such as DNA cleavage, could be built into these systems.

DNA-binding proteins contain domains with small structural motifs designed to discriminate specific DNA sequences through an ensemble of noncovalent interactions (*5*). Positioning of the α-helix in the major groove of DNA is a common motif found in both nonsequence-specific (*6*, *7*) and sequence-specific (*8*, *9*) DNA-binding proteins and yields a defined arrangement of noncovalent contacts between amino acid side chains of the protein and DNA bases and backbone structures. Typically, in protein structural motifs such as the helix-turn-helix (*10*−*13*), zinc-finger (*14*−*17*), or leucine zipper (*18*, *19*), the smallest DNA recognition element is an α-helix which binds to the DNA major groove. The remainder of the protein typically contains structural motifs, which orient this recognition element within the major groove, provide nonspecific DNA-binding affinity, and enable additional interactions such as dimerization or ligand binding.

Single zinc fingers have been studied for their ability to fold independently and bind native recognition sequences of the wild-type protein (*20*−*23*). Most proteins that recognize particular DNA sequences via zinc finger motifs have multiple zinc fingers in the DNA-binding domain (*24*). The contribution to sequence recognition of each zinc finger within the DNA binding domain of a protein can vary significantly (*25*). For example, in the glucocorticoid and estrogen receptor DNA-binding domains, the N-terminal finger contains all the interactions for sequence specificity, whereas the C-terminal zinc finger further stabilizes DNA binding by interacting nonspecifically. Finger-swapping studies, in which the N- and C-terminal zinc fingers of the estrogen receptor (ER)[1] and glucocorticoid receptor (GR)

† This work was supported by a donation from Phase Laser Systems Incorporated, Scottsdale, AZ.

* To whom correspondence should be addressed. Phone: (602) 965-3294. Fax: (602) 965-2747. E-mail: NWoodbury@asu.edu.

[1] Abbreviations: TO, thiazole orange; ZF, zinc finger; TOZF, thiazole orange-zinc finger conjugate; GR, glucocorticoid receptor protein; GR-DBD, glucocorticoid receptor protein-DNA-binding domain; GRE, glucocorticoid response element; ER, estrogen receptor protein; ERE, estrogen response element; dsDNA, double-stranded deoxyribonucleic acid.

10.1021/bi991907g CCC: $19.00
Published on Web 03/23/2000

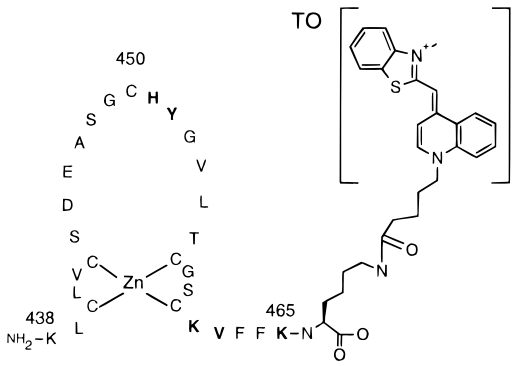


FIGURE 1: Schematic representation of the thiazole orange-zinc finger (TO−ZF) conjugate. The 29 amino acid single zinc finger tethered to the dye was derived from residues 438−465 of the native glucocorticoid receptor protein (GR). The residues in bold make direct contacts with the native glucocorticoid response element (GRE) in the wild-type GR (*38*).

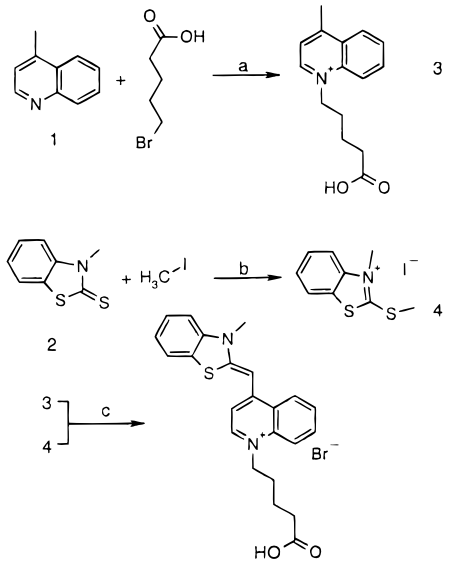
were switched, clearly show that sequence recognition was a function of the N-terminal zinc finger (*26*).

This has made it possible to produce short (less than 50 amino acid) single zinc finger peptides that bind to specific DNA sequences with moderately high affinity (*27*). Thus, it is possible to synthetically generate peptides that bind under normal physiological conditions to specific DNA sequences (*28*−*33*), opening the door for the production of a variety of different types of chemical reagents, which probe or modify DNA at specific sequences. One particularly interesting possibility is to combine the sequence specific binding properties of these oligopeptides with the DNA-specific fluorescence properties of intercalating dyes. There are a number of intercalating dyes which fluoresce almost exclusively when intercalated into DNA (*34*, *35*). By coupling these dyes to site-specific DNA-binding oligopeptides, a series of DNA stains could be produced which interact with dsDNA in its native state and only generate significant fluorescence when bound to the target sequence.

Another potentially attractive feature of such systems is the ability of a number of dye molecules to cleave DNA in the presence of light (*36*, *37*). By proper selection of peptide and dye, it may be possible to produce mixtures of photocleavers, where individual cleavage agents were activated by specific wavelengths of light. This could be useful in the automation of DNA mapping and analysis.

To further characterize such protein−dye conjugates and their interactions with DNA, we have covalently linked the amino-terminal zinc finger of the glucocorticoid receptor DNA-binding domain (GR-DBD) with the intercalating cyanine dye thiazole orange (Figure 1). GR is a protein containing two nonputative zinc finger motifs in its DNA-binding domain (*38*, *39*). The protein has been thoroughly characterized, both biochemically and structurally, and its DNA binding interactions are known in atomic detail (*38*). The single zinc finger used in these experiments is derived from residues 438−466 of the native GR. This short sequence was modeled after the hexameric half-site derived from the native glucocorticoid response element (GRE). The protein conjugate between the zinc finger and thiazole orange should bind sequence specifically to the major groove of the DNA, forcing the dye to intercalate in a defined region adjacent to the zinc finger. NMR studies of the binding of free thiazole orange to DNA show that the dye forms base-stacking interactions with the purine and pyrimidine bases of DNA (*40*−*42*). Thiazole orange was chosen to couple to the zinc finger, because the dye has a large emission enhancement upon dsDNA intercalation, binds DNA with moderate affinity ($K_d \approx 10^{-5}$ M), and is relatively sequence neutral in its DNA recognition (*34*, *43*).

Scheme 1: Conditions: (a) Dioxane, Reflux; (b) Absolute Ethanol, Reflux; (c) Absolute Ethanol, 2 equiv of Triethylamine



In addition to its established role as a nucleic acid stain (*44*−*46*), thiazole orange has been shown to photocleave dsDNA (*37*). This should make it possible to assay the specificity and yield of protein-directed photocleavage in this system. The ability of the cyanine dye intercalators to cleave DNA upon photoactivation furthermore provides a convenient assay for determining the preferred binding site of the designed complex.

MATERIALS AND METHODS

Materials. Lepidine, 5-bromovaleric acid, 3-methylbenzothiazole-2-thione, iodomethane, anhydrous ethanol, and triethylamine were purchased from Aldrich and used without further purification. F-moc amino acids and peptide synthesis reagents were purchased from Advanced Chemtech. The F-moc-Lys (Mtt)-OH was purchased from Anaspec. T4 polynucleotide kinase was from Promega, and [γ-^{32}P]dATP was purchased from Amersham Pharmacia Biotech.

Instrumentation. ^{1}H NMR spectra were recorded on a Varian 300 MHz spectrometer. Peptides were synthesized on a Millipore 9050 peptide synthesizer. Purification of synthetic peptides was performed on an HPLC. Absorbance measurements were performed on a Cary V spectrophotometer. Steady-state fluorescence polarization measurements were performed on PTI QuantaMaster QM-1. DNA sequencing gels were run on a Bio-Rad Sequi-Gen GT Sequencing Cell and analyzed on a Molecular Dynamics Storm-840 gel scanner.

Synthesis of 1-(4-Carboxybutyl)-4-(3-methyl-2,3-dihydro-(benzo-1,3-thiazole)-2-methylidene)-quinolinium Bromide (*Scheme 1*). The synthesis of the carboxylic acid derivative of thiazole orange was adapted from published procedures

(*47*, *48*). A total of 5.7 g (32 mmol) of 5-bromovaleric acid and 1.5 mL (11 mmol) of lepidine was refluxed in 30 mL of distilled dioxane under nitrogen overnight. The reaction mixture was allowed to cool to room temperature and filtered. The precipitate was washed twice with 10 mL of dioxane followed by 50 mL of petroleum ether to yield 2.6 g (70%) of pure 1-(4-carboxybutyl)-4-methyl-1,4-dihydro-quinolinium bromide.

A total of 2.0 g (11 mmol) of 3-methylbenzothiazole-2-thione and 3.5 mL (55 mmol) of iodomethane was refluxed in 50 mL of absolute ethanol for 4 h. Diethyl ether was added to the cooled mixture to precipitate the product. Recrystallization from ethanol:diethyl ether gave 3.2 g (90%) of pure 2-methylmercapto-3-methylbenzothiazole iodide.

A total of 1.0 g (3 mmol) of 1-(4-carboxybutyl)-4-methyl-1,4-dihydro-quinolinium bromide and 1.0 g (3 mmol) of 2-methylmercapto-3-methylbenzothiazole iodide were solubilized in 30 mL of absolute ethanol with slight heating for 5 min. To this solution, 840 μL (6 mmol) of triethylamine was added causing the reaction mixture to immediately turn a deep red color. Heating was discontinued, and the reaction was stirred for 60 min. Product was precipitated by addition of 150 mL of diethyl ether. Recrystallization from acetone: diethyl ether gave 1.4 g (80% yield) of pure 1-(4-carboxybutyl)-4-[3-methyl-2,3-dihydro-(benzo-1,3-thiazole)-2-methylidene]-quinolinium bromide.

Product characterization for each reaction was performed by thin-layer chromatography using EtOAc:AcOH:H_2O (1:2:2, v/v) as the solvent, MALDI-TOF mass spectrometry, and ^{1}H NMR. MALDI-TOF *m*/*z* 391.8 (391.5, calculated for $C_{23}H_{23}N_2O_2S^+$).

Peptide/*Conjugate Synthesis, Purification, and Characterization*. The peptides were synthesized on PAL-PEG-PS resin by automated solid-phase peptide synthesis using F-moc chemistry (*49*). The thiazole orange-labeled and unlabeled peptides were synthesized individually on the solid support. In the case of the thiazole orange-labeled peptide, the first amino acid coupled to the resin was an F-moc-Lys (Mtt)-OH. The Mtt protective group is selectively removed from the ϵ-amine of the lysine by 1% trifluoroacetic acid (TFA) in dichloromethane (*50*). 1-(4-Carboxybutyl)-4-[3-methyl-2,3-dihydro-(benzo-1,3-thiazole)-2-methylidene]-quinolinium was coupled to the ϵ-amine of the lysine using *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′,-tetramethyluronium hexafluorophosphate (HATU) activated coupling chemistry. After dye coupling, peptide synthesis proceeded under standard conditions. The peptides were removed from the resin and deprotected by TFA cleavage methods for 4 h. Thiazole orange is stable to standard deprotection and cleavage conditions.

The peptides were purified by RP-HPLC on a Zorbax C_8 column (9.4 mm × 25 cm) using a water (0.1% TFA) to acetonitrile (0.1% TFA) gradient. Identities of the peptides were confirmed by amino acid analysis and matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF). Quantitation of free thiols was performed with 5,5′-dithiobis-(2-nitrobenzoic acid) using L-cysteine as a reference (*51*). Peptide quantitation was performed using the Bradford Assay with a BSA reference (*52*).

Fluorescence Measurements. Steady-state fluorescence anisotropy measurements were carried out in 20 mM Tris-Cl (pH 7.9), 30 mM NaCl, 20 mM KCl, 1 mM $MgCl_2$, and 5% glycerol. The thiazole orange zinc finger conjugate (TO−ZF) or the zinc finger peptide (ZF) samples were prepared fresh by preincubating with $ZnCl_2$ on ice for 30 min. DNA samples used were either the nonlabeled sequences shown in Figure 2 or a fluorescein-labeled duplex with a sequence matching GRE2 (Figure 2).

```
                1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
GRE1: 5'-TCATACCACTAACTGTTCTATCA-3'
GRE2: 5'-TGATACGGCTGACTGTTCTATGA-3'
GRE3: 5'-TCATACATCTAACTATTCTATCA-3'
GRE4: 5'-TCATACATCTAACTGTCCTATCA-3'
ERE   5'-TCATACATCTAACTGACCTATCA-3'
```

FIGURE 2: The sequence of the five oligomers (23 base pairs each) used throughout these experiments. The native GRE target sequence is underlined. Base changes in the hexameric GRE half-site are indicated by an asterisk. All oligomers used in the photocleavage assays were 5′ end labeled with [γ-^{32}P]dATP. A sequence identical to GRE2 was also synthesized with a fluorescein molecule covalently linked to the 5′ end and was only used for steady-state fluorescence measurements. All labeled strands were annealed with a 2-fold excess of the complementary strand.

For nonlabeled DNA titrations (measuring the fluorescence from thiazole orange), the excitation wavelength was 510 nm and the emission was recorded at 535 nm with a 1.0 nm slit width. Glass cuvettes were used with a 0.3 cm path length. Titrations were performed using a constant concentration of unconjugated thiazole orange (TO) or TO−ZF and titrated with DNA as indicated.

For fluorescein-labeled DNA measurements, the excitation wavelength was 491 nm and the emission was recorded at 521 nm with a 1.0 nm slit width. Titrations were performed holding the concentration of fluorescein-labeled DNA constant and increasing the concentration of TO, ZF, or TO−ZF as indicated. To correct for the small amount of background fluorescence from TO in these measurements, control experiments were performed in parallel using DNA that was not fluorescein labeled, but was otherwise identical to the labeled DNA fragments. The thiazole orange fluorescence intensity measured in the control experiments was subtracted from the total fluorescence intensity for each data point. Under the excitation conditions of these measurements, the contribution to fluorescence by TO at 521 nm was less than 5% of the maximum fluorescence signal observed in the experiment. Even at the point in the titration where the maximum quenching of fluorescein fluorescence was observed (the point of lowest total fluorescence in the measurement), less than half of the fluorescence was due to TO, and this could accurately be removed by subtraction of the control value. Each point is an average of 60 measurements, and each titration was performed independently three times.

Anisotropy is calculated as $A = (I_{vv} - gI_{vh})/(I_{vv} + 2gI_{vh})$ (*53*, *54*), where I_{vv} is the fluorescence intensity with vertically polarized excitation and vertically polarized emission, I_{vh} is the fluorescence intensity with vertically polarized excitation and horizontally polarized emission, and $g(I_{hv}/I_{hh})$ is a factor correcting for the polarization dependence of the spectrometer. A two-state binding model was used for the analysis of zinc finger binding to F-GRE2. The data were fitted to eq 1 using a nonlinear least-squares algorithm. Here, $[GRE]_T$ and $[ZF]_T$ are the total concentrations as 50 nM F-GRE2 was titrated with the zinc finger. The parameters of the fit were the anisotropies of the free and bound DNA (A_F and A_B) and the dissociation constant (K_d).

$$A = A_f + (A_b - A_f)\left[\frac{[\mathrm{GRE}]_T + [\mathrm{ZF}]_T + K_d - \sqrt{([\mathrm{ZF}]_T + k_d + [\mathrm{GRE}]_T)^2 + 4[\mathrm{GRE}]_T[\mathrm{ZF}]_T}}{2[\mathrm{GRE}]_T}\right] \quad (1)$$

Total fluorescence data were collected in parallel to the anisotropy measurements and were calculated from raw data as the fluorescence intensity in the vertical plane plus two times the fluorescence intensity of the horizontal plane, (TF $= I_{vv} + 2I_{vh}$). These values were used to determine the percent of the maximum fluorescence for each point along the titration and fit to eq 2.

$$f = \frac{[\mathrm{TO(ZF)}]_T + [\mathrm{GRE}]_T + K_d - \sqrt{([\mathrm{GRE}]_T + K_d - [\mathrm{TO(ZF)}]_T)^2 + 4[\mathrm{TO(ZF)}]_T K_d}}{2[\mathrm{TO(ZF)}]_T} \quad (2)$$

In eq 2, f corresponds to the fraction of TO−ZF conjugate bound at a given concentration of DNA. The concentration given as $[\mathrm{TO(ZF)}]_T$ corresponds to either the unconjugated thiazole orange or the TOZF conjugate as indicated and $[\mathrm{GRE}]_T$ is the total concentration of GRE-binding sites. The parameter of the fit was the dissociation constant (K_d).

DNA Synthesis, Footprinting, and Photocleavage Experiments. All oligonucleotides were synthesized on a DNA synthesizer using phosphoramidite chemistry (*55*) and purified on a 15% denaturing polyacrylamide gel. Bands were cut out and recovered by soaking the crushed gel pieces in 100 mM Tris-Cl (pH 8.0) and 100 mM NaCl buffer for 4−6 h. Extractions were combined and the urea and salts removed by spin filtration. Oligos were 5′-^{32}P-end-labeled by reaction with [γ-^{32}P]dATP and T4 polynucleotide kinase at 37 °C for 4 h. Samples were heat denatured and spin filtered to remove unincorporated [γ-^{32}P]dATP. Each 5′-^{32}P-end-labeled oligo was annealed with a 2-fold excess of its unlabeled complement. The three 23-base GRE containing oligonucleotides used in this study are shown in Figure 2 (GRE1, GRE2, and GRE3).

The protocol for radical hydroxyl footprinting reactions was adapted from previously established methods (*56*, *57*). Reaction samples comprised of the TO−ZF and 5′-^{32}P-end-labeled GRE2 in 20 mM Tris-Cl (pH 7.9), 40 mM NaCl, 20 mM KCl, and 1 mM $MgCl_2$ were allowed to equilibrate in the dark at room temperature for 60 min. Equal volumes of a freshly prepared solution of 0.2 mM $[Fe(EDTA)]^{2-}$ was mixed with a 0.3% H_2O_2 solution and a 10 mM sodium ascorbate solution and added to the reaction sample. After a 1 min incubation, the reactions were quenched with an aqueous solution of 20 mM thiourea and 10 μg/mL tRNA. The footprinting reactions were analyzed on a 15% denaturing polyacrylamide gel.

The purified DNA strands shown in Figure 2 were 5′-^{32}P-end-labeled using T4 polynucleotide kinase and purified as described above. Oxygen removal was performed by bubbling argon through the samples. Irradiations were performed on 100 μL samples in a 1.0 mL glass cuvette fit with a septum using an argon-ion laser at 514 nm with a power of 300 μW/cm^2. The sample concentrations and irradiation times used are indicated in each figure caption. After irradiation, samples were divided in half, and one portion was treated with piperidine and the other immediately frozen. Piperidine treatment consisted of adding an equal volume of freshly prepared 20% piperidine in water to the sample followed by heating at 90 °C for 30 min. Both the treated and untreated fractions were dried under vacuum. To each dried sample was added 20 μL of 80% formamide in water. The products of the photocleavage experiments were analyzed on a 15% polyacrylamide (19:1, acrylamide: N,N′-methylene-bis-acrylamide) vertical sequencing gel with dimensions 21 cm × 40 cm × 0.4 mm. Maxam−Gilbert A+G and C+T sequencing reactions were used as a reference (*58*). Gels were run at 2000 V for 1.5−2 h. Gels were dried and exposed to a phosphorimaging screen for 4 h and analyzed on a Molecular Dynamics Storm-840 gel scanner in phosphorimaging mode.

RESULTS

Characterization of Thiazole Orange/Zinc Finger Conjugate. The purified 29 amino acid TO−ZF thiazole orange peptide conjugate was characterized using matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry. The expected mass of TO−ZF is 3501.8 and a sharp peak in its mass spectrum is observed for the singly charged species at (m/e) = 3501.9. By comparison, the expected mass of the zinc finger itself is 3154.2 and a sharp peak in its spectrum is observed for the singly charged species at (m/e) = 3155.6.

The conjugate was further characterized by reversed-phase HPLC and UV−vis absorption. An HPLC analysis of purified TO−ZF shows only one peak in the HPLC chromatograph, which confirms the purity of the preparation. The UV−vis absorbance spectrum of HPLC purified TO−ZF displays the additive spectral properties of thiazole orange and the isolated peptide.

The UV−vis spectrum of aqueous solutions of TO−ZF in the presence and absence of dsDNA is compared to that of the unconjugated thiazole orange molecule in Figure 3. The TO−ZF conjugate has two nearly equal absorbance maxima at 482 and 504 nm when coordinated with zinc or cadmium. In comparison, the absorbance maximum of unconjugated thiazole orange is at 503 nm with a much less prominent shoulder at 480 nm. After incubating the unconjugated thiazole orange with DNA for 30 min, the absorbance maxima redshifts from 503 to 510 nm. The redshift is due to thiazole orange intercalating into dsDNA. Similarly, the thiazole orange moiety of the TO−ZF conjugate shows a redshifted absorbance maxima as well when incubated with dsDNA. In this case, the longer wavelength of the two absorbance transitions becomes more prominent and shifts from 504 to 511 nm.

Fluorescence Changes upon Binding of TO, ZF, and TO-ZF to GRE Containing Duplex DNA. Equilibrium binding of the zinc finger alone (ZF), unconjugated thiazole orange (TO), and thiazole orange-zinc finger conjugate (TO−ZF) to fluorescein-labeled GRE2 sequence dsDNA was measured by both fluorescence anisotropy and total fluorescence. For fluorescein fluorescence, excitation was at 491 nm, and emission was detected at 521 nm. Each peptide sample was titrated into a constant concentration of fluorescein-labeled DNA. Changes in the steady-state anisotropy of the fluorescein fluorescence from the labeled GRE sequence containing DNA (F-GRE2) can arise from two sources. Binding of the zinc finger to the DNA slows the rotation time of the DNA and thus decreases the depolarization in the steady state (increasing the anisotropy). Binding of the TO dye to the DNA can also decrease the excited-state lifetime of the

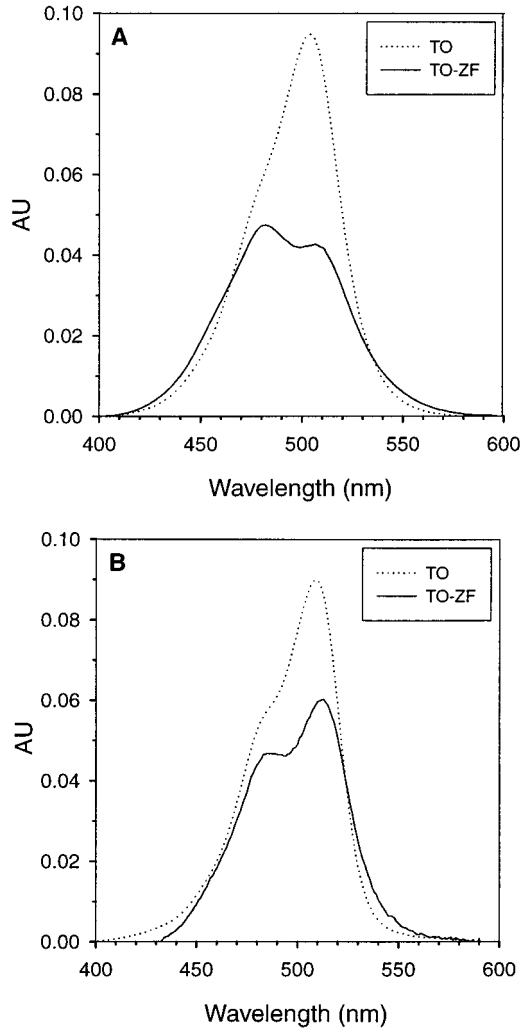


FIGURE 3: Absorbance spectra for the thiazole orange-zinc finger conjugate (TO−ZF) and unconjugated thiazole orange (TO) in 20 mM Tris-Cl (pH 7.9), 30 mM NaCl, 20 mM KCl, 1 mM $MgCl_2$, and 5% glycerol buffer in the absence (A) and presence (B) of 20 μM GRE1 sequence (see Figure 2). Sample concentrations were 2 μM for TO−ZF and TO. TO−ZF was incubated 30 min with 2.5 μM $ZnCl_2$ prior to mixing with DNA. The TO and TO−ZF samples were incubated for 30 min with 20 μM GRE1 sequence. All spectra were corrected for dilutions made upon DNA addition.

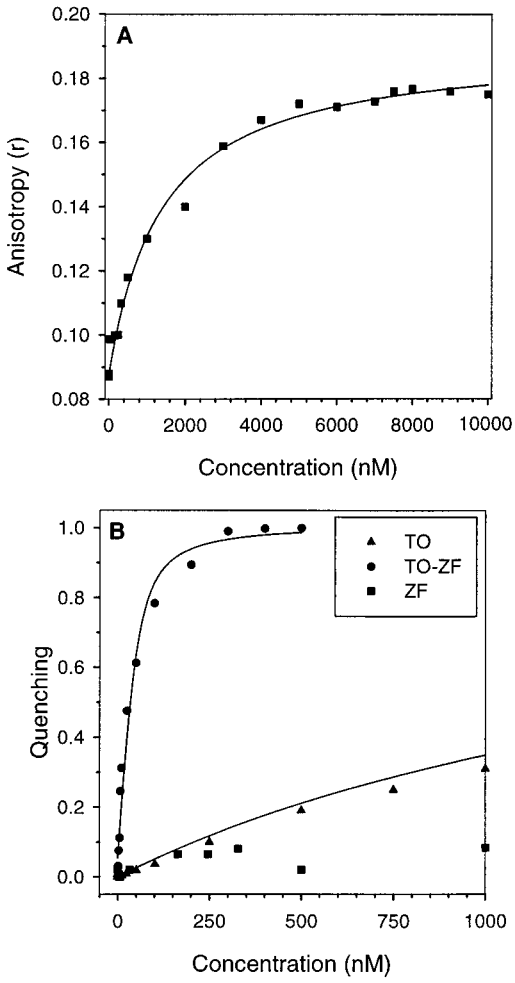


FIGURE 4: (A) Fluorescence anisotropy measurement of 50 nM fluorescein-labeled DNA titrated with the unconjugated zinc finger (ZF) binding protein (no TO attached). (B) Total fluorescence quenching [calculated as $[(F_{max} - F_i)/F_{max}]$, where F_{max} is the fluorescence in the absence of titrant, and F_i is the measured fluorescence during the titration. Fluorescein-labeled DNA was titrated with unconjugated thiazole orange [TO (▲)], the zinc finger alone [ZF (■)] or the peptide-dye conjugate [TO−ZF (●)]. For these experiments, a fluorescein-labeled sequence identical to GRE2 shown in Figure 2 (F-GRE2) was used at either 50 nM (TO and ZF titrations) or 1 nM (TO−ZF titrations). The solid lines represent a fit to eq 1 for the anisotropy of the zinc finger (A) and to eq 2 for the total fluorescence measurements (B).

fluorescein due to energy transfer from fluorescein to TO. This again decreases the steady-state depolarization of the dye (because the molecule does not rotate as much within the shorter excited-state lifetime) and therefore increases the anisotropy. Figure 4A shows the change in steady-state anisotropy upon binding of the sequence-specific zinc finger to the GRE sequence containing dsDNA. The equilibrium constant determined for ZF (not conjugated to TO) from a fit to eq 1 is 1.1 μM. By comparison, the 25 amino acid zinc finger Xfin-31 was determined to have a dissociation constant of approximately 1 μM (*59*), whereas, the 66 amino acid zinc finger from GATA-1 has a dissociation constant of 10 nM (*60*). Similar anisotropy measurements were also taken for TO and the TO−ZF conjugate, but energy transfer between the fluorescein attached to the DNA and the TO (see below) complicates the interpretation of the anisotropy ratio and these data are not presented.

Total fluorescence data, performed in parallel with the anisotropy measurements, are shown in Figure 4B for ZF as well as for TO and the TO−ZF conjugate. Essentially no change is seen in the total fluorescence over the course of the titration, when the ZF itself binds to the fluorescein-labeled GRE2 DNA. In contrast, both the TO−ZF conjugate and the TO alone result in large decreases in steady-state fluorescence. The decrease in fluorescence collected from fluorescein upon addition of TO or TO−ZF is due to energy transfer from fluorescein to thiazole orange (*61*). Contributions to the fluorescence signal at 521 nm by thiazole orange were only a few percent of the maximum fluorescence detected and were corrected for by performing control runs with unlabeled DNA as described in the Materials and Methods. In the case of the TO−ZF conjugate, the expected distance between the fluorescein and the intercalated thiazole orange is in the range of 20−30 Å. This distance range is short enough to allow efficient through-space energy transfer.

Fitting of the quenching curve (Figure 4B) to a two-state binding model (see Materials and Methods) resulted in a

dissociation constant for TO−ZF binding to the target GRE2 dsDNA-binding sequence of 14 nM. The dissociation constant for the unconjugated TO was found to be 2 orders of magnitude higher at 2.4 μM. However, one must be careful in interpreting the quenching curve for unconjugated TO directly in terms of a binding constant, since TO binds to many different sites on the DNA and thus its position relative to the fluorescein is distributed, causing the fluorescence energy transfer efficiency to be heterogeneous in the sample. The TO-binding constant can be estimated from binding curves in which the TO fluorescence is monitored directly, as described below.

Titrations of the TO fluorescence upon binding of TO−ZF to unlabeled DNA were measured for the dsDNA sequences shown in Figure 2. TO was excited at 510 nm and fluorescence was collected at 535 nm. The GRE1 DNA sequence contains the same binding site as the GRE2 DNA sequence used above; however, the GRE3 sequence contains a single mutation of the guanine in the native GRE target sequence (5′-TGTTCT-3′) to an adenine. The GRE2 and GRE3 sequences were generated for use in the photocleavage measurements described below. Finally, the GRE4 and ERE recognition sequences are from the nonspecific GRE half-site and the estrogen response element (ERE), respectively (*76*). These are used in conjunction with the other sequences to determine the sensitivity of this probe. The DNA is not fluorescently labeled in these measurements. Therefore, the fluorescence is exclusively from thiazole orange molecules intercalated into dsDNA.

Figure 5A shows the effect of altering the DNA sequence on the binding of TO−ZF. Fitting the binding curves to a two-state binding model (see Materials and Methods, eq 2) results in a dissociation constant of 25 nM for GRE1 and GRE2, which contain the native GRE. Titrations of the modified GRE sequences and the ERE sequence yielded dissociation constants of 57, 220, and 810 nM for the TO−ZF conjugate bound to GRE3, GRE4, and GRE3, and ERE, respectively.

As expected, the GRE1 and GRE2 sequences, which have the same target sequence, have essentially identical dissociation constants, in good agreement with the value determined for TO−ZF binding to the GRE2 sequence by fluorescein quenching above. However, as shown in Figure 5A, the TO−ZF conjugate binds a dsDNA sequence containing the native GRE target sequence approximately 30-fold more efficiently than the ERE target sequence (which contains two base changes relative to the GRE). More interestingly, the two different single base modifications within the native GRE sequence, GRE3 and GRE4, result in a 2-fold and a 10-fold increase in the dissociation constant relative to the native GRE, respectively. It should be noted that the modification in GRE3 is at the guanine where a base contact by the peptide was intentionally removed (used as the site of TO attachment) in the conjugate used here. Thus, this base change is expected to have only a small effect on binding, as observed. The base change in the GRE4 sequence, which is in the middle of the binding site for the zinc finger, shows a much larger effect on the dissociation constant (Figure 5A). The GRE4 results indicate that the zinc finger is able to dictate the binding specificity of the TO−ZF conjugate well enough to discriminate between binding sites with single base changes, at least in the more sensitive region of the sequence.

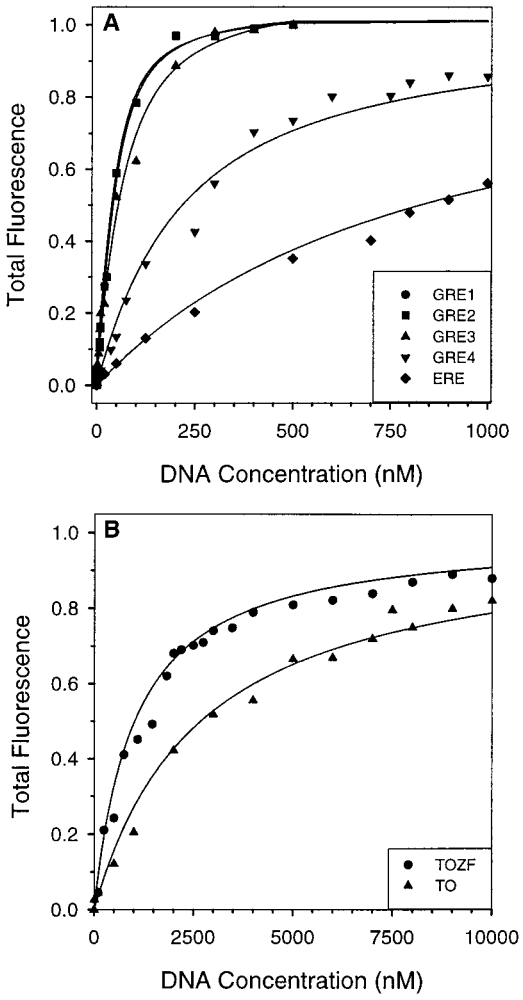


FIGURE 5: Comparison of total fluorescence (which is calculated as the sum of the parallel and 2 times the perpendicular fluorescence intensities) due to sequence specific and nonsequence specific binding of TO−ZF to various oligonucleotide sequences. Fluorescence measurements are given as a fraction of total fluorescence (F_i/F_{max}). (A) Total fluorescence from 1 nM TO−ZF when titrated with increasing concentrations of either GRE1 (●), GRE2 (■), GRE3 (▲), GRE4 (▼), or ERE (◆) dsDNA sequences (Figure 2). (B) Total fluorescence measurements of 50 nM TO−ZF (●) or 50 nM unconjugated TO (▲) titrated with dsDNA [poly(dG)-poly(dC)] not containing the GRE target sequence. For all data shown above, the solid lines represent a best fit to eq 2.

The overall specificity of TO−ZF binding to dsDNA containing the native GRE target sequence was investigated by comparing the data of Figure 5A to a titration of the TO−ZF conjugate with a dsDNA sequence that did not contain any target GRE sequence. Figure 5B compares the binding of unconjugated thiazole orange and TO−ZF conjugate to a poly(dG)-poly(dC) 20-mer. The contributions of the zinc finger toward nonspecific binding give the TO−ZF conjugate slightly enhanced binding over the unconjugated thiazole orange. The dissociation constants for the unconjugated thiazole orange and TO−ZF conjugate are 2.4 and 1.1 μM, respectively. Previous studies of unconjugated thiazole orange, using a 5-fold higher salt concentration determined the dissociation constant for binding to dsDNA to be approximately 3 μM (*62*). Parallel titrations with $(dGdC)_{20}$ show, within experimental error, the same results as the poly(dG)-poly(dC) 20-mer (data not shown). The dissociation

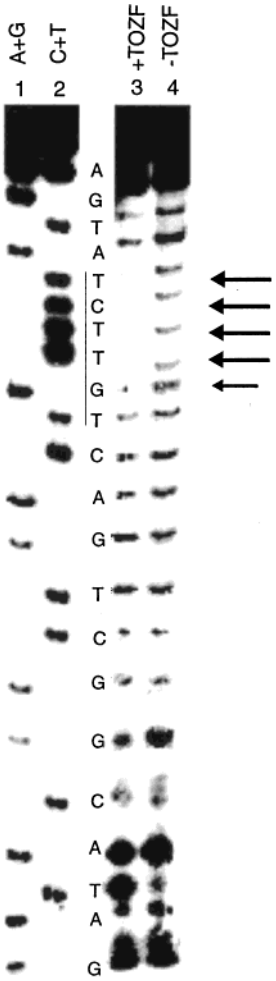


FIGURE 6: Radical hydroxyl footprint of the 5′-^{32}P-end-labeled GRE2 23-bp dsDNA bound to TO−ZF. Reaction samples contained 160 nM duplex DNA and 2.2 μM TO−ZF, 20 mM Tris (pH7.9), 40 mM NaCl, 20 mM KCl, and 1 mM $MgCl_2$. The samples were incubated for 60 min in the dark at 25 °C. The radical hydroxyl reaction proceeded for 60 s at 25 °C. Lanes 1 and 2, Maxam−Gilbert A+G and C+T sequencing reactions, respectively; lane 3, radical hydroxyl cleavage of GRE2 in the presence of TO−ZF; lane 4, radical hydroxyl cleavage of GRE2 in the absence of TO−ZF. Arrows to the right of the gel show schematically the radical hydroxyl cleavage protection data for the footprinting reaction of GRE2 with TO−ZF.

constant of unconjugated thiazole orange is essentially independent of the sequence, as expected for a sequence neutral intercalating dye.

Footprinting of TO−ZF Conjugate. DNA footprinting by the TO−ZF conjugate was performed to determine if the TO−ZF conjugate was binding the same sequence as the native glucocorticoid receptor protein DNA-binding domain. Radical hydroxyl footprinting of the TO−ZF conjugate is shown in Figure 6. A schematic representation showing the protection from radical hydroxyl cleavage that the zinc finger moiety provides is also shown in Figure 6 and it is clear that the binding is localized within the native GRE target sequence. The bound TO−ZF complex protects a four base pair region within the native GRE-binding sequence that correlates well with the location of the N-terminal zinc finger of native glucocorticoid receptor protein-DNA-binding domain that is shown in the crystal structure (*38*). A control experiment performed in the absence of TO−ZF shows no protection in this region.

Photocleavage of TO−ZF Conjugate. We examined the photoinduced cleavage of 5′-^{32}P-end-labeled dsDNA fragments containing the dsDNA sequences GRE1, GRE2, and GRE3 in the presence of either unconjugated TO or TO−ZF. Note that the data shown is only of the photocleavage observed in the DNA strands displayed in Figure 2. We have performed similar photocleavage measurements of the complementary strands of the dsDNA duplexes used in each case. However, the sequence dependence of the observed photocleavage positions is apparently the same on both strands. Therefore, we have concentrated on the photocleavage observed in the strands shown (Figure 2), while realizing that cleavage of the complementary strand is in all cases a competing reaction.

After irradiation of TO−ZF bound to dsDNA in Tris buffer solution at 514 nm (514 nm is in the main band of the absorbance spectrum of intercalated thiazole orange as shown in Figure 3), the photoproducts were denatured and separated using denaturing polyacrylamide gel electrophoresis (PAGE). The unconjugated cyanine dyes oxazole yellow and thiazole orange have been shown previously to cause single-stranded breaks in supercoiled DNA upon UV−vis irradiation. Irradiation of randomly intercalated thiazole orange at 514 nm leads to both alkali-dependent and direct cleavage upon visible illumination (data not shown). Piperidine treatment is a common method used to induce strand cleavage at abasic sites or oxidized guanines formed during a photochemical reaction, whereas direct cleavage does not require piperidine treatment (*36*). The mechanism of strand cleavage is not entirely clear for either alkali-dependent or direct cleavage. Like many fluorescent dyes (*63*, *64*), TO is thought to sensitize singlet oxygen formation, but the results of the photocleavage experiments described below suggest this is not the dominant reaction mechanism of photocleavage mediated by TO−ZF.

The thiazole orange molecule can extend up to 14 Å from the peptide backbone, due to the nature of the linkage used, permitting it to extend over a range of 1−4 base pairs on the 5′ side of the GRE. Photocleavage of GRE1 by TO−ZF causes piperidine-dependent cleavage to occur predominantly at the guanine within the GRE (5′-T<u>G</u>TTCT-3′). Figure 7 shows that, in the presence of a nearby guanine, the only observable photochemistry occurring above background levels is at the guanine and is dependent on piperidine treatment after irradiation. Presumably, electron transfer from the guanine to the excited singlet state of TO (possibly involving intermediate hole transfer) is the dominant excited-state decay pathway when a close-by guanine is available. Note that one reaches the same conclusion by analyzing the photoproducts of the complementary strand of the GRE1 dsDNA fragment, which also shows cleavage at nearby guanines with a preference for the GG sequence located seven base pairs from the GRE (data not shown).

Long-range guanine oxidation, particularly at GG sequences, was explored more rigorously using the GRE2 dsDNA sequence (Figure 2). The GRE2 sequence was designed to have a series of guanines both near and at some distance from the GRE target sequence, allowing for the possibility of long-range hole transfer between guanines. Also, a GG site was introduced which is known to have a lower oxidation potential than a single G (*65*). Photocleavage of GRE2 by TO−ZF shows remote guanine oxidation. Figure

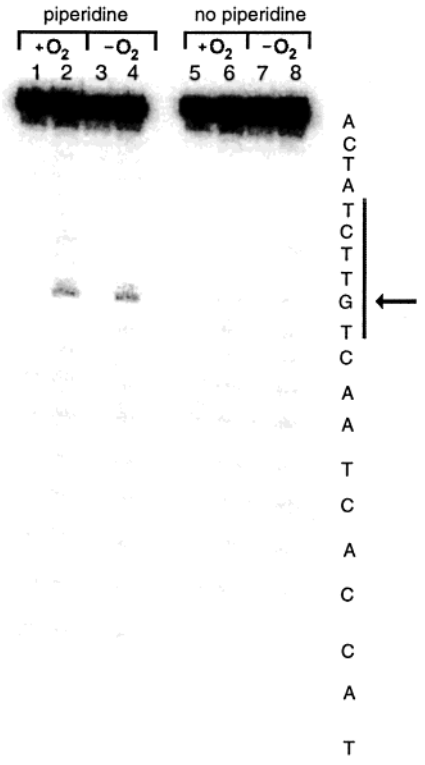


FIGURE 7: Photocleavage of GRE1 by TO−ZF in the presence and absence of oxygen. Reaction samples contained 160 nM duplex DNA and 200 nM TO−ZF in 20 mM Tris, pH 8.0. The samples were incubated for 15 min in the dark at 25 °C and irradiated at 514 nm for 5 min at 300 μW/cm^2 in the presence or absence of oxygen as indicated. Lanes 1, 3, 5, and 7 are irradiated reference samples in the absence of TO−ZF. Lanes 2, 4, 6 and 8 are irradiated samples in the presence of TO−ZF. Samples treated with piperidine are indicated. The GRE target recognition sequence is marked, and the guanine internal to this sequence is denoted with an arrow.

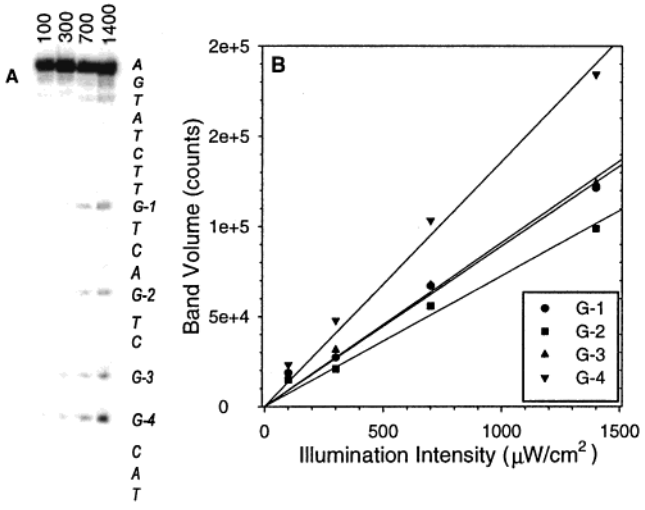


FIGURE 8: Photocleavage of GRE2 by TO−ZF showing the intensity dependence of guanine oxidation. Reaction samples contained 160 nM duplex DNA and 200 nM TO−ZF in 20 mM Tris-Cl, pH 8.0. The samples were incubated for 15 min in the dark at 25 °C and irradiated at 514 nm for 5 min at the indicated wattage. (A) Lanes 1−4 were irradiated at intensities of 100, 300, 700, and 1400 μW/cm^2, respectively. All samples were reacted in the absence of oxygen and treated with piperidine. (B) Plot showing the change in integrated band intensity with laser power for each observable guanine cleavage site.

8 shows that oxidation of remote guanines by excited-state thiazole orange occurs over a range 7−17 Å (17 Å was the farthest G in the DNA sequence used). A roughly linear increase in the product band formation at each guanine is observed as the intensity of the 514 nm argon ion laser was increased from 100 to 1400 μW/cm^2. Preferential cleavage is observed at the 5′-G of the GG sequence.

To explore other possible mechanisms of thiazole orange-mediated photocleavage in the region on the 5′ side of the GRE target sequence, the guanine located within the GRE (5′-TGTTCT-3′) in the GRE1 dsDNA fragment was changed to an adenine to give GRE3 (Figure 2). In the GRE3 dsDNA fragment, there are no Gs present on the strand shown (of course there are Gs on the complementary strand which can and do undergo competing photochemistry, but these are invisible in the results displayed here because the complementary strand was not labeled). Fluorescence measurements of TO−ZF binding to GRE1 and GRE3 target sequences (Figure 5A) show only a modest difference in the dissociation constant between the two sequences. Thus, the measured changes in the photocleavage products between GRE1 and GRE3 are not an artifact of strong differences in the binding affinity of TO−ZF between these sequences. Photoinduced cleavage of the GRE3 sequence by TO−ZF (Figure 9) shows a confined group of bands flanking the altered GRE sequence which are only partially piperidine dependent. The total yield of the piperidine-dependent reactions in these bands, measured as the number of fluorescence counts in product bands divided by the total counts of a given lane, increased 3-fold from 2% in GRE1 (Figure 7, lane 2) to 6% in GRE3 (Figure 9, lane 2) for samples reacted in the presence of oxygen. Piperidine-treated samples reacted in the absence of oxygen showed a 10−20% lower yield.

A comparison of the band patterns observed in the photocleavage products for GRE1 and GRE3 show that when a guanine is nearby, guanine oxidation and cleavage is the only photochemistry observed above background. In the absence of a local guanine, the location of the major bands are 2 and 3 bases from the 5′ end of the GRE under all photocleavage conditions described in Figure 9. Note that, in the GRE3 sequence, the GG pair which is present on the complementary strand in GRE1 is not present and thus cannot act as a trap for hole migration. Independent experiments in which the complementary strand was labeled show that the other Gs present on the complementary strand to GRE3 do act as electron donors and undoubtedly compete with direct photocleavage of the strand shown in Figure 2 (data not shown).

The role of singlet oxygen as a reactive intermediate was probed by using D_2O rather than H_2O as the solvent of parallel samples. Deuterium oxide increases the lifetime of singlet oxygen; and therefore, if singlet oxygen is created by excitation of the dye, its potential to cause damage will increase proportionally (*66*). However, in these experiments, strand cleavage shows that at most there is a 25% increase in band intensity in the presence of D_2O. This coupled with the small (10−20%) effects of removing oxygen imply that singlet oxygen mediated cleavage plays at most a minor role.

DISCUSSION

The fluorescent intercalating dye, thiazole orange (TO), has been tethered to the carboxy-terminal end of a single 29

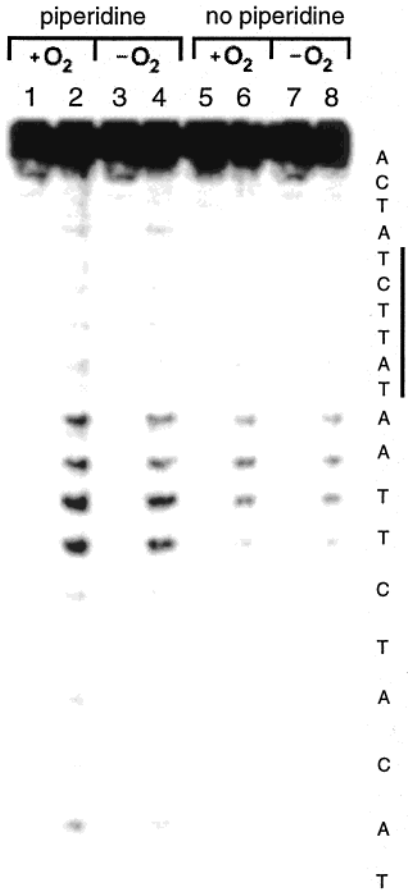


FIGURE 9: Photocleavage of the GRE3 dsDNA sequence by the TO−ZF conjugate. Reaction samples contained 160 nM duplex DNA and 200 nM TO−ZF in 20 mM Tris-Cl, pH 8.0. The samples were incubated for 15 min in the dark at 25 °C and irradiated at 514 nm for 5 min at 300 μW/cm^2 in the presence or absence of oxygen as indicated. Lanes 1, 3, 5 and 7 are irradiated reference samples in the absence of TO−ZF. Lanes 2, 4, 6 and 8 are irradiated samples in the presence of TO−ZF. Samples treated with piperidine are indicated. The altered GRE target recognition sequence is marked.

amino acid zinc finger derived from the DNA-binding domain of the glucocorticoid receptor. The thiazole orange molecule is linked covalently to the ϵ-amino group of the C-terminal lysine facilitating interaction of the dye in a region of 1−4 bases flanking the 5′ side of the GRE. The DNA sequence recognition characteristics of the conjugate have been examined by steady-state fluorescence and photocleavage assays. The sequence specificity of the peptide−dye conjugate reproduces features of response element recognition found in the native glucocorticoid receptor and binds relatively tightly with a dissociation constant of roughly 25 nM. In the bound state of the TO−ZF conjugate, TO associates with the dsDNA and fluoresces, but is essentially nonfluorescent in the unbound form. Thus, the TO−ZF conjugate described in this work represents a DNA sequence-specific fluorophore that can bind under physiological conditions.

The thiazole orange moiety also acts as a photocleavage agent at the binding site, showing direct cleavage in a short region flanking the GRE target sequence. The observed photocleavage pattern compliments data from radical hydroxyl footprinting of the TO−ZF bound to DNA, which shows a four base pair region within the GRE target sequence that is protected by the zinc finger moiety [in agreement with structural work on larger fragments of the binding domain (*67*, *68*)]. In addition to local direct photocleavage reactions, irradiation of the thiazole orange also results in preferential oxidation of the 5′-G of remote GG sequences over 15 Å away if treated with piperidine after irradiation.

Ground-State Absorbance Spectrum of TO Is Sensitive to Its Environment. Fluorescent dyes, such as fluorescein or rhodamine, are known to exhibit subtle changes in their photophysical properties upon changes in solvent, pH, or conjugation (*69*). The thiazole orange molecule exhibits more pronounced spectral changes than most other dyes, owing to two environment-dependent aspects of its absorbance spectrum (Figure 3). First, there is a clear red-shift of both of the transitions (near 480 and 503 nm) upon TO intercalation in dsDNA (*44*, *70*). In addition, the relative oscillator strength of these two transitions depends strongly on environment, which can be seen from the large relative decrease in the oscillator strength of the longer wavelength band upon conjugation of TO to the zinc finger. Thiazole orange dimer formation also causes an increase in the shoulder on the blue side of the thiazole orange absorbance band (*71*), and the oscillator strength of thiazole orange at 500 nm is higher in methanol than in aqueous buffer (*72*). This suggests that the relative oscillator strengths of the two bands may be a function of the polarity of the immediate environment. It is possible that this represents two distinct and interconvertible conformations of TO. Multiple spectroscopic forms of TO derivatives bound to DNA have been demonstrated previously (*73*). It may be that association with the protein surface results in even a more pronounced conformational heterogeneity.

The spectrum of the TO−ZF bound to the GRE sequence results in an absorbance spectrum much like unconjugated TO bound to DNA. This suggests that the thiazole orange moiety of the TO−ZF conjugate bound to the GRE sequence is experiencing an environment similar to that experienced by unconjugated, intercalated TO bound to DNA.

Binding of TO, ZF, and TO-ZF to dsDNA. Equilibrium binding was assessed by a combination of fluorescence anisotropy and total fluorescence measurements of TO, the zinc finger, and the TO−ZF conjugate. The TO−ZF conjugate showed a nearly 50-fold decrease in the dissociation constant relative to the zinc finger alone, due to the additional binding strength provided by the TO.

The fact that the dissociation constant of the TO−ZF is larger than the product of the individual dissociation constants from the zinc finger and the thiazole orange suggests the binding of TO and ZF in the TO−ZF conjugate is noncooperative in some way. It is likely that the conjoined binding events act noncooperatively, at least in part, because both the dye and the peptide individually alter the local DNA structure. For example, the dye has been shown to untwist the DNA approximately 20° (*42*), which may cause a decrease in the ability of the peptide to stay bound if the site of intercalation is close enough to perturb the noncovalent interactions between the peptide and the DNA. The peptide, on the other hand, alters multiple helical parameters in the region where it interacts (*74*). In addition, the peptide undoubtedly constrains the positions and orientations available for TO binding to the DNA. In fact, recent single molecule measurements of the TO−ZF complex with DNA

suggest that there may be multiple binding configurations of the tethered TO molecule (*75*).

The TO conjugated to ZF also proved an efficient energy-transfer acceptor from fluorescein attached to the end of the dsDNA fragment. As one might expect, energy transfer to TO−ZF was more efficient than energy transfer to randomly bound TO. The greater ability of the TO−ZF conjugate to quench fluorescein fluorescence over unconjugated thiazole orange is probably due to both distance and binding effects. The bound TO−ZF places the dye 25−35Å away from the fluorescein, whereas the unconjugated thiazole orange can intercalate anywhere over the entire dsDNA fragment, resulting in substantial distance heterogeneity. In addition, the higher dissociation constant of TO ($\sim 10^{-5}$ M) compared to TO−ZF decreases its effectiveness in energy transfer as well.

TO−ZF Conjugate Retains Binding Specificity. While covalent attachment of thiazole orange to the zinc finger has been shown to greatly increase the DNA-binding affinity, it does not adversely affect the sequence specificity. This system exhibits a relative ratio of specific vs nonspecific binding that is similar to the native glucocorticoid receptor protein−DNA-binding domain (GR-DBD). The native GR-DBD shows a 42-fold increase in binding affinity for the GRE sequence compared to calf-thymus DNA at salt concentrations similar to that used here (*76*). In general, agreement with such studies, the TO−ZF shows almost a 50-fold greater binding to dsDNA containing the GRE target sequence than to a [poly dG-poly dC] dsDNA sequence.

The binding affinity and specificity of TO−ZF is only moderately impaired by single-site replacement of a guanine near the 5′ end of the native GRE sequence (GRE1, Figure 2) with an adenine (GRE3, Figure 2; the dissociation constant for TO−ZF binding to GRE3 is about 2-fold greater than binding to GRE1 as shown in Figure 5). The insensitivity of the TO−ZF binding to a change at this position in the GRE recognition sequence is expected. In the wild-type zinc-finger, an arginine makes a base-specific contact with the guanine of the GRE toward the 5′ end of the recognition sequence (*77*). In the minimal structure used here, this arginine is replaced with the thiazole orange-derivatized lysine, thereby removing this base contact. In contrast, the single-site replacement of a thymine with a cytosine (GRE4, Figure 2) located toward the 3′ end of the GRE recognition sequence results in a 10-fold increase in the dissociation constant of TO−ZF (Figure 5). The reason for this is clearly illustrated in the crystal structure of the native DNA-binding domain, which shows the amino terminal zinc finger interacting toward the 3′ end of the recognition sequence (*78*). The ERE sequence (Figure 2), which differs from the native GRE recognition sequence in the two central base pairs, shows a roughly 30-fold increase in dissociation constant. In fact, the loss of native base contacts at these central two bases results in a dissociation constant that is comparable to nonspecific binding (Figure 5).

The results of the binding studies with different target sequences described above are qualitatively consistent with previous binding studies using a peptide containing both of the zinc fingers in the native DNA-binding domain of the GR (*69*). In that work, 4- and 16-fold increases in the dissociation constants were observed (relative to the native GRE recognition sequence) when the modified recognition sequences represented by GRE4 and ERE, respectively, were used. In fact, the specificity of the TO−ZF conjugate for the native GRE recognition sequence appears to be somewhat enhanced over the original peptide employing two zinc-fingers (one of which is involved primarily in nonspecific binding) (*68*).

Mechanisms of Photocleavage Mediated by TO-ZF. In the absence of a local guanine, direct cleavage at the site of intercalation appears to be the dominant photocleavage mechanism. Thus, by using the GRE3 sequence, it was possible to determine that the location of the thiazole orange moiety of the TOZF conjugate bound to the GRE target sequence in dsDNA was within a four base pair region flanking the 5′ side of the recognition sequence. Figure 9 shows cleavage occurring in the 1−4 base pair region next to the GRE where the bound peptide permits the thiazole orange moiety to interact with the DNA. The intensity distribution observed in the gel bands suggests that the most stable site of intercalation is between the second and third base pair from the GRE. This distance is far enough so that untwisting of the DNA by dye intercalation (*42*) probably has a reduced effect on local DNA structure in the GRE recognition sequence. Intercalation at a position closer to the GRE may perturb the local DNA structure, thus altering critical noncovalent contacts between the peptide side chains and the DNA bases and backbone structures. Intercalation of thiazole orange between base pairs further from the GRE is progressively less likely due to distance constraints imposed by the linker between the thiazole orange and zinc finger moieties.

In general, the mechanisms by which photocleavage occurs when thiazole orange associated with dsDNA is exposed to visible light are still unclear. On the basis of photocleavage experiments performed here (Figures 7−9), strand scission is a composite of several competing reaction mechanisms, including guanine oxidation followed by piperidine dependent cleavage and a direct photochemical reaction mechanism. Singlet oxygen appears to play only a minor role in the observed cleavage. The prevalence of singlet oxygen-mediated strand cleavage using this family of cyanine dyes has been shown to be dependent upon the mode of DNA binding. The external binding mode (binding to the major groove) results in oxygen-dependent photocleavage, which can be eliminated in the presence of scavengers, such as β-mercaptoethanol or by removing the oxygen with argon (*37*). As illustrated in Figures 7 and 9 both the effect of deuterium oxide, which should increase the lifetime of singlet oxygen from 4 μs to approximately 56 μs (*66*), and the effect of removing oxygen are minor. In comparison, other dyes, such as methylene blue, exhibit a D_2O effect resulting in a greater than 2-fold increase in the yield of DNA damage (*79*).

The thiazole orange moiety of the TO−ZF interacts with the π-electron system of the DNA base stack when it is intercalated into dsDNA. Because the thiazole orange is intercalated into a 1−4 base region 5′ to the GRE, the initial photochemistry between the dye and the DNA occurs in this region. This can be followed by long-range oxidation of remote guanines through the base stack (probably via a series of short electron-transfer reactions) (*80*), if guanines are present close enough to the TO and close enough to each other to promote hole migration. Enhanced cleavage of the 5′-G of 5′-GG-3′ sequences are signatures for charge migration through the DNA base stack (*79*). Figure 8 shows

that a radical cation introduced near the 5′ side of the GRE sequence can oxidize a guanine more than 15 Å away. Studies performed on other intercalators, such as anthraquinone (*81*−*83*) or ethidium bromide (*79*), suggest that this process is initiated by electron transfer from a DNA base to the excited state of the dye. This process is followed by localization and reaction of the radical cation (hole) at a reactive G (*81*, *82*). The remote guanine oxidation shown in Figure 8 must occur rapidly relative to the direct cleavage mechanism or mechanisms involving generation of diffusible species, such as singlet oxygen. A comparison of the photocleavage of the GRE1 sequence (Figure 7) and the GRE3 sequence (Figure 9) clearly shows an increase in the yield of photoproducts formed upon removal of the guanine. Consistent with previous studies, guanine oxidation may prevent accessibility to other reaction mechanisms simply because it is the faster reaction (*84*). The lower photocleavage yield of the GRE1 sequence, compared to the GRE2 sequence, may better be explained by previous studies that speculate that charge transfer to a single guanine can be reversible, whereas charge transfer to a double guanine site (GG) cannot (*80*). In the case of cleavage at distant guanines, the migration of a hole or an electron through the base stack of duplex DNA from its point of generation to a site of reaction has been used to explain several experimental observations (*80*, *85*). In fact, for the study of long-range G oxidation, this may be a more versatile system than the use of intercalators tethered to the DNA backbone as described by Barton and co-workers (*86*, *87*). Protein conjugates such as this one can be bound anywhere within a piece of DNA simply by introducing the GRE target sequence, and the range of dye positions is similar to that found for DNA−dye conjugates (Barton's work). Thus, they could be introduced into plasmids or even genomes by mutagenesis, without extensive synthetic manipulation of DNA−dye conjugates.

Clearly there are several possible mechanisms of strand cleavage by TO−ZF. It is possible that, within the lifetime of the DNA-bound zinc finger, the thiazole orange goes in and out of the base stack more than once. This could give rise to the sort of heterogeneity in fluorescence flux and dynamics that has been observed in single molecule measurements of the TO−ZF conjugate interacting with DNA (*75*).

The fact that photoinduced cleavage by the thiazole orange moiety occurs in a region 5′ to the GRE is in agreement with the footprinting data showing the location of peptide binding to be within the GRE target sequence. Crystal structure data and molecular dynamics simulations of the glucocorticoid receptor protein−DNA-binding domain with the recognition sequence show the N-terminal zinc finger positioned toward the 3′ end of the GRE (*38*, *78*). The footprint does not show protection in the flanking region because the intercalated thiazole orange does not provide protection from radical hydroxyl molecules attacking the deoxyribose of the DNA backbone.

Summary. The probe−peptide conjugate studied here binds to DNA with high affinity ($K_d \approx 10^{-8}$ M) and retains nativelike sequence specificity. A wide range of probe−peptide conjugates can be prepared using probes with fluorescent or photocleavage properties and DNA binding peptides that differ in sequence selectivity. The tethering of a sequence specific binding peptide onto a functional probe provides a viable strategy for construction of synthetic enzymes, repressors, and dynamic studies of protein−DNA interactions.

ACKNOWLEDGMENT

The authors are much indebted to Dr. Linda Bloom for the use of her fluorescence polarization instrument and discussions on anisotropy, Dr. Paul Liddel for assistance with dye synthesis, Dr. Dan Brune for assistance with the peptide synthesizer, and Dr. Scott Bingham for preparing the oligonucleotides.

REFERENCES

1. Beato, M. (1989) *Cell 56*, 335−344.
2. Evans, R. M. (1988) *Science 240*, 889−895.
3. Dervan, P. B. (1986) *Science 232*, 464−471.
4. Sardesai, N. Y., Zimmermann, K., and Barton, J. K. (1994) *J.Am. Chem. Soc. 116*, 7502−7508.
5. Pabo, C. O., and Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
6. Doherty, A. J., Serpell, L. C., and Ponting, C. R. (1996) *Nucleic Acids Res. 24*, 2488−2497.
7. Revzin, A., Ed. (1990) *The Biology of Non-Specific DNA-Protein Interactions*. CRC Press, Boca Raton, Florida.
8. Spolar, R. S., and Record, M. T. (1994) *Science 362*, 777−784.
9. Mack, D. P., Sluka, J. P., Shin, J. A., Griffin, J. H., Simon, M. I., and Dervan, P. B. (1990) *Biochemistry 29*, 6561−6567.
10. Brennan, R. G., and Matthews, B. W. (1989) *J. Biol. Chem. 264*, 1903−1906.
11. Zuiderweg, E. R. P., Billeter, M., Boelens, R., Scheek, R. M., Wuthrich, K., and Kaptein, R. (1984) *FEBS 174*, 243−247.
12. Zuiderweg, E. R. P., Kaptein, R., and Wuthrich, K. (1983) *Proc.Natl. Acad. Sci. U.S.A. 80*, 5837−5841.
13. Harrison, S. C., and Aggarwal, A. K. (1990) *Annu. Rev. Biochem. 59*, 933−969.
14. Klug, A., and Rhodes, D. (1987) *Trends Biochem. Sci. 12*, 464−469.
15. Berg, J. M. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 405−421.
16. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
17. Coleman, J. E. (1992) *Annu. Rev. Biochem. 61*, 897−946.
18. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science 254*, 539−544.
19. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science 243*, 538−542.
20. Bernstein, B. E., Hoffman, R. C., Horvath, S., Herriott, J. R., and Klevit, R. E. (1994) *Biochemistry 33*, 4460−4470.
21. Frankel, A. D., Berg, J. M., and Pabo, C. O. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4841−4845.
22. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L., and Berg, J. M. (1991) *J. Am. Chem. Soc. 113*, 4518−4523.
23. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) *Science 245*, 635−637.
24. Miller, J., McLachlan, A. D., and Klug, A. (1985) *EMBO J. 4*, 1609−1614.
25. Whyatt, D. J., deBoer, E., and Grosveld, F. (1993) *EMBO J. 12*, 4993−5005.
26. Green, S., and Chambon, R. (1987) *Nature 325*, 75−78.
27. Archer, T. K., Hager, G. L., and Omichinski, J. G. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7560−7564.
28. Sardesai, N. Y., and Barton, J. K. (1997) *J. Biol. Inorg. Chem. 2*, 762−771.
29. Oakley, M. G., Turnbull, K. D., and Dervan, P. B. (1994) *Bioconjugate Chem. 5*, 242−247.
30. Shin, J. A., Ebright, R. H., and Dervan, P. B. (1991) *Nucleic Acids Res. 19*, 5233−5236.
31. Graham, K. S., and Dervan, P. B. (1990) *J. Biol. Chem. 265*, 16534−16540.

32. Oakley, M. G., and Dervan, P. B. (1990) *Science 248*, 847−850.
33. Sluka, J. P., Horvath, S. J., Bruist, M. F., Simon, M. I., and Dervan, P. B. (1987) *Science 238*, 1129−1132.
34. Haugland, R. P. (1996) in *Handbook of Fluorescent Probes and Research Chemicals* (Spence, M. T. Z., Ed.) 6th ed., pp 144−156, Molecular Probes, Inc., Eugene, OR.
35. Glazer, A. N., and Rye, H. S. (1992) *Nature 359*, 859−861.
36. Armitage, B. (1998) *Chem. Rev. 98*, 1171−1200.
37. Akerman, B., and Tuite, E. (1996) *Nucleic Acids Res. 24*, 1080−1090.
38. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., and Yamamoto, K. R., Sigler, P. B. (1991) *Nature 352*, 497−505.
39. Tilborg, M. A. A. v., Bonvin, A. M. J. J., Hard, K., Davis, A. L., Maler, B., Boelens, R., Yamamoto, K. R., and Kaptein, R. (1995) *J. Mol. Biol. 247*, 689−700.
40. Jacobsen, J. P., Pedersen, J. B., Hansen, L. F., and Wemmer, D. E. (1995) *Nucleic Acids Res. 23*, 753−760.
41. Hansen, L. F., Jensen, L. K., and Jacobsen, J. P. (1996) *Nucleic Acids Res. 24*, 859−867.
42. Spielmann, H., Wemmer, D., and Jacobsen, J. (1995) *Biochemistry 34*, 8542−8553.
43. Netzel, T. L., Nafisi, K., Zhao, M., Lenhard, J. R., and Johnson, I. (1995) *J. Phys. Chem. 99*, 17936−17947.
44. Lee, L., Chen, C.-H., and Chiu, L. (1986) *Cytometry 7*, 508−517.
45. Hirons, G. T., Fawcett, J. J., and Crissman, H. A. (1994) *Cytometry 15*, 129−140.
46. Rye, H. S., and Glazer, A. N. (1995) *Nucleic Acids Res. 23*, 1215−1222.
47. Brooker, L. G. S., Keyes, G. H., and Williams, W. W. (1942) *J. Am. Chem. Soc. 64*, 199−210.
48. Rye, H., Yue, S., Wemmer, D., Quesada, M., Haugland, R., Mathies, R., and Glazer, A. (1992) *Nucleic Acids Res. 20*, 2803−2812.
49. Kent, S. B. H. (1988) *Annu. Rev. Biochem. 57*, 957−989.
50. Aletras, A., Barlos, K., Gatos, D., Koutsogianni, S., and Mamos, P. (1995) *Int. J. Peptide Protein Res. 45*, 488−496.
51. Ellman, G. L. (1996) *Arch. Biochem. Biophys. 82*, 70−77.
52. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
53. Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996) *Mol. Endocrinol. 10*, 607−612.
54. Carlsson, C., Larsson, A., Bjorkman, M., Jonsson, M., and Albinsson, B. (1997) *Biopolymers 41*, 481−494.
55. Caruthers, M. H., Barone, A. D., Beaucage, S. L., Dodds, D. R., Fisher, E. F., McBride, L. J., Matteucci, M., Stabinsky, Z., and Tang, J. Y., Eds. (1987) *Chemical Synthesis of Deoxyoligonucleotides by the Phosphoramidite Method. Methods in Enzymology*, Academic Press, San Diego.
56. Revzin, A., Ed. (1993) *Footprinting of Nucleic Acid-Protein Complexes. Separation, Detection and Characterization of Biological Macromolecules*. Academic Press, San Diego.
57. Nielsen, P. E. (1990) *J. Mol. Recognit. 3*, 1-25.
58. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
59. Lee, M. S., Gottesfeld, J. M., and Wright, P. E. (1991) *FEBS 279*, 289−294.
60. Omichinski, J. G., Trainor, C., Evans, T., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 1676−1680.
61. Förster, T. (1965) in *Modern Quantum Chemistry* (Sinanoglu, O., Ed.) Vol. III, pp 93−137, Academic Press, New York.
62. Nygren, J., Svanvik, N., and Kubista, M. (1998) *Biopolymers 46*, 39−51.
63. Morrison, H., Mohammad, T., and Kurukulasuriya, R. (1997) *Photochem. Photobiol. 66*, 245−252.
64. Tuite, E., and Kelly, J. (1993) *J. Photochem. Photobiol. B-Biol. 21*, 103−124.
65. Sugiyama, H., and Saito, I. (1996) *J. Am. Chem. Soc. 118*, 7063−7068.
66. Rodgers, M. A., and Snowden, P. T. (1982) *J. Am. Chem. Soc. 104*, 5541−5543.
67. Lunback, T., Cairns, C., Gustafsson, J.-A., Carlstedt-Duke, J., and Hard, T. (1993) *Biochemistry 32*, 5074−5082.
68. Lunback, T., Zilliacus, J., Gustafsson, J.-A., Carlstedt-Duke, J., and Hard, T. (1994) *Biochemistry 33*, 5955−5965.
69. Neckers, D. C., and Valdes-Aguilera, O. M., Eds. (1993) Photochemistry of Xanthine Dyes. In *Advances in Photochemistry*, John Wiley & Sons, Inc.
70. Bockstaele, D. R. V., and Peetermans, M. E. (1989) *Cytometry 10*, 214−216.
71. Nygren, J., Andarde, J. M., and Kubista, M. (1996) *Anal. Chem. 68*, 1706−1710.
72. Carlsson, C., Larsson, A., Jonsson, M., Albinsson, B., and Norden, B. (1994) *J. Phys. Chem. 98*, 10313−10321.
73. Milanovich, N., Suh, M., Hayes, J. M., and Small, G. J. (1996) *Biospectroscopy 2*, 125−129.
74. Eriksson, M. A. L., Hard, T., and Nilsson, L. (1995) *Biophys. J. 68*, 402−426.
75. Doug Daniel, personal communication.
76. Hard, T., Dahlman, K., Carlstedt-Duke, J., Gustafsson, J.-A., and Rigler, R. (1990) *Biochemistry 29*, 5358−5364.
77. Eriksson, M. A. L., Nilsson, L. (1995) *J. Mol. Biol. 253*, 453−472.
78. Harris, L. F., Sullivan, M. R., Popken-Harris, P., and Hickok, D. F. (1994) *J. Biomol. Struct. Dyn. 12*, 249−270.
79. Hall, D. B., Kelley, S. O., and Barton, J. K. (1998) *Biochemistry 37*, 15933−15940.
80. Giese, B., Meggers, E., and Michel-Beyerle, M. E. (1998) *J. Am. Chem. Soc. 120*, 12950−12955.
81. Schuster, G. B., and Breslin, D. T. (1996) *J. Am. Chem. Soc. 118*, 2311−2319.
82. Gasper, S. M., and Schuster, G. B. (1997) *J. Am. Chem. Soc. 119*, 12762−12771.
83. Breslin, D. T., Coury, J. E., Anderson, J. R., McFail-Isom, L., Kan, Y., Williams, L. D., Bottomly, L. A., and Schuster, G. B. (1997) *J. Am. Chem. Soc. 119*, 5043−5044.
84. Kelly, S. O., Holmlin, R. E., Stemp, E. D. A., Barton, J. K. (1997) *J. Am. Chem. Soc. 119*, 9861−9870.
85. Schuster, G. B., Armitage, B., Ly, D., Koch, T., Frydenlund, H., Orum, H., Batz, H. G. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12320−12325.
86. Barton, J. K. (1997) *Chem. Biol. 4*, 389−400.
87. Kelley, S. O., Barton, J. K. (1998) *Chem. Biol. 5*, 413−425.

BI991907G