A Type II DNA-Binding Protein Genetically Engineered for Fluorescence Spectroscopy: The "Arm" of Transcription Factor 1 Binds in the DNA Grooves†

Torleif Härd,‡,§ Michael H. Sayre,∥,⊥ E. Peter Geiduschek,∥,⊥ and David R. Kearns*,‡,⊥

Department of Chemistry, Department of Biology, and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0342

Received October 5, 1988; Revised Manuscript Received November 21, 1988

ABSTRACT: We examined the fluorescence properties of a mutant TF1 protein (transcription factor 1; a member of the type II class of DNA-binding proteins, DBPII) containing tryptophan in place of phenylalanine (TF1-W61) at position 61 in the "arms" of the protein dimer. The time-resolved fluorescence (excited at 295 nm) of Trp61 decays as a double exponential with lifetimes and amplitudes that are comparable to those found in other tryptophan-containing proteins and peptides, and the time-resolved fluorescence polarization decay indicates that the tryptophan residue possesses considerable internal flexibility, in agreement with crystal studies of the homologous HU protein. The tryptophan emission is quenched when TF1-W61 binds to DNA, and equilibrium studies based on fluorescence show that the nonspecific binding affinity of the TF1-W61 mutant to DNA is similar to that of wild-type TF1. Comparison of the time-resolved fluorescence decay and steady-state fluorescence intensity reveals at least two general classes of Trp61 in the DNA complexes. One class of tryptophans is partially quenched, and the extent of quenching in the complexes with various natural DNAs and synthetic double-stranded polynucleotides correlates with the spectral overlap between tryptophan emission and DNA absorption, indicating that through-space excitation energy transfer contributes to the observed quenching. Comparisons between experimentally determined energy transfer rates and model calculations suggest that the Trp61 is located in one of the DNA grooves at a distance of less than 7.5 Å from the DNA helix axis. The second class of Trp61 is "totally" quenched, and we attribute this to tryptophan residues that are in direct contact with the DNA bases. These results support an earlier proposal (Tanaka et al., 1984) that the arms of DBPII proteins directly interact with DNA in DNA–DBPII complexes.

The type II DNA-binding proteins (DBPII) are a family of homologous and ubiquitous procaryotic proteins that are thought to be involved in the formation of what is called "bacterial chromatin" (Gualerzi & Pon, 1986). Most of these proteins evidently bind nonspecifically to DNA, but two are known to display at least some selectivity: the integration host factor (IHF) of *Escherichia coli* (Craig & Nash, 1984) and the *Bacillus subtilis* phage SPO1 encoded TF1 (Johnson & Geiduschek, 1977; Greene & Geiduschek, 1985). The structure of the DBPII (or HU protein) from *Bacillus stearothermophilus* has been solved by X-ray crystallography (Tanaka et al., 1984). It was found to be dimeric with a tightly packed core from which two flexible arms extend (each arm contains two strands of antiparallel β-sheet). A structure of the HU–DNA complex, in which the arms serve as DNA-binding domains, was suggested on the basis of this crystal structure. Other DBPIIs are likely to have similar structures and DNA-binding motifs.

In previous work on TF1, we studied the fluorescence properties of the Tyr94 residue located close to its C-terminus (Härd et al., 1989). We found that the fluorescence of this residue is quenched in TF1–DNA complexes and that excited-state energy transfer from tyrosine to DNA contributes to the quenching. We estimated the distance between Tyr94 and the DNA helix axis to be 10–14 Å, by studying the observed quenching as a function of spectral overlap in TF1 complexes with various natural and synthetic DNAs.

In this work we have turned our attention to another part of the TF1 molecule, the already referred to arm. At the tip of this arm, the DBPII share a common sequence, R61-N-P-x-T (x is hydrophilic: Q, R, S, or K). N62, P63, and T65 are universal; R61 is conserved in every DBPII except TF1, in which it is nonconservatively replaced by F. Reasoning that the unique substitution at amino acid 61 of TF1 might be related to its special DNA-binding properties, some of us have analyzed the DNA-binding properties of variant TF1 proteins containing other amino acids in place of F61 (Sayre and Geiduschek, unpublished results). These variations have been generated by in vitro mutagenesis, and the corresponding proteins have been prolifically synthesized in *E. coli* through the use of an appropriate expression system. Realizing the potential utility of Trp as a reporting fluorophore, an F61 → W61 mutation has also been introduced. The experiments reported here deal with the fluorescence–spectroscopic properties of the protein containing a single Trp residue. Binding studies using gel electrophoresis (M. Sayre, unpublished results), as well as the fluorescence titrations that are presented below, show that this substitution has no dramatic quantitative effect on the affinity of TF1 for DNA. The fluorescence of the single tryptophan residue at position 61 can be studied separately from that of Tyr94 by selective excitation at 295 nm. We have used time-resolved and steady-state fluorescence measurements to show that the tryptophan fluorescence is quenched in DNA complexes and that tryptophan-to-DNA excitation energy transfer contributes to the quenching in analogy with our previous studies of the Tyr94 fluorescence.

† This work was supported by grants from the National Science Foundation (to D.R.K.) and the American Cancer Society and the National Institute of General Medical Sciences (to E.P.G.). T.H. acknowledges a postdoctoral fellowship of the Swedish Natural Sciences Research Council, and M.H.S. acknowledges a predoctoral fellowship of the National Science Foundation.

‡ Department of Chemistry.

§ Present address: Department of Medical Biophysics, Karolinska Institute, Box 60400, S-10401 Stockholm, Sweden.

∥ Department of Biology.

⊥ Center for Molecular Genetics.

0006-2960/89/0428-2813$01.50/0 © 1989 American Chemical Society

We use model calculations to estimate the effective distance between the tryptophan (indole) chromophore and the DNA bases and to test the hypothesis that the arms of this DBPII bind in one of the DNA grooves.

Materials and Methods

Materials. The tryptophan-containing variant of TF1 was generated by introducing a $\frac{TC}{AG} \rightarrow \frac{GG}{CC}$ mutation into codon 61 of the wild-type TF1 gene on plasmid mp9TF1, according to the general methods already described and referenced elsewhere (Sayre & Geiduschek, 1988). The F61 → W61 mutant gene was excised from this plasmid as a 524-bp *Hgi*AI–*Eco*RI fragment and resealed into the expression vector pKJB842 as described. The induction of synthesis of TF1-W61 in *E. coli* HB101 (pKJB842-TF1-W61) and the purification of overproduced protein essentially followed procedures that are, or will be, described elsewhere (Härd et al., 1989; Sayre and Geiduschek, unpublished results). The pure TF1-W61 and normal TF1 eluted from their respective final heparin–Sepharose columns at different points of the applied salt gradients. Calf thymus DNA was obtained from Sigma (type I) and purified as described previously (Härd et al., 1989). Double-stranded polynucleotides poly(dAdT), poly(dGdC), poly(dIdC), poly(dIBr^{5}dC), poly(dAdU), and poly(dAI^{5}dU) (Pharmacia) were dissolved directly in the experimental buffer. Polynucleotide concentrations were determined spectrophotometrically with extinction coefficients specified by the supplier. All experiments were carried out at ambient (room) temperature. The same buffer (50 mM NaCl, 10 mM Tris-HCl, and 2 mM Na_2EDTA at pH 7) was used throughout.

Fluorescence Measurements. Time-resolved fluorescence intensity decays were measured on a time-correlated photon counting instrument that has recently been described elsewhere (Skibsted et al., 1987). For excitation of tryptophan we used the 514.5-nm output of an actively mode-locked Ar^+ laser to pump a Rhodamine 6G dye laser tuned to 590 nm. The dye laser beam was passed through a cavity dumper to obtain a pulse repetition rate of 10 MHz and frequency doubled to 295 nm with a KDP crystal. Emission decay profiles (at 355 nm) were collected for parallel and perpendicular polarizations (relative to the polarization of the exciting light) of a Polaroid sheet polarizer. Residual background fluorescence in buffer or DNA/buffer samples was collected separately and subtracted from the protein fluorescence before fitting. The instrument response function (lamp pulse, width at half-maximum 200–300 ps) was collected from the Raman scattering of pure water. The fluorescence decays were deconvoluted with the instrument response function and fitted to theoretical decay functions with a nonlinear least-squares fitting program based on the Marquardt algorithm (Press et al., 1986). The accuracy of fits was judged on the basis of the χ_r^2 criterion and plots of weighted residuals (O'Connor & Phillips, 1984).

The time-resolved decay of the fluorescence polarization anisotropy (FPA), denoted $r(t)$, was calculated as

$$r(t) = \frac{I_{\parallel}(t) - GI_{\perp}(t)}{I_{\parallel}(t) + 2GI_{\perp}(t)} \quad (1)$$

where $I_{\parallel}(t)$ and $I_{\perp}(t)$ are the intensities polarized parallel and perpendicular to the polarization of the exciting beam, respectively, and G is an instrument-dependent constant used to correct for polarizing effects in the emission monochromator and detector (O'Connor & Phillips, 1984).

Steady-state fluorescence was measured on an Aminco SPF-500 spectrofluorometer operated in the ratio mode. The extent of quenching of Trp61 fluorescence in DNA–protein samples, $Q = (I_0 - I)/I_0$, was obtained from measurements of I_0 and I, the observed intensities in the absence and presence of DNA, respectively. The maximum quenching, Q_{max}, the observed quenching in a sample where the protein is quantitatively bound to DNA, was estimated from extrapolation of the observed quenching in "reverse" titrations of protein with DNA at constant protein concentration. The ratio of the fluorescence quantum yields in the free and bound states of the protein, ϕ_f/ϕ_b, is related to Q_{max} through $\phi_f/\phi_b = 1/(1 - Q_{max})$.

Analysis of Excitation Energy Transfer. The rate of excitation energy transfer (Förster, 1948; Lakowicz, 1984), k_{ET}, between an excited donor (e.g., Tyr94 in TF1) and an acceptor in the ground state (a DNA base) is given by

$$k_{ET} = \frac{\kappa^2 \phi_D J}{r^6 n^4 \tau_D} \times 8.71 \times 10^{23}\ s^{-1} \quad (2)$$

where ϕ_D is the quantum yield of the donor in the absence of the acceptor, n is the refractive index of the medium, r is the distance between donor and acceptor (in Å), κ^2 is an orientation factor reflecting the relative orientation of the donor and acceptor transition moments, τ_D is the measured fluorescence lifetime of the donor in the absence of the acceptor, and J is the spectral overlap between donor emission and acceptor absorption (in M^{-1} cm^3). The spectral overlap is defined as

$$J = \frac{\int_0^\infty F_D(\lambda)\epsilon_A(\lambda)\lambda^4\, d\lambda}{\int_0^\infty F_D(\lambda)\, d\lambda} \quad (3)$$

where $\epsilon_A(\lambda)$ is the extinction coefficient of the acceptor (DNA) and $F_D(\lambda)$ is the relative fluorescence intensity of the donor (Trp61 in TF1-W61) at the corresponding wavelength, λ. The critical distance, R_0, for energy transfer is defined as the donor–acceptor separation at which $k_{ET} = 1/\tau_D$ and is usually calculated from eq 2. In a DNA–protein complex several DNA bases might act as excitation energy acceptors, and it is therefore necessary to consider the observed rate of energy transfer as a sum of rate constants for energy transfer to individual bases.

If the structures of TF1-W61 complexes with different DNAs (with different absorption spectra) are similar, then a plot of the observed relative fluorescence quantum yields of free and bound TF1, ϕ_f/ϕ_b, versus the spectral overlap, J (Figure 7), is expected to be linear in J: described by

$$\frac{\phi_f}{\phi_b} = \frac{\Gamma/(\Gamma + k_f)}{\Gamma/(\Gamma + k_b + k_{ET})} = \frac{\Gamma + k_b}{\Gamma + k_f} + \tau_f k_{ET} = \frac{\Gamma + k_b}{\Gamma + k_f} + \sum_i \frac{\kappa_i^2}{r_i^6} \times \frac{\phi_f}{n^4} \times 8.71 \times 10^{23} \times J \quad (4)$$

where $1/\Gamma$ is the natural fluorescence lifetime, τ_f is the observed fluorescence lifetime of free TF1, k_f and k_b are nonradiative decay rates (other than energy transfer), ϕ_f and ϕ_b are (observed) fluorescence quantum yields in free and bound TF1, respectively, and k_{ET} is replaced by a sum of rate constants for energy transfer to all nearby DNA bases, as described above.

Results and Discussion

Fluorescence Properties of Trp61 in TF1. The Trp61 fluorescence in TF1-W61 can be separated from "overlapping" tyrosine (Tyr94) emission by exciting the tryptophan at 295 nm, where tyrosine absorption is negligible (Lakowicz, 1984). The Trp61 emission maximum is located at 355 nm (Figure 1), indicating that the tryptophan residue is exposed to a polar

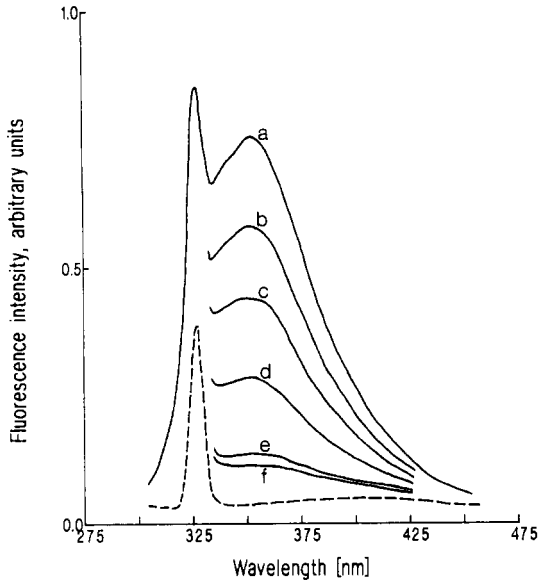

FIGURE 1: Fluorescence emission spectrum of TF1-W61 (0.85 μM dimers) in the absence (a) and presence (b–f) of SPO1 DNA. The DNA concentrations are 1.0 (b), 3.6 (c), 7.6 (d), 14.7 (e), and 23.3 (f) μM DNA bp. The excitation wavelength is 295 nm, and the dashed spectrum is the buffer background intensity with Raman scattering from water at 330 nm.

Table I: Time-Resolved Fluorescence Intensity Decays of TF1-W61[a]

sample	lifetimes (ns)[b]	amplitudes (%)[b]	χ_r^2	Q_{calcd}[f]
free TF1-W61[c]	2.8 ± 0.4	100	4.5	
	1.2 ± 0.2	49 ± 3	1.14	
	4.2 ± 0.2	51 ± 3		
TF1-W61–SPO1 DNA[d]	0.3 ± 0.2	78 ± 3	1.40	0.62
	3.6 ± 1.3	22 ± 3		
	0.2 ± 0.15	68 ± 2	0.97	0.62
	1.2 ± 0.2	19 ± 2		
	5.2 ± 2.0	13 ± 2		
	0.3 ± 0.2	72 ± 1		
	1.2[e]	12 ± 1	1.09	0.63
	4.2[e]	16 ± 1		

[a] Excitation and emission at 295 and 355 nm, respectively. [b] Uncertainties represent estimated standard deviations. [c] 2.0 μM TF1-W61 dimers. [d] 1.6 μM TF1-W61 dimers and 40 μM SPO1 DNA bp. [e] These lifetimes were kept fixed in the deconvolution. [f] $Q_{calcd} = (I_0 - I)/I_0$, where I_0 and I are calculated from measured lifetimes and amplitudes ($\sum \alpha_i \tau_i$) in the absence and presence of DNA, respectively. $Q_{obsd} = 0.80$ under the conditions in footnote *d*.

environment [the tryptophan emission is very sensitive to the solvent polarity (Konev, 1967)], as might have been expected on the basis of the crystal structure of the homologous HU protein where the arms (containing residue 61) protrude out from the protein core. The fluorescence quantum yield of Trp61 is somewhat higher ($\phi_f = 0.18$) than that of L-tryptophan in water at pH 7 [$\phi_f = 0.13$ (Chen, 1967; Lehrer, 1970)] as judged from fluorescence intensities obtained at equal optical densities (data not shown).

The time-resolved fluorescence decay of the Trp61 emission is well fitted by a double exponential (Table I) where the two components have lifetimes of 1.2 and 4.2 ns with approximately equal amplitudes. A biexponential decay with similar lifetimes is typical for tryptophan and tryptophan-containing peptides in water at pH 7 (Robbins et al., 1980; Szabo & Rayner, 1980; Petrich et al., 1983; Creed, 1984), although the amplitude of the longer lifetime component is usually larger. The nonexponential tryptophan fluorescence decays have been attributed to emission from several different rotational conformers (rotamers) about the C^{α}–C^{β} bond (Szabo & Rayner, 1980; Petrich et al., 1983; Philips et al., 1988). It is therefore possible that the distribution of Trp61 rotamers in TF1-W61 differs from that in other tryptophan-containing compounds with the result that the two lifetimes have about equal amplitudes.

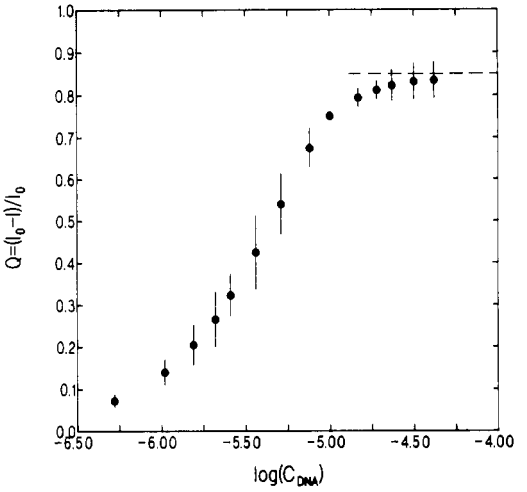

FIGURE 2: Quenching of TF1-W61 fluorescence as a function of DNA concentration for a reverse titration of TF1-W61 (0.8 μM dimers) with SPO1 DNA. Excitation and emission wavelengths are 295 and 355 nm, respectively. The data are corrected for attenuation of the excitation light due to DNA absorption at high DNA concentrations. The DNA concentration (C_{DNA}) refers to molar DNA bp. The dashed line indicates the extrapolated value of $Q_{max} = 0.85$. The experimental uncertainties in the individual data points are indicated by vertical bars.

Nanosecond molecular motions within the arm are evident in the time-resolved FPA decay (Figure 4), which is also composed of two components. A rapid component with a correlation time of <200 ps (as judged from the initial slope) can probably be attributed to local motions of the tryptophan residue and the arm domain, whereas a slower component with a correlation time of 3–4 ns presumably reflects overall rotational diffusion of the protein. The rapid component is responsible for most of the depolarization, indicating that the Trp61 undergoes large-amplitude internal motions. By using the "cone model" (Kinosita et al., 1977) and assuming that $r_0 = 0.35$ and that $r_\infty = 0.1$, an amplitude of ±50° can be calculated for the rapidly decaying component. The FPA decay in Figure 4 therefore supports the notion that the arms of TF1 (and presumably of HU and homologous proteins) possess conformational flexibility (Tanaka et al., 1984).

Quenching of Trp61 Fluorescence in DNA Complexes. Binding of TF1-W61 to SPO1 DNA quenches the tryptophan fluorescence without shifting the location of the emission maximum, and this suggests that the Trp61 residue is still located in a polar environment in the DNA complex. The fluorescence intensity in the bound state is only about 15% of that of the free protein, as judged from the value of Q_{max} obtained by extrapolation of the reverse titration shown in Figure 2. Reverse titrations can also be used to compare the stability of the TF1-W61 complex with SPO1 DNA to that of the wild-type TF1 protein. In Figure 3 we show Q/Q_{max} for reverse titrations of protein with SPO1 DNA at the same protein concentrations (4 μM protein dimers). [The data for wild-type TF1 refer to our previous study of the Tyr94 fluorescence, where we showed that $Q_{max} = 0.75$ for TF1 binding to SPO1 DNA (Härd et al., 1989).] It is obvious that the binding of TF1-W61 to SPO1 DNA is comparable to that of the wild-type TF1 protein.

The lifetimes and amplitudes for the time-resolved fluorescence decay of TF1-W61 in the bound state shown in

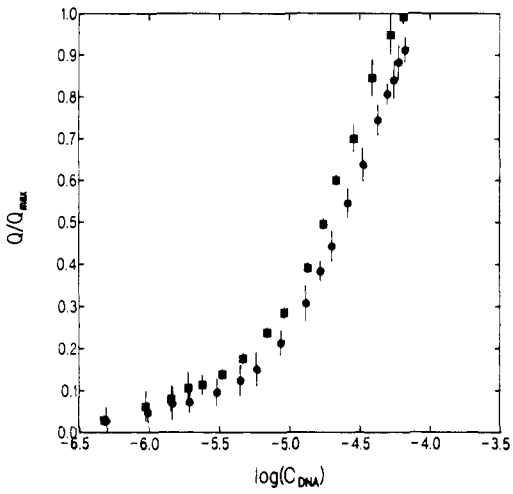

FIGURE 3: Comparison of the binding of TF1 (●) and TF1-W61 (■) to SPO1 DNA. The data represent reverse titrations of DNA on samples with fixed protein concentrations (4 μM dimers). The data for TF1 [from Härd et al. (1989)] represent quenching of the Tyr94 fluorescence (excitation and emission at 280 and 305 nm, respectively). The data for TF1-W61 refer to fluorescence of the Trp61 residue (excitation and emission at 295 and 355 nm, respectively). The observed quenching is "normalized" by division with Q_{max} = 0.75 and Q_{max} = 0.85 for TF1 and TF1-W61, respectively, to facilitate comparisons. The data are corrected for DNA absorption at high DNA concentrations.

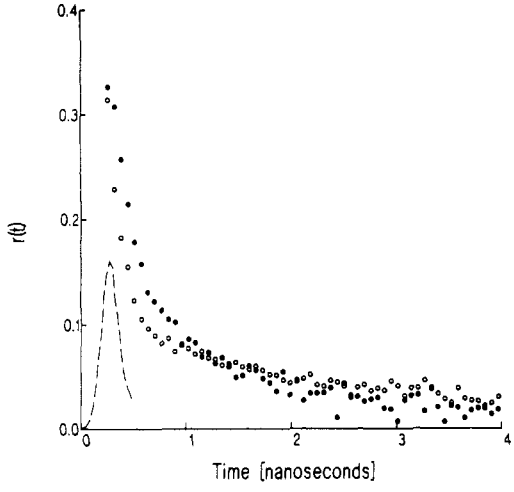

FIGURE 4: Time-resolved FPA decay, $r(t)$, for tryptophan emission in free TF1-W61 [(○) 2 μM dimers] and TF1-W61 bound to SPO1 DNA [(●) 1.6 μM TF1-W61 dimers and 40 μM DNA bp]. Steady-state fluorescence measurements on the same sample of TF1-W61 and DNA indicate that about 97% of the protein is bound to DNA. The dashed line is the instrument response function (lamp pulse).

Table I refer to a TF1-W61/SPO1 DNA sample for which steady-state intensity measurements indicate that Q/Q_{max} = 0.93; i.e., 93% of the protein in the sample is bound to DNA. The time-resolved fluorescence decay of this sample should therefore be dominated by short-lived (quenched) fluorescence with small-amplitude contributions from one or two components with longer lifetimes. However, when fitting the observed decay to both double and triple exponentials, we find that the quenching calculated directly from the measured amplitudes and lifetimes (Q = 0.62, Table I) is considerably less than the quenching ($Q \approx 0.8$) measured by steady-state fluorescence. Furthermore, a fit to a triple exponential decay, where two of the lifetimes are kept fixed to the values observed with the free protein, also yields an unacceptably large fraction of "long-lived" fluorescence (~30%) and a low value of the

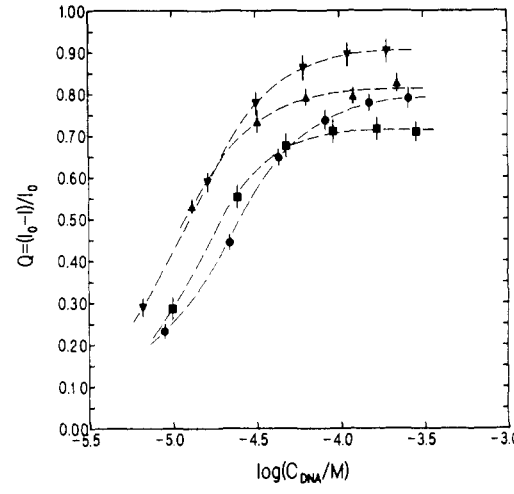

FIGURE 5: Reverse titrations of TF1-W61 (2 μM dimers) for determination of Q_{max} for complexes with some of the various DNAs used in this study. The data refer to titrations with poly(dAI^{5}dU) (▼), poly(dIdC) (▲), CT DNA (●), and poly(dAdT) (■). Experimental conditions and corrections are as in Figure 2.

calculated quenching (Q_{calcd} = 0.63). The apparent discrepancy between the steady-state intensity measurements and the fluorescence decay measurements can be resolved if we allow for the possibility that some tryptophan residues have been totally quenched (within experimental limitations of our lifetime measurements). Agreement between the value of Q = 0.80 found by steady-state fluorescence intensity measurements and the lifetime measurements can then be obtained if we assume that 48% of the tryptophan emission is completely quenched (see Table I), that 35% is shortened to ~0.2 ns, and that some of the remaining fluorescence intensity arises from free TF1. We caution that Q values calculated from the lifetime data are much less accurate than those derived from steady-state fluorescence measurements because the values of the lifetimes and intensities of the partially quenched tryptophan are rather uncertain. Nevertheless, these observations support the notion that in the bound state there are at least two different conformational states of Trp61 (e.g., rotamers): one that is "totally" quenched and another that is only partially quenched. Additional evidence for this interpretation is provided by a study of Trp61 fluorescence in TF1-W61 complexes with different DNAs (vide infra).

Excitation Energy Transfer from Tryptophan to DNA Bases Contributes to the Observed Quenching. In an earlier study of the Tyr94 fluorescence in TF1 we found that the extent of tyrosine fluorescence quenching was correlated with the spectral overlap between tyrosine emission and DNA absorption in various TF1–DNA complexes indicating that excitation energy transfer contributed to the quenching; we used this effect to calculate an effective distance between the Tyr94 residue and the DNA bases (Härd et al., 1989). It has been shown that the critical distance for Förster energy transfer between a tryptophan in the excited state and DNA bases in the ground state is about 5 Å (Montenay-Garestier, 1975). If the arms of TF1-W61 indeed bind in one of the DNA grooves, then we might expect Trp61 to be sufficiently close to DNA bases for noticeable energy transfer to take place.

To examine the extent to which energy transfer contributes to the quenching of Trp61 fluorescence, we measured the maximum quenching, Q_{max}, in TF1-W61 complexes with various DNAs and synthetic polynucleotides with different absorption properties (Figure 5) and also compared the spectral overlap between TF1-W61 emission and DNA absorption for these DNAs (Figure 6). Figure 7 shows a plot of the ratio

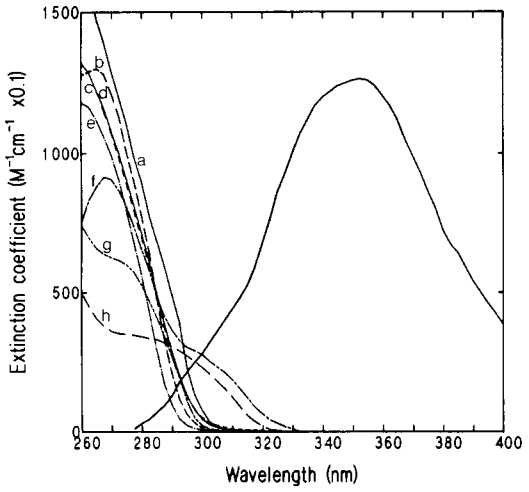

FIGURE 6: Absorption spectra of (a) poly(dGdC), (b) poly(dAdT), (c) SPO1 DNA (dashed line), (d) CT DNA (solid line), (e) poly(dAdU), (f) poly(dAI^{5}dU), (g) poly(dIdC), and (h) poly(dIBr^{5}dC) and emission spectrum of TF1-W61 illustrating the spectral overlap between tryptophan emission and DNA absorption.

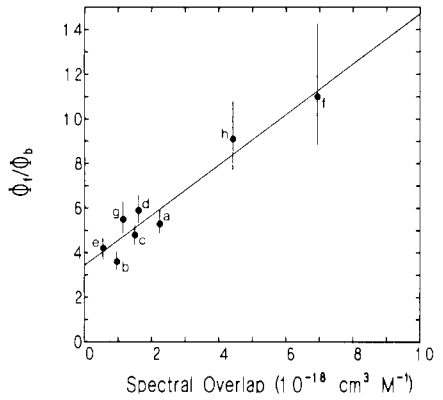

FIGURE 7: Plot of the relative fluorescence quantum yield of free and bound TF1-W61, ϕ_f/ϕ_b, (excitation and emission at 295 and 355 nm, respectively), versus spectral overlap, J, calculated from eq 3. [The relative fluorescence quantum yield is related to Q_{max} as $\phi_f/\phi_b = 1/(1 - Q_{max})$.] The data points are labeled as in Figure 6. Error bars are estimated standard deviations. The solid line represent a least-squares fit of the data to eq 4, yielding a slope of 11.3×10^{17} M cm^{-3} and an intercept at $\phi_f/\phi_b = 3.4$.

between the fluorescence quantum yields in the free and bound state as a function of the spectral overlap calculated from eq 3. It is evident that the fluorescence quantum yield of Trp61 in complexes with various DNAs is correlated with the spectral overlap between the DNA absorption spectrum and the Trp61 fluorescence spectrum and that energy transfer contributes to the quenching. [Note that efficient excitation energy transfer in the opposite direction, i.e., from DNA bases in the excited state to tryptophan in the ground state, is not possible because that fluorescence quantum yield is too low at room temperature [10^{-5}–10^{-4} (Anders, 1981; Aoki & Callis, 1982)] for DNA–tryptophan energy transfer to be noticeable.] The data in Figure 7 also suggest that the distance between the Trp61 fluorophore and the DNA bases is approximately the same in the different TF1-W61–DNA complexes, because the data points seem to lie close to the same straight line. The data in Figure 7 also show that excitation energy transfer is not the only quenching mechanism, because a linear least-squares fit to eq 4 does not extrapolate to $\phi_f/\phi_b = 1$ when $J = 0$. The fitted intercept ($\phi_f/\phi_b = 3.4$) actually indicates that about 70% of the fluorescence is lost by some process that results in total quenching of a fraction of the Trp61 residues. As noted above, we attribute this to Trp61 residues that are in direct contact with the DNA since this is known from earlier work to result in complete quenching of Trp fluorescence (Rajeswari et al., 1987).

Trp61 Is Located in One of the DNA Grooves in the TF1-W61–DNA Complex. A precise determination of the location of the partially quenched tryptophan residue with respect to the DNA is difficult for several reasons. In most cases one does not know the spectral overlaps with individual bases, but only the average for the two chromophores in a base pair. One can also expect the orientation factor, κ^2, to vary for the different bases surrounding the tryptophan residue, although the FPA decay in Figure 4 indicates that the Trp61 is undergoing motions on the nanosecond time scale and therefore that the individual orientation factors are probably close to an isotropic value (2/3) (Lakowicz, 1984). A third potential difficulty is that DNA is not cylindrically symmetric, and therefore, the rate of energy transfer cannot be expected to be the same at all positions at a fixed distance from the helix axis.

However, an upper limit of the distance between Trp61 and DNA bases can be estimated on the basis of simple model calculations. As a first approximation, we consider each base pair as a single "effective" chromophore, treat the DNA as a linear array of chromophores spaced 3.4 Å apart, and assume $\kappa^2 = 2/3$. With this model we can calculate the expected rate of energy transfer from a tryptophan residue at a given distance from the array of chromophores using eq 4 (with $n = 1.35$, $\phi_f = 0.18$, and $\kappa_i = 2/3$). We find that the calculated slope of a plot of ϕ_f/ϕ_b versus J is about 20×10^{17}, 4×10^{17}, and 1×10^{17} M cm^{-3} for distances of 5, 7.5, and 10 Å to the "central" chromophore in the array, respectively. Since the experimental value is 11×10^{17} M cm^{-3} for TF1-W61–DNA complexes (Figure 7), it would seem that the Trp61 residue is located at a distance of about 5–7.5 Å from the DNA base pairs. On the other hand, since our model is very simple, a more cautious interpretation would be that the distance to the closest DNA base is less than 7.5 Å.

A more rigorous calculation of the expected energy transfer in the TF1-W61–poly(dGdC) complex is possible because the spectral overlaps between cytidine and guanosine absorption and tryptophan emission are approximately equal in magnitude (Montenay-Garestier, 1975). In this case we placed the tryptophan on the DNA dyad axis at various distances from the helix axis (in the minor or major groove) or in direct contact with the phoshate backbone, as shown in Figure 8 and calculated the expected rate of energy transfer by summing the rates for transfer to all nearby cytidine and guanosine bases. The results (listed in the caption to Figure 8) should be compared with the observed rate ($k_{ET} = 7.5 \times 10^8$ s^{-1}) calculated from eq 4 and data in Figure 8 and with the assumption that the average fluorescence lifetime $\langle \tau_f \rangle = 2.8$ ns (Table I). A comparison shows that the most likely position of the Trp61 residue is in one of the DNA grooves at a distance from the helix axis that is somewhat less than 7.5 Å [$k_{ET} > (4–6) \times 10^8$ s^{-1}] and that a structure where it is located anywhere "outside" the DNA phosphates is less likely ($k_{ET} \approx 0.6 \times 10^8$ s^{-1}). On basis of these model calculations, we therefore conclude that Trp61 is located in one of the DNA grooves in TF1-W61–DNA complexes.

Summarizing Remarks

Our results demonstrate that in the complex with DNA the Trp61 residue in the arms of TF1-W61 is in close proximity to the DNA. The measurements of the fluorescence decay and of the quenching of the steady-state fluorescence intensity

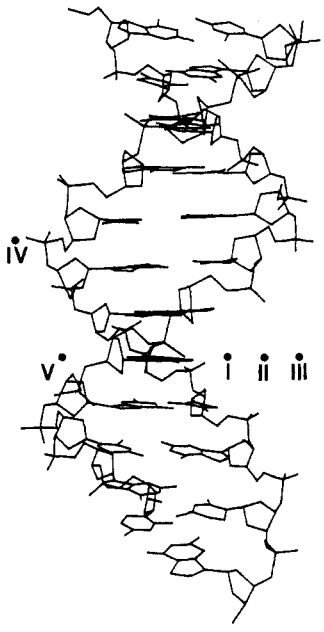

FIGURE 8: Various positions of Trp61 in the TF1-W61–poly(dGdC) complex for which theoretical rates of tryptophan-to-DNA excitation energy transfer were calculated (the DNA is drawn in a perspective view). The points (I–V) represent the center of the tryptophan (indole) chromophore. Points I–IV are located on the DNA pseudodyad axis at distances of 5.0 (I), 7.5 (II and IV), and 10 Å (III) from the helix axis, and point V is located at van der Waals distance from a backbone phosphate group. The calculations were carried out by measuring the distance from a specific position to the center of each surrounding cytidine or guanosine residue in DNA and then calculating k_{ET} as a sum of all individual rates (eq 4). It was assumed that $J = 1 \times 10^{-18}$ cm^3 M^{-1} for energy transfer to cytidine and guanosine, $\phi_f = 0.18$, and $n = 1.35$ and that all $\kappa_i = 2/3$. [The values of the spectral overlaps were based on the observations that the critical distances for tryptophan excitation energy transfer are approximately equal (Montenay-Garestier, 1975) and that the sum of the two spectral overlaps is 2×10^{-18} cm^3 M^{-1} (Figures 6 and 7).] The calculated values of k_{ET} for the different locations of the Trp61 are (I) 23×10^8 s^{-1}, (II) 4×10^8 s^{-1}, (III) 0.5×10^8 s^{-1}, (IV) 0.6×10^8 s^{-1}, and (V) 6×10^8 s^{-1}. These values should be compared to the observed rate, 7.5×10^8 s^{-1}, calculated from eq 4 and data in Figure 7 with $\langle \tau_f \rangle = 2.8$ ns (Table I).

in the SPO1–DNA complexes and the variation of the fluorescence quenching in the different DNA complexes are all consistent with Trp61 in at least two different local environments. In one environment, corresponding to 70% of the overall population, the Trp61 fluorescence is totally quenched in the DNA–TF1-W61 complexes, and we attribute this to direct interaction of Trp61 with the DNA bases. Within this population of totally quenched Trp61 residues, there could be more than one subpopulation, but we are not able to experimentally demonstrate this. For the remainder of the population, the Trp61 quenching that is observed can be understood in terms of through-space energy transfer from Trp61 residues located less than 7.5 Å from the DNA helix axis. These results therefore support the general model proposed for DNA–DBPII complexes in which the arms of the protein bind in one of the DNA grooves (Tanaka et al., 1984). Whether the different local environments are due entirely to different tryptophan rotamer populations or in part to protein–protein interactions in DNA–protein complexes cannot be resolved on the basis of the existing data.

The present study also has more general implications for the use of fluorescence to study DNA–protein complexes, because it shows that excitation energy transfer contains useful structural information (Härd et al., 1989). Fluorescence quenching is frequently observed when proteins and peptides bind to DNA (Brun et al., 1975; Hélène & Maurizot, 1981; Bujalowski & Lohman, 1987; de Jong et al., 1987), and it has been shown that energy transfer should contribute to the observed effect (Montenay-Garestier, 1975). Our studies of tyrosine and tryptophan fluorescence in TF1 show that the extent of quenching can be monitored as a function of spectral overlaps in complexes with different synthetic DNAs and that these data can be used to "measure" the distance between fluorescent amino acids (tyrosine and tryptophan) and DNA bases in DNA–protein complexes.

ACKNOWLEDGMENTS

Victor Hsu is acknowledged for help with computer graphics.

Registry No. L-Trp, 73-22-3.

REFERENCES

Anders, A. (1981) *Chem. Phys. Lett. 81*, 270–272.
Aoki, T. I., & Callis, P. R. (1982) *Chem. Phys. Lett. 92*, 327–332.
Brun, F., Toulme, J. J., & Hélène, C. (1975) *Biochemistry 14*, 558–563.
Bujalowski, W., & Lohman, T. M. (1987) *Biochemistry 26*, 3099–3106.
Chen, R. F. (1967) *Anal. Lett. 1*, 35–42.
Craig, N. L., & Nash, H. A. (1984) *Cell (Cambridge, Mass.) 39*, 707–716.
Creed, D. (1984) *Photochem. Photobiol. 39*, 537–562.
de Jong, E. A. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1987) *Biochemistry 26*, 2039–2046.
Förster, Th. (1948) *Ann. Phys. 2*, 55–79.
Greene, J. R., & Geiduschek, E. P. (1985) *EMBO J. 4*, 1345–1349.
Gualerzi, C. O., & Pon, C. H. (1986) in *Bacterial Chromatin*, Springer-Verlag, Berlin.
Härd, T., Hsu, V., Sayre, M. H., Geiduschek, E. P., Appelt, K., & Kearns, D. R. (1989) *Biochemistry 28*, 396–406.
Hélène, C., & Maurizot, J. C. (1981) *CRC Crit. Rev. Biochem. 10*, 213.
Johnson, G. G., & Geiduschek, E. P. (1977) *Biochemistry 16*, 1473–1485.
Kinosita, K., Kawato, S., & Ikegami, A. (1977) *Biophys. J. 20*, 289–305.
Konev, S. V. (1967) in *Fluorescence and Phosphorescence of Proteins and Nucleic Acids* (Udenfriend, S., Transl. Ed.), pp 1–204, Plenum Press, New York.
Lakowicz, J. R. (1984) in *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
Lehrer, S. S. (1970) *J. Am. Chem. Soc. 92*, 3459–3462.
Montenay-Garestier, T. (1975) *Photochem. Photobiol. 22*, 3–6.
O'Connor, D. V., & Phillips, D. (1984) in *Time-Correlated Single Photon Counting*, pp 180–189, Academic Press, Orlando, FL.
Petrich, J. W., Chang, M. C., McDonald, D. B., & Fleming, G. R. (1983) *J. Am. Chem. Soc. 105*, 3824–3832.
Philips, L. A., Webb, S. P., Martinez, S. J., Fleming, G. R., & Levy, D. H. (1988) *J. Am. Chem. Soc. 110*, 1352–1355.
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1986) in *Numerical Recipes*, Chapter 14, Cambridge University Press, Cambridge, England.
Rajeswari, M. R., Montenay-Garestier, T., & Hélène, C. (1987) *Biochemistry 26*, 6825–6831.
Robbins, R. J., Fleming, G. R., Beddard, G. S., Robinson, G. W., Thistlethwaite, P. J., & Woolfe, G. J. (1980) *J. Am. Chem. Soc. 102*, 6271–6278.

Sayre, M. H., & Geiduschek, E. P. (1988) J. Virol. 62, 3455–3462.
Skibsted, L. H., Hancock, M. P., Magde, D., & Sexton, D. A. (1987) Inorg. Chem. 26, 1708–1712.
Szabo, A. G., & Rayner, D. M. (1980) J. Am. Chem. Soc. 102, 554–563.
Tanaka, I., Appelt, K., Dijk, J., White, S. W., & Wilson, K. S. (1984) Nature (London) 310, 376–381.

Multinuclear NMR Studies of DNA Hairpins. 1. Structure and Dynamics of d(CGCGTTGTTCGCG)†

James R. Williamson‡ and Steven G. Boxer*
Department of Chemistry, Stanford University, Stanford, California 94305
Received September 8, 1988; Revised Manuscript Received December 8, 1988

ABSTRACT: The solution structure of the hairpin formed by d(CGCGTTGTTCGCG) has been examined in detail by a wide variety of NMR techniques. The hairpin was characterized by proton NMR to obtain interproton distances and torsion angle information. An energy-minimized model was constructed that is consistent with these data. The hairpin consists of a B-DNA stem of four C-G base pairs and a loop region consisting of five unpaired bases. Three bases in the 5′ of the loop are stacked over the 3′ end of the stem, and the other two bases in the 3′ of the loop are stacked over the 5′ end of the stem. The phosphorus NMR spectrum revealed a phosphate in the stem region with an unusual conformation, and two phosphates, P_9 and P_{10}, were found to undergo intermediate exchange between conformations. The hairpin was also synthesized with a carbon-13 label in each of the thymidine C6 carbons, and relaxation measurements were performed to determine the extent of internal motions in the loop region. The loop bases are more flexible than the stem bases and exhibit subnanosecond motions with an amplitude corresponding to diffusion in a cone of ~30°.

Hairpins are secondary structural elements for nucleic acids that are found in many structural and functional contexts. Hairpin loops are a particularly important feature of ribonucleic acids. The role of the hairpin loop is to negotiate a complete turn to allow base pairing of a single strand with itself. The properties of loop regions are of interest because they can serve as a site for tertiary interactions or as a recognition site for other molecules.

Advances in deoxyribonucleotide synthesis have made possible the preparation of large quantities of material of defined sequence required for NMR[1] studies (Gait, 1984). Only very recently has progress been made toward the routine synthesis of RNA oligonucleotides by chemical synthesis (Caruthers et al., 1987; Usman et al., 1987) and by in vitro transcription (Milligan et al., 1987). The first studies on nucleic acid hairpins were necessarily made on DNA oligomers as the best available model system for RNA structure. DNA hairpins have been observed from self-complementary oligonucleotides at low concentration (Wemmer et al., 1985; Patel et al., 1982), and the effect of loop size on thermodynamic stability of a series of hairpins has been investigated (Haasnoot et al., 1980, 1983). Distance geometry methods were used in conjunction with NMR to solve the solution structure of the hairpin formed by d(CGCGTTTTCGCG) (Hare & Reid, 1986). An NMR study of a similar hairpin, d(CGCGCGTTTTCGCGCG), has also been performed (Ikuta et al., 1986). An interesting hairpin containing only two bases in the loop region has recently been investigated (Orbons et al., 1987a,b).

We report here a detailed NMR investigation of the structure of the DNA oligomer d(CGCGTTGTTCGCG) that also adopts a hairpin conformation. This oligomer is similar in sequence to those described above, but is unique in that there are five bases in the loop region, one of which is a G residue. This sequence was originally designed to test the solution accessibility of guanosine residues in hairpin loops by using photo-CIDNP (McCord et al., 1984a,b). Initial studies on this molecule revealed several anomalous features in the NMR spectra, so we have attempted to characterize the possibilities of conformational flexibility and internal motions by as many NMR methods as possible. The oligomer has been characterized by proton NMR to provide interproton distances and torsion angle information and by phosphorus NMR to examine phosphate conformations. To probe dynamics in the loop region, the oligomer was resynthesized with thymidine residues labeled at C6 with carbon-13, and the relaxation behavior of the carbons was examined. A less detailed comparative study of two hairpins of similar sequence, d(CGCGTTTGTCGCG) and d(CTGCTCTTGTTGAGCAG), is presented in an accompanying paper. These sequences are related to that of the present study in that they contain the same stem and loop sequences, respectively, but differ in the remainder of the sequence. These changes lead to surprisingly large changes in the hairpin structure.

Materials and Methods

The DNA molecules d(CGCGTTGTTCGCG) and d(CGCGT*T*GT*T*CGCG), in which the thymidine C6

† This work was supported by a grant from the National Institutes of Health (R01 GM27738). S.G.B. is the recipient of a Presidential Young Investigator Award.
‡ Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309.

[1] Abbreviations: HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; NOESY, 2D NOE spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; NOE, nuclear Overhauser enhancement; AMBER, assisted model building with energy refinement; TMP, trimethyl phosphate; CSA, chemical shift anisotropy; T, Tesla; rms, root mean square.

0006-2960/89/0428-2819$01.50/0 © 1989 American Chemical Society