

Location of Cyanine-3 on Double-Stranded DNA: Importance for Fluorescence Resonance Energy Transfer Studies[†]

David G. Norman,[‡] Richard J. Grainger,^{‡,§} Dusan Uhrin,^{||} and David M. J. Lilley^{*,‡}

CRC Nucleic Acid Structure Research Group, Department of Biochemistry, The University of Dundee, Dundee DD1 4HN, U.K., and Department of Chemistry, The University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ, U.K.

Received December 23, 1999; Revised Manuscript Received March 31, 2000

ABSTRACT: Fluorescence resonance energy transfer provides valuable long-range distance information about macromolecules in solution. Fluorescein and Cy3 are an important donor–acceptor pair of fluorophores; the characteristic Förster length for this pair on DNA is 56 Å, so the pair can be used to study relatively long distances. Measurement of FRET efficiency for a series of DNA duplexes terminally labeled with fluorescein and Cy3 suggests that the Cy3 is close to the helical axis of the DNA. An NMR analysis of a self-complementary DNA duplex 5'-labeled with Cy3 shows that the fluorophore is stacked onto the end of the helix, in a manner similar to that of an additional base pair. This provides a known point from which distances calculated from FRET measurements are measured. Using the FRET efficiencies for the series of DNA duplexes as restraints, we have determined an effective position for the fluorescein, which is maximally extended laterally from the helix. The knowledge of the fluorophore positions can now be used for more precise interpretation of FRET data from nucleic acids.

Fluorescence resonance energy transfer (FRET) has been applied to the analysis of global conformation and folding transitions in nucleic acids with increasing frequency in the past decade. Potentially, the method provides distance information between known points that is typically in the range of 10–80 Å. This is a range that is inaccessible to other solution methods, and thus, the approach is complementary in principle to NMR in particular. FRET has been applied to the analysis of DNA structures (1–10) and more recently to the folding of a number of interesting RNA structures (11–18).

Determination of either relative or absolute distance information using FRET depends on the distance-dependent transfer of excitation energy from a donor to an acceptor fluorophore, resulting from the coupling between their transition dipoles. The efficiency of the transfer (E_{FRET}) depends inversely on the sixth power of the distance separating the fluorophores (19), and is given by

$$E_{\text{FRET}} = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \quad (1)$$

where R is the scalar distance between the fluorophores and R_0 is the characteristic Förster distance for the donor–

acceptor combination at which E_{FRET} equals 0.5. In general in our studies, we attach the fluorophores to the 5'-termini of known helical arms with a constant terminal sequence, and in this way, we can compare different end-to-end distances and study how these change upon the induction of conformational change.

Most FRET studies of nucleic acids have used fluorescein as the donor, and many have used tetramethylrhodamine as the acceptor. Fluorescein is generally very mobile when attached to DNA or RNA, and thus, orientation is not a problem in the interpretation of the measured FRET efficiencies. However, tetramethylrhodamine has some disadvantages, including the difficulty of coupling the fluorophore to the nucleic acid as a phosphoramidite. For this reason, we have largely changed to the use of indocarbocyanine-3 (Cy3) as the acceptor in our recent studies (Figure 1).

One of the major uncertainties in the use of distance information arising from FRET measurements is a lack of knowledge concerning the location of the fluorophores relative to the nucleic acid being studied. In this study, we have defined the position of Cy3 using a combination of fluorescence and NMR, and thereby also deduced the effective position of fluorescein.

MATERIALS AND METHODS

Construction of DNA Duplexes for Fluorescence and NMR Studies

DNA oligonucleotides were chemically synthesized using phosphoramidite chemistry using standard methods (20). Fully deprotected oligonucleotides for fluorescence spectroscopy were purified by gel electrophoresis as described previously (12). Fluorescein (PE-ABI) and Cy3 (Glen Research) were coupled to the 5'-termini as phosphoramid-

[†] This work was funded by the Cancer Research Campaign and the Biotechnology and Biological Sciences Research Council.

^{*} To whom correspondence should be addressed. D.G.N.: phone, 44-1382-344798; e-mail, dgnorman@bad.dundee.ac.uk. D.M.J.L.: phone, 44-1382-344246; e-mail, dmjlilley@bad.dundee.ac.uk; fax, 44-1382-201063.

[‡] The University of Dundee.

[§] Present address: Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, U.K.

^{||} The University of Edinburgh.

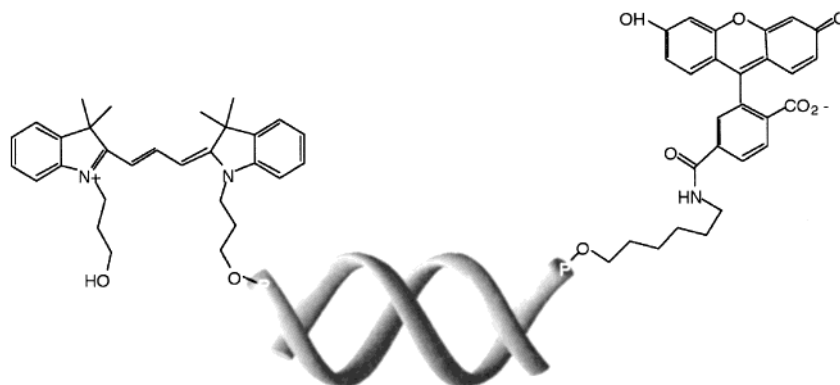


FIGURE 1: DNA duplex terminally labeled with fluorescein and Cy3. The FRET experiments employ DNA duplexes with fluorescein (donor) and Cy3 (acceptor) fluorophores covalently attached at the 5'-termini. The fluorescein is attached via a six-carbon tether, while the Cy3 is linked by a three-carbon tether. The fluorophores and the DNA are not drawn to scale in this schematic.

ites. Fluorescent DNA was further purified by reversed phase HPLC in 100 mM ammonium acetate (pH 7.5) using a C18 column (μ Bondapak, Waters). DNA was eluted with a linear gradient of acetonitrile. Stoichiometric quantities of fluorescein and Cy3-conjugated oligonucleotides were hybridized together in 90 mM Tris-borate (pH 8.3) and 50 mM NaCl (TBN buffer). The strands were heated to 90 °C for 3 min, slowly cooled to 4 °C, and purified by gel electrophoresis in TBN buffer. After 2-fold ethanol precipitation, the DNA was dialyzed against 90 mM Tris-borate (pH 8.3) and 100 mM NaCl at 4 °C before absorption and fluorescence measurements. Self-complementary DNA species for NMR studies were prepared by repeated chemical synthesis on a 1 μ mol scale, and purified by HPLC using a C18 reversed phase column (μ Bondapak, Waters) as described above.

Fluorimetry and Measurements of Fluorescence Resonance Energy Transfer

Fluorescent spectra were obtained at 4 °C using an SLM 8100 fluorimeter (SLM-Aminco, Urbana, IL) as described previously (12). Values of E_{FRET} were measured using the acceptor normalization method (1, 21). Fluorescence anisotropy (r) was determined from

$$r = (F_{\text{VV}} - GF_{\text{VH}})/(F_{\text{VV}} + 2GF_{\text{VH}}) \quad (2)$$

where G is the correction factor given by the ratio $F_{\text{HV}}/F_{\text{HH}}$, where subscripts V and H refer to fluorescence with vertical and horizontal polarizers, respectively, in the order excitation, followed by emission.

NMR spectroscopy and structure determination

Sample Preparation. NMR spectra were acquired from 200 A_{260} units of Cy3-conjugated or 150 A_{260} units of unmodified DNA dissolved in 0.6 mL of 0.1 M NaCl, 10 mM phosphate, and 1 mM EDTA solutions of either 99.96% D_2O or 90% $H_2O/10\%$ D_2O (v/v) at pH 7.0.

NMR Experiments. All spectra were acquired at 14 °C using a Varian INOVA 600 NMR spectrometer. A very weak RF field at the HOD frequency was applied during the relaxation delay of 1.5 s. Spectral widths in F_2 and F_1 were 8000 and 6100 Hz, respectively. The acquisition times in t_2 and t_1 dimensions were 256 and 168 ms, respectively. The sign discrimination in F_1 was achieved by the States-TPPI method (22). The number of scans per increment was 8 for

TOCSY (23) and 16 for NOESY (24) and DQF-COSY (25) experiments. In TOCSY experiments, the DIPSI-2 spin lock (26) was applied during 70 ms when $\gamma B/2\pi$ equaled 8.25 kHz. Two delays, each equal to one-quarter of the mixing time, were placed before and after the spin lock to compensate for the ROE peaks created during the spin lock. Pulsed field gradients of 1 ms and 10 G/cm were applied at the beginning of these delays. A series of NOESY spectra was acquired with mixing times of 50, 100, 150, and 200 ms. The NOESY experiment with a mixing time of 200 ms was repeated with the sample dissolved in H_2O . The spectral width in F_1 was increased to 8500 Hz, resulting in an acquisition time in the t_1 dimension of 120.5 ms. Soft-pulse WATERGATE (27) was used for water suppression in this instance.

Data Processing and Analysis. NMR data were processed, assigned, and quantified using Felix (MSI). NOESY data were zero filled in F_1 to 1024 points and apodized using a 90°-shifted sine bell function. NOESY cross-peaks from the D_2O spectra were quantified by volume integration at the four mixing times. COSY and TOCSY experiments were used for assignment purposes only. Spectra acquired in H_2O were assigned, but no quantification of NOE cross-peaks was attempted.

Structure Determination. Structures were calculated using restrained molecular mechanics and dynamics within the program X-PLOR version 3.1 (28). B-form DNA was constructed using Biopolymer within Insight II (MSI). Cy3 coordinates were taken from the crystal structure of (*S,S*)-(+)-1,3-bis(3-ethyl-1,3-dimethyl-2-indolyl)trimethinium iodide (29). Parameters for molecular dynamics were calculated using the XPLO2D program (30) and modified manually. Cy3 was initially placed on DNA in an extended conformation parallel to the helix axis. This was subjected to energy minimization and molecular dynamics in the absence of experimental restraints, but with the oligonucleotide conformation fixed, to generate a family of 10 starting structures. Structural refinement was carried out in several phases. A distance restraint list was generated from the measured NOE peaks. NOE peak volumes were converted to distances using an empirically derived constant with generous error ranges. For each of the starting structures, one round of refinement was carried out using distance-restrained simulated annealing (28) with the oligonucleotide conformation fixed. Several rounds of molecular dynamics using full relaxation matrix

refinement (31) were carried out with decreasing constraint of the oligonucleotide conformation, generating a final family of 10 structures. Because of symmetry, this yielded 20 unique structures for the Cy3–DNA interface, of which one was discarded due to high residual violations. An average structure was calculated and minimized using full relaxation restraints.

RESULTS AND DISCUSSION

Measurement of R_0 for the Fluorescein–Cy3 Fluorophore Pair from Spectral Overlap. In previous FRET experiments on DNA and RNA, we have found that when the fluorophores are attached to a 5′-CC sequence they are spectrally well-behaved (1, 2, 5, 12, 32). We have calculated the characteristic Förster distance (R_0) for the fluorescein–Cy3 pairing in this sequence environment, from the spectral properties of the fluorophores. R_0 is given by

$$R_0^6 = 8.8 \times 10^{-28} [\Phi_D \kappa^2 \eta^{-4} J(\lambda)] \quad (3)$$

where Φ_D is the fluorescent quantum yield of the donor, η is the refractive index of the medium, and κ is related to the relative orientation of the two transition dipole moments, i.e.

$$\kappa = \mathbf{D} \cdot \mathbf{A} - 3(\mathbf{R}_{DA} \cdot \mathbf{D})(\mathbf{R}_{DA} \cdot \mathbf{A}) \quad (4)$$

consisting of three scalar products between the unit vectors \mathbf{D} (the donor emission transition moment), \mathbf{A} (the acceptor absorption transition moment), and \mathbf{R}_{DA} (the donor–acceptor vector). Although κ^2 can take a value between 0 and 4, depending on the relative fluorophore orientation, this averages in a known way in the case of fluorophore mobility, and the value of $2/3$ is generally assumed under these conditions.

$J(\lambda)$ is the normalized spectral overlap integral, given by

$$J(\lambda) = \frac{\int_0^\infty \phi_D(\lambda) \epsilon_A \lambda^4 d\lambda}{\int_0^\infty \phi_D(\lambda) d\lambda} \quad (5)$$

where ϕ_D is the spectral shape of the fluorescent emission of the donor and ϵ_A is the molar absorbance of the acceptor at each wavelength (λ). Using the fluorescent emission spectrum of a 10 bp duplex DNA 5′-labeled with fluorescein (donor) and the absorption spectrum of the same DNA sequence 5′-labeled with Cy3 (acceptor), we calculated the overlap integral numerically as $4.1 \times 10^{32} \text{ nm}^6 \text{ mol}^{-1}$. In view of the low anisotropy of fluorescein attached to our DNA species (see below), we have taken a value for κ^2 of 0.67. Using values for Φ_D of 0.39 (3) and η of 1.33, we calculated a value for R_0 of 55.7 Å. This value is very close to that obtained in an earlier measurement by Jares-Erijman and Jovin (7) using the related acceptor sulfoindocarbocyanine-3 together with a different DNA sequence.

Efficiency of FRET for the Duplex Series as a Function of Length. We have studied the efficiency of energy transfer between fluorescein and Cy3 fluorophores attached at the 5′-termini of a series of DNA duplex species varying between 10 and 24 bp in length. These were based on the same sequences used in our previous study of energy transfer between fluorescein and tetramethylrhodamine (32). The sequences of the molecules ensure that both terminal 5 bp

Table 1: Anisotropy (r), FRET Efficiency Values, and Calculated Distances, for the Series of DNA Duplex Species^a

length (bp)	r_F	r_{Cy3}	E_{FRET}^b	calcd distance (Å) ^c
10	0.147	0.295	0.720 ± 0.020	47.6
12	0.146	0.300	0.644 ± 0.021	50.5
14	0.127	0.299	0.471 ± 0.029	56.8
18	0.104	0.303	0.257 ± 0.016	66.5
20	0.104	0.305	0.164 ± 0.010	73.1
24	0.093	0.305	0.084 ± 0.003	83.0

^a r values for doubly labeled DNA were measured in 90 mM Tris-borate (pH 8.3) and 100 mM NaCl using a λ_{ex} of 490 and a λ_{em} of 547 nm for fluorescein and a λ_{ex} of 520 and a λ_{em} of 560 nm for Cy3. Duplexes were constructed by hybridizing the following Cy3-labeled strands to their fluorescein-labeled complements: CCACCTAGG, CCACCTGGCTAGG, CCACCTGCTGCTAGG, CCACCTGCACGCTGCTAGG, CCACCTGCACTCGCTGCTAGG, and CCACCTGCACTGCTGCTGCTAGG. ^b Including standard errors on replicate measurements. ^c The calculated distances were derived using eq 1 with an R_0 of 55.7 Å.

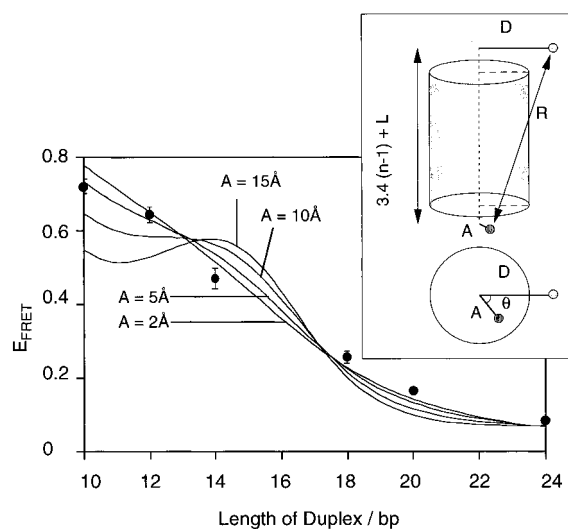


FIGURE 2: FRET efficiencies for the series of DNA duplex species. FRET efficiencies were measured for the DNA duplex series as a function of helical length. The experimental data are shown by the points (●). The error bars show standard deviations from replicate measurements. The lines were generated using the cylindrical model shown in the inset, using eq 5. The different curves are labeled with the values of the acceptor radial distance A . The values $D = 25$ Å, $L = 6$ Å, $P = 230^\circ$, and $R_0 = 56$ Å were taken for these simulations. Note the increase in the extent of modulation as the acceptor is placed further from the helical axis, which is not observed in the experimental data.

sections, and thus the local environment of the fluorophores, remain identical throughout the series.

The efficiency of energy transfer was measured using the acceptor normalization method (21), and the results are given in Table 1. Graphical presentation of the efficiencies as a function of duplex length shows the expected reduction in E_{FRET} as the length of the helix increases (Figure 2), as observed previously for the fluorescein–tetramethylrhodamine pair (32). However, there is a significant difference in the shape of the profiles obtained for the two fluorophore pairs. Whereas the data obtained using the fluorescein–tetramethylrhodamine pair exhibited a pronounced modulation of E_{FRET} with length, giving a clear shoulder at 12 bp, the data from the fluorescein–Cy3 pair display no such modulation. The helix length dependence obtained with the fluorescein–tetramethylrhodamine pair was ascribed to the variation in

the interfluorophore distance arising from the cylindrical geometry of the double helix (32); i.e., the distance is shorter when the fluorophores are on the same side of the helix, and longer when they are on opposite sides. However, if one of the dyes is close to the helix axis, this is no longer the case, and the modulation should be absent.

This is shown by simulating the data using the cylindrical model shown in Figure 2. The radial extensions of the donor and acceptor are D and A , respectively, and the axial separation corresponds to the length of the helix plus a further constant addition (L) since either or both fluorophores could be extended along this direction. The angular separation between the fluorophores is that of the phosphates to which they are attached plus a constant additional phase angle (P) to account for the angular difference between phosphates on opposite strands, and any nonradial orientation of the fluorophores. This leads to the calculated FRET efficiency of

$$E_{\text{FRET}} = [1 + \{[3.4(n-1) + L]^2 + [D - A \cos(\theta + P)]^2 + [A \sin(\theta + P)]^2\}^{1/2}/R_0]^6]^{-1} \quad (6)$$

where the rotation angle θ (in degrees) is given by $34(n-1)$. Values of E_{FRET} were calculated for different helix lengths by keeping the mean position of the donor in its normal radially extended position [$D = 25 \text{ \AA}$, consistent with the previous study using tetramethylrhodamine as the acceptor (32)] and varying the radial extension of the acceptor (A) from 2 to 15 \AA (Figure 2). As A increases, so does the modulation. Flexibility would be expected to smooth out this modulation to some degree. However, calculations incorporating torsional flexibility show that the modulation is still observed, as indicated by the earlier studies using fluorescein–rhodamine (32). The experimental data are in best agreement with the values calculated on the basis that the Cy3 lies close to the helical axis. As a result of this, we fixed the position of the Cy3 at 2 \AA from the axis, and fitted the experimental data with the value of R_0 as a free variable. The best fit corresponded to an R_0 value of 55.3 \AA . This is in good agreement with the value (55.7 \AA) obtained from the spectral overlap calculation.

Fluorescence Anisotropies of Fluorescein and Cy3. We have also measured the fluorescence anisotropy of both fluorescein and Cy3 for each member of the duplex series (Table 1). These were measured in the presence of 0, 1, and 5 mM magnesium ions, but the results were closely similar in each case; thus, we only present the data obtained in the absence of the divalent ion. In general, the anisotropy of fluorescein is in the range of 0.09–0.15, showing that the fluorophore is very mobile during the lifetime of the excited state. By contrast, the anisotropy of Cy3 is in the range of 0.30–0.31, indicating that movement of the fluorophore is extremely constrained by the DNA. Closer examination of the anisotropy values for fluorescein shows a systematic reduction with helix length. This is a direct result of the energy transfer. With an increase in FRET efficiency, the fluorescent lifetime is shortened, leaving less time for rotation of the fluorophore in the duration of the excited state and thus greater polarization of the emitted radiation. Thus, these measurements confirm the trend obtained by the direct measurement of FRET efficiency.

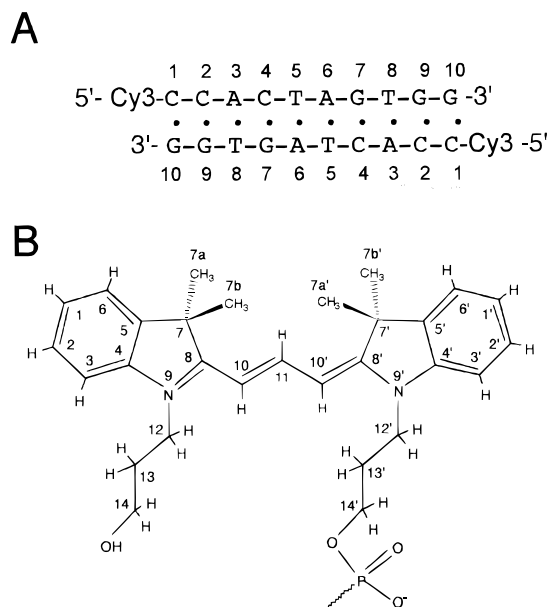


FIGURE 3: Numbering schemes of DNA and Cy3 used in the NMR study. (A) The sequence of the DNA duplex and the numbering of the bases. (B) The structure of Cy3, with the atomic numbering system used in this study.

Location of the Cy3 Suggested by the Fluorescence Data.

From the absence of periodicity in the FRET efficiencies of the duplex series, particularly when contrasted with that observed in earlier studies of a duplex series labeled with fluorescein and tetramethylrhodamine (32), it is apparent that the Cy3 must be positioned centrally, close to the long axis of the DNA duplex. While the negative charge on the fluorescein leads to repulsion by the ribose–phosphate backbone, and hence extension away from the DNA, Cy3 is positively charged and so could be held on the end of the duplex by a combination of hydrophobic and attractive electrostatic forces. A relatively fixed location of the Cy3 on the end of the helix would be consistent with the high anisotropy of this fluorophore, by comparison with the relatively mobile fluorescein.

NMR Studies of a Cy3-Labeled DNA Duplex. To define the position of the Cy3 on the DNA with greater precision, we have used NMR. We synthesized the self-complementary oligonucleotide 5'-CCACTAGTGG-3' with and without Cy3 attached via a three-carbon tether at the 5'-terminus. The numbering system for the DNA nucleotides is shown in Figure 3A. The DNA sequence was chosen to be directly comparable with that used in the fluorescence experiments.

Assignment of the DNA Proton Resonances. The nonexchangeable protons in the fluorophore-unconjugated DNA were assigned primarily by examination of the base to H1' cross-peaks in the 200 ms NOESY spectra (Figure 4B). The parent DNA gave a typical NOE pattern expected for a B-form duplex. The exchangeable protons were assigned from spectra acquired in H_2O at 4 $^\circ\text{C}$. Assignments were made by comparison with the nonexchangeable resonances. Comparison between the parent oligonucleotide and Cy3-labeled oligonucleotide spectrum showed only minor chemical shift differences in most cases. It was possible to assign the DNA proton resonances of the Cy3 oligonucleotide without assumptions about the identity of the peaks derived from the Cy3 adduct. The chemical shifts of assigned protons

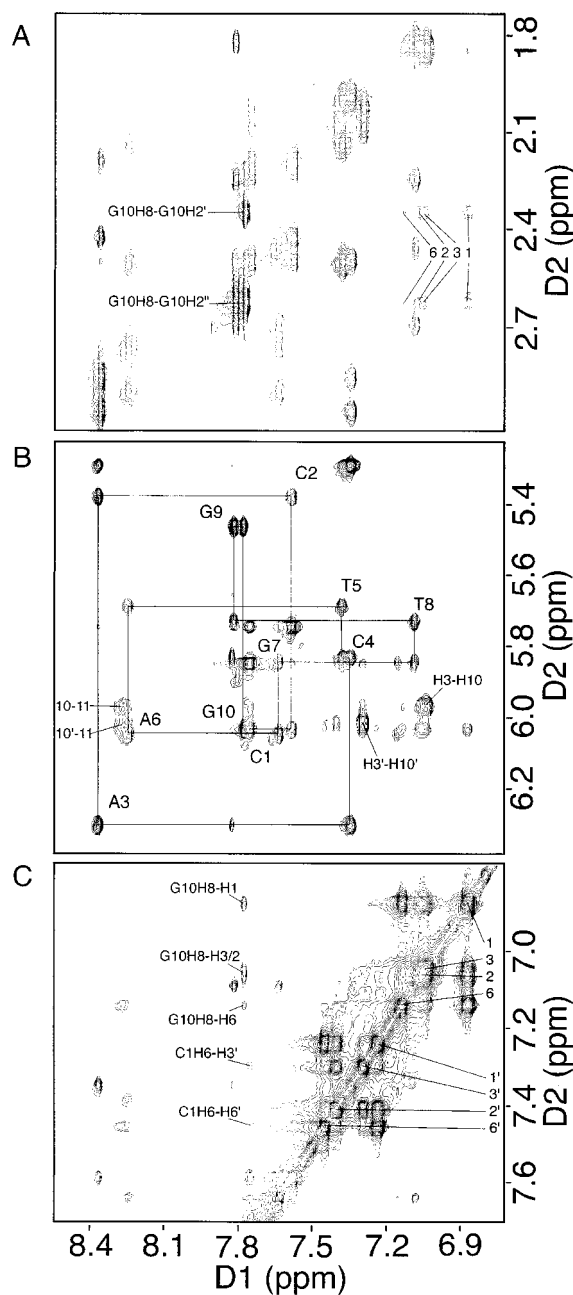


FIGURE 4: NMR spectra of the Cy3-labeled DNA duplex. Selected regions of a 200 ms NOESY spectrum recorded in D₂O at 15 °C. (A) The base to H2' and H2'' region, showing the positions of the G₁₀ H2' and H2'' protons together with cross-peaks arising from the Cy3 indole ring system. (B) The DNA base to H1' fingerprint region. The base to H1' connectivities are represented by lines and annotation. Some significant cross-peaks from the Cy3 are indicated. (C) The aromatic region of the two-dimensional NOESY spectrum, showing diagonal resonance positions arising from the two Cy3 indole ring systems as well as cross-peaks between the terminal base pair and Cy3 aromatic protons.

from both the parent oligonucleotide and the Cy3-conjugated oligonucleotide are presented in Table 2.

Assignment of Cy3 Proton Resonances. Our numbering system for Cy3 is shown in Figure 3B. The molecule comprises two indole ring systems connected by a planar trimethylene linker. The Cy3 is covalently connected to the 5'-phosphate terminus of the DNA via the tether at N9'; note that there is also an equivalent free chain attached at the corresponding N9 of the other indole ring, which we term the pseudotether.

COSY and NOESY (Figure 4C) spectra showed two sets of four non-DNA protons resonating between 6.87 and 7.44 ppm. The pattern of coupling and the size of NOE peaks identified these two sets of signals as belonging to Cy3 indole aromatic protons 1, 2, 3, and 6 and 1', 2', 3', and 6'. The numbering of these proton resonances was made possible by the observation of NOE interactions between the Cy3 methyl protons at ~1.5 ppm (data not shown) and the 6 and 6' positions. Scalar coupling within each ring spin system allowed the assignment of the other protons. The protons of the trimethylene linker were assigned by virtue of the strong scalar coupling between protons 10 or 10' and 11. These proton resonances could then be linked to the indole ring systems via strong NOE interactions with the intervening methyl groups.

Despite the internal assignment of indole protons, assignment of the spin systems between the indole rings with and without the tether was not possible directly. This problem was resolved by the identification of cross-peaks from one indole ring system to the G₁₀ H2' and H2'' sugar protons (Figure 4A), complemented by the observation of NOEs between the G₁₀ H8 proton and the same Cy3 ring system (Figure 4C). In contrast, protons from the other ring system exhibited no NOEs to G₁₀ but did show NOEs to the cytosine, sugar and base protons.

The Cy3 to DNA NOE interactions therefore permitted model-dependent specific assignment of the pseudosymmetrical Cy3. Assignment of protons in the tether and pseudotether were made by reference to scalar couplings observed in the COSY spectra, and to NOE cross-peaks between the Cy3 H12 and H12' protons and the Cy3 H3 and H3' and Cy3 H10 and H10' protons. Stereospecific assignment of the methylene protons of the tether chains was not possible.

Once all the Cy3 protons and the majority of the DNA protons had been assigned, the next task was to identify the NOE cross-peaks in the heavily overcrowded base to the H1' region of the spectrum. The resonance peaks for Cy3 H10' (Figure 4B), C1 H1', and G₁₀ H1' overlapped badly. The temperature at which the spectra had been recorded (15 °C) was optimized for the separation of these crucial resonance peaks. The cross-peak between Cy3 H3 and Cy3 H10 was clear of overlap, and therefore the cross-peak between Cy3 H3' and Cy3 H10' could be identified by analogy (Figure 4B). The prior assignment of NOE cross-peaks between Cy3 H1 and G₁₀ H2' and H2'' allowed a model-based identification of the NOE cross-peak between Cy3 H1 and G₁₀ H1' as well as those between Cy3 H2, H3, and H6 and G₁₀ H1'.

Structure Determination and Analysis. The structure of the Cy3–DNA adduct was determined using restrained molecular mechanics and dynamics, according to the protocol described in Materials and Methods. Refinement strategies were designed to determine the conformation of the Cy3 relative to the DNA, since the conformation of the oligonucleotide itself is not an issue, beyond confirming a B-form structure. A total of 336 NOE cross-peaks were measured by volume integration from spectra at four different mixing times.

Nineteen superimposed structures are shown in Figure 5B. The position of the planar heterocyclic body of Cy3 was highly defined by its large number of NOE contacts to the terminal C•G base pair of the DNA (Figure 5A) despite its

Table 2: Assignments and Chemical Shifts (ppm) of the Protons of the DNA with and without Attached Cy3

(A) Cy3-Conjugated DNA (5'-Cy3-CCACTAGTGG-3')										
DNA	H8/H6	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H1/H3	H4a	H4b
C1	7.74	5.85	6.04	2.22	2.50	4.82	4.17		8.31	7.08
C2	7.58	5.75	5.37	2.20	2.43	4.89	4.16		8.71	7.06
A3	8.36	7.82	6.31	2.85	2.96	5.11	4.51			
C4	7.34	5.29	5.83	1.99	2.50	4.72	4.37		8.23	6.90
T5	7.37	1.64	5.69	2.13	2.50	4.90	4.16	13.79		
A6	8.24	7.15	6.05	2.76	2.89	5.06	4.44			
G7	7.63		5.85	2.46	2.69	4.93	4.41	12.82		
T8	7.08	1.29	5.73	1.82	2.24	4.83	4.15	13.96		
G9	7.81		5.46	2.62	2.62	4.97	4.32	13.12		
G10	7.77		6.04	2.34	2.62	4.73	4.26	12.90		
(B) Unconjugated DNA (5'-CCACTAGTGG-3')										
DNA	H8/H6	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H1/H3	H4a	H4b
C1	7.79	5.97	6.02	2.10	2.54	4.70	4.14		8.05	7.10
C2	7.63	5.76	5.28	2.17	2.46	4.89	4.16		8.69	7.06
A3	8.38	7.89	6.32	2.682	2.96	5.09	4.48			
C4	7.34	5.30	5.72	1.99	2.50	4.71	4.36		8.19	6.89
T5	7.39	1.64	5.72	2.13	2.50	4.90	4.16	13.79		
A6	8.25	7.21	6.07	2.76	2.89	5.07	4.43			
G7	7.66		5.87	2.46	2.69	4.92	4.40	12.77		
T8	7.14	1.34	5.72	1.86	2.27	4.84	4.20	13.97		
G9	7.86		5.70	2.72	2.67	4.99	4.36	13.11		
G10	7.84		6.20	2.40	2.56	4.67	4.25	13.32		

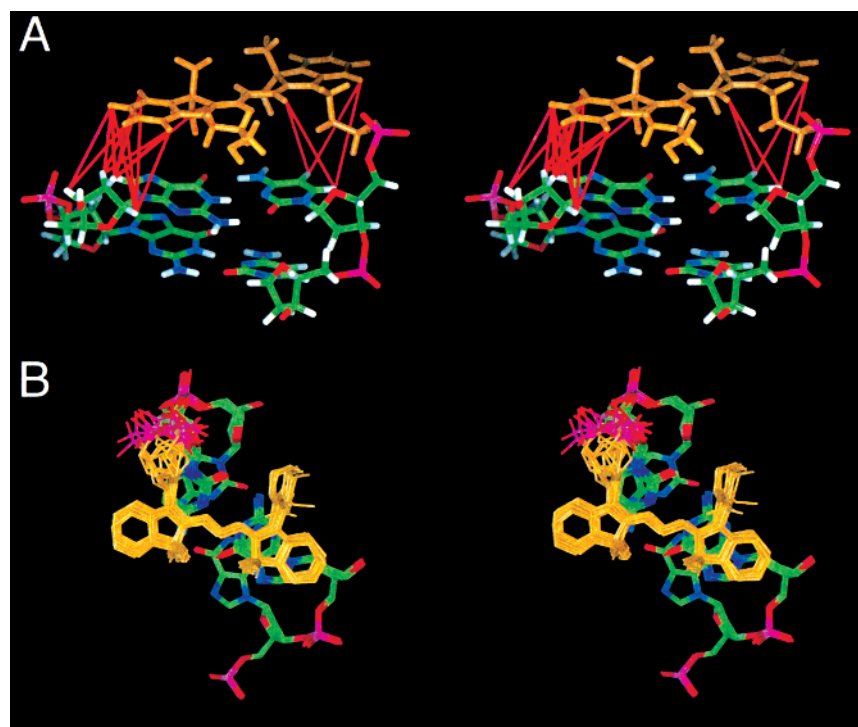


FIGURE 5: Stereopairs showing the conformation of the Cy3 fluorophore and the two end base pairs of the DNA decamer. Cy3 is colored orange, and the DNA is colored according to element. Structures were derived from 336 NOE-derived restraints, including 30 Cy3–DNA interactions. The pairwise rmsd for the structures shown is 0.415 Å, with a mean *R*-factor (33) of 0.0675 ± 0.0013 . (A) The minimized average conformation of Cy3 with many of the significant NOE interactions between Cy3 and the DNA indicated by red lines joining the participating protons. (B) An overlay of 19 structures viewed down the helix axis, showing the stacking interaction of Cy3 with the terminal base pair of the DNA.

terminal position. The fluorophore, connected to the terminal phosphate by a three-carbon tether, stacks across the terminal base pair (Figure 5B), making the overall structure of the Cy3–oligonucleotide adduct strikingly similar to that of a natural base pair. The observed rise of the Cy3 ring system is 5 Å, compared with an average value of 3.47 Å for the bases of the DNA. This nonoptimal stacking interaction is largely caused by the imposition of methyl group 7a between the Cy3 ring and G₁₀. The center of the Cy3 lies approximately 2 Å from the DNA helix axis. The helical twist is 45°, measured from cytosine₁ N1 to Cy3 N9', and 20°, measured from Cy3 N9 to G₁₀ N9. This compares with an average of 36.4° for the rest of the oligonucleotide. The indole rings are shown to be stacked on the terminal bases with different extents of overlap, with the G₁₀ base being the main contributor to this interaction.

Comparison of chemical shift and NOESY cross-peak intensities indicates that very little change in the DNA conformation results from attachment of the Cy3. NOESY spectra acquired in H₂O exhibited cross-peaks typical of Watson–Crick hydrogen bonding in all base pairs, with negative propeller twists throughout, and the sugars of the DNA remain in the C_{2'}-endo conformation. NOE cross-peaks from the terminal imino proton of G₁₀ were extremely weak or undetectable, but a strong exchange peak with water that is typical of a terminal base pair with some exposure to the solvent was observed. Significant chemical shift changes were observed between terminal exchangeable protons in DNA, with and without attached Cy3. The imino proton of G₁₀ was found to shift by 0.42 ppm, and the amino protons of cytosine₁ were found to shift by 0.26 and 0.02 ppm.

This structure is evidently dynamic to some degree. We observed the normal high level of proton exchange between the terminal imino proton and water in not only the parent oligonucleotide but also the Cy3 adduct, supporting a dynamic model. Nevertheless, we believe that the structure represented in Figure 5 is a good description of the molecule. Only one set of resonances is observed, ruling out alternative conformations in slow exchange, and the Cy3 proton line shapes do not support intermediate exchange rates. We observe a large number of NOEs between protons on the indole rings and on DNA; significant conformational dynamics would reduce this number and the observed intensities. Additionally, the structure that is generated fits the observed NOE buildups to a significantly small *R* value, consistent with a single structure. The location of the Cy3 stacked on the end of the helix explains the chemical shift changes of DNA protons lying beneath the indole rings (cytosine₁ H4a, Δ0.26 ppm; G₁₀ H1, Δ0.42 ppm; G₁₀ H8, Δ0.07 ppm; G₁₀ H1', Δ0.16 ppm; G₁₀ H2', Δ0.06 ppm; and G₁₀ H2'', Δ0.06 ppm). The constrained position of the fluorophore is also consistent with the high fluorescence anisotropy.

The three-carbon, aliphatic tethers are difficult to define structurally by virtue of their chemistry and the lack of stereospecific assignment. Although NOE buildup rates from tether protons have been included in the relaxation matrix refinement, they are nevertheless poorly defined in the final structures. The lack of definition in the conformation of the tether arms does not seem to have had a large impact on the definition of the indole ring systems in relation to the oligonucleotide.

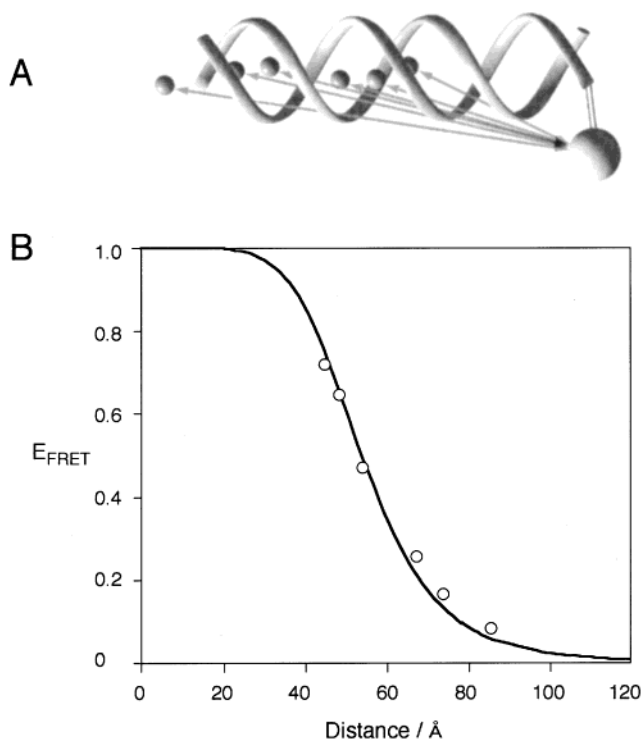


FIGURE 6: Determination of the effective position of fluorescein. (A) Positions of the Cy3 fluorophores (small spheres) and the effective position of fluorescein (large sphere). The arrows represent the distances measured by FRET and used to position fluorescein. (B) The fluorescein–Cy3 distances have been recalculated using the derived effective position of fluorescein. Plot of the E_{FRET} values vs these model-derived interfluorophore distances, fitted to the Förster equation. This gave an R_0 of 54.1 Å.

Effective Position of Fluorescein. Having determined the position of Cy3 on the end of the duplex, we have used the fluorescein–Cy3 distances obtained from the FRET data for the duplex series (using our calculated value of R_0) as constraints to determine an effective position of the fluorescein. We have used energy minimization to calculate the fluorescein position with a maximal tether distance of 17 Å from the terminal oxygen atom. The resulting position is illustrated in Figure 6. The extended position is consistent with our previous measurements of the FRET efficiencies of a series of DNA duplexes terminally labeled with fluorescein and tetramethylrhodamine (32), and with electrostatic repulsion of the negatively charged fluorescein. We have used the derived positions of fluorescein and Cy3 to recalculate the donor–acceptor distances for the duplex series, which were then used to fit to the Förster equation (eq 1) shown in Figure 6. The calculated fluorescein position satisfies the experimental and calculated distances with only small errors. The agreement between the experimental data and the calculated distances for the duplex series demonstrates empirically the quantitative reliability of the FRET approach.

Use of the Fluorescein–Cy3 Pair for the FRET Analysis of Nucleic Acid Structures. A knowledge of the location of one of the fluorophores is a big advantage in the structural interpretation of FRET data. The NMR data indicate clearly that the fluorophore Cy3 lies predominantly stacked onto the end of the helix close to the axis when attached to the 5'-terminus by a three-carbon tether. This provides a known point from which distances are measured. Attached fluorescein

cein is very much more mobile, but we have been able to define an effective position of this fluorophore using FRET distance constraints. As a result, we now have two measuring points that can be used to define the global structure of nucleic acids. Moreover, the value of the characteristic Förster distance is relatively long compared to those of many fluorophore pairings, resulting principally from the very good overlap between fluorescein emission and Cy3 excitation. The value of R_0 depends on the refractive index used in the calculation. We chose a η value of 1.33, corresponding to bulk water, because of the extended position of fluorescein. Calculations using larger values of refractive index (giving correspondingly shorter values of R_0) resulted in a failure to converge on a satisfactory solution to the set of data for the duplexes. The value for R_0 of 56 Å means that the dependence of E_{FRET} on distance will be most sensitive in the range of 35–85 Å (i.e., where E_{FRET} will range from 0.9 to 0.1). This is an extremely useful distance range for the study of the global structure and folding of nucleic acids, and we have exploited this in the determination of the global structure of the 3'-UTR element of U1A pre-mRNA (18). These long-range distance constraints will complement the shorter-range distances that are available from NMR. While many fewer such distances will be derived from FRET, these longer distances are nevertheless extremely valuable in helping to define the global architecture of branched or otherwise distorted nucleic acids.

The coordinates of the Cy3–DNA structure are available via <http://www.dundee.ac.uk/biochemistry/dgn.htm>.

ACKNOWLEDGMENT

We thank Scott Wassell for preliminary studies and Frank Stühmeier and Tim Wilson for valuable discussion.

REFERENCES

- Murchie, A. I. H., Clegg, R. M., von Kitzing, E., Duckett, D. R., Diekmann, S., and Lilley, D. M. J. (1989) *Nature* **341**, 763–766.
- Clegg, R. M., Murchie, A. I. H., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D. M. J. (1992) *Biochemistry* **31**, 4846–4856.
- Eis, P. S., and Millar, D. P. (1993) *Biochemistry* **32**, 13852–13860.
- Clegg, R. M., Murchie, A. I. H., Zechel, A., and Lilley, D. M. J. (1994) *Biophys. J.* **66**, 99–109.
- Gohlke, C., Murchie, A. I. H., Lilley, D. M. J., and Clegg, R. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11660–11664.
- Mergny, J. L., Boutorine, A. S., Garestier, T., Belloc, F., Rougee, M., Bulychyev, N. V., Koshkin, A. A., Bourson, J., Lebedev, A. V., Valeur, B., Thuong, N. T., and Helene, C. (1994) *Nucleic Acids Res.* **22**, 920–928.
- Jares-Erijman, E. A., and Jovin, T. M. (1996) *J. Mol. Biol.* **257**, 597–617.
- Yang, M. S., and Millar, D. P. (1996) *Biochemistry* **35**, 7959–7967.
- Stühmeier, F., Welch, J. B., Murchie, A. I. H., Lilley, D. M. J., and Clegg, R. M. (1997) *Biochemistry* **36**, 13530–13538.
- Mergny, J. L. (1999) *Biochemistry* **38**, 1573–1581.
- Tuschl, T., Gohlke, C., Jovin, T. M., Westhof, E., and Eckstein, F. (1994) *Science* **266**, 785–789.
- Bassi, G. S., Murchie, A. I. H., Walter, F., Clegg, R. M., and Lilley, D. M. J. (1997) *EMBO J.* **16**, 7481–7489.
- Murchie, A. I. H., Thomson, J. B., Walter, F., and Lilley, D. M. J. (1998) *Mol. Cell* **1**, 873–881.
- Walter, F., Murchie, A. I. H., Duckett, D. R., and Lilley, D. M. J. (1998) *RNA* **4**, 719–728.
- Walter, F., Murchie, A. I. H., Thomson, J. B., and Lilley, D. M. J. (1998) *Biochemistry* **37**, 14195–14203.
- Walter, F., Murchie, A. I. H., and Lilley, D. M. J. (1998) *Biochemistry* **37**, 17629–17636.
- Walter, N. G., Hampel, K. J., Brown, K. M., and Burke, J. M. (1998) *EMBO J.* **17**, 2378–2391.
- Grainger, R. J., Norman, D. G., and Lilley, D. M. J. (1999) *J. Mol. Biol.* **288**, 585–594.
- Förster, T. (1948) *Ann. Phys.* **2**, 55–75.
- Beaucage, S. L., and Caruthers, M. H. (1981) *Tetrahedron Lett.* **22**, 1859–1862.
- Clegg, R. M. (1992) *Methods Enzymol.* **211**, 353–388.
- Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.* **85**, 393–399.
- Rance, M. (1987) *J. Magn. Reson.* **74**, 557–564.
- Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys.* **71**, 4546–4553.
- Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117**, 479–485.
- Rucker, S. P., and Shaka, A. J. (1989) *Mol. Phys.* **68**, 509–517.
- Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2**, 661–665.
- Brünger, A. T. (1992) *X-PLOR (version 3.1) manual*, Yale University Press, New Haven, CT.
- Reichardt, C., Engel, H.-D., Allmann, R., Kucharczyk, D., and Krestel, M. (1990) *Chem. Ber.* **123**, 565–581.
- Kleywegt, G. J. (1995) *CCP4/ESF-EACBM Newsletter on Protein Crystallography* **31**, 45–50.
- Nilges, M., Habazettl, J., Brünger, A. T., and Holak, T. A. (1991) *J. Mol. Biol.* **219**, 499–510.
- Clegg, R. M., Murchie, A. I. H., Zechel, A., and Lilley, D. M. J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2994–2998.
- James, T. L., Gochin, M., Kerwood, D. J., Schmitz, U., and Thomas, P. D. (1991) *Computational aspects of the study of biological macromolecules by NMR* (Hoch, J. C., Ed.) Plenum Press, New York.

BI992944A