Sequence specificity of mutagen–nucleic acid complexes in solution: Intercalation and mutagen–base pair overlap geometries for proflavine binding to dC-dC-dG-dG and dG-dG-dC-dC self-complementary duplexes

(mutagen–tetranucleotide duplex/helix–coil transition/intercalation specificity at dC-dG sites)

DINSHAW J. PATEL AND LITA L. CANUEL

Bell Laboratories, Murray Hill, New Jersey 07974

Communicated by F. A. Bovey, April 29, 1977

ABSTRACT The complex formed between the mutagen proflavine and the dC-dC-dG and dG-dC-dC self-complementary tetranucleotide duplexes has been monitored by proton high resolution nuclear magnetic resonance spectroscopy in 0.1 M phosphate solution at high nucleotide/drug ratios. The large upfield shifts (0.5 to 0.85 ppm) observed at all the proflavine ring nonexchangeable protons on complex formation are consistent with intercalation of the mutagen between base pairs of the tetranucleotide duplex. We have proposed an approximate overlap geometry between the proflavine ring and nearest neighbor base pairs at the intercalation site from a comparison between experimental shifts and those calculated for various stacking orientations. We have compared the binding of actinomycin D, propidium diiodide, and proflavine to self-complementary tetranucleotide sequences dC-dC-dG-dG and dG-dC-dC by UV absorbance changes in the drug bands between 400 and 500 nm. Actinomycin D exhibits a pronounced specificity for sequences with dC-dC sites (dC-dC-dG-dC), while propidium diiodide and proflavine exhibit a specificity for sequences with dG-dC sites (dC-dC-dG-dC). Actinomycin D binds more strongly than propidium diiodide and proflavine to dC-dC-dG-dC (contains dC-dG and dG-dC binding sites), indicative of the additional stabilization from hydrogen bonding and hydrophobic interactions between the pentapeptide lactone rings of actinomycin D and the base pair edges and sugar-phosphate backbone of the tetranucleotide duplex.

The cationic acridine dyes form two types of complexes with nucleic acids dependent on the nucleotide to dye ratio (1). At high nucleotide/dye ratios, there is a red shift which has been attributed to intercalation of the dye between nucleic acid base pairs (2). At low nucleotide/dye ratios, there is a blue shift in the visible spectrum of the dye which is attributed to stacking of dye aggregates along the sugar-phosphate backbone (1). The published optical studies include a quantitative analysis of the binding of proflavine to polynucleotides and transfer RNA (3, 4), to DNA (5, 6), and to bacteriophage (7, 8).

The 2-2 complexes of 9-aminoacridine with the dinucleoside phosphates adenylic-(3'5')-uridic (9) and 5-iodocytidyl-(3'5')-guanose (10) have been solved by x-ray crystallography. The adenine bases form Hoogsteen type hydrogen bonds to the uracil bases in the former crystal structure (9), and the base pairs are stacked parallel to each other with the 9-aminoacridine sandwiched between them. The cytosine bases form Watson–Crick type hydrogen bonds to the guanosine bases in the latter crystals (10), with one 9-aminoacridine stacked on the exterior and the other molecule intercalated into the dinucleoside duplex.

Alden and Arnott have put forward a model based on stereochemical principles for the intercalation of proflavine (Fig. 1) into B-DNA (11). The model proposes a change in the torsion angles about C4'-C5' and O4-P backbone bonds to a trans conformation and a change in sugar puckers to C3' endo-(3'-5')-C2' endo configuration at the intercalation site (11). The B-DNA unwinds by ~18° for each proflavine intercalated into the duplex for this model (11).

We have recently reported on the melting transition of the dC-dC-dG-dC duplex and the dC-dG-dC-dC duplex by monitoring the base and sugar protons and backbone phosphorus nuclear magnetic resonances at superconducting fields (12). In contrast to DNA, these self-complementary tetranucleotide duplexes exhibit narrow resonances which can be monitored through the duplex-to-strand transition (12).

We report below on the 360 MHz proton nuclear magnetic resonance (NMR) spectra of the nonexchangeable protons of the proflavine + dC-dC-dG-dC complex and the proflavine + dG-dG-dC-dC complex between 5 and 95°C at high nucleotide/dye ratios in aqueous solution. The resonances of the mutagen and those of the tetranucleotide can be monitored through the melting transition of the complex. We shall demonstrate intercalation of proflavine into the tetranucleotide duplex and utilize the mutagen complexation shifts to deduce approximate overlap geometries between the proflavine ring and nearest neighbor base pairs at the intercalation site. The relative specifity of proflavine complexation at pyrimidine-(3'5')-purine and purine-(3'5')-pyrimidine sites has been evaluated from a comparison of the binding curves monitored at the mutagen UV absorbance band at 444 nm on addition of dC-dC-dG-dG and dG-dG-dC-dC tetranucleotide duplexes.

EXPERIMENTAL

The dC-dC-dG-dG (lot no: 634-61) and dG-dG-dC-dC (lot no: 634-62) tetranucleotide sequences were purchased as lyophilized ammonium salts from Collaborative Research, Waltham, MA.

Proflavine was purchased from Sigma Chemical Co. Sample concentrations are based on an extinction coefficient (444 nm) of 4.1 × 10^4 M^-1 cm^-1. Actinomycin D (Merck and Co.) and propidium diiodide (Sigma Co.) sample concentrations are based on extinction coefficients of (425 nm) 2.35 × 10^4 M^-1 cm^-1 and (494 nm) 5.85 × 10^4 M^-1 cm^-1, respectively.

High resolution proton NMR spectra were recorded on the Fourier transform mode on a Bruker XH-360 NMR spectrometer interfaced to a Nicolet BNC-12 computer. The chemical shifts are referenced relative to internal standard 2,2-dimethyl-2-silapentane-5-sulfonate. The temperature of the probe was calibrated with methanol and ethylene glycol.

UV visible spectra were recorded on a Cary 118C spectrometer. The tetranucleotide sample (passed twice through Biogel P2 columns to remove salt) concentrations are based on

Abbreviation: NMR, nuclear magnetic resonance.
extinction coefficients (in strands) at 70°C (no added salt) of
2.90 × 10^4 M⁻¹ cm⁻¹ for dC-dC-dG-dG, 3.05 × 10^4 M⁻¹ cm⁻¹
for dG-dG-dC-dC, and 3.03 × 10^4 M⁻¹ cm⁻¹ for dC-dG-dC-dG
at 260 nm.

RESULTS AND DISCUSSION

Proflavine + dC-dC-dG-dG Complex (Nucleotide/Dye = 28). The 360 MHz Fourier transform NMR spectrum of
proflavine + dC-dC-dG-dG complex (nucleotide/dye = 28;
10 mM strand concentration), between 5.5 and 8.5 ppm, in 0.1
M phosphate/D₂O at pH 6.9 at 55.5°C is presented in Fig. 2
upper. Since the tetranucleotide is in excess, the stronger signals
correspond to the nucleic acid and include the two guanine H-8
singlets (7.5-8.0 ppm), the two cytosine H-6 doublets (7.5-8.0
ppm), the two cytosine H-5 doublets (5.5-6.0 ppm), and theour sugar (H-1') triplets (5.5-6.5 ppm), while the weaker signals
designated by asterisks) correspond to the resonances of
the proflavine ring. We assign the proflavine singlet at higher
field to H₈ since it exhibits twice the area of the proflavine
singlet at lower field, which is assigned to H₈ (see Fig. 1 for
nomenclature). Resonances H₆ and H₇ are doublets, and we
tentatively assign the lower field doublet to H₇ since it is closer
to the central aromatic ring (larger deshielding ring current
effect) and meta to the amino group, while we assign the higher
field doublet to H₆ since it is further from the central aromatic
ring and ortho to the amino group (13, 14).

The mutagen and nucleic acid resonances are well resolved
for the proflavine + dC-dC-dG-dG complex (nucleotide/dye
= 28) (Fig. 2 upper). The temperature dependence of the
chemical shifts of the nucleic acid base protons and mutagen
protons in 0.1 M phosphate solution are plotted in Fig. 2 lower.
In the presence of excess tetranucleotide, the base and sugar
nucleic acid protons monitor a averaged width of the muta-

Table 1. Upfield complexation shifts at 26°C at nucleic acid base
and sugar protons of dC-dC-dG-dG and of dG-dG-dC-dC in the
presence of proflavine at nucleotide/dye ratio = 12

<table>
<thead>
<tr>
<th>Resonance*</th>
<th>Complexation shift, ppm‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dC-dC-dG-dG</td>
</tr>
<tr>
<td>G(H-8)</td>
<td>0.045 (7.855)</td>
</tr>
<tr>
<td>G(H-8)</td>
<td>0.055 (7.885)</td>
</tr>
<tr>
<td>C(H-6), int.</td>
<td>0.10 (7.755)</td>
</tr>
<tr>
<td>C(H-6), int.</td>
<td>0.09 (7.757)</td>
</tr>
<tr>
<td>C(H-5), term.</td>
<td>0.15 (5.97)</td>
</tr>
<tr>
<td>C(H-5), int.</td>
<td>0.085 (5.815)</td>
</tr>
<tr>
<td>(H-1')</td>
<td>0.09 (6.185)</td>
</tr>
<tr>
<td>(H-1')</td>
<td>0.105 (6.045)</td>
</tr>
<tr>
<td>(H-1')</td>
<td>0.045 (5.73)</td>
</tr>
<tr>
<td>(H-1')</td>
<td>0.16 (5.73)</td>
</tr>
</tbody>
</table>

* Differentiation between terminal (term.) and internal (int.) base
pairs is established in ref. 12.
† The values in parentheses correspond to the chemical shifts at 26°C,
in the absence of proflavine.

The chemical shifts of the nucleic acid (O) and proflavine (●) resonances between 0° and 100°C.

The chemical shifts of the nucleic acid (O) and proflavine (●) resonances between 0° and 100°C.

Fig. 2. The proflavine + dC-dC-dG-dG (nucleotide/dye = 28;
10 mM strand concentration) complex in 0.1 M phosphate/0.001 M
EDTA/D₂O at pH 6.9. (Upper) The 360 MHz proton NMR spectrum
(4.5-8.5 ppm, resolution improved by the application of convolution
filter function) at 55.5°C. (Lower) The chemical shifts of the nucleic
acid (O) and proflavine (●) resonances between 0° and 100°C.
The overlap contours for intercalation of proflavine at dC-dG site according to (A) the model of Alden and Arnott (11) and (B) a modified version of this model involving a greater overlap between the proflavine ring and base pair rings proposed in this paper. The symbols O, ◦, △, and □ represent the positions of H_a, H_b, H_c, and H_d ring protons of proflavine, respectively.

Proflavine Complexation Shifts. The upfield shifts of ~0.5 ppm (H_d), ~0.7 ppm (H_b), ~0.85 ppm (H_c), and ~0.7 ppm (H_a) for the proflavine ring protons on complex formation with the dC-dC-dC-dC duplex and with the dG-dG-dG-dG duplex at high nucleotide/dye ratios are consistent with intercalation of the mutagen between base pairs. These upfield shifts of the proflavine resonances primarily reflect ring current contributions (15) from base pairs adjacent to the intercalation site.

Overlap Geometry at Intercalation Site. We have evaluated the overlap geometries for proflavine intercalated at dC-dG, dG-dC, and dC-dC sites in the tetranucleotide duplexes based on the projections published by Alden and Arnott for proflavine intercalated into DNA (11). As an illustration, we have drawn the overlap contours for intercalation of proflavine at the dC-dG site (Fig. 4A). It becomes readily apparent that the base pairs are closer to H_d than they are to H_a, H_b, and H_c for this model. This would result in a much larger upfield shift on complexation at H_d compared to H_a, H_b, and H_c (Table 2), contrary to the experimental proflavine complexation shifts (Figs. 2 and 3) which are similar at H_b, H_c, and H_d and somewhat smaller at H_a. The NMR data require that there be much greater overlap between the base pairs and the proflavine ring, such as is shown in Fig. 4B, so that the base pairs are roughly equidistant from each other.

Proflavine + dG-dC-dC Complex (Nucleotide/Dye = 30). We have undertaken parallel NMR studies on the proflavine + dG-dC-dC complex (nucleotide/dye = 30; 10 mM strand concentration) in 0.1 M phosphate/D_2O at pH 6.9. The 360 MHz proton NMR spectrum of the complex between 5.5 and 8.5 ppm is presented in Fig. 3 upper; the chemical shifts of the nucleic acid and proflavine protons as a function of temperature are plotted in Fig. 3 lower. The proflavine resonances move upfield by 0.5–0.8 ppm on going from free (high temperature) to nucleic acid bound (low temperature) structures, with transition midpoints of 54° ± 2°C. The base and sugar complexation shifts for the proflavine + dG-dG-dC-dC complex (nucleotide/dye = 12) are summarized in Table 1.

Fig. 3. The proflavine + dG-dC-dC-dC (nucleotide/dye = 30; 10 mM strand concentration) complex in 0.1 M phosphate/0.001 M EDTA/D_2O at pH 6.9. (Upper) The 360 MHz proton NMR spectrum (4.5–8.5 ppm, resolution improved by the application of convolution filter function) at 55.5°C. (Lower) The chemical shifts of the nucleic acid (0) and proflavine (●) resonances between 0° and 100°C.

Table 2. Calculated upfield chemical shifts of the proflavine ring protons due to ring current and atomic diamagnetic anisotropy contributions of nearest neighbor base pairs (15) for intercalation at dC-dC, dC-dG, and dG-dC sites.

<table>
<thead>
<tr>
<th>Intercalation site</th>
<th>Calculated upfield shifts, ppm</th>
</tr>
</thead>
<tbody>
<tr>
<td>H_a</td>
<td>H_b</td>
</tr>
<tr>
<td>Ref. 11</td>
<td>0.3</td>
</tr>
<tr>
<td>dC-dC</td>
<td>0.35</td>
</tr>
<tr>
<td>dC-dG</td>
<td>0.35</td>
</tr>
<tr>
<td>This paper</td>
<td>0.45</td>
</tr>
<tr>
<td>dC-dC</td>
<td>0.5</td>
</tr>
<tr>
<td>dG-dC</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The calculations were undertaken for the overlap geometry proposed by Alden and Arnott (11) (see Fig. 4A) and a modified geometry proposed in this paper (see Fig. 4B).
Our results favor the intercalation model proposed by Lerman, in which the positive ring nitrogen of proflavine lies centrally over the base pairs and the amino side chains are directed towards the sugar-phosphate backbone (2). The alternate intercalation model of Pritchard et al. proposes that the proflavine lies over successive bases in the same polynucleotide strand with the proflavine positive ring nitrogen in close proximity to the polynucleotide phosphate of that strand (18).

The experimental observation of approximately equal chemical shifts at all the proflavine ring protons on intercalation into the tetranucleotide duplexes (Figs. 2 and 3) rules out a model for which the successive bases on the same polynucleotide strand can only partially overlap with the much larger plane of the proflavine ring (18).

**Nucleic Acid Complexation Shifts.** The upfield complexation shifts at the base and sugar resonances for proflavine complexes with the dC-dC-dG-dG duplex and the dG-dG-dC-dC duplex at nucleotide/dye = 12° at 26°C (Table 1) reflect, for the resonances of a given base pair, the difference between the ring current contribution due to the intercalating proflavine molecule (19) and the contribution from the neighboring base pair which is displaced from 3.4 to 6.8 Å as a result of intercalation. We observe complexation shifts to high field since the upfield ring current contributions at a distance of 3.4 Å from the proflavine ring (19) are greater than the corresponding contribution for a G+C base pair (15).

**Stabilization of Melting Transition.** The duplex-to-strand transitions of dC-dC-dG-dG and dG-dG-dC-dC tetranucleotide duplexes (20 mM strand concentration) exhibit midpoints for the chemical shifts plotted against temperature for internal base pair resonances of 37 to 42°C in 0.1 M phosphate solution (12). By contrast, the proflavine + dC-dC-dG-dG complex (nucleotide/dye = 28) and the proflavine + dG-dG-dC-dC complex (nucleotide/dye = 30) in 0.1 M phosphate (10 mM strand concentration) exhibit transition midpoints monitored by the mutagen resonances of 59.5° ± 3°C and 54° ± 2°C, respectively. These data demonstrate stabilization of the melting transitions of the tetranucleotide duplexes after intercalation of the mutagen proflavine between base pairs.

**Binding Specificity.** There has been considerable interest in the specificity of drug binding to nucleic acid duplexes at the dinucleoside monophosphate (20-26), self-complementary tetranucleotide (20, 27, 28), and hexanucleotide (29) duplex level in solution. At the dinucleoside phosphate level, the drug acts as a template on which the nucleic acid forms a miniature double helix (20-28). Since G+C sequences containing tetranucleotide (mM concentrations) form stable duplexes at low temperature in the absence of drugs, they serve as excellent models for the investigation of drug binding to stable nucleic acid duplexes (20, 27, 28). Further, the tetranucleotide duplexes contain different proportions of dG-dC, dC-dC, and dC-dC (equivalent to dG-dC) sites and are of great help in the elucidation of the specificity of drug binding at the stable duplex level (28).

The dC-dC-dG-dG and dG-dG-dC-dC self-complementary duplexes contain two dC-dC (equivalent to dG-dG) sites located between the terminal and internal base pairs. The former duplex contains a central dC-dG site (but no dG-dC site), while the latter duplex contains a central dG-dC site (but no dC-dG site), so that these two sequences are excellent models for differentiating pyrimidine-(3'-5')-purine specificity from pyrazine-(3'-5')-pyrimidine specificity associated with drug complexation.

We have monitored changes in the drug absorbance of actinomycin D (435 nm), propidium diiodide (494 nm), and
proflavine (430 nm) in 0.1 M phosphate solution, neutral pH, on gradual addition of dC-dC-dG-dC, dG-dC-dC-dC, and dC-dG-dC-dG at 21.5°C. Since proflavine and the other drugs aggregate in solution, we have worked with drug concentrations of 2.0 x 10^{-5} M for all three intercalating agents (5). We were able to obtain reproducible binding data for proflavine when the experiments were undertaken in dim red light.

A comparison of the tetranucleotide concentrations corresponding to a half-maximal change in the absorbance at 435 nm, demonstrates that actinomycin D binds more strongly to dG-dG-dC-dC and dC-dG-dC-dG compared to dC-dC-dG-dG in solution (Fig. 5A). This establishes a sequence specificity for complex formation of actinomycin D at purine-(3'-5')-pyrimidine sites at the stable duplex level, in agreement with earlier conclusions at the dinucleoside monophosphate level (20–23) and tetranucleotide (20, 27) and hexanucleotide duplex levels (29).

The binding curves for the complex of propidium diiodide with the tetranucleotide sequences demonstrate that propidium diiodide binds more strongly to dC-dC-dC-dG and dC-dG-dC-dG than to dG-dG-dC-dC in solution (Fig. 5B). This establishes a sequence specificity for complex formation of propidium diiodide at pyrimidine-(3'-5')-purine sites at the stable duplex level. Earlier studies had demonstrated a similar specificity for ethidium bromide at the dinucleoside monophosphate (24–26) and tetranucleotide duplex levels (28).

We can compare the tetranucleotide concentration corresponding to a half-maximal change in the absorbance for actinomycin D (435 nm) and for propidium diiodide (494 nm) on the addition of dC-dG-dC-dC, which contains both dC-dG and dG-dC binding sites (Fig. 5). It is readily apparent that actinomycin D binds more strongly to this tetranucleotide duplex than does propidium diiodide. Both antibiotics contain an intercalating chromophore so that the additional stabilization in the case of actinomycin D must reflect the hydrogen bonding and hydrophobic interactions between the two pentapeptide lactone rings of the antibiotic and the base pair edges and sugar-phosphate backbone of the tetranucleotide duplex.

We have monitored the binding of proflavine to the G+C tetranucleotide sequences by following the mutagen absorbance at 444 nm on addition of dC-dC-dG-dG, dG-dC-dC-dC, and dC-dG-dC-dG under dim red light (Fig. 5C). Proflavine binds more strongly to dC-dC-dG-dG and dC-dG-dC-dG than to dG-dG-dC-dG in solution. This establishes a sequence specificity for intercalation of proflavine at pyrimidine-(3'-5')-purine sites at the stable duplex level.

X-ray studies on the complexes between 9-amino-acridines and self-complementary dinucleoside phosphates demonstrate that the dyes intercalate between the miniature duplex of pyrimidine-(3'-5')-purine sequence (10) but stack with purine-(3'-5')-pyrimidine sequence (9).

The costs of publication of this article were defrayed in part by the payment of page charges from funds made available to support the research which is the subject of the article. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.