Spectral and physical characterization of the inverted terminal repeat DNA structure from adenoassociated virus 2

Jinsong Ren, Xiaogang Qu, Jonathan B. Chaires, James P. Trempe and John David Dignam.

Department of Biochemistry, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216-4505, USA and Department of Biochemistry and Molecular Biology, Block Health Science Building, Medical College of Ohio, 3035 Arlington Avenue, Toledo, OH 43614-5804, USA

Received January 25, 1999; Revised and Accepted March 10, 1999

ABSTRACT

An oligodeoxynucleotide (ODN) that includes elements found in secondary structures at the 5′- and 3′-ends of adenoassociated virus 2 virion DNA was synthesized by ligation of three overlapping ODNs. The most stable secondary structure was calculated to be branched, with a 61 bp duplex stem, terminating in a three-way junction with 9 bp arms. The electrophoretic mobility of the ODN is slower than expected for normal duplex DNA of the same size, suggesting a bent or branched conformation. CD spectra indicate that the ITR structure is largely B form DNA, although there is a slight blue shift compared to the spectra of the isolated stem and loop elements. Thermal melting experiments indicate that the hairpin is significantly more stable than the isolated stem and loop elements. Singular value decomposition of UV spectra obtained as a function of temperature indicates that four species contribute to changes in the spectra upon denaturation, indicating that the melting is not a simple two-state process. Characterization of the branched ODN by differential scanning calorimetry permits estimation of the enthalpy of melting by a model-free analysis, yielding ΔHcal = 614 kcal mol⁻¹. This value agrees with the enthalpy computed for the most stable secondary structure.

INTRODUCTION

Adenoassociated virus 2 (AAV2) belongs to the parvoviridae whose members contain a small (4000–5000 nt) single-stranded genomic DNA in the mature virion (1,2). This non-pathogenic virus has been the focus of efforts to develop a gene therapy vector, in part because of its ability to integrate into a specific site on human chromosome 19 (3,4). A hairpin-like structural element (inverted terminal repeat or ITR) at the ends of the linear virion DNA is essential both for replication of the virus and its capacity for site-specific integration (5). The duplex and hairpin nature of this structure was suggested in early physical and enzymatic studies of this element in virion DNA (6–9). The 4680 nt genome of AAV2 encodes two sets of overlapping polypeptides including four Rep proteins (Rep78, Rep68, Rep52 and Rep40) and three capsid proteins (10). The Rep proteins arise from the use of two transcription promoters (p5 and p19) and differential splicing (11–13), while the three capsid proteins arise from transcription from a single promoter with the two species resulting from differential splicing (13). The hairpin-like secondary structures (ITRs) at the ends (14) serve as initiation sites for synthesis of a complementary strand during viral DNA replication (15). The ITR structure in AAV2 consists of a three-way junction with two covalently closed hairpin arms. Sequence elements in the ITR structure are also essential for integration of the double-stranded replicative form DNA into a site on human chromosome 19 (3,4). Another feature of the ITR structure is a sequence element termed the terminal resolution site (trs) (16), which is cleaved in an endonucleolytic reaction by two of the Rep proteins (17,18), Rep68 and Rep78. AAV2 Rep78 and Rep68 bind specifically to a stem region of the structure and act on it both as a helicase and as a sequence-specific, ATP-dependent endonuclease. Although Rep78 and Rep68 have other functions in virion development and viral gene expression (1), one essential feature is the cleavage and unfolding of the hairpin structure during DNA replication, which allows for the resolution of dimeric intermediates that form during DNA replication (19).

We have examined a synthetic AAV2 hairpin structure, that is cleaved by Rep78/Rep68 to generate the mature ITR, by optical spectroscopic methods and by differential scanning calorimetry to obtain information concerning its shape in solution and the energetics of its unfolding.

MATERIALS AND METHODS

Materials

Oligodeoxyribonucleotides were synthesized commercially (Genosys Biotechnologies, Woodlands, TX) by the phosphoramidate method and purified by denaturing gel electrophoresis. T4 DNA ligase was obtained from New England Biolabs (Beverley, MA). The 53 nt loop of the hairpin (Fig. 1A, lane 4, and B, ‘loop’) was derived from a single oligonucleotide while the 61 bp stem (Fig. 1A, lane 5, and B, ‘stem’) was formed by annealing two complementary oligonucleotides.
Analytical procedures

Oligodeoxynucleotides (ODNs) and ligation products were examined by denaturing gel electrophoresis on 12% polyacrylamide gels containing 7 M urea in Tris–borate–EDTA buffer (20) or by native gel electrophoresis using gels of the same composition but lacking urea. Oligonucleotides used in UV melting and calorimetry experiments were dissolved in 8 mM NaHPO₄, 0.5 mM EDTA (pH 7) and exhaustively dialyzed against the same buffer. Mobility shift assays were performed as previously described (21) and analyzed on non-denaturing 8% polyacrylamide gels employing homogeneous Rep78 purified by a procedure that will be described elsewhere.

Spectroscopic analyses

CD spectra were recorded on a Jasco J500 spectropolarimeter (Jasco, Inc., Easton, MD) interfaced to an IBM personal computer. UV spectra and UV melting profiles were recorded on a Cary 3e spectrophotometer (Varian, Inc., Palo Alto, CA) equipped with a Peltier temperature controller and interfaced to a personal computer as described previously (22). In UV melting experiments, absorbance at 260 nm was monitored continuously with a heating rate of 0.2 °C min⁻¹. Singular value decomposition (SVD) of the data (23,24) from UV melting experiments was performed using the program MATLAB (The Mathworks, Inc., Natick, MA) as described by Haq et al. (22). Molar extinction coefficients at 260 nm of oligonucleotides were estimated from their UV absorption spectra by the method of Hirshman and Felsenfeld (25) or by the method of Seinor et al. (26). The following molar extinction coefficients at 260 nm were found for the ITR structure and its pieces: folded ITR, 7850 M(P)⁻¹ cm⁻¹; stem, 6900 M(P)⁻¹ cm⁻¹; loop, 4450 M(P)⁻¹ cm⁻¹.

Calorimetry

Scanning calorimetry was performed in a Microcal MC2 Differential Scanning Calorimeter (Microcal, Inc., Northampton, MA), using v.2 of the manufacturer’s software for instrument control and data acquisition.

Fluorescence polarization

The effective molar volumes of the hairpin, the stem and the loop were determined from Perrin plots of polarization data (27). Bisethidium (Molecular Probes, Eugene, OR) was used as an extrinsic probe and was added to the samples at 0.15 µM to give a 1:1 complex. Fluorescence intensity at 612 nm was measured parallel and perpendicular to the plane of polarization after excitation at 573 nm. Polarization measurements were taken over a temperature range of 10–36 °C. Values for the viscosity of solutions as a function of temperature were taken from the Handbook of Chemistry and Physics (CRC Press, Inc., Boca Raton, FL).

Synthesis and purification of AA V2 hairpin structure

The hairpin secondary structure (Fig. 1A, lane 3 and B, ‘ITR’) was synthesized by ligating three oligonucleotides corresponding to nucleotides 1–33, 34–120 and 121–165 of the AA V2 ITR structure; the positions where oligonucleotides are ligated are indicated by arrows in Figure 1B (‘ITR’). Oligonucleotides 1–33 and 34–120 were 5’-phosphorylated. The oligonucleotides (10 nmol of each) were combined in a 500 µl reaction containing 50 mM Tris–HCl (pH 7.5 at 25 °C), 50 mM NaCl and 0.5 mM EDTA, heated to 95 °C and allowed to cool slowly in a 3 l water bath over 3–5 h. The annealed oligonucleotides were ligated in a 2 ml reaction containing 50 mM Tris–HCl (pH 7.8), 10 mM MgCl₂, 10 mM dithiothreitol, 1 mM ATP and 1000 Weiss units of T4 DNA ligase. Ligation reactions were incubated at 16°C for 12 h and the reactions were terminated by the addition of EDTA to 20 mM and heating to 65 °C for 20 min. NaCl was added to 0.25 M and the reaction products were precipitated with 2.5 vol of 95% ethanol. The oligonucleotides were dissolved in 500 µl 50 mM Tris–HCl (pH 7.5 at 25 °C), 50 mM NaCl and 0.5 mM EDTA, and were purified by gel filtration chromatography on a Superose 12 HPLC column (1 cm × 30 cm, Pharmacia Biotech, Inc., Piscataway, NJ) equilibrated in the same buffer. The column was eluted at 0.4 ml/min. Fractions from Superose 12 chromatography were analyzed by denaturing gel electrophoresis and fractions containing the hairpin DNA were combined, precipitated with ethanol and rechromatographed on the same Superose 12 column equilibrated in 6 mM Na₂HPO₄, 2 mM NaH₂PO₄ and 0.5 mM EDTA. Fractions were analyzed by gel electrophoresis, and those free of contaminating oligonucleotides were combined and used for spectroscopic analysis and differential scanning calorimetry.

Computations

The most stable secondary structure was computed using the ‘mfold DNA’ algorithm developed by Michael Zuker (28) available on the worldwide web at http://www.ibc.wustl.edu/~zuker/SeqMan. The thermodynamic parameters from the John SantaLucia, Jr laboratory were used for the computation (29). DNA curvature and bendability were computed using the algorithms (30) available on the DNA Tools web site (http://www2.icgeb.trieste.it/~dna).

RESULTS

The most stable secondary structure computed for the folding of the 165 nt oligonucleotide is shown in Figure 2. The ‘mfold DNA’ algorithm predicts a branched structure, with a 61 bp duplex stem terminating in a three-way junction with two hairpin arms, each with a 9 bp stem and a 3 nt loop. This structure is similar to that proposed in other studies with the three bases at the ends of each arm not being hydrogen bonded. Assuming a two-state model, the algorithm predicts (in 1 M NaCl) the following values for the
folding of the oligonucleotide: \( T_m = 103.8^\circ C, \Delta G^\circ = -114.1 \) kcal mol\(^{-1}\), \( \Delta H^\circ = -647.2 \) kcal mol\(^{-1}\), \( \Delta S^\circ = -1718.5 \) cal (K mol\(^{-1}\)).

The results of native gel electrophoresis of oligonucleotides used in this study are shown in Figure 1A and correspond to elements of the ITR shown diagrammatically in Figure 1B. The electrophoretic mobility of the 61 bp stem (Fig. 1, lane 5) was slightly less than expected compared to duplex DNA markers (80 observed versus 61 predicted) while the 53 base loop (Fig. 1, lane 4) migrated near its expected size compared to a 32 bp marker shown in lane 2. However, the hairpin (Fig. 1A, lane 3, and B, 'ITR') migrated with a mobility corresponding to a size (~150 bp) that is significantly greater than that expected based on the size of the hairpin oligonucleotide (80–82 bp). All the DNA structures migrated at their expected sizes when examined by denaturing gel electrophoresis (not shown). Repeated heating of the hairpin DNA (ITR) and the loop did not alter their mobility on non-denaturing gels while the stem DNA migrated faster after heating as expected for the generation of single-stranded ODNs from duplex DNA. Hairpin DNA and the stem structure both bound to Rep78 as detected by gel mobility shift assays, while binding to the loop structure was not detected (not shown). The electrophoretic properties of the hairpin indicate that its structure deviates significantly from linear duplex DNA in a manner suggestive of a bent or branched conformation; DNAs containing a bend typically have electrophoretic mobility that is less than expected based solely on the criterion of size (30,31).

The structural differences among the whole ITR, the loop and the stem were examined further using spectroscopic methods. UV spectra (not shown) of the three structures were generally consistent with folded or duplex structures. Circular dichroic (CD) spectra are shown in Figure 3 for the three structures. All three oligonucleotides show CD spectra that are characteristic of B form DNA. The intact branched oligonucleotide (curve A in Fig. 3), however, has a CD spectrum with distinctive differences from that of the 61 bp duplex stem (Fig. 3B). The positive maximum is slightly blue shifted, and the intensity of the molar ellipticity is slightly decreased, as might be expected for a structure containing unpaired loop regions.

The stabilities of the hairpin and hairpin derived structures were examined in thermal denaturation studies as shown in Figure 4. The stem (curve B) melts with a single transition over ~20°C (\( T_m = 41.2^\circ C \)), while the loop (curve C) exhibits a broad transition over ~40°C (\( T_m = 46.2^\circ C \)). The melting curve of the whole ITR oligonucleotide is sharper and the structure is significantly more stable, showing a transition near 75°C. The thermal melts for the loop and whole ITR structures were reversible, giving the same melting curves (within experimental error) in a second melting experiment following cooling. Van’t Hoff enthalpy values for the melting of the stem, loop and whole ITR structures were estimated from derivative melting curves (not shown) using curve widths at the half heights and standard formulae (32). The results (which assumes a two-state transition) yields \( \Delta H_{VH} \) values (kcal mol\(^{-1}\)) of 23.0, 51.0 and 96.0 for the loop, stem and whole ITR structures, respectively. The results of the thermal denaturation studies are summarized in Table 1.

<table>
<thead>
<tr>
<th>Structure</th>
<th>( T_m ) (°C)</th>
<th>( \Delta H_{VH} ) (kcal mol(^{-1}))</th>
<th>( P_0 )</th>
<th>( V_0 ) (cm(^3) mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duplex stem</td>
<td>41.2</td>
<td>51.0</td>
<td>0.41</td>
<td>41.4</td>
</tr>
<tr>
<td>Loop</td>
<td>46.2</td>
<td>23.0</td>
<td>0.49</td>
<td>15.7</td>
</tr>
<tr>
<td>Whole ITR</td>
<td>75.6 (74.1)</td>
<td>96.0 (126.0)</td>
<td>0.42</td>
<td>52.2</td>
</tr>
</tbody>
</table>

*Obtained from UV melting studies (Fig. 4). Italicsized values in parentheses for the whole ITR were obtained from DSC experiments (Fig. 6).

To analyze the denaturation of the ITR structure in greater detail, and to rigorously test the two-state assumption used to extract thermodynamic data from melting curves, UV spectra were collected from 220 to 340 nm at temperatures ranging from
Figure 4. Thermal melts of the ITR (curve A), the stem (curve B) and the loop (curve C). Data is presented as fraction single strand concentration, calculated from absorbance data, as a function of temperature.

40 to 90°C at 1°C intervals. These results were analyzed by the method of SVD by constructing a data matrix consisting of absorbance values at different wavelengths and temperatures (22). SVD decomposes the data matrix into the product of three matrices: a U matrix containing the basis spectra, a diagonal S matrix containing the singular values and a V matrix containing amplitude vectors. These three matrices contain information that enables the number of significant spectral species that contribute to the family of spectra over the course of the melting transition (33). The results of this analysis revealed four significant singular values, as demonstrated by the magnitudes of singular values (Fig. 5B), autocorrelation of the columns of the U and V matrices, and by non-random nature of the residual plots (Fig. 5A). The four basis spectra derived from this analysis are shown in Figure 5C. The singular values and the autocorrelation values for the U and V matrices show a sharp decline between the fourth and fifth basis spectra. Singular values of ≤0.017 or autocorrelation values <0.8 were taken to indicate a random, non-significant spectral component. The analysis indicates that at least four species contribute to spectra seen in the thermal melt, suggesting the existence of the native structure, a denatured state and at least two intermediate species. Invoking the two-state assumption to calculate ΔH_{VH} values is clearly not valid.

The stability of the ITR structure was examined by differential scanning calorimetry in order to obtain thermodynamic data for the melting transition as directly as possible and by a model-free analysis that does not invoke the two-state assumption. Figure 6 shows the primary data from one determination. The DSC thermogram was found to be completely reversible, with rescans after cooling superimposing upon the original trace. From the maximum in the peak shown in Figure 6, we find that T_m = 74°C.

Integration of the peak shown in Figure 6 yields, with no assumptions, an enthalpy estimate of 614.6 kcal mol⁻¹ for the denaturation of the folded oligonucleotide. By transforming the data of Figure 6 into a plot of Cp/K versus temperature, the entropy for the denaturation process may be obtained by integration, yielding a value of 1771.8 cal (mol⁻¹K)⁻¹. As judged from multiple determinations, the enthalpy and entropy estimates have an error of 5–10%. From the shape and width of the DSC
Figure 6. Results from a differential scanning calorimetry experiment of the ITR structure. Solution conditions: BPE buffer (6 mM Na₂HPO₄, 2 mM NaH₂PO₄, 0.5 mM EDTA, pH 7.0), 2.1 µM total strand concentration.

The curve shown in Figure 6, \( \Delta H_{\text{VH}} = 126.0 \text{ kcal mol}^{-1} \) may be calculated, assuming (incorrectly) two-state behavior. That value is in fair agreement (within \( \sim 30\% \)) with estimates obtained from optical melting studies. The ratio \( \Delta H_{\text{VH}}/\Delta H_{\text{cal}} = 126.0/614.6 = 0.2 \), a value significantly <1.0, indicating a violation of the two-state assumption, and that intermediate species are significantly populated over the course of the transition (32). These conclusions support the results of the SVD analysis.

Information concerning the size of the hairpin was obtained by examining the dependence of polarization of a complex of the hairpin with an extrinsic fluorophor, bisethidium. Values for the effective molar volume obtained for the stem, the hairpin and the loop structures from the results of polarization studies shown in the Perrin plot in Figure 7 are summarized in Table 1 along with the value for the arithmetic difference between the hairpin and the stem structure. That the effective molar volume of the hairpin is not the arithmetic sum of the two components suggests that the joining of the stem and loop results in a significant structural alteration.

DISCUSSION

We have examined the physical characteristics of an ODN whose sequence was based on an element in the termini of the virion DNA of AAV2 to gain insight into its structure and stability. This structure is cleaved by Rep78/Rep68 to generate the mature ITR found in virion DNA.

The most stable secondary structure of the 165 nt DNA fragment was calculated to be a branched structure, with a 61 bp duplex stem terminating in a three-way junction with two hairpin arms, each with a 9 bp stem and a 3 nt loop. The magnitudes of the calculated enthalpy and entropy values for the folding of the DNA into this structure are in excellent agreement with experimental values determined as directly as is possible by differential scanning calorimetry. The enthalpy values of 647.2 kcal mol\(^{-1}\) (calculated) and 614.6 kcal mol\(^{-1}\) (experimental) agree to within 5% as do the entropy values of 1718.5 cal (\(^{\circ}\text{K}\cdot\text{mol}^{-1}\)) (calculated) and 1771.8 cal (\(^{\circ}\text{K}\cdot\text{mol}^{-1}\)) (experimental). The excellent agreement of calculated and experimental values strongly supports the notion that the branched, three-way junction structure exists in solution. We acknowledge that the calculated and experimental values refer to different total Na concentrations and reference temperatures, but invoke the common assumption that enthalpy values for melting are largely independent of salt concentration and have negligible heat capacity changes.

Melting of the branched ITR structure is not a simple two-state process, as analysis of spectroscopic data by SVD reveals. SVD suggests that at least four spectral species contribute to changes in the UV spectra over the course of the melting transition. For a strictly two-state melting process, there should be but two spectral species, corresponding to the folded and denatured structures. From calorimetry experiments, we found that \( \Delta H_{\text{VH}}/\Delta H_{\text{cal}} = 0.2 \). That this ratio is significantly <1.0 is a clear indication that intermediate states are populated in the melting transition, which is a violation of the two-state assumption. What these intermediate states might be cannot be said from these studies, since kinetic studies of the unfolding of the ITR structure would be required to characterize the intermediates, a task beyond the scope of this report. However, the intermediates might well involve tertiary interactions between the arms at the three-way junction. Such tertiary folding might unfold before the melting of the duplex.
regions of the secondary structure shown in Figure 2, and might have unique spectral properties.

The results of electrophoresis also indicate that the hairpin assumes a conformation similar to bent DNA or a three-way junction, resulting in reduced electrophoretic mobility compared to normal duplex DNA of the same size. There is significant interest in unusual three- and four-way DNA structures since they figure prominently in a number of biological phenomena, including recombination and DNA replication. The three-way junctions at the ends of the virion DNA of AAV2 and related viruses differ from other three-way junctions that have been examined in that the entire structure is composed of a single strand, whereas most studies of model three-way junctions have been performed on structures formed from three separate strands. Having the termini of two of the arms covalently closed would appear to place constraints on structure at the junction of the three arms. Studies of three-way junctions indicate that under some conditions of ionic strength and magnesium ion concentration, the arms of a three-way junction are not structurally equivalent, the differences being reflected in sensitivity to a variety of chemical and enzymatic probes. In addition, there is preferential base stacking at the junction of such three-armed structures.

ACKNOWLEDGEMENTS

This work was supported by grants CA35635 (J.B.C.) and CA58743 (J.P.T.) from the National Cancer Institute (J.P.T.) and a grant from the Ohio Cancer Research Associates (J.D.D.).

REFERENCES