Preparation and melting of single strand circular DNA loops

David E. Wemmer* and Albert S. Benight

Department of Chemistry, BG-10, University of Washington, Seattle, WA 98195, USA

Received 6 August 1985; Revised 28 October 1985; Accepted 29 October 1985

ABSTRACT

A method for preparation of single strand DNA circles of almost arbitrary sequence is described. By ligating two sticky ended hairpins together a linear duplex is formed, closed at both ends by single stranded loops. The melting characteristics of such loops are investigated using optical absorbance and NMR. It is shown by comparison with the corresponding linear sequence (closed circle minus the end loops) that the effects of end fraying and the strand concentration dependence of the melting temperature are eliminated in the circular form. Over the concentration range examined (0.5 to 2.0 micromolar strands), the circular DNA has a monophasic melting curve, while the linear duplex is biphasic, probably due to hairpin formation. Since effects of duplex to single strands dissociation do not contribute to melting of the circular molecules (dumbbells), these DNAs present a realistic experimental model for examining local thermal stability in DNA.

INTRODUCTION

Thermal denaturation studies of short DNA restriction fragments (less than 500 base pairs in length) have clearly demonstrated that the helix-coil transition of these molecules is dominated by the dissociation of the duplex to single strands¹⁻⁴. For short fragments having a heterogeneous sequence, melting from the ends of the duplex is seen preceding complete dissociation. This preferential "end fraying" complicates physical studies aimed at evaluating thermodynamic parameters which govern base pair stability. For most oligonucleotides "end fraying" and the ensuing strand dissociation makes the mode of melting, and the melting temperature, strongly dependent on DNA strand concentration. Consequently, results from NMR measurements cannot be reliably compared with those from optical melting curves because the optimum DNA concentrations are very different in the two experiments. One way to eliminate the dependence of short DNA melting on the DNA concentration is by linking the duplex strands so they can never completely separate.

We have developed a general procedure for synthesizing covalently closed single strand loops of DNA, of almost arbitrary sequence, using ligation of

sticky end dimers. These are chosen to form stable duplexes in a central segment, with the loops on either end, as shown in Figure 1. Covalent closure of the DNA in this way almost completely removes "end fraying" and makes the melting independent of DNA concentration. These DNA circles provide a new approach for evaluating contributions of the internal molecular degrees of freedom to DNA melting. Similar circles of the alternating copolymer $d(AT)_n$ have been made and studied, though they were generated by closure of a single DNA strand, rather than the combination of two separate strands, and cannot be generalized for arbitrary sequences.

MATERIALS AND METHODS

DNA oligomers were synthesized using solid-phase phosphite triester chemistry⁸ on an Applied Biosystems automated DNA synthesizer. After cleavage of the DNA from the solid support it was purified using a sequence of TLC, hydroxylapatite and Sephadex chromatography. The DNA was then lyophilized and redissolved in appropriate buffers for experiments as described below. Phosphorylation of the 5' deoxyribose was carried out using T4 polynucleotide kinase. 750 micrograms of the 20 base pair DNA were reacted in 70mM Tris.HCl (pH=7.6), 10mM $MgCl_2$, 100mM KCl, 5mM DTT, 1.6 mM ATP using 400 units of kinase (Bethesda Research Laboratories) in a volume of 5 ml for 22 hours at 37°C. Subsequent ligation was carried out in 50mM Tris.HCl (pH=7.8), 10mM $MgCl_2$, 20mM DTT, 1.6mM ATP, 150 micrograms/ml BSA using 8000 units of T4 DNA ligase (New England Biolabs) in a total volume of 10 ml for 48 hours at 22°C. The product was purified by extraction twice with phenol, then twice with ether. The solution was then dialysed against a large volume of buffer, and desalted over a Sephadex G-15 column. Analysis of the ligated DNA was performed using electrophoresis on a 20% polyacrylamide, 0.75% bisacrylamide, 8M urea gel. Samples were ethanol precipitated, redissolved in 90% formamide, 1X TBE (tris-borate EDTA gel buffer) and heated to 100°C for 4 minutes, then chilled on ice and quickly loaded onto the gel. The gels were stained with Ethidium bromide to visualize the DNA. Ligated DNA showed marked retardation compared to unreacted material (see Figure 4). Ligation was quantitative, to the extent that no band was observed on the gel for unreacted molecules.

NMR studies of imino protons were carried out on a Bruker WM-500 NMR spectrometer. Spectra were collected in 90% H_2O /10% D_2O solutions, at DNA concentrations of 0.3 to 0.8 millimolar, using a Redfield 214 sequence⁹. 8k data points were collected with an 11400 Hz spectral width and a delay of 1 sec between acquisitions. Data were apodized with a Lorentz-Gaussian

function¹⁰ (LB = -6, GB = 0.125) before Fourier transformation. NOE experiments for assignment of resonances were carried out using an interleaved direct difference collection (8 scans on resonance irradiation followed by 8 scans off resonance subtracted), with an irradiation time of 0.7 sec, and a total time of ca. 1 sec between acquisitions. Optical melting curves were measured on a modified Acta CV double beam spectrophotometer interfaced to an IBM XT computer. Absorbance at 268 nm was monitored while constantly raising the temperature at 6.4°C/hour. Temperature and absorbance were sampled by the computer every 20 seconds. The sample temperature was determined with a platinum resistance thermometer sealed directly in the sample cuvette. Data were transferred to a VAX 11/780 for plotting.

RESULTS AND DISCUSSION

It is well known that hairpin loops form in DNAs which have complementary ends, but noncomplementary central sequences¹¹. The best studied of the loop sequences is dT₄, which has been examined in detail with a variety of "stem" sequences by Haasnoot et al.⁹, and using the distance geometry approach with NOE distances by Hare & Reid¹³. By extending one of the stem strands with a self-complementary sequence, we can form "sticky end"

Figure 1. Sequences of DNAs used in the present work, with a schematic drawing of a "sticky end" dimer as it exists in solution at low temperatures.

Figure 2. Imino and aromatic regions of the 500 MHz ^1H NMR spectrum of GATCTTTGATCGGCC (sequence I of Fig. 1) in 90% H_2O /10% D_2O solution at 10°C , with NOE difference spectra obtained from irradiation of peaks A and E, middle and top spectra respectively. The DNA concentration was about 1 mM in strands, 10 mM phosphate buffer, no added salt.

dimers of the hairpin loop sequence, Figure 1. To confirm that such dimers are formed in solution, we have studied the sequence 5'-GATCTTTGATCGGCC 3'. The imino proton spectrum is shown in Figure 2, together with NOE experiments which lead to assignments. The spectrum of a hairpin with four base pairs would consist of at most four lines. The observation of five sharp lines in the downfield region (the lowest field peak has intensity two, see also fig. 3 where these are partially resolved) immediately shows that the dimer is formed in solution, with two-fold symmetry as expected. The loop thymines give rise to broad peaks slightly upfield (ca 11 ppm) of the normal hydrogen bonded imino region, as has been observed previously. NOE experiments^{14,15}, Figure 2b and c, corresponding to saturation of peaks A and E, are also shown. The NOEs to aromatic protons in Figure 2b show that peak A corresponds to both of the A-T base pairs in the sequence, at positions 2 and 3, with neighboring base pairs corresponding to resonances D and E. Saturation of peak E, Figure

Figure 3. Temperature dependence of the imino region of GATCTTTTGATCGGCC, sample as in Figure 2. Peaks are identified with the position, counting in from the loop end shown in Figure 1 sequence I.

2c, shows that in addition to A it has C as a neighbor (the weak peak at D is due to irradiation spillover), identifying E as base pair 4. The broad peak at ca 8.4 ppm is an amino proton from this G-C base pair, while the sharp line at ca 8 ppm is the C2H from base pair 3, as expected⁸. By elimination peak D must correspond to base pair, 1, C to base pair 5, and B to base pair 6. The existence of base pairs 5 and 6 establishes unequivocally that a stable dimer is formed in solution, while the NOE from E to C shows that the two base pairs adjoining the "gap" in the backbone are stacked in solution making effectively one 12 base pair helix with loop ends. The behavior of the imino spectrum with increasing temperature is shown in Figure 3. It is clear from the broadening of peaks B, C and E that base pairs at or near the "gap" melt out first, with those in the stem of the hairpin melting about 5-10°C higher. The effective melting temperature of the four base pair sticky end is much higher than would ordinarily be observed for a tetramer, the extra stabilization

Figure 4. Gel electrophoresis of DNAs: lane A is a 21 base pair reference (CGTATCACCGCAAGGGATACG + complement); lane B is the ligated form of sequence III, Figure 1; lane C is the unreacted form of sequence III, Figure 1; lane D is the 16 base pair linear version of the ligated sequence, sequence IV, Figure 1; lane E is a 17 base pair reference (the 17 central base pairs of the lane A sequence). The gel origin is at the top of the photograph.

clearly arising from the additional stacking with the neighboring base pairs.

After the phosphorylation and ligation reactions, on sequence III of Figure 1, the mobility of the DNA on polyacrylamide gels is significantly altered, Figure 4. The new position on the gel has lower mobility than the original 20 base hairpin loops, but between 17 and 21 base pair linear DNAs, consistent with our conclusion that a covalently closed circle of 40 bases, making a 16 base duplex with loop ends, is formed. The 16 base pair linear DNA with the same sequence shows anomalously high mobility on the gel, probably due to formation of hairpins, see also Fig. 6. The imino proton spectrum of the circular DNA is shown in Figure 5. For the two fold symmetric sequence, predicted to have sixteen central bases pairs, we indeed observed eight resonances. These have been assigned by analogy to sequences II (for which base pairs 1-4 should be essentially identical) and sequence IV (for which base pairs 4-8 should be nearly identical) for which sequential NOE assignments were obtained, as for sequence I described above. The behavior of the spectrum with increasing temperature is also shown in Figure 5. The behavior is quite distinct from that seen for the unligated dimer in solution. The first resonance to show broadening from solvent exchange is

Figure 5. Temperature dependent imino proton spectra of the ligated DNA prepared from the sequence III in Figure 1. The comparable spectra of the self-complementary linear version of the same sequence are shown in Figure 7. The DNA concentration was about 0.3mM, with 10mM phosphate buffer, pH=6.8 and 200mM NaCl. Small glitches from the spectrometer occur between 13.6 and 14 ppm.

from the end base pair nearest the loop, and becomes measurable only above 40°C. For a normal oligomer exchange broadening may well extend several base pairs in from the end at the same temperature (see Figure 7)^{16,17,18}. At yet higher temperature broadening of other resonances occurs. Notably all of the AT resonances, including those at the center of the molecule broaden before the bulk of the GC resonances do. Such non-sequential "melting" (or more properly local opening) of AT base pairs has been observed previously in normal oligomers^{19,20}, though end fraying has complicated the analysis in some cases. The UV absorbance curve measured at 268 nm is shown in Figure 6. The transition has a midpoint of 84°C at low salt, (data not shown, 2mM NaCl, 10mM phosphate buffer pH=6.9) which increases as expected with increasing salt concentration to ca 92°C at 200mM salt. These transition

Figure 6. UV melting curves (baseline uncorrected) for the circular (right) and linear (left) forms of the sequence GTATCCGTACGGATAC (sequence IV of Fig. 1). The biphasic melting of the linear form probably represents formation, then melting, of a hairpin conformation of this self-complementary sequence. The DNA concentration was about 1 micromolar, 10mM phosphate buffer, 200mM NaCl in each case.

temperatures were determined to be independent of DNA concentration over the range of 0.5 to 2.0 micromolar in DNA strands. The absorbance changes seen are fully reversible, and there is no evidence for degradation of the DNA seen on gels run after melting experiments.

For comparison we have also examined the same central 16 base pairs, as a normal two strand duplex. The imino proton spectrum is shown in Figure 7 as a function of temperature. The terminal base pair is clearly much less stable against exchange than in the circular DNA. At intermediate temperatures the AT base pairs at positions 2,3, and 4 show clear effects of "end fraying". The temperature at which broadening of all resonances is observed has also been reduced by ca 10°C. The melting behavior detected optically for this sequence at much lower strand concentrations is also more complex showing two transitions (Figure 6). The first, at low temperature, is almost certainly a transition from duplex to single strand hairpins, as has been described in detail for CGCGTATACGCC¹⁹, and has been seen in other sequences as well²⁰. The second transition corresponds to the melting of the hairpins into random coils. Without added salt there are still two transitions, though the relative absorbance change is different, and the temperature of each transition is decreased. A detailed statistical thermodynamic analysis of the melting of the circular form, and its comparison with the same central

Figure 7. Imino proton spectra of the linear DNA hexadecamer, GTATCCGTACGGATAC (sequence IV of Fig. 1), as a function of temperature. The melting behavior, particularly the presence of end fraying, is to be contrasted with the covalently closed circular form with the same central sequence, Figure 5. The DNA concentration was 0.5mM, about 30 fold more concentrated than for the UV melt shown in Figure 6, with 10mM phosphate buffer and 200mM NaCl.

sequence, but linear, will be presented in a separate publication²³.

Although in the present work we have described the preparation of a two-fold symmetric DNA circle, this is not an essential requirement of the method described. By making the "sticky end" region of the molecule non-selfcomplementary, and adding a second strand which has the appropriate complementary dangling end, "hetero" duplexes will form in solution which can be ligated after phosphorylation. Since the final circle is made from the two strands, it is possible to generate almost any sequence desired. The end loops must be noncomplementary with other segments of the molecule to prevent alternate locations for hairpin formation. In studies of hairpin loops we have used TTTT and TAAT with good success (unpublished results), while others

have used other sequences, and increased the loop length to 5 or 6 bases¹², There is no evidence that the choice of the loop sequence significantly affects either the structure or dynamics of the duplex part of the molecule. Shorter loops, such as TTT may slightly destabilize the end of the stem segment¹², though this has not yet been examined in detail. Sequences may also be divided into two pieces of very different length to construct sequences with central symmetric sequences, but with asymmetric ends. The hairpins with dangling ends described here also offer the opportunity to study the association of "sticky end" sequences in solution, without requiring four strands in the complex, as would be required in using normal duplex oligomers.

ACKNOWLEDGEMENTS

We would like to thank Drs. Patrick Chou and Foon Lee for synthesis and purification of the DNA used in these experiments, and Dr. Dennis Hare for help with interfacing the computer to the spectrophotometer. We also thank Profs. J.M. Schurr and B.R. Reid for helpful discussions. Support through NIH Program Project Grant GM 32681 (D.E.W. and J.M. Schurr) and instrumentation grants from the Murdock Foundation, NIH (GM 2874-0151) and NSF (PCM 80-18053) are gratefully acknowledged.

*Present address: Department of Chemistry, University of California, Berkeley, CA 94720

REFERENCES

1. Benight, A.S., Wartell, R.M. and Howell, D.K. (181) *Nature* **289**, 203.
2. Hillen, W., Goodman, T.C., Benight, A.S., Wartell, R.M. and Wells, R.D. (1981) *J. Biol. Chem.* **256**, 2761.
3. Benight, A.S. and Wartell, R.M. (1983) *Biopolymers* **22**, 1409.
4. Wartell, R.M. and Benight, A.S. (1985) *Physics Reports*, in press.
5. Olivera, B.M., Scheffler, I.E. and Lehman, I.R. (1968) *J. Mol. Bio.* **36**, 275.
6. Scheffler, I.E., Elson, E.L. and Baldwin, R.L. (1970) *J. Mol. Bio.* **48**, 145
7. Baldwin, R.L. (1971) *Accts. Chem. Res.* **4**, 265.
8. Chou, S.-H., Hare, D.R., Wemmer, D.E. and Reid, B.R. (1983) *Biochem.* **22**, 3037.
9. Redfield, A.G., Kunz, S.D. and Ralph, E.K. (1975) *J. Magn. Res.* **19**, 114.
10. Ferrige, A.G. and Linden, J.C. (1978) *J. Magn. Res.* **31**, 337.
11. Haasnoot, C.A.G., den Hartog, J.H.J., de Rooij, J.F.M., van Boom, J.H. and Altona, C. (1980) *Nuc. Acids. Res.* **8**, 169.
12. Haasnoot, C.A.G., de Bruin, S.H., Berendsen, R.G., Janssen, H.G.J.M., Binnendijk, T.J.J., Hilbers, C.W., van der Marel, B.A. and van Boom, J.H. (1983) *J. biomolec. Stereodyn.* **1**, 115.
13. Hare, D.R. and Reid, B.R. in preparation.
14. Roy, S., Papstauros, M.Z. and Redfield, A.G. (1982) *Biochem.* **14**, 6081
15. Hare, D.R. and Reid, B.R. (1982) *Biochem.* **21**, 1835.
16. Patel, D.J. and Hilbers, C.W. (1975) *Biochem.* **14**, 2651.

17. Kan, L.S., Borer, P.N., and T'so, P.O. (1975) *Biochem.* 14, 4864.
18. Early, T.A., Kearns, D.R., Hillen, W. and Wells, R.D. (1981) *Biochem.* 20, 3756.
19. Cheung, S., Arndt, K. and Lu, P. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3665.
20. Chou, S.-H., Wemmer, D.E., Hare, D.R. and Reid, B.R. (1984) *Biochem.* 23, 2257.
21. Wemmer, D.E., Chou, S.-H., Hare, D.R. and Reid, B.R. (1985) *Nuc. Acids. Res.* 13, 3755.
22. Marky, L.A., Blumenfeld, K.S., Kozlowsk, S. and Breslauer, K.J. (1983) *Biopolymers* 22, 1247.
23. Benight, A.S. and Wemmer, D.E., in preparation.