Nucleic Acids Research, 1994, Vol. 22, No. 21 4361–4363

Two-base DNA hairpin-loop structures *in vivo*

Angus Davison and David R.F.Leach*
Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, UK

Received August 17, 1994; Revised and Accepted September 24, 1994

ABSTRACT

***In vitro* studies have revealed that DNA hairpin-loops usually contain four unpaired bases. However, a small subset of sequences can form two-base loops. We have previously described an *in vivo* assay that is sensitive to tight loop formation and have set out to test whether DNA sequences known to form two-base loops *in vitro* also form tight loops *in vivo*. It is shown that the sequences 5'dCNNG and 5'dTNNA behave as predicted if they favour two-base loop formation *in vivo*, a result that is consistent with previously described *in vitro* studies. The ability of specific DNA sequences to form tight loops *in vivo* has implications for their potential to form transient structures involved in gene regulation, recombination and mutagenesis.**

INTRODUCTION

It is increasingly evident that alternative DNA secondary structures including hairpin-loops play a role in gene regulation (1–3). Hairpin-loops are also known to form as intermediates in immunoglobulin VDJ recombination (4). Additionally, long palindromes and certain repeated sequences associated with human disease may be unstable as a result of slippage during replication through hairpin-loop structures (5, 6). However, no evidence exists for the manner in which DNA sequence affects loop structure and folding *in vivo*. In contrast, several studies of hairpin-loop structure have been performed *in vitro* (7–10). In particular, thermal transition (UV melting) and molecular modelling studies of the inverted repeat oligonucleotides 5'dGGTA(CGCG)TACC, 5'dGGTA(CGAG)TACC and 5'dGGTA(AGCG)TACC (central four bases shown in parentheses for clarity) have suggested that the first two sequences can adopt two-base hairpin-loops while the last forms a loop with four unpaired bases (11, 12). Likewise, NMR and thermal transition studies of the inverted repeat oligonucleotides 5'dATCCTA(CTTG)TAGGAT, 5'dATCCTA(TTTA)TAGGAT, 5'dATCCTA(GTTC)TAGGAT and 5'dATCCTA(ATTT)TAGGAT have been interpreted to indicate that only the former pair can form two base hairpin-loops (13, 14). Until now, the significance of these results to the dynamics of DNA *in vivo* has not been known. Here, we use a plaque assay (15) to assess the effects of the same loop sequence differences on the viability of a λ phage containing a long palindrome. Our previous work has revealed that the DNA sequence at the centre of a long palindrome affects plaque size on an *sbcC* mutant host as predicted by its ability to fold in a tight loop; tight loops result in small plaques. In that study, we observed interactions between the bases just outside the central two which suggested that some sequences might form two-base loops. We have therefore set out to test whether DNA sequences known to form two-base loops *in vitro* behave as if they do so *in vivo*. The conclusion, consistent with the previous *in vitro* studies (11–14), is that the sequences 5'dCNNG and 5'dTNNA favour two-base loop formation *in vivo*.

MATERIALS AND METHODS

Media, bacterial strains, and bacteriophage strains

Casitone-agar (CA) medium contained 10g Bacto-casitone, 10g Bacto-agar (both from Difco Laboratories), and 14.7g NaCl (250mM final concentration) per litre. The added salt was found to accentuate the differences between bacteriophage strains. All plates were poured at 46°C, and to a volume of exactly 50ml. λ plating took place 3 days after pouring, and before use each stack of plates was randomly reordered. These precautions were necessary because small variations between plates can markedly affect the size of plaques on a lawn. CA top medium contained half the quantity of Bacto-agar (5g/litre).

The bacterial strain used for the plaque area assay was N2364 (AB1157 *sbcC*201 *phoR*::Tn*10*), and was obtained from R.G.Lloyd. The palindrome containing phages were all derived from DRL167 λ *pal spi*6 *cI*857 χ*C*153 (15). The palindrome in DRL167 is a 462-bp perfect inverted repeat, flanked by two *Eco*RI sites and with a unique central *Sac*I site. The *Sac*I site of DRL167 was used to insert oligonucleotide sequences obtained from the Oswel DNA service (T.Brown, Edinburgh). To aid the identification of new clones, the inserts were designed so that the cloning procedure disrupted the *Sac*I site and for positive identification another restriction site was introduced into the centre. The oligonucleotides were not phosphorylated so as to avoid insertion of multimers.

Bacteriophage plating

An overnight culture of *E.coli sbcC* was diluted 1:10 in L-broth supplemented with 5 mM $MgSO_4$/0.2% maltose and grown for 140 min at 37°C. An equal volume of 10mM Tris–HCl (pH

*To whom correspondence should be addressed

8), 10 mM $MgSO_4$ was added. The diluted cultures were stored at 4°C and used within three days. Aliquots of 0.25 ml of the cell suspension were incubated with a suitable dilution of phage at 37°C for 20 min, and then poured onto CA plates in 2 ml of molten CA top medium. These conditions were chosen to maximise pre-adsorption of bacteriophage. After an overnight incubation at 37°C, the area of individual plaques on the cell lawn was measured using a Quantimet 970 Image Analyser (C.Jeffree, Edinburgh). Approximately 60–100 plaques per plate were analysed (8 plates for each bacteriophage strain). The method has been described previously (15)

RESULTS

Construction of a set of palindromes in a λ derivative

To initiate the investigation of loop structure *in vivo*, a series of long DNA palindromes that differ only in their central sequence was constructed in bacteriophage λ (Fig. 1). One strand of the central six bases for each palindrome was identical to a member of the two sets of DNA sequences used to study the structure of DNA hairpin-loops *in vitro* (11–14). However, the palindrome arms were much longer so that the full size was 480bp.

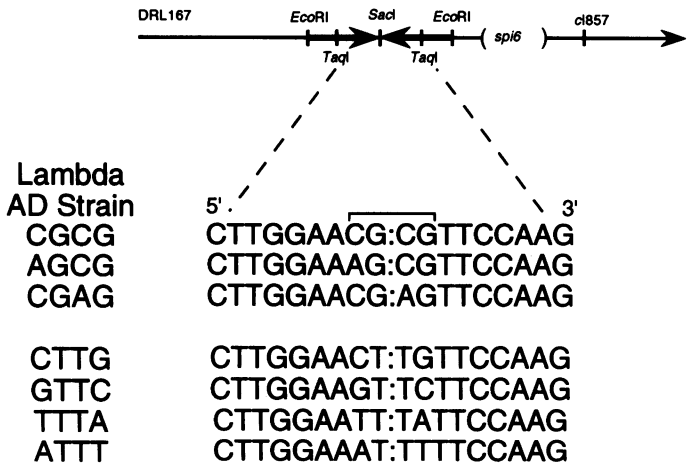

Figure 1. Bacteriophage λ constructs. **Top**: Partial map of DRL167, the parental phage used in this assay. **Bottom**: The sequences at the centre of the 480 base pair palindromes used in this assay. The central 18 base pairs are represented but changes were made only to the central four. The phage were constructed by inserting oligonucleotides into the *SacI* site of DRL167 as described previously (15) and in the materials and methods section. Positive identification of the derivative containing the oligonucleotide was facilitated by using sequences that generated *XmnI* sites at the centres (target $GAAN_4TTC$).

Plaque area assay for loop structure

The *in vivo* assay for loop structure relies on the observation that plaque size on an *E.coli sbcC* host correlates inversely with the ability to form intra-strand hairpin-loop base pairs (15). The assay

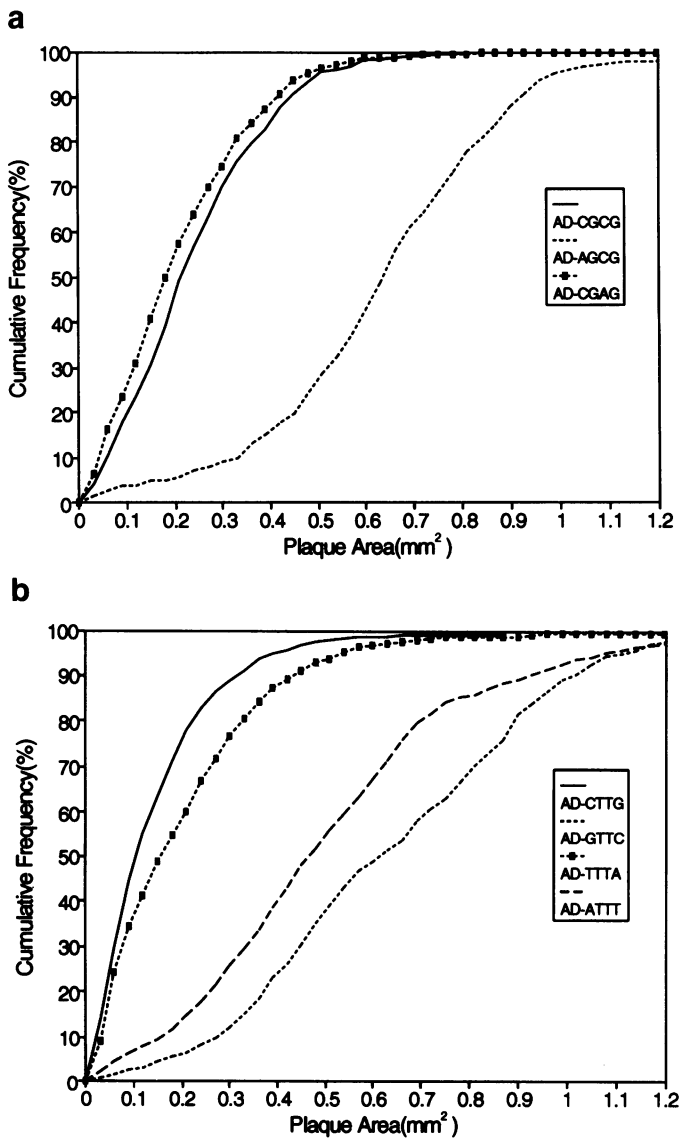

Figure 2. Cumulative frequency distribution of plaque sizes for palindrome containing λ phages. The median was taken as the point at which 50% cumulative frequency intersected a line to the x-axis. (**a**) Comparison of central sequences 5'dCGCG, 5'dAGCG and 5'CGAG.(**b**) Comparison of central sequences 5'dCTTG, 5'dGTTC, 5'dTTTA and 5'dATTT. The method of quantification of plaque area has been described previously (15) and is in the Materials and Methods section.

Table 1. Median plaque areas of palindromic phage*

	Experiment 1			Experiment 2			
Phage λAD	CGCG+	AGCG	CGAG+	CTTG+	GTTC	TTTA+	ATTT
Area (/mm²)	0.22	0.63	0.18	0.10	0.61	0.16	0.47

*λAD-CGCG, CGAG, CTTG and TTTA produce relatively small plaques. λAD-AGCG, GTTC, and ATTT produce relatively large plaques.
+Equivalent sequences have two-base hairpin-loops *in vitro* (11–14).

involves measuring the area of plaques formed on an *sbcC* host by a λ phage derivative containing a 480 bp palindrome into which different central sequences have been inserted. The results are plotted in Fig. 2(a) and (b). Table 1 and Fig. 2 show that the four l derivatives with central sequences that can form two base loops *in vitro* (CGCG, CGAG, CTTG, TTTA centres; 11−14), make small plaques *in vivo*. Conversely, the three sequences that form four base loops *in vitro* (AGCG, GTTC, ATTT centres; 11−14) confer significantly larger plaque size *in vivo*.

DISCUSSION

We have previously shown (15) that the plaque area of a bacteriophage λ derivative containing a long palindrome, when plated on an *sbcC* host, is acutely sensitive to the palindrome's central sequence. Small plaque formation was correlated with the strength of intra-strand pairing in the stem and the ability to form small loops. The base pairs having the most effect on plaque size were those adjacent to the loop. This work was interpreted in terms of a model where intra-strand base pairing and the formation of a small loop enhance the kinetics of hairpin or cruciform formation, by stabilising a tight transition state with a minimum number of unpaired bases. Furthermore, it was suggested that the results were consistent with two-base loop formation for some sequences.

In order to test this hypothesis, sequences known to form two and four-base loops *in vitro* have been compared in the plaque-size assay. Fig. 2(a) compares the plaque areas of the λ phage with the three centres 5'dCGCG, 5'dAGCG and 5'dCGAG (positions 5'1,2,3,4). It can be seen that the ability to form a base pair between dC in position 1 and dG in position 4 is essential for small plaque formation but that complementarity between bases in positions 2 and 3 is not required.

In order to test the effect of all four different complementary bases in positions 1 and 4 the plating behaviour of phage with central sequences 5'dCTTG, 5'dGTTC, 5'dTTTA and 5'dATTT were compared. As can be seen in Figure 2(b), centres with pyrimidines (dC or dT) in position 1 resulted in significantly smaller plaques than centres with purines (dG or dA) in this position. This is consistent with the observation that a 5' pyrimidine in position 1, and a 3' complementary purine in position 4, allows two-base loop formation for these sequences *in vitro* (13, 14). Both *in vivo* and *in vitro*, the same rule applies: Two-base loop formation is favoured by the sequence 5'dCNNG or 5'dTNNA.

The assay for tight loop formation used here measures the effect of a small DNA sequence at the centre of a long palindrome. Therefore, the structure resulting in alteration of plaque size is likely to be a long hairpin or cruciform. Nevertheless, it is the interactions within the short central sequence that are crucial in determining plaque size. There is no reason to believe that these interactions may not also occur in a normal, non-palindromic, DNA sequence context. According to this interpretation, the long palindromic arms act as a monitor for the dynamics of the central sequence.

The ability of certain DNA sequences to form tight loops may relate to their roles in gene regulation, recombination or mutagenesis. This plaque assay for DNA structure will allow the testing of intrinsic loop potential for specific sequences believed to mediate their effect via intra-strand base pairing.

ACKNOWLEDGEMENTS

We thank E.Okely for technical assistance and C.Jeffree for assistance with the image analysis. Funding for this work has come from the Medical Research Council

REFERENCES

1. Müller, U. R. and Fitch, W. M. (1982) *Nature*, **298**, 582−585.
2. Dimri, G. P. and Das, H. K. (1990) *Nucleic Acids Res.*, **18**, 151−156.
3. van Holde, K. and Zlatanova, J. (1994) *BioEssays*, **16**, 59−67.
4. Roth, D. B., Menetski, J. P., Nakajima, P. B. Bosma, M. J. and Gellert, M. (1992) *Cell*, **70**, 983−991.
5. Sinden, R. R. and Wells, R. D. (1992) *Curr. Opin. Biotechnol.*, **3**, 612−622.
6. Leach, D. R. F. *BioEssays* (In press).
7. van de Ven, F. J. M. and Hilbers, C. W. (1988) *Eur. J. Biochem.*, **178**, 1−38.
8. Senior, M. M., Jones, R. A. and Breslauer, K. J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6242−6246.
9. Antao, V. P., Lai, S. Y., and Tinoco Jr., I. (1991) *Nucleic Acids Res.*, **19**, 5901−5905.
10. Pieters, J. M. L., deVroom, E., van der Marel, G. A., van Boom, J. H., Koning, T. M. G., Kaptein, R. and Altona, C. (1990) *Biochemistry*, **29**, 788−799.
11. Raghunathan, G., Jernigan, R. L., Miles, H. T. and Sasisekharan, V. (1991)*Biochemistry*, **30**, 782−788.
12. Howard, F. B., Chen, C.-Q., Ross, P. D. and Miles, H. T. *Biochemistry*, **30**, 779−782.
13. Blommers, M. J. J., Walters, J. A. L. I., Haasnoot, C. A. G., Aelen, J. M. A., van der Marel, G. A., van Boom, J. H. and Hilbers, C. W. (1989) *Biochemistry*, **28**, 7491−7498.
14. Blommers, M. J. J., van de Ven, F. J. M., van der Marel, G. A., van Boom, J. H. and Hilbers, C. W. (1991) *Eur J. Biochem.*, **201**, 33−51.
15. Davison, A. and Leach, D. R. F. (1994) *Genetics*, **137**, 361−368.