DNA tertiary structures formed *in vitro* by misaligned hybridization of multiple tandem repeat sequences

Lesley W.Coggins* and Margaret O'Prey

Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK

Received July 6, 1989; Revised and Accepted August 18, 1989

ABSTRACT

DNA tertiary structures are shown to be formed by denaturation and reannealing in vitro of molecularly-cloned DNA containing multiple tandem repeat sequences. Electron microscopy of homoduplex DNA molecules containing the human c-Harvey-ras gene revealed knot-like structures which mapped to the position of the 812 bp variable tandem repeat (VTR) sequence. We propose that the structures result from slipped-strand mispairing within the VTR and hybridisation of homologous repetitive sequences in the single-stranded loops so produced. Similar structures were also found in freshly-linearised supercoiled plasmids. More complex knot-like structures were found in homoduplexes of a 4 kb tandem array from the hypervariable region 3′ to the human α-globin locus. Formation of such DNA tertiary structures in vitro also provides a practical method for identifying and mapping direct tandem repeat arrays that are at least 800 bp long.

INTRODUCTION

The conformation of a DNA molecule in vitro and in vivo depends on a number of factors, including its nucleotide sequence and whether the DNA is under torsional constraint, as recently reviewed (1). Several classes of sequences found in the eukaryotic genome are multiple tandem repeats of simple DNA sequences. They include satellite sequences, some of which are localised in centromeric heterochromatin (2), and telomeric DNA (3). Hypervariable regions (HVR) are arrays of tandemly-repeated simple DNA sequences, some of which are dispersed throughout the genome as 'minisatellite' sequences (4). The size of an array is inherited, but subject to a high rate of mutation (5) so that members of a population exhibit different copy numbers of the sequence. Resulting restriction fragment length polymorphisms (RFLPs) provide useful markers for mapping genes.

One such sequence, termed the variable tandem repeat (VTR), lies about 1 kilobase (kb) to the 3′ side of the human cellular Harvey ras (c-Ha-ras) oncogene (6). Its function is not known, but c-Ha-ras sequences lacking the

©IRL Press

VTR have 5-10 fold lower transforming activity (7) and a reduced expression of the p21 ras protein (8), so it may act as an enhancer. Particular VTR size alleles may (7, 9) or may not (10) be correlated with susceptibility to certain tumours, as are RFLPs for Taq I restriction sites within the VTR (11). A different HVR at the 3′ side of the human α-globin gene locus (12) provides a genetic marker for adult polycystic kidney disease (13).

In negatively-supercoiled DNA, direct tandem repeats can form slipped-strand structures (1). Slipped-strand mispairing has been proposed as a mechanism for the evolution of tandem repeats of simple sequences and a cause of frameshift mutations in such sequences (14, 15). In this paper we show that both the HVRs described above can give rise to knot-like structures as a result of denaturation and reannealing in vitro, and postulate that they are derived by slipped-strand mispairing followed by secondary hybridisation of the single-stranded DNA loops so produced.

MATERIALS AND METHODS

Recombinant Clones

The activated Ha-ras oncogene in a human genomic 6.46 kb Bam HI fragment, isolated from T24 bladder carcinoma cell DNA, was cloned in pBR322 (16). It was recloned in pUC8 to produce plasmids pT24a and pT24b (17) in which the inserts are in opposite orientations with respect to the vector. This was confirmed both by restriction endonuclease mapping and by heteroduplex mapping to the viral Ha-ras sequence BS-9 (17, 18). Plasmids pT24a and b were linearised with Sca I, which cuts once in the pUC8 sequence to give vector arms of 0.9 and 1.6 kb.

The plasmid pα3′HVR.64 (19) contains an HVR which lies 3′ to the human α-globin locus (12, 20) cloned in the Hinc II site of plasmid pSP64. Purified plasmid DNA was generously provided by Dr. G. Lanyon, Department of Medical Genetics, Yorkhill Hospital, Glasgow and used with kind permission of Dr D.R. Higgs, M.R.C. Molecular Haematology Unit, John Radcliffe Hospital, Oxford. pα3′HVR.64 was linearised with Xmn I which cuts once in the pSP64 sequence to give vector arms of 1.0 and 2.0 kb.

Electron microscopy

Native and Sca I-cut pT24a DNA were spread for electron microscopy in 30% formamide, 0.1 M Tris HCl pH 8.5, 0.01 M EDTA (1 x TE), 0.1 μg/ml cytochrome c on a hypophase of 5% formamide, 0.1 x TE. For homoduplex formation, linearised plasmids were denatured at 80°C for 2 min in 50% formamide, 1 x TE (21) and reannealed at 37°C for 20-30 min. Samples were

either placed on ice and spread immediately after addition of cytochrome c, or were kept at room temperature before addition of cytochrome c and spreading as described above.

The percentage of molecules with knot-like structures in a preparation was determined from observations of a number (n) of full-length, untangled and unbroken homoduplex molecules, scored sequentially with a tally counter in a unidirectional search of an E.M. grid. Length measurements of molecules were made on electron micrographs, using a Summagraphics SummaSketch Plus digitizer with a Jandel SigmaScan programme, and are displayed as mean ± S.D. for data on n molecules. One strand in each knot-like structure was traced, assuming it followed a path such as those shown in Fig. 3.

RESULTS

Tertiary Structures in pT24

The recombinant plasmids pT24a and b (Fig. 1a) each carry a 6.46 kb insert which contains a human c-Ha-ras gene and flanking sequences (6). In preparations containing Sca I-cut pT24a which had been denatured, reannealed and immediately spead for electron microscopy, it was found that a small knot-like structure (Fig. 2, a and b) was present in 11.6% (n=500) of homoduplex molecules. Although several different types of structure were observed (described in more detail below), all were of similar size, and

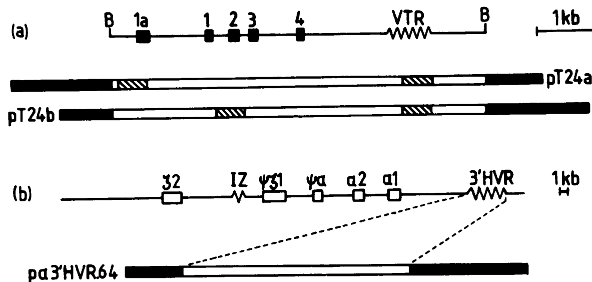

Fig. 1. (a) Map of human Bam HI (B) fragment containing the activated human c-Ha-ras gene (exons 1a, 1-4) and the VTR sequence (6). Sca I-cut pT24a and b plasmids are displayed with inserts in the same orientation, and vector arms in black. Cross-hatched boxes show the two possible locations of the knot-like structure in each plasmid. (b) Map of human α-globin gene locus showing position of genes and pseudogenes (open boxes), and the inter-zeta (IZ) and 3' HVRs (zig-zag lines) (12). Map of Xmn I-cut pα3'HVR.64, which contains a 4 kb 3' HVR, is shown with pSP64 vector arms in black, and at the same scale as pT24 maps.

consistently positioned from 20.64 (± 0.90)% to 25.99 (± 1.27)% along the length of the Sca I-cut molecules (n=25). In a denatured and reannealed preparation of pT24b, 9.0% (n=500) of homoduplexes displayed a knot-like structure (Fig. 2, c and d), in this case positioned from 29.59 (± 1.09)% to 35.39 (± 1.80)% of the length of the Sca I-cut molecules (n=25). One of the two possible locations in each orientation of the cloned T24 insert coincided at the position of the VTR (Fig. 1a).

The appearance of the structure differed in individual pT24 homoduplex molecules. In general, it looked like a knot in the homoduplex DNA, and usually exhibited bilateral symmetry. Theta-shaped (Fig. 2, a-c) structures were most frequently seen, followed by figures-of-eight or a pair of small loops (Fig. 2d), then pretzel-shaped and other conformations. In theta-shaped and figure-of-eight structures, one loop was sometimes larger than the other, with the most extreme case resulting in a structure looking like a single small ring of duplex DNA on one side of a linear homoduplex molecule.

In order to determine the stability of the structures, aliquots of the pT24a homoduplex preparation were kept at room temperature for different lengths of time before spreading. Whereas 11.6% of pT24a homoduplexes spread directly after reannealing contained a tertiary structure, this decreased to 7.6% in homoduplexes spread after 15 min, 4.8% after 30 min, and 2.4% after 1 h (n=500 for each time point).

Restriction endonuclease-treated supercoiled plasmid preparations were routinely examined by electron microscopy to determine whether they were completely cut before use in heteroduplex analysis. It was found that 5.0% (n=500) of Sca I-cut 'native' pT24a plasmid DNA contained knot-like structures (Fig. 2e) similar to those described above. The structures were from 20.61 (± 0.89)% to 27.28 (± 1.52)% along the cut molecules (n=25) and so were in the same position as those in reannealed molecules. Uncut supercoiled molecules were too tightly twisted to determine unequivocally whether tertiary structures were present.

Tertiary Structures in pα3'HVR.64

Homoduplexes of the plasmid pα3'HVR.64 (Fig. 1b) showed larger knotted structures often consisting of several loops (Fig. 2, f and g), and not simply enlarged versions of the structures found in pT24 homoduplexes. Their complexity made interpretation difficult as individual strands could not be followed. Two duplex DNA arms emerged from the central knot, and were of at least the length expected for the Xmn I cut vector segments.

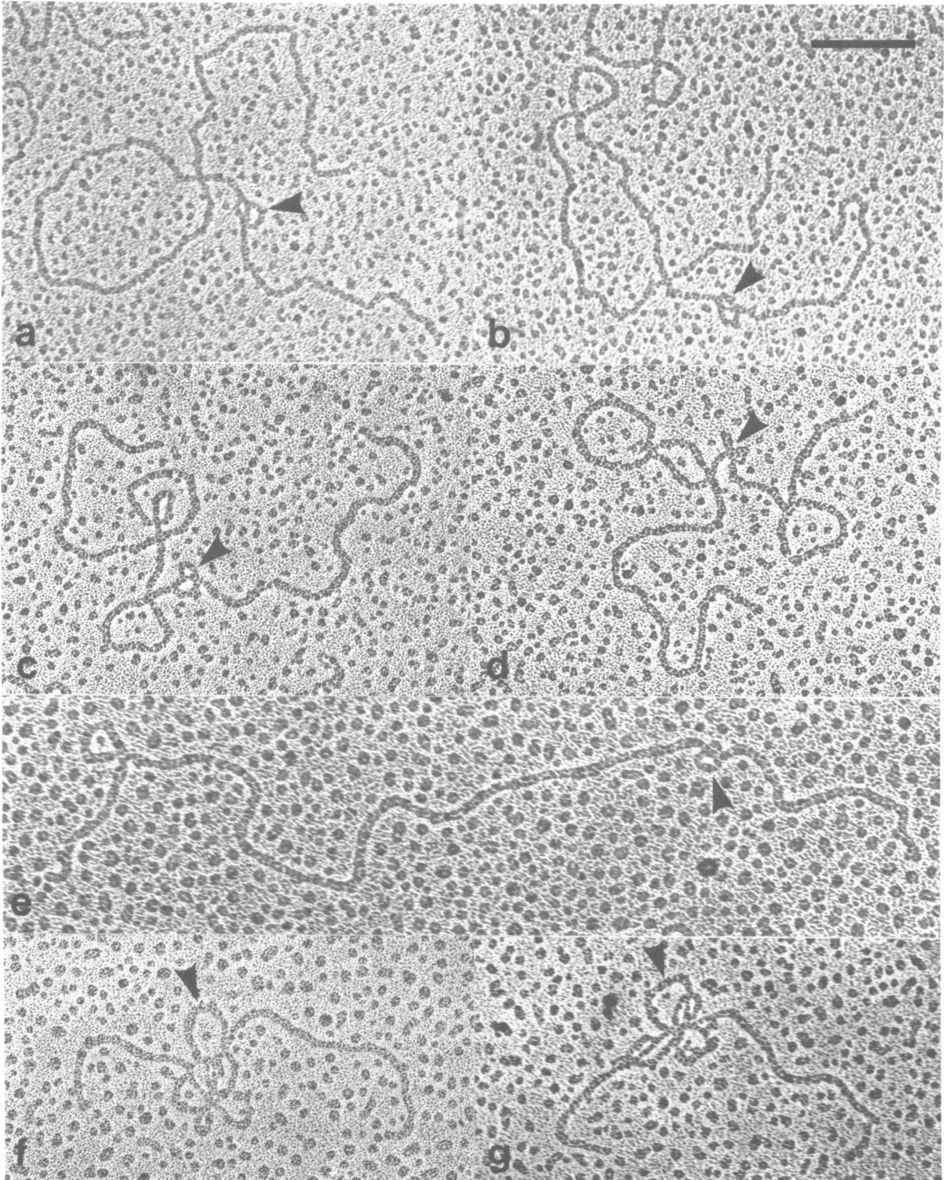

Fig. 2. Electron micrographs showing tertiary structures (arrows) in: (a, b) Sca I-cut pT24a homoduplexes, (c, d) Sca I-cut pT24b homoduplexes, (e) pT24a molecule spread immediately after cutting supercoiled plasmids with Sca I and (f, g) Xmn I-cut pα3'HVR.64 homoduplexes. All molecules are shown at the same magnification and the bar represents 0.25 μm.

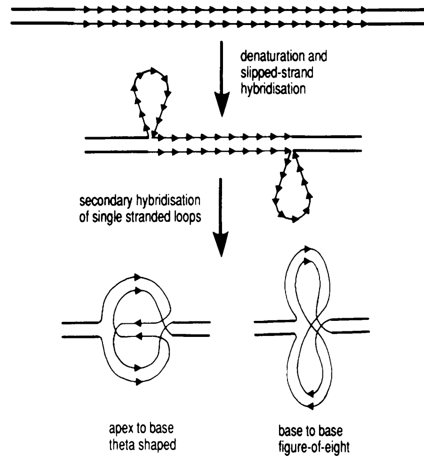

Fig. 3. Proposed mechanisms of formation of theta-shaped and figure-of-eight structures. Arrow heads indicate one end of each repeat sequence in a hypothetical tandem array. For clarity, only some arrow heads are shown in the knot-like structures.

DISCUSSION

Proposed Mechanism of Formation of Tertiary Structures

We propose that the knot-like tertiary structures described above in reannealed pT24 molecules result from misaligned hybridisation of the VTR, which in the cloned T24 sequence is a tandem array of 29 copies of a 28 base pair (bp) consensus sequence (6). The percentage of tertiary structures observed is presumed to reflect their frequency of formation and their stability during subsequent procedures. Formation would be initiated by nucleation, a second order reaction, occurring within the tandem array. This is expected to be subject to previously-established factors (22) such as the length, complexity and GC content of the array compared to the rest of molecule. If nucleation occurs between two different members of the VTR, out-of-register hybridisation proceeding from the site results in a single-stranded DNA loop at each end of the array (Fig. 3).

If the loops translocate along the array by single-strand branch migration (23, 24) and then merge, an opportunity for in-register hybridisation occurs. For a metastable knot-like structure to form, we postulate that secondary, out-of-register hybridisation between the complementary repetitive sequences in the loops takes place instead. The appearance of the resulting tertiary structure depends on the degree of slippage at nucleation, and hence the relative sizes of the region that

hybridised first and of the loops, and on the alignment of the loops when they pair. We propose that the most commonly-observed structures are formed as shown in Fig. 3. Intermediate alignments, and partial hybridisation due to constraints on the formation of double-stranded DNA in the topologically-closed loops (21, 24), might give rise to pretzel-shaped structures and other conformations.

The recombinant plasmid pα3'HVR.64 (Fig. 1b) contains 228 tandem copies of a 17 bp consensus sequence (19). This 4 kb array formed complex tertiary structures containing a number of duplex loops of different sizes. This may be due to more than one nucleation event and hence secondary hybridisation of more than two single-stranded DNA loops, possibly under torsional constraint (24).

Even with the additional hybridisation between the loops, the tertiary structures might be removed by local denaturation and in-register rehybridisation ('breathing'), or by branch migration (23, 24). Both processes increase at higher temperatures (22, 25) and we have found that the structures are progressively lost when kept at room temperature. The topology of the structures could affect their stability; in addition, longer duplex sequences with high GC content and no mismatched sequences would be expected to be most stable. Sequence divergence (22) occurs within both tandem arrays studied (6, 11, 19).

Use in Mapping Tandem Repeat Sequences

In practical terms, our observations show that electron microscopy can identify multiple tandem repeats of a simple-sequence DNA and map them with respect to restriction sites in molecularly-cloned DNA. The number and length of individual repeats in the array must be determined by sequencing.

Some of the structures described (for example, pretzel-shaped) are similar to those formed in heteroduplexes by translocated sequences (26-28), but in those cases the loops are of constant length and can only hybridise with their bases aligned. The variety of knot-like conformations described in this paper are also clearly distinct from the simple single-stranded loop formed when two DNA species differing by a deletion are hybridised (21). We do not believe that the tertiary structures just result from copy number differences between cloned plasmid molecules forming homoduplexes, because they are also found in cut supercoiled pT24 molecules where no reannealing has occurred. In the molecularly-cloned VTR, we have found no evidence in heteroduplex preparations (17) for any significant loss

of tandem repeat sequences. Loss of copies from the HVR during cloning has been reported (20), and the HVR homoduplex preparation did contain some molecules with simple deletion loops (data not shown). In this case, deletions may confer an additional level of complexity on some of the knot-like structures in the preparation.

The structures formed by the VTR are close to the minimum size that can be identified in a spread preparation containing cytochrome c. Their measured length of about 500 bp is less than the 812 bp of the cloned VTR, possibly due to translocation of the loops as described above. A molecule with a metastable knot-like tertiary structure might also be expected to have increased mobility in gel electrophoresis, especially if it is a low molecular weight fragment, and this may explain some minor bands noted in high-resolution electrophoretic analysis of VTR alleles (29).

Tertiary Structures in Supercoiled DNA In Vitro and In Vivo

In supercoiled plasmids, direct repeats flanking a TATA box in the adenovirus late promotor (30) and in both mouse and chicken α2(I) collagen gene promotor regions (31) have been found to form S1 nuclease-sensitive sites, which were ascribed to single-stranded loops at the termini of slipped-strand structures. Furthermore, the pattern of S1 nuclease digestion in a direct repeat sequence upstream of Drosophila heat shock genes has been attributed to slipped-strand mispairing, followed by hybridisation of complementary sequences in the loops (32). We invoke a similar mechanism in vitro for larger direct repeat sequence arrays.

We have not confirmed whether any or all knot-like tertiary structures are favoured by negative supercoiling. If they are, some supercoiled as well as linear pT24 plasmids should exhibit sites which are sensitive to S1 nuclease and other probes for conformational changes in DNA. A novel conformation, anisomorphic DNA, has been proposed to explain the properties in supercoiled plasmids of tandem repeat arrays of the Herpes simplex virus type 1 (HSV-1) DR2 sequence (33), but the deduced structure is different from the one described here. The DR2 array is part of a sequence which has been implicated in genomic segment inversion and cleavage of newly-replicated concatemeric HSV-1 DNA (34-36). Intriguingly, the 12 bp DR2 core sequence from HSV-1 strain F [GGGGGAggAGCG - the complementary sequence to that given in (33)] shows homology with part of the 17 bp 3′ HVR consensus sequence [gGGGGGAacAGCGacac (19)]. GGGGA and ACAC motifs are also found in the VTR sequence (6). All three sequences are GC-rich. They have a strand bias in the distribution of purines and pyrimidines (DR2 and

3'HVR), which can cause DNA bending (1), or alternating blocks of purines and pyrimidines (VTR). It would be interesting to determine whether direct tandem repeat sequences can form knot-like tertiary structures in vivo, under what conditions, and whether this provides a mechanism for any of the diverse functions in which such sequences have been implicated.

ACKNOWLEDGEMENTS

The Beatson Institute is supported by the Cancer Research Campaign.

*To whom correspondence should be addressed

REFERENCES

1. Wells, R.D. (1988) J. Biol. Chem., 263, 1095-1098.
2. Singer, M.F. (1982) Int. Rev. Cytol., 76, 67-112.
3. Blackburn, E.H. and Szostak, J.W. (1984) Ann. Rev. Biochem., 53, 163-194.
4. Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Nature (London), 314, 67-73.
5. Jeffreys, A.J., Royle, N.J., Wilson, V. and Wong, Z. (1988) Nature (London), 332, 278-281.
6. Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. and Goeddel, D.V. (1983) Nature (London), 302, 33-37.
7. Krontiris, T.G., DiMartino, N.A., Colb, M. and Parkinson, D.R. (1985) Nature (London), 313, 369-374.
8. Seeberg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V. and Levinson, A.D. (1984) Nature (London), 312, 71-73.
9. Heighway, J., Thatcher, N., Cerny, T. and Hasleton, P.S. (1986) Brit. J. Cancer, 53, 453-457.
10. Thein, S.L., Oscier, D.G., Flint, J. and Wainscoat, J.S. (1986) Nature (London), 321, 84-85.
11. Radice, P., Pierotti, M.A., Borrello, M.G., Illeni, M.T., Rovini, D. and Della Porta, G. (1987) Oncogene, 2, 91-95.
12. Higgs, D.R., Wainscoat, J.S., Flint, J., Hill, A.V.S., Thein, S.L., Nicholls, R.D., Teal, H., Ayyub, H., Peto, T.E.A., Falusi, A.G., Jarman, A.P., Clegg, J.B. and Weatherall, D.J. (1986) Proc. Natl. Acad. Sci. USA, 83, 5165-5169.
13. Reeders, S.T., Breuning, M.H., Davies, K.E., Nicholls, R.D. Jarman, A.P., Higgs, D.R. and Weatherall, D.J. (1985) Nature (London), 317, 542-544.
14. Levinson, G. and Gutman, G.A. (1987) Molec. Biol. Evol., 4, 203-221.
15. Murphy, G.L., Connell, T.D., Barritt, D.S., Koomey, M. and Cannon, J.G. (1989) Cell, 56, 539-547.
16. Santos, E., Tronick, S.R., Aaronson, S.A., Pulciani, S. and Barbacid, M. (1982) Nature (London), 298, 343-347.
17. McCaffery, R., Coggins, L.W., Kennedy, I., O'Prey, M., McColl, L. and Campo, M.S. (1989) Oncogene (in press).
18. Ellis, R.W., DeFeo, D., Maryak, J.M., Young, H.A., Shih, T.Y., Chang, E.M., Lowy, D.R. and Scolnick, E.M. (1980) J. Virol., 36, 408-420.
19. Jarman, A.P., Nicholls, R.D., Weatherall, D.J., Clegg, J.B. and Higgs, D.R. (1986) EMBO J., 5, 1857-1865.

20. Goodbourn, S.E.Y., Higgs, D.R., Clegg, J.B. and Weatherall, D.J. (1983) Proc. Natl. Acad. Sci. USA, 80, 5022-5026.
21. Coggins, L.W. (1987) In Sommerville, J. and Scheer, U. (eds.) Electron Microscopy in Molecular Biology: A Practical Approach. IRL Press, Oxford and Washington, D.C., pp. 1-60.
22. Wetmur, J.G. and Davidson, N. (1968) J. Mol. Biol., 31, 349-370.
23. Kim, J.-S., Sharp, P.A. and Davidson, N. (1972) Proc. Natl. Acad. Sci. USA, 69, 1948-1952.
24. Ferguson, J. and Davis, R.W. (1978) In Koehler, J.K. (ed.) Advanced Techniques in Biological Electron Microscopy, Springer-Verlag, Berlin and Heidelberg, Vol. 2, pp. 123-171.
25. Thompson, B.J., Camien, M.N. and Brown, R.C. (1976) Proc. Natl. Acad. Sci. USA, 73, 2299-2303.
26. Malamy, M.H., Fiandt, M. and Szybalski, W. (1972) Molec. gen. Genet., 119, 207-222.
27. Mosharrafa, E., Pilacinski, W., Zissler, J., Fiandt, M. and Szybalski, W. (1976) Molec. gen. Genet., 147, 103-109.
28. Broker, T.R., Soll, L. and Chow, L.T. (1977) J. Mol. Biol., 113, 579-589.
29. Gejman, P.V., Sitaram, N., Hsieh, W.-T., Gelernter, J. and Gershon, E.S. (1988) Appl. Theor. Electrophor., 1, 29-34.
30. Yu, Y.-T. and Manley, J.L. (1986) Cell, 45, 743-751.
31. McKeon, C., Schmidt, A. and de Crombrugghe, B. (1984) J. Biol. Chem., 259, 6636-6640.
32. Mace, H.A.F., Pelham, H.R.B. and Travers, A.A. (1983) Nature (London), 304, 555-557.
33. Wohlrab, F., McLean, M.J. and Wells, R.D. (1987) J. Biol. Chem., 262, 6407-6416.
34. Varmuza, S.L. and Smiley, J.R. (1985) Cell, 41, 793-802.
35. Chou, J. and Roizman, B. (1985) Cell, 41, 803-811.
36. Deiss, L.P., Chou, J. and Frenkel, N. (1986) J. Virol, 59, 605-618.