

Common determinants in DNA melting and helicase-catalysed DNA unwinding by papillomavirus replication protein E1

Sandrine Castella, Gregg Bingham and Cyril M. Sanders*

Institute for Cancer studies, University of Sheffield, Beech Hill Road, Sheffield S10 2RX, UK

Received February 7, 2006; Revised April 26, 2006; Accepted May 3, 2006

ABSTRACT

E1 and T-antigen of the tumour viruses bovine papillomavirus (BPV-1) and Simian virus 40 (SV40) are the initiator proteins that recognize and melt their respective origins of replication in the initial phase of DNA replication. These proteins then assemble into processive hexameric helicases upon the single-stranded DNA that they create. In T-antigen, a characteristic loop and hairpin structure (the pre-sensor 1 β hairpin, PS1 β H) project into a central cavity generated by protein hexamerization. This channel undergoes large ATP-dependent conformational changes, and the loop/PS1 β H is proposed to form a DNA binding site critical for helicase activity. Here, we show that conserved residues in BPV E1 that probably form a similar loop/hairpin structure are required for helicase activity and also origin (*ori*) DNA melting. We propose that DNA melting requires the cooperation of the E1 helicase domain (E1HD) and the origin binding domain (OBD) tethered to DNA. One possible mechanism is that with the DNA locked in the loop/PS1 β H DNA binding site, ATP-dependent conformational changes draw the DNA inwards in a twisting motion to promote unwinding.

INTRODUCTION

T-antigen and E1 both function as hexameric helicases in the replication of Simian virus 40 (SV40) (1,2) and papillomavirus (3,4) DNA, respectively, but also perform the initiation function required for helicase loading, generating the single-stranded origin (*ori*) DNA from which processive unwinding can proceed. The assembly and structure of SV40 and bovine papillomavirus (BPV-1) initiator complexes are now well understood. In BPV, the E1 helicase forms in a stepwise manner and requires the assistance of the viral transcription factor E2 (5,6). *In vitro*, dimers of E1 and E2 first

bind cooperatively to *ori* where E1 and E2 sites are side by side (7–9). The function of this E1E2–*ori* complex is to generate the specificity and affinity for origin DNA binding that E1 alone otherwise lacks (5). This is achieved, in part, by the ability of E2 to modulate the non-specific DNA binding activity of the E1 helicase domain (E1HD) (10). In a subsequent reaction that requires ATP, E2 is displaced and more E1 molecules are recruited to *ori*. The function of this E1–*ori* complex appears to be stable and specific origin binding (11), although structural changes in the DNA ultimately leading to strand separation may begin (12,13). This complex is a tetramer and the recruitment of more E1 molecules, mediated via the E1 sequence-specific DNA binding domain, leads to *ori* melting either side of the E1 binding site (11,12). Recent work suggests that the *ori*-melting complex can be viewed as two functional trimers that ultimately become two hexameric helicases (14). However, despite this detail, the challenge of relating structure to function remains.

The recent crystal structures of a fragment of SV40 T-antigen have revealed features that may aid in the understanding of DNA binding, melting and unwinding by the protein (15,16). T-antigen (residues 251–627) forms a two-tiered hexameric structure with a longitudinal cavity that opens to form a wide central chamber. The channel is strongly positively charged, and it has been proposed that strand separation can occur within the chamber, as double-stranded DNA is threaded through the complex. There are also six positively charged holes formed between monomers on the side-walls of the large chamber, which may form an exit for single-stranded DNA (15). The nucleotide hydrolysis cycle appears to induce large expansions and contractions of the channel, which may be important for origin melting and helicase activity (16). Studies with the electron microscope have demonstrated the presence of DNA in the central channel, and its binding also appears to induce structural changes in the protein complex (17). Positively charged and ring-shaped aromatic amino acids on a β -hairpin (the pre-sensor 1 β -hairpin, PS1 β H) and a loop structure that undergo large NTP-dependent movements are key for T-antigen unwinding (18). A phenylalanine (on the loop) and a lysine (on the PS1 β H) are absolutely conserved in helicase superfamily 3,

*To whom correspondence should be addressed. Tel: +1 14 2712482; Fax: +1 14 2713892; Email: c.m.sanders@sheffield.ac.uk

and in papillomavirus there is an extended conserved sequence motif, FWL (amino acids 464–466 in BPV E1), on the loop (Figure 1a). Figure 1b illustrates the positioning of the conserved residues in the T-antigen hexamer. Information for a fragment of HPV 18 E1 (19) suggests that these features are structurally conserved between SV40 and papillomavirus, and a structural representation of the BPV sequence modelled on the available structural information is shown in Figure 1c. Here we have asked if the conserved residues of the β -hairpin and aromatic loop in BPV E1 are required for helicase activity and also *ori* DNA melting. Our findings demonstrate a requirement for the same amino acid in both these processes, indicating a mechanistic relationship between DNA melting and processive unwinding.

MATERIALS AND METHODS

Protein expression and purification

Expression and purification of the full-length E1 protein were as described previously (7). To express the E1HD, DNA encoding the C-terminal 307 amino acids of the BPV-1 E1 open reading frame (ORF) was cloned as a glutathione

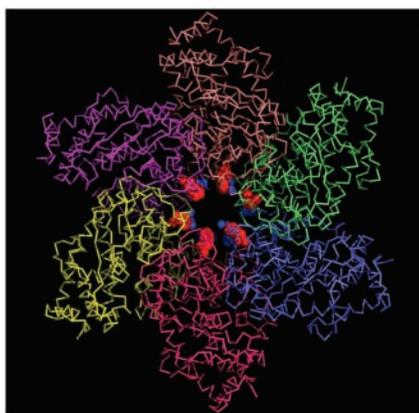
S-transferase (GST)–ORF fusion in pET11c as for full-length E1. Mutagenesis was performed by overlapping primer extension, and all sequences verified in full using fluorescent primer or chain terminator technologies (LiCore/ABI).

Transformed *Escherichia coli* BL21(DE3) cultures were grown at 22°C and induced with IPTG for 6–8 h after reaching an A_{600} 0.8–1. All purification procedures were performed at 4°C. Frozen cells (–80°C) were lysed by sonication in 50 mM Tris–HCl (pH 8.0, 4°C), 0.9 M NaCl, 10 mM EDTA, 5 mM DTT, 10% glycerol and 1 mM phenylmethanesulphonyl fluoride (PMSF) following treatment with lysozyme (1 mg/ml) for 30 min. Lysates were cleared at 25 000 g and incubated with glutathione–Sephacrose (50 μ l/g of cells) for ~16 h, and the beads processed as described for E1 (7). Following thrombin cleavage, the protein extract was diluted with 5 vol of 20 mM sodium phosphate buffer (pH 6.8), 5 mM DTT, 10% glycerol and 1 mM PMSF, and applied to a 1 ml Source-S column (Amersham Bioscience) equilibrated with the same buffer containing 50 mM NaCl. The column was developed over a gradient from 50 to 350 mM NaCl and peak fractions pooled. The protein was purified further by anion exchange chromatography (Source-Q, 1 ml column) as described for full-length E1, and the peak fractions pooled,

(a)

	Walker A	FWL	Walker B	RKH	Sensor 1
HPV16	FLTALKRFLQGI	CAANDGRSLFGMSLMKFLQGSV	ICVFNKSH	EWLQPLA	AKIIGMLDDATVPCWNYIDD
HPV18	FLGALKSFLKGT	CPANQKSYF	FGMSFIHFIQ	GAVISFVNSTSH	EWLEPLT
HPV57	FLAAFKSFLKGV	CPANQKSYF	FGMSFIHFIQ	GAVISFVNSTSH	EWLEPLT
HPV6	FLSKFKLWLHG	CPANQKSYF	FGMSFIHFIQ	GAVISFVNSTSH	EWLEPLT
BPV	FINALKLWLKGI	CPANQKSYF	FGMSFIHFIQ	GAVISFVNSTSH	EWLEPLT
HPV8	FLAALKDPLHSV	CPANQKSYF	FGMSFIHFIQ	GAVISFVNSTSH	EWLEPLT
AAV	-----	-----	-----	-----	-----
SV40	-----	-----	-----	-----	-----

(b)



(c)



Figure 1. Conserved amino acids in helicase superfamily 3, structure of SV40 T-antigen and structural model for the BPV helicase domain. (a) Protein sequence alignments in the helicase superfamily 3 ATPase domain. The alignment was generated using CLUSTAL W (1.82) multiple sequence alignment software. The BPV amino acids shown are from 412 to 529. Absolute conservations are highlighted in red and conservative substitutions in bold red script. The ATPase regions (Walker A and B) and Sensor 1 motif are boxed, as are the conserved/semi-conserved amino acids of the hydrophobic loop (FWL) and PS1 β HP (RKH) regions. The accession numbers are: T-antigen SV40, AF316141_5; adeno-associated-virus 2, AAK76418.1; BPV, AAB35071.1; HPV8, P06420; HPV6, AAF00059.1; HPV16, AAV91678.1; HPV18, CAA28666.1; HPV57, P22153. (b) Structure of the SV40 T-antigen hexamer is shown, viewed from the face of the large tier, with the α backbones of individual subunits coloured. The conserved phenylalanine (F459) in each subunit is highlighted as a red space-filling representation and the conserved lysine in the PS1 β H (K512, positioned behind F459) likewise in blue. (c) Structural representation of the BPV helicase domain modelled using Geno3D (Geno3D: automatic comparative molecular modelling of protein, <http://geno3d-pbil.ibcp.fr>). A monomer is shown in a similar orientation to the subunits of the T-antigen hexamer in (b), with similar secondary structure features coloured the same. The conserved residues F464 and K506, positioned on loops, are shown as a red or blue space-filling representation, respectively.

concentrated and stored at -80°C . Protein concentrations were determined by BioRad assay with BSA as a standard, and purity assessed on Coomassie stained SDS-PAGE gels.

ATPase assays and helicase assays

ATPase assays were performed in 20 mM HEPES-NaOH, 135 mM NaCl, 1 mM DTT, 0.01% NP-40, 7.5 mM ATP, 8.5 mM MgCl_2 containing 35 nmol/ μl [γ - ^{32}P]ATP (7000 Ci/mmol) and 4 μM protein. Reactions were incubated at 30°C and $^{32}\text{P}_i$ release determined by the charcoal-binding assay described by Iggo and Lane (20).

Helicase activity was determined by the ability of protein to displace a 70 bp DNA strand from a substrate with a 45 bp 3' overhang. The 115 base strand was generated from a PCR product produced with one phosphorylated primer and treated with λ exonuclease. The initial PCR product was derived from pUC19 constructs bearing BPV origin fragments described previously (5), using the following primers: 5'-GACTATGTATTTTTTCCAGTGTG (top strand, unphosphorylated) and 5'-GCTATGACCATGATTACGCC (bottom strand, phosphorylated). The 5' ^{32}P -labelled primer 5'-CCGCTTGAAAACCGGCAACGGTGT was annealed to the 115 bp strand and extended with *Taq* polymerase to complete synthesis of the substrate before gel purification. The helicase reaction was performed in 25 mM HEPES-NaOH, 20 mM NaCl, 1 mM DTT, 1 mM ATP, 3 mM MgCl_2 and 1 nM DNA substrate. Reactions were incubated for 60 min at 22°C and terminated by adjusting the reactions to 20 mM EDTA, 0.1% SDS, 10% glycerol and 0.13% (w/v) bromophenol blue. Products were separated on an 8% polyacrylamide/TBE gel containing 0.05% (w/v) SDS, and gels exposed to phosphorimager plates (Fujifilm) for imaging and quantification (Fuji FLA3000, image gauge V3.3 software).

Gel filtration assays

Gel filtration was performed using a Superdex S-200 column (Amersham Bioscience) equilibrated in 20 mM sodium phosphate (pH 7.2), 200 mM NaCl, 10% glycerol, 1 mM PMSF, 1 mM DTT, 1 mM ATP and 3 mM MgCl_2 . The column was calibrated using a standard range of high molecular weight markers. Proteins were pre-incubated at 2.25 mg/ml in the presence of 5 mM ATP/ Mg^{2+} (100 μl reactions, 4°C) for 10 min before sample injection and development of the column at a flow rate of 0.5 ml/min. Protein elution was monitored at 280 nm and chromatograms analysed using computer software (Unicorn V4.0, Amersham Bioscience).

DNA binding reactions and gel-shift assays

For single-strand (ss) DNA binding assays the substrate was a 30 bp oligonucleotide (5'-TGCTCTAGAATGCAAACACTACTGAACGAG) end labelled (^{32}P) with polynucleotide kinase. Reactions were heat inactivated and desalted using a G25 spin column. Interactions with the wild-type BPV origin of replication were investigated as described previously (5). Binding reactions were performed in 20 mM sodium phosphate (pH 7.2), 135 mM NaCl, 10% glycerol, 0.1% NP-40, 0.1 mg/ml BSA, 1 mM PMSF, 1 mM DTT and 125 ng/ μl poly(dA-dT). In some reactions ATP/ Mg^{2+} was at 5 mM. ssDNA binding reactions contained 1 nM probe and dsDNA binding reactions 0.1 nM probe. Products

of ssDNA binding reactions were resolved on 5% 80:1 acrylamide/bis-acrylamide gels using $0.25\times$ TBE as the electrophoresis buffer and origin binding reactions on agarose/TAE gels following glutaraldehyde cross-linking of reaction products (5). Imaging and quantification were as described above.

DNA footprinting

DNA binding reactions were assembled as described above, except that glycerol was omitted from reactions for hydroxyl radical footprinting (11). KMnO_4 footprinting was performed in the presence of excess Mg^{2+} ions (25 mM), conditions that increase the reactivity of the reagent, essentially as described previously (11). For each assay, products were analysed on an 8% urea-polyacrylamide gel with sequence ladders (G) generated by standard chemical sequencing techniques (21). Imaging and quantification were as described above.

RESULTS

Nucleotide hydrolysis and oligomerization of BPV-1 E1HD and mutants

The E1HD and proteins with substitutions of conserved or semi-conserved amino acids (Table 1) were generated and purified similarly. ATPase assays (Figure 2) demonstrated that a subset of the mutants retained activities at or close to wild-type levels, but all were active in the presence of a high concentration of a short ssDNA oligonucleotide. This result is surprising since we observed no significant effect of ssDNA on activity of the wild-type protein. The mutant R505E hydrolysed ATP at rates close to wild-type only with ssDNA (Figure 2a). Mutation of the conserved amino acid K506 to Ala (K506A) resulted in significantly reduced ATPase without ssDNA but increased activity in its presence. K506E was ~ 3 –4-fold more active than wild-type, with or without ssDNA (Figure 2a). H507 was mutated more extensively, since this amino acid is conserved between SV40 and BPV. Without ssDNA, H507A and H507L hydrolysed ATP at a rate similar to wild-type. H507A ATPase was stimulated in the presence of ssDNA, but the activity of H507L did not change significantly (Figure 2b). H507F was more active than wild-type and not stimulated, but inhibited, by ssDNA. For the hydrophobic loop 'FWL' motif, F464A and F464L hydrolysed ATP at a rate 5–10-fold less than wild-type when no ssDNA was present but were active like wild-type in its presence (Figure 2c). The ATPase activity of F464W was moderately increased relative to wild-type with and without ssDNA. W465A, W465F and L466A were

Table 1. Residues and mutations made in E1/E1HD used in this study

Residue	Substitution
F464	A L W
W465	A F
L466	A (F)
R505	E
K506	A E (R)
H507	A L F (M K)

Those in brackets failed to be of use or were used only for limited analysis.

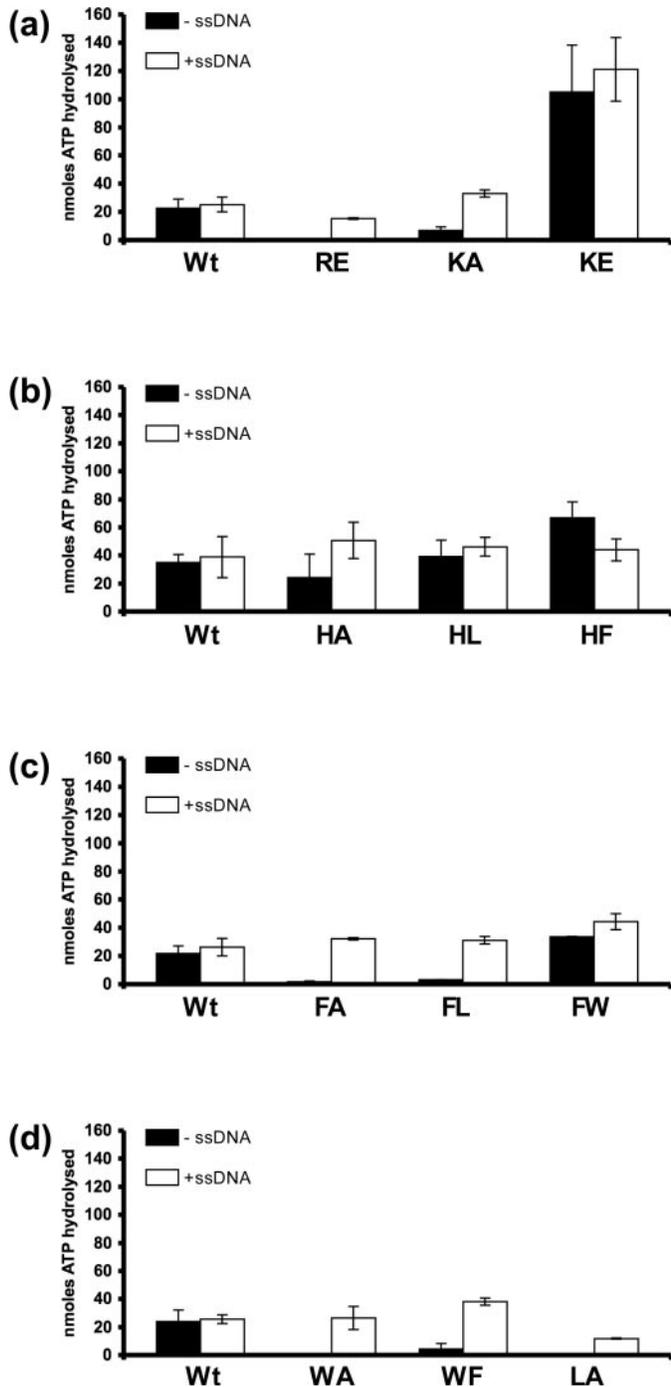


Figure 2. ATPase activity of E1HD and mutants. Reactions were assembled with 4 μ M final protein without and with a single-stranded oligonucleotide (1:3::ssDNA:protein), and released 32 P_i determined as a function of time (2.5–40 min). The data for P_i released (nmol of ATP hydrolysed) after 20 min without ssDNA are given as black columns, and with ssDNA as unfilled columns. (a) Mutants R505E, K506A and K506E (RE, KA, KE). (b) H507 mutants (HA, HL, HF). (c) F464 mutants (FA, FL, FW). (d) Mutants W465A, W465F and L466A (WA, WF, LA).

inactive or showed minimal ATPase activity without ssDNA but were all active in its presence (Figure 2d). Mutants in the Walker A and sensor 1 motifs that are essential for ATP binding and hydrolysis (K439A and N523A) were ATPase inactive, with or without DNA (data not shown).

The active form of the BPV E1 helicase is a hexamer (22). We next asked if E1HD and selected mutants were proficient for oligomerization using gel filtration. Wild-type protein was predominantly hexameric, based on the elution of protein standards, as were H507A, K506A and F464W (Figure 3). W465A and L466A were completely deficient for hexamerization, while R505E, F464A and W465F (data not shown) had an intermediate phenotype in terms of the ratios of monomer to hexamer formed, as indicated by the values given below in Figure 3. We did not observe discernable peaks corresponding to intermediate forms from monomer through to hexamer with wild-type or any of the mutants, indicating that hexamerization is a highly cooperative event. The trend that emerged was that deficiency in oligomerization appears to correlate with reduced ATPase activity in the absence of DNA.

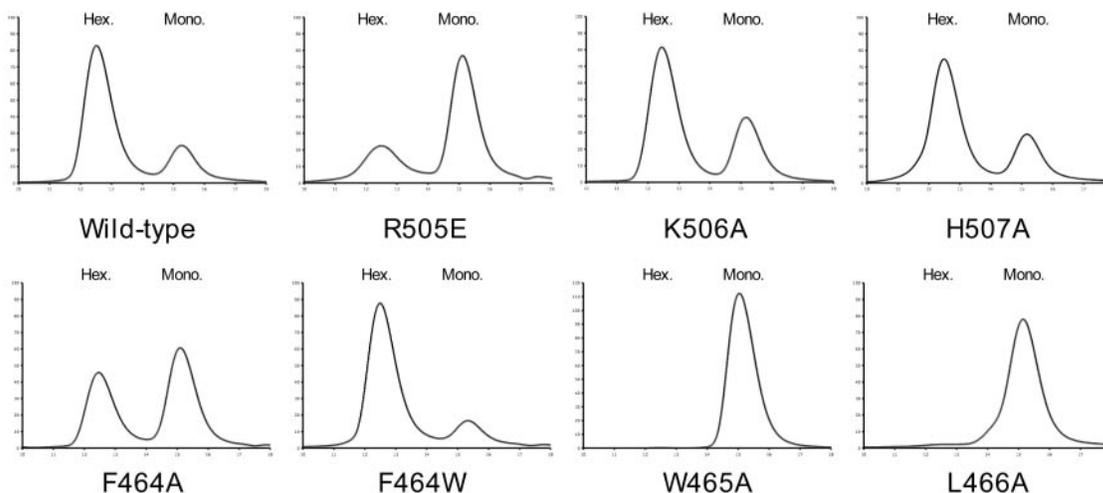
Helicase activity of E1HD and mutants

We next asked how the RKH/FWL amino acid substitutions affect helicase activity using a 70 bp DNA strand annealed to a 115 bp strand with a 45 bp 3' overhang as a substrate. Representative data is shown in Figure 4a and b, while the results are summarized in Figure 4c (500 nM E1HD). Wild-type E1HD unwound DNA (Figure 4, lanes 3–6), but high concentrations of protein were inhibitory. R505E was active and showed no inhibition of unwinding at high protein concentration (lanes 7–10), while glutamic acid (K506E, lanes 11–14) or alanine substitutions (data not shown) of K506 were inactive. The H507F substitution unwound DNA at a similar level to wild-type, while the activity of H507A and H507L were reduced (lanes 15–26 compared with lanes 3–6; Figure 4c). A definitive correlation with size of the amino acid side chain did not emerge, as the mutant H507K was only 10% active (data not shown). All substitutions in the FWL region were inactive or virtually inactive for helicase activity (Figure 4b and c), even the conservative substitutions F464W and W465F.

ssDNA binding

The PS1 β H and hydrophobic loop form a ringed array in the central channel of T-antigen, through which DNA is proposed to thread as it is unwound. Wild-type E1HD bound ssDNA to a greater extent in the presence of ATP (ATP/Mg²⁺; herein referred to simply as ATP) compared with its absence (Figure 5a, lanes 2–4, top compared with bottom). R505E did not bind DNA without ATP but was active in its presence (lanes 5–7, top compared with bottom). K506E and K506A were inactive for DNA binding (lanes 8–13, top and bottom), with or without ATP, except for formation of a fast-migrating species for K506A with ATP (lane 13, bottom). For the H507 substitutions (lanes 14–22), ssDNA binding was dependent on the bulk of the amino acid side chain (HF > wild-type > HL > HA), and in each case augmented in the presence of ATP. The gel-shift pattern for H507L and H507A with ATP indicated a level of complex instability, judged by some partial retardation of probe (lanes 15 and 16, 18 and 19).

In the hydrophobic loop, all substitutions inhibited ssDNA binding without ATP, except the conservative substitution F464W that bound oligonucleotide at wild-type levels (Figure 5b, lanes 11–13, upper panel). All mutants showed increased binding in the presence of ATP



Protein	Wild-type	R505E	K506A	H507A	F464A	F464W	W465A	L466A
Hex:Mono	81:19	25:75	70:30	75:25	43:57	86:14	0:100	0:100

Figure 3. Oligomerization of E1HD and PS1 β H and hydrophobic loop mutants. Selected mutants, as indicated, were analysed by gel filtration on a superdex S200 column in the presence of 1 mM ATP/3 mM Mg²⁺. Reactions (2.25 mg/ml protein) were pre-incubated with 5 mM ATP/Mg²⁺ for 10 min at 4°C before injection. Protein elution was monitored by A₂₈₀ as a function of volume, hexameric and monomeric peaks are indicated. The ratio of peak areas (Hex/Mono) are indicated below.

(Figure 5b, lower panel), with the substitution W465F exhibiting significantly increased activity compared with all others (lanes 11–13).

Double-stranded DNA (dsDNA) binding and the roles of PS1 β H and hydrophobic loop residues in *ori*-melting

In the presence of ATP an E1-*ori* complex forms where the HD contacts the DNA either side of the E1 binding site (10) and melts the origin DNA (5,13). To test the effects of helicase domain mutations on origin binding and DNA melting, selected mutations were constructed in full-length E1 and proteins expressed and purified. We analysed the Ala, Leu and Phe/Trp mutations of residues H507 and F464 that are conserved between SV40 and BPV, and the R505E, K506E, W465A and L466A mutations (the lysine and leucine are also conserved between BPV and SV40).

Mutation of the conserved residues in both the PS1 β H and hydrophobic loop altered DNA binding in the presence of ATP, and sample data are presented as part of the analysis of *ori* melting described subsequently (see below; Figure 6). In the PS1 β H the K506E mutant was most severely defective for dsDNA binding, requiring ~10-fold higher concentrations of protein to achieve complete binding site occupancy. Binding of R505E and H507A was reduced ~2-fold, while H507A and H507F were similar in activity to wild-type. In the hydrophobic loop, binding of F464A, W465A and L466A were all reduced 2–3-fold, while increasing the bulk of the side chain of amino acid 464 restored binding to near wild-type levels (Wild-type = F464W > F464L > F464A).

For wild-type E1, efficient unwinding is only achieved at 100% site occupancy, which correlates with the generation of a distinct hydroxyl-radical cleavage pattern (14, see below). Accordingly, we established E1 binding reactions in the presence of ATP with a carefully determined titration to achieve complete site occupancy for all proteins tested. DNA melting was then probed using potassium permanganate sensitivity, and melted DNA regions were revealed on sequencing gels. A sample of the reaction processed for gel-shift/site occupancy analysis is shown below each potassium permanganate melting assay, illustrating the results for dsDNA binding described above. All tested substitutions in the PS1 β H (Figure 6a) melted DNA in the AT-rich region and over the E2 binding site except K506E (and the Ala substitution, data not shown) that was inactive (lanes 2–4 compared with lanes 8–10). For R505E (lanes 5–7) and H507A (lanes 11–13) melting over the A/T-rich region was reduced ~2–2.5-fold, but melting downstream of the E1 binding site at nucleotide 20 (nt. 20, indicated by an arrow) was similar to wild-type. Amino acid substitutions at position 507 with increasingly bulky side chains favoured DNA melting (H507F = H507L > H507A; lanes 11–20). For DNA melting, the hydrophobic loop was less tolerant to amino acid substitution (Figure 6b). F464A was completely inactive for DNA melting (lanes 5–7 compared with lanes 2–4), while increasing the bulk of the amino acid side chain recovered minimal but detectable activity (lanes 11–20). W465A was inactive for *ori* melting (lanes 14–16), while L466A supported reduced activity (lanes 17–19).

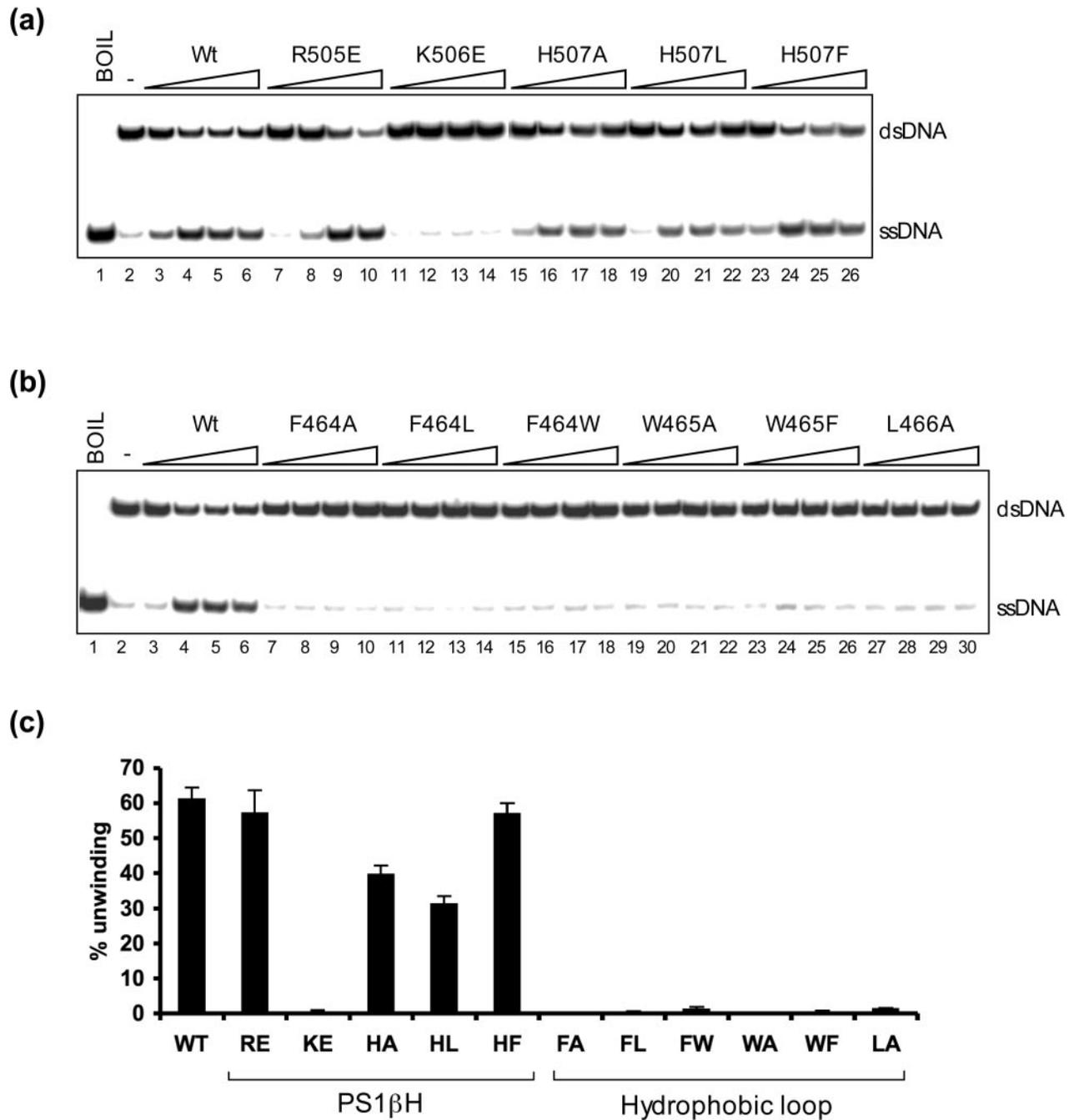
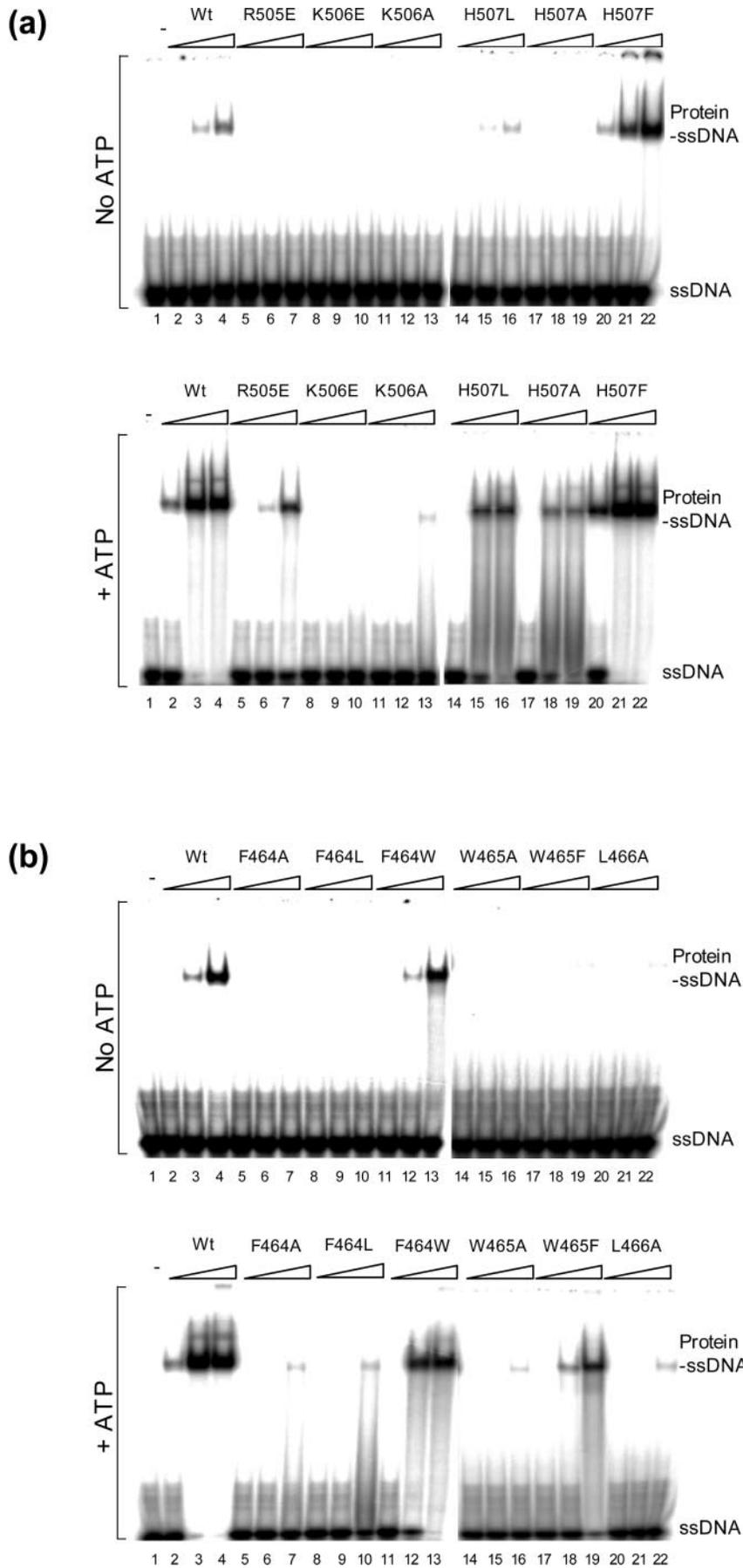


Figure 4. Helicase activity of E1HD and PS1βH and hydrophobic loop mutants. Helicase activity was measured by the ability of protein to displace a radiolabelled 70 bp DNA strand annealed to a 115 bp strand incorporating a 45 bp 3' overhang. Reaction products were separated on native polyacrylamide gels. (a) Helicase activity of mutants in the PS1βH motif of BPV. Lane 1, the boiled substrate indicating mobility of the displaced 70mer. Lane 2, native probe, no protein. Substrate was incubated with an increasing protein titration (100, 250, 500, and 1000 nM) and displaced oligo measured after 60 min: Lanes 3–6, wild type. Lanes 7–10, R505E, Lanes 11–14, K506E. Lanes 15–18, H507A. Lanes 19–22, H507L. Lanes 23–26, H507F. For active mutants, a comparison of product release as a function of time substantiated the results shown above (not shown). (b) Helicase activity of mutants in the hydrophobic FWL motif. Lanes 1 and 2, boiled substrate and native probe, respectively. Lanes 3–6, wild-type protein. Lanes 7–10, F464A. Lanes 11–14, F464L. Lanes 15–18, F464W. Lanes 19–22, W465A. Lanes 23–26, W465F. Lanes 27–30, L466A. All mutants in this region were inactive or virtually inactive. (c) Summary of helicase activity for mutants shown in a and b, determined at 500 nM protein, 60 min incubation.

High-resolution footprinting of E1 and mutants

As described above, *ori* DNA melts when more E1 molecules are loaded on a minimal E1-*ori* complex in the presence of ATP, and this correlates with the generation of a distinct

DNA footprint. We determined whether defects in DNA melting also correlate with defects in DNA binding using high-resolution hydroxyl-radical (OH) footprinting. In Figure 7a, lane 3 demonstrates the periodic protection on



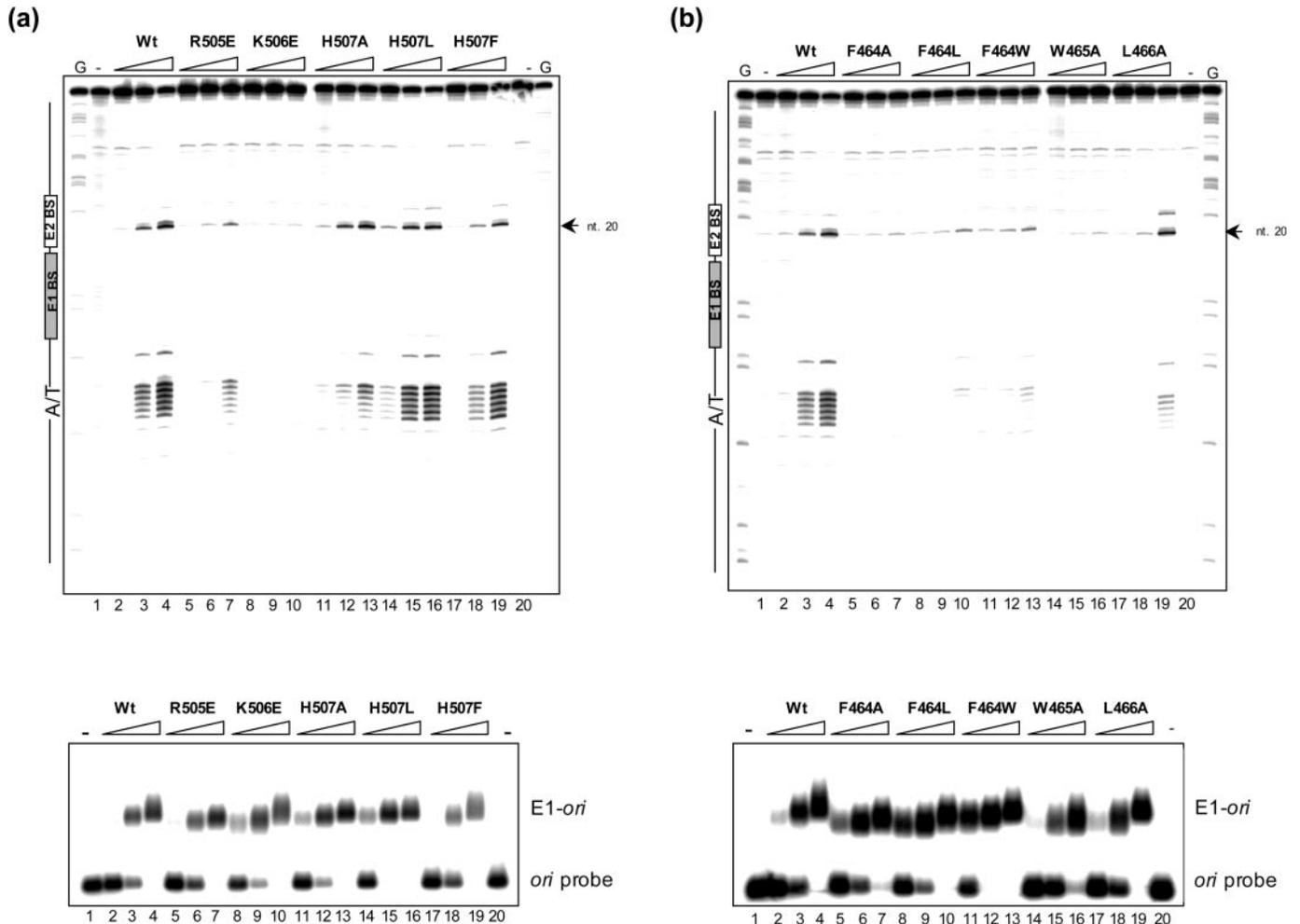


Figure 6. Origin DNA melting by E1 and selected E1HD mutants. (a) Melting by E1 and PS1βH mutants as indicated above the lanes. Reactions were assembled and a sample processed for gel-shift, shown below, in parallel with the KMnO₄ melting assay (equivalent lane/reaction numbering). Lanes 1 and 20 show the reactivity of free probe. Lanes 2–4 show melting by wild-type E1 (10–50 nM) over the A/T-rich region and E2 binding site (20 nt indicated). Lanes 5–7, R505E, 25–100 nM. Lanes 8–10, K506E, 150–600 nM. Lanes 11–13, H507A, 20–80 nM. Lanes 14–16, H507L, 25–100 nM. Lanes 17–19, H507F, 10–50 nM, as wild-type. G is a Maxam–Gilbert DMS marker ladder. (b) *Ori* melting by E1 and hydrophobic loop mutants as indicated. Lanes 1 and 20, reactivity of free probe; 2–4 wild-type E1, as above. Lanes 5–7, F464A, 25–125 nM. Lanes 8–10, F464L, 25–125 nM. Lanes 11–13, F464W, 25–125 nM. Lanes 14–16, W465A, 25–125 nM. Lanes 17–19, L466A, 25–125 nM.

ori generated by E1 binding in the absence of ATP. This protection is identical to that resulting from E1 binding at low protein concentrations with ATP and, therefore, demonstrates the signature protection of the precursor to the higher-order E1–*ori* melting complex (14). DNA binding at high E1 concentrations with ATP results in complete protection over and beyond the E1 binding site, which diminishes at the flanks (Figure 7a, lane 7 compared with lane 3). When footprints of mutants in the PS1βH motif (lanes 8–23) are compared with wild-type, it becomes clear that failure to melt DNA

results from an impairment in progressive E1 binding. K506E, the only mutant inactive for DNA melting, generates a periodic protection pattern indicating that E1 loading is stalled at the level of the precursor E1–*ori* complex (lane 13 compared with lanes 7 and 3). R505E and all H507 mutants melt DNA, albeit at a reduced extent, and generate the block protection centred over the E1 binding site (lanes 10, 16, 19 and 22, compared with lane 7). We do note, however, a subtle difference at the boundaries of the H507A and H507L protections, which do not extend to the same extent as

Figure 5. Binding of E1HD and PS1βH and hydrophobic loop mutants to ssDNA, without and with ATP/Mg²⁺ (indicated simply as ATP), top and bottom panel, respectively. The ssDNA substrate was a 30mer chosen at random, the protein titration series is 250, 1000 and 2000 nM. (a) ssDNA binding of E1HD and PS1βH mutants as indicated above the lanes. Wild-type is on the left, lanes 2–4; lane 1 is free probe (ssDNA). The ssDNA binding activity of wild-type was enhanced in the presence of ATP (lanes 2–4 lower panel), and R505E recovered partial binding activity. K506A was marked by the appearance of a complex with aberrant mobility (lane 13). Without ATP, H507 substitutions to Leu and Ala incrementally abolished ssDNA binding (lanes 14–19), while the H507F substitution (lanes 20–22) improved activity compared with wild-type. With ATP, ssDNA binding of H507L and H507A improved but was indicative of complex instability. (b) ssDNA binding of E1HD and hydrophobic loop mutants as indicated above the lanes. Without ATP, DNA binding was practically abolished for all mutants except the conservative substitution F464W. With ATP, below, activity was enhanced for all mutants but most noticeably the conservative substitutions F464W and W465F.

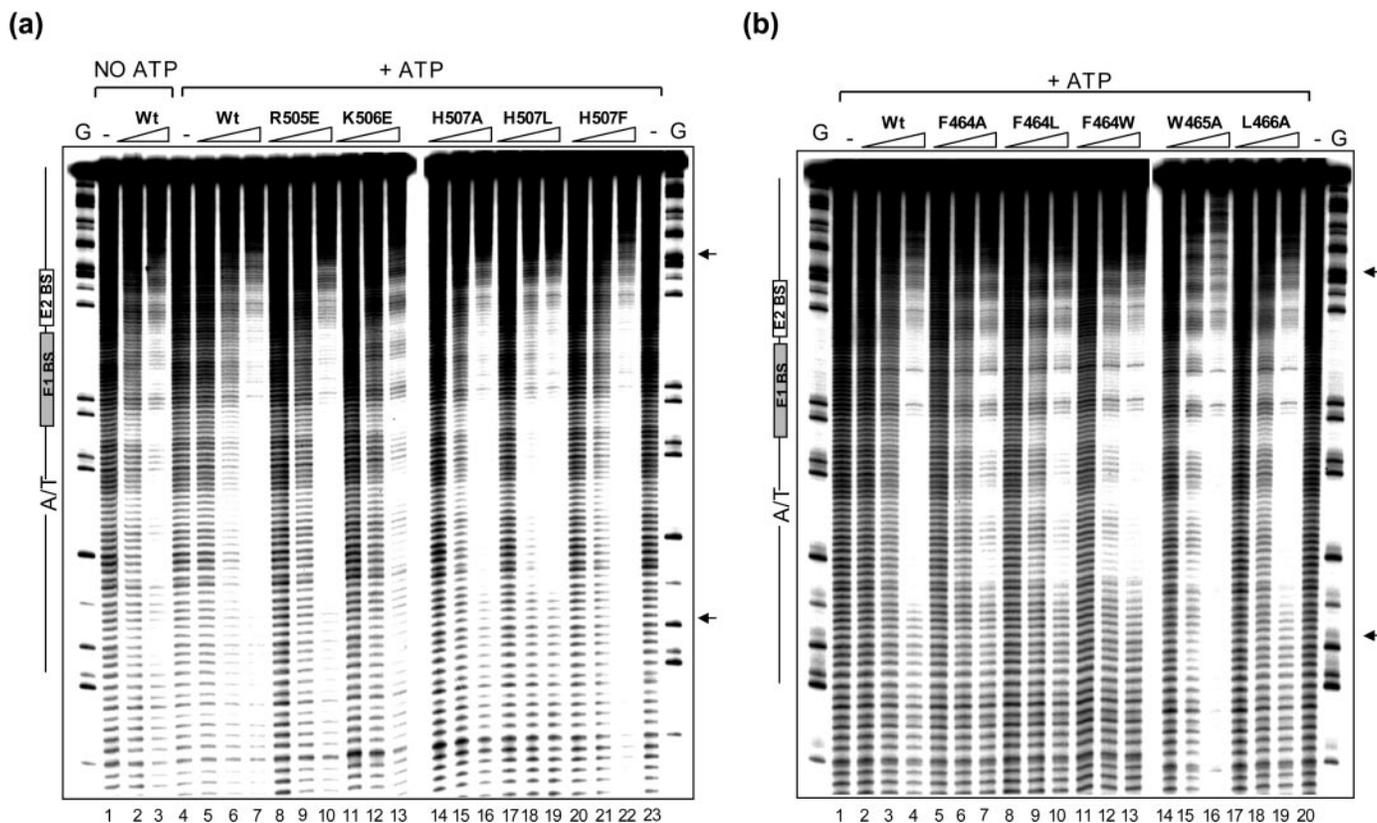


Figure 7. Hydroxyl radical footprinting of E1 and selected E1HD mutants. **(a)** Footprints of E1 with and without ATP and PS1 β H mutants as indicated above the lanes. Lanes 1–3, reactions without ATP; free probe, 25 and 50 nM wild-type E1. All other reactions contained ATP/Mg²⁺ (5 mM). Lanes 4 and 23, reactivity of free probe. Lanes 4–7, wild-type E1, 10–50 nM. Lanes 8–10, R505E, 25–125 nM. Lanes 11–13, K506E, 75–300 nM. Lanes 14–16, H507A, 20–80 nM. Lanes 17–19, H507L, 25–100 nM. Lanes 20–22, H507F, 15–60 nM. G is a Maxam–Gilbert DMS marker ladder. **(b)** Footprints of hydrophobic loop mutants as indicated above the lanes. Lanes 1 and 20, reactivity of free probe. Lanes 2–4, wild-type E1, 10–50 nM. Lanes 5–7, F464A, 25–125 nM. Lanes 8–10, F464L, 25–125 nM. Lanes 11–13, F464W, 25–100 nM. Lanes 14–16, W465A, 50–200 nM. Lanes 17–19, L466A, 30–125 nM.

wild-type (indicated by the two arrows, compare lanes 7 and 19 for example). For H507F, the flanking protection is enhanced compared with wild-type (lane 7 compared with lane 22), and this is probably a result of promiscuous DNA binding.

Mutants in the hydrophobic loop behaved similarly (Figure 7b). The periodic protection pattern emerges for F464A (lane 7 compared with lane 4), a mutant completely inactive for origin melting (Figure 6b). As the bulk of the 464 side chain is increased (F464L and F464W, lanes 10 and 13), the periodic pattern of protection diminishes and this correlates with a recovery of melting activity, although poor. Mutant W465A appeared to behave differently from all other mutants. Although DNA binding activity is reduced with ATP, the mutant exhibits a promiscuous mode of DNA binding, protecting the DNA sequences beyond the initial site of binding. The mutant L466A bound DNA in a manner indistinguishable from wild-type (lane 19 compared with lane 4), and this correlated with significant detectable DNA melting activity (Figure 6b).

DISCUSSION

We have investigated whether or not conserved amino acids in the BPV E1HD are required for both site specific DNA melting and processive DNA unwinding. The roles of two

sequences, corresponding to the PS1 β H and a hydrophobic loop that project into the central cavity of the T-antigen hexamer (Figure 1), have been probed by mutagenesis. Our results show that the conserved amino acids of both these features are critical for processive DNA unwinding (helicase activity) by E1. As summarized in Table 2, in the PS1 β H only mutants K506A and K506E failed to unwind DNA. A defect in ssDNA binding provides the most likely explanation for their loss of function. We were unable to test the effect of the conservative substitution K506R since this mutant failed to express. All H507 mutants unwind DNA, even the equivalent Ala substitution that is inactive in T-antigen (18). Again, a defect in ssDNA binding may account for the reduced activity of H507L and H507A. Likewise, the activity of mutants in the hydrophobic loop supports the notion that it forms part of a DNA binding site that is intolerant even to subtle structural changes in order to be functional in processive unwinding. This is best illustrated by the proteins with conservative substitutions, F464W and W465F. They retain ssDNA binding activity but have barely detectable helicase activity. The inactivity of mutant L466A also supports the proposal. The equivalent residue L461 in T-antigen is positioned in an internal hydrophobic structural core rather than at the surface of the hexameric channel. The Ala substitution would be predicted to shrink the protein core, altering the positioning of the hydrophobic loop in a

Table 2. Summary of activities of E1HD mutants relative to wild-type

E1HD mutant	Helicase activity	ATPase		ssDNA binding		Oligomerization	dsDNA binding ^a	Ori melting ^a
		-ssDNA	+ssDNA	-ATP	+ATP			
R505E	+++	-	++	-	+	+	++	+
K506A	-	+	>wt	-	-	++	-/+	-
K506E	-	>wt	>wt	-	-	n.d.	-/+	-
H507A	++	++	>wt	-	++	++	++	+
H507L	++	>wt	+++	+	++	n.d.	++	+++
H507F	+++	+++	+++	>wt	+++	n.d.	+++	+++
F464A	-	-/+	+++	-	-/+	++	++	-
F464L	-	-/+	>wt	-	-/+	n.d.	++	-/+
F464W	-/+	>wt	>wt	+++	++	+++	+++	+
W465A	-	-	+++	-	-/+	-	++	-
W465F	-/+	+	>wt	-	++	+++	++	n.d.
L466A	-	-	++	-	-/+	-	++	+

n.d., not determined; wt, wild-type.

^aAll activities were measured in the context of the E1HD except dsDNA binding and *ori* melting.

way that could impair DNA binding. Together, these results indicate that K506 of the PS1 β H and F464 and W465 of the hydrophobic loop are the principal components of a composite DNA binding structure in the BPV E1 helicase.

Among all the mutants tested the helicase activity of R505E is surprising since it is impaired in all activities, including ssDNA binding, required for unwinding. Furthermore, high protein concentrations did not inhibit unwinding. Many helicases, including bacteriophage T7 helicase gp4A' (23), require a substrate with an ssDNA tail that is excluded from the protein complex in order to initiate unwinding *in vitro*, suggesting that a structural feature of the tailed substrate is recognized and bound by the helicase. This is also true of T-antigen (24) and the E1HD (C. Sanders and S. Castella, unpublished data), suggesting that the proficiency of R505E could be explained in terms of this recognition reaction rather than oligomerization on ssDNA. Oligomerization of E1HD on the ssDNA strand that must otherwise be excluded during unwinding could explain why helicase activity is inhibited at high protein concentrations (Figure 4) but not in the case of R505E. Also of note is the fact that many mutants only demonstrated significant ATP hydrolysis in the presence of high ssDNA concentrations (2000 times higher than that used in ssDNA binding assays, 3:1 molar ratio of protein:DNA). Since it has been demonstrated that ssDNA promotes hexamerization and ATPase activity of full-length E1 (22), it is likely that high DNA concentrations can drive hexamerization and, hence, ATPase activity of the oligomerization defective mutants.

Our data indicate that there are mechanistic similarities between *ori* DNA melting and unwinding by E1, since they implicate the same residues required for helicase activity in origin DNA binding (Table 2). Furthermore, subtle differences were observed in origin DNA binding with and without ATP (data not shown), consistent with cofactor (ATP)-dependent modulation of a non-specific dsDNA binding site in E1-*ori*. When we probed *ori* melting by the PS1 β H/hydrophobic loop mutants (Figure 6) it was clear that the same residues involved in processive unwinding/dsDNA binding were required for *ori* melting, and, hence, the activity of each mutant in DNA melting could be explained also by impaired DNA binding site function. Generation of potassium

permanganate hypersensitivity correlates with progression of the hydroxyl radical protection from a periodic pattern to the block protection associate with the E1-*ori* melting complex. A defect in E1 binding events is clearly demonstrated for the mutants K506E and F464A that are inactive for melting (Figure 6a and b) and whose footprints compare only with the E1-*ori* complex that forms in the absence of ATP (Figure 7), or low E1 concentrations with ATP (11). In the E1-*ori* melting complex, the DNA over the E1 binding site is not melted but binding of additional E1 molecules, via the origin binding domain (OBD), is evident by the change in OH footprint. The implications of these observations are that binding of the E1 molecules that generate the final unwound structure and DNA melting are coordinated ATP-dependent events and that the E1HD and OBD cooperate to melt DNA.

We have, therefore, linked the processes of DNA melting and helicase unwinding in BPV E1 based on the common requirement for amino acids in a non-specific DNA binding site. Structures of the related T-antigen helicase in the ATP-bound, ADP-bound and nucleotide-free forms show a high degree of conformational flexibility. Encircling of DNA in a central channel is considered key in explaining the enzyme's action. In the nucleotide-free form the channel opening is at its widest and at its narrowest with ATP. The smaller tier or domain 1 (D1) of the T-antigen hexamer is linked to the larger tier (D2/D3) via a helical 'spring', H5, and the interface between monomers is flexible allowing rotation of the tiers. The PS1 β H and hydrophobic loop also move longitudinally upon ATP hydrolysis (16), and low-resolution models of T antigen binding to the SV40 *ori* show two hexamers bound in a head to tail array with the hydrophobic loop/PS1 β H positioned near the extremities (25,26). Although there are no structures of hexameric helicases bound to DNA, crystal structures of PcrA and Rep helicases bound to DNA show conserved aromatic residues binding directly to ssDNA by base stacking and interactions of electropositive residues with the DNA phosphodiester backbone (27,28). It is, therefore, likely that the PS1 β H and hydrophobic loop bind the DNA in a similar manner, allowing it to be drawn into the hexameric chamber where strand separation could occur (16). In BPV it is proposed that the *ori*-melting

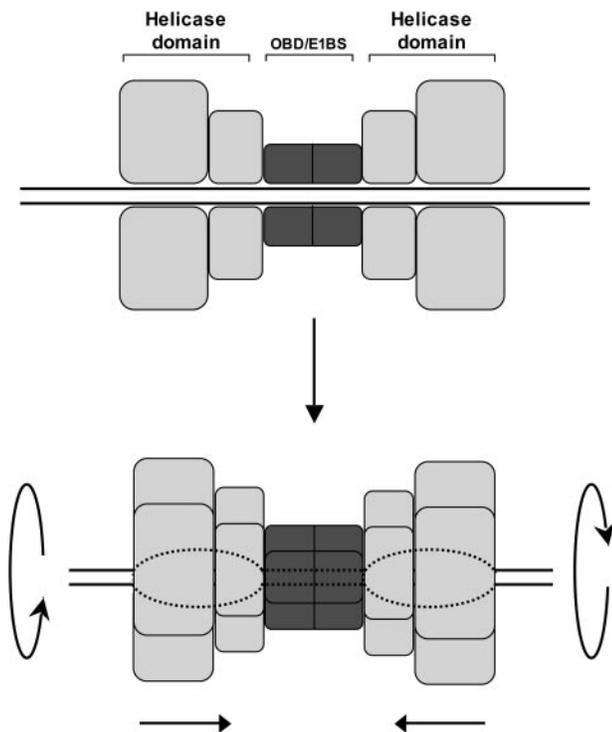


Figure 8. Proposed model for DNA melting by the E1-ori initiator complex. With ATP, E1 binds *ori* initially as a tetramer then as a double trimer (14). Progression from the tetramer to the double trimer is accompanied by a characteristic change in the hydroxyl radical footprint pattern (11). In the *ori*-melting complex the specific origin binding domain is bound to the E1 binding site and the helicase domain engages the DNA upstream and downstream of the E1 site. E1 domains are shown in a simple cross-sectional view on a representation of the *ori* DNA. Rotation of the two tiers of the E1HD (assumed on the basis of the SV40 T-antigen structure) and the inward movement of the PS1 β H/hydrophobic loop draws the DNA inward to unwind it. The large central chamber accommodates the unwound DNA.

complex is a double trimer of E1 molecules arranged in a head-to-tail array on the origin DNA (14). By analogy with T-antigen it is possible to imagine the two E1 trimers immobilized by the OBD rotating against each other, while the DNA is locked in the hydrophobic loop/PS1 β H binding site, drawn inwards and unwound (Figure 8). The mutational analysis that we have described, coupled with the ability to assemble the E1 helicase stepwise *in vitro*, should make this model eminently testable.

ACKNOWLEDGEMENTS

This work was supported by grants from Yorkshire Cancer Research and BBSRC to C.M.S. We thank Jon Sayers and Fred Antson for help with molecular modelling, and Mark Meuth and Jon Sayers for critical reading of the manuscript. Funding to pay the Open Access publication charges for this article was provided by Yorkshire Cancer Research.

Conflict of interest statement. None declared.

REFERENCES

- Dean, F.B., Bullock, P., Murakami, Y., Wobbe, C.R., Weissbach, L. and Hurwitz, J. (1987) Simian virus 40 (SV40) DNA replication: SV40

- large T antigen unwinds DNA containing the SV40 origin of replication. *Proc. Natl Acad. Sci. USA*, **84**, 16–20.
- Wold, M.S., Li, J.J. and Kelley, T.J. (1987) Initiation of simian virus 40 DNA replication *in vitro*: large-tumour antigen- and origin-dependent unwinding of the template. *Proc. Natl Acad. Sci. USA*, **84**, 3643–3647.
- Yang, L., Mohr, I., Fouts, E., Lim, D.A., Nohaile, M. and Botchan, B. (1993) The E1 protein of bovine papillomavirus 1 is an ATP-dependent DNA helicase. *Proc. Natl Acad. Sci. USA*, **90**, 5086–5090.
- Seo, Y.-S., Muller, F., Lusky, M. and Hurwitz, J. (1993) Bovine papillomavirus (BPV)-encoded E1 protein contains multiple activities for BPV DNA replication. *Proc. Natl Acad. Sci. USA*, **90**, 702–706.
- Sanders, C. and Stenlund, A. (1998) Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor. *EMBO J.*, **17**, 7044–7055.
- Lusky, M., Hurwitz, J. and Seo, Y.-S. (1994) The bovine papillomavirus E2 protein modulates the assembly of but is not stably maintained in a replication-competent multimeric E1-replication origin complex. *Proc. Natl Acad. Sci. USA*, **91**, 8895–8899.
- Sedman, T., Sedman, J. and Stenlund, A. (1997) Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus. *J. Virol.*, **71**, 2887–2896.
- Thorner, L., Lim, D. and Botchan, M. (1993) DNA-binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects. *J. Virol.*, **67**, 6000–6014.
- Winokur, P.L. and McBride, A.A. (1996) The transactivation and DNA binding domains of the BPV-1 E2 proteins have different roles in cooperative origin binding with the E1 protein. *Virology*, **221**, 44–53.
- Stenlund, A. (2003) E1 initiator DNA-binding specificity is unmasked by selective inhibition of non-specific DNA binding. *EMBO J.*, **22**, 954–963.
- Sanders, C.M. and Stenlund, A. (2000) Transcription factor dependent-loading of the E1 initiator reveals modular assembly of the papillomavirus origin melting complex. *J. Biol. Chem.*, **275**, 3522–3524.
- Chen, G. and Stenlund, A. (2002) Sequential and ordered assembly of E1 initiator complexes on the papillomavirus origin of DNA replication generates progressive structural changes related to melting. *Mol. Cell. Biol.*, **22**, 7712–7720.
- Gillette, T.G., Lusky, M. and Borowiec, J.A. (1994) Induction of structural changes in the bovine papillomavirus type 1 origin of replication by the viral E1 and E2 proteins. *Proc. Natl Acad. Sci. USA*, **91**, 8846–8850.
- Schuck, S. and Stenlund, A. (2005) Assembly of a double hexameric helicase. *Mol. Cell*, **20**, 377–389.
- Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Joachimiak, A., Szakonyi, G. and Chen, X. (2004) Structure of the replicative helicase of the oncoprotein SV40 tumour antigen. *Nature*, **423**, 512–518.
- Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X. (2004) Mechanism of conformational change for a replicative hexameric helicase of SV40 large tumour antigen. *Cell*, **119**, 47–60.
- VanLoock, M.S., Alexandrov, A., Yu, X., Cozzarelli, N.R. and Egelman, E.H. (2002) SV40 T antigen structure: domain organisation and DNA-induced conformational changes. *Curr. Biol.*, **12**, 472–476.
- Shen, J., Gai, D., Patrick, A., Greenleaf, W.B. and Chen, X. (2005) The roles of the residues on the channel β -hairpin and loop structures of simian virus 40 hexameric helicase. *Proc. Natl Acad. Sci. USA*, **102**, 11248–11253.
- Abbate, E.A., Berger, J.M. and Botchan, M.R. (2004) The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes. Dev.*, **18**, 1981–1996.
- Iggo, R.D. and Lane, D.P. (1989) Nuclear protein p68 is an RNA-dependent ATPase. *EMBO J.*, **8**, 1827–1831.
- Maxam, A.M. and Gilbert, W. (1979) A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA*, **74**, 560–564.
- Sedman, J. and Stenlund, A. (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J. Virol.*, **72**, 6893–6897.
- Hacker, K.J. and Johnson, K.A. (1997) A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. *Biochemistry*, **36**, 14080–14087.
- SenGupta, D. and Borowiec, J.A. (1992) Strand-specific recognition of a synthetic DNA replication fork by the SV40 large tumour antigen. *Science*, **256**, 1656–1661.

25. Gai,D., Li,D., Finkielstein,C.V., Ott,R.D., Taneja,P., Fanning,E. and Chen,X.S. (2004) Insights into the oligomeric states, conformational changes, and helicase activities of SV40 large tumour antigen. *J. Biol. Chem.*, **279**, 38592–38959.
26. Weisshart,K., Friedl,S., Taneja,P., Nasheuer,H-P., Schlott,B., Grosse,F. and Fanning,E. (2004) Partial proteolysis of simian virus 40 T antigen reveals intramolecular contacts between domains and conformation changes upon hexamer assembly. *J. Biol. Chem.*, **279**, 38943–38951.
27. Velankar,S.S., Soutanas,P., Dillingham,M.S., Subramanya,H.S. and Wigley,D. (1999) Crystal structure of complexes of PcrA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.
28. Korolev,S., Hsieh,J., Gauss,G.H., Lohman,T.M. and Waksman,G. (1997) Major domain swivelling revealed by the crystal structure of complexes of *E.coli* Rep helicase bound to single-stranded DNA and ADP. *Cell*, **90**, 635–647.