Infection by viruses, including herpes simplex virus-1 (HSV-1), and cellular stresses cause widespread disruption of transcription termination (DoTT) of RNA polymerase II (RNAPII) in host genes. However, the underlying mechanisms remain unclear. Here, we demonstrate that the HSV-1 immediate early protein ICP27 induces DoTT by directly binding to the essential mRNA 3’ processing factor CPSF. It thereby induces the assembly of a dead-end 3’ processing complex, blocking mRNA 3’ cleavage. Remarkably, ICP27 also acts as a sequence-dependent activator of mRNA 3’ processing for viral and a subset of host transcripts. Our results unravel a bimodal activity of ICP27 that plays a key role in HSV-1-induced host shutoff and identify CPSF as an important factor that mediates regulation of transcription termination. These findings have broad implications for understanding the regulation of transcription termination by other viruses, cellular stress and cancer.
RNA polymerase II (RNAPII) transcription termination is an essential step in eukaryotic gene expression, but its mechanism remains elusive. Two models have been proposed, the allosteric model and the torpedo model. In both models, poly(A) site (PAS) and the mRNA 3′ processing machinery are essential. Most mammalian PAS consists of an A(U/A)/UAAA hexamer, a U/GU-rich downstream element, and various auxiliary sequences. These RNA sequences recruit the mRNA 3′ processing factors CPSF, CstF, and CFIm, CFIm and the poly(A) polymerase to assemble the mRNA 3′ processing complex. Within this complex, mRNA 3′ processing takes place in two steps, an endonucleolytic cleavage and the subsequent addition of a poly(A) tail. Many but not all mRNA 3′ processing factors are required for transcription termination.

Recent studies have revealed that RNAPII transcription termination is regulated. Through nascent RNA sequencing, we first demonstrated that lytic herpes simplex virus-1 (HSV-1) infection induces widespread disruption of transcription termination (DoTT) in host genes. The Steitz group subsequently reported that osmotic stress induces transcription downstream of human genes (DoGs). Further studies demonstrated that influenza virus infection and other cellular stresses, such as heat shock, also cause widespread DoTT and that the genes affected by virus- and stress-induced DoTT are highly correlated.

Importantly, pervasive DoTT was observed in renal cell carcinoma due to the loss of the histone methyltransferase SETD2. The mechanisms of virus- or stress-induced DoTT/DoGs, however, remain poorly defined.

HSV-1 is a DNA virus that causes a number of diseases ranging from painful skin lesions to encephalitis. Like other viruses, HSV-1 hijacks host factors to facilitate productive infection while efficiently shutting down host gene expression. Its tegument protein VP16 recruits the host transcription machinery to express immediate early genes, including ICP0, ICP4, ICP27, and ICP47. ICP4, in turn, activates the transcription of early and late genes.

ICP27 regulates viral gene expression through several mechanisms, including splicing, 3′ processing, and mRNA export. Earlier studies provided evidence that ICP27 inhibits splicing. As most viral genes are intronless, ICP27-mediated splicing inhibition was proposed to specifically block host gene expression. Recent transcriptomic analyses, however, revealed that neither HSV-1 infection nor ICP27 overexpression globally inhibits splicing, but rather modulates alternative splicing of a subset of cellular genes, suggesting that splicing regulation is not a major mechanism for HSV-1-induced host shut-off.

Here we dissect the molecular mechanism of HSV-1-induced DoTT. Our data show that ICP27 blocks transcription termination of host genes by inhibiting mRNA 3′ processing. Meanwhile ICP27 can act as a sequence-dependent activator of mRNA 3′ processing for viral transcripts. These results reveal a critical role for ICP27 in mediating host shut-off during HSV-1 infection.

**Results**

The role of immediate early proteins in HSV-1-mediated DoTT. Our previous transcriptome analysis of HSV-1-infected cells detected widespread DoTT in host genes. To characterize the timing of DoTT, we monitored transcriptional activities at different time points post-infection by 4-thiouridine (4sU) pulse-labeling and high-throughput sequencing (4sU-seq). Significant DoTT was observed as early as 2 h post-infection (h.p.i.) and steadily increased thereafter. To identify the responsible viral gene(s), we infected primary human fibroblast cells with mutant viruses lacking key viral regulators (immediate early genes and vhs) and quantified DoTT by 4sU-seq. While knockout mutants of ICP0 and ICP22 as well as the key viral host shut-off protein vhs still induced DoTT, this was substantially reduced upon infection with viruses deficient for ICP4 (ΔICP4 or TsK), a temperature-sensitive mutant of ICP4) or ICP27 (ΔICP27) (Fig. 1b). Among these viral genes, deletion of ICP27 resulted in the greatest decrease in DoTT (Fig. 1b). Previously, we observed a strong linear correlation between DoTT and the variance in cellular mRNA expression changes during infection. This correlation was essentially maintained for the mutant viruses including Δvhs, ΔICP22, ΔICP0 as well as ΔICP4 and TsK (Fig. 1c). DoTT induced by infection with the latter two mutant viruses matched DoTT observed 3–4 h.p.i. with wild-type (WT) HSV-1. In contrast, DoTT for ΔICP27 was substantially reduced compared to the observed alterations in host gene expression infected with the WT HSV-1 (Fig. 1c), suggesting that ICP27 plays a major role in HSV-1-induced DoTT. Two specific examples of DoTT in cells infected with the various mutant viruses are shown in Fig. 1d (SRSF3) and Supplementary Fig. 1 (DDX5). In contrast to uninfected cells, transcription continued over 30 kilobases (kb) downstream of the PAS in both genes. While DoTT was observed for all mutant viruses including ΔICP4, TsK and ΔICP27, read-through transcription protruded substantially less in ΔICP27 infection. These data suggest that multiple factors contribute to HSV-1-induced DoTT and that ICP27 plays a major role in this process.

In absence of ICP4, viral gene expression is restricted to the four remaining immediate early genes (ICP0, ICP22, ICP27, and ICP47). To directly assess the role of ICP27 in DoTT, we employed a range of deletion mutants that lack ICP4 in combination with various other immediate early gene.

We quantified DoTT for two marker genes (SRSF3 and DDX5) by qRT-PCR upon infection of primary human fibroblasts with the respective mutant viruses (Fig. 1e). While deletion of ICP4 only resulted in a modest (~2-fold) reduction in read-through, knockout of ICP27 significantly reduced read-through transcription by >20-fold. This was particularly prominent when comparing the two mutants d103 (expressing only ICP0 and ICP27) and d106 (expressing only ICP0), which only differ by the expression of ICP27. Importantly, DoTT could not be rescued by increasing the dose of infection of both d106 and d107 from a multiplicity of infection (MOI) of 5–20, thereby excluding dose effects on the observed differences. We conclude that ICP27 mediates HSV-1 induced DoTT.

ICP27 is sufficient to induce DoTT. We next asked whether ectopic expression of ICP27 alone was sufficient to induce DoTT. To this end, we expressed ICP27 in HeLa cells by transient transfection and performed 4sU-seq. Interestingly, we observed significant DoTT upon transfection of ICP27 but not control (empty vector) in many host genes, including KDM4C and SPTSSA (Fig. 2a), similar to HSV-1-induced DoTT. Genome-wide analysis confirmed extensive transcriptional activity downstream of the normal transcript end site (TES) in ICP27-expressing cells (Fig. 2b), albeit that the 4sU-seq signal density was significantly less compared to that in HSV-1 infected cells (Fig. 2b). In addition to the extent of DoTT, we also analyzed the pattern of genes that displayed DoTT induced by HSV-1 or transient transfection of ICP27. Of note, 65% (701 genes) of genes with significant DoTT (5-fold change in 4sU-seq signal downstream/upstream of PAS) in ICP27-expressing cells also displayed similar defect in HSV-1-infected cells (Fig. 2c). Together, these results strongly suggest that ICP27 by itself is sufficient for inhibiting RNAPII transcription termination and is a major contributor of HSV-1-induced DoTT.

ICP27 interacts specifically with CPSF. To understand how ICP27 inhibits RNAPII transcription termination, we first
Fig. 1 ICP27 is a major contributor for HSV-1-induced disruption of transcription termination. a Boxplots showing the distribution of read-through at different time points post-infection during HSV-1 wild-type infection (time-course). b Boxplots showing the distribution of read-through in cells infected with wild-type and various mutant viruses at 7–8 h (WT, Δvhs, ΔICP4, ΔICP27) and 11–12 h (ΔICP22, ΔICP0) post-infection. Please note that read-through at 7–8 h post-infection shown here is from a repeat experiment performed together with the mutant viruses. Read-through at 0–1 h post-infection from the WT time-course is also shown. c Median read-through values for each condition and time-point are plotted against the variance in log2 gene expression (=gene FPKM) fold-changes. The gray curve indicates the result of linear fit on all HSV-1 infection time-points from the time-course experiment. d Mapped 4sU-seq reads for SRSF3 in cells infected with wild-type or various mutant HSV-1 strains. The region where transcription termination occurs in mock-infected cells is shaded. e Primary human fibroblasts infected were infected for 8 h with mutant viruses lacking various immediate early genes. The genes expressed ICP27 co-precipitated CPSF subunits, but not the precipitated proteins by western blotting analysis of the IP samples. Consistent with our IP analyses of HSV-1-infected cells, ectopically expressed ICP27 co-precipitated CPSF subunits, including CstF and CFIm subunits, were not detected at significant levels (Fig. 3b), suggesting that ICP27 specifically interacts with the CPSF complex. Splexekin is a core component of the mRNA 3′ processing machinery and it associates with both CPSF and CstF complex27,28. Intriguingly, little to no splexekin was detected in ICP27 IP samples (Fig. 3b). By contrast, other mRNA 3′ processing factors, including CstF and CFIm subunits, were not detected at significant levels (Fig. 3b), suggesting that ICP27 specifically interacts with the CPSF complex. To test this hypothesis, we characterized the interactions between ICP27 and CPSF. We first validated the interactome mapping results by immunoprecipitating ICP27 following HSV-1 infection and analyzing the precipitated proteins by western blotting. Consistent with our mass spectrometry results, CPSF subunits including CPSF160, CPSF100, CPSF73, CPSF30, Wdr33, and Fip1, all co-precipitated with ICP27 (Fig. 3b). By contrast, other mRNA 3′ processing factors, including CstF and CFIm subunits, were not detected at significant levels (Fig. 3b).
other mRNA 3′ processing factors (Supplementary Fig. 2b and Supplementary Data 1). Furthermore, symplekin again did not co-precipitate with ICP27 (Supplementary Fig. 2b), indicating that ICP27 and symplekin may bind to CPSF complex in a mutually exclusively manner. Together these results demonstrate that ICP27 specifically interacts with CPSF and that this interaction does not require other viral proteins.

We next determined which CPSF subunit(s) directly interact with ICP27. We purified recombinant GST-ICP27 and various mRNA 3′ processing factors, and performed GST pulldown assays. Among the tested 3′ processing factors, GST-ICP27 pulled down Fip1 and CPSF73 of the CPSF complex, but not CstF (Fig. 3c), indicating that ICP27 directly interacts with multiple subunits of the CPSF complex. ICP27 contains a number of known functional domains, including the Leucine-rich region, the RGG box, and the C-terminal zinc finger domain (Fig. 3d, top panel)29. In order to map the domain(s) that mediate CPSF interaction, we expressed Flag-tagged WT or various mutant ICP27 by transient transfection, and performed IP with anti-Flag antibodies. The mutant ICP27 included deletion mutants in which specific domain/regions were deleted, such as dLeu, d1-2, d2-3, d4-5, and d5-6 (Fig. 3d). In addition, m15 is a substitution of PG to LE at amino acids 465, 466, and m16 is a C to L substitution at residue 48830. n406 has a stop codon in all three reading frames at amino acid 406 and n504 has a stop codon in all three reading frames at amino acid 50431. Interestingly, the CPSF subunits, Fip1 and CPSF73, were efficiently co-precipitated by ICP27 mutants lacking the N-terminal region, the NLS, and the RGG box. However, little or no Fip1 or CPSF73 were co-precipitated by ICP27 mutants in which the C-terminal regions were deleted or mutated (m15, m16, n406, and n504, Fig. 3d). We conclude that the C-terminal domain of ICP27 is necessary for mediating interactions with CPSF.

ICP27 represses mRNA 3′ end processing. To test whether ICP27 affects mRNA 3′ processing, we performed in vitro 3′ processing assays. We incubated recombinant ICP27 with nuclear extract (NE) from HeLa cells and assessed mRNA 3′ cleavage of L3, a commonly used PAS RNA derived from adenovirus major late transcript. MBP-MS2, another RNA-binding protein, served as a control. Interestingly, ICP27 inhibited mRNA 3′ cleavage of L3 in a dose-dependent manner while MBP-MS2 had no effect (Fig. 4a). We also tested several other human and viral PASs using the same assay and found that ICP27 inhibited the cleavage of all tested PASs including the PAS of the ICP27 gene itself (Fig. 4b)32. In this reporter, Renilla (Rluc) and Firefly luciferase (Fluc) genes are expressed from a bicistronic mRNA in which translation of the downstream Fluc gene is driven by an internal ribosomal entry site (IRES). A candidate PAS to be tested is incorporated between the two luciferase genes. If cleavage/polyadenylation occurs inefficiently at the candidate PAS, only Rluc is efficiently co-precipitated by ICP27 mutants in which the C-terminal regions were deleted or mutated (m15, m16, n406, and n504, Fig. 3d). We conclude that the C-terminal domain of ICP27 is necessary for mediating interactions with CPSF.

To test the impact of ICP27 on mRNA 3′ processing in living cells, we took advantage of a dual luciferase reporter pPASPORT (Fig. 4b)32. In this reporter, Renilla (Rluc) and Firefly luciferase (Fluc) genes are expressed from a bicistronic mRNA in which translation of the downstream Fluc gene is driven by an internal ribosomal entry site (IRES). A candidate PAS to be tested is inserted between the two luciferase genes. If cleavage/polyadenylation occurs efficiently at the candidate PAS, only Rluc is expressed. In contrast, inefficient 3′ processing will lead to transcription read-through and termination at the strong

---

Fig. 2 ICP27 is sufficient to inhibit RNAPII transcription termination. a 4sU-seq tracks of KDM4C and SPTSSA genes in cells transfected with vector or an ICP27-expressing plasmid. For comparison, 4sU-seq tracks for cells infected with WT or ΔICP27 HSV-1 were also included. Two replicates for each condition are shown. b Metagene analysis of 4sU-seq signals at the transcript end site (TES) in cells transfected with vector or an ICP27-expressing plasmid or infected with HSV-1. c Venn diagram showing the overlap of genes with significant termination defects in cells infected with HSV-1 or transfected with ICP27 overexpression.
pulldown samples were analyzed by western blotting. Dots are factors with known functions in mRNA export and the blue dots are mRNA 3′ processing factors. The plot was based on STRING. Cells infected with WT HSV-1 (KOS) were harvested at different time points post-infection and subjected to immunoprecipitation with an anti-ICP27 antibody. Input and IP samples were analyzed by western blotting.

We next mapped the domain/regions of ICP27 that are necessary for inhibiting mRNA 3′ processing. We over-expressed the WT or various mutant ICP27 and checked their effect on mRNA 3′ processing using the reporter assay. Our results demonstrate that mutant ICP27 with N-terminal deletions, such as dLeu and d1-2, inhibited mRNA 3′ processing to a similar degree as the WT ICP27 while the C-terminal mutants (m15 and n504) lost this ability (Fig. 4d). Additionally, we also compared the activities of the WT and mutant ICP27 in inhibiting host gene transcription termination by qRT-PCR. Consistent with the reporter assay results, the N-terminal deletion mutants of ICP27 significantly suppressed mRNA 3′ processing to a similar degree as the WT ICP27 while the C-terminal mutants (m15 and n504) lost this ability (Fig. 4d). Together these data suggest that the C-terminal domain of ICP27 is required for its interactions with CPSF and its inhibition of mRNA 3′ processing and transcription termination.

downstream β-globin PAS, thereby resulting in the expression of both Rluc and Fluc. The Rluc/Fluc ratio thus provides a quantitative measurement of RNA cleavage/polyadenylation efficiency at the candidate PAS.22–34.

Consistent with our in vitro data, expression of ICP27 significantly suppressed mRNA 3′ processing of all seven different PASs (three human, four viral) we tested (Fig. 4b). As ICP27 is known to regulate mRNA export and translation,22,35, we wanted to determine if the effect of ICP27 observed in our reporter assay was due to its direct role in regulating mRNA 3′ processing. We over-expressed ICP27 and recombinant Fip1, CPSF73, and CstF complex. The input and pulldown samples were analyzed by western blotting. Source data are provided as a Source Data file.
**ICP27 interferes with mRNA 3' processing complex assembly.**

To assess which step of mRNA 3' processing ICP27 interferes with, we first monitored the assembly of the mRNA 3' processing complex, the earliest step of mRNA 3' processing. L3 or ICP27 PAS-containing RNAs were incubated with NE under cleavage/polyadenylation conditions and then resolved by native PAGE. In the absence of ICP27 (Fig. 4e, lanes 1 and 5), the mRNA 3' processing complex (P complex) efficiently assembled on both RNA substrates. In the presence of ICP27, however, the P complex migrated faster on the gel and appeared more diffused (Fig. 4e, compare lanes 2–3 to lane 1 and compare lanes 6–7 to lane 5) although complex formation (overall signals) seemed unaffected. The control protein, MBP-MS2, had no effect. This suggests that ICP27 induces the assembly of an aberrant mRNA 3' processing complex that may be compositionally and/or conformationally distinct from the normal complex.

To directly compare the composition of the mRNA 3' processing complexes assembled in the presence or absence of ICP27, we purified these complexes using an RNA affinity approach. Briefly, PAS-containing RNAs were fused to three tandem copies of the MS2 hairpin (the MS2 hairpins are ~60 nt upstream of the cleavage site). These RNA substrates were first incubated with MBP-MS2, which binds to the MS2 hairpins, and then with NE to allow for the assembly of mRNA 3' processing complexes. These complexes were then pulled down by using amylose beads and analyzed by western blotting. Source data are provided as a Source Data file.

---

**Fig. 4 ICP27 inhibits mRNA 3' processing.**

(a) In vitro cleavage assay with HeLa nuclear extract (NE) using L3 (top panel) or ICP27 (lower panel) PAS pre-mRNA substrate. 5' cleavage product is marked. Different amounts of recombinant MBP-ICP27 (labeled as ICP27) or MBP-MS2 (Ctrl) were added to NE.

(b) Dual luciferase reporter assay. The top panel is a diagram for pPASPORT vector and the site where the PAS to be tested is inserted is marked as PAS (in a red box). Seven distinct PASs were cloned into pPASPORT, co-transfected with a vector or ICP27-expression plasmid. The Rluc/Fluc ratio, which measures cleavage/polyadenylation efficiency at the test PAS, are plotted as mean ± s.d.

(c) pPASPORT assays with L3 and ICP27 PAS. Empty vector, ICP27 or its various mutants were co-expressed with the pPASPORT constructs and renilla/firefly luciferase ratio was plotted as mean ± s.d.

(d) In vitro cleavage/polyadenylation reactions (similar to a) were resolved on a native gel and visualized by phosphorimaging. The mRNA 3' processing complex (P complex) and the heterogenous complex (H complex) are marked.

(e) 3' processing ICP27 interferes with mRNA 3' processing complex assembly. L3 or ICP27 PAS-containing RNAs were incubated with MBP-MS2, which binds to the MS2 hairpins, and then with NE to allow for the assembly of mRNA 3' processing complexes. These complexes were then pulled down by using amylose beads and analyzed by western blotting. By comparing the mRNA 3' processing complexes assembled on L3 and ICP27 PAS with or without recombinant ICP27 protein (Fig. 4f), we confirmed that ICP27 associates with the mRNA 3' processing complex (Fig. 4f, lanes 4 and 8). Furthermore, nearly all mRNA 3' processing factors assembled at similar levels in the mRNA 3' processing complex in the presence or absence of ICP27 (Fig. 4f, compare lanes 3 and 4 or 7 and 8). However, symplekin and CstF64 seemed to display a modest decrease (Fig. 4f, compare lanes 3 and 4 or 7 and 8). Together with our earlier observation that the ICP27-associated CPSF complex lacks symplekin (Fig. 3b), these results indicate that ICP27 interaction with CPSF prevents its stable association with symplekin and CstF. Furthermore, when we resolved the affinity purified P complexes assembled in the absence or presence of ICP27 by glycerol
gradient sedimentation, we observed that the P complexes assembled in the presence of ICP27 failed to reach higher density fractions and spread out into more gradient fractions (Supplementary Fig. 4). This is consistent with our previous observation that the P complex assembled in the presence of ICP27 not only migrated faster but also appeared more diffuse on native gels (Fig. 4e). We conclude that ICP27 does not block the assembly of a functional mRNA 3′ processing complex but rather interferes with mRNA 3′ processing by preventing the assembly of a functional mRNA 3′ processing complex.

ICP27 activates viral mRNA 3′ processing. HSV-1 induces DoTT in thousands of host genes, while transcription termination for HSV-1 as well as some host genes remains mostly unaffected. However, ICP27 inhibits mRNA 3′ processing of all PASs that we tested, including viral PASs (Fig. 4 and Supplementary Fig. 3). To understand the molecular basis for the specificity of HSV-1-induced DoTT, we selected four PAS for further studies. These included two cellular PASs displaying strong HSV-1 induced DoTT (SPTSSA and KDM4C, see Fig. 2a). In addition, we included a host PAS (POLR2A) and a viral PAS (ICP27), both of which showed normal transcription termination in HSV-1-infected cells (Supplementary Fig. 5). When we tested the core sequences of these PASs (from −100 nt to +100 nt relative to the cleavage site) in the pPASPORT assays, we again observed strong inhibition of mRNA 3′ processing by ICP27 (Fig. 5a). As the core PAS sequences were apparently not sufficient to explain the specificity of ICP27-induced transcription termination, we asked if sequences upstream of the PAS (UPS) could determine such specificity. We thus cloned the four aforementioned PASs with 1440 nt UPS from the endogenous genes and tested these extended PASs using the pPASPORT assays. Again, ICP27 strongly inhibited mRNA 3′ processing at the extended SPTSSA and KDM4C PAS (Fig. 5b). By contrast, ICP27 now failed to inhibit mRNA 3′ processing at the extended POLR2A PAS. Strikingly, ICP27 even activated mRNA 3′ processing at the extended ICP27 PAS by over 2-fold (Fig. 5b). We performed 3′ RACE to confirm that, although ICP27 inhibited the usage of the core ICP27 PAS, it stimulated its usage when UPS was included (Supplementary Fig. 6). These data suggest that the UPS plays an important role in determining how ICP27 influences mRNA 3′ processing.

Next, we tested if the UPSs alone are sufficient to determine the mode of ICP27 action. To this end, we generated chimeric PAS between KDM4C and ICP27, in which the UPSs and the PAS were exchanged (Fig. 5c). Although ICP27 inhibited 3′ processing at the core and extended KDM4C PASs (Fig. 5a–b), it had no significant effect when the KDM4C PAS was fused to ICP27 UPS.
ICP27-RNA interactions modulate its effect on 3′ processing.

How do UPSs determine ICP27 activities? It has been proposed that ICP27 binds to GC-rich sequences in the viral mRNAs37. The GC content of the ICP27 UPS was substantially higher than that of the KDM4C UPS (70.5% vs. 44.1%) (Fig. 5c). This is a general feature of viral vs. cellular UPS (68.3% vs. 40.1%) (Fig. 5d).

Additionally, we compared the sequences of host genes that displayed significant DoTT in HSV-1-infected cells to those without significant defects. PAS and surrounding sequences of host genes without DoTT had significantly higher GC content than those with termination defect (Fig. 5e, p value = 8.8e−38 by K-S test). These data provide evidence that there is an anti-correlation between the GC-content of PAS and their susceptibility to ICP27-mediated inhibition. We thus hypothesized that ICP27 does not inhibit mRNA 3′ processing when bound near a PAS.

To test this hypothesis, we mapped the global ICP27-RNA interaction landscape by CLIP-seq (UV crosslinking and immunoprecipitation followed by high throughput sequencing)38. We performed this analysis both in HSV-1-infected and ICP27-overexpressing (O/E) cells (Supplementary Fig. 7). We found that ICP27 preferentially binds cellular mRNAs close to PAS or TESs and, to a lesser degree, at transcription start sites (TSS) (Fig. 6a).

Motif analysis of ICP27 RNA binding sites revealed an enrichment of GC-rich sequences (Fig. 6b), consistent with previous in vitro studies39. Additionally, AATAAA was enriched in ICP27-associated sequences, consistent with the observed strong ICP27-RNA interactions near PAS/TES (Fig. 6a). These results suggest that ICP27 preferentially binds to GC-rich UPS of PAS.

We next determined whether ICP27 differentially binds to PAS of host genes with or without HSV-1-induced DoTT. While low levels of ICP27–RNA interactions were detected near the PAS of genes with DoTT (Fig. 6c, red line), ICP27 CLIP-seq signals were detected more frequently at PAS without DoTT (Fig. 6c, green line). As exemplified by the SRSF3 gene that displayed strong DoTT in HSV-1-infected cells (Fig. 1d, 6e), no significant ICP27 CLIP signal was detected near the SRSF3 PAS (Fig. 6e, left panel).

The fact that ICP27 CLIP signals were observed in the 5′ quarter of the mRNA excludes low expression levels or low sensitivity to be responsible for the lack of signal near PAS. By contrast, GAPDH showed little or no DoTT in HSV-1-infected cells, but strong and selective ICP27–RNA interactions were detected upstream of its PAS (Fig. 6e, right panel). Together our 4S-U-seq and CLIP-seq analyses demonstrate that ICP27 preferentially binds to PASs in genes without DoTT (Fig. 6c).

Since HSV-1 specifically blocks transcription termination of host genes without affecting viral genes, we compared the ICP27–RNA interactions at host and viral PASs. As shown in Fig. 6d, the overall ICP27 CLIP signals (from −1 kb to +1 kb relative to the cleavage site, normalized by RNA levels) were significantly higher at viral PASs than at host PASs, suggesting that ICP27 predominantly binds to viral transcripts. This is consistent with the higher GC-content of the viral genome (Fig. 5d). The anti-correlation between ICP27–RNA interaction and HSV-1-induced DoTT suggests that, once bound to the UPS of a PAS, ICP27 does not induce DoTT.

We directly tested this concept in the following three experiments. First, we synthesized a 100-nt RNA fragment from the UPS of ICP27 (GC content: 63%), which showed no DoTT in HSV-1-infected cells, and another fragment from KDM4C (GC content: 41%), which displayed strong DoTT following HSV-1 infection (Fig. 2a). In vitro gel mobility shift assays demonstrated that recombinant ICP27 strongly binds to the GC-rich ICP27 RNA fragment while much weaker binding was observed for the KDM4C fragment (Supplementary Fig. 8a). Next we inserted one or two tandem copies of these RNA fragments upstream of a human PAS that is inhibited by ICP27 and tested how these fragments modulate the inhibitory effect of ICP27 using pPASPORT assay (Supplementary Fig. 8b). The ICP27 UPS RNA fragment alleviated ICP27-mediated inhibition of mRNA 3′ processing while the KDM4C UPS fragment had little effect (Supplementary Fig. 8b), suggesting that ICP27 interactions with RNA sequences upstream of PASes alleviate ICP27-mediated inhibition of mRNA 3′ processing.

Secondly, we tested if the RNA-binding domain of ICP27, the RGG box, is necessary for its activation function. We repeated the reporter assay as described in Fig. 5e with either the WT ICP27 or the d4-5 mutant, which lacks the RGG box. Consistent with our earlier results, ICP27 activated the extended ICP27 PAS and had little effect on the KDM4C fused to the ICP27 UPS (Fig. 6f). Interestingly, however, the d4-5 mutant failed to activate the extended ICP27 PAS and even inhibited both PAS (Fig. 6f), strongly suggesting that the activation function of ICP27 requires its RGG domain and therefore its RNA-binding activity.

Finally, we tested our model using a tethering assay. We inserted two copies of the BoxB hairpin upstream of a host gene (BASPI) or ICP27 core PAS (172 nt and 121 nt upstream of the cleavage site, respectively) in pPASPORT (Supplementary Fig. 8c).

We then co-expressed these reporters with an empty vector, Flag-ICP27, or ΔN-Flag-ICP27-expressing plasmids. The expression of Flag-ICP27 significantly repressed mRNA 3′ processing at both PASs (Fig. 6g), consistent with our earlier results (Fig. 4b, 5a). By contrast, the expression of ΔN-Flag-ICP27, which was tethered to the BoxB-containing PAS through its interactions with the BoxB sequences39, had no inhibitory effect (Fig. 6g), suggesting that interactions with RNA prevent ICP27 from inhibiting mRNA 3′ processing. Taken together, we conclude that ICP27 is a general inhibitor of mRNA 3′ processing by disrupting the assembly of a functional mRNA 3′ processing complex. However, by interacting with GC-rich RNA sequences upstream of the PAS, it can activate mRNA 3′ processing, most likely by recruiting the CPSF complex to its target mRNAs.

Discussion

Recent studies revealed that transcription termination is regulated by viral infections and cellular stresses13–15,40, and that increased DoTT is associated with cancer16. To understand the underlying mechanisms, we carried out an in-depth study of HSV-1-induced DoTT. We found that the HSV-1 immediate early protein ICP27 is sufficient for inducing DoTT when ectopically expressed in human cells. Mechanistically, ICP27 causes DoTT by inducing the assembly of an aberrant mRNA 3′ processing complex via its interactions with the CPSF complex and blocking/delaying mRNA 3′ cleavage. Importantly, ICP27 can also function as a sequence-dependent activator of 3′ end processing, thereby explaining efficient transcription termination of viral transcripts (Fig. 7). It does so through binding to GC-rich sequences upstream of PAS and most likely promoting recruitment of CPSF and other mRNA 3′ processing factors. As transcription termination significantly impacts mRNA production and the subsequent translation13, the bimodal activity of ICP27 plays a major role in HSV-1-mediated host shutoff.
mRNA 3′ processing factors are also targeted by other viruses. For example, influenza virus infection or ectopic expression of the influenza viral protein NS1 causes transcription termination defects of host genes40,41. NS1 inhibits host mRNA 3′ processing by specifically binding to the CPSF subunit CPSF30 (ref. 42). CPSF contains an RNA-binding module (CPSF30, Wdr33, CPSF160, and Fip1) and a cleavage module (CPSF73, CPSF100, and symplekin)43–45. As CPSF30 directly binds to the poly(A) signal, AAUAAA, NS1–CPSF30 interaction blocks PAS recognition43,44,46,47. Here, we showed that the HSV-1 protein ICP27 directly binds to both modules of CPSF. In contrast to the influenza virus mechanism, however, ICP27 does not block...
EHV, and KSHV, do not have a high GC content (<60%). Third, although ICP27 is conserved among herpes viruses, ICP27 homologs in HCMV (UL69), EHV (SM protein), and KSHV (ORF57) do not have the same RNA-binding domain (RGG box) and their RNA binding specificities have not been fully defined. Taken together, these results suggest that herpesviruses are quite divergent and further studies are needed to fully elucidate their interactions with RNAs and host factors as well as in their impact on transcription termination.

**Methods**

**Cell culture, viruses, and infection.** Human foreskin fibroblasts (HFF, #86031405, purchased from ECACC) were cultured in DMEM, high glucose, pyruvate (Thermo Fisher #14966052) supplemented with 1× MEM non-essential amino acids (Thermo Fisher #11140050), 1 mM additional sodium pyruvate (Thermo Fisher #11360070), 10% (v/v) fetal bovine serum (FBS, Biochrom #S0115), 200 IU/ ml penicillin (pen), and 200 µg/ml streptomycin (strep). HFF293 and HeLa cell lines were cultured in Dulbecco’s modified Eagle medium (DMEM, with 10% FBS. All cells were incubated at 37°C in a 5% (v/v) CO2-enriched incubator. Virus stocks for WT HSV-1 strain 17 as well as the null mutants of vhs52, ICP0 (FXE, strain 17)56, ICP22 (R325, strain F)57, and TsK (strain 17)58 were produced on baby hamster kidney (BHK) cells as described. The temperature-sensitive mutant TsK was grown at 31°C. Infections were subsequently performed at 39°C to render ICP4 inactive. Stocks of the ICP27 null mutant were produced on complementing Vero 2-2 cells. Null mutants of multiple immediate early genes (d100, d103, d106, d107) were a kind gift of Neal DeLuca, grown on complementing cell lines and titrated as described. Cells were infected with an MOI of 10 unless otherwise specified and incubated at 37°C until cells were harvested at the specified time points.

**qRT-PCR.** Total RNA was isolated using Trizol reagent following the manufacturer’s instructions. Reverse transcription was performed using All-in-One cDNA Synthesis Supermix (Biotool) including a mix of hexanucleotide random primers and poly-dT primers. qRT-PCR was performed using the SYBR Green 2x dye (Life Technologies) and an ABI 7500 Real-Time PCR System according to the manufacturer’s instructions. A total of 700 ng of total RNA, 70 µl of 10× Biotinylation Buffer, and 1 µl of thiol-specific biotinylation reaction contained 70 µg of total RNA, 70 µl of 10× Biotinylation Buffer were used in the 25 µl reaction. cDNA was diluted by a factor of 1000 and 1 µl was used to do first round PCR with 3′RACE-R1, Renilla-F1, and Firefly-F1 primers in one reaction. A total of 1 µl of PCR product from first round PCR was used to carry out the second round PCR with primers 3′RACE-R2, Renilla-F2, and Firefly-F2. PCR products were resolved on 2% agarose gels.

**Reporter assay.** PAS sequences were cloned into the multiple cloning sites in pPASPORT. Cells were transfected with plasmids encoding GFP and the indicated mutant ICP27 using polyethyleneimine (PEI). Luciferase activities were measured by using the same method. All reporters assays were carried out with three replicates.

**In vitro cleavage assay.** PAS substrates were synthesized by in vitro transcription with T7 polymerase in the presence of [-α-32P]-UTP. In vitro assays for cleavage/polyadenylation were done without any cellular factors. The ICP27 cDNA was cloned in pGEX-4T-1 for GST-ICP27. In vitro assays for coupled cleavage/polyadenylation were performed with HeLa NE as previously described.

**Protein purification.** The ICP27 cDNA was cloned in pGEX-4T-1 for GST-ICP27 and PET-51B (+) for Strep-ICP27. MBP-ICP27, GST-ICP27, and Strep-ICP27 fusion proteins were expressed in Escherichia coli and purifed using amylase beads (NEB), glutathione sepharose (GE Healthcare), and Strep-Tactin™ Sepharose (IBA Lifesciences), respectively. Fip1, CPSF73, and Csf1DNDs were cloned into pFastBac. The pFastBac and MultiBac constructs were used to produce recombinant baculoviruses using standard procedures. Recombinant baculoviruses were used to infect S9 cells and cells were harvested 2 days post-infection. Recombinant proteins were purified with Cobalt beads per manufacturer’s instructions (Fisher).
Read distribution data around termination sites or gene body were generated using deepTools. Peak distribution data were extracted by Homer software25. The visualization and any post-processing step then were done in Python.

Termination defect was studied based on the ratio of reads mapped on 5 kb downstream of the termination site to the expression of the gene. This ratio (termination ratio) was calculated for each gene and in all conditions, and its change was used as the measure of termination defect. Bedtools26 was used to extract the downstream region, and featureCounts25 generated the read counts. The read counts were normalized by the length of the genes. Genes were then ranked based on the change in termination ratio and then based on their coverage and defect and not defected ones were chosen from this list. We applied a threshold of 5-fold change for the termination ratio and 5 reads per kilobase for coverage to find the significant changes in termination. Genes with a small change in their termination ratio (<1.1 fold) were marked as the ones with no significant change.

Bedtools was used to extract the sequence for each region under study when necessary. Any post processing then was done in Python. Motif analysis was done using DREMI25.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

4u-seq and CLIP-seq data have been deposited to the GEO database (accession number: GSE128753). The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium with the dataset identifier PXD012837. The Source Data for Figs. 3b–d, 4a, e, f, and Supplementary Figs. 2a, b, 3a, b, 4a, b, 6–8 are provided as a Source Data file.

Received: 16 May 2019; Accepted: 11 December 2019;
Published online: 15 January 2020

**References**

1. Richard, P. & Manley, J. L. Transcription termination by nuclear RNA polymerases. Genes Dev. 23, 1247–1269 (2009).
2. Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science 352, aa9926 (2016).
3. Burzowski, S. Connections between mRNA 3' end processing and transcription termination. Curr. Opin. Cell Biol. 17, 257–261 (2005).
4. Luo, W. & Bentley, D. A ribonucleolytic rat torpedoes RNA polymerase II. Mol. Cell 11, 911–914 (2004).
5. Logan, J., Falck-Pedersen, E., Darnell, J. E. & Shenk, T. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta major-globin gene. Proc. Natl Acad. Sci. USA 84, 8306–8310 (1987).
6. Connelly, S. & Manley, J. L. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. 2, 440–452 (1988).
7. West, S., Gromak, N. & Proudfoot, N. J. Human 3' → 5' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature 432, 522–525 (2004).
8. Kim, M. et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature 432, 517–522 (2004).
9. Fong, N., Ohman, M. & Bentley, D. L. Fast ribosome cleavage releases transcripts from RNA polymerase II and aborts co-transcriptional pre-mRNA processing. Nat. Struct. Mol. Biol. 16, 916–922 (2009).
10. Fong, N. et al. Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. Mol. Cell 60, 256–267 (2015).
11. Chan, S., Choi, E. A. & Shi, Y. Pre-mRNA 3' end processing complex assembly and function. Wiley Interdiscip. Rev. RNA 2, 321–335 (2011).
12. Bircz, C. E., Minvielle-Sebastia, L., Lee, R. A., Keller, W. & Proudfoot, N. J. Coupling termination of transcription to messenger RNA maturation in yeast. Mol. Cell 16, 298–301 (1998).
13. Rutfowski, A. J. et al. Widespread disruption of host transcription termination in HSV-1 infection. Nat. Commun. 6, 7126 (2015).
14. Vilborg, A., Passarelli, M. C., Yario, T. A., Tyrosowski, K. T. & Steitz, J. A. Widespread inducible transcription downstream of human genes. Mol. Cell 59, 449–461 (2015).
15. Hennig, T. et al. HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. PLoS Pathog. 14, e1006954 (2018).
16. Grosso, A. R. et al. Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *elife* 4, e08721 (2015).

17. Whitley, R. J. & Roizman, B. Herpes simplex virus infections. *Lancet* 357, 1513–1518 (2001).

18. Sandri-Goldin, R. M. & Mendoza, G. E. A herpesvirus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. *Genes Dev.* 6, 484–486 (1992).

19. Tang, S., Patel, A. & Krause, P. R. Herpes simplex virus ICP27 regulates alternative pre-mRNA polyadenylation and splicing in a sequence-dependent manner. *Proc. Natl. Acad. Sci. USA* 113, 12256–12261 (2016).

20. McGregor, F., Phelan, A., Dunlop, J. & Clements, J. B. Regulation of herpes simplex virus poly (A) site usage and the action of immediate-early protein ICP4 in the early-late switch. *J. Virol.* 70, 1931–1940 (1996).

21. Chen, I.-H. B., Scabica, K. S. & Sandri-Goldin, R. M. ICP27 interacts with the RNA export factor Aly/REF to direct Herpes Simplex Virus Type 1 Intronless mRNAs to the TAP export pathway. *J. Virol.* 76, 12877–12889 (2002).

22. Sandri-Goldin, R. M. ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding viral intronless RNAs through an RGG motif. *Genes Dev.* 12, 868–879 (1998).

23. Hobbs, W. E., DeLuca, N. A. & DeLuca, N. A. Perturbation of cell cycle progression and cellular gene expression as a function of herpes simplex virus ICP0. *J. Virol.* 73, 8245–8255 (1999).

24. Samaniego, L. A. 3’RNA sequester host pre-mRNA metabolism. *PLoS Pathog.* 11, e1004652 (2015).

25. Nojima, T. et al. Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 161, 526–540 (2015).

26. Eaton, J. D. et al. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CTF73 activity. *Genes Dev.* 32, 127–139 (2018).

27. Takagaki, Y. & Manley, J. L. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* 20, 1515–1525 (2000).

28. Shi, Y. et al. Molecular architecture of the human pre-mRNA 3’ processing complex. *Mol. Cell* 33, 363–376 (2009).

29. Tunnudzid, R. B. et al. The structure of the folded domain from the signature multifunctional protein ICP27 from herpes simplex virus-1 reveals an intertwined dimer. *Sci. Rep.* 5, 11234 (2015).

30. Rice, S. A., Long, M. C., Lam, V. & Spencer, Ca RNA polymerase II is aberrantly phosphorylated and localized to viral replication compartments following herpes simplex virus infection. *J. Virol.* 68, 988–1001 (1994).

31. Rice, S. A., Su, L. S. & Knipe, D. M. Herpes simplex virus alpha protein ICP27 possesses separable positive and negative regulatory activities. *J. Virol.* 63, 3399–3407 (1989).

32. Yao, C. et al. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc. Natl. Acad. Sci. USA* 106, 18773–18778 (2009).

33. Zhu, Y. et al. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Genes Dev.* 29, 889–897 (2015).

34. Shui, Y. & Manley, J. L. The end of the message: multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29, 1000–1005 (2015).

35. Sun, Y. et al. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc. Natl. Acad. Sci. USA* 201718723, https://doi.org/10.1073/pnas.1718723115 (2017).

36. Clerici, M., Faini, M., Muckenfuss, L. M., Aebersold, R. & Jinek, M. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat. Struct. Mol. Biol.* 25, 135–138 (2018).

37. Di Giammartino, D. C., Shi, Y. & Manley, J. L. PARP1 represses PAP and inhibits polyadenylation during heat shock. *Mol. Cell* 49, 7–17 (2013).

38. Kolev, N. G. & Steitz, J. A. Symplekin and multiple other polyadenylation factors participate in 3’-end maturation of histone mRNAs. *Genes Dev.* 19, 2533–2539 (2005).

39. Xing, H., Mayhew, C. N., Cullen, K. E., Park-Sarge, O. K. & Sarge, K. D. HSFI modulation of Hsp70 mRNA polyadenylation via interaction with symplekin. *J. Biol. Chem.* 279, 10551–10555 (2004).

40. Gatherer, D. et al. High-resolution human cytomegalovirus transcriptome. *Proc. Natl. Acad. Sci. USA* 108, 19755–19760 (2011).

41. Marciniowska, L. et al. Real-time transcriptional profiling of cellular and viral gene expression during lytic cytomegalovirus infection. *PLoS Pathog.* 8, e1002908 (2012).

42. Sei, E., Wang, T., Hunter, O. V., Xie, Y. & Conrad, N. K. HTS-CLIP analysis uncovers a link between the Kapoosi’s sarcoma-associated herpesvirus ORF57 with a spliceosome component. *Cell Rep.* 7, 1–11 (2016).

43. Hirtart, E. et al. A region of the Epstein-Barr Virus (EBV) mRNA export factor EB2 containing an arginine-rich motif mediates direct binding to RNA. *J. Biol. Chem.* 278, 37790–37798 (2003).

44. Fenwick, M. L. & Everett, R. D. Transfer of UL41, the gene controlling virion-associated host cell shut-off, between different strains of herpes simplex virus. *J. Virol.* 71, 411–418 (1997).

45. Boultell, C., Sadis, S. & Everett, R. D. Herpes simplex virus type 1 immediate-early protein ICP0 and is isolated RING finger domain act as ubiquitin E3 ligases in vitro. *J. Virol.* 76, 841–850 (2002).

46. Ackermann, M., Sarmiento, M. & Roizman, B. Application of antibody to synthetic peptides for characterization of the intact and truncated alpha 22 protein specified by herpes simplex virus 1 and the B23 alpha 22 deletion mutant. *J. Virol.* 56, 207–215 (1985).

47. Preston, C. M. Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant tsK. *J. Virol.* 32, 357–369 (1979).

48. Sekulovich, R. E., Leary, K. & Sandri-Goldin, R. M. The herpes simplex virus type 1 alpha protein ICP27 can act as a trans-repressor or a trans-activator in combination with ICP4 and ICP0. *J. Virol.* 62, 4510–4522 (1988).

49. Dobin, A. et al. STAR: ultra-fast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).

50. Guo, Y., Mahony, S. & Gifford, D. K. High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. *PLoS Comput. Biol.* 8, e1002638 (2012).

51. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 578–589 (2010).

52. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842 (2010).

53. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930 (2014).

54. Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* 27, 1653–1659 (2011).

Acknowledgements
We would like to thank David M. Knipe for providing reagents, Neal deLuca for providing his extended set of HSV-1 deletion mutants and UCI GHTF for sequencing. This study was supported by the following grants: NIH GM090056 and GM138441 to Y.S. ERC Consolidator award (ERC-2016-CoG 721016-HERPES) and DFG grant DFG (1275/6-1 to L.D.). A.W.W. was the recipient of a generous grant from the Alexander von Humboldt Foundation.

Author contributions
X.W., R.M.S.-G., L.D. and Y.S. conceived and designed the experiments. X.W., T.H., A.W.W., W.H., and L.E. performed the majority of the experiments. F.E., C.C.F., E.F., L.E., B.K.P., and Y.C. contributed to data analyses. X.W., L.D., and Y.S. wrote the paper with input from all authors.

Competing interests
The authors declare no competing interests.
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-14109-x.

Correspondence and requests for materials should be addressed to R.M.S.-G., L.D. or Y.S.

Peer review information Nature Communications thanks William Marzluff and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020