Conserved Outer Tegument Component UL11 from Herpes Simplex Virus 1 Is an Intrinsically Disordered, RNA-Binding Protein

Claire M. Metrick,a,b* Andrea L. Koenigsberg,a,c Ekaterina E. Heldweina

aDepartment of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts, USA
bGraduate Program in Biochemistry, Tufts School of Graduate Biomedical Sciences, Tufts University School of Medicine, Boston, Massachusetts, USA
cGraduate Program in Molecular Microbiology, Tufts School of Graduate Biomedical Sciences, Tufts University School of Medicine, Boston, Massachusetts, USA

ABSTRACT A distinguishing morphological feature of all herpesviruses is the multi-protein tegument layer located between the nucleocapsid and lipid envelope of the virion. Tegument proteins play multiple roles in viral replication, including viral assembly, but we do not yet understand their individual functions or how the tegument is assembled and organized. UL11, the smallest tegument protein, is important for several distinct processes in replication, including efficient virion morphogenesis and cell-cell spread. However, the mechanistic understanding of its role in these and other processes is limited in part by the scant knowledge of its biochemical and structural properties. Here, we report that UL11 from herpes simplex virus 1 (HSV-1) is an intrinsically disordered, conformationally dynamic protein that undergoes liquid-liquid phase separation (LLPS) in vitro. Intrinsic disorder may underlie the ability of UL11 to exert multiple functions and bind multiple partners. Sequence analysis suggests that not only all UL11 homologs but also all HSV-1 tegument proteins contain intrinsically disordered regions of different lengths. The presence of intrinsic disorder, and potentially, the ability to form LLPS, may thus be a common feature of the tegument proteins. We hypothesize that tegument assembly may involve the formation of a biomolecular condensate, driven by the heterogeneous mixture of intrinsically disordered tegument proteins.

IMPORTANCE Herpesvirus virions contain a unique tegument layer sandwiched between the capsid and lipid envelope and composed of multiple copies of about two dozen viral proteins. However, little is known about the structure of the tegument or how it is assembled. Here, we show that a conserved tegument protein UL11 from herpes simplex virus 1, a prototypical alphaherpesvirus, is an intrinsically disordered protein that undergoes liquid-liquid phase separation in vitro. Through sequence analysis, we find intrinsically disordered regions of different lengths in all HSV-1 tegument proteins. We hypothesize that intrinsic disorder is a common characteristic of tegument proteins and propose a new model of tegument as a biomolecular condensate.

KEYWORDS herpesvirus, tegument, viral protein, viral assembly, intrinsically disordered protein (IDP), RNA-binding protein, small-angle X-ray scattering (SAXS), conformational flexibility, structural model, liquid-liquid phase separation (LLPS), biomolecular condensate

Herpesviruses are large enveloped viruses that infect nearly all species of vertebrates and some invertebrates. In humans, these viruses cause lifelong infections that are generally asymptomatic or cause mild symptoms, such as orofacial lesions (cold sores) and genital lesions, but can also cause severe disease in certain subpopulations.


Editor Thomas Shenk, Princeton University
Copyright © 2020 Metrick et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.
Address correspondence to Ekaterina E. Heldwein, katya.heldwein@tufts.edu.
* Present address: Claire M. Metrick, Biogen, Cambridge, Massachusetts, USA.
This article is a direct contribution from Ekaterina E. Heldwein, a Fellow of the American Academy of Microbiology, who arranged for and secured reviews by Russell Diefenbach, Macquarie University, and Stephen Graham, University of Cambridge.
Received 1 April 2020
Accepted 7 April 2020
Published 5 May 2020
Nine human herpesviruses belong to three subfamilies, alpha-, beta-, and gammaherpesviruses, that differ in tropism and pathogenesis yet share many conserved replication mechanisms, including viral assembly and morphogenesis.

Herpesvirus virions are composed of an icosahedral capsid encasing the double-stranded DNA (dsDNA) genome and are coated with lipid envelopes containing viral glycoproteins. Between the capsid and the envelope lies the tegument layer that is unique to herpesviruses and is a dense network of dozens of viral proteins that is structurally maintained even following removal of the viral envelope (1). In the prototype alphaherpesvirus herpes simplex virus 1 (HSV-1), the tegument contains multiple copies of 24 viral proteins and several host proteins (2). Tegument proteins are assigned to either the inner layer or the outer layer depending on their association with the capsid or viral envelope, respectively (3). While the inner tegument layer is anchored on the nucleocapsid, and the viral proteins in this layer are present in stoichiometric quantities (4), the outer tegument layer is more heterogeneous and appears amorphous (5). The assembly of the tegument onto the capsids occurs largely in the cytoplasm during the process of secondary envelopment wherein capsids gain their external envelope by budding into host vesicles derived from the Golgi/endosomal network (6, 7). Many tegument proteins are found associated with these vesicles in infected cells and are thought to regulate the secondary envelopment process by a mechanism yet to be identified (reviewed in reference 3). In addition to their importance in virion structure and morphogenesis, many tegument proteins perform other functions. For example, inner tegument protein UL37 is thought to bridge the viral capsid and envelope by binding capsid-associated UL36 and membrane-anchored glycoprotein K (gK) and UL20 (8). UL37 may also target capsids to the correct site of secondary envelopment, based on the structural similarities between CATCHR proteins and UL37 from HSV-1 and pseudorabies virus (PRV) (9, 10). Different tegument proteins with known structural roles also have regulatory functions, including modulation of host intrinsic and innate immune responses, upregulation of viral gene expression, downregulation of host gene expression, and capsid trafficking through the cytoplasm (reviewed in references 3, 11, and 12).

The UL11 gene is conserved across all herpesvirus subfamilies (Fig. 1A) and encodes a small, membrane-associated outer tegument protein that localizes primarily to the Golgi (13). UL11 homologs from HSV-1, herpes simplex virus 2 (HSV-2), Epstein-Barr virus (EBV), cytomegalovirus (CMV), and likely all others are N-terminally myristoylated (13–21). UL11 homologs from HSV-1, EBV, and likely, HSV-2 are also palmitoylated (17, 22–24) (Fig. 1A). These two modifications enable UL11 to bind to cytoplasmic membranes (13, 14, 22) and associate with detergent-resistant membrane microdomains, or lipid rafts (23, 24). Palmitoylation is also required for both Golgi targeting and strong membrane interactions (22). Surprisingly, myristoylation and palmitoylation are not essential for replication because UL11 mutants lacking these modifications can partially rescue replication-deficient UL11-null HSV-1 (25).

In many herpesviruses, UL11 homologs are important for efficient viral replication, and their deletion results in reductions in viral titer that range in magnitude depending on the virus (15, 25–29). Deletion of UL11 homologs causes accumulation of nonenveloped capsids in the cytoplasm (29–32). Thus, UL11 is thought to participate in secondary capsid envelopment through an unclear mechanism (3). UL11 binds capsid-associated tegument protein UL16 (33–36) using a cluster of acidic residues within its N terminus (19, 27, 37, 38) (Fig. 1A). This conserved interaction may help bridge the capsid and envelope during secondary envelopment (3). In addition to its role in secondary envelopment, UL11 appears to have other functions. In alphaherpesviruses, UL11 forms a ternary complex with UL16 and UL21 (39) that binds the cytoplasmic tail of glycoprotein E (gE) (40) (Fig. 1A) and modulates its function in cell-cell fusion (39). Finally, UL11 localizes to the nucleus (41) for reasons that are unclear.

Despite the importance of UL11 in multiple viral processes, sparse biochemical and structural information limits both our mechanistic understanding of its roles in viral replication and tegument assembly. Here, we employed circular dichroism (CD), limited
proteolysis, small-angle X-ray scattering (SAXS), and light microscopy to characterize the structure and solution properties of HSV-1 UL11. We report that UL11 is an intrinsically disordered protein (IDP) that undergoes liquid-liquid phase separation (LLPS) in vitro. We propose that intrinsic disorder underlies the ability of UL11 to exert multiple functions and bind multiple partners, which include UL16, gE, and potentially, RNA. Sequence analysis highlights that all UL11 homologs and, more broadly, all HSV-1 tegument proteins contain intrinsically disordered regions of different lengths, suggesting that the presence of disorder is a common feature of the tegument. Based on the ability of heterogeneous mixtures of intrinsically disordered proteins to assemble into biological condensates in vivo, we hypothesize that formation of similar assemblies by outer tegument proteins at the sites of secondary envelopment may contribute to viral morphogenesis.

RESULTS

Expression and purification of UL11. To characterize the structure and biochemical properties of HSV-1 UL11, we expressed several constructs in Escherichia coli.
UL11 containing N-terminal His6 tag (H6-UL11) was purified in soluble form (Fig. 2C), but the yield was low. To increase expression and to allow removal of the affinity tag, an N-terminal glutathione S-transferase (GST) tag followed by an HRV3C (PreScission) protease cleavage site was added, generating the GST-UL11 construct. Cleavage with PreScission protease yielded UL11 with an N-terminal GPLGS linker sequence (Fig. 2A). The GST-UL11 construct had higher protein yields but could not be purified to homogeneity (Fig. 2C). We then sought an alternative affinity tag that would not attract nonspecific contaminants. Therefore, a small, C-terminal Strep-tag II (StII) was introduced to improve both the yield and purity (UL11-StII) (Fig. 2A). We placed the new tag at the C terminus, away from the predicted functional sites (Fig. 1A) or any predicted structure at the N terminus (Fig. 1B). This construct had a higher yield and could be purified to homogeneity (Fig. 2C). UL11 lacks tryptophans and has only one tyrosine, which complicates concentration estimates based on absorption at 280 nm. Addition of the StII tag (WSHPQ FKE), which contains a tryptophan, also improved spectrophotometric tracking. Since all constructs displayed similar characteristics once purified, the UL11-StII construct was used in the majority of subsequent studies.
UL11 copurifies with E. coli RNA. During initial purification of UL11-StII, the eluate from streptactin affinity resin appeared pure by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). However, the $A_{260}/A_{280}$ ratio of 1.9 indicated the presence of approximately 60% nucleic acids (NAs) (95) (Fig. 3A). To identify which NAs copurified with UL11, we used a nuclease digestion assay. After streptactin affinity purification, bound NAs were isolated from UL11/NA complex samples using phenol-chloroform extraction and incubated with DNase or RNase or left untreated. UL11-bound NAs were sensitive to RNase but not DNase treatment (Fig. 3B). Furthermore, the size and banding pattern of the UL11-bound NAs were consistent with rRNA (96). Whether UL11 binds rRNA with some degree of specificity or copurifies with rRNA because these are the major RNA species in E. coli remains unclear. In any case, our results suggest that UL11 binds not only UL16 and gE but also some RNA species. The addition of a heparin affinity step (Fig. 3A) to the UL11-StII purification scheme (Fig. 2B) to separate UL11 from copurifying NAs reduced the $A_{260}/A_{280}$ ratio to 0.8, which corresponds to 2% NA content (Fig. 3A). All further experiments with UL11-StII in this study were performed with samples free of bound NAs (“NA-free”).

UL11 displays hallmarks of an intrinsically disordered protein. Despite the calculated molecular masses ranging from 10.9 to 12.7 kDa (Fig. 2A), UL11 constructs consistently migrated as 15- to 20-kDa species on SDS-polyacrylamide gels (SDS-PAGs) (Fig. 2C). Slow electrophoretic mobility is typically observed for proteins that are posttranslationally modified (44), but E. coli-expressed UL11 is unlipidated because prokaryotes lack endogenous N-myristoyltransferase activity (45). Moreover, matrix-assisted laser desorption ionization—time of flight mass spectrometry (MALDI-TOF MS) analysis of UL11-StII indicated the molecular mass of 11.4 kDa (Fig. 2A and 4A). Abnormal electrophoretic mobility, manifesting as a 1.2- to 1.8-times-higher apparent molecular mass than expected from the sequence, is also observed for intrinsically disordered proteins (IDPs) (46, 47; reviewed in references 48 and 49). IDPs are thought
to bind less SDS than globular proteins due to a distinct sequence composition, e.g., high net charge and low hydrophobicity, which would cause them to migrate more slowly than expected (reviewed in reference 48). Higher proline content, typical of IDPs, increases protein rigidity, which further reduces electrophoretic mobility (reviewed in reference 49).

During purification, all three UL11 constructs eluted from size exclusion chromatography (SEC) at a volume consistent with that of an ~30-kDa globular protein, much larger than the molecular mass predicted from the UL11 sequence (Fig. 4B). This observation held true not only for H6-UL11 and GPLGS-UL11, which potentially contained copurifying NAs, but also for NA-free UL11-StII (Fig. 4B). IDPs typically have reduced SEC mobility due to reduced protein compaction (reviewed in reference 49). However, SEC cannot distinguish an elongated monomer from an oligomer. To probe the oligomeric state, NA-free UL11-StII was cross-linked with bis(sulfosuccinimidyl)suberate (BS3), a homobifunctional cross-linking reagent that reacts with primary amines. In the presence of increasing amounts of cross-linker, no bands of higher molecular mass were detected on SDS-PAGs (Fig. 4C). Instead, cross-linked UL11-StII migrated faster (Fig. 4C) probably due to the presence of intramolecular cross-links that prevented complete unfolding under SDS-PAGE conditions. Thus, UL11 is monomeric in solution under these experimental conditions.

Sequence analysis of HSV-1 UL11 identified several features characteristic of disorder-containing proteins (reviewed in reference 49). First, the amino acid sequence is enriched in disorder-promoting residues P, G, E, A, S, R, K, and Q (50) (56% total...
residues) and depleted of order-promoting residues W, Y, F, C, I, L, and N (19% total residues) (Fig. 1A). The low content of aromatic residues in particular results in a very low extinction coefficient at 280 nm (51), 1,490 M$^{-1}$ cm$^{-1}$ for HSV-1 UL11. Second, HSV-1 UL11 has little predicted secondary structure (~15% $\beta$ strand and no $\alpha$ helices) (Fig. 1A); a predicted disordered C terminus, residues 61 to 96, as estimated by FoldUnfold server (Fig. 1B); and no predicted globular domains, according to the IUPred2A server. Finally, it has a hypervariable sequence such that in the Pfam database (42), the UL11 family is represented only by residues 1 to 39, and even these do not align well across UL11 homologs (Fig. 1A). Collectively, these features, along with aberrant electrophoretic and chromatographic mobility, suggested the presence of structural disorder in UL11.

**UL11 undergoes LLPS in vitro.** During affinity purification of H$_6$-UL11 at 4°C, the elution fraction appeared cloudy, yet no pellet formed after centrifugation. Furthermore, after SEC at 4°C, the most concentrated fractions of the UL11-containing peak also appeared cloudy. Under the microscope, such cloudy samples contained droplets reminiscent of liquid-liquid phase separation (LLPS) (Fig. 5A). This phenomenon was observed not only in the purification buffer but also with a wide range of additives, including ~1/3 of 752 sitting-drop crystal screening conditions (Fig. 5A). Similar droplet formation was observed with GPLGS-UL11 and UL11-StI samples (Fig. 5A).

A common property of LLPS is its temperature-dependent reversibility (43, 52). Purified NA-free UL11-StI samples appeared turbid when cooled down to 4°C but became clear when warmed up by hand. Subsequent cooling to 4°C restored turbidity (Fig. 5B). Observing these warmed and cooled samples of UL11-StI under a light microscope confirmed that turbid samples had droplets whereas clear samples did not. Furthermore, these droplets displayed the characteristic liquid-like properties of merging and surface wetting (Fig. 5C). The ability of UL11 to undergo LLPS was also dependent on protein concentration, salt concentration, and the presence of molecular crowders polyethylene glycol 350 monomethyl ether (PEG 350 MME) and bovine serum albumin (BSA) (Fig. 5D). Notably, the addition of molecular crowders to the most concentrated UL11 samples at room temperature immediately prompted LLPS formation, which manifested as turbidity. We hypothesize that LLPS is an intrinsic property of UL11 because all constructs manifested LLPS regardless of the presence of copurifying NAs, the identity of the affinity tag, or its placement.

**UL11 is mostly disordered in solution.** HSV-1 UL11 has little predicted secondary structure, ~15% $\beta$ strand and no $\alpha$ helices (Fig. 1A). To assess the secondary structure of NA-free UL11-StI, we used far-UV circular dichroism (CD) at three concentrations (Fig. 6A). The CD spectrum of UL11-StI at the lowest concentration resembled the spectrum of a random-coil polypeptide, with a large negative molar ellipticity near 200 nm and a negative shoulder in the 220- to 240-nm region. Such spectra are typical of IDPs (reviewed in reference 49). Thus, UL11 is largely disordered in solution. Unexpectedly, with increasing concentration, the absolute magnitude of the CD signal decreased (Fig. 6A), which is often associated with protein aggregation, although no precipitation was observed in the samples. Instead, increasing LLPS was seen with increasing concentration of UL11 (Fig. 6A), and this phenomenon occurred at a lower concentration of UL11 in CD buffer relative to purification buffer. While droplet formation was not obvious under the light microscope at the UL11 concentrations used in the CD experiment, the droplets could have been too small to detect by eye. Thus, LLPS could, in principle, account for the decrease in absolute magnitude of the CD signal, as has been reported for phase-separating proline-arginine-rich peptides (52). Although the CD spectra of UL11-StI had random-coil characteristics at higher protein concentrations, the large negative peak shifted toward a higher wavelength (Fig. 6A). The CDSSTR algorithm (53) as implemented in Dichroweb estimated that the proportion of the regularly structured elements (including regular $\alpha$ helices and regular $\beta$ strands) increased from 13% to 23% with increasing protein concentration (Table 1).
This suggested that some regular secondary structural elements within UL11 may form as the concentration increases.

To further probe the disordered nature of UL11, we subjected NA-free UL11-StII to limited proteolysis with a panel of proteases either preferentially targeting specific residues—basic (trypsin), acidic (V8), or bulky hydrophobic (chymotrypsin) residues—or having low specificity (proteinase K). Only digestion with V8 resulted in a proteolytically resistant fragment (Fig. 6B and C). This fragment was identified by N-terminal sequencing and MALDI-TOF MS (see Fig. S1 in the supplemental material) and reflects the locations of the cleavage sites. Such high proteolytic susceptibility suggests a lack of structured regions to protect against cleavage, i.e., the presence of intrinsic disorder.

FIG 5 Ul11 undergoes liquid-liquid phase separation. (A) Microscopic images of Ul11 constructs in solution show representative images of liquid-liquid phase separation (LLPS) seen under a variety of conditions, including purification buffer and for H4-Ul11, crystallization screens (1:1 mix of protein-solution in vapor diffusion chamber). (B) Macroscopic images of Ul11-StII in solution in purification buffer show reversibility of Ul11-StII phase separation (cloudiness) upon temperature cycling. (C) Time lapse microscopic images of Ul11-StII condensates in purification buffer display hallmark liquid-like properties of surface wetting and droplet merging (selected example in orange box). (D) LLPS (separation, orange circles; clear solution, blue triangles) occurs with increasing protein concentration, increasing salt concentration, and increasing molecular crowding.
UL11 is an extended, conformationally dynamic molecule. To characterize the conformational state and flexibility of UL11 in solution, we performed small-angle X-ray scattering coupled with size exclusion chromatography (SEC-SAXS). SEC-SAXS profiles were collected for NA-free UL11-StII (Fig. 7A), and SAXS data were processed by averaging frames with a consistent radius of gyration ($R_g$) value and subtracting averaged buffer frames to generate subtracted curves (Fig. 7B). These data were selected only from frames corresponding to the SEC peak to limit analysis to homo-

**FIG 6** UL11 lacks defined secondary structure and is sensitive to proteolysis. (A) CD spectra of UL11-StII at several concentrations show primarily random-coil characteristics. UL11-StII undergoes LLPS (LLPS, orange circles; clear solution, blue triangles) with increasing concentration in CD buffer at ambient temperature. See also Table 1. (B) Limited proteolysis of UL11-StII using trypsin, chymotrypsin, V8 protease, or proteinase K with increasing amounts of protease. (C) Sequence of UL11-StII with potential proteolysis sites colored (green, chymotrypsin; pink, trypsin; purple, V8). V8 fragment identified by MALDI-TOF mass spectrometry and N-terminal sequencing (see Fig. S1 in the supplemental material) is underlined in purple. The StII sequence is double underlined.

<table>
<thead>
<tr>
<th>UL11 concn (mg/ml)</th>
<th>Reference set or parameter</th>
<th>NRMSD$^b$</th>
<th>% of residues$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Helix1</td>
<td>Helix2</td>
</tr>
<tr>
<td>0.46</td>
<td>SP175</td>
<td>0.043</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Set 6</td>
<td>0.056</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Set 7</td>
<td>0.041</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>0.23</td>
<td>SP175</td>
<td>0.034</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Set 6</td>
<td>0.023</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Set 7</td>
<td>0.024</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>0.13</td>
<td>SP175</td>
<td>0.019</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Set 6</td>
<td>0.022</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Set 7</td>
<td>0.023</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td></td>
<td>0.02</td>
</tr>
</tbody>
</table>

$^a$The reference set is the set of curves used in CDSSTR (53) to estimate secondary structure content from experimental curves.

$^b$NRMSD, normalized root mean square deviation between calculated and experimental CD spectra.

$^c$The percentage of residues in a regular helix (93) (Helix1), distorted helix (Helix2), regular strand (Strand1), distorted strand (Strand2), and beta turns (Turns) and the percentage of residues lacking secondary structure (Helix1, Helix2, Strand1, Strand2, and Turns) (Unordered).
The molecular mass for UL11-StII calculated by the volume of correlation method (54), 15.8 kDa, fell between the expected values for a monomer and a dimer (Fig. 2A and Table 2). Although no oligomeric species were observed in the cross-linking experiments (Fig. 4C), these experiments were done at a lower protein concentration than that used in the SEC-SAXS experiment. Unfortunately, concentration-independent methods currently in use cannot accurately estimate the molecular mass of a disordered protein from SEC-SAXS curves. A Guinier plot was assembled from the subtracted SAXS curve (Fig. 7B), and the linear region was used to calculate the $R_g$ of the sample (Table 2). Linearity in the Guinier region ($qR_g < 1.3$) additionally indicates a lack of intermolecular attraction (aggregation) or repulsion that would interfere with parameter calculation. We hypothesize that UL11-StII is monomeric under the conditions of SEC-SAXS experiment but cannot rule out the presence of higher-molecular mass species.

To calculate more accurate values of $R_g$ and the maximum dimension ($D_{\text{max}}$) of UL11-StII, a pair distance distribution function $P(r)$ was generated from the subtracted curve (Table 2). The $P(r)$ derived from UL11-StII SAXS data was asymmetrical and trailed.
This profile was observed regardless of whether the function was preemptively forced to zero or not, and is characteristic of an elongated and/or flexible molecule (Fig. 7C) (55, 56). Thus, assuming the sample was fully monomeric, UL11 adopts an extended conformation, with $R_g = 24.7$ Å and $D_{max} = 76$ Å.

To estimate conformational flexibility, a dimensionless Kratky plot was calculated (56) (Fig. 7D). In contrast to the expected parabolic shape associated with a globular, ordered protein (54), the UL11-StII Kratky plot remained elevated in value after the $qR_g$ inflection point (Fig. 7D), which is characteristic of IDPs (56, 57).

Modeling UL11 as a conformational ensemble. Due to disorder, IDPs sample multiple conformations in solution and cannot be modeled as single conformers using ab initio bead modeling. Indeed, AMBIMETER reported hundreds of compatible shape categories for UL11-StII (Table 2). Instead, it is more appropriate to model IDPs as conformational ensembles—groupings of multiple static structures that represent possible conformations (reviewed in reference 49). Conformational ensembles do not provide a unique description of the conformer population because any given IDP likely exists as a very large number of conformers. Instead, they provide possible, nonunique solutions consistent with the experimental data that are useful for representing the conformational diversity of an IDP under study.

To generate conformational ensembles for UL11, we used ensemble optimization modeling (EOM), a method commonly used to model SAXS data for conformationally dynamic systems (58, 59). EOM quantifies flexibility of the system using the $R_{flex}$ parameter, where $R_{flex}$ is 0% for a fully rigid system and 100% for a fully flexible system (59). The $R_{flex}$ of a random pool representing 10,000 possible conformations is usually ~85 to 90%. For a rigid system, e.g., a macromolecule that exists as a small number of conformers, $R_{max}$ would be expected to decrease significantly after ensemble optimization modeling. However, for UL11-StII, there was no appreciable decrease in the $R_{flex}$

### TABLE 2 SAXS structural parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value(s) for UL11-StII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guinier analysis</td>
<td></td>
</tr>
<tr>
<td>$I(0)$ (cm$^{-1}$)$^{a}$</td>
<td>0.00477 ± 0.000041</td>
</tr>
<tr>
<td>$R_g$ (Å)$^{b}$</td>
<td>24.29 ± 2.0</td>
</tr>
<tr>
<td>$q_{min}$ (Å$^{-1})^{c}$</td>
<td>0.01254</td>
</tr>
<tr>
<td>$qR_g$ max$^{d}$</td>
<td>1.30239</td>
</tr>
<tr>
<td>Coefficient of correlation, $R^2$</td>
<td>0.938</td>
</tr>
<tr>
<td>$P(r)$ analysis</td>
<td></td>
</tr>
<tr>
<td>(forced to 0/not forced)$^{f}$</td>
<td></td>
</tr>
<tr>
<td>$I(0)$ (cm$^{-1}$)</td>
<td>0.004748 ± 0.00003566/0.004866 ± 0.00009504</td>
</tr>
<tr>
<td>$R_g$ (Å)</td>
<td>24.74 ± 0.1599/26.29 ± 1.712</td>
</tr>
<tr>
<td>$D_{max}$ (Å)$^{g}$</td>
<td>76/120</td>
</tr>
<tr>
<td>$q$ range (Å$^{-1})^{h}$</td>
<td>0.0074 – 0.2802/0.0074 – 0.2802</td>
</tr>
<tr>
<td>$\chi^2$ (total estimate from GNOM)$^{i}$</td>
<td>1.0385 (0.8640)/0.9935 (0.7203)</td>
</tr>
<tr>
<td>AMBIMETER analysis of $P(r)$</td>
<td></td>
</tr>
<tr>
<td>(forced to 0/not forced)$^{f}$</td>
<td></td>
</tr>
<tr>
<td>Compatible shape categories$^{j}$</td>
<td>720/153</td>
</tr>
<tr>
<td>Ambiguity score$^{k}$</td>
<td>2.857/2.185</td>
</tr>
<tr>
<td>Uniqueness$^{l}$</td>
<td>Highly ambiguous/might be ambiguous</td>
</tr>
</tbody>
</table>

$^{a}$I(0), scattering intensity at zero angle.
$^{b}R_g$, radius of gyration.
$^{c}q_{min}$, lowest $q$ value used in the Guinier analysis.
$^{d}qR_g$ max, maximum $q$ value used in the Guinier analysis × $R_g$.
$^{e}R^2$, linear fit of the Guinier plot.
$^{f}$(forced to 0/not forced) = $P(r)$ forced to 0 at $D_{max}$ in DATGNOM (91)/$P(r)$ not forced to 0.
$^{g}D_{max}$, maximum dimension calculated from $P(r)$.
$^{h}q$ range, $q$ values used in $P(r)$ plot.
$^{i}\chi^2$ fit of $P(r)$ to data. The total quality estimate from GNOM (91) is shown in the parentheses.
$^{j}$The number of shape topologies, $M$, compatible with $P(r)$ evaluated by AMBIMETER (94).
$^{k}$The ambiguity score is log $M$ and correlates with probability of finding a false-positive result in a three-dimensional (3D) reconstruction (94).
$^{l}$Uniqueness of 3D reconstruction of $P(r)$ evaluated by AMBIMETER (94).
between the random pool and optimized ensembles of models representing the SAXS data (Table 3), which indicates that UL11 is highly flexible and exists as multiple conformers of different shape.

EOM generates probability density distributions, which represent the probability that the protein adopts a certain conformation, and the relative ratios of those conformations. Three independently optimized conformational ensembles representing UL11-StII identified three populations in the histogram: compact (~50%), intermediate (~40%), and extended (~10%) (Fig. 8 and Table 3), which define three major conformational ensembles of UL11-StII. Overall, SAXS measurements suggest that UL11 samples multiple conformations in solution.

**DISCUSSION**

In addition to being structural components of the herpesviral virions, tegument proteins perform numerous functions during herpesvirus replication: they regulate viral morphogenesis, mediate intracellular capsid trafficking, control gene expression, and suppress the host antiviral response, to mention a few. Our limited mechanistic understanding of their roles in viral replication would greatly benefit from knowing the structures and the biochemical properties of tegument proteins. However, structures of only 6 out of 24 HSV-1 tegument proteins are available (10, 60–65), and structural information on tegument proteins in other herpesviruses is even scarcer. Expanding our structural knowledge of the tegument would inform not only mechanistic studies of individual tegument proteins but also models of how they are arranged within the tegument layer.

Here, we investigated the biochemical properties, solution behavior, and structural characteristics of HSV-1 UL11, the smallest conserved tegument protein that is impor-

**TABLE 3** Dimensional and flexibility characteristics of optimized ensembles

<table>
<thead>
<tr>
<th>Pool or ensemble</th>
<th>Parameter or conformational state</th>
<th>( R_g ) (Å)</th>
<th>( D_{\text{max}} ) (Å)</th>
<th>Proportion (%)</th>
<th>( R_{\text{max}} ) (%)</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random pool</td>
<td>Avg</td>
<td>27.16</td>
<td>82.86</td>
<td>100</td>
<td>86.57</td>
<td></td>
</tr>
<tr>
<td>Optimized ensemble 1</td>
<td>Compact</td>
<td>20.75</td>
<td>59.61</td>
<td>44.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>27.35</td>
<td>81.195</td>
<td>44.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extended</td>
<td>35.95</td>
<td>111.43</td>
<td>11.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>25.48</td>
<td>74.24</td>
<td>100</td>
<td>83.96</td>
<td>0.983</td>
</tr>
<tr>
<td>Optimized ensemble 2</td>
<td>Compact</td>
<td>21.38</td>
<td>61.47</td>
<td>54.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>28.59</td>
<td>79.97</td>
<td>36.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extended</td>
<td>37.55</td>
<td>109.55</td>
<td>9.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>25.44</td>
<td>72.2</td>
<td>100</td>
<td>84.97</td>
<td>0.983</td>
</tr>
<tr>
<td>Optimized ensemble 3</td>
<td>Compact</td>
<td>21.94</td>
<td>63.63</td>
<td>55.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>27.78</td>
<td>81.53</td>
<td>33.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extended</td>
<td>35.79</td>
<td>107.3</td>
<td>11.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg</td>
<td>25.55</td>
<td>75.43</td>
<td>100</td>
<td>84.30</td>
<td>0.983</td>
</tr>
<tr>
<td>Avg ensemble</td>
<td>Compact</td>
<td>21.36</td>
<td>61.57</td>
<td>51.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>27.91</td>
<td>80.90</td>
<td>38.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extended</td>
<td>36.43</td>
<td>109.43</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

\( R_g \) is radius of gyration of structural models.

\( D_{\text{max}} \) is maximum dimension of structural models.

Proportion is the percentage of curves representing each category within the pool or ensemble.

\( R_{\text{max}} \) is a measurement of information entropy, correlated with conformational flexibility, in the pool or ensemble. \( R_{\text{max}} \) would be 100% for a totally flexible system versus 0% for an ideal rigid system.

\( \chi^2 \) is a measurement of fit between the SAXS curve generated from the optimized ensemble of models and experimental data.

The random pool is 10,000 structural models representing UL11-StII generated by RANCH within EOM 2.0 (58, 59).

Optimized ensembles 1, 2, and 3 are from the random pool that together represent the experimental data as chosen by GAJOE within EOM 2.0 (58, 59).

Compact, category of models with \( R_g < 24 \) and \( D_{\text{max}} < 71 \).

Intermediate, category of models with \( 26 < R_g < 31 \) and \( 77 < D_{\text{max}} < 83 \).

Extended, category of models with \( R_g > 35 \) and \( D_{\text{max}} > 107 \).
tant for efficient virion morphogenesis and cell-cell spread. We show that recombinant, unlipidated HSV-1 UL11 is a conformationally dynamic, intrinsically disordered protein (IDP). Several characteristics of disorder-containing proteins were evident from sequence analysis: high content of disorder-promoting residues along with low content of order-promoting residues, a hypervariable sequence across homologs, little predicted secondary structure, and no predicted globular domains. Experimentally, the intrinsically disordered nature of UL11 was evident from the predominantly random-coil character of its CD spectrum, high proteolytic sensitivity, reduced electrophoretic and chromatographic mobility, and IDP signatures within the SAXS spectra (reviewed in references 56 and 57).

All experiments reported here used E. coli-expressed unlipidated UL11, whereas in HSV-1-infected cells or in uninfected cells overexpressing UL11, UL11 is both myristoylated and palmitoylated (13–15, 22, 23). Nevertheless, we propose that the IDP characteristics of UL11 observed in vitro reflect its behavior in vivo. For example, UL11 homologs isolated from transfected and infected cells migrate with a higher apparent molecular mass on SDS-PAGs (14, 27, 66) similarly to E. coli-expressed unlipidated UL11.
While slower electrophoretic mobility was previously attributed to posttranslational modifications (14, 27, 66), one report showed that unlipidated HSV-1 UL11 migrated just as slowly as the lipidated UL11 (22). Therefore, we propose that the slow electrophoretic mobility of UL11 in vivo is instead due to the intrinsic disorder. Another in vivo phenotype that could be attributed to the intrinsic disorder is that in the absence of lipid modifications or the binding partner UL16, UL11 and its homologs are relatively unstable in cells, presumably due to rapid turnover (17, 25, 67, 68). Such behavior is typical of IDPs because they are more susceptible to proteolysis than globular proteins. Association with a binding partner such as UL16 or a membrane could potentially protect UL11 from proteolysis and increase its lifetime.

UL11 homologs have hypervariable sequences even within the N-terminal UL11 Pfam motif (Fig. 1A). Nonetheless, UL11 homologs share conserved functional motifs within their N-terminal portions, such as lipidation sites and acidic clusters. The UL11/UL16 interaction has been demonstrated for their homologs from HSV-1 (19), PRV (19), CMV (68), and Kaposi’s sarcoma herpesvirus (28), and thus, the UL16-binding site located within the UL11 N terminus (37) is likely conserved. Even the more divergent C-terminal portions share certain sequence characteristics (overabundance of disorder-promoting residues and basic residues) and may thus share conserved functions, e.g., the ability to bind gE (67) or negatively charged binding partners such as nucleic acids (this paper). These sequence characteristics are not limited to UL11 homologs from human herpesviruses but extend to animal herpesviruses from all three subfamilies (Fig. 1A).

We found that E. coli-produced UL11 bound endogenous RNA, potentially through interactions with its basic C terminus. Overabundance of basic residues is frequently found in low-complexity sequences where they are thought to promote disorder (48), and as a result, many proteins containing basic intrinsically disordered regions (IDRs) bind NAs either specifically or promiscuously (69). The banding pattern of the observed RNA was consistent with rRNA (96). rRNA makes up a significant portion of the total E. coli RNA pool (70), so a UL11/rRNA interaction could be nonspecific. However, another HSV-1 tegument protein, UL21, which also copurifies with E. coli RNA, does not show any preference for binding rRNA (61). UL11 could, thus, bind rRNA with some degree of specificity. Many other tegument proteins bind RNA (61, 71), so future studies should explore RNA as a potential UL11-binding partner and the effect of RNA binding by UL11 on UL11/gE interaction (40).

Ensemble modeling of SAXS data suggests that despite its highly dynamic nature, UL11 samples three distinct conformational states: extended, intermediate, and compact. Transient compaction of UL11 between these distinct states or simply within its flexible landscape of conformations could be driven by intramolecular interactions, for example, electrostatic interactions between the N-terminal acidic cluster and the basic residues within the C terminus. Although we did not see any evidence of oligomerization at lower concentrations, we could not rule out the presence of higher-molecular-mass UL11 species under the conditions of the SEC-SAXS experiment. Therefore, one or more of the observed conformational states could correspond to a UL11 oligomer. Regardless, the ability to adopt multiple conformations could allow UL11 to take on many functions, with each conformation mediating distinct functions and interactions, such as binding to protein partner UL16 or gE, which, in turn, may stabilize particular conformations. Binding different RNA species could also induce or stabilize a particular structure of UL11, as has been observed with HIV Tat (reviewed in reference 69), increasing the conformational and thus functional landscape available to UL11. Finally, a conformational shift from unstructured to more structured could occur at higher UL11 concentrations, e.g., in our CD (Fig. 6A and Table 1) and SAXS (Fig. 7 and 8) experiments or in lipid rafts (24).

Many IDPs and IDR-containing proteins can undergo phase separation not only in vitro (72) but also in vivo, forming biomolecular condensates also known as proteinaceous membrane-less organelles (PMLOs) (73). For example, FUS, an RNA-binding protein associated with several neurodegenerative diseases, forms “spherical droplets”
in the neuronal nuclei, while its recombinant, purified form forms similar phase-separated “liquid droplets” *in vitro* (74). Biomolecular condensates are heterogeneous mixtures enriched in IDPs and IDR-containing proteins that engage in many multivalent specific and nonspecific interactions to ensure that condensates remain stable even when individual components are lost (72). UL11 readily undergoes phase separation *in vitro* (this work) and may do so *in vivo*, judging by reports that it localizes to punctate cytosolic structures (19, 38) and nuclear speckles (41). Phase separation is influenced by many factors, including protein concentration, salt concentration, binding partners, and other solution components (43, 52; reviewed in references 75 and 76). Although *in vitro*, the UL11 condensates primarily formed in cold conditions and dissipated before reaching physiological temperature, in the presence of molecular crowders, concentrated UL11 underwent phase separation at room temperature. Thus, specific conditions may promote formation of UL11 biomolecular condensates *in vivo*, for example, binding to gE, UL16, or RNA or clustering within lipid rafts (24). Reevaluating UL11 puncta formed *in vivo* in the context of biomolecular condensates will be important for elucidating the mechanisms of phase separation by UL11 *in vivo* and their dependence on binding partners.

Along with other tegument proteins, UL11 is important for virion assembly during capsid budding at the cytoplasmic membranes. However, how the tegument assembles around the capsid and how the tegument proteins are arranged within the tegument layer are unclear. It is tempting to speculate that rather than forming an ordered structure, the outer tegument exists as a biomolecular condensate, a heterogeneous milieu composed of many partially disordered molecules that make multiple redundant nonspecific interactions. Indeed, outer tegument proteins create a complex network that has both specific (3) and nonspecific interactions that may be important for tegument formation (6). In fact, several outer tegument proteins, including UL21 (61) and UL48 (63), have been reported to contain short or long IDRs, and our analysis suggests that all 24 HSV-1 tegument proteins contain disordered regions of variable lengths (Fig. 9). A biomolecular condensate model of the outer tegument could explain why deletions of individual tegument proteins typically do not disrupt the tegument layer and cause only modest defects in viral replication. Puzzling reports that virions missing individual outer tegument proteins have increased amounts of other tegument proteins, RNA, or even host proteins (2, 77) can be explained by the latter serving as “space fillers” within the biological condensate of the tegument. The ability of the outer tegument to form a biomolecular condensate, rather than a strictly ordered protein network, would allow for productive replication even when the availability of one or more outer tegument proteins fluctuates. Furthermore, based on their liquid-like characteristics, phase-separated droplets are spherical (reviewed in reference 72). Formation of such phase-separated droplets, driven by membrane-associated tegument proteins such as UL11, could promote budding during secondary envelopment by causing the spherical bud to pinch off into *trans*-Golgi network (TGN)-derived membranes, akin to a lava lamp. Such a budding mechanism would not require the capsid, which is consistent with the formation of capsid-less L particles during normal HSV-1 infection (1). Indeed, phase separation has recently been suggested as a mechanistic mediator of membrane deformation during endocytosis (78). Furthermore, viral inclusions formed during replication and assembly of viruses such as rabies virus (79) and influenza virus (80) share characteristics of membrane-less organelles. Finally, it has been proposed that herpesviral replication compartments in the nucleus—the sites of viral gene expression and replication—may be formed by a mechanism that involves LLPS (81). We hypothesize that secondary envelopment is a new example of the use of LLPS in herpesviral replication. While this proposed mechanism does not rule out contributions of host factors or undiscovered viral mediators of membrane deformation and scission, it would ensure envelopment regardless of whether these mediators are present. Future studies of the tegument structure and assembly mechanisms should, thus, take into account the presence of structural disorder in tegument proteins.
MATERIALS AND METHODS

Sequences and analyses. Sequences for UL11 homologs from eight human herpesviruses: herpes simplex virus 1 (HSV-1) strain 17 UL11 (RefSeq accession no. YP_009137085.1), HSV-2 strain HG52 UL11 (RefSeq YP_009137162.1), vesicular stomatitis virus (VSV) strain Dumas ORF49 (RefSeq NP_040171.1), Epstein-Barr virus (EBV) strain B95-8 BBLF1 (RefSeq YP_401686.1), cytomegalovirus (CMV) strain AD169 UL99 (RefSeq P13200.3), human herpesvirus 6A (HHV-6A) strain Uganda-1102 U71 (RefSeq NP_042964.1), HHV-6B strain Z29 U71 (RefSeq NP_050250.1), HHV-7 strain JI U71 (RefSeq YP_073811.1), Kaposi’s sarcoma herpesvirus (KSHV) strain GK18 ORF38 (RefSeq YP_001129391.1), Marek’s disease virus (MDV) strain Md5 UL11 (RefSeq YP_001033939.1), pseudorabies virus (PRV) composite strain UL11 (RefSeq YP_068364.1), murine cytomegalovirus (MCMV) strain Smith UL99 (RefSeq YP_214100.1), saimiriine herpesvirus 2 (SaHV-2) ORF38 (RefSeq NP_040240.1), and equine herpesvirus 2 (EHV-2) strain 86/67 myristoylated tegument protein (RefSeq NP_042635.1) were aligned in Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) and adjusted manually. ExPASy ProtParam (http://web.expasy.org/protparam/) was used to calculate protein molecular masses, amino acid content, and extinction coefficients. CSSpalm (http://csspalm.biocuckoo.org) was used to predict palmitoylation sites. Myristylation was marked on glycine 2 based on previous studies (13–15) and conservation. JPred (82) (http://www.compbio.dundee.ac.uk/jpred/) and PsiPred (83) (bioinf.cs.ucl.ac.uk/psipred/) were used to predict secondary structure. FoldUnfold (84) (http://bioinfo.protres.ru/ogu/) was used with default settings to predict disorder. IUPred2A (85) (https://iupred2a.eel.hu/) was used with default settings to predict structured domains.

Cloning and expression constructs. A plasmid encoding full-length UL11 from HSV-1 strain 17 preceded by a His6 tag and a thrombin cleavage site (His6-UL11) in a pET28 vector was a gift from John Wills (Pennsylvania State University). This plasmid contains an A8T mutation in UL11, but because this region of protein is poorly conserved (Fig. 1A), the mutation was not considered problematic. Furthermore, UL11 with this mutation successfully rescues a UL11-null virus (25). UL11 was amplified from the pET28 vector using primers CM121 and CM122 (see Table S1 in the supplemental material), digested with NcoI and XhoI, and subcloned into the NcoI/XhoI-digested pGEX-6P-1 backbone to produce plasmid pCM054 which encodes UL11 preceded by a glutathione S-transferase (GST) tag, a human rhinovirus...
(HRV) 3C (PreScission) protease site, and a Leu-Gly-Ser linker, GST-HRV3C-UL11. To add a C-terminal Strep-tag II (StII), single overlap extension (SOE) PCR was performed on pCM054 using primers pGEX3, pGEX5, CM131, and CM132 (Table S1). The insert was digested with Ncol and Xhol and subcloned into the Ncol/Xhol-digested pCM054 vector to produce plasmid pCM056 encoding GST-HRV3C-UL11-StII. Finally, the same insert was subcloned into the original pET-28 vector, again using Ncol and Xhol, to produce plasmid pCM055 encoding UL11-StII.

Recombinant protein expression and purification. Hs-UL11 was expressed in E. coli strain T7 cells in LB for 4 to 8 h at 37°C using induction with 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) at an optical density at 600 nm (OD600) of 0.6 to 1.0. GST-HRV3C-UL11 was expressed in E. coli strain Rosetta (Novagen) cells using an autoinduction protocol (86). Briefly, a small culture of cells, grown overnight at 37°C in terrific broth (TB) with 1% glucose and 2 mM MgSO4, was inoculated at a ratio of 1:100 into TB supplemented with 0.2% lactose and 2 mM MgSO4 and grown for 4 h at 37°C and then for 18 to 22 h at 25°C. St-II-tagged constructs were expressed in E. coli strain Rosetta or LOBSTR (87) cells at 37°C using IPTG induction as described above.

Cells were harvested by centrifugation at 5,000 × g for 30 min at 4°C, resuspended in buffer A [50 mM HEPES, 100 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) (pH 7.5)] with added 0.1 mM phenyl-formaldehyde and run at 75 V in 1/1000 MOPS buffer [200 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 30 mM sodium acetate trihydrate]. After heating for 1 to 3 min at 85°C and cooling, 25% FORMAzol (Molecular Research Center, Inc.), and 0.05 mg/ml ethidium bromide. After heating for 1 to 3 min at 85°C and cooling to 4°C, samples were immediately loaded onto a 1.2% agarose gel prepared in 1× MOPS buffer with 5% formaldehyde and run at 75 V in 1× MOPS buffer.
Crystal screening. Crystallization trials of UL11 constructs at several concentrations (12 to 18 mg/ml) were set up using nine screens (Classics Suite [Qiagen], Index [Hampton Research], in-house Harrison Lab Grid Screen, Peg/Ion [Hampton Research], Protein Complex Suite [Qiagen], SaltRx [Hampton Research], Top 96 [Anatrace], Wizard 1-4 [Rigaku]) in 96-well sitting-drop trays with drops containing 0.2 μl protein and 0.2 μl crystallization solution dispensed by Phoenix liquid handling robot (Art Robbins). Crystallization plates were stored at ambient temperature and evaluated using a stereo microscope daily for a week and periodically during subsequent months.

Phase separation assays. Macroscopic and microscopic evaluation of UL11 liquid-liquid phase separation (LLPS) was performed using UL11-StII purified as described above in buffer A at various concentrations. In the experiment shown in Fig. 5B and C, UL11-StII was diluted from 18 mg/ml to 9 mg/ml, chilled in a 4°C chill block to induce droplet formation, warmed by hand to clear droplet formation, and either viewed by eye in a tube or with a stereo light microscope as a 2-μl drop on a glass coverslip. The LLPS behavior was also seen at the 18 mg/ml concentration and after one cycle of freezing and thawing. In the experiment shown in Fig. 5D, purified UL11-StII was concentrated to 10 mg/ml, flash frozen, and stored at ~80°C. Samples were thawed and diluted to marked concentrations in buffer A with different amounts of NaCl or the addition of molecular crowders (100 mg/ml bovine serum albumin [BSA] or 30% polyethylene glycol 350 monomethylether [PEG 350 MME]). For the LLPS experiment in circular dichroism (CD) buffer (Fig. 6A), UL11-StII was subjected to size exclusion chromatography in buffer A containing no NaCl (50 mM HEPES, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) [pH 7.5]) before frozen storage and diluted to marked concentrations in CD buffer (20 mM sodium phosphate, 100 mM sodium fluoride [pH 8.0]) after thawing. All samples in the LLPS assay were incubated in a 4°C chill block for 2 h, and aliquots were transferred at 4°C to a 96-well plate and stored for 24 h at either 4°C or ambient temperature before visualization with a stereo light microscope. Similar results were seen after 24-h storage in either condition.

Circular dichroism. CD was used to investigate secondary structure content of UL11. Purified, nucleic acid-free UL11-StII in buffer A was exchanged into CD buffer (20 mM sodium phosphate, 100 mM sodium fluoride [pH 8.0]) using a PD SpinTrap G-25 column (GE Healthcare Life Sciences) and diluted to various concentrations. After storage at 4°C, far-UV spectral scans from 185 to 300 nm were taken in a 1-mm cuvette at 20°C on a JASCO J-815 spectropolarimeter, continuously scanning at 50 nm/min with 1-nm bandwidth and 1-s data integration time. Five spectra were collected and averaged, and buffer was subtracted for each sample. Machine data were converted to mean residue ellipticity (MRE) using the equation MRE = (θ)0 (MRW × i)/(10 × c × l), where MRW is mean residue weight = molecular mass/(number of amino acids − 1), θ(0) is ellipticity (millidegrees), c is concentration in grams per liter, and l is path length (centimeters).

Limited proteolysis, mass spectrometry, and N-terminal sequencing. Limited proteolysis was conducted as described previously (88). Briefly, 3 μg purified UL11-StII (at 1.5 mg/ml) was incubated for 1 h at room temperature in buffer A with Nα-p-tosyl-L-lysine chloromethyl ketone (TLCK)-treated chymotrypsin/UL11 at a ratio of 1:2,000, 1,000, 500, 200, 100, 50, or 25; tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin/UL11 at a ratio of 1:3,200, 1,600, 800, 400, 200, 100, 50, or 25; proteinase K/UL11 at a ratio of 1:2,000, 1,000, 500, 250, 100, 50, 25, or 5; V8 protease/UL11 at a ratio of 1:2,000, 1,000, 500, 100, 50, 25. All proteases were from Worthington. Samples were boiled for ~5 min at 95°C with 1× loading buffer (2% SDS, 4% glycerol, 40 mM Tris [pH 6.8], 2% β-mercaptoethanol) to stop the digest and analyzed by SDS-PAGE. For mass spectrometry analysis, 55 μg (at 1.5 mg/ml) purified UL11-StII was digested with V8 protease at a 1:50 protease-to-protein mass ratio. After an hour, phenylmethanesulfonyl fluoride (PMSF) was added to a final concentration of 10 mM to stop the reaction. The samples were analyzed by matrix-assisted laser desorption ionization—time of flight mass spectrometry (MALDI-TOF MS) matrix a dihydroxybenzoic acid (DBH) matrix on a Voyager DE-PRO instrument at the Tufts University Core Facility. For N-terminal sequencing, 33 μg (at 1.5 mg/ml) purified UL11-StII was incubated with V8-protease at a 1:50 protease/protein ratio. The sample was separated on a 4 to 15% SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membrane in transfer buffer (50 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS) [pH 10.5], 10% methanol) using a semidry transfer apparatus. The membrane was stained in 0.05% Coomassie blue R-250 dissolved in 40% methanol and destained first in 50% methanol and then in water. Bands were cut from the air-dried membrane and subjected to N-terminal sequencing by Edman degrada tion at the Tufts University Core Facility.

Cross-linking. Chemical cross-linking was performed as described previously (86). One hundred picomoles of purified UL11-StII in water was incubated with bis(sulfosuccinimidyl) carbonate (BS5) (Fisher Scientific) at 0, 1, 2, 4, 8, or 10 molar ratios of cross-linker to protein (1.5 mg/ml) in amber tubes in the dark at room temperature for 20 min. To stop the reaction, 1 M Tris-HCl (pH 8.0) was added to the final concentration of 50 mM and incubated for 15 min. Half of the reaction was loaded onto an SDS-PAG and visualized by Coomassie blue staining.

Size exclusion chromatography—small angle X-ray scattering. Small-angle X-ray scattering coupled with size exclusion chromatography (SEC-SAXS) experiments were conducted at the BioSAXS beamline G1 at the Cornell High Energy Synchrotron Source (Ithaca, NY). UL11-StII was centrifuged at 16,000 × g for 5 min at 4°C to remove aggregates, and a 100-μl sample was injected onto a Superdex 200 Increase 5/150 (GE Healthcare Life Sciences) in buffer A at 0.3 ml/min and 4°C and fed directly into a proprietary SEC-SAXS flow cell with a glass window. To ensure an adequate signal, the concentration of UL11-StII was ~18 mg/ml. Samples were irradiated with a 9.9496-keV (1.246122-Å) beam with 5.7 × 10^11 photons/s flux and diameter of 250 μm by 250 μm, and 2-s images were recorded through the duration of the run on a Pilatus 100K-S detector (Dectris) at a sample-to-detector distance of 1,521.09 mm
resulting in a q range of 0.006 Å⁻¹ < q < 0.8 Å⁻¹ where scattering vector q = 4πsin(θ)/λ. SANS data were processed using the RAW software package (https://sourceforge.net/projects/bioxtasraw/) (89) by selecting and averaging the scattering curves from buffer data frames (evaluated by integrated scattering intensity correlated with background) subtracting this averaged buffer curve from each subsequent curve and calculating the resulting radius of gyration (Rg), selecting and averaging sample curves with a consistent Rg and finally subtracting the buffer average curve from the sample average curve. These curves were further analyzed using AUTORG, DATGNOM, AMBIMETER (90), and EOM 2.0 (59) as implemented in RAW and the ATSAS software package (https://www.embl-hamburg.de/biosaxs/software.html) (91). AUTORG was used to define the Guinier region, calculate Rg and I(0) parameters, and evaluate uncertainty in these parameters. DATGNOM (as implemented in RAW) was used to generate the pair distance distribution function [P(r)] from the subtracted experimental data and automatically evaluate it prior to manually adjusting the automatically assigned Dmax value to improve P(r) shape. AMBIMETER was used to predict whether ab initio models generated from the P(r) would be unique. EOM 2.0 was used to generate a pool of 10,000 macromolecules randomly sampling the conformational space available to the UL11-StII sequence. Theoretical SANS curves were generated for random combinations of these models, and the subensemble of conformers coexisting in solution that yielded the best fit to the experimental SANS data was selected. Subensemble selection, or ensemble optimization, was performed three times for the same pool of random models. No symmetry information or structures were specified as constraints in EOM, which allowed for the generation of random configurations of the backbone based only on sequence. Rg and Dmax values calculated from the species in each group were represented in histograms.

Data availability. Small-angle scattering data have been deposited into the SASBDB under code SASDEX4.

SUPPLEMENTAL MATERIAL
Supplemental material is available online only.

FIG S1, PDF file, 1.5 MB.

TABLE S1, PDF file, 0.03 MB.

ACKNOWLEDGMENTS
We thank MacCHESS BioSAXS staff, especially Jesse Hopkins, for their advice on SANS data collection and analysis and for helpful discussions. We thank John Wills (Penn State University) for the gift of the plasmid encoding H6-UL11, Peter Cherepanov (Francis Crick Institute) for the gift of the GST-PreScission protease expression plasmid, and Thomas Schwartz (Massachusetts Institute of Technology) for the gift of LOBSTR cells. We also thank Jose Martin Ramirez and Michael Thorsen for assistance with protein expression and purification, Xuanzong Guo for purifying PreScission protease, Marta Gaglia (Tufts University) for reagents used in RNA experiments, and Michael Berne and the Tufts University Core Facility for performing mass spectrometry and N-terminal protein sequencing. Circular dichroism experiments were performed at the Center for Macromolecular Interactions in the Department of Biological Chemistry and Molecular Pharmacology at Harvard Medical School. All software was installed and maintained by SBGrid (92).

This work was funded by the Investigators in the Pathogenesis of Infectious Disease Award from Burroughs Wellcome Fund (E.E.H.), by the NIH grant R01 GM111795 (E.E.H.), and Faculty Scholar grant 55108533 from Howard Hughes Medical Institute (E.E.H.). C.M.M. was supported by the NIH fellowship F31 GM115121. This work is based on research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation and the National Institutes of Health/National Institute of General Medical Sciences under NSF award DMR-0936384, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award GM-103485 from the National Institutes of Health through the National Institute of General Medical Science.

We declare that we have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES


