Received: 28 June 2023 | Revised: 8 September 2023 | Accepted: 17 September 2023

DOI: 10.1002/jcb.30480

RESEARCH ARTICLE

Journal of Cellular Biochemistry WILEY

Comprehensive analysis of liquid–liquid phase separation propensities of HSV-1 proteins and their interaction with host factors

Sushma Subedi[1] | **Niharika Nag**[1] | **Harish Shukla**[1] | **Aditya K. Padhi**[2] | **Timir Tripathi**[1,3]

[1]Molecular and Structural Biophysics Laboratory, Department of Biochemistry, North-Eastern Hill University, Shillong, India

[2]Laboratory for Computational Biology & Biomolecular Design, School of Biochemical Engineering, Indian Institute of Technology (BHU), Varanasi, India

[3]Department of Zoology, North-Eastern Hill University, Shillong, India

Correspondence
Timir Tripathi, Molecular and Structural Biophysics Laboratory, Department of Biochemistry, North-Eastern Hill University, Shillong- 793022, India; Department of Zoology, North-Eastern Hill University, Shillong, India.
Email: timir.tripathi@gmail.com

Funding information
Indian Council of Medical Research (ICMR), Grant/Award Number: 52/06/2020/BIO/BMS; Department of Health Research (DHR), Grant/Award Number: R.11013/47/2021-GIA/HR

Abstract

In recent years, it has been shown that the liquid–liquid phase separation (LLPS) of virus proteins plays a crucial role in their life cycle. It promotes the formation of viral replication organelles, concentrating viral components for efficient replication and facilitates the assembly of viral particles. LLPS has emerged as a crucial process in the replication and assembly of herpes simplex virus-1 (HSV-1). Recent studies have identified several HSV-1 proteins involved in LLPS, including the myristylated tegument protein UL11 and infected cell protein 4; however, a complete proteome-level understanding of the LLPS-prone HSV-1 proteins is not available. We provide a comprehensive analysis of the HSV-1 proteome and explore the potential of its proteins to undergo LLPS. By integrating sequence analysis, prediction algorithms and an array of tools and servers, we identified 10 HSV-1 proteins that exhibit high LLPS potential. By analysing the amino acid sequences of the LLPS-prone proteins, we identified specific sequence motifs and enriched amino acid residues commonly found in LLPS-prone regions. Our findings reveal a diverse range of LLPS-prone proteins within the HSV-1, which are involved in critical viral processes such as replication, transcriptional regulation and assembly of viral particles. This suggests that LLPS might play a crucial role in facilitating the formation of specialized viral replication compartments and the assembly of HSV-1 virion. The identification of LLPS-prone proteins in HSV-1 opens up new avenues for understanding the molecular mechanisms underlying viral pathogenesis. Our work provides valuable insights into the LLPS landscape of HSV-1, highlighting potential targets for further experimental validation and enhancing our understanding of viral replication and pathogenesis.

KEYWORDS
AlphaFold, disordered proteins, herpes simplex virus-1, liquid–liquid phase separation, protein–protein interaction, structure prediction, viral proteins

Sushma Subedi and Niharika Nag contributed equally to the work.

1 | INTRODUCTION

Herpes simplex virus-1 (HSV-1) is a highly pathogenic double-stranded DNA virus that infects humans worldwide. It is responsible for various clinical manifestations, including oral and genital herpes, encephalitis and ocular infections. The structure of HSV-1 comprises four distinct components: a core containing DNA, an icosahedral capsid that encases the genetic material, a surrounding layer around the capsid known as the tegument and an outer envelope enclosing the entire assembly, within which the viral glycoproteins are embedded.[1,2] The 152 kb HSV-1 genome contains ~80 open reading frames (ORFs) that encode proteins involved in various stages of the viral life cycle.[3] These genes are classified into three classes: immediate-early (IE), early (E) and late (L).[4] The IE proteins are the first viral proteins synthesized following viral infection and play critical roles in initiating viral gene expression and modulating host cellular processes. IE protein ICP0 is an E3 ubiquitin ligase and targets host factors involved in antiviral responses and chromatin remodelling for degradation, enabling efficient viral replication and immune evasion.[5] The E proteins are synthesized after IE proteins and primarily function in viral DNA replication, transcriptional regulation and modulation of host responses. The E protein ICP8, a major HSV single-strand DNA-binding protein, is essential for viral DNA replication.[6] It has been found to interact with multiple viral and host factors involved in DNA synthesis and repair processes. The L proteins are synthesized during the late phase of viral replication and are predominantly structural proteins involved in virion assembly and maturation. A significant focus of recent research has been on deciphering the intricate interactions and regulatory networks between IE, E and L proteins and host factors. HSV-1 proteins interact with host factors, influencing viral pathogenesis and immune responses.[7–10]

Intrinsically disordered proteins (IDPs) and their associated intrinsically disordered regions (IDRs) lack a fixed, stable structure. Instead, they exist as an ensemble and comprise a dynamic collection of conformations that fluctuate over time and among different populations.[11–13] An IDR is defined as a region containing a continuous set of at least 30 disordered amino acid residues.[11,14,15] A protein is considered an IDP if it contains at least one IDR.[11] Liquid–liquid phase separation (LLPS) of proteins is a fascinating phenomenon observed in cells, where certain proteins undergo self-assembly to form liquid-like compartments. This process plays a crucial role in cellular organization, as these compartments facilitate the spatial segregation of cellular components and regulate various biological processes.[16–18] Recent research has revealed that several viral proteins are intrinsically disordered and can undergo LLPS, forming viral replication compartments or inclusion bodies.[19–23] LLPS of viral proteins is increasingly recognized as a fundamental mechanism for the assembly and function of viral replication complexes.[24,25] Various viruses, including RNA viruses, DNA viruses and retroviruses, exploit LLPS to create specialized membrane-less compartments or membraneless organelles (MLOs) to promote efficient viral replication and particle assembly.[26] These compartments provide an environment that facilitates efficient viral genome replication, transcription and translation. The formation of these replication compartments is often mediated by specific viral proteins that possess IDRs or low-complexity domains. The LLPS of viral proteins relies on several factors, including protein–protein interactions (PPIs), posttranslational modifications (PTMs) and environmental conditions.[27] Viral proteins contain specific domains or motifs that promote LLPS formation, such as prion-like domains, multivalent protein-interaction domains or RNA-binding domains. Interactions between these domains can lead to the formation of dynamic liquid-like droplets, which can subsequently mature into more stable and membrane-less compartments. The presence of viral nucleic acids and host factors further modulate the phase behaviour of viral protein droplets. LLPS of viral proteins not only facilitates viral replication but also impacts viral pathogenesis.[28] The formation of viral replication compartments sequesters viral components and protects them from host antiviral defences, including the immune system and cellular restriction factors. The dynamic nature of viral protein droplets also allows for rapid and localized viral genome replication and gene expression. Furthermore, the recruitment of host factors into viral protein droplets can modulate cellular processes, leading to alterations in host cell physiology and immune responses.

Recent studies have revealed that certain HSV-1 proteins, such as infected cell protein 4 (ICP4 or RS1) and myristylated tegument protein UL11, are capable of undergoing LLPS.[21,29] These droplets serve as sites for the concentration of viral replication machinery, viral genome transcription and the sequestration of viral components, protecting them from host antiviral defences.[21] However, a comprehensive analysis of the HSV-1 proteome to assess their potential to undergo LLPS has not been performed. In this study, we conducted an in-depth examination of the HSV-1 proteome to explore the potential of its proteins to undergo LLPS.

2 | METHODS

2.1 | Retrieval of protein sequences

The sequences of the HSV-1 proteins were retrieved from NCBI. A total of 74 protein sequences were retrieved and used further for predictions and analysis (Supporting Information: Table S1). The names and NCBI Protein IDs of the proteins used in this study are myristylated tegument protein UL11 (YP_009137085.1), transcriptional regulator ICP4 RS1 (YP_009137149.1), capsid maturation protease UL26 (YP_009137100.1), capsid scaffold protein UL26.5 (YP_009137101.1), envelope glycoprotein C (gC) UL44 (YP_009137119.1), large tegument protein UL36 (YP_009137111.1), multifunctional expression regulator UL54 (YP_009137130.1), tegument protein US11 (YP_009137147.1), tegument protein VP11/12 (YP_009137121.1) and ubiquitin E3 ligase ICP0 RL2 (YP_009137074.1).

2.2 | Compositional profiling

The amino acid composition of a protein provides valuable insights into its structural and conformational properties. We utilized a freely available web server called the Composition Profiler to analyse the composition profile.[30] This profiler calculates the fractional difference in composition, $(C_x - C_{\text{order}})/(C_{\text{order}})$, for each amino acid within the protein query set (HSV-1 proteins) and a reference set of disordered proteins obtained from the DisProt database. C_x represents the amino acid content in the protein query and reference sets in this formula. At the same time, C_{order} denotes the standard value derived from a set of ordered proteins obtained from the PDB Select25.

2.3 | Per residue disorder propensity analysis

The disorder propensities of the HSV-1 proteins were analysed using the protein disorder predictor tools such as IUPred3 (long, short and ANCHOR2),[31] PONDR (VLXT and VSL2),[32] PSIPRED[33] and DISOPRED3.[34] IUPred3 is a comprehensive web interface that integrates the functionalities of IUPred2 and ANCHOR2. It enables the identification of both disordered protein regions using IUPred2 and disordered binding regions using ANCHOR2. By combining these two prediction methods, IUPred3 provides a unified platform for analysing and characterizing both intrinsic disorder and binding-induced disorder-to-order transitions in protein sequences. PONDR VXLT and PONDR VSL2 are protein disorder prediction methods that utilize neural network-based approaches and sequence-based features to predict intrinsic disorder. PONDR VXLT considers features such as amino acid composition and predicted secondary structure, whereas PONDR VSL2 incorporates an ensemble of neural networks trained on diverse feature sets, including composition, predicted secondary structure and sequence conservation, to enhance prediction accuracy. PSIPRED is a method for predicting protein secondary structure that employs a two-stage neural network approach. In the first stage, a feed-forward neural network predicts the probabilities of each residue belonging to helix, strand or coil secondary structure classes. In the second stage, a second neural network improves the predictions by incorporating information from neighbouring residues to assign more precise secondary structure labels to each residue. DISOPRED3 is an updated machine-learning approach that excels in detecting IDRs in proteins. It integrates a neural network and a nearest neighbour classifier, trained on extensive data from the PDB and DisProt databases, to identify long disordered regions. The output of the predictors was in the form of a score value between 0 and 1 for the individual residue of a particular protein. The residues having a score above 0.5 were considered disordered.

2.4 | Structure prediction of the shortlisted HSV-1 disordered proteins

The 10 disordered proteins, which showed promising potential for phase separation, underwent structure prediction due to the unavailability of their complete structures. To gain a deeper understanding of their disorderedness and grasp their entire structural topology, AlphaFold v2.0, a state-of-the-art tool, was employed.[35] The run_docker.py Python script was utilized to specify the FASTA file containing the protein sequences, while setting the max_template_date to a recent date ensured the inclusion of the most up-to-date templates for comprehensive model construction source: https://github.com/deepmind/alphafold.

2.5 | Prediction of LLPS potential

To predict the phase separation potential of the HSV-1 proteins, we used sequence-based predictor tools such as PSPredictor,[36] catGranule[37] and FuzDrop.[38] PSPredictor is a machine learning-based tool that integrates with the LLPSDB database. It enables the prediction of protein phase separation potential and provides information about

similar proteins in the LLPSDB database and their phase separation conditions. The output is a score between 0 and 1, where a score above 0.5 suggests a more substantial likelihood of phase separation. catGranule utilizes structural disorder, nucleic acid-binding properties and primary amino acid sequence composition to predict the propensity of a protein to undergo LLPS formation. A score >0 indicates a likelihood of undergoing LLPS, whereas a score >1 indicates a strong ability to phase separate. The FuzDrop server estimates the likelihood of proteins undergoing LLPS. It employs a sequence-based approach to identify regions within proteins that promote droplet formation and regions that contribute to the aggregation of proteins within these droplets. The scores can range from 0 to 1, where a score closer to 1 suggests a stronger likelihood of phase separation.

2.6 | Molecular recognition features (MoRF) analysis

Prediction of MoRF sites for each protein was carried out using different predictors like DISOPRED3,[34] MoRFpred[39] and MoRFchibi SYSTEM.[40] Each predictor uses a different algorithm to predict the MoRF site in a protein sequence. DISOPRED3 identifies the disordered regions in a protein sequence and annotates them as protein binding regions through an additional support vector machine (SVM) classifier. It is more sensitive to IDRs longer than twenty amino acid residues. In MoRFpred, the annotations generated by sequence alignment are fused with the predictions generated by SVM, which uses a custom set of sequence-derived features. The MoRFchibi SYSTEM is a set of software tools that includes three predictors: MoRFCHiBi, a fundamental predictor component in other applications; MoRFCHiBi_Light for high-throughput prediction; and MoRFCHiBi_Web for high accuracy predictions.

2.7 | Identification of nucleic acid-binding residues

The messenger RNA (mRNA) and DNA-binding sites on each protein of HSV-1 were predicted using DisoRPDbind and DRNApred web servers. Both are sequence-based predictive tools that identify binding residues within IDRs capable of interacting with RNA, DNA and proteins.[41,42] DisoRPDbind implements a runtime-efficient multilayered design incorporating information from various sources, such as amino acid physiochemical properties, sequence complexity, putative secondary structure and disorder, and sequence alignment. By leveraging these features, DisoRDPbind aims to predict potential binding sites in disordered regions that are traditionally challenging to identify using conventional structure-based methods. Residues lacking any annotations are represented as 'x'. DRNApred incorporates a novel two-layered architecture in its predictive model. This architecture enables the method to capture intricate features and patterns associated with DNA- and RNA-binding residues, leading to improved prediction performance. The output for both predictors is annotated as 0 and 1, where '1' represents binding residues and '0' nonbinding residues.

2.8 | PPI networks of HSV-1 proteins

The PPI of HSV-1 proteins with the human protein was analysed using Viruses.STRING.[43] Viruses.STRING is a database focusing on interactions between proteins, specifically designed to address interactions between viruses themselves and between viruses and their host organisms. By integrating information from experimental studies and text mining sources, this database offers consolidated probabilities for protein interactions between viruses and their hosts. In the network, the nodes represent the protein, and the edges represent the predicted functional associations. An edge may be drawn with up to seven differently coloured lines that represent the existence of the seven types of evidence used in predicting the associations.

3 | RESULTS

Seventy-four protein sequences were retrieved from the NCBI database and subject to the prediction of disorder using seven predictors, namely IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. The analysis of all 74 proteins revealed that 27 proteins were predicted to be disordered (Supporting Information: Table S2). These proteins contained at least one disordered stretch comprising more than 30 amino acid residues and exhibited an overall sequence disorderedness exceeding 30%. Out of 27 disordered proteins, 10 that demonstrated the potential to undergo phase separation were selected for further analysis.

3.1 | Myristylated tegument protein (UL11)

The myristylated tegument protein (UL11) localizes to the Golgi and attaches to membranes through myristoylation and palmitoylation (Table 1). UL11 is a 96 amino acid

TABLE 1 Functional roles of HSV-1 proteins used in this study.

Protein	Functional role
Myristylated tegument protein (UL11)	• Golgi-localized, membrane-associated outer-tegument protein.[44] • UL11 homologues in herpesviruses, such as HSV-1, undergo N-terminal myristoylation[45] and palmitoylation,[46] crucial for membrane attachment and lipid raft association.[21,47] • It is an IDP and undergoes LLPS in vitro.[21]
Transcriptional regulator ICP4 (ICP4 or RS1)	• It is vital for viral transcription regulation during infection.[48] • It interacts with cellular RNA pol II machinery to control gene expression, including TBP, TFIIB, and TAF1.[20]
Capsid maturation protease (UL26)	• It contributes to viral capsid assembly.[49] • UL26's proteolytic activity processes itself and ICP35.[50] • HSV-1 capsid proteins form B-capsids in the nucleus, with VP22a, VP21 and VP24 scaffolding proteins.[51–53] • It has a role to play in DNA packaging.[54]
Capsid scaffold protein (ICP35 or UL26.5)	• ICP35 shapes viral capsids and is processed by the UL26 gene's protease.[55] • It also participates in VP5 nuclear transport.[56]
Envelope gC (UL44)	• It aids viral entry by binding to C3b, inhibiting complement immunity.[57] • It interacts with cell surface proteoglycans, enabling entry and immune evasion.[57] • Its absence in HSV-1's mutant leads to the lack of a C3b receptor.[58] • Its role in immune shielding may contribute to HSV-1 persistence and is a target for antibody neutralization, highlighting its significance in concealing vital viral components from immune recognition.[57,59]
Large tegument protein (UL36)	• UL36 cross-links capsid, tegument and membrane proteins during viral assembly.[60] • Upon cell entry, it releases capsids from the outer tegument and envelope. UL36-deficient capsids fail to target membranes for assembly and nuclear pores for genome uncoating.[60]
Multifunctional expression regulator (ICP27 or UL54)	• It serves diverse functions: viral mRNA processing and export, nuclear protein quality control, cell cycle regulation, stress signalling activation and antiapoptosis.[61] • It interacts with Nup62 in the host, inhibiting the host nucleocytoplasmic transport pathways.[62]
Tegument protein (US11)	• It is linked to host translational arrest pathways and found in HSV-1-infected cell-derived vesicles.[63] • It interacts with kinesin's heavy chain and cellular PAT1 polypeptide, suggesting involvement in unenveloped capsid trafficking.[64,65]
Tegument protein 11/12 (VP11/12 or UL46)	• It is a phosphoprotein that boosts viral gene transcription via VP16 and activates cell survival while suppressing innate immunity.[66] • It interacts with Src family kinase through C-terminal tyrosine-based motifs (YETV and YEEI), triggering Src activation and VP11/12 phosphorylation.[67,68]
Ubiquitin E3 ligase (ICP0 or RL2)	• ICP0 in HSV-1, initially identified as a nuclear phosphoprotein, is pivotal in low MOI cell cultures.[10] • It serves as a transactivator of genes that are introduced by transfection or infection.[69–72] • It also curbs innate immunity, affects endocytosis and performs other roles.[10]

Abbreviations: gC, glycoprotein C; HSV-1, herpes simplex virus-1; ICP4, infected cell protein 4; IDP, intrinsically disordered protein; LLPS, liquid–liquid phase separation; MOI, multiplicity of infection; mRNA, messenger RNA; RNA pol II, RNA polymerase II; TBP, TATA-Box-binding protein.

residue protein with a molecular weight of ~10.5 kDa and a theoretical pI of 5.2. UL11 comprises 22.43% disorder-promoting amino acid residues (Figure 1A), with a mean predicted percentile of intrinsic disorder ($PPID_{Mean}$) amounting to 64.58% (Table 2 and Figure 1C). An analysis of the compositional profiling of the protein revealed its enrichment in disorder-promoting amino acids Pro and Ser, whereas it exhibited a deficiency in order-promoting amino acids Trp, Ile and Phe (Figure 1B). Detailed information about the amino acid composition of the protein is shown in Figure 1A. The mean disorder propensity analysis of the protein showed that UL11

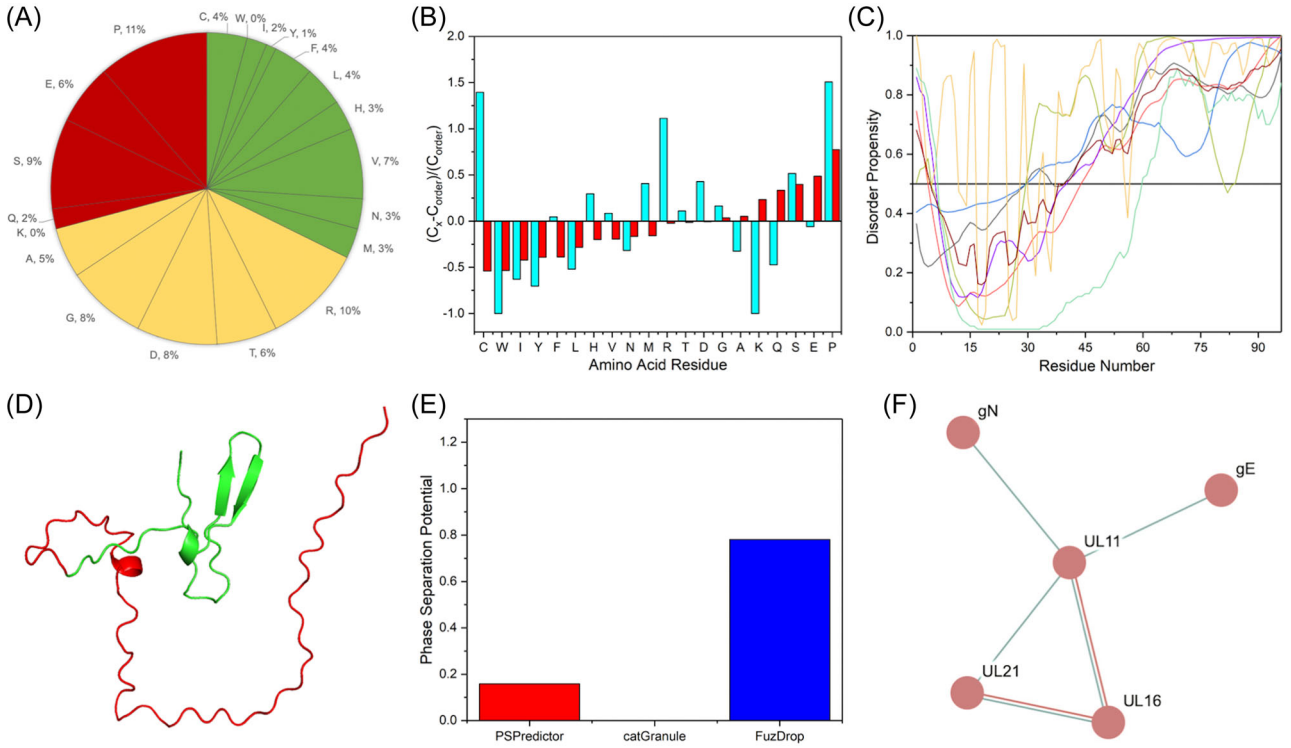

FIGURE 1 Analysis of myristylated tegument protein (UL11). (A) Amino acid composition. Red indicates disorder-promoting amino acids, green indicates order-promoting amino acids and yellow indicates neutral amino acids. (B) Compositional analysis. UL11 protein sequence was compared to a set of protein sequences from DisProt (red bars). Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. Their predictions are depicted using black, red, blue, green, purple, orange and teal lines, respectively. The mean of the predicted scores is depicted by the maroon line. (D) AlphaFold-based structural modelling. Model building and predicted structure of UL11 were generated using AlphaFold v2.0. The red colour depicts the disordered region, while the green colour depicts the ordered region of the protein. The colour scheme is based on the disorder propensity prediction. (E) Liquid–liquid phase separation propensities of UL11. Prediction scores by PSPredictor are shown in red bars, catGranule scores in green bars and FuzDrop scores in blue bars. (F) Virus.STRING analysis. The red nodes indicate herpes simplex virus-1 proteins, whereas the teal nodes indicate host proteins. Lines of various colours indicate different types of evidence for interactions, including fusion evidence (pink line), neighbourhood evidence (green line), co-occurrence evidence (dark blue line), experimental evidence (red line), text mining evidence (teal line), database evidence (orange line), co-expression evidence (black line) and protein homology (light blue line).

contains one disordered region with >30 amino acid residues spanning from amino acid residues 40–96 (total 57 residues) (Table 3).

Upon inspection of the UL11 structure using AlphaFold, it was observed that the per-residue predicted local distance difference test (pLDDT) scores for the residues 20–40 were close to 90, indicating a high level of model accuracy (Figure S1A). The residues 40–96 exhibited pLDDT scores between 40 and 60, signifying lower confidence and a less favourable backbone configuration. This observation aligns with our prediction of disordered regions (Figure 1D).

UL11 is a protein known to undergo LLPS.[21] FuzDrop predicted the protein to have a high LLPS potential with a score of 0.78. However, PSPredictor and catGranule predicted no LLPS potential, with a score of 0.16 and −0.56, respectively (Figure 1E and Supporting Information: Table S3). Notably, FuzDrop also predicted one droplet-promoting region spanning residues 63–96. In terms of MoRF, both DISOPRED3 and MoRFchibi SYSTEM predicted the presence of three MoRF regions each, whereas MoRFpred predicted two MoRF regions. The specific details of the predicted MoRF regions are listed in Table 4. Furthermore, DisoRPDbind and DRNApred provided insights into UL11's binding properties. DisoRPDbind predicted the protein to contain one RNA-binding residue and no DNA-binding residues, whereas DRNApred predicted the absence of both RNA- and DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, UL11 was found to interact with four other HSV-1 proteins, namely

TABLE 2 PPID of the selected HSV-1 proteins.

S. no	Protein	$PPID_{IUPred3\ long}$	$PPID_{IUPred3\ short}$	$PPID_{ANCHOR2}$	$PPID_{PONDR\ VLXT}$	$PPID_{PONDR\ VSL2}$	$PPID_{PSIPRED}$	$PPID_{DISOPRED3}$	$PPID_{Mean}$
1	Myristylated tegument protein (UL11)	66.67	59.375	69.79	71.88	63.54	82.29	44.79	64.58
2	Transcriptional regulator ICP4 (RS1)	52.00	43.84	50.69	60.09	75.42	60.86	45.22	56.16
3	Capsid maturation protease (UL26)	54.02	37.00	44.41	53.86	63.94	64.72	45.67	53.07
4	Capsid scaffold protein (ICP35/UL26.5)	88.75	56.53	75.08	71.43	91.49	76.60	62.31	81.76
5	Envelope gC (UL44)	44.42	28.96	22.50	51.47	40.31	57.53	31.90	37.18
6	Large tegument protein (UL36)	36.00	31.73	31.35	51.04	52.88	–	–	41.10
7	Multifunctional expression regulator (ICP27/UL54)	57.23	50.78	57.03	48.05	57.03	61.52	48.05	48.24
8	Tegument protein (US11)	100.00	100.00	72.67	93.17	100.00	98.14	85.71	100.00
9	Tegument protein VP11/12 (UL46)	35.52	32.72	36.07	39.14	46.80	52.51	32.45	39.55
10	Ubiquitin E3 ligase ICP0 (RL2)	76.26	65.55	67.48	76.65	74.84	82.32	65.68	71.74

Note: PPIDs are obtained from the seven disorder predictors used. PPIDMean represents the mean PPID calculated from the result of all the predictors utilized.

Abbreviations: gC, glycoprotein C; HSV-1, herpes simplex virus-1; ICP4, infected cell protein 4; PPID, predicted percentile of intrinsic disorder.

TABLE 3 Disordered regions in the selected HSV-1 proteins under study.

Protein	Disordered regions (stretch of >30 amino acid residues)
Myristylated tegument protein (UL11)	40–96
Transcriptional regulator ICP4 (RS1)	1–295, 311–342, 486–518, 590–621, 700–829, 1213–1251, 1262–1298
Capsid maturation protease (UL26)	259–325, 339–378, 406–481, 492–617
Capsid scaffold protein (ICP35/UL26.5)	33–80, 102–174, 187–312
Envelope glycoprotein C (gC/UL44)	19–132
Large tegument protein (UL36)	260–507, 743–786, 945–1036, 1242–1331, 1910–1944, 2269–2306, 2491–3076
Multifunctional expression regulator (ICP27/UL54)	1–247
Tegument protein (US11)	1–161
Tegument protein VP11/12 (UL46)	429–614, 659–699
Ubiquitin E3 ligase ICP0 (RL2)	1–109, 223–636

Abbreviations: gC, glycoprotein C; HSV-1, herpes simplex virus-1; ICP4, infected cell protein 4.

glycoprotein N, responsible for mediating viral fusion,[73] and glycoprotein E, UL21 and UL16, all of which play roles in linking capsids to membranes[74] (Figure 1F).

3.2 | Transcriptional regulator ICP4 (RS1)

The transcriptional regulator ICP4 (RS1) in HSV-1 plays a crucial role in the regulation of viral transcription by interacting with components of the cellular RNA polymerase II (RNA pol II) machinery (Table 1). ICP4 is a 1298 amino acid residue protein with a molecular weight of ~133 kDa and a theoretical pI of 6.12. ICP4 encompasses 24.73% of disorder-promoting amino acid residues (Figure 2A), with a $PPID_{Mean}$ of 56.16% (Table 2 and Figure 2C). An analysis of the compositional profiling analysis of the protein revealed its enrichment in the disorder-promoting amino acid Pro, whereas it exhibited a deficiency in the order-promoting amino acids Ile, Met and Asn (Figure 2B). Detailed information about the amino acid composition of the protein is shown in Figure 2A. The mean disorder propensity analysis of the protein showed that ICP4 contains seven disordered regions, each comprising >30 amino acid residues. These regions are located at amino acid positions 1–295, 311–342, 486–518, 590–621, 700–829, 1213–1251 and 1262–1298, with the longest stretch spanning 295 amino acid residues (Table 3).

Upon examining the ICP4 structure using AlphaFold, it became evident that the per-residue pLDDT scores for most residues exceeded 90, indicating a high level of model accuracy (Supporting Information: Figure S1B). However, the N-terminus (up to ~300), residues around 500, those spanning 700–800 and the C-terminus (from ~1200) displayed pLDDT scores ranging from 20 to 60, suggesting lower confidence and a less favourable backbone configuration. This observation aligns with our prediction of disordered regions (Figure 2D).

The protein exhibited LLPS potential according to all three predictors, with a PSPredictor score of 0.57, a catGranule score of 1.13 and a FuzDrop score of 0.99 (Figure 2D and Supporting Information: Table S3). FuzDrop also predicted nine droplet-promoting regions spanning from 1–288, 308–342, 478–623, 663–680, 700–914, 932–962, 1137–1159, 1215– 1238 and 1282–1298. In terms of MoRF, DISOPRED3 predicted the presence of two MoRF regions, MoRFchibi SYSTEM predicted nine MoRF regions and MoRFpred predicted five MoRF regions. The details of the predicted MoRF regions are listed in Table 4. Additionally, DisoRPDbind predicted the presence of multiple RNA-binding and DNA-binding residues in the protein, whereas DRNApred predicted several RNA-binding but no DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, ICP4 was cur other HSV-1 proteins: gC, UL4, which colocalizes with isoforms of ICP22[75], ICP27 and US3, a Ser/Thr kinase. (Figure 2F). Furthermore, the protein exhibited potential interactions with six host proteins, namely, T protein (T-box protein), a transcription factor, early growth response-1 (EGR1), which plays a role in vascular dysfunction and disease, TATA-Box-binding protein (TBP), TAF1, transcription factor SP1 and general transcription factor II B (GTF2B).

TABLE 4 Prediction of MoRF regions in the selected HSV-1 proteins using the DISOPRED3, MoRFchibi SYSTEM and MoRFpred predictors.

	MoRF predictors					
	DISOPRED3		MoRFchibi SYSTEM		MoRFpred	
Protein	No. of MoRFs	Position	No. of MoRFs	Position	No. of MoRFs	Position
Myristylated tegument protein (UL11)	3	1–6, 10–82, 89–96	3	13–19, 28–79, 87–95	2	11–15, 88–94
Transcriptional regulator ICP4 (RS1)	2	1148–1158, 1293–1298	9	79–100, 108–125, 162–174, 209–220, 249–253, 262–275, 301–312, 1210–1214, 1287–1297	5	622–632, 693–697, 931–937, 1204–1210, 1278–1283
Capsid maturation protease (UL26)	3	1–6, 8–14, 378–382	1	21–26	4	22–27, 119–125, 328–333, 627–633
Capsid scaffold protein (ICP35/UL26.5)	2	1–13, 119–124	2	18–30, 140–147	4	21–28, 93–97, 138–147, 321–327
Envelope gC (UL44)	1	475–511	2	154–161, 505–510	3	10–18, 71–75, 155–160
Large tegument protein (UL36)	NA	NA	0	0	13	22–27, 866–874, 1297–1303, 2490–2496, 2546–2554, 2618–2626, 2660–2664, 2689–2697, 2713–2717, 2778–2782, 2838–2847, 2875–2885, 3132–3137
Multifunctional expression regulator (ICP27/UL54)	0	0	3	1–37, 105–110, 123–129	2	7–15, 505–512
Tegument protein (US11)	1	1–12	3	16–24, 26–31, 94–98	1	153–161
Tegument protein VP11/12 (UL46)	2	1–23, 706–718	7	6–15, 516–524, 544–569, 572–585, 649–662, 683–695, 697–713	0	0
Ubiquitin E3 ligase ICP0 (RL2)	4	1–15, 174–181, 184–188, 762–775	9	1–47, 63–73, 310–320, 359–364, 434–438, 453–459, 488–508, 521–525, 619–624	2	23–33, 69–75

Abbreviations: gC, glycoprotein C; HSV-1, herpes simplex virus-1; ICP4, infected cell protein 4; MoRF, molecular recognition feature; NA, not applicable.

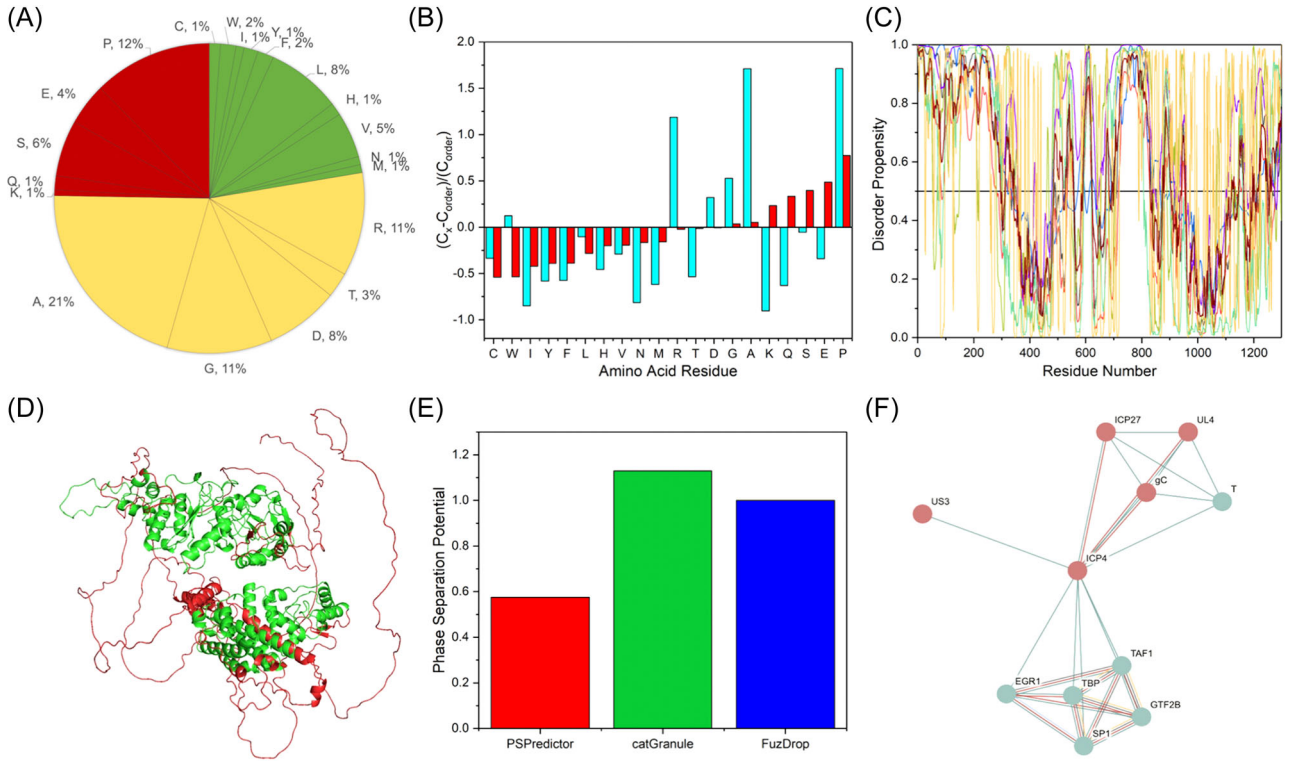

FIGURE 2 Analysis of infected cell protein 4 (ICP4). (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED, and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of ICP4. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

3.3 | Capsid maturation protease (UL26)

The capsid maturation protease (UL26) facilitates viral capsid assembly and maturation (Table 1). It is a 635 amino acid residue protein with a molecular weight of ~66 kDa and a theoretical pI of 6.21. UL26 functions as a protease self-cleaving at two specific positions to generate two capsid proteins, VP21 and VP24.[76] Additionally, it also cleaves the *UL26.5* gene product, resulting in the formation of VP22a.[76,77] UL26 has 28.66% disorder-promoting amino acid residues (Figure 3A) and a $PPID_{Mean}$ of 53.07% (Table 2 and Figure 3C). According to the compositional profiling analysis of the protein, it was found to be enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Trp, Ile, Phe and Asn (Figure 3B). Detailed information about the amino acid composition of the protein is shown in Figure 3A. The mean disorder propensity analysis of the protein showed that it contained four disordered regions with >30 amino acid residues. These regions are located at amino acid positions 259–25, 339–378, 406–481 and 492–617, with the longest stretch spanning 126 amino acid residues (Table 3).

Upon examining UL26 using AlphaFold, it was observed that the per-residue pLDDT scores for residues from 1 to 200 exceeded 90, indicating a high level of model accuracy (Supporting Information: Figure S1C). In contrast, the remaining residues displayed pLDDT scores between 20 and 40, with short stretches around 400 and 500 having pLDDT scores of ~80. These observations align with our prediction of disordered regions (Figure 3D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.63, a catGranule score of 0.35 and a FuzDrop score of 0.99 (Figure 3E and Supporting Information: Table S3). As per the MoRF analyses, DISOPRED3 predicted the protein to have three MoRF regions, MoRFchibi SYSTEM predicted one MoRF region and MoRFpred predicted four MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to have several RNA-binding residues and no DNA-binding residues, whereas DRNApred predicted the protein to have one RNA-binding residue and several DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

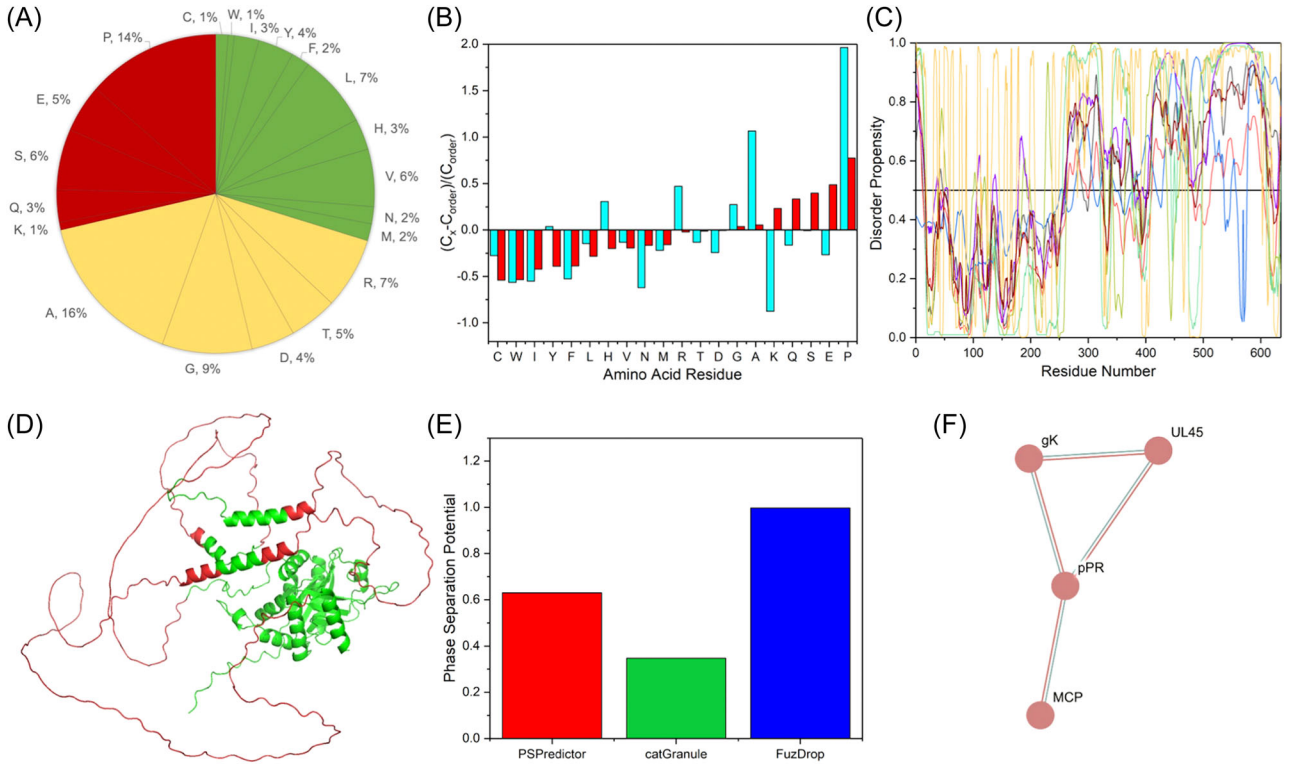

FIGURE 3 Analysis of UL26. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of UL26. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

Based on the STRING analysis, UL26 was found to interact with three other HSV-1 proteins, namely glycoprotein K (gK), responsible for mediating membrane fusion;[78] the membrane protein UL45, involved in fusion events;[79] and the major capsid protein (MCP) (Figure 3F).

3.4 | Capsid scaffold protein (ICP35)

The capsid scaffold protein (UL26.5 or ICP35) is a 329 amino acid residue assembly protein with a molecular weight of ~34 kDa and a theoretical pI of 6.13 (Table 1). The ICP35 ORF shares a frame with and overlaps the carboxy-terminal portion of the UL26 ORF.[77] The ICP35 protein is not generated through proteolytic cleavage of the UL26 polypeptide but is synthesized independently, as it possesses its own promoter within the UL26 ORF.[77] As every amino acid sequence found in the *UL26.5* gene product is also present within the UL26 polypeptide, the UL26 protease identifies the identical cleavage site near the carboxy terminus in both proteins. Furthermore, ICP35 has 33.74% disorder-promoting amino acid residues (Figure 4A) and a $PPID_{Mean}$ of 81.76% (Table 2 and Figure 4C). According to the compositional profiling analysis of the protein, it was found to be significantly enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Trp, Ile, Phe, Leu and Asn (Figure 4B). Detailed information about the amino acid composition of the protein is shown in Figure 4A. The mean disorder propensity analysis of the protein showed that it contained three disordered regions with >30 amino acid residues. These regions are located at amino acid positions 33–80, 102–174 and 187–312, with the longest stretch spanning 126 amino acid residues (Table 3).

The inspection of ICP35 using AlphaFold revealed that the per-residue pLDDT scores for the residues around 70 to 100 and 160 to 190 exceeded 90, indicating a high level of model accuracy (Supporting Information: Figure S1D). The N-terminus (up to ~50) and C-terminus (from ~300) exhibited a moderate backbone configuration. However, the remaining residues had pLDDT scores close to 40. The modelled disordered regions align with the sequence prediction of disordered regions (Figure 4D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.87, a

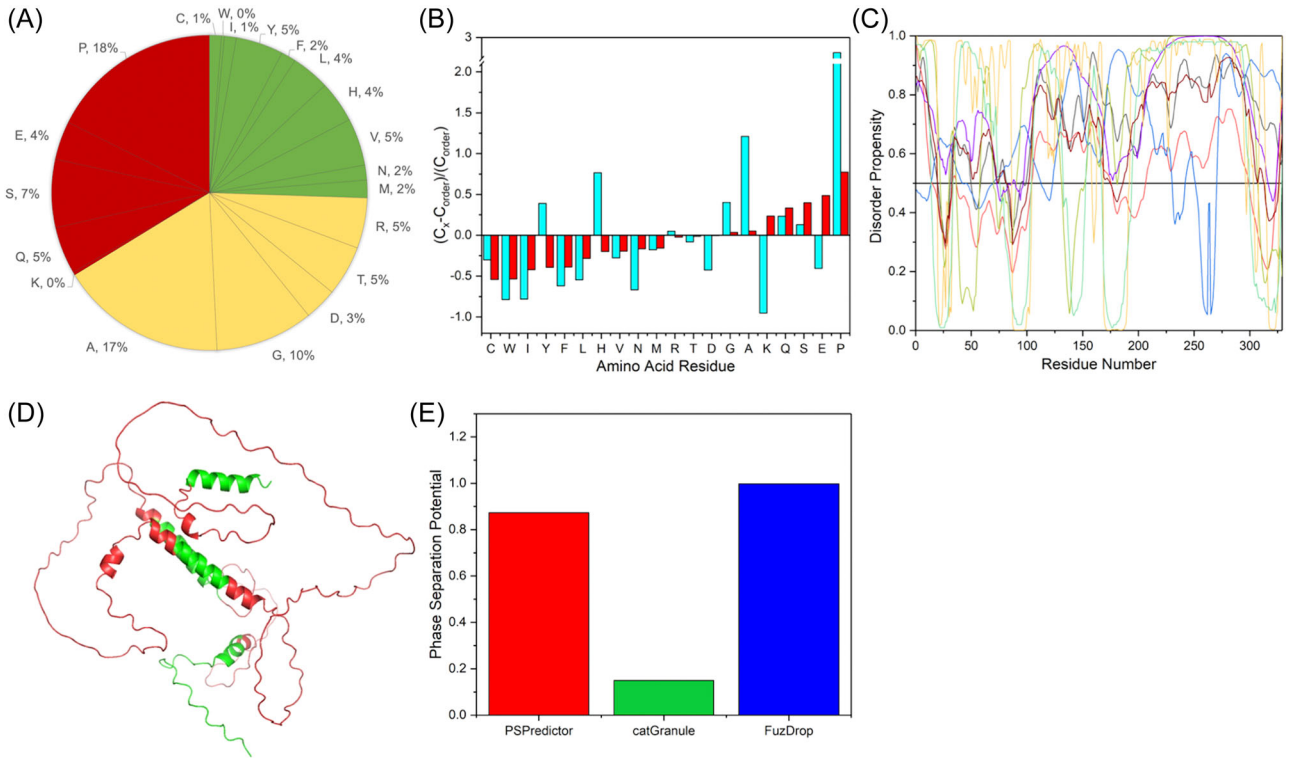

FIGURE 4 Analysis of ICP35. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED, and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of ICP35. The colour coding is the same as Figure 1.

catGranule score of 0.15 and a FuzDrop score of 0.99 (Figure 4E and Supporting Information: Table S3). As per the MoRF analyses, DISOPRED3 and MoRFchibi SYSTEM predicted the protein to have two MoRF regions and MoRFpred predicted four MoRF regions. The details of the predicted MoRF regions are listed in Table 3. DisoRPDbind predicted the protein to have a few RNA-binding residues and no DNA-binding residues, whereas DRNApred predicted the protein to have three RNA-binding residues and three DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

There was no data available in the STRING database for ICP35.

3.5 | Envelope gC (UL44)

The envelope gC (UL44) is a 511 amino acid residue protein with a molecular weight of ~55 kDa and a theoretical pI of 7.66. It is a vital component of HSV-1 with a multifaceted role (Table 1). gC has 29.35% disorder-promoting amino acid residues (Figure 5A) and a $PPID_{Mean}$ of 37.18% (Table 2 and Figure 5C). According to the compositional profiling analysis of the protein, it was found to be significantly enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Ile and Asn (Figure 5B). Detailed information about the amino acid composition of the protein is shown in Figure 5A. The mean disorder propensity analysis of the protein showed that it contained one disordered region with >30 amino acid residues. This region is located at amino acid position 19–132, spanning 114 residues (Table 3).

The AlphaFold structure prediction of gC revealed that the per-residue Plddt scores for most residues exceeded 90, indicating a high level of model accuracy, except for those in the N-terminus (up to ~100) (Supporting Information: Figure S1E). The residues within the N-terminus (up to ~100) displayed pLDDT scores ranging from 20 to 60, signifying lower confidence and a less favourable backbone configuration. This observation aligns with our prediction of disordered regions (Figure 5D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.62, a catGranule score of 0.05, and a FuzDrop score of 0.78 (Figure 5E and Supporting Information: Table S3). FuzDrop also predicted three droplet-promoting regions: 1–126, 312–324, and 446–470. As per the

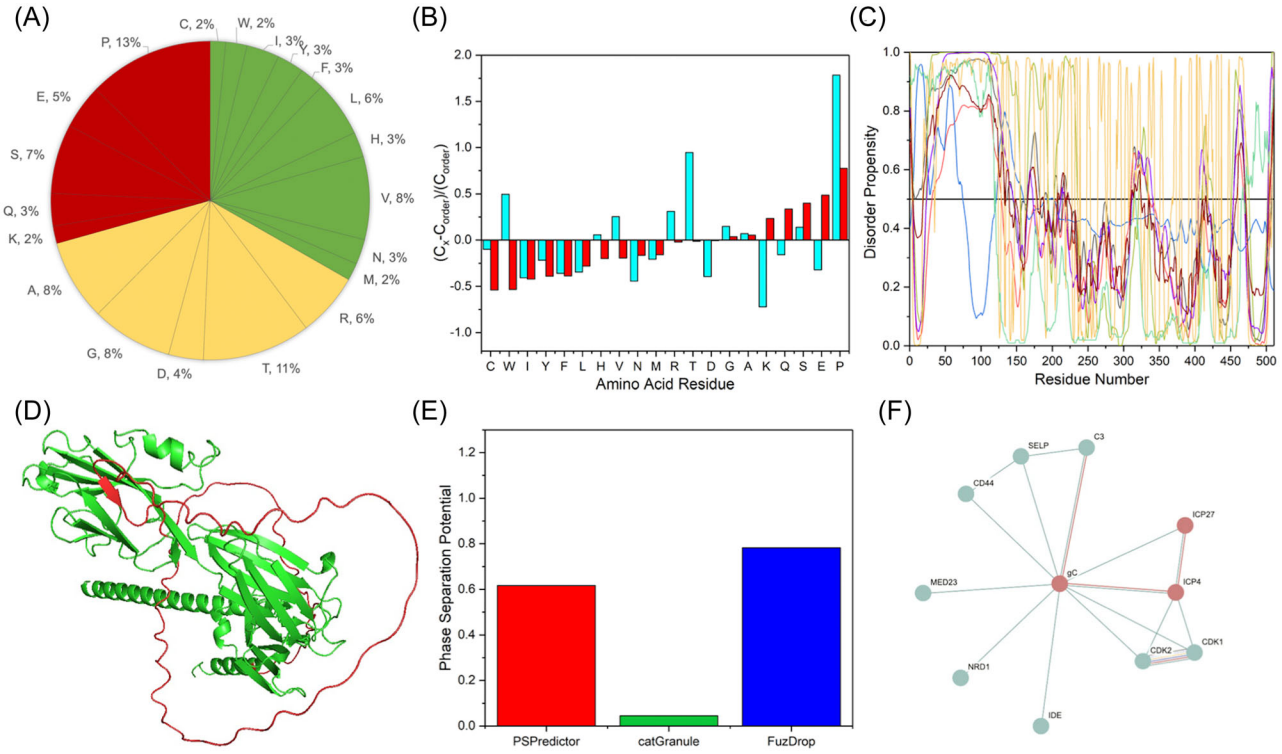

FIGURE 5 Analysis of glycoprotein C (gC). (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of gC. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

MoRF analyses, DISOPRED3 predicted the protein to have one MoRF region, MoRFchibi SYSTEM predicted two MoRF regions, and MoRFpred predicted three MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to have a few RNA-binding residues and two DNA-binding residues, whereas DRNApred predicted the protein to have a few RNA-binding residues and no DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, gC was found to interact with two other HSV-1 proteins, namely, ICP27 and ICP4 (Figure 5F). The protein was also found to interact with eight host proteins, namely, complement component 3 (C3), SELP, which is a Ca^{2+}-dependent receptor that mediates the interaction of activated endothelial cells or platelets with leukocyte, CD44 involved in cell adhesion,[80] mediator complex subunit 23 (MED23), an antiviral component, NRD1 or nardilysin that cleaves peptide substrates on the N-terminus of arginine residues, IDE, an insulin-degrading enzyme, and cyclin-dependent kinases CDK1 and CDK2.

3.6 | Large tegument protein (UL36)

The large tegument protein (UL36), also known as VP1/2, is a 3139 amino acid residue protein with a molecular weight of ~333 kDa and a theoretical pI of 5.84. Studies have demonstrated its role in DNA packaging and transport within HSV-1. Specifically, the absence of the UL36-encoded polypeptide results in the accumulation of numerous cytosolic capsids containing DNA, which fail to mature into enveloped viruses[81] (Table 1). UL36 has 28.89% disorder-promoting amino acid residues (Figure 6A) and a $PPID_{Mean}$ of 41.10% (Table 2 and Figure 6C). According to the compositional profiling analysis of the protein, it was found to be enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Cys, Ile, Phe and Asn (Figure 6B). Detailed information about the amino acid composition of the protein is shown in Figure 6A. The mean disorder propensity analysis of the protein showed that it contained seven disordered regions with >30 amino acid residues. These regions are located at amino acid positions 260–507, 743–786, 945–1036, 1242–1331, 1910–1944, 2269–2306, and 2491–3076, with the longest stretch spanning 586 amino acid residues (Table 3).

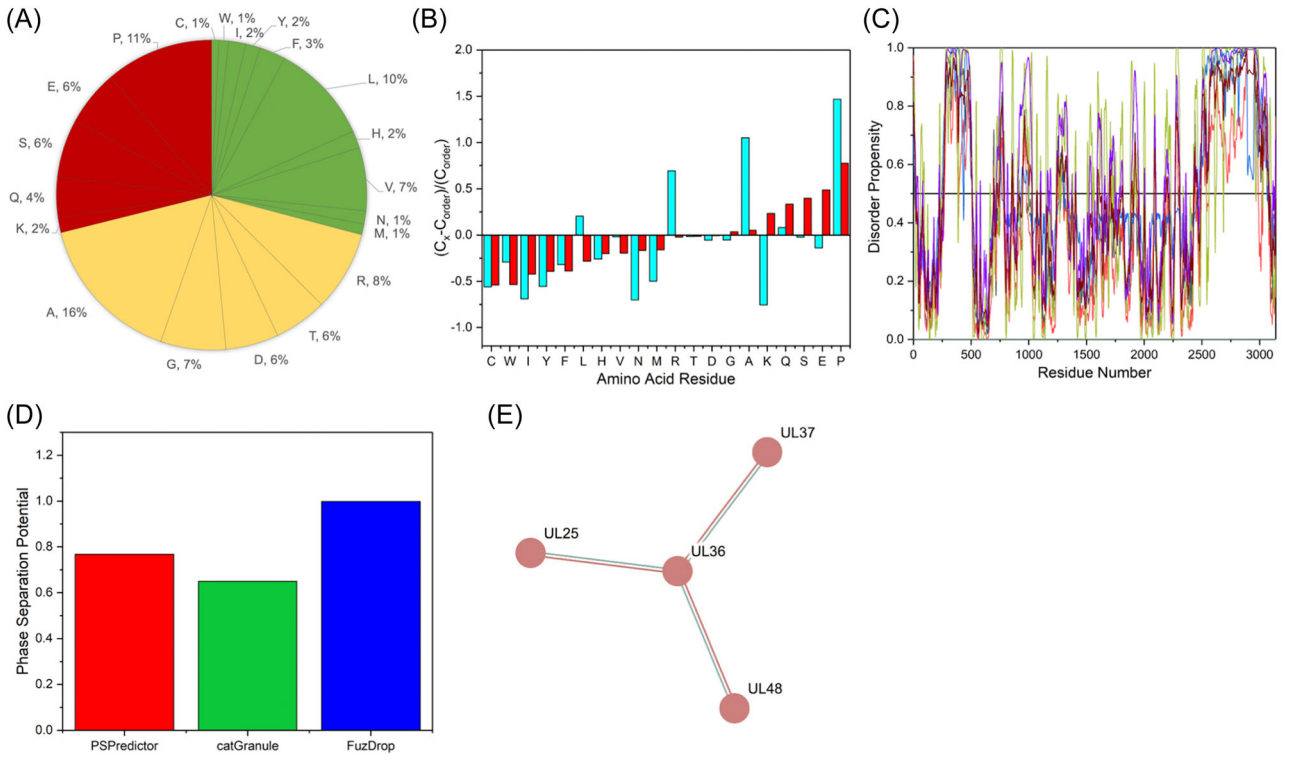

FIGURE 6 Analysis of UL36. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) Liquid–liquid phase separation propensities of UL36. (E) Virus.STRING analysis. The colour coding is the same as Figure 1.

The protein could not be modelled using AlphaFold due to the large sequence length of UL36 and the limited availability of RAM in our computing system.

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.77, a catGranule score of 0.65, and a FuzDrop score of 0.99 (Figure 6D and Supporting Information: Table S3). FuzDrop also predicted 17 droplet-promoting regions: 1–14, 258–500, 683–704, 732–778, 937–970, 1058–1070, 1265–1288, 1399–1410, 1693–1714, 1748–1759, 1907–1960, 2074–2096, 2191–2212, 2259–2297, 2486–2546, 2551–3016 and 3024–3069. As per the MoRF analyses, MoRFchibi SYSTEM predicted the protein to have no MoRF regions, and MoRFpred predicted 13 MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to several RNA-binding and DNA-binding residues, whereas DRNApred predicted the protein to have no RNA-binding and several DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, UL36 was found to interact with three other HSV-1 proteins, namely, UL25 and UL37, essential for viral growth and assembly[82,83] and UL48, crucial for replication and infection of HSV-1[84] (Figure 6E).

3.7 | Multifunctional expression regulator (ICP27)

The multifunctional expression regulator (ICP27 or UL54) is a 512 amino acid residue protein with a molecular weight of ~55 kDa and a theoretical pI of 6.48. It plays crucial roles in various stages of viral mRNA biogenesis, including transcription, RNA processing, nuclear export and translation[85] (Table 1). ICP27 is characterized by the presence of 28.52% disorder-promoting amino acid residues (Figure 7A), with a $PPID_{Mean}$ of 48.24% (Table 2 and Figure 7C). According to the compositional profiling analysis of the protein, it was found to be enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Tyr and Asn (Figure 7B). Detailed information about the amino acid composition of the protein is shown in Figure 7A. The mean disorder propensity analysis of the protein showed that it contained one disordered region with >30 amino acid residues. This region is located at the amino acid positions 1–247, spanning 247 amino acid residues (Table 3).

Upon inspection of ICP27 using AlphaFold, it was observed that the per-residue pLDDT scores for the residues of the C-terminal half of the protein (from ~250) were consistently >90, indicating a high level of model

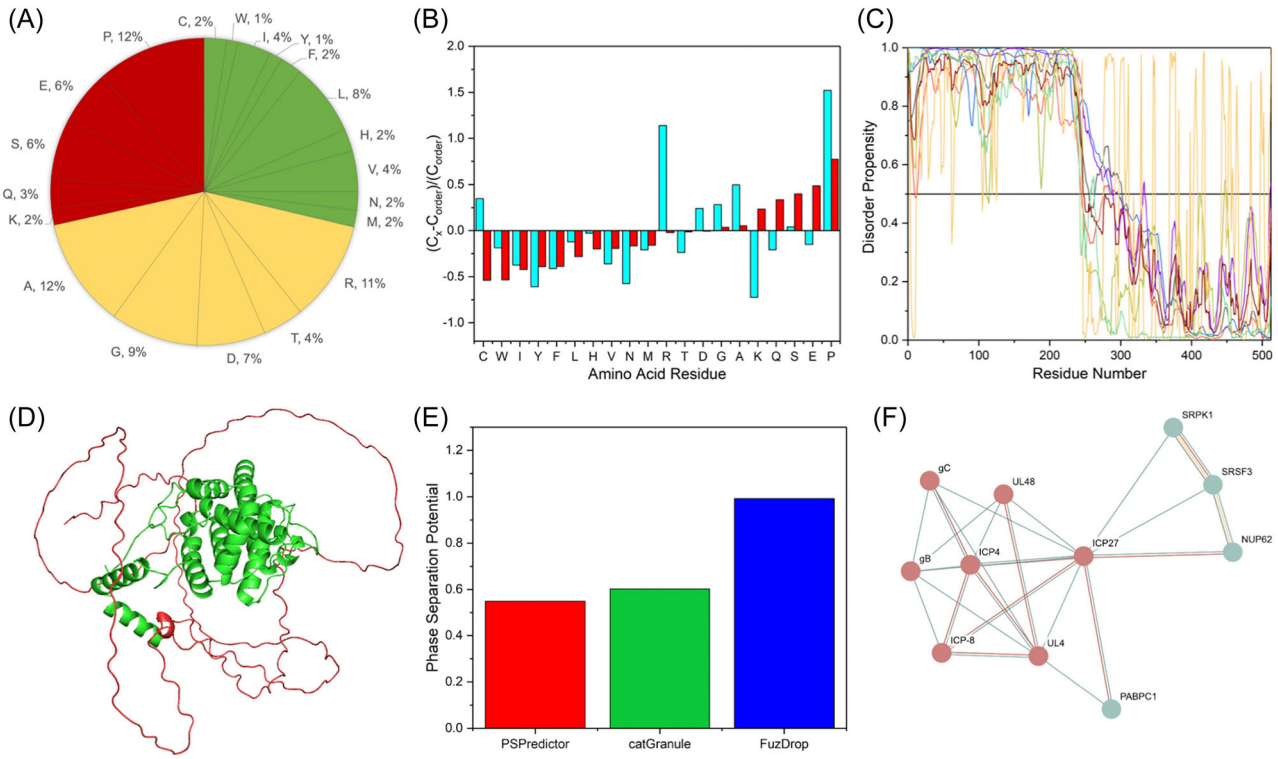

FIGURE 7 Analysis of ICP27. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of ICP27. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

accuracy (Supporting Information: Figure S2A,). Conversely, the residues of the N-terminal half of the protein (up to ~250) exhibited pLDDT scores ranging from 20 to 60, signifying lower confidence and a less favourable backbone configuration. This observation aligns with our prediction of disordered regions (Figure 7D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.55, a catGranule score of 0.60 and a FuzDrop score of 0.99 (Figure 7E and Supporting Information: Table S3). FuzDrop also predicted two droplet-promoting regions: 1–247 and 263–305. As per the MoRF analyses, DISOPRED3 predicted the protein to have zero MoRF regions, MoRFchibi SYSTEM predicted three MoRF regions, and MoRFpred predicted two MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to have several RNA-binding and three DNA-binding residues, whereas DRNApred predicted the protein to have several RNA-binding and a few DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, ICP27 was found to interact with six other HSV-1 proteins, namely UL48, gC, gB, ICP8, a single-stranded DNA-binding protein, UL4 and ICP4 (Figure 7F). The protein was also found to interact with four host proteins, namely, SRPK1 or SRSF protein kinase 1, SRSF3 or Serine/arginine-rich splicing factor 3, NUP62, an important disordered protein participating in the permeability barrier of the nuclear pore complex, and PABPC1 (poly-A binding protein cytoplasmic 1), an RNA-binding protein.[86]

3.8 | Tegument protein (US11)

The tegument protein (US11) is a 161 amino acid residue protein with a molecular weight of ~17.8 kDa and a theoretical pI of 11.74. Its primary function is to act as an antagonist of the IFN pathway, thereby inhibiting the shutdown of protein translation and facilitating virus production[87] (Table 1). US11 has 40.37% disorder-promoting amino acid residues (Figure 8A) and a $PPID_{Mean}$ of 100% (Table 2 and Figure 8C). According to the compositional profiling analysis of the protein, it was found to be significantly enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Cys, Trp, Ile, Phe, Tyr, Leu and

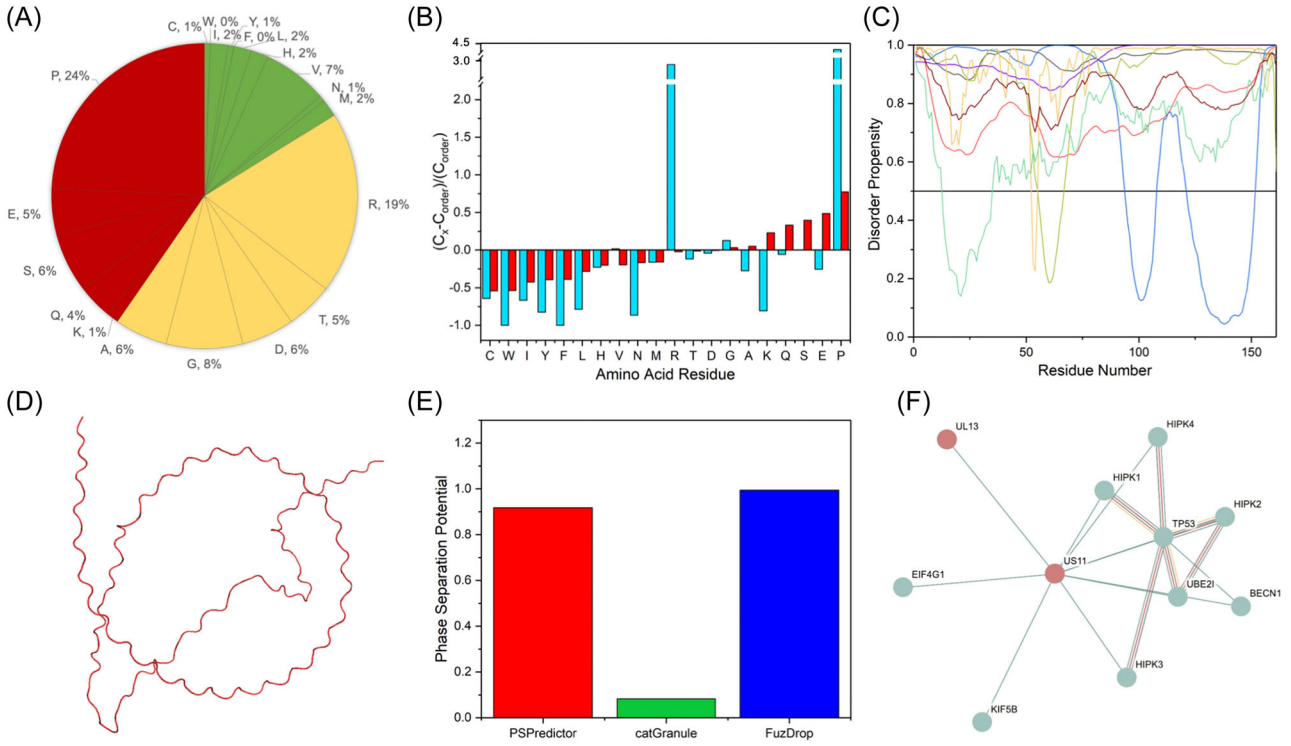

FIGURE 8 Analysis of US11. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of US11. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

Asn (Figure 8B). Detailed information about the amino acid composition of the protein is shown in Figure 8A. The mean disorder propensity analysis of the protein showed that it contained one disordered region with >30 amino acid residues. This region is located at amino acid position 1–161, spanning 161 amino acid residues (Table 3).

The structure prediction of US11 using AlphaFold revealed that the per-residue pLDDT scores for all residues of the protein fell within the range of 20–80, indicating lower confidence and a less favourable backbone configuration (Supporting Information: Figure S2B). This observation aligns with our prediction of disordered regions (Figure 8D).

The protein was predicted to possess LLPS potential by all three predictors, with a PSPredictor score of 0.92, a catGranule score of 0.08 and a FuzDrop score of 0.99 (Figure 8E and Supporting Information: Table S3). Interestingly, FuzDrop predicted the entire length of the protein to be a droplet-promoting region. As per the MoRF analyses, DISOPRED3 and MoRFpred predicted the protein to have one MoRF region, and MoRFchibi SYSTEM predicted three MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to a few RNA-binding and DNA-binding residues, whereas DRNApred predicted the protein to have no RNA-binding and DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, US11 was found to interact with one other HSV-1 protein, namely, the protein kinase UL13[88] (Figure 8F). The protein was also found to interact with nine host proteins, namely eukaryotic translation initiation factor 4γ (EIF4G1), Kinesin family member 5B (KIF5B), a microtubule-dependent motor, ubiquitin-conjugating enzyme E2I (UBE21), Beclin1 (BECN1) involved in autophagy, tumour protein p53 (TP53) and homeodomain interacting protein kinase 1, 2, 3 and 4 (HIPK1, HIPK2, HIPK3 and HIPK4), which are Ser/Thr protein kinases.

3.9 | Tegument protein VP11/12 (UL46)

The tegument protein VP11/12 (UL46) is a 718 amino acid residue protein with a molecular weight of ~78 kDa and a theoretical pI of 9.14. Its role involves the

stimulation of signal transduction in the PI3-Akt pathway during infection[67] (Table 1). VP11/12 is characterized by the presence of 24.93% disorder-promoting amino acid residues (Figure 9A), with a $PPID_{Mean}$ of 39.55% (Table 2 and Figure 9C). According to the compositional profiling analysis of the protein, it was found to be enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Ile and Phe (Figure 9B). Detailed information about the amino acid composition of the protein is shown in Figure 9A. The mean disorder propensity analysis of the protein showed that it contained two disordered regions with >30 amino acid residues. These regions are located at amino acid positions 429–614 and 659–699, with the longest stretch spanning 186 amino acid residues (Table 3).

The analysis of the VP11/12 structure using AlphaFold showed that the per-residue pLDDT scores for the residues of the N-terminal half of the protein (up to ~400) were consistently >90, indicating a high level of model accuracy (Supporting Information: Figure S2C). Conversely, the residues of the C-terminal half of the protein (from ~400) exhibited pLDDT scores ranging from 20 to 60, signifying lower confidence and a less favourable backbone configuration. This observation aligns with our prediction of disordered regions (Figure 9D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.83, a catGranule score of 0.23 and a FuzDrop score of 0.94 (Figure 9E and Supporting Information: Table S3). FuzDrop also predicted six droplet-promoting regions: 140–152, 256–286, 421–647, 553–576, 593–609 and 661–683. As per the MoRF analyses, DISOPRED3 predicted the protein to have two MoRF regions, MoRFchibi SYSTEM predicted seven MoRF regions and MoRFpred predicted zero MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to have several RNA-binding and no DNA-binding residues, whereas DRNApred predicted the protein to have no RNA-binding and several DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, VP11/12 was found to interact with seven other HSV-1 proteins, namely, US3, UL33, involved in viral growth and DNA packaging[89], gK, UL45, involved in fusion events,[78] ICP0, UL48 and structural protein[90] UL47 (Figure 9F).

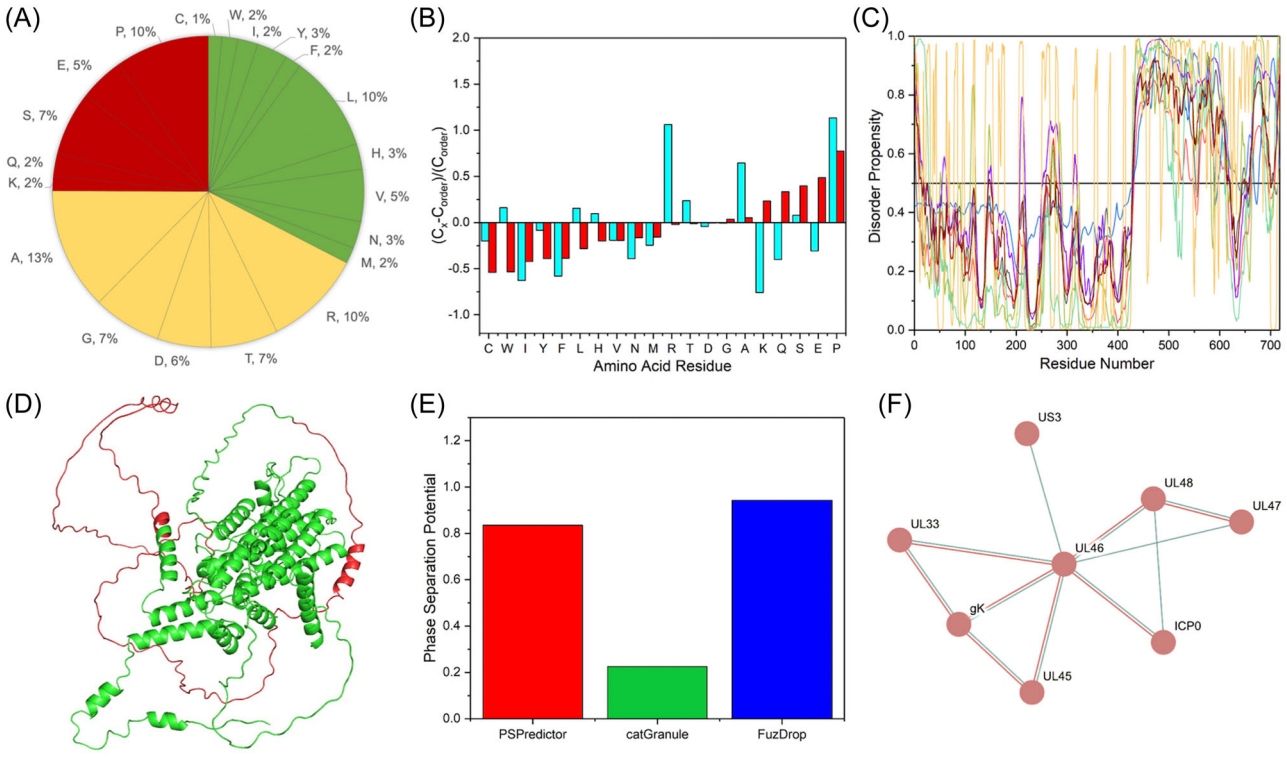

FIGURE 9 Analysis of VP11/12. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of VP11/12. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

3.10 | Ubiquitin E3 ligase ICP0 (RL2)

The ubiquitin E3 ligase (ICP0 or RL2) is a 775 amino acid residue protein with a molecular weight of ~78.5 kDa and a theoretical pI of 6.95. ICP0 performs diverse functions, including triggering transcription activation and chromatin remodelling, evading antiviral responses, influencing the cell cycle, disrupting DNA damage responses, and impacting endocytosis processes[10] (Table 1). ICP0 has 31.61% disorder-promoting amino acid residues (Figure 10A) and a $PPID_{Mean}$ of 71.74% (Table 2 and Figure 10C). According to the compositional profiling analysis of the protein, it was found to be significantly enriched in the disorder-promoting amino acid Pro and deficient in the order-promoting amino acids Ile, Tyr and Phe (Figure 10B). Detailed information about the amino acid composition of the protein is shown in Figure 10A. The mean disorder propensity analysis of the protein showed that it contained two disordered regions with >30 amino acid residues. These regions are located at amino acids 1–109 and 223–636, with the longest stretch spanning 414 residues (Table 3).

The analysis of the ICP0 structure using AlphaFold showed that the per-residue pLDDT scores for residues between ~100-200 were consistently >90, indicating a high level of model accuracy (Supporting Information: Figure S2D). All other residues exhibited pLDDT scores ranging from 20 to 60, signifying lower confidence and a less favourable backbone configuration. This finding aligns with our prediction of disordered regions (Figure 10D).

The protein was predicted to have LLPS potential by all three predictors, with a PSPredictor score of 0.93, a catGranule score of 1.24 and a FuzDrop score of 0.99 (Figure 10E and Supporting Information: Table S3). FuzDrop also predicted three droplet-promoting regions: 1–117, 212–639 and 757–775. As per the MoRF analyses, DISOPRED3 predicted the protein to have four MoRF regions, MoRFchibi SYSTEM predicted nine MoRF regions and MoRFpred predicted two MoRF regions. The details of the predicted MoRF regions are listed in Table 4. DisoRPDbind predicted the protein to have several RNA-binding and DNA-binding residues, whereas DRNApred predicted the protein to have several RNA-binding and no DNA-binding residues. The RNA- and DNA-binding residues are enlisted in Supporting Information: Tables S4 and S5, respectively.

Based on the STRING analysis, ICP0 was found to interact with four other HSV-1 proteins, namely,

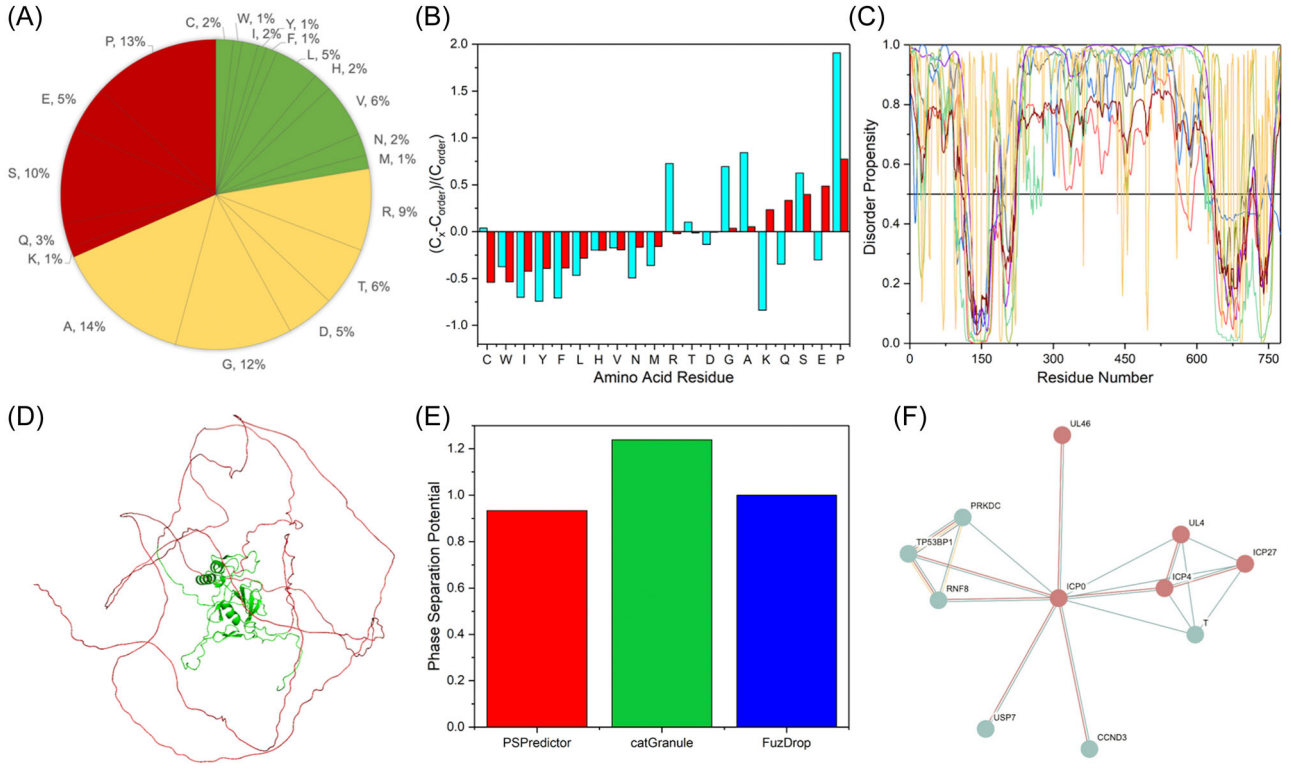

FIGURE 10 Analysis of ICP0. (A) Amino acid composition. (B) Compositional analysis. Positive values indicate an abundance of amino acid residues, whereas negative values show depleted amino acid residues. (C) Predicted disorder predisposition. The disorder profiles were generated using IUPred3 (long and short), ANCHOR2, PONDR (VLXT and VSL2), PSIPRED and DISOPRED3. (D) AlphaFold-based structural modelling. (E) Liquid–liquid phase separation propensities of ICP0. (F) Virus.STRING analysis. The colour coding is the same as Figure 1.

VP11/12, UL4, ICP4 and ICP27 (Figure 10F). The protein was also found to interact with six host proteins, namely T protein, the Ser/Thr Protein kinase DNA-activated catalytic polypeptide (PRKDC), tumour protein p53-binding protein 1 (TP53BP1), which plays a vital role in response to DNA damage, ring finger protein 8 (RNF8), an E3 ubiquitin-protein ligase crucial in DNA damage signalling, human ubiquitin-specific protease 7 (USP7), an enzymatic protein that eliminates ubiquitin and polyubiquitin clusters, and cyclin D3 (CCND3).

4 | DISCUSSION

Biomolecular LLPS underlies the functional basis of many processes such as signal transduction, stress response, autophagy, DNA repair, RNA metabolism and so on.[91–94] Recently, studies have revealed that LLPS plays a vital role in viral growth and infection.[19,24,95,96] There is evidence that steps of viral lifestyle can be performed by viruses in MLOs built de novo[97,98] as well as interfere with pre-existing MLOs such as host stress granules and procession bodies, thus interfering with the signalling of the host's innate immunity.[99,100] In many viruses, such MLOs have been found to play a role in every step of their life cycle, including viral entry, genome replication, assembly and viral packaging.[27] Influenza A virus has been found to utilize LLPS for entry into host cells,[101] whereas other viruses employ LLPS to form replication factories after entry in cells,[24,95] genome packaging[102] and viral assembly.[103] The viral ribonucleoproteins undergo LLPS to form inclusions called Negri bodies, which are used as hubs for the rabies virus genome replication and transcription.[24] The nucleocapsid protein (N) of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus undergo LLPS and form condensates with other viral and host proteins for RNA genome assembly during packaging.[102,104] The HIV-1 nucleocapsid protein also undergoes LLPS, which is modulated by RNA-binding domains and ions such as zinc and copper.[103,105] MLOs are also modulated and regulated by interaction with RNA,[106–108] pH, salt concentration and various PTMs.[109]

Considering the significance of LLPS and the formation of MLOs in various steps of the viral life cycles, we explored to obtain detailed insights into the compositional profiles, disorder propensity, LLPS potential, MoRF analysis, nucleic acid-binding sites and PPI networks of the identified LLPS-prone proteins in HSV-1. Out of the 74 protein sequences available, we focussed our analysis on proteins that were predicted to be disordered and subsequently have the potential to undergo LLPS. The compositional profiling of the amino acids of the selected proteins depicted that the proteins revealed a significant fraction of disorder-promoting and neutral amino acids. The MoRF regions indicate residues that may participate in interactions with host factors. Our MoRF analysis identified multiple MoRFs within the selected proteins that could be involved in dynamic interactions with other binding partners. It is known that the interaction of MLOs with nucleic acids can influence the behaviour of MLOs and, hence, we also predicted nucleic acid-binding sites in the proteins. Using two predictors, we identified several RNA-binding amino acid residues in the proteins, while fewer residues were predicted to be involved in DNA binding. Experimental evidence has shown multiple instances of RNA interacting with MLOs in viruses, including HSV-1.[21]

We observed a high degree of consistency when we compared predicted disordered regions using sequence-based bioinformatics tools with AlphaFold-modelled structures. These disordered regions, highlighted in red, exhibited strong alignment across all the proteins discussed in our manuscript. This alignment underscores the reliability and robustness of both approaches in identifying disordered regions of different proteins. Although the sequence-based predictors are readily available, typically faster and computationally less intensive, making them suitable for high-throughput analysis of large protein data sets, they sometimes miss detailed structural information, which can limit their accuracy in capturing fine-grained disorder patterns. Moreover, poorly annotated or divergent sequences may reduce the sensitivity and accuracy of disordered predictions. On the other hand, structural methods like AlphaFold leverage deep learning and structural information, offering a higher level of structural detail and accuracy in predicting disordered regions. These methods not only provide disordered region predictions but also insights into the overall three-dimensional structure of proteins, enabling a more holistic understanding. However, it is worth noting that these methods can have high computational demands for large sequences and often rely on access to pretrained models, which may not cover all protein sequences. Therefore, it is reasonable to consider the convergence of sequence-based predictors and AlphaFold-based predictions (or similar structure-based predictors) in identifying disordered regions. Such an integrated approach combines the advantages of both methods and enhances the accuracy of disorder region identification.

The Viruses.STRING analysis tool was employed to explore interactions between HSV-1 and human proteins, providing insights into the intricate interplay between the virus and its host. The analysis unveiled numerous interactions between HSV-1 and human proteins,

highlighting extensive cross-talk between the viral and host cellular machinery. ICP4 plays a crucial role in regulating viral transcription during infection. As viral genes are transcribed by cellular RNA pol II, ICP4 interacts with various components of the RNA pol II machinery to effectively control viral gene expression. STRING analysis reveals the interaction of ICP4 with transcription factors such as EGR1,[110] TBP[111] and GTF2B,[111] providing additional evidence of its involvement in transcriptional regulation. gC, responsible for viral entry, was found to bind to complement component C3b, inhibiting complement-mediated host immunity.[112] The PPI network also revealed an interaction between gC and selectin P, an adhesion molecule expressed in the endothelial cells of HSV-1 infected cells.[113] ICP27 was found to interact with SRPK1 and Aly/REF.[114,115] These interactions were shown to be regulated by arginine methylation, influencing ICP27 export and its interaction with these cellular proteins. Additionally, ICP27 exhibited an interaction with Nup62, which impairs nucleocytoplasmic transport in the host.[62] The analysis also revealed interactions involving ICP0 with CCND3 (cyclin D3),[116] USP7[117] and RNF8.[118] These interactions highlight the involvement of ICP0 in cell cycle regulation, protein degradation and DNA damage signalling pathways. US11, a tegument protein involved in the translational arrest pathways of the host, has been found to interact with KIF5B,[64] EIF4G1[119] and HIPK3.[120] These interactions highlight the involvement of US11 in cellular processes such as intracellular transport, translation regulation and transcriptional control. Proteins such as IDE, NRD1 and T-box were found to show interactions with HSV-1 proteins during the STRING analysis; however, we did not find any literary evidence.

STRING analysis also revealed numerous interactions among HSV-1 proteins, highlighting their ability to form functional complexes and exploit synergistic interactions. These interactions play crucial roles in various stages of the viral life cycle, including viral entry, replication, assembly and evasion of host immune responses. Studies have also reported highly conserved interactions among HSV-1 proteins, such as the interaction between HSV-1 UL33 and the nuclear egress proteins UL31/UL34.[121] Despite the generally low sequence similarity observed between orthologous proteins, these interactions remained conserved, suggesting that the functional aspects of these interactions are more likely to be preserved over evolutionary time, emphasizing that the conservation of function may outweigh the conservation of sequence. The analyses also unveiled a diverse array of interactions between HSV-1 proteins and host proteins, shedding light on the functional roles of viral proteins during infection. These interactions spanned various cellular processes, including immune response, transcriptional regulation, protein degradation and intracellular transport. The understanding of these interactions enhances our knowledge of HSV-1 pathogenesis and holds the potential to guide the development of targeted antiviral strategies.

The understanding that LLPS plays a crucial role in virus infection and life cycle opens the possibility of targeting the phenomenon for antiviral therapy. Risso-Ballester et al.[122] and Wang et al.[123] achieved success in this field in combating viral infection. Risso-Ballester et al. discovered compounds that can inhibit the replication of the respiratory syncytial virus by disrupting and hardening viral condensates, forming inclusion bodies.[122] Wang et al.[123] identified a peptide that targets the dimerization domain of the nucleocapsid protein of SARS-CoV-2, which is crucial for the protein to undergo LLPS with RNA. They found that treatment with this peptide enhanced the innate antiviral response both in vitro and in vivo. HSV-1 is known to cause oral and genital herpes, encephalitis and ocular infections, and it also plays a role in the pathology of neurodegenerative Alzheimer's disease.[124] Elucidating the mechanism of LLPS and targeting it in HSV-1 proteins holds the potential to discover a cure for the diseases associated with HSV-1 infections. In conclusion, further analysis of LLPS and MLOs of HSV-1 proteins will provide valuable insights into the intricate molecular mechanisms underlying HSV-1 replication and assembly. Further exploration of these processes holds great potential for developing novel antiviral strategies and expanding our understanding of virus–host interactions.

AUTHOR CONTRIBUTIONS

Sushma Subedi: Methodology; formal analysis; investigation; data curation; validation; writing—original draft. **Niharika Nag**: Methodology; formal analysis; investigation; data curation; validation; writing—original draft. **Harish Shukla**: Formal analysis; writing—review. **Aditya K. Padhi**: Methodology; formal analysis; writing—review. **Timir Tripathi**: Conceptualization; supervision; writing—review & editing.

ACKNOWLEDGEMENTS

Timir Tripathi acknowledges the support of project grants from the Indian Council of Medical Research (ICMR) (Grant 52/06/2020/BIO/BMS) and Department of Health Research (DHR) (Grant R.11013/47/2021-GIA/HR), Government of India, India. Sushma Subedi would like to thank the DHR for providing Junior Research Fellowship and North-Eastern Hill University for providing Non-NET fellowship. Niharika Nag would like to thank DST-INSPIRE for providing fellowship.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in NCBI at https://www.ncbi.nlm.nih.gov, reference number Refseq ID: NC_001806.2. These data were derived from the following resources available in the public domain: Human herpesvirus 1 strain 17, complete genome, https://www.ncbi.nlm.nih.gov/nuccore/NC_001806.2/. All data supporting the findings of this study are available within the paper and its Supplementary Information.

ORCID

Sushma Subedi https://orcid.org/0000-0002-0824-0186
Niharika Nag http://orcid.org/0000-0003-1236-9289
Harish Shukla http://orcid.org/0000-0002-5919-6808
Aditya K. Padhi http://orcid.org/0000-0001-6732-2547
Timir Tripathi http://orcid.org/0000-0001-5559-289X

REFERENCES

1. Steven A, Spear P. Herpesvirus capsid assembly and envelopment. In: Viruses WC, Burnett R, Garcea R, eds. *Structural biology of viruses*, 1997:312-351.
2. Wildy P, Watson DH. Electron microscopic studies on the architecture of animal viruses. *Cold Spring Harb Symp Quant Biol*. 1962;27:25-47.
3. Jacob RJ, Roizman B. Anatomy of herpes simplex virus DNA VIII. Properties of the replicating DNA. *J Virol*. 1977;23(2):394-411.
4. Honess RW, Roizman B. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. *J Virol*. 1974;14(1):8-19.
5. Boutell C, Sadis S, Everett RD. Herpes simplex virus type 1 immediate-early protein ICP0 and its isolated RING finger domain act as ubiquitin E3 ligases in vitro. *J Virol*. 2002;76(2):841-850.
6. Weller SK, Coen DM. Herpes simplex viruses: mechanisms of DNA replication. *Cold Spring Harbor Perspect Biol*. 2012;4(9):a013011.
7. Griffiths SJ, Koegl M, Boutell C, et al. A systematic analysis of host factors reveals a Med23-interferon-λ regulatory axis against herpes simplex virus type 1 replication. *PLoS Pathog*. 2013;9(8):e1003514.
8. Lang FC, Li X, Vladmirova O, et al. Selective recruitment of host factors by HSV-1 replication centers. *Dongwuxue Yanjiu*. 2015;36(3):142-151.
9. Amin I, Vajeeha A, Younas S, et al. HSV-1 infection: role of viral proteins and cellular receptors. *Crit Rev Eukaryot Gene Expr*. 2019;29(5):461-469.
10. Dogrammatzis C, Waisner H, Kalamvoki M. "Non-essential" proteins of HSV-1 with essential roles in vivo: a comprehensive review. *Viruses*. 2020;13(1):17.
11. van der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 2014;114(13):6589-6631.
12. Lieutaud P, Ferron F, Uversky AV, Kurgan L, Uversky VN, Longhi S. How disordered is my protein and what is its disorder for? A guide through the "dark side" of the protein universe. *Intrinsically Disord Proteins*. 2016;4(1):e1259708.
13. Dunker AK, Babu MM, Barbar E, et al. What's in a name? Why these proteins are intrinsically disordered: why these proteins are intrinsically disordered. *Intrinsically Disord Proteins*. 2013;1(1):e24157.
14. Pentony MM, Jones DT. Modularity of intrinsic disorder in the human proteome. *Proteins Struct Funct Bioinf*. 2010;78(1):212-221.
15. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol*. 2004;337(3):635-645.
16. Zhang H, Ji X, Li P, et al. Liquid-liquid phase separation in biology: mechanisms, physiological functions and human diseases. *Sci China Life Sci*. 2020;63(7):953-985.
17. Zeng M, Chen X, Guan D, et al. Reconstituted postsynaptic density as a molecular platform for understanding synapse formation and plasticity. *Cell*. 2018;174(5):1172-1187.
18. Nag N, Sasidharan S, Uversky VN, Saudagar P, Tripathi T. Phase separation of FG-nucleoporins in nuclear pore complexes. *Biochim Biophys Acta Mol Cell Res*. 2022;1869(4):119205.
19. Brocca S, Grandori R, Longhi S, Uversky V. Liquid-liquid phase separation by intrinsically disordered protein regions of viruses: roles in viral life cycle and control of virus-host interactions. *Int J Mol Sci*. 2020;21(23):9045.
20. Lester JT, DeLuca NA. Herpes simplex virus 1 ICP4 forms complexes with TFIID and mediator in virus-infected cells. *J Virol*. 2011;85(12):5733-5744.
21. Metrick CM, Koenigsberg AL, Heldwein EE. Conserved outer tegument component UL11 from herpes simplex virus 1 is an intrinsically disordered, RNA-binding protein. *mBio*. 2020;11(3):e00810-20.
22. Lyngdoh DL, Shukla H, Sonkar A, Anupam R, Tripathi T. Portrait of the intrinsically disordered side of the HTLV-1 proteome. *ACS Omega*. 2019;4(6):10003-10018.
23. Xue B, Blocquel D, Habchi J, et al. Structural disorder in viral proteins. *Chem Rev*. 2014;114(13):6880-6911.
24. Nikolic J, Le Bars R, Lama Z, et al. Negri bodies are viral factories with properties of liquid organelles. *Nat Commun*. 2017;8(1):58.
25. Savastano A, Ibáñez de Opakua A, Rankovic M, Zweckstetter M. Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat Commun*. 2020;11(1):6041.
26. York A. Targeting viral liquid-liquid phase separation. *Nat Rev Microbiol*. 2021;19(9):550.
27. Etibor T, Yamauchi Y, Amorim M. Liquid biomolecular condensates and viral lifecycles: review and perspectives. *Viruses*. 2021;13(3):366.
28. Wei W, Bai L, Yan B, et al. When liquid-liquid phase separation meets viral infections. *Front Immunol*. 2022;13:985622.
29. Seyffert M, Georgi F, Tobler K, et al. The HSV-1 transcription factor ICP4 confers liquid-like properties to viral replication compartments. *Int J Mol Sci*. 2021;22(9):4447.

30. Vacic V, Uversky VN, Dunker AK, Lonardi S. Composition profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics*. 2007;8:211.
31. Erdős G, Pajkos M, Dosztányi Z. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res*. 2021;49(W1):W297-W303.
32. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. *Proteins*. 2001;42(1):38-48.
33. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404-405.
34. Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*. 2015;31(6):857-863.
35. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.
36. Chu X, Sun T, Li Q, et al. Prediction of liquid-liquid phase separating proteins using machine learning. *BMC Bioinformatics*. 2022;23(1):72.
37. Bolognesi B, Lorenzo Gotor N, Dhar R, et al. A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell Rep*. 2016;16(1):222-231.
38. Hardenberg M, Horvath A, Ambrus V, Fuxreiter M, Vendruscolo M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proc Natl Acad Sci*. 2020;117(52):33254-33262.
39. Disfani FM, Hsu WL, Mizianty MJ, et al. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. *Bioinformatics*. 2012;28(12):i75-i83.
40. Malhis N, Jacobson M, Gsponer J. MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences. *Nucleic Acids Res*. 2016;44(W1):W488-W493.
41. Peng Z, Kurgan L. High-throughput prediction of RNA, DNA and protein binding regions mediated by intrinsic disorder. *Nucleic Acids Res*. 2015;43(18):e121.
42. Yan J, Kurgan L. DRNApred, fast sequence-based method that accurately predicts and discriminates DNA- and RNA-binding residues. *Nucleic Acids Res*. 2017;45(10):gkx059.
43. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362-D368.
44. Bowzard JB, Visalli RJ, Wilson CB, et al. Membrane targeting properties of a herpesvirus tegument protein-retrovirus Gag chimera. *J Virol*. 2000;74(18):8692-8699.
45. MacLean CA, Clark B, McGeoch DJ. Gene UL11 of herpes simplex virus type 1 encodes a virion protein which is myristylated. *J Gen Virol*. 1989;70(Pt):3147-3157.
46. Loomis JS, Bowzard JB, Courtney RJ, Wills JW. Intracellular trafficking of the UL11 tegument protein of herpes simplex virus type 1. *J Virol*. 2001;75(24):12209-12219.
47. Yang L, Wang M, Cheng A, et al. Features and functions of the conserved herpesvirus tegument protein UL11 and its binding partners. *Front Microbiol*. 2022;13:829754.
48. DeLuca NA, McCarthy AM, Schaffer PA. Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. *J Virol*. 1985;56(2):558-570.
49. Newcomb WW, Homa FL, Thomsen DR, et al. Assembly of the herpes simplex virus capsid: characterization of intermediates observed during cell-free capsid formation. *J Mol Biol*. 1996;263(3):432-446.
50. Gao M, Matusick-Kumar L, Hurlburt W, et al. The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth. *J Virol*. 1994;68(6):3702-3712.
51. Liu F, Roizman B. Characterization of the protease and other products of amino-terminus-proximal cleavage of the herpes simplex virus 1 UL26 protein. *J Virol*. 1993;67(3):1300-1309.
52. Liu FY, Roizman B. The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. *J Virol*. 1991;65(10):5149-5156.
53. Newcomb WW, Brown JC. Structure of the herpes simplex virus capsid: effects of extraction with guanidine hydrochloride and partial reconstitution of extracted capsids. *J Virol*. 1991;65(2):613-620.
54. Sheaffer AK, Newcomb WW, Brown JC, Gao M, Weller SK, Tenney DJ. Evidence for controlled incorporation of herpes simplex virus type 1 UL26 protease into capsids. *J Virol*. 2000;74(15):6838-6848.
55. Matusick-Kumar L, Hurlburt W, Weinheimer SP, Newcomb WW, Brown JC, Gao M. Phenotype of the herpes simplex virus type 1 protease substrate ICP35 mutant virus. *J Virol*. 1994;68(9):5384-5394.
56. Goshima F, Watanabe D, Takakuwa H, et al. Herpes simplex virus UL17 protein is associated with B capsids and colocalizes with ICP35 and VP5 in infected cells. *Arch Virol*. 2000;145(2):417-426.
57. Sari TK, Gianopulos KA, Nicola AV. Glycoprotein C of herpes simplex virus 1 shields glycoprotein B from antibody neutralization. *J Virol*. 2020;94(5):e01852-19.
58. Friedman HM, Cohen GH, Eisenberg RJ, Seidel CA, Cines DB. Glycoprotein C of herpes simplex virus 1 acts as a receptor for the C3b complement component on infected cells. *Nature*. 1984;309(5969):633-635.
59. WuDunn D, Spear PG. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J Virol*. 1989;63(1):52-58.
60. Schipke J, Pohlmann A, Diestel R, et al. The C terminus of the large tegument protein pUL36 contains multiple capsid binding sites that function differently during assembly and cell entry of herpes simplex virus. *J Virol*. 2012;86(7):3682-3700.
61. Sandri-Goldin RM. The many roles of the highly interactive HSV protein ICP27, a key regulator of infection. *Future Microbiol*. 2011;6(11):1261-1277.
62. Malik P, Tabarraei A, Kehlenbach RH, et al. Herpes simplex virus ICP27 protein directly interacts with the nuclear pore complex through Nup62, inhibiting host nucleocytoplasmic transport pathways. *J Biol Chem*. 2012;287(15):12277-12292.
63. Deschamps T, Kalamvoki M. Extracellular vesicles released by herpes simplex virus 1-infected cells block virus replication in recipient cells in a STING-dependent manner. *J Virol*. 2018;92(18):e01102-18.
64. Diefenbach RJ, Miranda-Saksena M, Diefenbach E, et al. Herpes simplex virus tegument protein US11 interacts with

conventional kinesin heavy chain. *J Virol*. 2002;76(7):3282-3291.

65. Benboudjema L, Mulvey M, Gao Y, Pimplikar SW, Mohr I. Association of the herpes simplex virus type 1 Us11 gene product with the cellular kinesin light-chain-related protein PAT1 results in the redistribution of both polypeptides. *J Virol*. 2003;77(17):9192-9203.
66. Wagner MJ, Smiley JR. Herpes simplex virus requires VP11/12 to induce phosphorylation of the activation loop tyrosine (Y394) of the Src family kinase Lck in T lymphocytes. *J Virol*. 2009;83(23):12452-12461.
67. Wagner MJ, Smiley JR. Herpes simplex virus requires VP11/12 to activate Src family kinase-phosphoinositide 3-kinase-Akt signaling. *J Virol*. 2011;85(6):2803-2812.
68. Strunk U, Saffran HA, Wu FW, Smiley JR. Role of herpes simplex virus VP11/12 tyrosine-based motifs in binding and activation of the Src family kinase Lck and recruitment of p85, Grb2, and Shc. *J Virol*. 2013;87(20):11276-11286.
69. Alazard-Dany N, Nicolas A, Ploquin A, et al. Definition of herpes simplex virus type 1 helper activities for adeno-associated virus early replication events. *PLoS Pathog*. 2009;5(3):e1000340.
70. Cai WZ, Schaffer PA. Herpes simplex virus type 1 ICP0 plays a critical role in the de novo synthesis of infectious virus following transfection of viral DNA. *J Virol*. 1989;63(11):4579-4589.
71. Cheung P, Panning B, Smiley JR. Herpes simplex virus immediate-early proteins ICP0 and ICP4 activate the endogenous human alpha-globin gene in nonerythroid cells. *J Virol*. 1997;71(3):1784-1793.
72. Everett RD. Trans activation of transcription by herpes virus products: requirement for two HSV-1 immediate-early polypeptides for maximum activity. *EMBO J*. 1984;3(13):3135-3141.
73. El Kasmi I, Lippé R. Herpes simplex virus 1 gN partners with gM to modulate the viral fusion machinery. *J Virol*. 2015;89(4):2313-2323.
74. Meckes, Jr. DG, Wills JW. Dynamic interactions of the UL16 tegument protein with the capsid of herpes simplex virus. *J Virol*. 2007;81(23):13028-13036.
75. Wigington CP, Williams KR, Meers MP, Bassell GJ, Corbett AH. Poly(A) RNA-binding proteins and polyadenosine RNA: new members and novel functions. *WIREs RNA*. 2014;5(5):601-622.
76. Person S, Laquerre S, Desai P, Hempel J. Herpes simplex virus type 1 capsid protein, VP21, originates within the UL26 open reading frame. *J Gen Virol*. 1993;74(Pt):2269-2273.
77. Liu FY, Roizman B. The promoter, transcriptional unit, and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in frame with the UL26 open reading frame. *J Virol*. 1991;65(1):206-212.
78. Jambunathan N, Clark C, Musarrat F, Chouljenko V, Rudd J, Kousoulas K. Two sides to every story: herpes simplex type-1 viral glycoproteins gB, gD, gH/gL, gK, and cellular receptors function as key players in membrane fusion. *Viruses*. 2021;13(9):1849.
79. Haanes EJ, Nelson CM, Soule CL, Goodman JL. The UL45 gene product is required for herpes simplex virus type 1 glycoprotein B-induced fusion. *J Virol*. 1994;68(9):5825-5834.
80. Crosby HA, Lalor PF, Ross E, Newsome PN, Adams DH. Adhesion of human haematopoietic (CD34+) stem cells to human liver compartments is integrin and CD44 dependent and modulated by CXCR3 and CXCR4. *J Hepatol*. 2009;51(4):734-749.
81. Desai PJ. A null mutation in the UL36 gene of herpes simplex virus type 1 results in accumulation of unenveloped DNA-filled capsids in the cytoplasm of infected cells. *J Virol*. 2000;74(24):11608-11618.
82. Kelly BJ, Fraefel C, Cunningham AL, Diefenbach RJ. Functional roles of the tegument proteins of herpes simplex virus type 1. *Virus Res*. 2009;145(2):173-186.
83. Bucks MA, Murphy MA, O'Regan KJ, Courtney RJ. Identification of interaction domains within the UL37 tegument protein of herpes simplex virus type 1. *Virology*. 2011;416:42-53.
84. Mossman, KL, Sherburne, R, Lavery, C, Duncan, J, Smiley, JR. Evidence that herpes simplex virus VP16 is required for viral egress downstream of the initial envelopment event. *J Virol*. 2000;74:6287-6299. doi:10.1128/jvi.74.14.6287-6299.2000
85. Park D, Lalli J, Sedlackova-Slavikova L, Rice SA. Functional comparison of herpes simplex virus 1 (HSV-1) and HSV-2 ICP27 homologs reveals a role for ICP27 in virion release. *J Virol*. 2015;89(5):2892-2905.
86. Jahedi S, Markovitz NS, Filatov F, Roizman B. Colocalization of the herpes simplex virus 1 UL4 protein with infected cell protein 22 in small, dense nuclear structures formed prior to onset of DNA synthesis. *J Virol*. 1999;73(6):5132-5136.
87. Xing J, Wang S, Lin R, Mossman KL, Zheng C. Herpes simplex virus 1 tegument protein US11 downmodulates the RLR signaling pathway via direct interaction with RIG-I and MDA-5. *J Virol*. 2012;86(7):3528-3540.
88. Kato A, Yamamoto M, Ohno T, et al. Herpes simplex virus 1-encoded protein kinase UL13 phosphorylates viral Us3 protein kinase and regulates nuclear localization of viral envelopment factors UL34 and UL31. *J Virol*. 2006;80(3):1476-1486.
89. Beilstein F, Higgs MR, Stow ND. Mutational analysis of the herpes simplex virus type 1 DNA packaging protein UL33. *J Virol*. 2009;83:8938-8945. doi:10.1128/jvi.01048-09
90. Loret S, Guay G, Lippe R. Comprehensive characterization of extracellular herpes simplex virus type 1 virions. *J Virol*. 2008;82:8605-8618. doi:10.1128/JVI.00904-08
91. Case LB, Zhang X, Ditlev JA, Rosen MK. Stoichiometry controls activity of phase-separated clusters of actin signaling proteins. *Science*. 2019;363(6431):1093-1097.
92. Kilic S, Lezaja A, Gatti M, et al. Phase separation of 53BP1 determines liquid-like behavior of DNA repair compartments. *EMBO J*. 2019;38(16):e101379.
93. Zhang G, Wang Z, Du Z, Zhang H. mTOR regulates phase separation of PGL granules to modulate their autophagic degradation. *Cell*. 2018;174(6):1492-1506.
94. Antonicka H, Shoubridge EA. Mitochondrial RNA granules are centers for posttranscriptional RNA processing and ribosome biogenesis. *Cell Rep*. 2015;10(6):920-932.
95. Heinrich BS, Maliga Z, Stein DA, Hyman AA, Whelan SPJ. Phase transitions drive the formation of vesicular stomatitis virus replication compartments. *mBio*. 2018;9(5):e02290-17.

96. Di Nunzio F, Uversky VN, Mouland AJ. Biomolecular condensates: insights into early and late steps of the HIV-1 replication cycle. *Retrovirology*. 2023;20(1):4.
97. Alenquer M, Vale-Costa S, Etibor TA, Ferreira F, Sousa AL, Amorim MJ. Influenza A virus ribonucleoproteins form liquid organelles at endoplasmic reticulum exit sites. *Nat Commun*. 2019;10(1):1629.
98. Chen H, Cui Y, Han X, et al. Liquid-liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. *Cell Res*. 2020;30(12):1143-1145.
99. Peng Q, Wang L, Qin Z, et al. Phase separation of Epstein-Barr virus EBNA2 and its coactivator EBNALP controls gene expression. *J Virol*. 2020;94(7):e01771-19.
100. Jobe F, Simpson J, Hawes P, Guzman E, Bailey D. Respiratory syncytial virus sequesters NF-κB subunit p65 to cytoplasmic inclusion bodies to inhibit innate immune signaling. *J Virol*. 2020;94(22):e01380-20.
101. Yamauchi Y. Influenza A virus uncoating. *Adv Virus Res*. 2020;106:1-38.
102. Iserman C, Roden CA, Boerneke MA, et al. Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol Cell*. 2020;80(6):1078-1091.
103. Monette A, Mouland AJ. Zinc and copper ions differentially regulate prion-like phase separation dynamics of Pan-virus nucleocapsid biomolecular condensates. *Viruses*. 2020;12(10):1179.
104. Wang J, Shi C, Xu Q, Yin H. SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation into stress granules through its N-terminal intrinsically disordered region. *Cell Discov*. 2021;7(1):5.
105. Monette A, Niu M, Chen L, Rao S, Gorelick RJ, Mouland AJ. Pan-retroviral nucleocapsid-mediated phase separation regulates genomic RNA positioning and trafficking. *Cell Rep*. 2020;31(3):107520.
106. Alshareedah I, Moosa MM, Raju M, Potoyan DA, Banerjee PR. Phase transition of RNA-protein complexes into ordered hollow condensates. *Proc Natl Acad Sci*. 2020;117(27):15650-15658.
107. Guo Q, Shi X, Wang X. RNA and liquid-liquid phase separation. *Non-coding RNA Res*. 2021;6(2):92-99.
108. Perdikari TM, Murthy AC, Ryan VH, Watters S, Naik MT, Fawzi NL. SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation stimulated by RNA and partitions into phases of human ribonucleoproteins. *bioRxiv*. 2020;65:615-620.
109. Sasidharan S, Nag N, Tripathi T, Saudagar P. Chapter 12 - Interactions and interplay of MLOs with classical membrane-bound organelles. In: Uversky VN, ed. *Droplets of Life*. Academic Press; 2023:375-395.
110. Bedadala GR, Pinnoji RC, Hsia SCV. Early growth response gene 1 (Egr-1) regulates HSV-1 ICP4 and ICP22 gene expression. *Cell Res*. 2007;17(6):546-555.
111. Gu B, Kuddus R, DeLuca NA. Repression of activator-mediated transcription by herpes simplex virus ICP4 via a mechanism involving interactions with the basal transcription factors TATA-binding protein and TFIIB. *Mol Cell Biol*. 1995;15(7):3618-3626.
112. Friedman HM, Wang L, Pangburn MK, Lambris JD, Lubinski J. Novel mechanism of antibody-independent complement neutralization of herpes simplex virus type 1. *J Immunol*. 2000;165(8):4528-4536.
113. Etingin OR, Silverstein RL, Hajjar DP. Identification of a monocyte receptor on herpesvirus-infected endothelial cells. *Proc Natl Acad Sci*. 1991;88(16):7200-7203.
114. Souki SK, Hernandez FP, Sandri-Goldin RM. Arginine methylation of the RGG box does not appear to regulate ICP27 import during herpes simplex virus infection. *J Virol*. 2011;85(13):6809-6813.
115. Souki SK, Sandri-Goldin RM. Arginine methylation of the ICP27 RGG box regulates the functional interactions of ICP27 with SRPK1 and Aly/REF during herpes simplex virus 1 infection. *J Virol*. 2009;83(17):8970-8975.
116. Kawaguchi Y, Van Sant C, Roizman B. Herpes simplex virus 1 alpha regulatory protein ICP0 interacts with and stabilizes the cell cycle regulator cyclin D3. *J Virol*. 1997;71(10):7328-7336.
117. Pozhidaeva AK, Mohni KN, Dhe-Paganon S, et al. Structural characterization of interaction between human ubiquitin-specific protease 7 and immediate-early protein ICP0 of herpes simplex virus-1. *J Biol Chem*. 2015;290(38):22907-22918.
118. Lilley CE, Chaurushiya MS, Boutell C, et al. A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and DNA damage responses. *EMBO J*. 2010;29(5):943-955.
119. Charron AJ, Ward SL, North BJ, Ceron S, Leib DA. The US11 gene of herpes simplex virus 1 promotes neuroinvasion and periocular replication following corneal infection. *J Virol*. 2019;93(9):e02246-18.
120. Giraud S, Diaz-Latoud C, Hacot S, Textoris J, Bourette RP, Diaz JJ. US11 of herpes simplex virus type 1 interacts with HIPK2 and antagonizes HIPK2-induced cell growth arrest. *J Virol*. 2004;78(6):2984-2993.
121. Fossum E, Friedel CC, Rajagopala SV, et al. Evolutionarily conserved herpesviral protein interaction networks. *PLoS Pathog*. 2009;5(9):e1000570.
122. Risso-Ballester J, Galloux M, Cao J, et al. A condensate-hardening drug blocks RSV replication in vivo. *Nature*. 2021;595(7868):596-599.
123. Wang S, Dai T, Qin Z, et al. Targeting liquid–liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. *Nat Cell Biol*. 2021;23(7):718-732.
124. Itzhaki RF. Herpes simplex virus type 1 and Alzheimer's disease: increasing evidence for a major role of the virus. *Front Aging Neurosci*. 2014;6:202.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Subedi S, Nag N, Shukla H, Padhi AK, Tripathi T. Comprehensive analysis of liquid–liquid phase separation propensities of HSV-1 proteins and their interaction with host factors. *J Cell Biochem*. 2023;1-24. doi:10.1002/jcb.30480