Retrovirology

REVIEW

Open Access

Biomolecular condensates: insights into early and late steps of the HIV-1 replication cycle

Francesca Di Nunzio[1], Vladimir N. Uversky[2] and Andrew J. Mouland[3,4,5*]

Abstract

A rapidly evolving understanding of phase separation in the biological and physical sciences has led to the redefining of virus-engineered replication compartments in many viruses with RNA genomes. Condensation of viral, host and genomic and subgenomic RNAs can take place to evade the innate immunity response and to help viral replication. Divergent viruses prompt liquid–liquid phase separation (LLPS) to invade the host cell. During HIV replication there are several steps involving LLPS. In this review, we characterize the ability of individual viral and host partners that assemble into biomolecular condensates (BMCs). Of note, bioinformatic analyses predict models of phase separation in line with several published observations. Importantly, viral BMCs contribute to function in key steps retroviral replication. For example, reverse transcription takes place within nuclear BMCs, called HIV-MLOs while during late replication steps, retroviral nucleocapsid acts as a driver or scaffold to recruit client viral components to aid the assembly of progeny virions. Overall, LLPS during viral infections represents a newly described biological event now appreciated in the virology field, that can also be considered as an alternative pharmacological target to current drug therapies especially when viruses become resistant to antiviral treatment.

Keywords Retrovirus, HIV-1, Liquid–liquid phase separation (LLPS), Biomolecular condensates (BMC), Membraneless organelles (MLO), Intrinsically disordered protein (IDP), Droplet-promoting region, Intrinsically disordered region (IDR), Nuclear pore complex, Integration, Transcription, Virus assembly, Protein–protein interactions, Coacervation, Posttranslational modification, Molecular recognition feature

Background

Liquid–liquid phase separation (LLPS) relies on the condensation of proteins and often (but not always) RNAs or DNAs (or other biological polymers) into membraneless organelles (MLOs), also known as biomolecular condensates (BMCs), allowing select molecules to become concentrated while excluding others. This process, which represents "coacervation", if it is driven by the polyelectrostatic interactions between macromolecules undergoing LLPS, compartmentalizes fundamental processes in the nucleus and in the cytoplasm. One should keep in mind that coacervation is one of several types of LLPS which occurs in solutions of oppositely charged macromolecular species including proteins, nucleic acids,

*Correspondence:
Andrew J. Mouland
andrew.mouland@mcgill.ca
[1] Advanced Molecular Virology Unit, Department of Virology, Institut Pasteur, Université Paris Cité, 75015 Paris, France
[2] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA
[3] Lady Davis Institute at the Jewish General Hospital, Montréal, QC H3T 1E2, Canada
[4] Department of Microbiology and Immunology, McGill University, Montréal, QC H3A 2B4, Canada
[5] Department of Medicine, McGill University, Montréal, QC H4A 3J1, Canada

polymers, and colloids. The other mechanisms driving LLPS include weak pi-pi, cation-pi, and hydrophobic interactions [1]. Compartmentalization of function relies on MLOs such as the nucleolus, nuclear speckle, neuronal RNA trafficking granules, stress granules, and processing bodies [2, 3] enables key gene expression steps in which multiple components must come into proximity for function [4–15]. The presence of intrinsically disordered regions (IDR) in proteins and multivalent macromolecular interactions between proteins or between proteins and nucleic acids, lead to liquid demixing and phase separation [16–22]. One of the theories describing this process in the cell is the "stickers-and-spacers" model [23] which considers protein and RNA as flexible polymers where sequence motifs, the "stickers," determine attractive interactions between different molecules while other regions act as inert "spacers" between them. Stickers on the protein/RNA chain, once a critical concentration threshold is reached, can induce a separation into a denser phase that coexists with a more dilute phase.

It is commonly accepted that RNA played a key role in the evolution of life. In the "RNA world" theory, the primitive RNA arose spontaneously using specific features of the early Earth environment. RNA can serve as template to code for proteins. Experiments in vitro show that RNA molecules perform a wide range of catalytic reactions. Furthermore, RNA and proteins carrying IDRs can assemble into the biocondensates, which are formed by LLPS. These coacervates have been implicated in the origins of life, acting as bioreactors, where nucleotides, magnesium, and RNA molecules accumulate and increase in local concentration [24, 25]. Indeed, one of the most interesting questions in science is how life first appeared. In 1938, the Soviet biochemist Alexander Oparin in "The Origin of Life" proposed that life originated as coacervate drops of organic materials [26]. His theory was based on simple observation that droplets of organic molecules merge spontaneously from an otherwise diluted solution. The inability to demonstrate the presence of membrane barriers accounted for the loss of credibility in Oparin's coacervate idea. Although in recent years this theory has been reexamined due to the discovery of numerous intracellular, membraneless BMCs [27]. Such liquid droplets or MLOs are generated by LLPS. LLPS ensures the tightly controlled performance of complex cellular functions in a limited space and in a spatiotemporal manner. LLPS is based on multivalent interactions among IDRs and/or modular interacting domains of some components [5, 28, 29]. LLPS is a widespread phenomenon that can explain numerous cellular functions [30–32], illness states [33], and host-pathogens coexistence [34]. It is not surprising that viruses learned to usurp LLPS [35]. "Viral factories" and other "viral replication centers" in infected cells are central to virus replication and are thus defined as BMCs [36–40].

The oligomerization of IDR-rich proteins is key to driving phase separation. Furthermore, there are several examples in nature, such as nascent ribosomal RNA (rRNA) transcripts [41, 42] or long non-coding RNA (lncRNA) such as Neat1 [43] and other types of RNAs are associated with specific DNA loci that may act as scaffolds for locally enriching self-interacting IDPs [44]. Several proteins appear to self-oligomerize leading to phase separation, like NPM1 that pentamerizes to form a radial array of IDRs and RNA-binding domains important for phase separation and nucleolus assembly [30, 45]. The oligomerization of the protein G3BP1 is also important for stress granule condensation [46]. It is also clear that the LLPS-driven biogenesis of BMCs can be dysregulated by pathological gene mutations [47] that lead to neurodegenerative disease [48–50] and to tumorigenesis [48–50].

The importance of LLPS and related BMCs/MLOs in relation to infectious diseases has recently gained widespread attention due in part to focussed research efforts on the RNA virus, SARS-CoV-2, etiologic to COVID-19. SARS-CoV-2 expresses a (N)nucleocapsid protein that phase separates and co-condenses with several host and viral proteins in the assembly of its RNA genome to aid viral packaging [11, 51, 52]. Indeed, N acts as a filter to select vRNA full-length genome excluding subgenomic vRNA from BMCs. Likewise, HIV-1 and other retroviral proteins possess intrinsically disordered domains that are juxtaposed with RNA- and Zn-binding domains (RBD and ZnF) [11, 53, 54] that promote protein co-condensation and confer liquid-like behaviour of resulting condensates [55]. HIV-1 Gag BMCs exhibit fluid-like behaviour characterized by fluid movement, fusion and fission [56]. The fluidity of BMCs can also be enhanced by RNA thereby promoting more dynamic interactions [57, 58], likely playing a role for viruses with RNA genomes. Endogenous retroviruses (ERV) also elicit BMC assemblies for replication [59, 60]. A recent study indicates that ERV retrotransposons have the ability to hijack biomolecular condensates enriched of transcriptional regulatory proteins in pluripotent cells. ERVs reactivation seems to compete with super-enhancers for transcriptional condensates in pluripotent cells. Thus, the repression of ERVs is important for the correct cellular fates This mechanism explains why thousands of transposition-incapable ERVs are repressed in mammals [60].

Retroviral replication: implications for membraneless organelles generated by phase separation

Fifty-two years ago, the discovery of reverse transcription [61, 62] amended the central dogma showing that not only the DNA but also the RNA can contain the genetic information to be coded into protein through a reverse transcription process that converts the RNA in DNA, the process, which is the hallmark characteristic retroviruses. The discovery of reverse transcription links RNA to the DNA world. Before the discovery of reverse transcription, the first evidence of the existence of retroviruses had been obtained from Robin Weiss who was able to demonstrate that Env protein is expressed from a retroviral genome integrated in the host DNA [63].

Retroviruses are a large family of RNA viruses, and they are classified in orthovirinae and spumavirinae, the former with six genera, the latter with five [64]. Among the family of orthoretrovirinae, there is the subfamily of lentiviruses that has the ability to infect both dividing and non-dividing cells. Early and late phases define the retroviral replication cycle (Fig. 1). The early steps begin with viral fusion to the cellular membrane followed by the cytoplasmic viral journey of the viral core to reach the host chromatin for the viral integration into the host genome. As a result, the host cell becomes a chimera, and the host transcriptional mechanism takes over the transcription of viral genes in a similar way to its own genes, which must be highly expressed. Viral gene transcription, nuclear export of viral RNAs, mRNA translation and the assembly of new viral particles represent key late steps of the retroviral replication cycle.

Condensates are formed during both the early and late viral cycle phases to create microenvironments to facilitate and optimize viral replication and likely, to shield the viral genome from anti-viral innate immune host responses (Fig. 1). Often, condensates serve to organize cellular biochemistry by bringing the correct partners together in a confined environment to accelerate reaction rates, to quickly reply to stress, or to act as filters. These are new biological phenomena that have only recently been recognized as a part of key steps in the replication cycle of the HIV-1, etiological to the AIDS pandemic. Certain proteins function as scaffolds and are critical elements involved in driving the biogenesis of condensates (drivers or scaffolds), as it is the case for host proteins, SRRM2 and SON, that are essential for the formation of nuclear speckles, which are also considered MLOs [65, 66]. The implications for

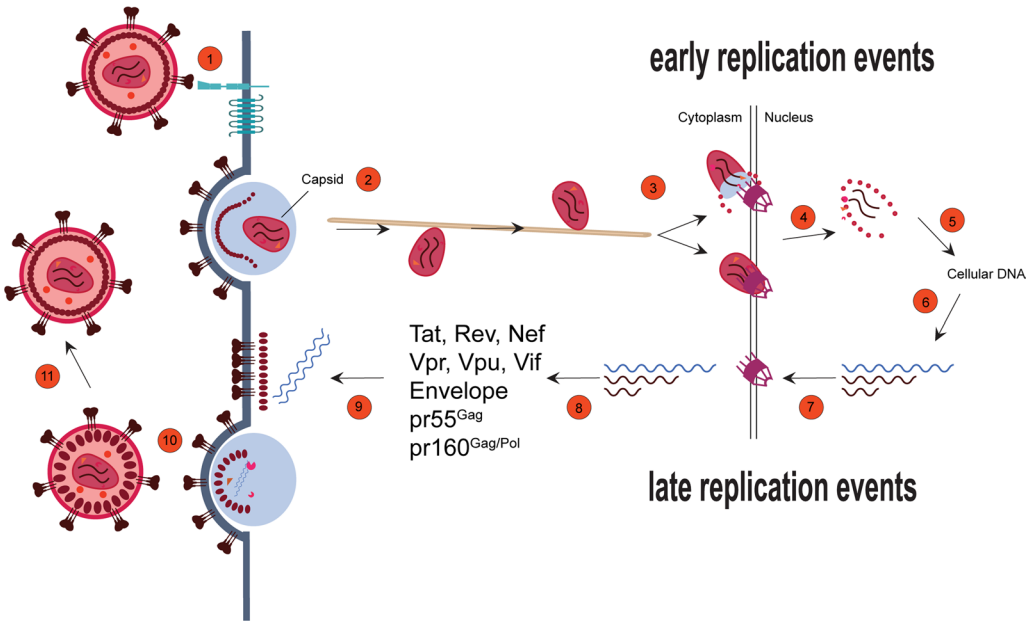


Fig. 1 Evolving and potential roles for biomolecular condensates at various steps of the retroviral replication cycle. 1, engagement with host cell receptors and viral entry; 2, membrane fusion and intracellular trafficking of the viral capsid; 3, nuclear entry; 4, uncoating and completion of reverse transcription; 5, integration into host DNA; 6, transcription, intranuclear trafficking of the viral ribonucleoprotein; 7, nucleocytoplasmic mRNA export through the nuclear pore with completely spliced mRNAs exported first then the singly-spliced and unspliced mRNAs; 8, mRNA engagement with the host translation machinery to generate early (Tat, Rev and Nef) and late ($pr55^{Gag}$, $pr160^{Gag/Pol}$, Vpr, Vpu and Vif) viral gene products, as indicated, and genomic RNA selection; 9, genomic RNA and viral protein trafficking to virus assembly sites; 10, virus budding; and 11, virus maturation mediated by the viral protease substrates, $pr55^{Gag}$ and $pr160^{Gag/Pol}$. Please see text for details on condensation at virus replication steps

MLOs in retroviral replication based on recent literature and bioinformatic analyses are discussed below.

Early events of the retroviral replication cycle

Nuclear translocation

The nuclear pore complex (NPC) is the path used by HIV-1 to enter the nucleus of both dividing and non-dividing cells**.** The viral capsid docks at the NPC through its interaction with the cytoplasmic filaments of the pore, mainly formed by the nucleoporin RanBP2 [67, 68] and next, translocates through the NPC channel helped by an intranuclear nucleoporin the Nup153 [69, 70]). Thus, the viral capsid is the determinant responsible for the viral nuclear entry because it interacts with specific nucleoporins to first dock at the NPC and then translocate through it [67–71]. Once inside the nucleus, the virus will enter in contact with host nuclear factors as well as with an euchromatin environment located all around the nuclear basket of the NPC. The nuclear basket of the NPC is surrounded by euchromatin [71, 72], contrary to the chromatin under the nuclear envelope, which is heterochromatin, known as lamin associated domains (LADs). HIV-1 favors integration in active chromatin, especially in the chromatin near the NPC [72, 73]) and others near nuclear speckles [74, 75]. On the other hand, HIV-1 disfavors integration in LADs that are silent chromatin regions [70, 74]. How does the partitioning between euchromatin and heterochromatin underneath the NE occur? Some nucleoporins, especially the ones carrying IDRs, like Nup153 that has phenylalanine-glycine (FG) domains could assist in forming phase-separated transcriptional compartments [76]. Perhaps the proximity to the nuclear pore and the abundance of proteins with IDRs at super-enhancers could recruit Nups, which could assist in phase separation through mediating proximity to nuclear speckles (NSs), which are surrounded by active chromatin, and NPCs. Nup153 is an important component of the nuclear basket and it is also part of a free intranuclear population and holds 29 FG repeats that are highly disordered [77]. Figure 2 and Table 1 represent results of functional disorder analysis of human Nup153 and shows that this host protein contains high levels of intrinsic disorder regions as evidenced by the per-residue disorder profile generated by a set of commonly used disorder predictors PONDR® VLXT, PONDR® VL3, PONDR® VSL2, PONDR® FIT, IUPred_long and IUPred_short) (Fig. 2A) and a functional disorder profile generated by the D^2P^2 platform (Fig. 2B). Therefore, both analyses show that most of the protein is expected to be disordered. Furthermore, Fig. 2B shows that Nup153 contains multiple disorder-based binding sites (i.e., IDRs that are capable of folding at interaction with specific binding partners and known as molecular recognition features, MoRFs). Interestingly, Nup153 has been found to mediate chromatin structure and influence transcription by modulating CTCF and cohesin binding across *cis*-regulatory elements and topologically associating domains (TAD) boundaries [78]. Note that MoRFs

(See figure on next page.)

Fig. 2 Functional disorder analysis of human Nup153 protein (UniProt ID: P49790). **A** Per-residue disorder profile generated by RIDAO [138]. Per-residue intrinsic disorder profiles generated by PONDR® VLXT, PONDR® VL3, PONDR® VSL2, PONDR® FIT, IUPred_long and IUPred_short. Mean disorder profile (MDP) is calculated by averaging of the disorder profiles of individual predictors. Light pink shade represents MDP error distribution. The thin black line at the disorder score of 0.5 is the threshold between order and disorder, where residues/regions above 0.5 are disordered and residues/regions below 0.5 are ordered. The thin dashed line at the disorder score of 0.15 is the threshold between order and flexibility, where residues/regions above 0.15 are flexible and residues/regions below 0.15 are highly ordered. **B** Functional disorder profile generated by the D^2P^2 platform [139]. Here, the IDR localization predicted by IUPred, PONDR® VLXT, PONDR® VSL2, PrDOS, PV2, and ESpritz are shown by 9 differently colored bars on the top of the plot, whereas the blue-green-white bar in the middle of the plots shows the agreement between the outputs of these disorder predictors, with disordered regions by consensus being shown by blue and green. The two lines with colored and numbered bars above the disorder consensus bar show the positions of functional SCOP domains [140, 141] predicted using the SUPERFAMILY predictor [142]. Positions of the predicted disorder-based binding sites (MoRF regions) identified by the ANCHOR algorithm are shown by yellow zigzagged bars [143]. Locations of the sites of different PTMs identified by the PhosphoSitePlus platform [144] are shown at the bottom of the plot by the differently colored circles. Here, phosphorylation, ubiquitylation, acetylation, glycosylation and methylation sites are shown by red, violet, yellow, orange and blue circles with the letters P, U, A, G, and M, respectively. An interactive profile generated for human Nup153 protein by D^2P^2 can be found at the following link: https://d2p2.pro/view/sequence/up/P49790. **C** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]. Top plot represents sequence distribution of the probability of forming a droplet state through liquid–liquid phase separation (p_{LLPS}). Proteins with $p_{LLPS} \geq 0.60$ are droplet-drivers, which can spontaneously undergo liquid–liquid phase separation. Droplet-client proteins have $p_{LLPS} < 0.60$, but possess droplet-promoting regions (DPRs), which can induce their partitioning into condensates. Positions of DPRs (which contain residues with $p_{DP} \geq 0.60$ capable to promote liquid–liquid phase separation) are shown by blue bar. Positions of aggregation hot-spots that drive aggregation of condensates are displayed by orange boxes. These regions are droplet-promoting ($p_{DP} \geq 0.60$) and exhibit high interaction mode divergence ($S_{BIND} \geq 2.2$) making them sensitive to the cellular context. The bottom plot shows sequence distribution of the S_{BIND} values that characterize the ability of residues to switch between different binding modes. These are computed as Shannon entropy values of the binding mode distribution [146]. Plot also shows regions with context-dependent interactions that change protein interaction behavior and binding modes under different cellular conditions. These residues can be ordered or disordered with $S_{BIND} \geq 2.25$. **D** Model of the CPSF6 3D structure generated by AlphaFold [147]. Structure is colored based on the AlphaFold-generated per-residue confidence score (predicted local distance difference test, pLDDT) values, where orange, yellow, cyan, and blue colors correspond to the segments predicted by AlphaFold with very high very low (pLDDT < 50), low (70 > pLDDT > 50), high (90 > pLDDT > 70), and (pLDDT > 90) confidence

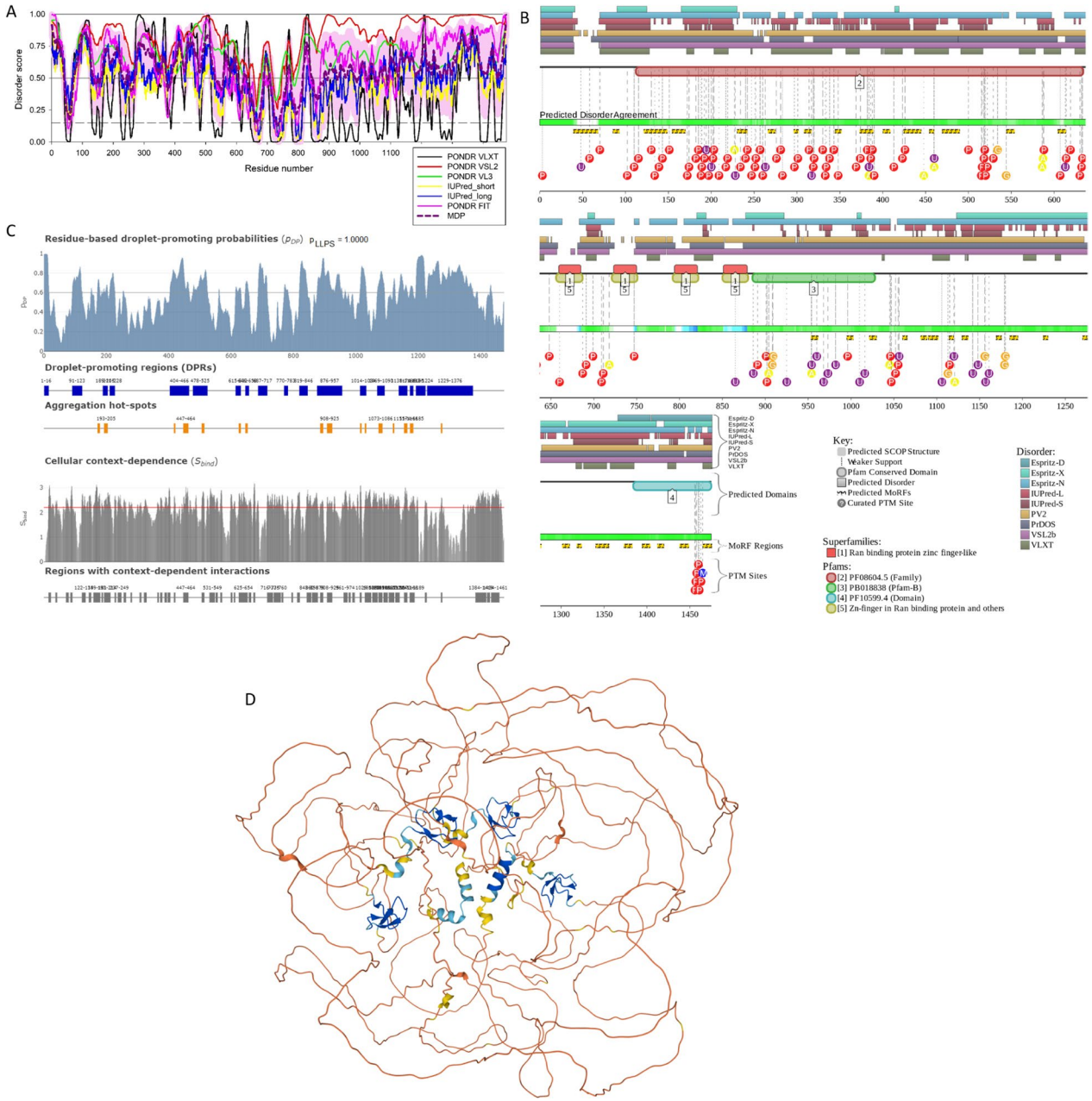


Fig. 2 (See legend on previous page.)

and post-translational modifications (PTMs) are distributed throughout the entire sequence of this protein. Figure 2C and Table 1 suggest that Nup153 has a very strong propensity to phase separate (its LLPS propensity evaluated by FuzDrop is equal to 1, and possesses numerous droplet-promoting regions (DPRs), in line with Nup153 gel formation properties [79]. This indicates that Nup153 can serve as a driver (or a scaffold) controlling the formation of Nup153-containing BMCs. Finally, Fig. 2D provides an overview of the expected 3D structure of this protein generated by AlphaFold and shows that a very significant part of Nup153 is highly disordered.

Viral post-nuclear entry steps

Since the formation of various MLOs represents one cellular stress response mechanism, it is not surprising to find that HIV-1 infection prompts the formation of MLOs. These intranuclear viral-host structures defined as HIV-1 MLOs [75] characterize viral post-nuclear entry steps. In these organelles, the scaffold

Table 1 Levels of intrinsic disorder and LLPS predispositions of host and viral proteins related to HIV-1 replication cycle

	$PPIDR_{PONDR\ VSL2}$ (%)	P_{LLPS}
Viral proteins		
CA	22.94	0.1607
$pr55^{Gag}$	43.40	0.7147
RT	19.11	0.1892
IN	17.79	0.1681
NC	100.00	0.6502
Host proteins		
Nup153	94.64	1.0000
CPSF6	79.31	0.9976
LEDGF	77.74	0.9963

protein is the cleavage and polyadenylation specific factor 6 (CPSF6) that contains mixed-charge domains (MCDs), which are protein domains or regions/clusters characterized by high enrichment for positive and negatively charged amino acids, where the relative numbers of basic to acidic components may be balanced or skewed [80–83], at the *C*-terminus region of the protein. MCDs seem to be critical for its capacity to phase separate [82]. CPSF6 binds the hydrophobic pockets of the viral core [84, 85]. Interestingly, the human CPSF6 is characterized by high levels of intrinsic disorder (see Fig. 3A, B and Table 1) and contains multiple disorder-based binding sites (i.e., disordered regions capable of disorder-to-order transition at interaction with specific partners) and multiple sites of various PTMs (see Fig. 3B). Furthermore, this protein shows high LLPS propensity ($p_{LLPS} = 0.9976$) and contains two long droplet-promoting regions and multiple regions with context-dependent interactions (see Fig. 3C). These observations indicate that CPSF6 can serve as a driver of LLPS (see Table 1). The fact that CPSF6 is a highly disordered protein is further supported by Fig. 3D representing the 3D structure of this protein generated by AlphaFold. Note that the most structured regions of this protein that are predicted with highest confidence are the RNA-binding motif (RRM) domain (residues 81–161) and the segment required for the RNA binding (residues 405–551) [86] (regions in blue and in cyan).

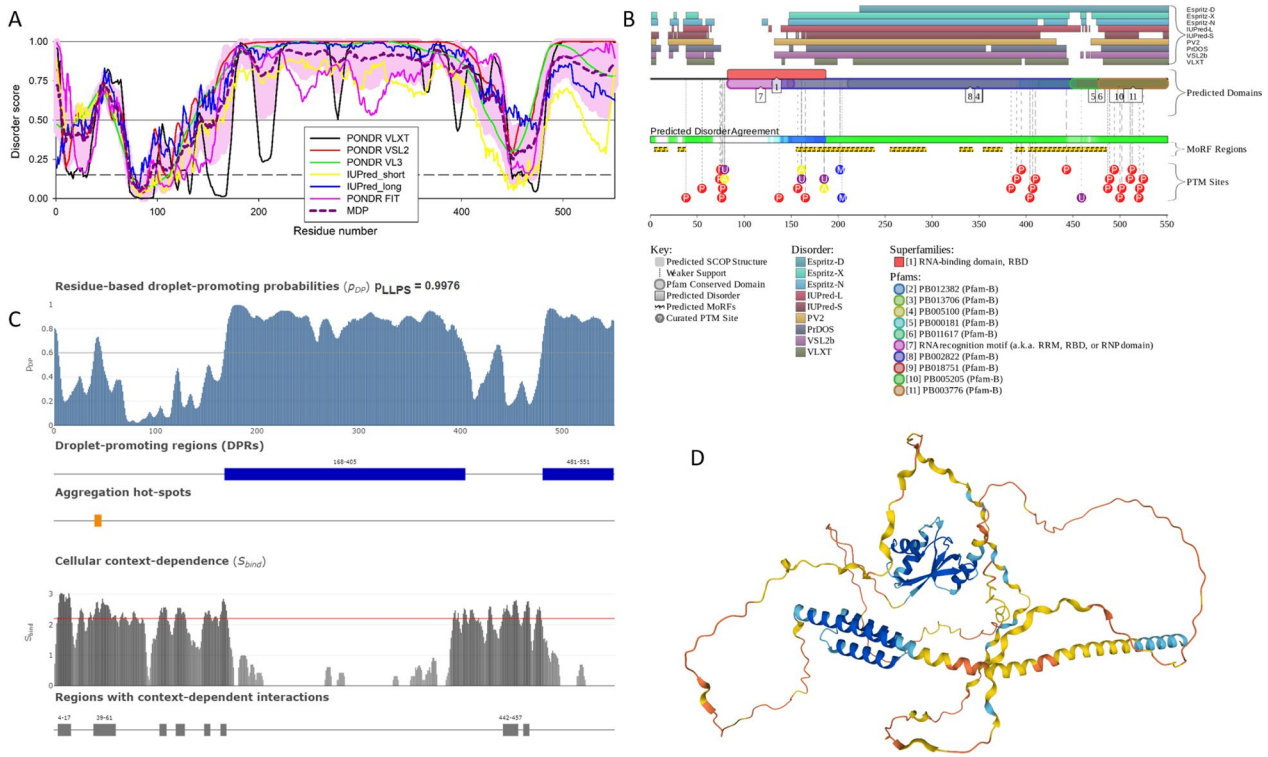


Fig. 3 Functional disorder analysis of human CPSF6 protein (UniProt ID: Q16630). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Functional disorder profile generated by the D^2P^2 platform [139]. Here, phosphorylation, ubiquitylation, acetylation, and methylation sites are shown by red, violet, yellow, and blue circles with the letters P, U, A, and M, respectively. A D^2P^2-generated interactive profile of the human CPSF6 protein can be found at the following link: https://d2p2.pro/view/sequence/up/Q16630. **C** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]. **D** Model of the CPSF6 3D structure generated by AlphaFold [147]

The presence of several IDRs and DPRs can account in part for the formation of HIV-1 MLOs in combination with the presence of the viral capsid in the nucleus [75]. It is possible that CPSF6 condensates act as a scaffold to recruit the viral capsid as a client. Of note, client molecules are not able to form condensates by their own, but through direct interaction with scaffolds can associate [5, 87] with condensates [87]. In line with these considerations, Fig. 4 and Table 1 show that although capsid is predicted to contain noticeable levels of disordered and flexible regions and possesses five regions with context-dependent interactions, this protein is incapable of driving LLPS. However, although being characterized by a low p_{LLPS} value of 0.1607 (see Fig. 4B, and Table 1), capsid is predicted by FuzDrop to contain one DPR. These observations suggest that the capsid can serve as an MLO client, which is in line with experimental data [92]. It is possible that the DPR could be exposed when

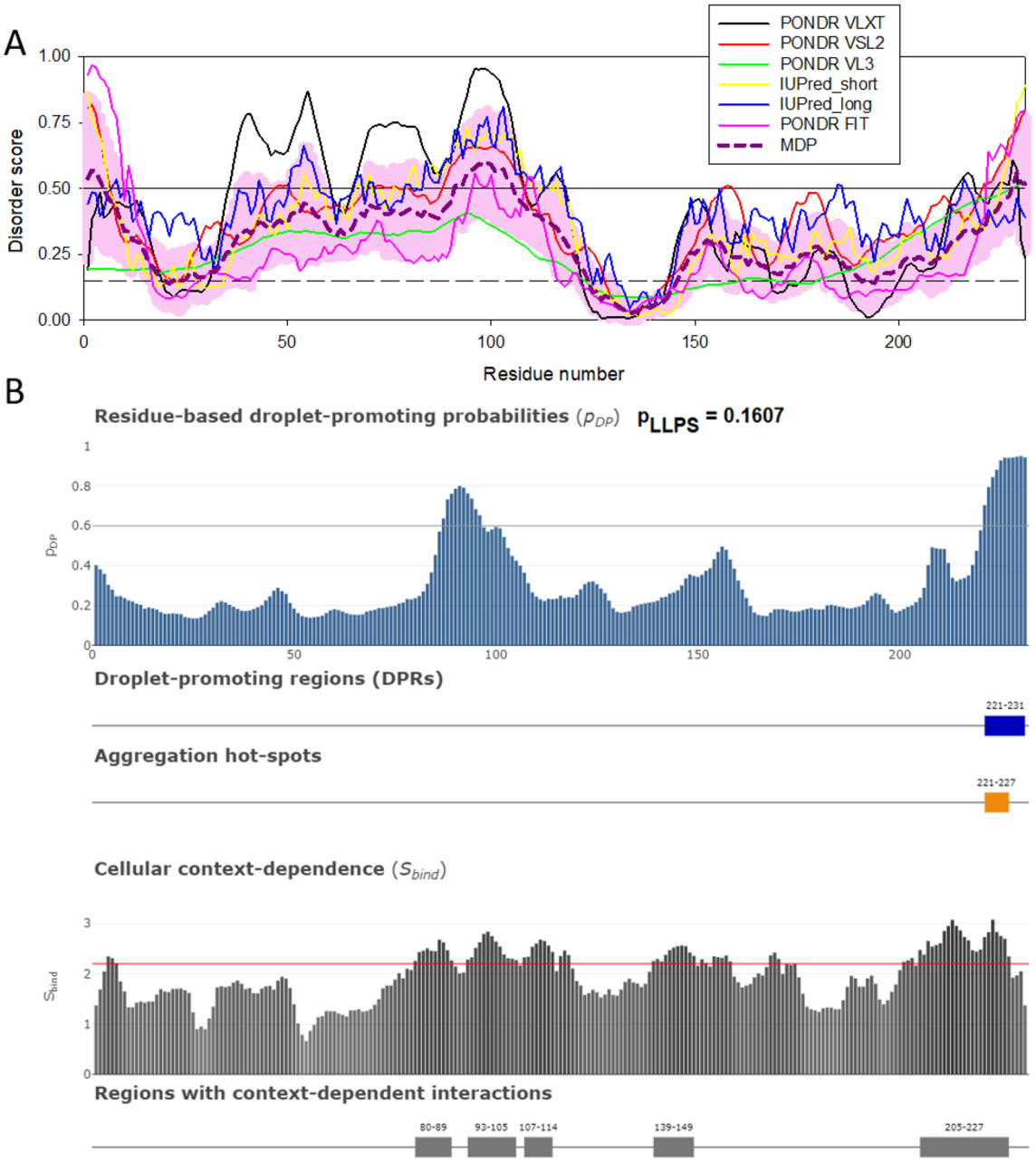


Fig. 4 Disorder analysis of the HIV-1 capsid protein (UniProt ID: P12493; residues 133–363). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]

the fragmented viral capsid while transiting inside the nucleus [88, 89], another possibility is that the fragments of the viral capsid or CA monomers formed after viral uncoating can have a role in maintaining the HIV-MLOs once established in the nucleus.

Reverse transcription

HIV-1 MLOs contain the viral RNA genome that is reverse transcribed in these organelles to generate the viral DNA [75, 90] the only viral form capable of integration into the host chromatin. Intriguingly, HIV-1 MLOs could act as bioreactor to preserve viral cores intact that represent a confined space where the reverse transcription has been shown to occur in vitro [91] or to keep multiple reverse transcription (p66/p51) heterodimers together, likely, acting in the context of a condensate. Another possibility is that HIV-1 MLOs keep multiple cores together and that the reverse transcription occurs within them. Running bioinformatics analysis, we observed that the reverse transcriptase enzyme is expected to contain several IDRs (Fig. 5A) and possesses three droplet-promoting regions (Fig. 5B). However, the overall droplet-promoting probability of this protein is low (below the threshold of 0.6, see Table 1). Therefore, these observations suggest that reverse transcriptase has a low probability to induce LLPS but likely serves as a client of liquid droplets, as shown recently [92].

Integration

HIV-MLOs also act as sites to pilot the HIV-1 genome in active chromatin regions. All around HIV-1 MLOs there is active chromatin similarly to the open chromatin that surrounds NSs. Active proviruses have been found to mainly localize in lens epithelium–derived growth factor (LEDGF) enriched chromatin regions near HIV-1 MLOs. LEDGF appears as clusters in the nucleus in macrophage-like cells and they are mainly distributed near HIV-1 MLOs when cells are infected [75]. This is in line with the calculated P_{LLPS} value ~ 1 (Fig. 6 and Table 1). Contrary to LEDGF, IN clusters are only visible inside HIV-1-MLOs but due to their low tendency to form condensates (IN p_{LLPS} value ~ 0.1), (Fig. 7 and Table 1) they likely serve as MLO clients [92] condensed within the HIV-MLO cores, while they are not associated with LEDGF when observed by imaging [75]. Indeed, highly active chromatin regions (type A compartment) are physically associated with NSs [93, 94], while repressive chromatin regions (type B compartment) associate with the nuclear periphery or the nucleolus [95, 96]. Importantly, HIV-1 proviruses have been found in speckle-associated domains (SPADs) by genomic [97] and nuclear spatial imaging analysis [75]. In line with these findings, it is known that active proviruses are excluded from the type B chromatin compartment, indeed HIV-1 integration in the human genome favours active genes and local hotspots [98]. A recent study shows that LLPS plays a key role in selecting HIV-1 hotspots to ensure efficient viral replication [75].

Late replication steps in HIV-1

Transcription and proviral DNA reactivation

The engagement of *cis*-acting transcription factors with their cognate sequences in promoter sequences has taken on a new understanding. While transcriptional initiation generally involves several sequential events, it now is understood to involve inhibitory and stimulatory complexes [6, 99–101]. The late steps of retroviral replication involve proviral gene transcription driven by the long-terminal repeat (LTR) that harbours multiple *cis*-acting sequences and responsive to in large part host transcription factors. Several recent indicators show that proviral DNA transcription from latently infected cells also relies on phase separation of multi-component complexes. LLPS is also a key factor in the reactivation of HIV-1 proviral DNA transcription reactivation from latency, implicating components of the host histone chaperone chromatin assembly factor 1 and polycomb repressive complex 1 at the HIV-1 promoter [102]. Interestingly, phase separating fused-in-sarcoma and transcription super elongation complex component, AF4/FMR2 family member 4 (AFF4), both known to act via phase separation [103], act to suppress HIV-1 reactivation [104]. Few details exist on the precise involvement of condensates in HIV-1 gene transcription in actively replication cells, but in light of work in other systems that show the engagement of RNA Pol II processivity via positive transcription factor b (pTEFb) and AFF4 condensation [103] is suggestive of a role in HIV-1 DNA transcription. Co-transcriptional retroviral RNA condensates have not been identified except that there is some visible evidence of Rev-responsive element (RRE)-containing RNA-CRM1-Rev complexes that are able to recover following photobleaching using fluorescence recovery after photobleaching techniques [105] and recent work shows that Gag-containing unspliced RNA complexes are found in the cell nucleus [106]. These condensate-like ribonucleoproteins make their way into the cytoplasm through the nuclear pore complex, which also represents a liquid barrier compartmentalizing the cytoplasm from the nucleoplasm [107], as described above. Interestingly, studies on unrelated viruses show that cellular stress provokes activation of persistent viral infection mediated by biomolecular condensates in the host cytoplasm. It has been recently reported that phosphoprotein of mumps virus

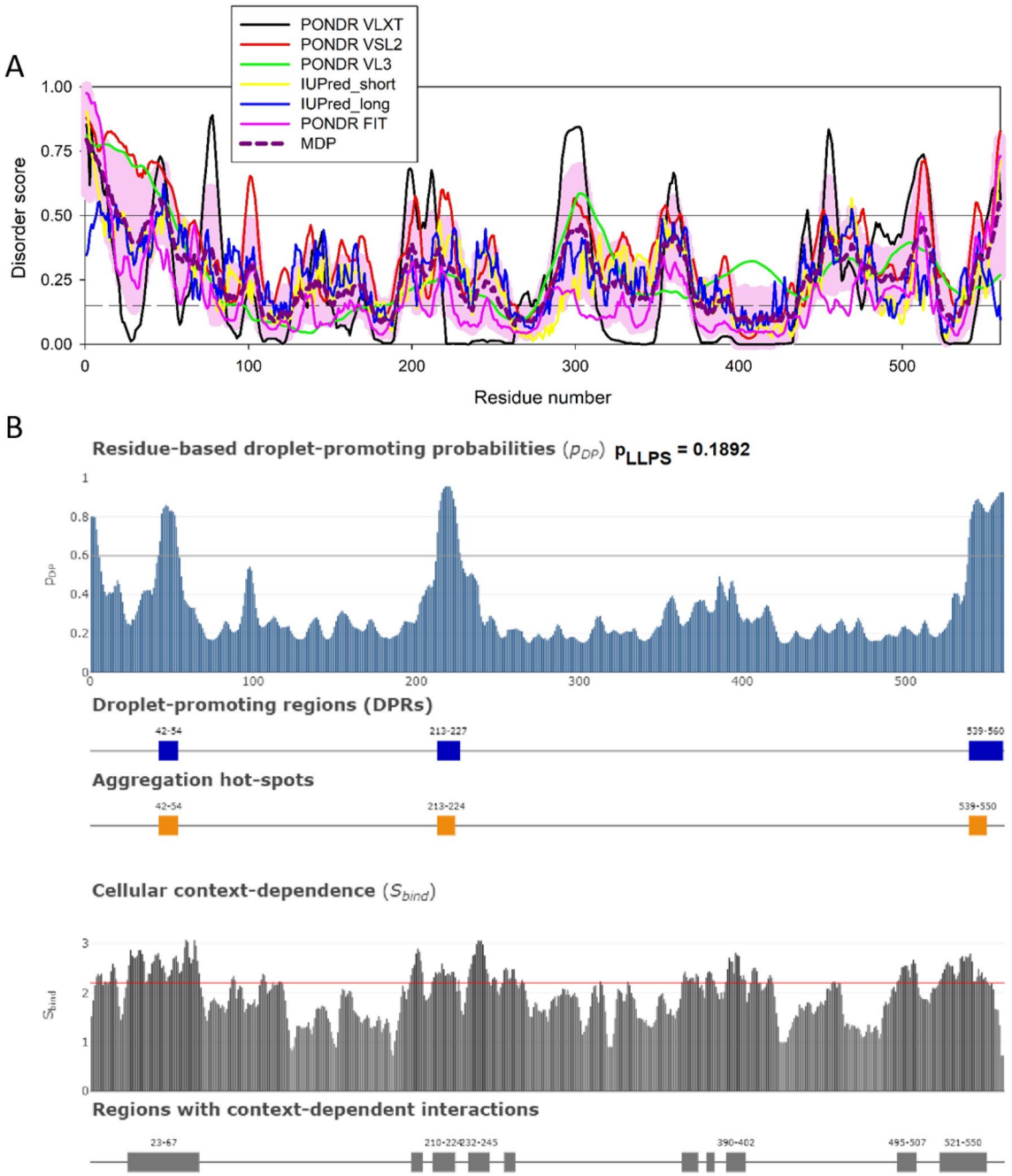


Fig. 5 Disorder analysis of the HIV-1 reverse transcriptase (UniProt ID: P12497; residues: 588–1147). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]

coincides with the phase separation of viral polymerase into pre-formed biomolecular condensates and the formation of a stable replication machinery [108]. Future work should reveal additional means by which retroviruses assemble condensates via LLPS for their nuclear intermediate replication steps.

Assembly, maturation and budding

Classically, following nuclear export, the viral mRNAs are translated to viral structural, polymerase, accessory (auxiliary), and regulatory gene products for final assembly of virus particles. Depending on cell type, these components need to traffic to sites of assembly on membranes

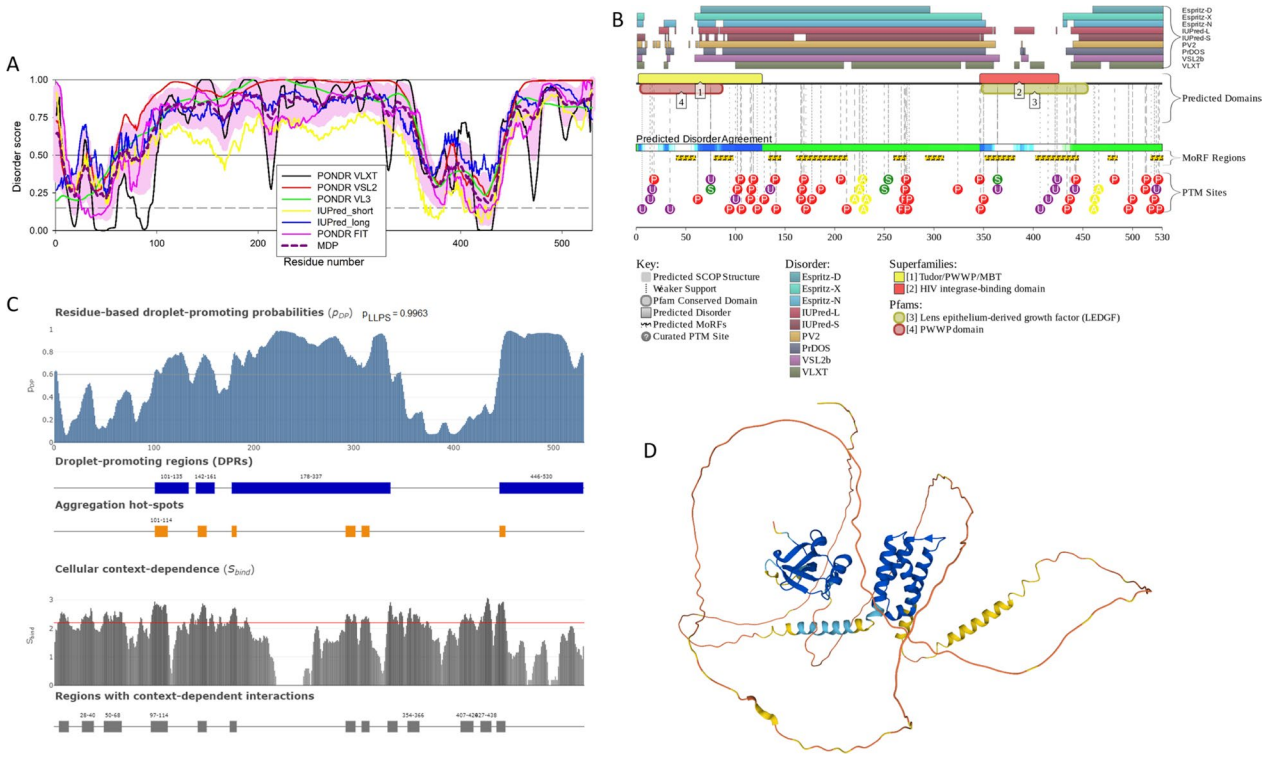


Fig. 6 Functional disorder analysis of human LEDGF/p75 protein (UniProt ID: O75475). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Functional disorder profile generated by the D^2P^2 platform [139]. Here, phosphorylation, ubiquitylation, sumoylation, and acetylation sites are shown by red, violet, green and yellow circles with characters P, U, S, and A, respectively. An interactive profile generated for this protein by D^2P^2 is available at https://d2p2.pro/view/sequence/up/O75475. **C** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]. **D** Model of the LEDGF/p75 structure generated by AlphaFold [147]

and assemble in a stochiometric manner. Few details are understood about the implication of phase separation in the assembly of these proteins; nevertheless bioinformatic analyses have revealed a propensity of viral proteins to phase separate and form phase separated droplets [109]. Full-length, precursor 55 kDa Gag (pr55Gag) on its own displays several intrinsically disordered regions that experimentally, is reflected in droplet formation in vitro [53] (Fig. 8 and Table 1), that bear characteristics of phase separated compartments such that they fuse and fission on coverslips or in the case of Rous sarcoma virus, form liquid phase, mobile droplets in the nucleus [110, 111].

Later during virus assembly, as Gag recruits ESCRT complexes to induce membrane curvature, the viral protease auto-activates and begins processing the viral precursor proteins pr55Gag and pr160Gag/Pol during virus maturation. Interestingly, the addition of recombinant proteins making up a viral core (NC, IN, RT, vRNA, CA) co-assemble to attain a geometry characteristic of a *bona fide* core [92]. Processing of pr55Gag in cells with a recombinant and active viral protease resulted in abundant mobile, condensates in cells [92]. While NC was shown to co-assemble PR-containing condensates late at the end of budding, in elegant work [112], NC condensates exhibited strong "scaffolding" activity in this context (Fig. 9 and Table 1), with the core proteins RT, IN, PR (and the viral genomic RNA) as clients, relying on the scaffolding behaviour of HIV-1 NC in cells [92] consistent with work from the Lyonnais and Mirambeau labs [112] and predicted high capacity to form droplets (p_{LLPS} = 0.6502, Fig. 9B and Table 1). In this particular case, bioinformatic predictions are reflected by in vitro observations made in cells [92].

What's a virus without its host? Intrinsic disorder and LLPS propensities of host factors

Analysis of the intrinsic disorder propensity and LLPS potential of the 4,792 human proteins interacting with HIV-1 proteins, revealed that a very significant part of these HIV-1-interacting partners (> 36%) is predicted to be highly disordered, with even more of these proteins predicted to be moderately disordered (Fig. 10). Although these values are rather comparable to those found for the entire human proteome (e.g., see [113]), the HIV-1 interactome contains significantly less highly ordered proteins (0.02% vs. 0.04% and 2.98% vs. 5.1% in

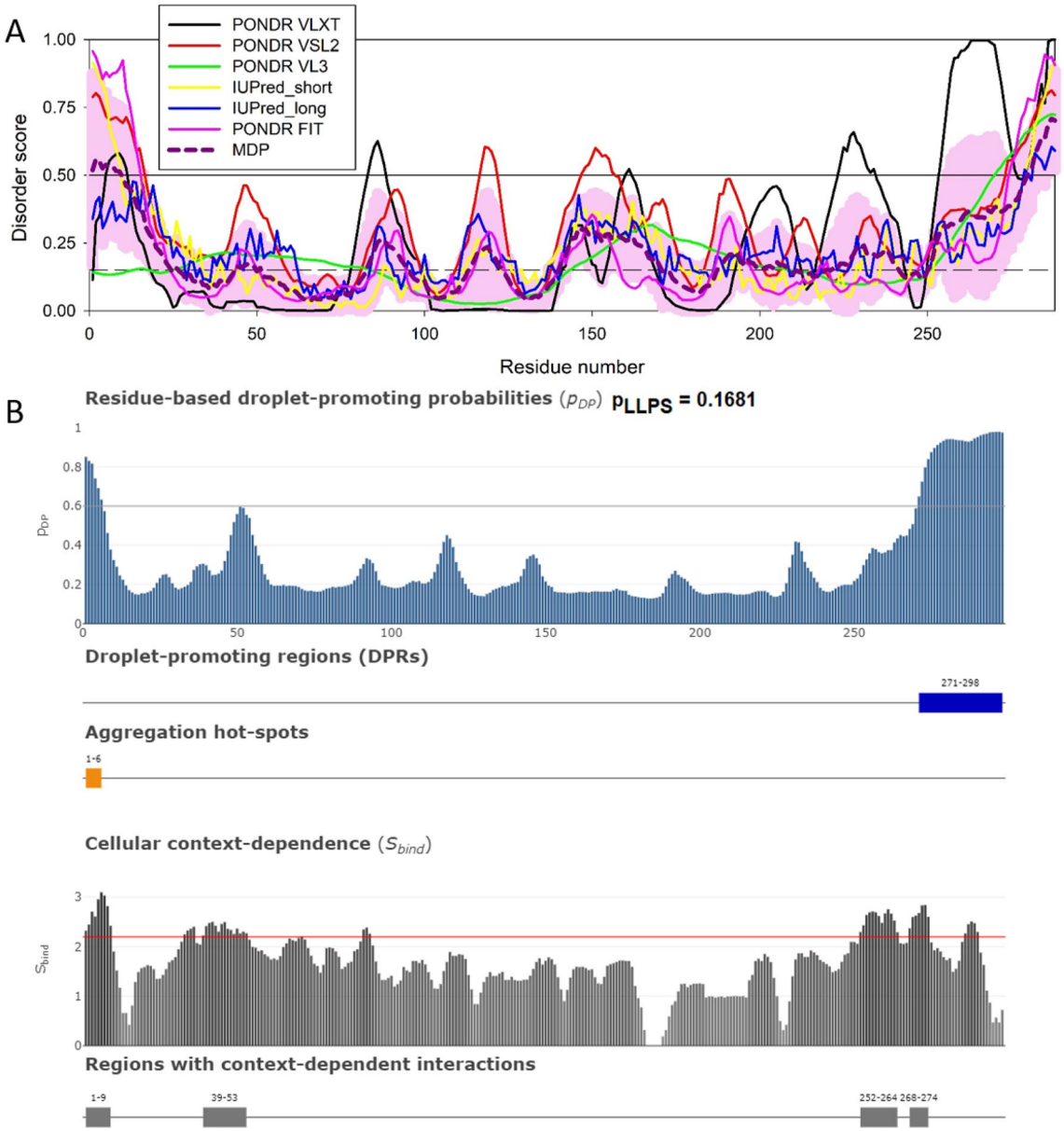


Fig. 7 Disorder analysis of the HIV-1 integrase IN protein (UniProt ID: Q79666; residues 1149-1446). **A**. Per-residue disorder profile generated by RIDAO [122]. **B**. Evaluation of LLPS predisposition, the presence of droplet-promoting regions (DPRs), and existence of regions with context-dependent disorder [129]

the dark-blue and light-blue segments of HIV-1 interactors and whole human genome, respectively), suggesting that the functionality associated with intrinsic disorder might be important for these host proteins. Table 2 represents the results of the evaluation of global intrinsic disorder predisposition of human proteins from several datasets related to the HIV-1 infection. Analyzed datasets include whole human proteome (20,317 proteins; https://www.uniprot.org/proteomes/UP000005640, accessed on 09.24.2022), 4792 human proteins interacting with HIV-1, 252 CA interactors, 221 Gag-Pol interactors, 210 Integrase interactors, 657 Pr44Gag interactors, 106 Integrase interactors, 1667 Tat interactors, and 657 Vpr interactors. Information about human proteins in these datasets was retrieved from the HIV-1 Interaction Database accessible at the National Institutes of Health: National Center for Biotechnology Information (see https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/, accessed on 12.01.2022).

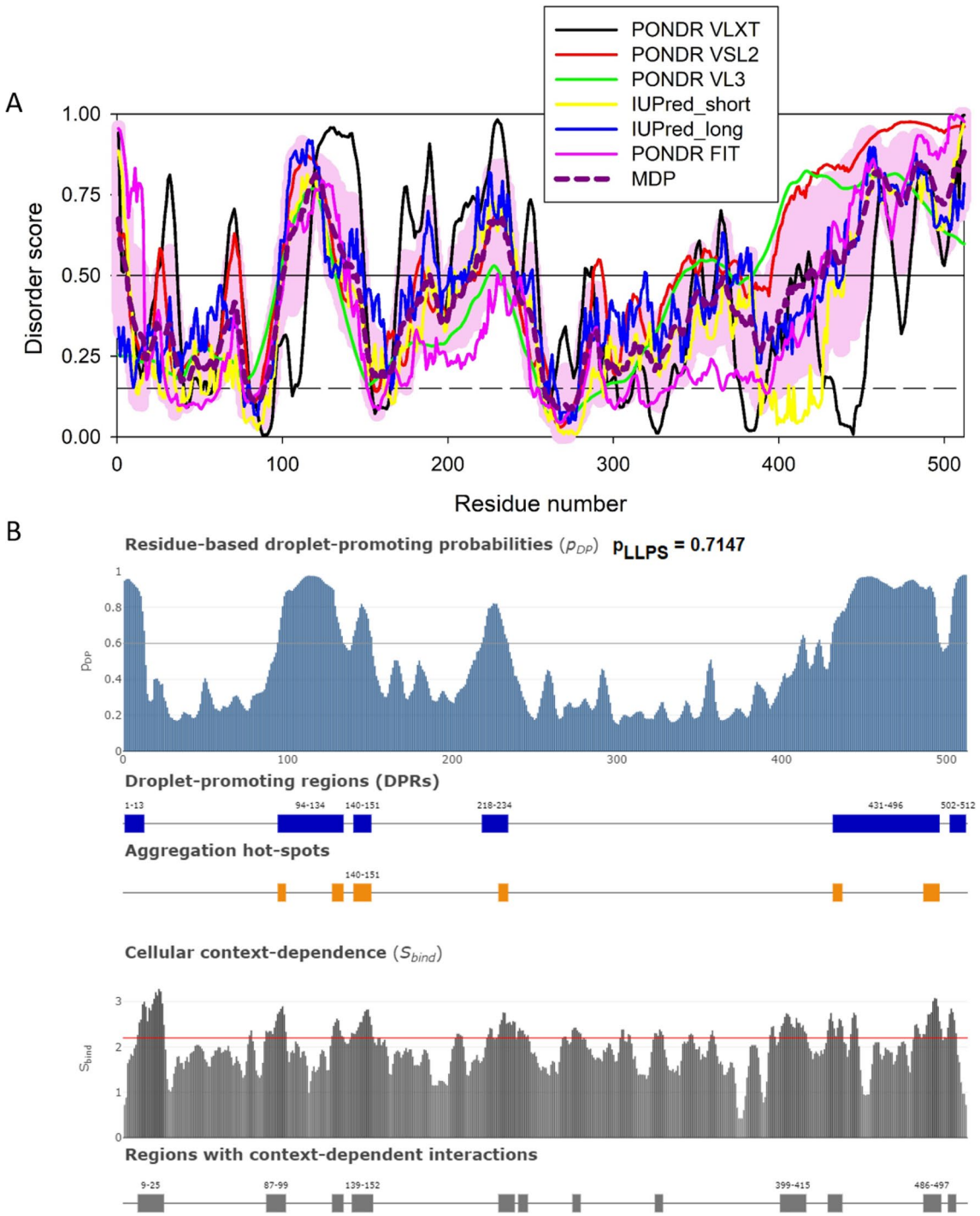


Fig. 8 Disorder analysis of the HIV-1 precursor 55 kDa Gag protein (UniProt ID: P12493). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]

As it was already emphasized, one of the crucial functions of intrinsically disordered proteins and regions is their involvement in LLPS and associated biogenesis of MLOs [16, 18, 19, 114, 115]. Figure 11 shows that 1670 human proteins interacting with HIV-1 (34.8%) are predicted to have propensity for phase separation higher than the 0.6 threshold, and therefore are expected to serve as scaffolds or drivers of phase separation. About

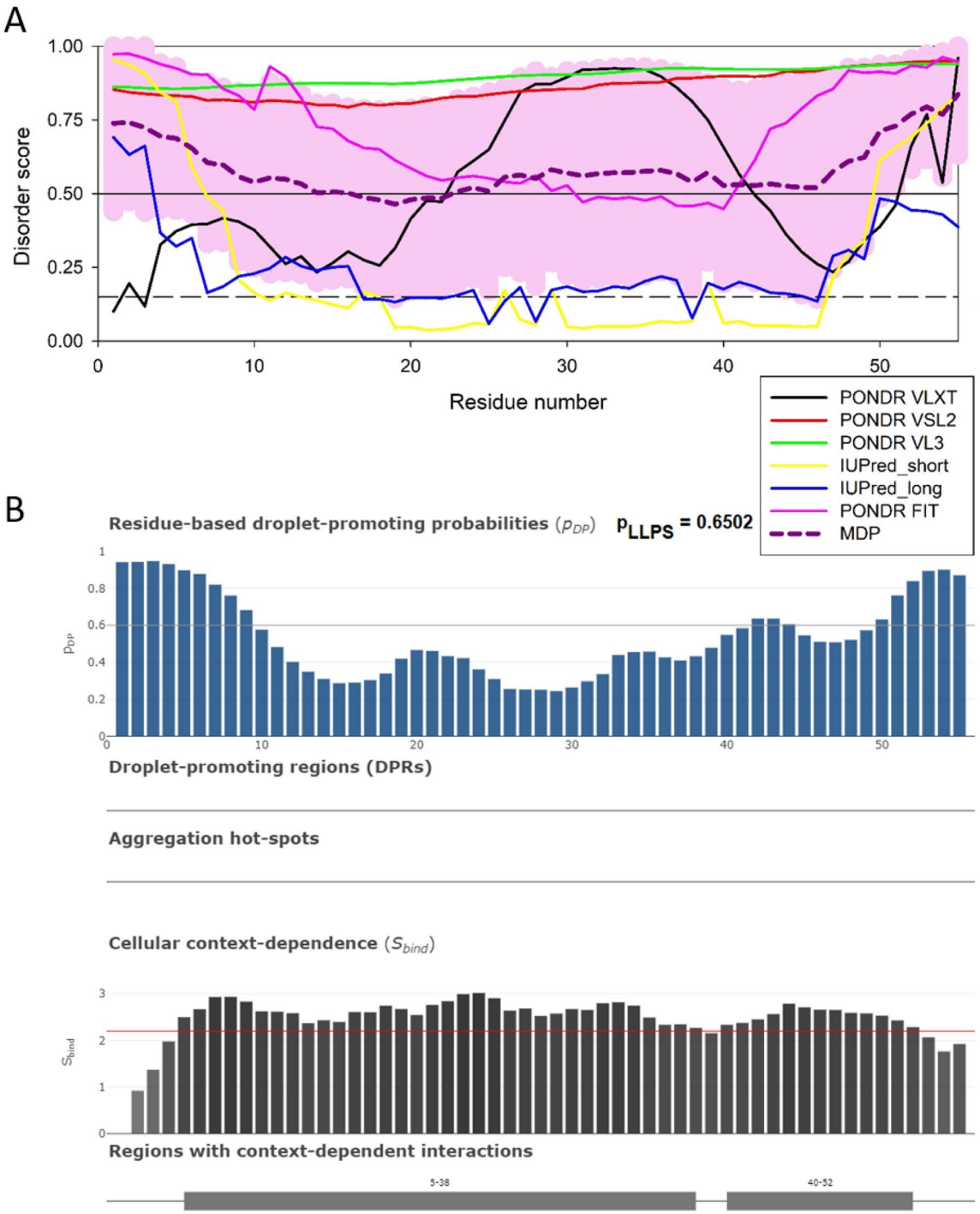
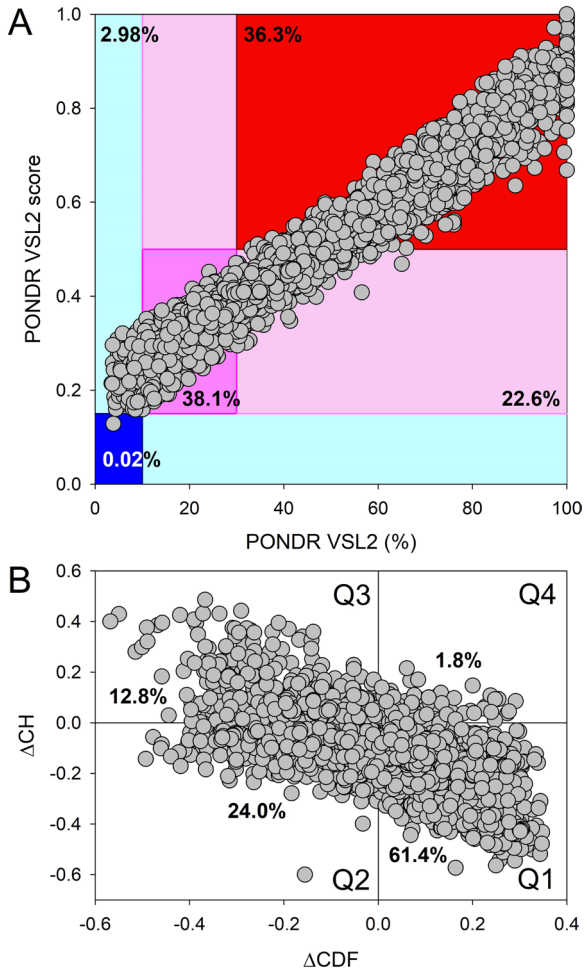


Fig. 9 Disorder analysis of the HIV-1 nucleocapsid NC protein (UniProt ID: P12497; residues 378–432). **A** Per-residue disorder profile generated by RIDAO [138]. **B** Evaluation of LLPS predisposition and existence of regions with context-dependent disorder [145]

40% of these LLPS drivers are predicted to be highly disordered (their mean disorder scores exceed the threshold of 0.5), and most of them (1037 or 62%) are expected to contain more than 30% of disordered residues. Most of the LLPS drivers contain DPRs (in fact, only 14 of these proteins or 0.8% do not have DPRs), with number of DPRs can be as high as 48 (84 LLPS drivers have at least 10 DPRs). More than half of the remaining 3122 HIV-1-interactors (1674 or 53.5%) were predicted do not have any DPRs, whereas 1448 proteins from this dataset



◀ **Fig. 10** Overall disorder analysis of 4792 human proteins interacting with all HIV-1 proteins. Information about these host proteins was retrieved from the HIV-1 Interaction Database accessible at the National Institutes of Health: National Center for Biotechnology Information (see https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/). **A** PONDR® VSL2 mean disorder score vs. PONDR® VSL2 (%) plot. Here, each point corresponds to a query protein, coordinates of which are evaluated from the corresponding PONDR® VSL2 data as its mean disorder score (MDS) and percent of the predicted intrinsically disordered residues (PPIDR). MDS was calculated for each query protein as a protein length-normalized sum of all the per-residue disorder scores, whereas PPIDR was calculated as a protein length-normalized number of residues with the disorder scores of at least 0.5 multiplied by 100%. Color blocks are used to visualize proteins based on the accepted classification, with red, pink/light pink, and blue/light blue regions containing highly disordered, moderately disordered, and ordered proteins, respectively. Classification is based on the accepted practice, where a PPIDR value of less than 10% is taken to correspond to a highly ordered protein, PPIDR between 10 and 30% is ascribed to moderately disordered protein, and PPIDR greater than 30% corresponds to a highly disordered protein [148, 149]. Proteins can also be grouped based on their corresponding MDS values, being classified as highly ordered (MDS < 0.15), moderately disordered of flexible (MDS between 0.15 and 0.5) and highly disordered (MDS ≥ 0.5). Dark blue or pink regions correspond to the regions, where PPIDR agrees with MDS, whereas areas in which only one of these criteria applies are shown by light blue or light pink. Numbers within these segments reflect their protein content. **B** CH-CDF plot, where coordinates for a protein are calculated as the average distance of its CDF curve from the CDF boundary (X axis) and its distance from the CH boundary. Protein classification is based on the quadrant, where it is located: Q1, protein predicted to be ordered by both predictors. Q2, protein predicted to be ordered to by CH-plot and disordered by CDF. Q3, protein predicted to be disordered by both predictors. Q4, protein predicted to be disordered by CH-plot and ordered by CDF. CH-CDF analysis combines results from the charge-hydropathy (CH) and cumulative distribution function (CDF) plots, which are both binary predictors of disorder. For CH, net charge is plotted versus hydropathy for each protein [150]. Due to the observation that disordered proteins tend to have high net charge and low hydropathy, disordered and ordered proteins cluster two regions of the plot. A linear boundary separates disordered and ordered proteins [150, 151]. Proteins that are disordered appear above the boundary while ordered proteins appear below [150, 151]. In the CDF-plot predictor, PONDR scores for each residue of a single protein is plotted against their frequency within the sequence. If a CDF curve of a given protein is below the order–disorder boundary, this protein is considered disordered, and protein is ordered if the CDF curve is located above this boundary [151]. Data generated by CH- and CDF-plots are then combined to generate ΔCH-ΔCDF plot, which enables rapid discrimination between flavors of disorder [152, 153]. To this end, for each query protein, ΔCH, the vertical distance of the corresponding point in CH plot from the boundary, is calculated, whereas ΔCDF is computed as the average distance between the order–disorder boundary and the CDF curve. Then, ΔCH is plotted against ΔCDF resulting in a CH-CDF plot. Proteins in the top-left quadrant are predicted to be disordered by both CH and CDF, the ones in the bottom-left are predicted to be ordered by CH and disordered by CDF, the ones in the top-right are predicted to be disordered by CH and ordered by CDF, and in the bottom-right quadrant predicted to be ordered by both [152, 153]

(or 30.2% of the HIV-1 interactome) are predicted to have 1 to 5 DPRs and therefore are expected to serve as droplet clients. In other words, more than 65% of HIV-1-interactors are expected to be engaged in LLPS. On the other hand, in the entire human proteome containing 20,367 proteins), 7570 and 5252 proteins (37.2% and 25.8%) are expected to act as LLPS drivers and droplet clients, respectively. Therefore, the HIV-1 interactome contains less LLPS drivers and more droplet clients than those found in the human proteome. Curiously, 1448 droplet clients found in HIV-1 interactome are expected to be more disordered than 1674 HIV-1 interactors not involved in droplet biogenesis. This prediction is in agreement with the comparison between IN and CPSF6 clusters in the infected cells for their ability to form condensates. BMCs are highly dynamic, and this feature can be measured by fluorescence recovery after photobleaching. Results clearly show that IN proteins are unable to recover after photobleaching, contrary to CPSF6 proteins

Table 2 Evaluation of global intrinsic disorder predisposition of human proteins from several datasets related to HIV-1 infection

Dataset (number of proteins)	PONDR® VSL2 analysis				CH-CDF analysis			
	Mostly ordered[a]	Moderately disordered[b]	Mostly disordered[c]	Highly disordered[d]	Q1	Q2	Q3	Q4
Whole proteome (20,317)	83(0.4%) 1030(5.1%)	6841(33.7%) 4267(21.0%)	8095(39.8%)	7381(36.3%)	12,014 (59.1%)	5177 (25.5%)	2501 (12.3%)	625 (3.1%)
All human interactors (4792)	1(0.02%) 143(2.98%)	1828(38.1%) 1082(22.6%)	1738(36.3%)	1534(32.0%)	2945 (61.4%)	1149 (24.0%)	612 (12.8%)	86 (1.8%)
Capsid interactors (252)	0(0.0%) 12(4.7%)	98(38.9%) 68(27.0%)	74(29.4%)	65(25.8%)	180 (71.4%)	49 (19.5%)	21 (8.3%)	2 (0.8%)
GagPol interactors (221)	0(0.0%) 6(2.7%)	99(44.8%) 32(14.5%)	84(38.0%)	75(33.9%)	122 (55.2%)	35 (15.8%)	59 (22.7%)	5 (2.3%)
Integrase interactors (210)	0(0.0%) 8(3.8%)	95(45.2%) 50(23.8%)	57(27.2%)	54(25.7%)	143 (68.1%)	37 (17.6%)	28 (13.3%)	2 (1.0%)
Pr44Gag interactors (657)	0(0.0%) 20(3.0%)	264(40.2%) 126(19.2%)	247(37.6%)	226(34.4%)	374 (56.9%)	140 (21.3%)	130 (19.8%)	13 (2.0%)
RT Interactors (106)	0(0.0%) 2(1.9%)	58(54.7%) 21(19.8%)	25(23.6%)	21(19.8%)	78 (73.6%)	19 (17.9%)	9 (8.5%)	0 (0.0%)
Tat Interactors (1667)	0(0.0%) 42(2.5%)	617(37.0%) 420(25.2%)	588(35.3%)	508(30.5%)	1034 (62.0%)	380 (22.8%)	234 (14.0%)	19 (1.2%)
Vpr interactors (657)	0(0.0%) 19(2.9%)	272(41.4%) 166(25.3%)	200(30.4%)	173(26.3%)	432 (65.8%)	140 (21.3%)	73 (11.1%)	12 (1.8%)

[a] Mostly ordered proteins are those found in the dark blue and light blue areas of the mean disorder score (MDS) vs. percent of predicted intrinsically disordered residues (PPIDR) plots; i.e., areas containing proteins with PPIDR < 10% and MDS < 0.15, where dark blue correspond to area where these two parameters agree (upper values in the corresponding cells), whereas light blue color show areas in which only one of these criteria applies (bottom values in the corresponding cells)

[b] Moderately disordered proteins are those found in the dark pink and light pink areas of the mean disorder score (MDS) vs. percent of predicted intrinsically disordered residues (PPIDR) plots; i.e., areas containing proteins with 10% ≤ PPIDR < 30% and 0.15 ≤ MDS < 0.1, where dark pink correspond to the area where these two parameters agree (upper values in the corresponding cells), whereas light pink color show areas, in which only one of these criteria applies (bottom values in the corresponding cells)

[c] Mostly disordered proteins are those found in the red area of the mean disorder score (MDS) vs. percent of predicted intrinsically disordered residues (PPIDR) plots; i.e., area containing proteins with PPIDR ≥ 30% and MDS ≥ 0.5

[d] Highly disordered proteins are a subset of mostly disordered proteins that are characterized by PPIDR ≥ 50% and MDS ≥ 0.5

that show high mobility giving rise to the restoration of the fluorescent droplet after photobleaching [75].

Therapeutics and perspectives of targeting MLOs

Biomolecular condensates formed during virus replication or those formed aberrantly in other human diseases represent new and emerging targets for therapeutic intervention [116–118] (for recent reviews on phase separation and viruses, please see: [34, 39, 119–121]). Further understanding of LLPS in retroviral replication will lay the foundation for future study of retroviral MLOs that exhibit fluid-like behaviour during replication. Long-term goals in LLPS virology research include the screening for compounds that modulate the stability of viral condensates [122–126], such as those identified to rigidify respiratory syncytial virus BMCs [127] or those that target scaffolding SARS-CoV-2 N condensates to evade immune signalling [128–130] or finding mutants in HIV-1 core proteins that lead to aberrant condensation such as those found to be etiologic for human disease [118, 131]. Based on our current understanding of HIV-1 biology for example, it follows that the HIV-1 BMC makes its way through a nuclear pore condensate and then into a nuclear condensate, the nuclear speckle to complete reverse transcription [90] (Fig. 1). Integration is then achieved with the help of LEDGF/p75 [132, 133] that form condensates near HIV-1 MLOs in macrophage-like cells [75]. LEDGF is a host protein that possesses intrinsically disordered domains to form clusters (Fig. 6). While several retroviral NC proteins form droplets in vitro [53], it remains to be determined if condensation through LLPS plays a fundamental role in other retroviruses. Future perspectives for HIV-1 will include the study on how condensation plays roles in early pre-integration, integration steps [34, 75, 90] as well late in HIV-1 replication in transcription, considering that recently characterized BMCs involving host histone chaperone chromatin assembly factor 1 and polycomb repressive complex 1 and their influence on HIV-1 reactivation from latency, acting near the HIV-1 promoter [102, 134]. Targeting viruses with BMC therapeutics will be multi-pronged since virus infection, including HIV-1, exacerbates the outcomes of neurodegenerative

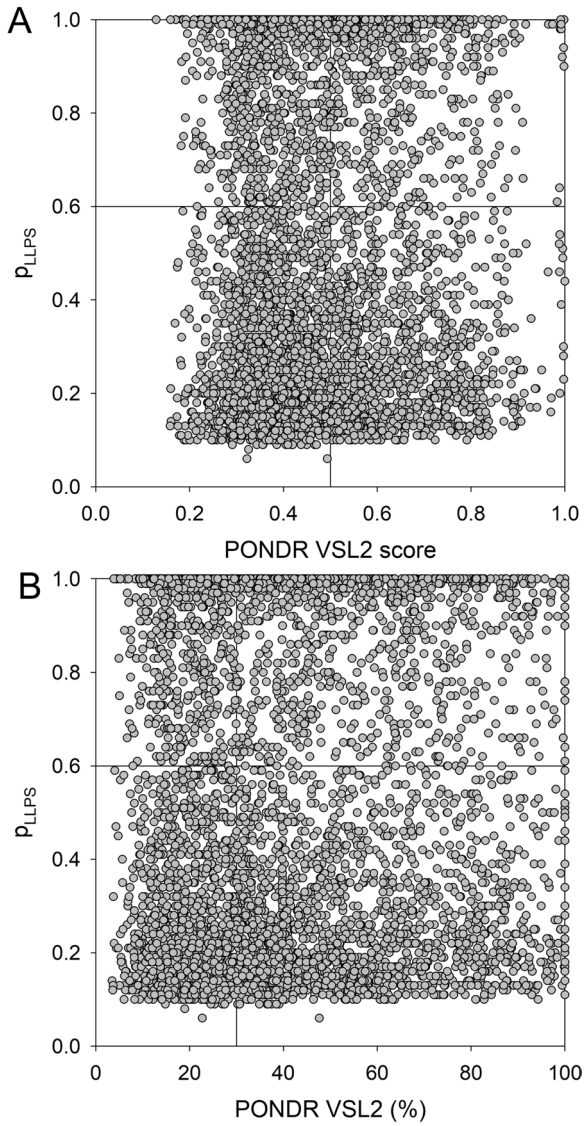


Fig. 11 Intrinsic disorder and phase separation propensity of 4792 human proteins interacting with all HIV-1 proteins. **A** p_{LLPS} vs. PONDR® VSL2 disorder score dependence. In these analyses, proteins with $p_{LLPS} \geq 0.6$ are expected to serve as LLPS drivers, whereas proteins with the disorder scores ≥ 0.5 are predicted as mostly disordered. Among the 4792 human proteins interacting with all HIV-1 proteins, 1738 proteins (36.27%) are expected to be mostly disordered, and 1669 (34.83%) are expected to serve as LLPS drivers. Among the LLPS drivers, 655 proteins (39.24%) are mostly disordered, indicating that 13.67% of HIV-1-interactors are mostly disordered proteins driving LLPS. **B** Dependence of the p_{LLPS} on the percent of disordered residues predicted by PONDR® VSL2. Proteins with $p_{LLPS} \geq 0.6$ are expected to serve as LLPS drivers, whereas proteins with the PPIDR $\geq 30\%$ are predicted as highly disordered. Based on the outputs of these analyses, 2820 human proteins from HIV-1 interactome (58.84%) are expected to be highly disordered. Among 1669 of LLPS drivers, 1038 proteins (62.19%) are expected to be highly disordered, indicating that 21.66% of HIV-1-interactors are highly disordered LLPS drivers

diseases that are often dependent on aberrant BMC assemblies [135, 136].

Conclusions

Ultimately, the characterization of biomolecular condensates in many viruses, and not just retroviruses, have revealed mechanistic ways in which viruses mimic, highjack or usurp host membraneless organelles. In retroviruses, many of the bioinformatics analyses on disorder and whether viral proteins are drivers/scaffold or client of biomolecular condensates shown here are reflected by recent experimental observations. On the other hand, many of the proposed roles for condensates proposed here await confirmation by further in vitro and in cell assays, and biophysical analyses including roles for condensation through phase separation during viral assembly on membranes, during nuclear entry of core components and early and late nuclear events. The plasticity of BMCs should seamlessly orchestrate various replication steps (Fig. 1), especially enzymatic activities (e.g., PR, RT, IN) [112] and may allow for virus-host assemblies and interactions through engagement and disengagement [137] that are obligatory throughout the replication cycles of retroviruses and other viruses.

Abbreviations

AFF4	AF4/FMR2 family member 4
BMCs	Biomolecular condensates
CA	Capsid protein
COVID-19	Coronavirus disease-2019
CPSF6	Cleavage and polyadenylation specific factor 6
CRM1	Chromosomal maintenance 1 (or Exportin 1)
DNA	Deoxyribonucleic acid
DPR	Droplet-promoting region
Env	Envelope protein
FG	Phenylalanine-glycine
G3BP1	Ras GTPase-activating protein-binding protein 1
Gag	Group specific antigen or Gag polyprotein
HIV-1	Human immunodeficiency virus-type 1
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IN	Integrase protein
LAD	Lamin associated domain
LEDGF/p75	Lens epithelium-derived growth factor
lncRNA	Long non-coding RNA
LLPS	Liquid–liquid phase separation
LTR	Long-terminal repeat
MCD	Mixed charge domain
MDS	Mean disorder score
MLO	Membraneless organelles
MoRF	Molecular recognition feature
N	Nucleocapsid (SARS-CoV-2)
NC	Nucleocapsid protein (HIV-1)
NPC	Nuclear pore complex
NSs	Nuclear speckles
Nup	Nucleoporin
Nup153	Nucleoporin 153
p_{LLPS}	Probability of liquid–liquid phase separation
pr55Gag	Precursor 55 kDa Gag
PONDR	Predictor of natural disordered regions
PPIDR	Predicted intrinsically disordered residues

pTEFb	Positive transcription elongation factor b
PTM	Posttranslational modification
RBD	RNA-binding domain
RNA	Ribonucleic acid
RRE	Rev-responsive element
RRM	RNA-binding motif
rRNA	Ribosomal RNA
RT	Reverse transcriptase enzyme
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SPAD	Speckle-associated domain
vRNA	Viral RNA
ZnF	Zinc finger domain

Acknowledgements
We thank Bao-An Chau and Venessa Chen for artwork.

Author contributions
All authors contributed equally to this review in the writing and the revision. All authors read and approved the final manuscript.

Funding
F.D.N. is supported by *L'Agence nationale de recherche sur le SIDA et les hépatites virales—maladies infectieuses émergentes* (ANRS-MIE; Grant# ECTZ192036; ECTZ88162) and A.J.M. is supported by Grant FRN-162447 from the Canadian Institutes of Health Research.

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors consent to publish.

Competing interests
The authors declare no competing interests.

Received: 7 December 2022 Accepted: 16 March 2023
Published online: 07 April 2023

References

1. Alberti S, Gladfelter A, Mittag T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. Cell. 2019;176:419–34.
2. Anderson P, Kedersha N. RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat Rev Mol Cell Biol. 2009;10:430–6.
3. Kedersha N, Anderson P. Mammalian stress granules and processing bodies. Methods Enzymol. 2007;431:61–81.
4. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, et al. Protein phase separation: a new phase in cell biology. Trends Cell Biol. 2018;28:420–35.
5. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18:285–98.
6. Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science. 2018;361:eaar3958.
7. Stortz M, Pecci A, Presman DM, Levi V. Unraveling the molecular interactions involved in phase separation of glucocorticoid receptor. BMC Biol. 2020;18:59.
8. Ong JY, Torres JZ. Phase separation in cell division. Mol Cell. 2020;80:9–20.
9. Wei MT, Chang YC, Shimobayashi SF, Shin Y, Strom AR, Brangwynne CP. Nucleated transcriptional condensates amplify gene expression. Nat Cell Biol. 2020;22:1187–96.
10. Brodsky S, Jana T, Mittelman K, Chapal M, Kumar DK, Carmi M, Barkai N. Intrinsically disordered regions direct transcription factor in vivo binding specificity. Mol Cell. 2020;79:459-471.e454.
11. Roden C, Gladfelter AS. RNA contributions to the form and function of biomolecular condensates. Nat Rev Mol Cell Biol. 2021;22:183–95.
12. Slobodin B, Dikstein R. So close, no matter how far: multiple paths connecting transcription to mRNA translation in eukaryotes. EMBO Rep. 2020;21:e50799.
13. Noda NN, Wang Z, Zhang H. Liquid-liquid phase separation in autophagy. J Cell Biol. 2020;219:e202004062.
14. Espinosa JR, Joseph JA, Sanchez-Burgos I, Garaizar A, Frenkel D, Collepardo-Guevara R. Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. Proc Natl Acad Sci USA. 2020;117:13238–47.
15. Lyon AS, Peeples WB, Rosen MK. A framework for understanding the functions of biomolecular condensates across scales. Nat Rev Mol Cell Biol. 2021;22:215–35.
16. Antifeeva IA, Fonin AV, Fefilova AS, Stepanenko OV, Povarova OI, Silonov SA, Kuznetsova IM, Uversky VN, Turoverov KK. Liquid-liquid phase separation as an organizing principle of intracellular space: overview of the evolution of the cell compartmentalization concept. Cell Mol Life Sci. 2022;79:251.
17. Nesterov SV, Ilyinsky NS, Uversky VN. Liquid-liquid phase separation as a common organizing principle of intracellular space and biomembranes providing dynamic adaptive responses. Biochim Biophys Acta Mol Cell Res. 2021;1868:119102.
18. Uversky VN. Recent developments in the field of intrinsically disordered proteins: intrinsic disorder-based emergence in cellular biology in light of the physiological and pathological liquid-liquid phase transitions. Annu Rev Biophys. 2021;50:135–56.
19. Darling AL, Zaslavsky BY, Uversky VN. Intrinsic disorder-based emergence in cellular biology: physiological and pathological liquid-liquid phase transitions in cells. Polymers. 2019;11:990.
20. Turoverov KK, Kuznetsova IM, Fonin AV, Darling AL, Zaslavsky BY, Uversky VN. Stochasticity of biological soft matter: emerging concepts in intrinsically disordered proteins and biological phase separation. Trends Biochem Sci. 2019;44:716–28.
21. Uversky VN. Protein intrinsic disorder-based liquid-liquid phase transitions in biological systems: complex coacervates and membrane-less organelles. Adv Colloid Interface Sci. 2017;239:97–114.
22. Uversky VN, Kuznetsova IM, Turoverov KK, Zaslavsky B. Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. FEBS Lett. 2015;589:15–22.
23. Choi JM, Holehouse AS, Pappu RV. Physical principles underlying the complex biology of intracellular phase transitions. Annu Rev Biophys. 2020;49:107–33.
24. Koga S, Williams DS, Perriman AW, Mann S. Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. Nat Chem. 2011;3:720–4.
25. Poudyal RR, Pir Cakmak F, Keating CD, Bevilacqua PC. Physical principles and extant biology reveal roles for RNA-containing membraneless compartments in origins of life chemistry. Biochemistry. 2018;57:2509–19.
26. Oparin AI. Evolution of the concepts of the origin of life, 1924–1974. Orig Life. 1976;7:3–8.
27. Keating CD. Aqueous phase separation as a possible route to compartmentalization of biological molecules. Acc Chem Res. 2012;45:2114–24.
28. Bergeron-Sandoval LP, Safaee N, Michnick SW. Mechanisms and consequences of macromolecular phase separation. Cell. 2016;165:1067–79.
29. Shin Y, Brangwynne CP. Liquid phase condensation in cell physiology and disease. Science. 2017;357:eaaf4382.
30. Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP. Coexisting liquid phases underlie nucleolar subcompartments. Cell. 2016;165:1686–97.

31. Brangwynne CP, Mitchison TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. Proc Natl Acad Sci USA. 2011;108:4334–9.
32. Aumiller WM Jr, Keating CD. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. Nat Chem. 2016;8:129–37.
33. Ramaswami M, Taylor JP, Parker R. Altered ribostasis: RNA-protein granules in degenerative disorders. Cell. 2013;154:727–36.
34. Scoca V, Di Nunzio F. Membraneless organelles restructured and built by pandemic viruses: HIV-1 and SARS-CoV-2. J Mol Cell Biol. 2021;13:259–68.
35. 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;21:9045.
36. Nikolic J, Le Bars R, Lama Z, Scrima N, Lagaudriere-Gesbert C, Gaudin Y, Blondel D. Negri bodies are viral factories with properties of liquid organelles. Nat Commun. 2017;8:58.
37. Heinrich BS, Maliga Z, Stein DA, Hyman AA, Whelan SPJ. Phase transitions drive the formation of vesicular stomatitis virus replication compartments. MBio. 2018;9:e02290-e2317.
38. 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;10:1629.
39. Etibor TA, Yamauchi Y, Amorim MJ. Liquid biomolecular condensates and viral lifecycles: review and perspectives. Viruses. 2021;13:366.
40. Wubben JM, Atkinson SC, Borg NA. The role of protein disorder in nuclear transport and in its subversion by viruses. Cells. 2020;9:2654.
41. Berry J, Weber SC, Vaidya N, Haataja M, Brangwynne CP. RNA transcription modulates phase transition-driven nuclear body assembly. Proc Natl Acad Sci USA. 2015;112:E5237-5245.
42. Falahati H, Pelham-Webb B, Blythe S, Wieschaus E. Nucleation by rRNA dictates the precision of nucleolus assembly. Curr Biol. 2016;26:277–85.
43. West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, Wang PI, Tolstorukov MY, Kingston RE. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol Cell. 2014;55:791–802.
44. Dundr M, Misteli T. Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol. 2010;2:a000711.
45. Mitrea DM, Kriwacki RW. Phase separation in biology; functional organization of a higher order. Cell Commun Signal. 2016;14:1.
46. Tourriere H, Chebli K, Zekri L, Courselaud B, Blanchard JM, Bertrand E, Tazi J. The RasGAP-associated endoribonuclease G3BP assembles stress granules. J Cell Biol. 2003;160:823–31.
47. Banani SF, Afeyan LK, Hawken SW, Henninger JE, Dall'Agnese A, Clark VE, Platt JM, Oksuz O, Hannett NM, Sagi I, et al. Genetic variation associated with condensate dysregulation in disease. Dev Cell. 2022;57:1776-1788.e1778.
48. Li W, Hu J, Shi B, Palomba F, Digman MA, Gratton E, Jiang H. Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. Nat Cell Biol. 2020;22:960–72.
49. Song MS, Grabocka E. Stress granules in cancer. Rev Physiol Biochem Pharmacol. 2023;185:25–52. https://doi.org/10.1007/112_2020_37.
50. Peng Q, Tan S, Xia L, Wu N, Oyang L, Tang Y, Su M, Luo X, Wang Y, Sheng X, et al. Phase separation in cancer: from the impacts and mechanisms to treatment potentials. Int J Biol Sci. 2022;18:5103–22.
51. 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;7:5.
52. Iserman C, Roden CA, Boerneke MA, Sealfon RSG, McLaughlin GA, Jungreis I, Fritch EJ, Hou YJ, Ekena J, Weidmann CA, et al. Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. Mol Cell. 2020;80:1078-1091.e1076.
53. 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;31:107520.
54. Monette A, Mouland AJ. Zinc and copper ions differentially regulate prion-like phase separation dynamics of pan-virus nucleocapsid biomolecular condensates. Viruses. 2020;12:1179.
55. Gotor NL, Armaos A, Calloni G, Torrent Burgas M, Vabulas RM, De Groot NS, Tartaglia GG. RNA-binding and prion domains: the Yin and Yang of phase separation. Nucleic Acids Res. 2020;48:9491–504.
56. Milev MP, Brown CM, Mouland AJ. Live cell visualization of the interactions between HIV-1 Gag and the cellular RNA-binding protein Staufen1. Retrovirology. 2010;7:41.
57. Alberti S, Dormann D. Liquid-liquid phase separation in disease. Annu Rev Genet. 2019;53:171–94.
58. Alshareedah I, Moosa MM, Raju M, Potoyan DA, Banerjee PR. Phase transition of RNA-protein complexes into ordered hollow condensates. Proc Natl Acad Sci USA. 2020;117:15650–8.
59. Newton JC, Naik MT, Li GY, Murphy EL, Fawzi NL, Sedivy JM, Jogl G. Phase separation of the LINE-1 ORF1 protein is mediated by the N-terminus and coiled-coil domain. Biophys J. 2021;120:2181–91.
60. Asimi V, Sampath Kumar A, Niskanen H, Riemenschneider C, Hetzel S, Naderi J, Fasching N, Popitsch N, Du M, Kretzmer H, et al. Hijacking of transcriptional condensates by endogenous retroviruses. Nat Genet. 2022;54:1238–47.
61. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature. 1970;226:1211–3.
62. Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature. 1970;226:1209–11.
63. Weiss RA. The discovery of endogenous retroviruses. Retrovirology. 2006;3:67.
64. Coffin J, Blomberg J, Fan H, Gifford R, Hatziioannou T, Lindemann D, Mayer J, Stoye J, Tristem M, Johnson W, Ictv Report C. ICTV Virus taxonomy profile: retroviridae 2021. J Gen Virol. 2021;102:001712.
65. Fare CM, Villani A, Drake LE, Shorter J. Higher-order organization of biomolecular condensates. Open Biol. 2021;11:210137.
66. Ilik IA, Malszycki M, Lubke AK, Schade C, Meierhofer D, Aktas T. SON and SRRM2 are essential for nuclear speckle formation. Elife. 2020;9:e60579.
67. Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, Hue S, Fletcher AJ, Lee K, KewalRamani VN, et al. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. PLoS Pathog. 2011;7:e1002439.
68. Di Nunzio F, Danckaert A, Fricke T, Perez P, Fernandez J, Perret E, Roux P, Shorte S, Charneau P, Diaz-Griffero F, Arhel NJ. Human nucleoporins promote HIV-1 docking at the nuclear pore, nuclear import and integration. PLoS ONE. 2012;7:e46037.
69. Matreyek KA, Yucel SS, Li X, Engelman A. Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. PLoS Pathog. 2013;9:e1003693.
70. Di Nunzio F, Fricke T, Miccio A, Valle-Casuso JC, Perez P, Souque P, Rizzi E, Severgnini M, Mavilio F, Charneau P, Diaz-Griffero F. Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. Virology. 2013;440:8–18.
71. Krull S, Dorries J, Boysen B, Reidenbach S, Magnius L, Norder H, Thyberg J, Cordes VC. Protein Tpr is required for establishing nuclear pore-associated zones of heterochromatin exclusion. EMBO J. 2010;29:1659–73.
72. Lelek M, Casartelli N, Pellin D, Rizzi E, Souque P, Severgnini M, Di Serio C, Fricke T, Diaz-Griffero F, Zimmer C, et al. Chromatin organization at the nuclear pore favours HIV replication. Nat Commun. 2015;6:6483.
73. Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, Pongor S, Luzzati R, Recchia A, Mavilio F, et al. Nuclear architecture dictates HIV-1 integration site selection. Nature. 2015;521:227–31.
74. Li W, Singh PK, Sowd GA, Bedwell GJ, Jang S, Achuthan V, Oleru AV, Wong D, Fadel HJ, Lee K, et al. CPSF6-Dependent targeting of speckle-associated domains distinguishes primate from nonprimate lentiviral integration. MBio. 2020;11:e02254-e2320.
75. Scoca V, Morin R, Collard M, Tinevez JY, Di Nunzio F. HIV-induced membraneless organelles orchestrate post-nuclear entry steps. J Mol Cell Biol. 2022. https://doi.org/10.1093/jmcb/mjac060.
76. Pascual-Garcia P, Capelson M. Nuclear pores in genome architecture and enhancer function. Curr Opin Cell Biol. 2019;58:126–33.
77. Lyngdoh DL, Nag N, Uversky VN, Tripathi T. Prevalence and functionality of intrinsic disorder in human FG-nucleoporins. Int J Biol Macromol. 2021;175:156–70.
78. Kadota S, Ou J, Shi Y, Lee JT, Sun J, Yildirim E. Nucleoporin 153 links nuclear pore complex to chromatin architecture by mediating CTCF and cohesin binding. Nat Commun. 2020;11:2606.
79. Frey S, Gorlich D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell. 2007;130:512–23.

80. Akinyi MV, Frilander MJ. At the intersection of major and minor spliceosomes: crosstalk mechanisms and their impact on gene expression. Front Genet. 2021;12:700744.
81. Brendel V, Karlin S. Association of charge clusters with functional domains of cellular transcription factors. Proc Natl Acad Sci USA. 1989;86:5698–702.
82. Greig JA, Nguyen TA, Lee M, Holehouse AS, Posey AE, Pappu RV, Jedd G. Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation. Mol Cell. 2020;77:1237-1250.e1234.
83. Miyagi T, Yamazaki R, Ueda K, Narumi S, Hayamizu Y, Uji IH, Kuroda M, Kanekura K. The patterning and proportion of charged residues in the arginine-rich mixed-charge domain determine the membrane-less organelle targeted by the protein. Int J Mol Sci. 2022;23:7658.
84. Price AJ, Jacques DA, McEwan WA, Fletcher AJ, Essig S, Chin JW, Halambage UD, Aiken C, James LC. Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. PLoS Pathog. 2014;10:e1004459.
85. Buffone C, Martinez-Lopez A, Fricke T, Opp S, Severgnini M, Cifola I, Petiti L, Frabetti S, Skorupka K, Zadrozny KK, et al. Nup153 unlocks the nuclear pore complex for HIV-1 nuclear translocation in nondividing cells. J Virol. 2018;92:e00648-e718.
86. Dettwiler S, Aringhieri C, Cardinale S, Keller W, Barabino SM. Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. J Biol Chem. 2004;279:35788–97.
87. Ditlev JA, Case LB, Rosen MK. Who's in and who's out-compositional control of biomolecular condensates. J Mol Biol. 2018;430:4666–84.
88. Blanco-Rodriguez G, Gazi A, Monel B, Frabetti S, Scoca V, Mueller F, Schwartz O, Krijnse-Locker J, Charneau P, Di Nunzio F. Remodeling of the core leads HIV-1 preintegration complex into the nucleus of human lymphocytes. J Virol. 2020;94:e00135-e220.
89. Francis AC, Melikyan GB. Single HIV-1 imaging reveals progression of infection through CA-dependent steps of docking at the nuclear pore, uncoating, and nuclear transport. Cell Host Microbe. 2018;23(536–548):e536.
90. Rensen E, Mueller F, Scoca V, Parmar JJ, Souque P, Zimmer C, Di Nunzio F. Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. EMBO J. 2021;40:e105247.
91. Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O, Sundquist WI. Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. Science. 2020;370:eabc8420.
92. Monette A, Niu M, Nijhoff Asser M, Gorelick RJ, Mouland AJ. Scaffolding viral protein NC nucleates phase separation of the HIV-1 biomolecular condensate. Cell Rep. 2022;40:111251.
93. Quinodoz SA, Ollikainen N, Tabak B, Palla A, Schmidt JM, Detmar E, Lai MM, Shishkin AA, Bhat P, Takei Y, et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. Cell. 2018;174(744–757):e724.
94. Chen W, Yan Z, Li S, Huang N, Huang X, Zhang J, Zhong S. RNAs as proximity-labeling media for identifying nuclear speckle positions relative to the genome. iScience. 2018;4:204–15.
95. Nemeth A, Conesa A, Santoyo-Lopez J, Medina I, Montaner D, Peterfia B, Solovei I, Cremer T, Dopazo J, Langst G. Initial genomics of the human nucleolus. PLoS Genet. 2010;6:e1000889.
96. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature. 2008;453:948–51.
97. Francis AC, Marin M, Singh PK, Achuthan V, Prellberg MJ, Palermino-Rowland K, Lan S, Tedbury PR, Sarafianos SG, Engelman AN, Melikyan GB. HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. Nat Commun. 2020;11:3505.
98. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002;110:521–9.
99. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A phase separation model for transcriptional control. Cell. 2017;169:13–23.
100. Sharp PA, Chakraborty AK, Henninger JE, Young RA. RNA in formation and regulation of transcriptional condensates. RNA. 2022;28:52–7.
101. Sawyer IA, Bartek J, Dundr M. Phase separated microenvironments inside the cell nucleus are linked to disease and regulate epigenetic state, transcription and RNA processing. Semin Cell Dev Biol. 2019;90:94–103.
102. Wu L, Pan T, Zhou M, Chen T, Wu S, Lv X, Liu J, Yu F, Guan Y, Liu B, et al. CBX4 contributes to HIV-1 latency by forming phase-separated nuclear bodies and SUMOylating EZH2. EMBO Rep. 2022;23:e53855.
103. Guo C, Che Z, Yue J, Xie P, Hao S, Xie W, Luo Z, Lin C. ENL initiates multivalent phase separation of the super elongation complex (SEC) in controlling rapid transcriptional activation. Sci Adv. 2020;6:eaay4858.
104. Krasnopolsky S, Marom L, Victor RA, Kuzmina A, Schwartz JC, Fujinaga K, Taube R. Fused in sarcoma silences HIV gene transcription and maintains viral latency through suppressing AFF4 gene activation. Retrovirology. 2019;16:16.
105. Nawroth I, Mueller F, Basyuk E, Beerens N, Rahbek UL, Darzacq X, Bertrand E, Kjems J, Schmidt U. Stable assembly of HIV-1 export complexes occurs cotranscriptionally. RNA. 2014;20:1–8.
106. Tuffy KM, Maldonado RJK, Chang J, Rosenfeld P, Cochrane A, Parent LJ. HIV-1 Gag forms ribonucleoprotein complexes with unspliced viral RNA at transcription sites. Viruses. 2020;12:1281.
107. Celetti G, Paci G, Caria J, VanDelinder V, Bachand G, Lemke EA. The liquid state of FG-nucleoporins mimics permeability barrier properties of nuclear pore complexes. J Cell Biol. 2020;219:e201907157.
108. Zhang X, Sridharan S, Zagoriy I, Oegema CE, Ching C, Pflaesterer T, Fung HKH, Poser I, Müller CW, Hyman AW, et al. Molecular mechanisms of stress-induced reactivation in mumps virus condensates. bioRxiv. 2021. https://doi.org/10.1101/2021.07.10.451879
109. Goh GK, Dunker AK, Uversky VN. Protein intrinsic disorder toolbox for comparative analysis of viral proteins. BMC Genomics. 2008;9(Suppl 2):S4.
110. Lochmann TL, Bann DV, Ryan EP, Beyer AR, Mao A, Cochrane A, Parent LJ. NC-mediated nucleolar localization of retroviral gag proteins. Virus Res. 2013;171:304–18.
111. Maldonado RJK, Rice B, Chen EC, Tuffy KM, Chiari EF, Fahrbach KM, Hope TJ, Parent LJ. Visualizing association of the retroviral gag protein with unspliced viral RNA in the nucleus. MBio. 2020;11:e00524-e620.
112. Lyonnais S, Sadiq SK, Lorca-Oro C, Dufau L, Nieto-Marquez S, Escriba T, Gabrielli N, Tan X, Ouizougun-Oubari M, Okoronkwo J, et al. The HIV-1 nucleocapsid regulates its own condensation by phase-separated activity-enhancing sequestration of the viral protease during maturation. Viruses. 2021;13:2312.
113. Mohammed AS, Uversky VN. Intrinsic disorder as a natural preservative: high levels of intrinsic disorder in proteins found in the 2600-year-old human brain. Biology. 2022;11:1704.
114. Fonin AV, Antifeeva IA, Kuznetsova IM, Turoverov KK, Zaslavsky BY, Kulkarni P, Uversky VN. Biological soft matter: intrinsically disordered proteins in liquid-liquid phase separation and biomolecular condensates. Essays Biochem. 2022;66:831–47.
115. Darling AL, Liu Y, Oldfield CJ, Uversky VN. Intrinsically disordered proteome of human membrane-less organelles. Proteomics. 2018;18:e1700193.
116. Alberti S, Hyman AA. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. Nat Rev Mol Cell Biol. 2021;22:196–213.
117. Ahn JH, Davis ES, Daugird TA, Zhao S, Quiroga IY, Uryu H, Li J, Storey AJ, Tsai YH, Keeley DP, et al. Phase separation drives aberrant chromatin looping and cancer development. Nature. 2021;595:591–5.
118. Mathieu C, Pappu RV, Taylor JP. Beyond aggregation: pathological phase transitions in neurodegenerative disease. Science. 2020;370:56–60.
119. Wu C, Holehouse AS, Leung DW, Amarasinghe GK, Dutch RE. Liquid phase partitioning in virus replication: observations and opportunities. Annu Rev Virol. 2022;9(1):285–306.
120. Li H, Ernst C, Kolonko-Adamska M, Greb-Markiewicz B, Man J, Parissi V, Ng BW. Phase separation in viral infections. Trends Microbiol. 2022;30(12):1217–1231.
121. Sagan SM, Weber SC. Let's phase it: viruses are master architects of biomolecular condensates. Trends Biochem Sci. 2023;48(3):229–243.
122. Joshi P, Vendruscolo M. Druggability of intrinsically disordered proteins. Adv Exp Med Biol. 2015;870:383–400.
123. Viny AD, Levine RL. Drug modulation by nuclear condensates. Science. 2020;368:1314–5.

124. Dolgin E. Drug startups coalesce around condensates. Nat Biotechnol. 2021;39:123–5.
125. Lopez N, Camporeale G, Salgueiro M, Borkosky SS, Visentin A, Peralta-Martinez R, Loureiro ME, de Prat-Gay G. Deconstructing virus condensation. PLoS Pathog. 2021;17:e1009926.
126. Mitrea DM, Mittasch M, Gomes BF, Klein IA, Murcko MA. Modulating biomolecular condensates: a novel approach to drug discovery. Nat Rev Drug Discov. 2022;21(11):841–862
127. Risso-Ballester J, Galloux M, Cao J, Le Goffic R, Hontonnou F, Jobart-Malfait A, Desquesnes A, Sake SM, Haid S, Du M, et al. A condensate-hardening drug blocks RSV replication in vivo. Nature. 2021;595:596–9.
128. Wang S, Dai T, Qin Z, Pan T, Chu F, Lou L, Zhang L, Yang B, Huang H, Lu H, Zhou F. Targeting liquid-liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. Nat Cell Biol. 2021;23:718–32.
129. Zhao D, Xu W, Zhang X, Wang X, Ge Y, Yuan E, Xiong Y, Wu S, Li S, Wu N, et al. Understanding the phase separation characteristics of nucleocapsid protein provides a new therapeutic opportunity against SARS-CoV-2. Protein Cell. 2021;12:734–40.
130. Zhao M, Yu Y, Sun L-M, Xing J-Q, Li T, Zhu Y, Wang M, Yu Y, Xue W, Xia T, et al. GCG inhibits SARS-CoV-2 replication by disrupting the liquid phase condensation of its nucleocapsid protein. Nat Commun. 2021;12:2114.
131. Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, Stoynov S, Mahamid J, Saha S, Franzmann TM, et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. Cell. 2015;162:1066–77.
132. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J Biol Chem. 2003;278:372–81.
133. Lapaillerie D, Lelandais B, Mauro E, Lagadec F, Tumiotto C, Miskey C, Ferran G, Kuschner N, Calmels C, Metifiot M, et al. Modulation of the intrinsic chromatin binding property of HIV-1 integrase by LEDGF/p75. Nucleic Acids Res. 2021;49:11241–56.
134. Ma X, Chen T, Peng Z, Wang Z, Liu J, Yang T, Wu L, Liu G, Zhou M, Tong M, et al. Histone chaperone CAF-1 promotes HIV-1 latency by leading the formation of phase-separated suppressive nuclear bodies. EMBO J. 2021;40:e106632.
135. Bellmann J, Monette A, Tripathy V, Sojka A, Abo-Rady M, Janosh A, Bhatnagar R, Bickle M, Mouland AJ, Sterneckert J. Viral infections exacerbate FUS-ALS phenotypes in iPSC-derived spinal neurons in a virus species-specific manner. Front Cell Neurosci. 2019;13:480.
136. Leblanc P, Vorberg IM. Viruses in neurodegenerative diseases: More than just suspects in crimes. PLoS Pathog. 2022;18:e1010670.
137. Garabedian MV, Wang W, Dabdoub JB, Tong M, Caldwell RM, Benman W, Schuster BS, Deiters A, Good MC. Designer membraneless organelles sequester native factors for control of cell behavior. Nat Chem Biol. 2021;17:998–1007.
138. Dayhoff GW 2nd, Uversky VN. Rapid prediction and analysis of protein intrinsic disorder. Protein Sci. 2022;31(12):e4496.
139. Oates ME, Romero P, Ishida T, Ghalwash M, Mizianty MJ, Xue B, Dosztanyi Z, Uversky VN, Obradovic Z, Kurgan L, et al. D(2)P(2): database of disordered protein predictions. Nucleic Acids Res. 2013;41:D508-516.
140. Andreeva A, Howorth D, Brenner SE, Hubbard TJ, Chothia C, Murzin AG. SCOP database in 2004: refinements integrate structure and sequence family data. Nucleic Acids Res. 2004;32:D226-229.
141. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995;247:536–40.
142. de Lima Morais DA, Fang H, Rackham OJ, Wilson D, Pethica R, Chothia C, Gough J. SUPERFAMILY 1.75 including a domain-centric gene ontology method. Nucleic Acids Res. 2011;39:D427-434.
143. Meszaros B, Simon I, Dosztanyi Z. Prediction of protein binding regions in disordered proteins. PLoS Comput Biol. 2009;5:e1000376.
144. Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, Latham V, Sullivan M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. 2012;40:D261-270.
145. 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 USA. 2020;117:33254–62.
146. Horvath A, Miskei M, Ambrus V, Vendruscolo M, Fuxreiter M. Sequence-based prediction of protein binding mode landscapes. PLoS Comput Biol. 2020;16:e1007864.
147. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–9.
148. Rajagopalan K, Mooney SM, Parekh N, Getzenberg RH, Kulkarni P. A majority of the cancer/testis antigens are intrinsically disordered proteins. J Cell Biochem. 2011;112:3256–67.
149. Uversky VN. Analyzing IDPs in interactomes. Methods Mol Biol. 2020;2141:895–945.
150. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins. 2000;41:415–27.
151. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. Comparing and combining predictors of mostly disordered proteins. Biochemistry. 2005;44:1989–2000.
152. Mohan A, Sullivan WJ Jr, Radivojac P, Dunker AK, Uversky VN. Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. Mol Biosyst. 2008;4:328–40.
153. Huang F, Oldfield C, Meng J, Hsu WL, Xue B, Uversky VN, Romero P, Dunker AK. Subclassifying disordered proteins by the CH-CDF plot method. Pac Symp Biocomput. 2012;128–39. https://doi.org/10.1142/9789814366496_0013.

Publisher's Note