Review

The Significance of the Intrinsically Disordered Regions for the Functions of the bHLH Transcription Factors

Aneta Tarczewska and Beata Greb-Markiewicz *

Department of Biochemistry, Faculty of Chemistry, Wroclaw University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wroclaw, Poland; aneta.tarczewska@pwr.edu.pl
* Correspondence: beata.greb-markiewicz@pwr.edu.pl

check for updates

Abstract: The bHLH proteins are a family of eukaryotic transcription factors regulating expression of a wide range of genes involved in cell differentiation and development. They contain the Helix-Loop-Helix (HLH) domain, preceded by a stretch of basic residues, which are responsible for dimerization and binding to E-box sequences. In addition to the well-preserved DNA-binding bHLH domain, these proteins may contain various additional domains determining the specificity of performed transcriptional regulation. According to this, the family has been divided into distinct classes. Our aim was to emphasize the significance of existing disordered regions within the bHLH transcription factors for their functionality. Flexible, intrinsically disordered regions containing various motives and specific sequences allow for multiple interactions with transcription co-regulators. Also, based on in silico analysis and previous studies, we hypothesize that the bHLH proteins have a general ability to undergo spontaneous phase separation, forming or participating into liquid condensates which constitute functional centers involved in transcription regulation. We shortly introduce recent findings on the crucial role of the thermodynamically liquid-liquid driven phase separation in transcription regulation by disordered regions of regulatory proteins. We believe that further experimental studies should be performed in this field for better understanding of the mechanism of gene expression regulation (among others regarding oncogenes) by important and linked to many diseases the bHLH transcription factors.

Keywords: bHLH; IDP; IDR; LLPS; disorder prediction; LLPS prediction; transcription; phase separation

1. Introduction

The bHLH (basic Helix-Loop-Helix) proteins are the important family of transcription factors (TFs) present in all eukaryotes: from yeasts [1,2] and fungi [3] to plants [4] and metazoans [5–10]. All family members contain the HLH domain responsible for dimerization [11]. This domain is usually preceded by a stretch of basic residues which enable DNA binding [12]. The bHLH TFs recognize tissue-specific enhancers containing E-box sequences which regulate expression of a wide range of genes involved in cell differentiation and development [13].

Currently, a few independent classification systems of the bHLH proteins exists: evolutionary classification based on the phylogenetic studies of the bHLH proteins, which classify the bHLH family members into six A–F classes [7,8,14], and a new one based on the complete amino acid sequence analyses, classifying the bHLH proteins into six clades without assumptions about gene function [15]. Contrary to the previous methods, natural method of classification proposed by Murre [12], which divides the bHLH proteins into seven classes, is based on the presence of additional domains, expression

patterns and performed transcriptional function [10]. For purposes of clarity, some attempts to revise and systematize different classification systems were undertaken [16]. In this review we present classification of bHLH proteins according to Murre [12], with some short description of presented classes (Table 1).

Table 1. Classification of bHLH proteins based on [5,7,8,10,12,14,16].

Structural Motif Dimerization	Representative Members	Short Description
class I (E proteins)/ group A		
bHLH, homo- and heterodimerization	Vertebrate: E12, E47 [17], HEB [18,19], TCF4 [20] Invertebrate: Daughterless	transcription activators, ubiquitous expression, neurogenesis, immune cell development, sex development, gonadogenesis
class II/ group A		
bHLH, preferred heterodimerization with class I partners	Vertebrate: MYOD, Myogenin, MYF5-6, Ngn1-3, ATOH, NeuroD, NDRF, MATH, MASH, ASCL1 [21], TAL1/SCL [22], OLIG1-3 [23] Invertebrate: TWIST [24], AS-C	transcription activators, tissue specific expression, muscle development, neuro-genesis, generation of autonomic and olfactory neurons, development of granule neurons and external germinal layer of cerebellum, oligodendrocyte development, specification of blood lineage and maturation of several hematopoietic cells, pancreatic development
class III/ group B		
bHLH-LZ	Vertebrate: MYC [25], USF, TFE3, SREBP1-2 *Drosophila*: MYC Plants: MYC2	transcription activators/represors, oncogenic transformation, apoptosis, cellular differentiation, proliferation, cholesterol-mediated induction of the low-density lipoprotein receptor, jasmonate signaling (plants)
class IV/ group B		
bHLH, heterodimerisation with each other and MYC proteins	Vertabrate: MAD, MAX [26], MXI1 *Drosophila*: MNT, MAX	transcription regulators lacking transactivation domain (TAD)
class V/ group D		
HLH (no basic region)	Vertebrate: ID1-4 [27] Invertebrate:EMC	negative transcription regulators of class I and II (group A) proteins, no DNA binding, regulation by sequestration.
class VI/ group B		
bHLH-O, (presence of proline in basic region)	Vertebrate: HES, HEY1-3 [28], STRA13, HERP1-2 [29] *Drosophila*: HAIRY [30], E(spI)	negative transcription regulators interacting with corepressors (Groucho); neurogenesis, vasculogenesis, mesoderm segmentation, myogenesis, T lymphocyte development, cardiovascular development and homeostasis; effectors of Notch signalling [28]; in *Drosophila*: regulation of differentiation, anteroposterior segmentation and sex determination
class VII/ group C - subclass I		
bHLH-PAS, heterodimerization with subclass II	Vertebrate: AHR [31], HIF1-3α [32], SIM1-2 [33], CLOCK [34], NPAS1-4 [35–39] *Drosophila*: MET [40], GCE, SIMA, TRH	transcription regulation in response to physiological and environmental signals: xenobiotics, hypoxia, development, circadian rhytms
class VII/ group C - subclass II		
bHLH-PAS, homo- and heterodimerization with subclass I	Vertebrate: ARNT [41], ARNT2, BMAL1, BMAL2 *Drosophila*: TANGO, CYCLE	general partners for subclass I bHLH-PAS proteins

Both class I (known as E proteins) and class II of the bHLH TFs do not possess domains additional to the bHLH. Contrary to the class I which is expressed in many tissues, the class II proteins expression is tissue specific. Members of the class II are dimerization partners for the class I transcription factors. Class III comprises proteins possessing Leucine-zipper (LZ) motif in addition to the bHLH. Important members of the class III are proteins belonging to the Myc subfamily, which regulate oncogenic transformation, apoptosis, and cellular differentiation. To class IV belong MAD and MAX which can dimerize with MYC and regulate its activity. Also, MAD/MAX are able to create homo- and heterodimers with each other. Although these TFs do not possess transcription activation domain (TAD), MAD/MAX dimers can influence the transcription in a differentiated way. Class V contains

transcriptional inhibitors ID1-3 which are not able to bind DNA and act by the other bHLH proteins sequestration. Interestingly, the fourth member of this class- ID4 function as inhibitor of ID1-3 [42]. Class VI comprise proteins containing additional Orange domain adjacent C-terminally to the bHLH domain (bHLH-O). Transcription factors from the described classes perform regulatory function in various developmental processes including cells differentiation and maintaining pluripotency. For this reason they are often linked to cancer development. Class VII comprise transcription factors which possess PAS (Period-Aryl hydrocarbon receptor nuclear translocator-Single minded) domain located C-terminally to the bHLH domain. PAS domain is crucial for the bHLH-PAS proteins specifity [43]. Structurally, the C-terminal PAS domain is often associated with PAC (C-terminal to PAS) motif [44,45]. bHLH-PAS transcription factors are responsible for sensing environmental signals like the presence of xenobiotics (AHR), hypoxia (HIF) or setting of circadian rhythms of organism (CLOCK, CYCLE, BMAL). The members of subclass II of bHLH-PAS TFs -ARNT proteins are general dimerization partners of the subclass I members.

2. The Role of the bHLH Proteins in Transcription

The regulation of genes expression by multiple transcription factors, cofactors and chromatin regulators establish and maintains a specific state of a cell. Inaccurate regulation of transmitted signals can results in diseases and severe disorders [46]. Therefore, transcription requires balanced orchestration of adjustable complexes of proteins. A key regulator of transcription is Mediator, a multi-subunit Mediator complex which interacts with RNA polymerase II (Pol II), and coordinates the action of numerous co-activators and co-repressors [47–50]. Function of the Mediator is conserved in all eukaryotes, though, the individual subunits have diverged considerably in some organisms [51,52].

Up to date, for some bHLH family representatives, interactions with subunits of the Mediator and/or chromatin remodeling histone acetyltransferases/deacyltransferase, were reported. In plants, the Mediator complex is a core element of transcription regulation important for their immunity [53]. It was shown, that in *Arabidopsis thaliana* important jasmonate signaling and resistance to fungus *Botrytis cinerea*, is dependent on the interaction between MED25 subunit of the Mediator and MYC2 [54–56], and interaction of MED8 subunit of the Mediator with FAMA belonging to the bHLH family [57]. Sterol regulatory element binding proteins (SREBPs) the class II bHLH TFs (Table 1) are transcription activators critical for regulation of cholesterol and fatty acid homeostasis in animals. It was shown that human SREBPs bind CBP/p300 acetyltransferase [58] and MED15 subunit of the Mediator to activate target genes [59]. Also yeast Ino2 was shown to bind MED15 subunit of the Mediator tail [60].

The representative of class II TFs TAL1 (Table 1) is required for the specification of the blood lineage and maturation of several hematopoietic cells. TAL1/SCL is considered as a master TF delineating the cell fate and the identity of progenitor and normal hematopoietic stem cells (HSCs). It regulates other hematopoietic TFs thus has a potential for cell reprogramming [22]. TAL1 also binds CBP/p300 acetyltransferase [61,62]. Similarly MyoD—a myogenic regulatory factor which controls skeletal muscle development binds CBP and recruits histone acetyltransferase to activate myogenic program [63]. Cao et al. showed that of MyoD modify the myoblasts chromatin structure and accessibility [64]. ASCL1 (class II, Table 1) was shown to be a pioneer factor which promotes chromatin accessibility and enables chromatin binding by others TFs [65]. Recently, also AHR (bHLH-PAS, Table 1) was suggested to be a pioneer factor which regulates DNA methylation during embryonic developments in unknown way [66]. In clear cell renal cell carcinoma (ccRCC), the most frequent mutation causes the von Hippel-Lindau (VHL) tumor suppressor inactivation leading to genome-wide enhancer and super-enhancer remodeling. This process is mediated by the interaction of HIF2α and HIF1β (bHLH-PAS, Table 1) with histone acetyltransferase p300 [67]. CLOCK, the other bHLH-PAS subfamily member (Table 1) was shown to mediate histone acetylation in a circadian time-specific manner [68].

Interestingly, the bHLH-O proteins members (class VI, Table 1) HEY proteins can function as transcription repressors as well as transcription activators. They were shown to bind directly DNA and interact with histone deacetylases and other TFs [28,69]. On the other hand, gene activation by

HEY is regulated in an indirect way. Multiple HEY binding sites located downstream and close to the transcriptional start site, resulted in a hypothesis that HEY influence the pausing/elongation switch of Pol II [70]. Interestingly, though most of TFs stimulate transcription initiation, MYC (class III, Table 1) was shown to stimulate transcription elongation by recruitment of the elongation factor [71]. The presented studies indicate that the crucial role of the bHLH proteins in maintaining transcriptional regulation of important developmental (e.g., cell differentiation) and oncogenic pathways is dependent on the multiple interactions with basal transcriptional machinery.

3. The bHLH Transcription Factors as IDPs

Intrinsically disordered proteins (IDPs) discovered in 1990s obliterate the paradigm derived from Anfinsen's work, stating that functional proteins must possess a well-defined, ordered, three dimensional structure [72]. Currently it is known, that a large number of proteins is perfectly functional or even multifunctional in a disordered state in which a polypeptide chain undergoes rapid conformational fluctuations [73–76]. Intrinsic disorder can be spread throughout the whole polypeptide chain, or it can be limited to intrinsically disordered regions (IDRs) of various length, which are accompanied by well folded domains [77]. The unique properties of disordered proteins originate from their unusual amino acids composition [78]. IDPs/IDRs are depleted in order promoting amino acid residues (hydrophobic, aromatic, aliphatic side chains). In contrast, they possess unusually high content of charged and hydrophilic amino acid residues [79–81]. As a consequence, disordered polypeptide chains have extremely high net charge and low hydrophobicity [82]. IDPs are pliable and highly dynamic molecules of interconvertible conformations. They may completely or almost completely lack the regular secondary structures. However, the content of secondary structure may also be quite significant and molecules can exist in a molten globule state [83–85]. Various in silico analyses indicated that the proportion of disordered proteins is drastically higher in eukaryotes comparing to prokaryotes [86]. This disproportion reflect the complexity of signaling pathways in which IDPs/IDRs play a crucial role [87]. Due to the flexible and dynamic nature, IDPs/IDRs can form fuzzy complexes, adopting various conformations [88]. According to this, one IDP can form multiple interactions with various partners. Due to a large accessibility of particular residues in a disordered chain, the interaction pattern can be easily modified by posttranslational modifications [89]. For that reason IDPs/IDRs often serve as molecular hubs, modulators and sensors of cellular signals [85].

bHLH TFs are responsible for a control of developmental processes like retinal development, proliferation of progenitors, neurogenesis and gliogenesis. Importantly, this is due to a direct interaction between bHLH TFs and interaction of bHLH TFs with homeodomain factors which create complexes that bind to the specific promoters [90,91]. Transcription of muscle-specific genes during skeletal muscle development is also dependent on the interactions between specific bHLH TFs: MyoD, Myogenin, Myf5 and MRF4 with ubiquitously expressed bHLH E-proteins (E12, E47, TCF4, HEB). Interestingly, it was shown that MyoD interacts with two isoforms of HEB: HEBα and HEBβ. which regulate differentially transcriptional activity of MyoD not only on different, but also on the same promoter [92]. Also interesting is the ability of ID4 to recruit multiple ID proteins to assemble higher order complexes. ID4 restores DNA binding by E47 protein even in the presence of repressing ID1 and ID2. Additionally, the ID proteins can interact with non-bHLH partners expanding regulatory network of ID4 [42]. As a consequence, the ID proteins are proposed as a 'hub' for coordination of multiple cancer events [27]. These examples illustrate the possibility of bHLH TFs to interact with many partners in differentiated way. We suggest that these is related to the disordered character of the bHLH proteins. This hypothesis is substantiated by some experimental studies. Neurogenic bHLH transcriprion factor Neurogenin 2 (Ngn2) was shown to possess long IDR which phosphorylation regulates the activity of the protein [93]. Interestingly, though the bHLH domain was considered as a stable, well ordered structure, partially disordered character of this domain was presented for NeuroD [94], MYC and MAX [95]. We performed in silico analyses to predict the presence of intrinsic disorder and get an insight into the degree of flexibility of bHLH proteins representing all established classes (see Table 1): hHEB (class I), hMYOD

(class II), hMYC and atMYC2 (class III) (Figure 1); hMAD1 and hMAX (class IV), hID4 (class V), hHES (class VI) (Figure 2); hAHR, hHIF-1α, hCLOCK and hARNT (class VII) (Figure 3). We used PONDR-VLXT [96,97], http://www.pondr.com/ for the disorder prediction and DynaMine [98,99], http://dynamine.ibsquare.be/submission/ for prediction of the flexibility of proteins backbone.

A representative of the class I, human HEB shows a high content of predicted as disordered and flexible sequences. The only highly ordered/rigid region appears between 577–630 aa which comprise the bHLH domain (Figure 1A). Based on prediction results, we assume HEB as IDP. Also hMyoD, the class II TFs presents a high content of flexible IDRs especially in the C-terminal part of the protein (Figure 1B). As the representatives of the class III we have chosen hMYC (Figure 1C) (for which partial disorder of the bHLH domain was experimentally documented [95]) and *Arabidopsis thaliana* MYC2 (Figure 1D). For both proteins the presence of flexible IDRs was predicted, though they locations were different.
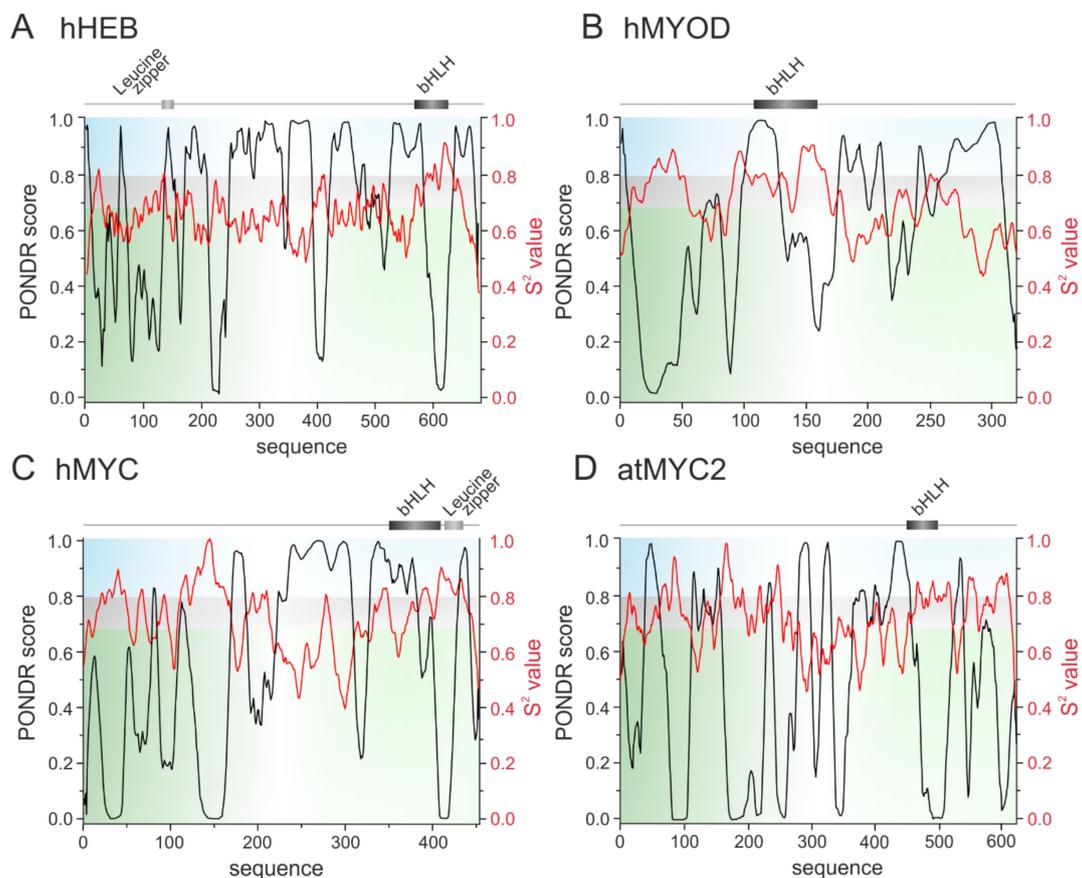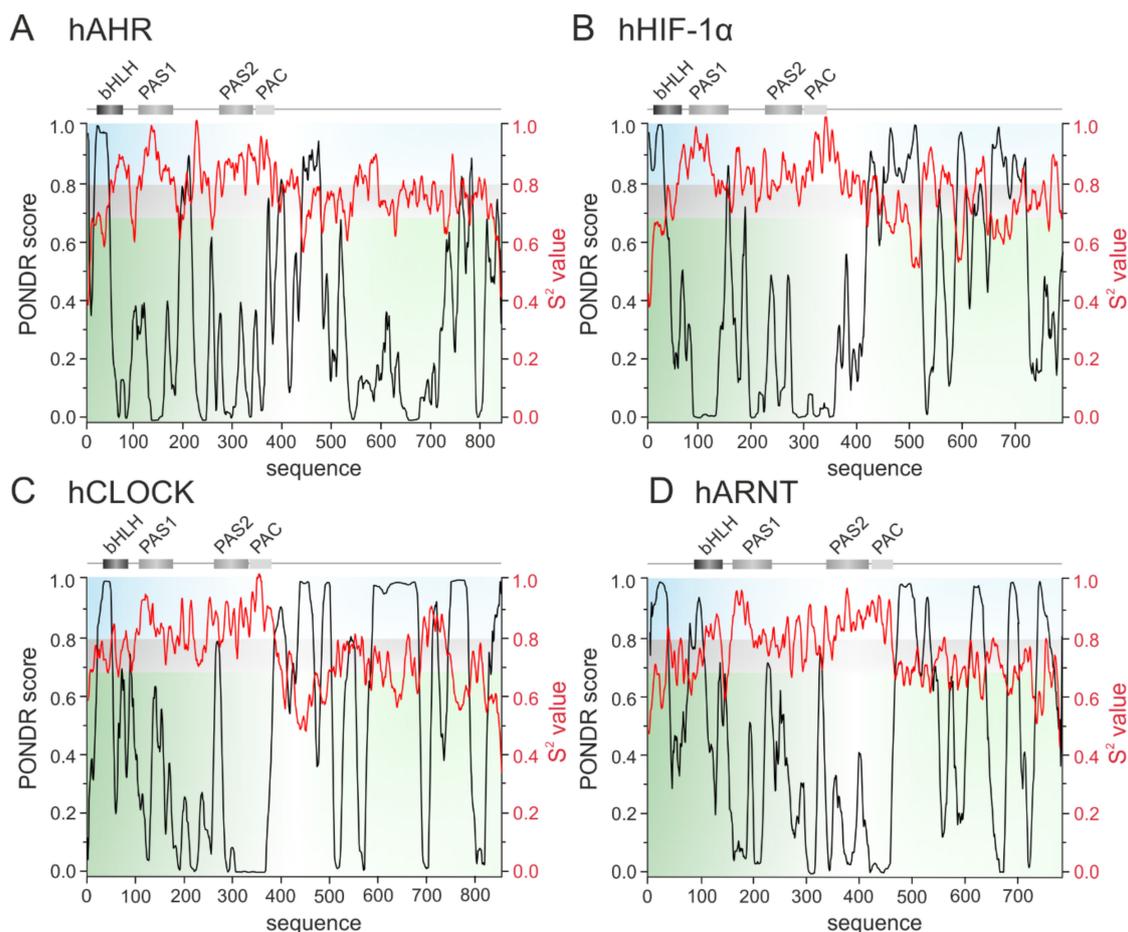

Figure 1. Prediction of intrinsically disordered regions. The top panel presents the domain structure of the analyzed bHLH proteins. Dark grey rectangle indicates the position of bHLH domain, the light grey Leucine zipper. The bottom panel presents a prediction of intrinsically disordered and flexible regions based on the amino acid sequence of proteins. Prediction were performed using PONDR-VLXT (left Y axis) and DynaMine (right Y axis) software. For PONDR prediction, a score above 0.5 indicates disorder. For DynaMine, a S^2 value above 0.8 (blue zone) indicates rigid conformation, 0.69-0.8 (grey zone) is context dependent and a value below 0.69 (green zone) indicates flexible conformation. (**A**) class I human HEB [Q99081], (**B**) class II human MYOD [P15172], (**C**) class III human MYC [P01106-2] and (**D**) *Arabidopsis thaliana* MYC2 [Q39204].
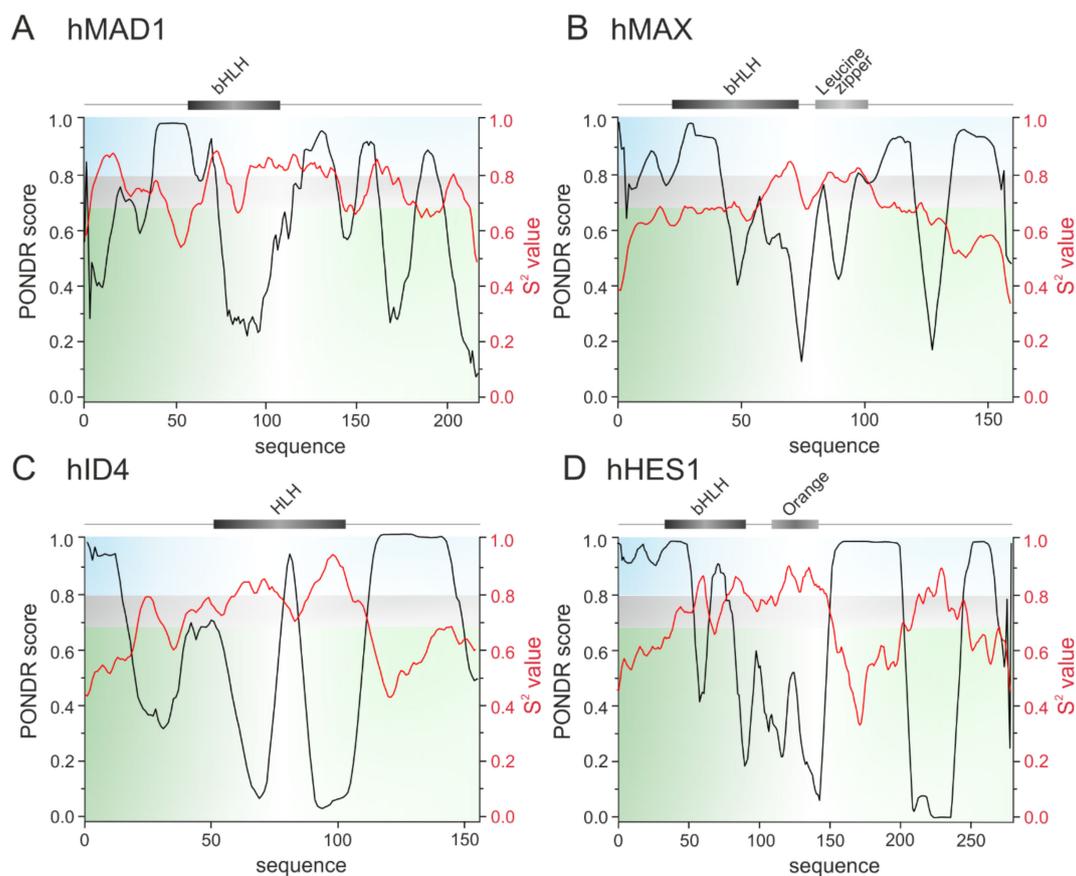
Figure 2. Prediction of intrinsically disordered regions. The top panel presents the domain structure of the analyzed bHLH proteins. Dark grey rectangle indicates the bHLH domain, light grey indicates Leucine zipper or Orange domain. The bottom panel presents a prediction of intrinsically disordered and flexible regions, based on the amino acid sequence of proteins. Predictions were performed using PONDR-VLXT (left Y axis) and DynaMine (right Y axis) software. For PONDR prediction, a score above 0.5 indicates disorder. For Dynamine, a S^2 value above 0.8 (blue zone) indicates rigid conformation, 0.69–0.8 (grey zone) is context dependent and a value below 0.69 (green zone) indicates flexible conformation. (**A**) class IV human MAD [Q9Y6D9] and (**B**) human MAX [P61244], (**C**) class V human ID4 [P47928], (**D**) class VI human HES1 [Q14469].

The representative of the class IV, human MAD1 also shows high content of predicted as disordered and flexible sequences (Figure 2A). Interestingly IDRs of hMAX which belongs to the same class IV are located in the N- and C- protein termini, while the middle part is predicted as possessing more rigid structure (Figure 2B). Also, ID4 belonging to the class V of transcriptional inhibitors presents flexible IDR in the C-terminal part of protein and a shorter one in the N-terminal part (Figure 2C). In addition to similarly located the N- and C-terminal IDRs in the class VI member, human HES1 analysis shows high flexibility/disorder in the central part of protein (Figure 2D).

The class VII proteins comprise the bHLH-PAS subfamily, which additionally to the bHLH domain possess a PAS domain responsible for ligands and co-factors binding. Importantly, their C-termini are usually responsible for the regulation of the protein and created complexes activity [100]. Human AHR, HIF1-α, and CLOCK belong to the subclass I of specialized factors, while human ARNT (the subclass II) is one of the general partners which dimerize with the subclass I proteins and is important for their activity. In contrast to the hAHR, for which relatively short IDRs were predicted within the middle, the N- and the C-terminal part of the protein (Figure 3A), other bHLH-PAS members contain longer IDRs which comprise most of the C-terminal half of proteins and are predicted as highly flexible (hHIF-1α, Figure 3B; hCLOCK, Figure 3C; hARNT, Figure 3D).

Figure 3. Prediction of intrinsically disordered regions of the class VII bHLH-PAS proteins. The top panel presents the domain structure of the analyzed bHLH–PAS proteins. Dark grey rectangle indicates the bHLH domain, light grey indicates PAS/PAC domains. The bottom panel presents a prediction of intrinsically disordered and flexible regions based on the amino acid sequence of proteins. Prediction were performed using PONDR-VLXT (left Y axis) and DynaMine (right Y axis) software. For PONDR prediction, score above 0.5 indicate disorder. For Dynamine, a S^2 value above 0.8 (blue zone) indicates rigid conformation, 0.69–0.8 (grey zone) is context dependent and a value below 0.69 (green zone) indicates flexible conformation. (**A**) human AHR [P35869], (**B**) human HIF-1α [Q16665], (**C**) human CLOCK [O08785], (**D**) human ARNT [P27540].

To date, the only report, concerning the structure of the full-length bHLH protein is the mentioned study showing Neurogenin as IDP [93]. Based on the presented predictions and our own experience with expression of the selected bHLH proteins (not published), we assume that this is due to the relatively high content of IDRs. This makes overexpression and purification process extremely difficult because of propensity to aggregation and high sensitivity to proteases.

4. The Role of IDPs in Maintaining/Creation of LLPS

Over the last decade, since the pioneering work regarding physical nature of P-bodies was published by Hyman and co-workers [101], many molecular biologists and biophysicists have focused on the significance of spontaneous thermodynamically driven liquid-liquid phase separation (LLPS) in biological systems. LLPS leads to formation of dense, liquid condensates that stably coexist in diluted phase [101,102]. At the molecular level it was shown that LLPS is forced by multiple weak and transient interactions which engage IDPs/IDRs [101,103–106]. Repetitively distributed within IDRs highly charged regions of opposite charges, short motifs such as YG/S-, FG-, RG-, GY-, KSPEA-, SY- and Q/N-rich regions form multivalent interactions between condensate components [107]. A

model for the condensate formation and composition proposes that some proteins act as the scaffolds, while others as the clients. The scaffolds are the modular proteins which contain repeated motives that enable heterotypical scaffold-scaffold interaction. As they undergo spontaneous LLPS they are essential for the structural integrity of a condensate [108,109]. Directly interacting sequences called stickers are usually multivalent, whereas the interval sequences which separate stickers, called spacers are responsible for the properties of a condensate [110]. Highly charged and flexible IDRs are in fact frequently identified as scaffolds [108,111]. The clients participate into the condensates by binding to the free, unoccupied scaffold sites [108]. A growing number of evidences indicate that LLPS constitute a fundamental mechanism to compartmentalize the intracellular space. LLPS form the functional centres for biochemical reactions in cytoplasm and membrane-surrounded organelles including nucleus.

The structural and functional organisation of the interior of the nucleus was believed to rely solely on the rigid insoluble nuclear matrix [112]. The rich in A and T DNA sequences known as scaffold/matrix associated regions (S/MARs) attach to nuclear matrix and organise chromatin into higher-order structures which comprise distinct loops and functional units attached to the matrix [113]. That concept is now giving way to a new concept, were dynamic, spontaneously formed condensates, such as nucleolus, splicing speckles, Cajal bodies, PML bodies are the key structural and functional components of the nuclear interior. The barrier-free character of liquid condensates allows for rapid exchange of their components with surrounding so they form an ideal environment for biochemical reactions. On the other hand, nuclear condensates have a stable inert, well-defined structure and can be purified by biochemical methods [114]. It was shown, that the concentration of nucleolar components is close to saturation [115]. It means that small changes in the nucleus can drive spontaneous LLPS. In fact association/dissociation events of nuclear condensates regulate many processes related to gene expression [116] including chromatin structure organisation [117], RNA processing [118], ribosome biogenesis [119]. Importantly, LLPS was shown to be involved in formation of some functional condensates that regulate genes transcription [76,120–122].

5. The Transcription Regulation and LLPS

The genes transcription process require tight regulation to ensure physiological balance of the cell. Knowledge regarding the mechanism of transcription is quite advanced, however some aspects of regulation remains unexplored. Recent findings indicate that regulatory mechanism may tightly depends on the spontaneous LLPS. Transcription of tissue specific gene is initiated at the specific genome regions called super-enhancers (SE). SE first described in embryonic stem cells (ESC) [123] are dense multicomponent assemblies different from typical enhancers [124]. Recently Hnisz [125] performed computational simulation to obtain the probable explanation for typical features of SE. Simulations led to conclusion that formation, activity and unique properties of SE such as sensitivity to concentration of its components, sensitivity to posttranslational modifications, extremely high frequency bursting [126–128] may originate from the fact that SE are liquid condensates assembled/disassembled via spontaneous LLPS [125]. Hnisz and co-workers were the first who point connection and strong dependence between the regulation of transcription initiation at SE and LLPS. Although not experimentally proven, the model serves as the conceptual framework for further research. Recently, Sabari et al. [121] showed that largely disordered BRD4 and MED1 subunit of the Mediator are in close spatial proximity to one another within SE in murine ESC and co-localised puncta show characteristic features of phase separated condensates Moreover, MED1 condensates can incorporate BRD4 and Pol II from nuclear extract [121]. MED1 subunit interacts also with other major pluripotency TFs e.g., OCT-4 [129] and estrogen receptor (ER) [130] forming liquid-like puncta at SE of the key pluripotency genes [121,122]. MED1 condensates depends on the OCT-4 occupancy [122], which are crucial for initiation of tissue specific genes transcription at SE [122,131]. In vitro analyses pointed that formation of MED1-OCT4 liquid condensates occurs via the electrostatic interactions and involves acidic residues enriched in disordered activation domain of the OCT-4 [122]. Interestingly, ER interact with the MED1 subunit by LXXLL motif [132] which is located in the ordered ligand binding domain.

This interaction is regulated by estrogen what means that not only disordered-disordered regions interaction but also disordered-ordered regions interactions play a role in transcription regulation forced by LLPS [122]. Wu et al. [120] showed that largely disordered transcription co-activator TAZ protein forms liquid condensates in vitro and *in vivo*. TAZ condensates compartmentalize DNA binding cofactor TEAD4 and other components of transcription initiation machinery including BRD4, MED1 and CDK9. Importantly, deletion mutant, that is not able to undergo spontaneous LLPS cannot initiate transcription though is able to bind TAZ partners such TEAD4.

Importantly, there are some evidences that not only the initiation, but also the elongation of transcription depends on LLPS. For the transcription elongation essential is hyper-phosphorylation of the YSPTSPS consensus sequence which is repeated multiple times in the disordered C-terminal domain (CTD) of Pol II [133–136]. pTEFb which begins the elongation phase consists of CDK9 kinase associated with cyclin T1 (CycT1). Lu with co-workers [76] concentrated on the function of the lengthy C-terminal IDR of CycT1 in regulation of CDK9 activity. They revealed that a histidine-rich domain (HRD) located in the IDR of CycT1 (residues 480–550) is directly involved in the regulation of the kinase activity [76]. Interestingly, HRD is present also in some other kinases, for example Dyrk1A which phosphorylates CTD of Pol II. Importantly, a homologues kinase Dyrk3 was shown to be responsible for disassembly of stress granules [137] and other cellular condensates during cell division [138]. In vitro studies using a set of recombinant IDRs of the CycT1 and Dyrk1A revealed that the regions can undergo phase separation in a HRD dependent manner. HRD was shown to form condensates which compartmentalize the kinases and the substrate what enables efficient reactions resulting in the hyper-phosphorylation of the CTD of Pol II [76]. Interestingly, the CTD of Pol II can undergo spontaneous LLPS in vitro only in a non-phosphorylated state. The weak CTD-CTD interaction keeps the enzymes molecules in hubs within nucleoplasm. Phosphorylation change the interaction pattern allowing CTD to engage in new multivalent interactions with selected partners [139]. These results indicate that LLPS allows for the condensation of cofactors, that in turn triggers posttranslational modifications leading to the reorganization of the condensate components. Pol II escapes from the promoter site and enables the entry into active elongation stage [76].

Currently not much is known about proteins responsible for formation of the condensates which are important for transcription regulation. The question still remains unanswered which proteins are the scaffolds and which are the clients. Importantly, also not much is known about the involvement of the bHLH TFs in the LLPS process, though they are key players involved in many important cell differentiation and organisms development pathways. As we discussed in previous section, bHLH proteins possess long IDRs which could interact with different partners and be engaged in LLPS. This hypothesis is substantiated by an experimental verification of MyoD possibility to create LLPS [122], and discussed in previous section possibility of some bHLH TFs to interact with the Mediator subunits or other elements of the mechanism which modifies the chromatin accessibility. Interestingly, regulation of circadian clock by BMAL1 comprises binding of CBP, which occurs in discrete nuclear foci. This led to a hypothesis that formation of nuclear bodies containing BMAL1/CBP provides transcriptionally active sites of target genes, like *Per1-2* [34]. Taking the above into consideration, we asked the question if the ability to undergo LLPS is a more general property of the bHLH TFs. As we got positive results for the previously performed prediction of disorder, which was shown to be important for LLPS initiation [76,121,122], we decided to perform in silico analyses to predict if members of the bHLH family comprise putative sequences able to create liquid condensates. We used catGranule program, (http://service.tartaglialab.com/update_submission/216885/dd56e32a89) for computational analyses of the putative propensity to undergo LLPS [140] for the bHLH proteins representing all established classes (see Table 1). Prediction results showed that hHEB (class I), hMyoD (class II), hMYC and 84atMYC2 (class III) (Figure 4) contain sequences with a positive score of propensity to LLPS formation. Interestingly, proteins from the class IV regulators which do not possess TAD: hMAD1 and hMAX, similarly like transcription repressors: hID4 (class V) and hHES (class VI) present very low or even negative score within the whole protein sequence (Figure 5). bHLH-PAS transcription

factors representing the class VII, hAHR, hHIF-1α, hCLOCK and hARNT were predicted as containing some sequences with high propensity score (Figure 6). Especially interesting is the observation that the transcription repressors show a very low propensity scoreto undergo LLPS in contrast to the transcription activators such as hHEB or atMYC2. It is possible that the bHLH repressors inhibit transcription by preventing spontaneous phase separation required to form a complete initiation complex. This hypothesis is substantiated by the observation for TAZ mutants [120], discussed in the previous section.
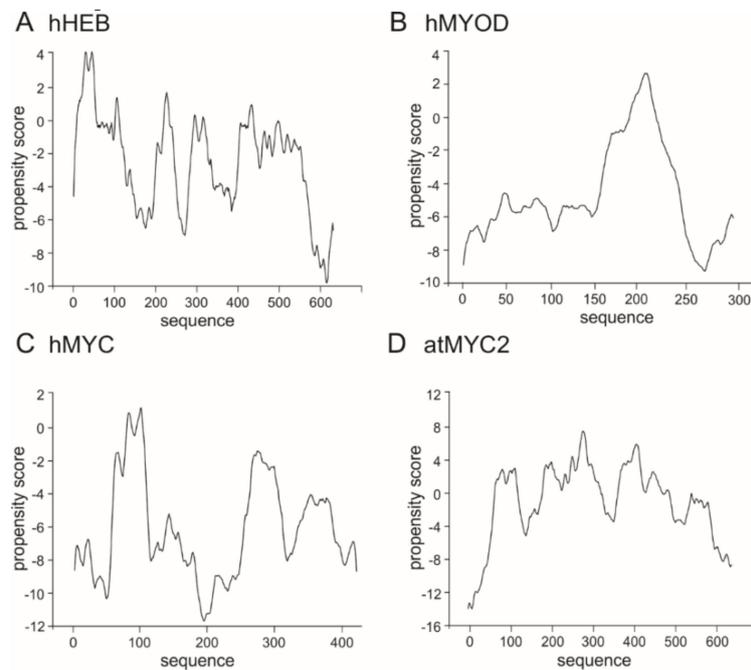
Figure 4. Prediction of propensity of LLPS formation. (**A**) class I human HEB [Q99081], (**B**) class II human MYOD [P15172], (**C**) class III human MYC [P01106-2] and (**D**) *Arabidopsis thaliana* MYC2 [Q39204].
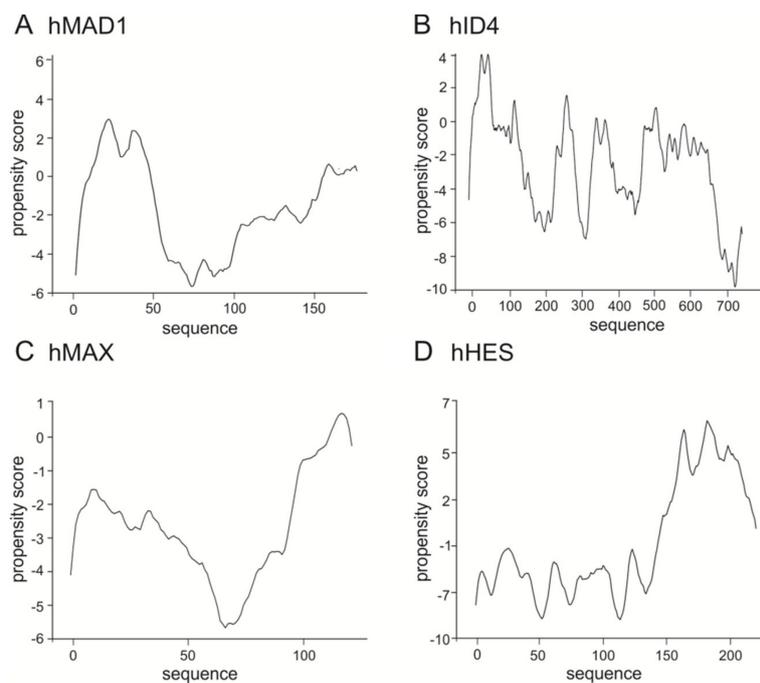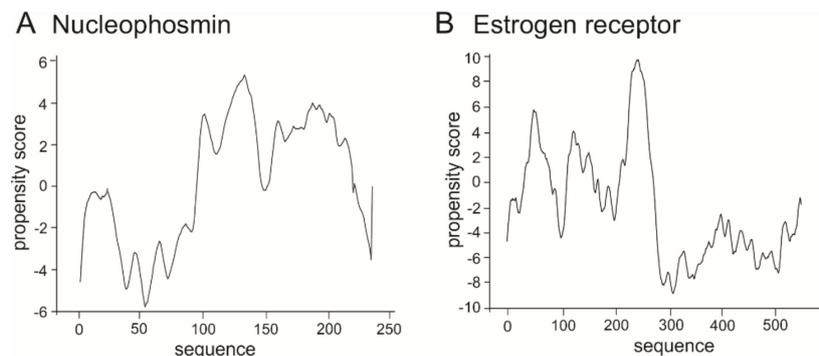

Figure 5. Prediction of propensity of LLPS formation. (**A**) class IV human MAD [Q05195] and (**B**) human MAX [P61244], (**C**) class V human ID4 [P47928], (**D**) class VI human HES1 [Q14469].
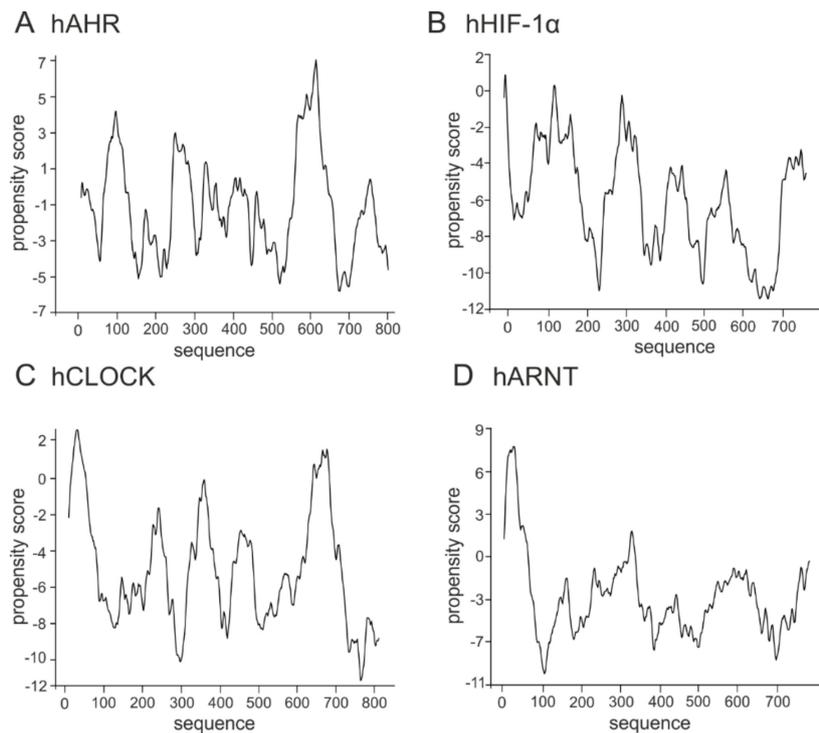
Figure 6. Prediction of propensity of LLPS formation for bHLh-PAS proteins. (**A**) human AHR [P35869], (**B**) human HIF-1α [Q16665], (**C**) human CLOCK [O08785], (**D**) human ARNT [P27540].

As the range of the propensity score is not determined precisely, as a control we performed catGranule prediction for proteins known to create LLPS: nucleophosmin (Figure 7A) and estrogen receptor (Figure 7B) which are deposited in the recently published PhaSePro database (https://phasepro.elte.hu) [141].

Figure 7. Prediction of propensity of LLPS formation for representative LLPS-enabled proteins. (**A**) nucleophosmin [P06748], (**B**) estrogen receptor [P03372].

Results of performed in silico analyses in comparison to the control show that the selected bHLH proteins have regions that might be involved in multivalent interaction leading to formation of liquid condensates. What would be their role in condensates formation and how would mutations and wrong dimerization/interaction influence formation of the bHLH TFs containing condensate remains a puzzle, however we believe that such an important family of TFs engaged in the crucial pathways and related to many severe disorders like cancer should be the subject of research in this field.

6. Concluding Remarks and Future Perspectives

In eukaryotic cells, regulation of transcription is a dynamic process which requires very precise temporal and spatial coordination of proteins assembling functional complexes. The bHLH family comprises a large group of TFs which utilize conserved DNA binding domain to interact with DNA, but also additional, often disordered domains and motives that allows formation of complex interacting network with various transcription co-factors. It is possible that flexible disordered regions of the bHLH proteins play a role in formation of liquid condensates via LLPS and contribute in this way to regulation of transcription process. Up to date however, there is a lack of experimental evidences. Also recently published PhaSePro database for LLPS does not contain any bHLH TF [141]. We believe that this is due to difficulties with the experimental studies of the bHLH proteins mentioned previously and we expect that some bHLH proteins will be appended in future.

Presented in the previous section predictions may give a hint about the link between LLPS by the bHLH proteins and transcription regulation. This raise a question about functional relevance of this discrepancy between family members. An interesting observation is the predicted low propensity score to form LLPS in the case of transcriptional repressors in contrast to proteins acting as activators. This raise a question about the functional relevance of this discrepancy between family members. Importantly, connection between LLPS and transcription regulation is not limited to the direct interaction between transcription regulators at the active transcription sites. LLPS form nuclear bodies, that maintain, store and modify transcription regulators. Examples include nuclear speckles, polyleukemia bodies, nucleolus, histone locus and others [142]. Within LLPS-formed condensates proteins can undergo acetylation/deacetylation or sumoylation, proteasome-dependent degradation and other posttranslational modifications that influence their functionality [143–145]. Importantly, barrier-free character of these phase separated condensates allows shuttling of its component between the condensates and nucleoplasm, and whenever needed molecules can be recruited from these compartments to the active transcriptionally sites. The discovery that LLPS which is well known in polymer chemistry can play an important role in molecular biology has definitely brought us closer to understanding the cell functionality and regulation of fundamental cellular processes such as transcription. However, our understanding and detailed knowledge is still residual. Many important questions regarding a LLPS concept in transcription regulation remain without answer. We do not know, which components drive association/dissociation events at the active sites. Which molecules serves as a scaffold conditioning formation of liquid condensates and which are just clients. How the type of client molecules influence the function of the phase separated condensates? Also, we do not know which factors and in which way alter LLPS leading to the pathological processes. What would be the role of the bHLH TFs in a condensates formation, and how mutations and incorrect dimerization/interaction of these proteins would impact formation and function of condensates? These questions, as well as many other ones await experimental verification. We believe that such important family of transcription factors which is engaged in crucial pathways and related to many severe diseases like cancer and neurodegenerative disorders, should be the subject of further intensive studies.

Author Contributions: A.T. and B.G.-M. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AHR	Aryl hydrocarbon receptor
AS-C	Achaete scute complex
ARNT	Aryl hydrocarbon receptor nuclear translocator
bHLH	Helix–loop–helix
ccRCC	Clear cell renal cell carcinoma
CLOCK	Circadian locomotor output cycles protein kaput
CTD	C-terminal domain
CycT1	Cyclin T1
EMC	Extramacrochaetae
E(spI)	Enhancer of split
ER	Estrogen receptor
ESC	Embryonic stem cells
GCE	Germ cell-expressed protein
GRO	Groucho
HAT	Histone transacetylase
HIF	Hypoxia-inducible factor
HSCa	Hematopoietic stem cells
HRD	histidine reach domain
ID	Inhibitor of DNA binding
IDPs	Intrinsically disordered proteins
IDRs	Intrinsically disordered regions
LLPS	liquid-liquid phase separation
LZ	Leucine zipper motif
MET	Methoprene-tolerant protein
MXI1	Max interacting protein
Ngn2	Neurogenin
NPAS	Neuronal PAS domain-containing protein
PAS	Period-arylhydrocarbon nuclear translocator-single minded domain
Pol II	RNA polymerase II
SE	Super-enhancer
SIM	Single-minded protein
SIMA	Similar protein
S/MARs	Scaffold/matrix associate regions
SREBP	Sterol-responsive element-binding protein
TAD	Transactivation domain
TAZ	Tafazzin
TFs	Transcription factors
TRH	Tracheales protein
USF	Upstream stimulatory factor
VHL	Von Hippel-Lindau tumor suppressor

References

1. Robinson, K.A. A network of yeast basic helix-loop-helix interactions. *Nucleic Acids Res.* **2000**, *28*, 4460–4466. [CrossRef] [PubMed]
2. Robinson, K.A. SURVEY AND SUMMARY: Saccharomyces cerevisiae basic helix-loop-helix proteins regulate diverse biological processes. *Nucleic Acids Res.* **2000**, *28*, 1499–1505. [PubMed]
3. Sailsbery, J.K.; Atchley, W.R.; Dean, R.A. Phylogenetic analysis and classification of the fungal bHLH domain. *Mol. Biol. Evol.* **2012**, *29*, 1301–1318. [CrossRef] [PubMed]
4. Pires, N.; Dolan, L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol. Biol. Evol.* **2010**, *27*, 862–874.
5. Massari, M.E.; Murre, C. Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **2000**, *20*, 429–440. [CrossRef]

6. Simionato, E.; Ledent, V.; Richards, G.; Thomas-Chollier, M.; Kerner, P.; Coornaert, D.; Degnan, B.M.; Vervoort, M. Origin and diversification of the basic helix-loop-helix gene family in metazoans: Insights from comparative genomics. *BMC Evol. Biol.* **2007**, *7*, 33. [CrossRef]

7. Ledent, V.; Paquet, O.; Vervoort, M. Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biol.* **2002**, *3*, 1–18. [CrossRef]

8. Wang, Y.; Chen, K.; Yao, Q.; Zheng, X.; Yang, Z. Phylogenetic Analysis of Zebrafish Basic Helix-Loop-Helix Transcription Factors. *J. Mol. Evol.* **2009**, *68*, 629–640.

9. Sailsbery, J.K.; Dean, R.A. Accurate discrimination of bHLH domains in plants, animals, and fungi using biologically meaningful sites. *BMC Evol. Biol.* **2012**, *12*, 154.

10. Murre, C. Helix–loop–helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes Dev.* **2019**, *33*, 6–25.

11. Murre, C.; McCaw, P.S.; Baltimore, D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **1989**, *56*, 777–783. [CrossRef]

12. Murre, C.; Bain, G.; van Dijk, M.A.; Engel, I.; Furnari, B.A.; Massari, M.E.; Matthews, J.R.; Quong, M.W.; Rivera, R.R.; Stuiver, M.H. Structure and function of helix-loop-helix proteins. *Biochim. Biophys. Acta (BBA) Gene Struct. Expr.* **1994**, *1281*, 129–135. [CrossRef]

13. Ephrussi, A.; Church, G.; Tonegawa, S.; Gilbert, W. B lineage–specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* **1985**, *227*, 134–140. [CrossRef] [PubMed]

14. Atchley, W.R.; Fitch, W.M. A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5172–5176. [CrossRef] [PubMed]

15. Skinner, M.K.; Rawls, A.; Wilson-Rawls, J.; Roalson, E.H. Basic helix-loop-helix transcription factor gene family phylogenetics and nomenclature. *Differ. Res. Biol. Divers.* **2010**, *80*, 1–8. [CrossRef]

16. Jones, S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* **2004**, *5*, 226. [CrossRef]

17. Wöhner, M.; Tagoh, H.; Bilic, I.; Jaritz, M.; Poliakova, D.K.; Fischer, M.; Busslinger, M. Molecular functions of the transcription factors E2A and E2-2 in controlling germinal center B cell and plasma cell development. *J. Exp. Med.* **2016**, *213*, 1201–1221. [CrossRef]

18. Yi, S.; Yu, M.; Yang, S.; Miron, R.J.; Zhang, Y. Tcf12, A Member of Basic Helix-Loop-Helix Transcription Factors, Mediates Bone Marrow Mesenchymal Stem Cell Osteogenic Differentiation In Vitro and In Vivo. *Stem Cells* **2017**, *35*, 386–397. [CrossRef]

19. Li, Y.; Brauer, P.M.; Singh, J.; Xhiku, S.; Yoganathan, K.; Zúñiga-Pflücker, J.C.; Anderson, M.K. Targeted Disruption of TCF12 Reveals HEB as Essential in Human Mesodermal Specification and Hematopoiesis. *Stem Cell Rep.* **2017**, *9*, 779–795. [CrossRef]

20. Quednow, B.B.; Brzózka, M.M.; Rossner, M.J. Transcription factor 4 (TCF4) and schizophrenia: Integrating the animal and the human perspective. *Cell. Mol. Life Sci.* **2014**, *71*, 2815–2835. [CrossRef]

21. Huang, C.; Chan, J.A.; Schuurmans, C. Proneural bHLH Genes in Development and Disease. In *Current Topics in Developmental Biology*; Academic Press: Cambridge, MA, USA, 2014; Volume 110, pp. 75–127.

22. Tan, T.K.; Zhang, C.; Sanda, T. Oncogenic transcriptional program driven by TAL1 in T-cell acute lymphoblastic leukemia. *Int. J. Hematol.* **2019**, *109*, 5–17. [CrossRef] [PubMed]

23. Choudhury, S. Genomics of the OLIG family of a bHLH transcription factor associated with oligo dendrogenesis. *Bioinformation* **2019**, *15*, 430–438. [CrossRef] [PubMed]

24. Bouard, C.; Terreux, R.; Honorat, M.; Manship, B.; Ansieau, S.; Vigneron, A.M.; Puisieux, A.; Payen, L. Deciphering the molecular mechanisms underlying the binding of the TWIST1/E12 complex to regulatory E-box sequences. *Nucleic Acids Res.* **2016**, *44*, 5470–5489. [CrossRef] [PubMed]

25. Whitfield, J.R.; Beaulieu, M.-E.; Soucek, L. Strategies to Inhibit Myc and Their Clinical Applicability. *Front. Cell Dev. Biol.* **2017**, *5*, 10. [CrossRef] [PubMed]

26. Amati, B.; Land, H. Myc-Max-Mad: A transcription factor network controlling cell cycle progression, differentiation and death. *Curr. Opin. Genet. Dev.* **1994**, *4*, 102–108. [CrossRef]

27. Lasorella, A.; Benezra, R.; Iavarone, A. The ID proteins: Master regulators of cancer stem cells and tumour aggressiveness. *Nat. Rev. Cancer* **2014**, *14*, 77–91. [CrossRef]

28. Weber, D.; Wiese, C.; Gessler, M. Hey bHLH transcription factors. In *Current Topics in Developmental Biology*; Academic Press: Cambridge, MA, USA, 2014; Volume 110, pp. 285–315.

29. Iso, T.; Kedes, L.; Hamamori, Y. HES and HERP families: Multiple effectors of the notch signaling pathway. *J. Cell. Physiol.* **2003**, *194*, 237–255. [CrossRef]

30. Saha, T.T.; Shin, S.W.; Dou, W.; Roy, S.; Zhao, B.; Hou, Y.; Wang, X.-L.; Zou, Z.; Girke, T.; Raikhel, A.S. Hairy and Groucho mediate the action of juvenile hormone receptor Methoprene-tolerant in gene repression. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E735–E743. [CrossRef]

31. Wright, E.J.; Pereira De Castro, K.; Joshi, A.D.; Elferink, C.J. Canonical and non-canonical aryl hydrocarbon receptor signaling pathways. *Curr. Opin. Toxicol.* **2017**, *2*, 87–92. [CrossRef]

32. Semenza, G.L. Hypoxia-inducible factors in physiology and medicine. *Cell* **2012**, *148*, 399–408. [CrossRef]

33. Moffett, P.; Pelletier, J. Different transcriptional properties of mSim-1 and mSim-2. *FEBS Lett.* **2000**, *466*, 80–86. [CrossRef]

34. Lee, Y.; Lee, J.; Kwon, I.; Nakajima, Y.; Ohmiya, Y.; Son, G.H.; Lee, K.H.; Kim, K. Coactivation of the CLOCK-BMAL1 complex by CBP mediates resetting of the circadian clock. *J. Cell Sci.* **2010**, *123*, 3547–3557. [CrossRef] [PubMed]

35. Teh, C.H.L.; Lam, K.K.Y.; Loh, C.C.; Loo, J.M.; Yan, T.; Lim, T.M. Neuronal PAS domain protein 1 is a transcriptional repressor and requires arylhydrocarbon nuclear translocator for its nuclear localization. *J. Biol. Chem.* **2006**, *281*, 34617–34629. [CrossRef] [PubMed]

36. Ohsawa, S.; Hamada, S.; Kakinuma, Y.; Yagi, T.; Miura, M. Novel function of neuronal PAS domain protein 1 in erythropoietin expression in neuronal cells. *J. Neurosci. Res.* **2005**, *79*, 451–458. [CrossRef]

37. Gilles-Gonzalez, M.-A.; Gonzales, G. Signal transduction by heme-containing PAS-domain proteins. *J. Appl. Physiol.* **2004**, *96*, 774–783. [CrossRef] [PubMed]

38. Kamnasaran, D. Disruption of the neuronal PAS3 gene in a family affected with schizophrenia. *J. Med Genet.* **2003**, *40*, 325–332. [CrossRef] [PubMed]

39. Ooe, N.; Saito, K.; Mikami, N.; Nakatuka, I.; Kaneko, H. Identification of a Novel Basic Helix-Loop-Helix-PAS Factor, NXF, Reveals a Sim2 Competitive, Positive Regulatory Role in Dendritic-Cytoskeleton Modulator Drebrin Gene Expression. *Mol. Cell. Biol.* **2003**, *24*, 608–616. [CrossRef]

40. Li, M.; Mead, E.A.; Zhu, J. Heterodimer of two bHLH-PAS proteins mediates juvenile hormone-induced gene expression. *Proc. Natl. Acad. Sci. USA* **2011**, *2011*, 638–643. [CrossRef]

41. Brunnberg, S.; Pettersson, K.; Rydin, E.; Matthews, J.; Hanberg, A.; Pongratz, I. The basic helix-loop-helix-PAS protein ARNT functions as a potent coactivator of estrogen receptor-dependent transcription. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6517–6522. [CrossRef]

42. Sharma, P.; Chinaranagari, S.; Chaudhary, J. Inhibitor of differentiation 4 (ID4) acts as an inhibitor of ID-1, -2 and -3 and promotes basic helix loop helix (bHLH) E47 DNA binding and transcriptional activity. *Biochimie* **2015**, *112*, 139–150. [CrossRef]

43. Zelzer, E.; Wappner, P.; Shilo, B.-Z. The PAS domain confers target gene specificity of Drosophila bHLH/PAS proteins. *Genes Dev.* **1997**, *11*, 2079–2089. [CrossRef] [PubMed]

44. Cusanovich, M.A.; Meyer, T.E. Photoactive yellow protein: A prototypic PAS domain sensory protein and development of a common signaling mechanism. *Biochemistry* **2003**, *42*, 4759–4770. [CrossRef] [PubMed]

45. Mö Glich, A.; Ayers, R.A.; Moffat, K. Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Struct. Fold. Des.* **2009**, *17*, 1282–1294. [CrossRef] [PubMed]

46. Roeder, R.G. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* **2005**, *579*, 909–915. [CrossRef]

47. Ansari, S.A.; Morse, R.H. Mechanisms of Mediator complex action in transcriptional activation. *Cell. Mol. Life Sci.* **2013**, *70*, 2743–2756. [CrossRef]

48. Conaway, R.C.; Conaway, J.W. Function and regulation of the Mediator complex. *Curr. Opin. Genet. Dev.* **2011**, *21*, 225–230. [CrossRef]

49. Poss, Z.C.; Ebmeier, C.C.; Taatjes, D.J. The Mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 575–608. [CrossRef]

50. Quevedo, M.; Meert, L.; Dekker, M.R.; Dekkers, D.H.W.; Brandsma, J.H.; van den Berg, D.L.C.; Ozgür, Z.; van IJcken, W.F.J.; Demmers, J.; Fornerod, M.; et al. Mediator complex interaction partners organize the transcriptional network that defines neural stem cells. *Nat. Commun.* **2019**, *10*, 2669. [CrossRef]

51. Malik, S.; Roeder, R.G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* **2010**, *11*, 761–772. [CrossRef]

52. Malik, N.; Agarwal, P.; Tyagi, A. Emerging functions of multi-protein complex Mediator with special emphasis on plants. *Crit. Rev. Biochem. Mol. Biol.* **2017**, *52*, 475–502. [CrossRef]

53. An, C.; Mou, Z. The function of the Mediator complex in plant immunity. *Plant Signal. Behav.* **2013**, *8*, e23182. [CrossRef] [PubMed]

54. Chen, R.; Jiang, H.; Li, L.; Zhai, Q.; Qi, L.; Zhou, W.; Liu, X.; Li, H.; Zheng, W.; Sun, J.; et al. The Arabidopsis Mediator Subunit MED25 Differentially Regulates Jasmonate and Abscisic Acid Signaling through Interacting with the MYC2 and ABI5 Transcription Factors. *Plant Cell* **2012**, *24*, 2898–2916. [CrossRef] [PubMed]

55. An, C.; Li, L.; Zhai, Q.; You, Y.; Deng, L.; Wu, F.; Chen, R.; Jiang, H.; Wang, H.; Chen, Q.; et al. Mediator subunit MED25 links the jasmonate receptor to transcriptionally active chromatin. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8930–E8939. [CrossRef] [PubMed]

56. Liu, Y.; Du, M.; Deng, L.; Shen, J.; Fang, M.; Chen, Q.; Lu, Y.; Wang, Q.; Li, C.; Zhai, Q. Myc2 regulates the termination of jasmonate signaling via an autoregulatory negative feedback loop[open]. *Plant Cell* **2019**, *31*, 106–127. [CrossRef]

57. Li, X.; Yang, R.; Chen, H. The arabidopsis thaliana mediator subunit MED8 regulates plant immunity to botrytis cinerea through interacting with the basic helix-loop-helix (bHLH) transcription factor FAMA. *PLoS ONE* **2018**, *13*, e0193458. [CrossRef]

58. Oliner, J.D.; Andresen, J.M.; Hansen, S.K.; Zhou, S.; Tjian, R. SREBP transcriptional activity is mediated through an interaction with the CREB-binding protein. *Genes Dev.* **1996**, *10*, 2903–2911. [CrossRef]

59. Yang, F.; Vought, B.W.; Satterlee, J.S.; Walker, A.K.; Jim Sun, Z.Y.; Watts, J.L.; DeBeaumont, R.; Mako Saito, R.; Hyberts, S.G.; Yang, S.; et al. An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **2006**, *442*, 700–704. [CrossRef]

60. Pacheco, D.; Warfield, L.; Brajcich, M.; Robbins, H.; Luo, J.; Ranish, J.; Hahn, S. Transcription Activation Domains of the Yeast Factors Met4 and Ino2: Tandem Activation Domains with Properties Similar to the Yeast Gcn4 Activator. *Mol. Cell. Biol.* **2018**, *38*, e00038-18. [CrossRef]

61. Lazrak, M.; Deleuze, V.; Noel, D.; Haouzi, D.; Chalhoub, E.; Dohet, C.; Robbins, I.; Mathieu, D. The bHLH TAL-1/SCL regulates endothelial cell migration and morphogenesis. *J. Cell Sci.* **2004**, *117*, 1161–1171. [CrossRef]

62. Huang, S.; Qiu, Y.; Stein, R.W.; Brandt, S.J. p300 functions as a transcriptional coactivator for the TAL1/SCL oncoprotein. *Oncogene* **1999**, *18*, 4958–4967. [CrossRef]

63. Puri, P.L.; Avantaggiati, M.L. p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription arrest in the G 0 phase, irreversible exit from the cell cycle. *EMBO J.* **1997**, *16*, 369–383. [CrossRef] [PubMed]

64. Cao, Y.; Yao, Z.; Sarkar, D.; Lawrence, M.; Sanchez, G.J.; Parker, M.H.; MacQuarrie, K.L.; Davison, J.; Morgan, M.T.; Ruzzo, W.L.; et al. Genome-wide MyoD Binding in Skeletal Muscle Cells: A Potential for Broad Cellular Reprogramming. *Dev. Cell* **2010**, *18*, 662–674. [CrossRef] [PubMed]

65. Raposo, A.A.S.F.; Vasconcelos, F.F.; Drechsel, D.; Marie, C.; Johnston, C.; Dolle, D.; Bithell, A.; Gillotin, S.; van den Berg, D.L.C.; Ettwiller, L.; et al. Ascl1 coordinately regulates gene expression and the chromatin landscape during neurogenesis. *Cell Rep.* **2015**, *10*, 1544–1556. [CrossRef] [PubMed]

66. Ko, C.-I.; Puga, A. Does the aryl hydrocarbon receptor regulate pluripotency? *Curr. Opin. Toxicol.* **2017**, *2*, 1–7. [CrossRef] [PubMed]

67. Yao, X.; Tan, J.; Lim, K.J.; Koh, J.; Ooi, W.F.; Li, Z.; Huang, D.; Xing, M.; Chan, Y.S.; Qu, J.Z.; et al. VHL deficiency drives enhancer activation of oncogenes in clear cell renal cell carcinoma. *Cancer Discov.* **2017**, *7*, 1284–1305. [CrossRef] [PubMed]

68. Doi, M.; Hirayama, J.; Sassone-Corsi, P. Circadian regulator CLOCK is a histone acetyltransferase. *Cell* **2006**, *125*, 497–508. [CrossRef] [PubMed]

69. Harada, A.; Ohkawa, Y.; Imbalzano, A.N. Temporal regulation of chromatin during myoblast differentiation. *Semin. Cell Dev. Biol.* **2017**, *72*, 77–86. [CrossRef]

70. Heisig, J.; Weber, D.; Englberger, E.; Winkler, A.; Kneitz, S.; Sung, W.-K.; Wolf, E.; Eilers, M.; Wei, C.-L.; Gessler, M. Target gene analysis by microarrays and chromatin immunoprecipitation identifies HEY proteins as highly redundant bHLH repressors. *PLoS Genet.* **2012**, *8*, e1002728. [CrossRef]

71. Rahl, P.B.; Lin, C.Y.; Seila, A.C.; Flynn, R.A.; McCuine, S.; Burge, C.B.; Sharp, P.A.; Young, R.A. c-Myc Regulates Transcriptional Pause Release. *Cell* **2010**, *141*, 432–445. [CrossRef]

72. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230. [CrossRef]

73. Uversky, V.N. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci. A Publ. Protein Soc.* **2013**, *22*, 693–724. [CrossRef] [PubMed]

74.　Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208. [CrossRef] [PubMed]

75.　Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [CrossRef]

76.　Lu, H.; Yu, D.; Hansen, A.S.; Ganguly, S.; Liu, R.; Heckert, A.; Darzacq, X.; Zhou, Q. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **2018**, *558*, 318. [CrossRef] [PubMed]

77.　Uversky, V.N.; Dunker, A.K. Understanding protein non-folding. *Biochim. Biophys. Acta (BBA) Proteins Proteom.* **2010**, *1804*, 1231–1264. [CrossRef]

78.　Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol.* **2005**, *347*, 827–839. [CrossRef]

79.　Campen, A.; Williams, R.; Brown, C.; Meng, J.; Uversky, V.; Dunker, A. TOP-IDP-Scale: A New Amino Acid Scale Measuring Propensity for Intrinsic Disorder. *Protein Pept. Lett.* **2008**, *15*, 956–963. [CrossRef]

80.　Williams, R.M.; Obradovi, Z.; Mathura, V.; Braun, W.; Garner, E.C.; Young, J.; Takayama, S.; Brown, C.J.; Dunker, A.K. The protein non-folding problem: Amino acid determinants of intrinsic order and disorder. *Pac. Symp. Biocomput.* **2001**, *2001*, 89–100.

81.　Vacic, V.; Uversky, V.N.; Dunker, A.K.; Lonardi, S. Composition Profiler: A tool for discovery and visualization of amino acid composition differences. *BMC Bioinform.* **2007**, *8*, 211. [CrossRef]

82.　Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins Struct. Funct. Genet.* **2000**, *41*, 415–427. [CrossRef]

83.　Dunker, A.K.; Obradovic, Z. The protein trinity—Linking function and disorder. *Nat. Biotechnol.* **2001**, *19*, 805. [CrossRef] [PubMed]

84.　Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* **2005**, *579*, 3346–3354. [CrossRef] [PubMed]

85.　Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R.J.; Daughdrill, G.W.; Dunker, A.K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D.T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631. [CrossRef] [PubMed]

86.　Peng, Z.; Yan, J.; Fan, X.; Mizianty, M.J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V.N.; Kurgan, L. Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* **2014**, *72*, 137–151. [CrossRef]

87.　Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [CrossRef]

88.　Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **2008**, *33*, 2–8. [CrossRef]

89.　Iakoucheva, L.M.; Radivojac, P.; Brown, C.J.; O'Connor, T.R.; Sikes, J.G.; Obradovic, Z.; Dunker, A.K. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **2004**, *32*, 1037–1049. [CrossRef]

90.　Hatakeyama, J.; Kageyama, R. Retinal cell fate determination and bHLH factors. *Semin. Cell Dev. Biol.* **2004**, *15*, 83–89. [CrossRef]

91.　Sölter, M.; Locker, M.; Boy, S.; Taelman, V.; Bellefroid, E.J.; Perron, M.; Pieler, T. Characterization and function of the bHLH-O protein XHes2: Insight into the mechanism controlling retinal cell fate decision. *Development* **2006**, *133*, 4097–4108. [CrossRef]

92.　Parker, M.H.; Perry, R.L.S.; Fauteux, M.C.; Berkes, C.A.; Rudnicki, M.A. MyoD Synergizes with the E-Protein HEB To Induce Myogenic Differentiation. *Mol. Cell. Biol.* **2006**, *26*, 5771–5783. [CrossRef]

93.　McDowell, G.S.; Hindley, C.J.; Lippens, G.; Landrieu, I.; Philpott, A. Phosphorylation in intrinsically disordered regions regulates the activity of Neurogenin2. *BMC Biochem.* **2014**, *15*, 24. [CrossRef] [PubMed]

94.　Aguado-Llera, D.; Goormaghtigh, E.; de Geest, N.; Quan, X.-J.; Prieto, A.; Hassan, B.A.; Gómez, J.; Neira, J.L. The basic helix-loop-helix region of human neurogenin 1 is a monomeric natively unfolded protein which forms a "fuzzy" complex upon DNA binding. *Biochemistry* **2010**, *49*, 1577–1589. [CrossRef] [PubMed]

95. Panova, S.; Cliff, M.J.; Macek, P.; Blackledge, M.; Jensen, M.R.; Nissink, J.W.M.; Embrey, K.J.; Davies, R.; Waltho, J.P. Mapping Hidden Residual Structure within the Myc bHLH-LZ Domain Using Chemical Denaturant Titration. *Structure* **2019**, *27*, 1537–1546. [CrossRef] [PubMed]

96. Romero, P.; Obradovic, Z.; Dunker, A.K. Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. *Genome Inform. Workshop Genome Inform.* **1997**, *8*, 110–124.

97. Li, X.; Romero, P.; Rani, M.; Dunker, A.; Obradovic, Z. Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome Inform. Workshop Genome Inform.* **1999**, *10*, 30–40.

98. Cilia, E.; Pancsa, R.; Tompa, P.; Lenaerts, T.; Vranken, W.F. From protein sequence to dynamics and disorder with DynaMine. *Nat. Commun.* **2013**, *4*, 2741. [CrossRef]

99. Cilia, E.; Pancsa, R.; Tompa, P.; Lenaerts, T.; Vranken, W.F. The DynaMine webserver: Predicting protein dynamics from sequence. *Nucleic Acids Res.* **2014**, *42*, W264–W270. [CrossRef]

100. Furness, S.G.B.; Lees, M.J.; Whitelaw, M.L. The dioxin (aryl hydrocarbon) receptor as a model for adaptive responses of bHLH/PAS transcription factors. *FEBS Lett.* **2007**, *581*, 3616–3625. [CrossRef]

101. Brangwynne, C.P.; Eckmann, C.R.; Courson, D.S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A.A. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **2009**, *324*, 1729–1732. [CrossRef]

102. Molliex, A.; Temirov, J.; Lee, J.; Coughlin, M.; Kanagaraj, A.P.; Kim, H.J.; Mittag, T.; Taylor, J.P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **2015**, *163*, 123–133. [CrossRef]

103. Nott, T.J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T.D.; Bazett-Jones, D.P.; Pawson, T.; Forman-Kay, J.D.; et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* **2015**, *57*, 936–947. [CrossRef] [PubMed]

104. Elbaum-Garfinkle, S.; Kim, Y.; Szczepaniak, K.; Chen, C.C.H.; Eckmann, C.R.; Myong, S.; Brangwynne, C.P. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7189–7194. [CrossRef] [PubMed]

105. Lai, J.; Koh, C.H.; Tjota, M.; Pieuchot, L.; Raman, V.; Chandrababu, K.B.; Yang, D.; Wong, L.; Jedd, G. Intrinsically disordered proteins aggregate at fungal cell-to-cell channels and regulate intercellular connectivity. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15781–15786. [CrossRef] [PubMed]

106. Mitrea, D.M.; Cika, J.A.; Guy, C.S.; Ban, D.; Banerjee, P.R.; Stanley, C.B.; Nourse, A.; Deniz, A.A.; Kriwacki, R.W. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife* **2016**, *5*, e13571. [CrossRef] [PubMed]

107. Brangwynne, C.P.; Tompa, P.; Pappu, R.V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015**, *11*, 899–904. [CrossRef]

108. Banani, S.F.; Rice, A.M.; Peeples, W.B.; Lin, Y.; Jain, S.; Parker, R.; Rosen, M.K. Compositional Control of Phase-Separated Cellular Bodies. *Cell* **2016**, *166*, 651–663. [CrossRef]

109. Ditlev, J.A.; Case, L.B.; Rosen, M.K. Who's In and Who's Out—Compositional Control of Biomolecular Condensates. *J. Mol. Biol.* **2018**, *430*, 4666–4684. [CrossRef]

110. Harmon, T.S.; Holehouse, A.S.; Rosen, M.K.; Pappu, R.V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *eLife* **2017**, *6*, e30294. [CrossRef]

111. Posey, A.E.; Holehouse, A.S.; Pappu, R.V. Phase Separation of Intrinsically Disordered Proteins. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2018; ISBN 9780128156490.

112. Berezney, R.; Coffey, D.S. Identification of a nuclear protein matrix. *Biochem. Biophys. Res. Commun.* **1974**, *60*, 1410–1417. [CrossRef]

113. Linnemann, A.K.; Platts, A.E.; Krawetz, S.A. Differential nuclear scaffold/matrix attachment marks expressed genes. *Hum. Mol. Genet.* **2009**, *18*, 645–654. [CrossRef]

114. Engelke, R.; Riede, J.; Hegermann, J.; Wuerch, A.; Eimer, S.; Dengjel, J.; Mittler, G. The quantitative nuclear matrix proteome as a biochemical snapshot of nuclear organization. *J. Proteome Res.* **2014**, *13*, 3940–3956. [CrossRef] [PubMed]

115. Weber, S.C.; Brangwynne, C.P. Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.* **2015**, *25*, 641–646. [CrossRef] [PubMed]

116. Berry, J.; Weber, S.C.; Vaidya, N.; Haataja, M.; Brangwynne, C.P. RNA transcription modulates phase transition-driven nuclear body assembly. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5237–E5245. [CrossRef] [PubMed]

117. Gibson, B.A.; Doolittle, L.K.; Jensen, L.E.; Gamarra, N.; Redding, S.; Rosen, M.K. Organization and Regulation of Chromatin by Liquid-Liquid Phase Separation. *bioRxiv* **2019**. [CrossRef]

118. Duronio, R.J.; Marzluff, W.F. Coordinating cell cycle-regulated histone gene expression through assembly and function of the Histone Locus Body. *RNA Biol.* **2017**, *14*, 726–738. [CrossRef]

119. Mitrea, D.M.; Kriwacki, R.W. Phase separation in biology; Functional organization of a higher order Short linear motifs—The unexplored frontier of the eukaryotic proteome. *Cell Commun. Signal.* **2016**, *14*, 1–20. [CrossRef]

120. Wu, T.; Lu, Y.; Gutman, O.; Lu, H.; Zhou, Q.; Henis, Y.I.; Luo, K. Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression. *bioRxiv* **2019**. [CrossRef]

121. Sabari, B.R.; Dall'Agnese, A.; Boija, A.; Klein, I.A.; Coffey, E.L.; Shrinivas, K.; Abraham, B.J.; Hannett, N.M.; Zamudio, A.V.; Manteiga, J.C.; et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **2018**, *361*, eaar3958. [CrossRef]

122. Boija, A.; Klein, I.A.; Sabari, B.R.; Dall'Agnese, A.; Coffey, E.L.; Zamudio, A.V.; Li, C.H.; Shrinivas, K.; Manteiga, J.C.; Hannett, N.M.; et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **2018**, *175*, 1842–1855.e16. [CrossRef]

123. Chen, X.; Xu, H.; Yuan, P.; Fang, F.; Huss, M.; Vega, V.B.; Wong, E.; Orlov, Y.L.; Zhang, W.; Jiang, J.; et al. Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. *Cell* **2008**, *133*, 1106–1117. [CrossRef]

124. Whyte, W.A.; Orlando, D.A.; Hnisz, D.; Abraham, B.J.; Lin, C.Y.; Kagey, M.H.; Rahl, P.B.; Lee, T.I.; Young, R.A. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **2013**, *153*, 307–319. [CrossRef] [PubMed]

125. Hnisz, D.; Shrinivas, K.; Young, R.A.; Chakraborty, A.K.; Sharp, P.A. A Phase Separation Model for Transcriptional Control. *Cell* **2017**, *169*, 13–23. [CrossRef] [PubMed]

126. Mansour, M.R.; Abraham, B.J.; Anders, L.; Berezovskaya, A.; Gutierrez, A.; Durbin, A.D.; Etchin, J.; Lee, L.; Sallan, S.E.; Silverman, L.B.; et al. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **2014**, *346*, 1373–1377. [CrossRef] [PubMed]

127. Brown, J.D.; Lin, C.Y.; Duan, Q.; Griffin, G.; Federation, A.J.; Paranal, R.M.; Bair, S.; Newton, G.; Lichtman, A.H.; Kung, A.L.; et al. Nf-kb directs dynamic super enhancer formation in inflammation and atherogenesis. *Mol. Cell* **2014**, *56*, 219–231. [CrossRef] [PubMed]

128. Lovén, J.; Hoke, H.A.; Lin, C.Y.; Lau, A.; Orlando, D.A.; Vakoc, C.R.; Bradner, J.E.; Lee, T.I.; Young, R.A. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **2013**, *153*, 320–334. [CrossRef] [PubMed]

129. Apostolou, E.; Ferrari, F.; Walsh, R.M.; Bar-Nur, O.; Stadtfeld, M.; Cheloufi, S.; Stuart, H.T.; Polo, J.M.; Ohsumi, T.K.; Borowsky, M.L.; et al. Genome-wide Chromatin Interactions of the Nanog Locus in Pluripotency, Differentiation, and Reprogramming. *Cell Stem Cell* **2013**, *12*, 699–712. [CrossRef] [PubMed]

130. Manavathi, B.; Samanthapudi, V.S.K.; Gajulapalli, V.N.R. Estrogen receptor coregulators and pioneer factors: The orchestrators of mammary gland cell fate and development. *Front. Cell Dev. Biol.* **2014**, *12*, 34. [CrossRef]

131. Kagey, M.H.; Newman, J.J.; Bilodeau, S.; Zhan, Y.; Orlando, D.A.; Van Berkum, N.L.; Ebmeier, C.C.; Goossens, J.; Rahl, P.B.; Levine, S.S.; et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **2010**, *467*, 430. [CrossRef]

132. Heery, D.M.; Kalkhoven, E.; Hoare, S.; Parker, M.G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **1997**, *387*, 733. [CrossRef]

133. Cagas, P.M.; Corden, J.L. Structural studies of a synthetic peptide derived from the carboxyl-terminal domain of RNA polymerase II. *Proteins Struct. Funct. Bioinform.* **1995**, *21*, 149–160. [CrossRef]

134. Portz, B.; Lu, F.; Gibbs, E.B.; Mayfield, J.E.; Rachel Mehaffey, M.; Zhang, Y.J.; Brodbelt, J.S.; Showalter, S.A.; Gilmour, D.S. Structural heterogeneity in the intrinsically disordered RNA polymerase II C-terminal domain. *Nat. Commun.* **2017**, *8*, 15231. [CrossRef] [PubMed]

135. Corden, J.L.; Cadena, D.L.; Ahearn, J.M.; Dahmus, M.E. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 7934–7938. [CrossRef] [PubMed]

136. Dahmus, M.E. Reversible Phosphorylation of the C-terminal Domain of RNA Polymerase II. *J. Biol. Chem.* **1996**, *271*, 19009–19012. [CrossRef]

137. Wippich, F.; Bodenmiller, B.; Trajkovska, M.G.; Wanka, S.; Aebersold, R.; Pelkmans, L. Dual specificity kinase DYRK3 couples stress granule condensation/ dissolution to mTORC1 signaling. *Cell* **2013**, *152*, 791–805. [CrossRef] [PubMed]

138. Rai, A.K.; Chen, J.X.; Selbach, M.; Pelkmans, L. Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* **2018**, *559*, 211. [CrossRef] [PubMed]

139. Boehning, M.; Dugast-Darzacq, C.; Rankovic, M.; Hansen, A.S.; Yu, T.; Marie-Nelly, H.; McSwiggen, D.T.; Kokic, G.; Dailey, G.M.; Cramer, P.; et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **2018**, *25*, 833. [CrossRef] [PubMed]

140. Bolognesi, B.; Lorenzo Gotor, N.; Dhar, R.; Cirillo, D.; Baldrighi, M.; Tartaglia, G.G.; Lehner, B. A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell Rep.* **2016**, *16*, 222–231. [CrossRef] [PubMed]

141. Mészáros, B.; Erdős, G.; Szabó, B.; Schád, É.; Tantos, Á.; Abukhairan, R.; Horváth, T.; Murvai, N.; Kovács, O.P.; Kovács, M.; et al. PhaSePro: The database of proteins driving liquid–liquid phase separation. *Nucleic Acids Res.* **2019**. [CrossRef]

142. Darling, A.L.; Liu, Y.; Oldfield, C.J.; Uversky, V.N. Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **2018**, *18*, 1700193. [CrossRef]

143. Hyman, A.A.; Weber, C.A.; Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58. [CrossRef]

144. Uversky, V.N. 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. [CrossRef] [PubMed]

145. Shin, Y.; Brangwynne, C.P. Liquid phase condensation in cell physiology and disease. *Science* **2017**, *357*, eaaf4382. [CrossRef] [PubMed]