RNA BIOLOGY
2019, VOL. 16, NO. 11, 1604–1621
https://doi.org/10.1080/15476286.2019.1650612

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

An IRES-dependent translation of HYPK mRNA generates a truncated isoform of the protein that lacks the nuclear localization and functional ability

Debasish Kumar Ghosh[a,b] and Akash Ranjan[a]

[a]Computational and Functional Genomics Group, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, Telangana, India; [b]Graduate studies, Manipal Academy of Higher Education, Manipal, Karnataka, India

ABSTRACT

Different mechanisms of translation initiation process exist to start the protein synthesis from various viral and eukaryotic mRNA. The cap-independent and tertiary structure directed translation initiation of mRNAs forms the basis of internal ribosome entry site (IRES) mediated translation initiation that helps in cellular protein production in different conditions. HYPK protein sequesters different aggregation-prone proteins to help in the cellular proteostasis. HYPK mRNA is differentially translated from an internal start/initiation codon to generate an amino terminal-truncated isoform (HSPC136) of HYPK protein. In this study, we report that an IRES-dependent translation initiation of HYPK mRNA results in the formation of the HSPC136/HYPK-ΔN isoform of HYPK protein. The IRES-driven translation product, HYPK-ΔN, lacks the N-terminal tri-arginine motif that acts as the nuclear localization signal (NLS) in the full-length HYPK protein. While the full-length HYPK protein translocates to the nucleus and prevents the aggregation of the mutant p53 (p53-R248Q) protein, the HYPK-ΔN lacks this activity. The NLS of HYPK is not evolutionarily conserved and its exclusive presence in the HYPK of higher eukaryotic animals imparts additional advantage to the HYPK protein in tackling the cytosolic as well as nuclear protein aggregates. The presence of the NLS in full-length HYPK also allows this protein to modulate the cell cycle. These results provide a mechanistic detail of HYPK mRNA's translation initiation control by an IRES that dictates the formation of HYPC136/HYPK-ΔN which lacks the nuclear localization and functional ability.

ARTICLE HISTORY
Received 11 April 2019
Revised 20 July 2019
Accepted 25 July 2019

KEYWORDS
HYPK; internal ribosome entry site; translation; nuclear localization signal; nuclear proteostasis

Introduction

Huntingtin interacting protein K (HYPK) is an important protein that maintains cellular proteostasis by preventing the aggregation of several aggregation-prone proteins [1], while also functioning in the cell cycle regulation process [2]. Though HYPK was first identified as ahuntingtin interacting protein [3], subsequent studies have shown that it is also able to bind other aggregation-prone proteins, like α-synuclein-A53T and superoxide dismutase 1-G93A [1]. HYPK is known to possess chaperone-like activity [4]. It also shows the property of sequestering toxic protein aggregates by using its annular-shaped structural complexes that are known as H-granules [1]. The formation of the annular shaped H-granules of HYPK occurs due to its self-association into different hierarchical structures. Self-association of HYPK depends upon its C-terminal hydrophobic patches and a low complexity region (LCR) [1]. In fact, the annular structures of HYPK result from the self-association of the seeds of its C-terminal ubiquitin-associated domain (UBA) [5]. A metastable state of the HYPK-UBA seed is responsible for its aggregation [5]. In the physiological condition, the self-association of HYPK is prevented by the activity of its N-terminal region. The N-terminus of HYPK is a disordered nanostructure that loops towards the C-terminus to shield the positive charges of the LCR [6]. Such an intra-molecular interaction prevents the HYPK aggregation. HYPK is reported to associate with the N-acetyl transferase (Nat) proteins at the translating mature ribosome [7]. Very recently, the co-complex structure of HYPK-NatA has been solved [8]. Other than the Nat proteins, HYPK also interacts with a wide range of cytoplasmic and nuclear proteins [2]. These interactions are presumed to regulate the critical stages of cell cycle, cell growth and cell death phenomena [2]. In *Caenorhabditis elegans*, HYPK aggregates form proteasome blocking complexes [9] that finally regulate the ageing process of the organism [10]. There are not many reports that describe the regulation of HYPK expression, except two independent studies that state the transcriptional regulation of HYPK gene by HSF1 transcription factor [11,12]. Human HYPK protein is expressed as different isoforms. The longest isoform (isoform-1) of HYPK is 129 amino acids long. While the isoform-2 is shorter (81 amino acids long), there is the expression of another N-terminal truncated form of HYPK isoform-1 (121 amino acids long) which is annotated as HSPC136. The cellular localization of HYPK is also intriguing. Though HYPK mostly remains as a cytosolic protein, it also has the ability to be transported to the nucleus. Expression of variable isoforms of HYPK and differential cellular localization of HYPK warrants more studies to understand the post-transcriptional expression regulation of HYPK that correlates to this protein's specific cellular localization. In

CONTACT Akash Ranjan akash@cdfd.org.in Computational and Functional Genomics Group, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, Telangana 500039, India

this study, we find that the translation initiation of the HYPK mRNA to generate the N-terminal truncated isoform-1 of HYPK (i.e. HSPC136) is driven by the internal ribosome entry site (IRES)-dependent translation and this protein is incapable of showing nuclear presence. We also explicitly identify the nuclear localization signal (NLS) in the full-length HYPK and the functional consequences of HYPK's nuclear localization.

IRES-mediated translation initiation is a major mode of non-canonical translation initiation [13] that ensures the cap-independent continuous translation of different mRNAs in the physiological and stressful conditions [14]. The increasing number of mRNAs, that undergo translation initiation through IRES, signify the evolutionary importance of this mechanism in the cellular translation process. The IRES elements are highly structured regions that are found in the 5′UTR (5′ untranslated region) of several viral RNAs and eukaryotic mRNAs [15,16]. IRESs are primarily characterized by the GC-rich structured regions in the 5′UTR of mRNAs. The secondary (like, hairpin-loop) and tertiary (formed by the tertiary folding of the secondary structures) structures of 5′UTRs can serve as an alternative site of deposition of the 40S subunit of the eukaryotic ribosome. Several of the eukaryotic translation initiation factors (eIFs), like eIF2, eIF3, eIF4A, eIF4B, eIF4G, etc., can help the deposition of 40S subunit to the IRES site [17]. The 40S subunit then scans the mRNA to start the translation from the original or an alternative start codon [18]. The structural motifs of IRES elements vary considerably in different mRNAs. While some of the IRESs are present as a continuous single stretch of nucleotide sequence in the respective mRNAs, some IRESs are non-contiguous in the mRNA sequences. Some of the IRESs of viral RNAs are structurally compact and functionally robust [19]. Translation initiation activity by such viral IRESs does not require the functions of eIFs [20]. On the other hand, some of the IRESs of eukaryotic and viral mRNAs are situated in an extended region of the RNAs and they can acquire the compact folded structure only during the initiation of the translation process that is coupled to binding of some of the eIFs to the ribosome [21]. Few of the IRESs are located in the extended region of RNAs and they are structurally more flexible [22]. The flexibility of IRES elements requires the binding of most of the eIFs to ribosome at the time of initiation of the translation process. Not only the eIFs, but the noncanonical IRES trans-acting factors (ITAFs) also play a significant role in initiating translation from several of the IRES elements of eukaryotic mRNAs. It is believed that ITAFs help in properly folding of the IRESs to their functional forms. The optimum function of IRES depends upon the environmental and cellular conditions. The activity of viral IRESs depends upon the host responses, like anti-viral response, protease activity, altered cell cycle conditions, etc. On the contrary, activity of IRESs of eukaryotic mRNAs depends upon the cellular stress (like hypoxia, growth factor depletion, heat shock, proteotoxicity, apoptosis, etc.) and regulated interaction of the IRES elements with various eIFs [23]. The IRES-dependent translation has been selected and maintained as an evolutionary beneficial process. It enables the cells to continue the translation process even in a condition in which the cap-dependent translation initiation is compromised due to environmental or pathological stresses.

The nuclear localization signal (NLS) is an indispensable signature sequence present in many proteins that are destined to be transported to the nucleus [24]. Though most of the nuclear proteins contain at least one NLS, some of the other proteins can also be transported to the nucleus by binding to other nuclear proteins [25] or by the Ran-GTPase cycle [26]. Usually, the NLS contains a stretch of sequence of multiple positively charged lysine and/or arginine residues. These positively charged residues of NLS are known to be important to mediate the binding of the proteins to importin-α/β [27]. The importin-α/β proteins are required to translocate the bound proteins to the nucleus [27]. The NLS can be a single stretch of sequence or it can be bipartite [25]. In the latter case, the stretches can be two nearby regions or they can be two segments residing in distant regions of the proteins. Nonetheless, the folding of the proteins allows the NLSs to work coherently.

In this study, we report that the 5′UTR region of HYPK mRNA can be subjected to IRES-dependent translation initiation. The IRES-mediated translation initiation of the HYPK mRNA results in the formation of an N-terminal truncated isoform1 of HYPK, i.e. HSPC136/HYPK-ΔN. IRES of the 5′UTR of HYPK mRNA acquires a tertiary structure. The IRES-dependent translation product of HYPK (HYPK-ΔN) lacks the N-terminal eight amino acids. Since there is an NLS, represented by a arginine repeat, in the N-terminal stretch of the full-length HYPK protein, HYPK-ΔN lacks the ability of being transported to the nucleus. On the other hand, the full-length isoform-1 of HYPK has the NLS and it can shuttle to the nucleus by binding to the importin-α protein. The full-length HYPK protein maintains the nuclear proteostasis by preventing aggregation of p53-R248Q inside the nucleus. The nuclear presence of HYPK also allows it to show its cell cycle regulatory activity. Taken together, these results find the novel mode of IRES-dependent translation initiation of HYPK mRNA, along with the determination of mechanism of the nuclear localization process of HYPK.

Results

The 5'UTR region of HYPK mRNA contains an IRES element

The human Huntingtin interacting protein K (HYPK) is expressed as two major isoforms. The longer isoform of HYPK, i.e. isoform-1 (NCBI accession number: NP_057484.3), is a 129 amino acid long protein, whereas the shorter isoform (isoform-2; NCBI accession number: NP_001186814.1) is an 81 amino acid long protein. The isoform-2 of HYPK is formed from a shorter HYPK mRNA that results due to an alternative splicing of the full-length mature HYPK mRNA. The open reading frame (ORF) of HYPK's isoform-1 is 390 nucleotides long in the mature mRNA (1351 nucleotides). Other than these two isoforms, expression of another form of HYPK's isoform-1 is reported. Though this protein is named as HSPC136 (NCBI accession number: AAF29100.1), it is actually an N-terminal truncated form of HYPK's isoform-1. This protein lacks the N-terminal eight amino acids (MRRRGEID) in comparison to the isoform-1 of HYPK. To simplify the naming process, we will denote the full-length HYPK isoform-1 as HYPK and HSPC136

as HYPK-ΔN in the rest of the manuscript (Figure 1a). The HYPK-ΔN protein translation does not start from the first start codon (AUG) which is located at the 179th-181st position of the HYPK mRNA (Figure 1b). Instead, translation of HYPK-ΔN starts from the second start codon which is located at the 203rd-205th position of the mRNA. This forms the basis of the question of how the first start codon is skipped and the second start codon is used to incorporate the first methionine during the translation of HYPK-ΔN. We hypothesized that the translation of HYPK-ΔN was not mediated by the conventional cap-dependent translation process. Instead, we believed that the translation of HYPK-ΔN occurred through an internal ribosome entry site (IRES)-dependent manner of translation initiation.

To investigate if HYPK-ΔN translation initiation was mediated by an IRES element, we first checked if the 5′UTR of HYPK mRNA (1-202nd nucleotides) had any internal ribosome entry site. Theoretical prediction by MFOLD algorithm [28] showed that the 5′UTR sequence of HYPK mRNA was capable of forming highly complex tertiary structure (Figure 1c). Formation of the tertiary structure of the 5′UTR was associated with a high free energy of formation (ΔG = −69.40 kcal/mol). It indicated that the predicted tertiary structure of the 5′UTR was highly stable. The three-dimensional projection of the 5′UTR structure by RosettaCommons had also shown the presence of different stem-loop structures in the 5′UTR of HYPK mRNA (Figure 1c). Complex stem-loop secondary and tertiary structures were previously reported to be parts of IRES elements of different viral RNAs [29] and eukaryotic mRNAs [30]. Hence, the predicted secondary/tertiary structure of the 5′UTR of HYPK mRNA stimulated us to find if the 5′UTR of HYPK mRNA had any IRES activity.

To validate the existence of the IRES element in the 5′UTR of HYPK mRNA, we generated several constructs on the backbone of the pEGFP-N1 and pcDNA3.1(+) vectors. In pEGFP-N1, mCherry was cloned upstream of the enhanced green fluorescent protein (EGFP)-ORF to generate a bicistronic construct. Since this construct had no specific region in the intercistronic space, it was named as mCherry-null-EGFP (Figure 2a). The stop codon at the end of mCherry-ORF assured that the EGFP would not be translated from the bicistronic mCherry-null-EGFP construct (Figure 2b). In the intercistronic region of the mCherry-null-EGFP construct, we inserted the 5′UTR and the coding sequence of first nine amino acids (MRRRGEIDM) of full-length HYPK. This construct was named as mCherry-HYPK

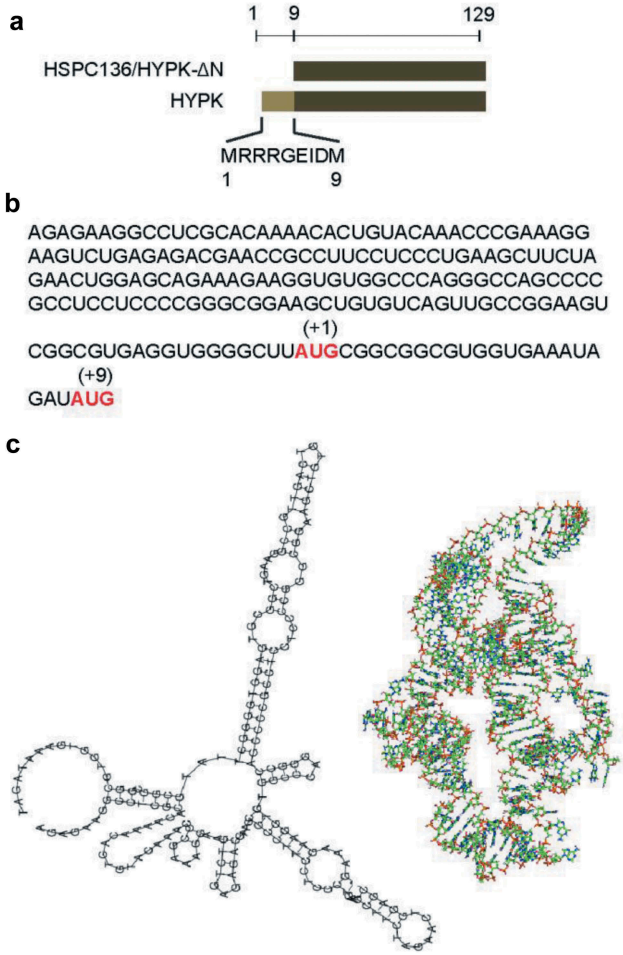

Figure 1. The 5′UTR of HYPK mRNA forms tertiary structure.

(a) Diagrammatic representation of HYPK and HSPC136/HYPK-ΔN. (b) The sequence of 5′UTR of HYPK mRNA. The +1 and +9 numbers denote the first and second AUG codon positions in the open reading frame of HYPK mRNA, i.e. the first and second AUG codons are positioned at the first and ninth codon positions, respectively. (c) MFOLD predicted secondary structure of HYPK 5′UTR and the three dimensional/tertiary projection of the secondary structure of the 5′UTR of HYPK mRNA. The first and second AUG codons start at 179th and 203rd nucleotides of the HYPK mRNA.

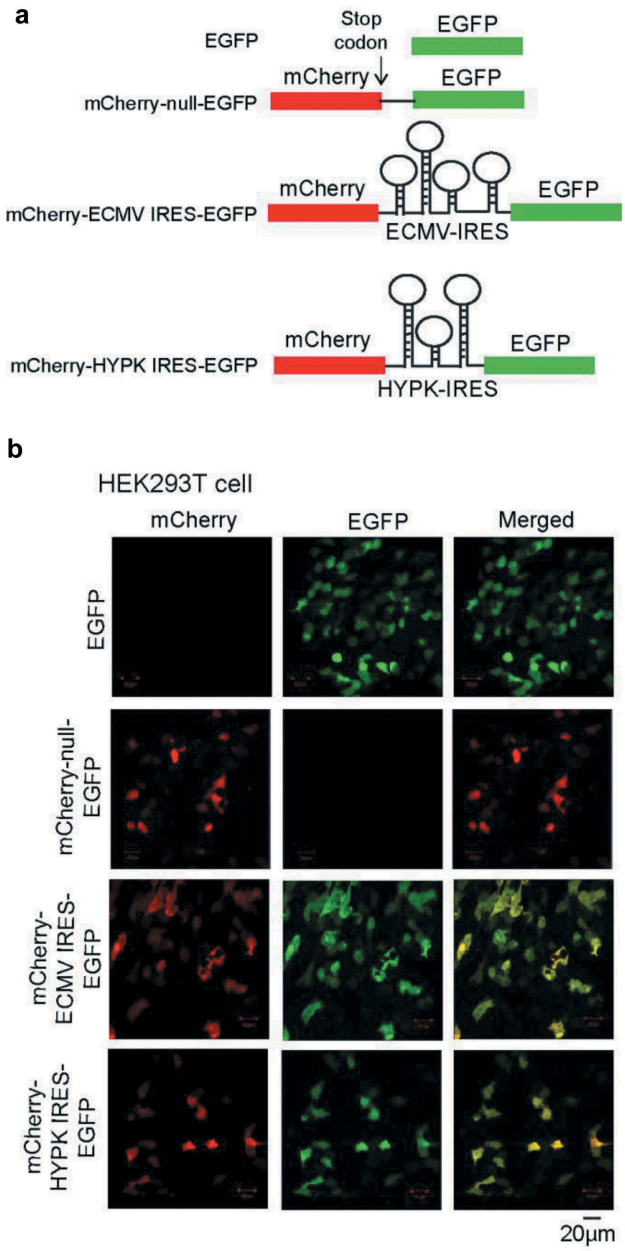

Figure 2. HYPK mRNA's 5′UTR contains an internal ribosome entry site.

(a) Schematic representation of the constructs that are used in the study. [For the details of the constructs, see results.] (b) Transfection of different constructs in the HEK293T cell shows the expression of the EGFP from the mCherry-HYPK IRES-EGFP construct which proves the fact that the 5′UTR of HYPK mRNA functions as an IRES element.

IRES-EGFP (Figure 2a). In the mCherry-HYPK IRES-EGFP construct, the putative HYPK IRES contained the first two AUG codons of HYPK-ORF [the AUG(+1) codon coded the first methionine and the AUG(+9) codon coded the ninth methionine in the full-length HYPK protein]. It was to be noted that the position numbering of the AUG codons was done based on the codon numbers in the ORF of the full-length HYPK protein. The numbering was not done as the nucleotide numbers in the mRNA of full-length HYPK. As a positive control, we also separately inserted the encephalomyocarditis virus (ECMV) IRES element in the intercistronic space of mCherry-null-EGFP. This construct was named as mCherry-ECMV IRES-EGFP (Figure 2a). In the mCherry-HYPK IRES-EGFP construct, the expression of mCherry was mediated by cap-dependent translation. The expression of EGFP was only possible if the inter-cistronic 5′UTR of HYPK acted as an IRES element that could drive the cap-independent translation initiation of EGFP. Indeed, it was observed that both mCherry and EGFP were expressed in the HEK293T cells that were transfected with the mCherry-HYPK IRES-EGFP construct (Figure 2b). The positive control mCherry-ECMV IRES-EGFP construct also showed similar kind of expression of both mCherry and EGFP in the HEK293T cells (Figure 2b). This result indicated that the 5′UTR of HYPK mRNA had actually acted as an IRES element that could lead to cap-independent translation initiation. In the mCherry-HYPK IRES-EGFP transfected cells, all the cells, which expressed mCherry, had also expressed EGFP. This observation led us to rule out the possibility of ribosomal read-through. Had

it been a situation of ribosomal read-through, we could expect that some of the transfected cells would express both the mCherry and EGFP due to the sporadic events of ribosome read-through. Since such was not the case and all mCherry-HYPK IRES-EGFP transfected cells showed the expression of both mCherry and EGFP, we were certain about the IRES property of the 5′UTR of HYPK mRNA.

Since the mCherry-HYPK IRES-EGFP construct had both the first two AUG codons of HYPK mRNA, it was necessary to identify which of the two start codons could lead to the initiation of translation of the downstream EGFP open reading frame in IRES-dependent manner. The possibilities were that the translation started at first AUG codon (at position +1) or at the second AUG codon (at position +9). We separately mutated the methionine coding AUG codons to proline coding CCU to generate the mCherry-HYPK IRES-EGFP AUG(+1)>CCU(+1) and mCherry-HYPK IRES-EGFP AUG (+9)>CCU(+9) constructs (Figure 3a). In the mCherry-HYPK IRES-EGFP AUG(+1)>CCU(+1), IRES-dependent translation could not start from the first AUG(+1) as it was mutated to CCU. On the contrary, IRES-dependent translation in mCherry-HYPK IRES-EGFP AUG(+9)>CCU(+9) could not start from the second AUG(+9) codon due to the mutation of this AUG codon to CCU. If the IRES-dependent translation of EGFP had to start in the mCherry-HYPK IRES-EGFP AUG (+1)>CCU(+1), it had to start from the second AUG(+9)

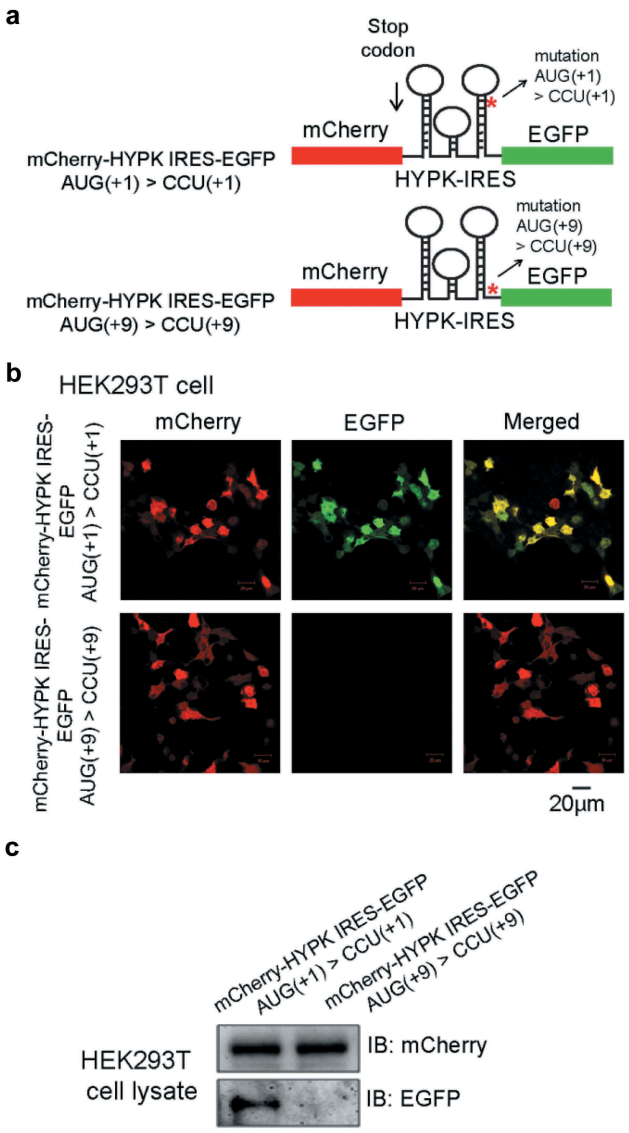

Figure 3. IRES-dependent translation of HYPK mRNA starts from the second AUG codon.

(a) Schematic representation of the HYPK IRES constructs that are mutated either in the first AUG codon [+1] or in the second AUG codon [+9]. The methionine coding AUG codons are mutated to proline coding CCU codons. In mCherry-HYPK IRES-EGFP AUG(+1) > CCU(+1) construct, the first AUG codon is mutated to CCU codon. In the mCherry-HYPK IRES-EGFP AUG(+9) > CCU(+9) construct, the second AUG codon is mutated to CCU codon. (b) In HEK293T cells, transfection of mCherry-HYPK IRES-EGFP AUG(+1) > CCU(+1), but not the mCherry-HYPK IRES-EGFP AUG(+9) > CCU(+9), shows expression of both mCherry and EGFP. This shows that the IRES-dependent translation of HYPK mRNA starts from the second AUG [+9] codon. (c) Immunoblots of HEK293T cell lysates of mCherry-HYPK IRES-EGFP AUG(+1) > CCU(+1) and mCherry-HYPK IRES-EGFP AUG(+9) > CCU(+9) transfected cells show that the EGFP is only expressed in mCherry-HYPK IRES-EGFP AUG(+1) > CCU(+1) transfected HEK293T cells. Expression of EGFP from mCherry-HYPK IRES-EGFP AUG(+1) > CCU(+1) shows that the second AUG [+9] is used in IRES-dependent translation initiation of HYPK mRNA.

codon. IRES-dependent translation initiation in mCherry-HYPK IRES-EGFP AUG(+9)>CCU(+9) could only happen from first AUG(+1) codon. We observed that the expression of EGFP occurred in the mCherry-HYPK IRES-EGFP AUG (+1)>CCU(+1) transfected cells, and not in the mCherry-HYPK IRES-EGFP AUG(+9)>CCU(+9) transfected cells (Figure 3(b,c). This implied that IRES-dependent translation of EGFP in mCherry-HYPK IRES-EGFP started from the second AUG(+9) codon. The first AUG(+1) codon of mCherry-HYPK IRES-EGFP was not used to initiate IRES-dependent translation. It was noted that the first AUG(+1) codon was located in the stem of a stem-loop structure of the HYPK-IRES. On the other hand, the second AUG(+9) codon was positioned in a linear region of HYPK-IRES. Since the AUG codon in a secondary structure of mRNA in often not conducive for initiation of IRES-dependent translation, the first AUG(+1) codon of HYPK-IRES was not involved in initiating the translation for subsequent codons. The second AUG(+9) codon was more conducive to initiate the IRES-dependent translation due to its better accessibility in the linear region of mRNA.

We conducted the quantitative estimation of the IRES activity of HYPK mRNA's 5′UTR by the Luciferase assays. We constructed *Renilla* Luciferase (Rluc)-Firefly Luciferase (Fluc) bicistronic construct on the backbone of the pcDNA3.1(+) plasmid. The generation scheme of the construct was similar to what was described earlier for the bicistronic mCherry-EGFP constructs. We generated three constructs, namely Rluc-null-Fluc (contained no specific sequence in the inter-cistronic region), Rluc-HYPK IRES-Fluc (contained the HYPK IRES in the intercistronic space of Rluc and Fluc) and Rluc-ECMV IRES-Fluc (contained the ECMV IRES in the intercistronic region) (Figure 4a). In the HEK293T and HeLa cells, the activities of Fluc of the Rluc-HYPK IRES-Fluc were seen to be 8.63 (s.d. = ±3.31) and 6.77 (s.d. = ±1.69) fold higher than the Fluc of the Rluc-null-Fluc construct (Figure 4b). High fold increase of the Fluc activity in the Rluc-HYPK IRES-Fluc transfected cells had quantitatively shown the higher translation initiation activity by the HYPK IRES. The Rluc-ECMV IRES-Fluc transfected cells had also shown high Fluc activity [10.27 (s.d. = ±1.58) and 8.27 (s.d. = ±2.65) fold higher in HEK293T and HeLa cell, respectively] compared to the Fluc activity of the Rluc-null-Fluc construct transfected cells (Figure 4b). These phenomena confirmed that the 5′UTR of HYPK mRNA had an intrinsic IRES activity that could help in starting the translation process from the second start codon at the 203rd-205th position of the HYPK mRNA.

It was essential to rule out the possibility that a monocistronic Fluc mRNA could generate from the Rluc-HYPK IRES-Fluc mRNA by alternative splicing. In order to know if such a splice-site variant Fluc mRNA was generated, we conducted reverse transcription-polymerase chain reaction (RT-PCR) of the RNA(s) that was/were generated from the Rluc-HYPK IRES-Fluc construct. We used different pairs of primers/oligos to ascertain the length/quality of the RNA(s) that could have derived from the Rluc-HYPK IRES-Fluc (Figure 5a). Use of these oligos in the RT-PCR analysis did not show the presence of smaller-sized splice-variant RNA(s) (Figure 5a). Moreover, the northern blot against the Fluc containing mRNA(s) (isolated from Rluc-HYPK IRES-Fluc transfected HEK293T cells) also showed the presence of only one bicistronic Rluc-HYPK IRES-Fluc mRNA, and no other smaller RNA (Figure 5b). These results proved the fact that no alternative splicing event had happened in the Rluc-HYPK IRES-Fluc mRNA, and the expression of Fluc in Rluc-HYPK IRES-Fluc transfected cells had occurred due to the IRES functioning of the intercistronic 5′UTR of HYPK mRNA. We also checked if there was any cryptic promoter in the Rluc-HYPK IRES-Fluc that could lead to the generation of monocistronic Fluc mRNA and the subsequent activity of the Fluc protein. To eliminate the possibility of the presence of any cryptic promoter, we created a construct in which the Fluc-HYPK IRES-Rluc segment was inserted in a bacterial vector, pET21b, which did not contain any eukaryotic promoter, but contained a T7 promoter. When transfected in HEK293T cells, this construct (pET21b-Rluc-HYPK IRES-Fluc) showed very low/negligible luciferase activity for both Rluc and Fluc (Figure 5c). This showed the absence of any cryptic promoter in the 5′UTR of HYPK mRNA that could initiate the transcription of a monocistronic Fluc mRNA. When the pET21b-Rluc-HYPK IRES-Fluc was co-transfected in HEK293T cells with another plasmid that contained the T7 RNA polymerase clone (pcDNA3.1-T7 RNA polymerase), both the Rluc and Fluc activities were very high (Figure 5c). These findings indicated that there was no cryptic promoter located in the 5′UTR of HYPK mRNA.

HYPK-ΔN lacks the nuclear localization capacity due to the absence of N-terminal nuclear localization signal

In the next part of our study, we tried to find if the HYPK-ΔN had differential cellular localization and functioning property than the full-length HYPK protein. Our previous studies had shown that the HYPK localization was predominantly cytoplasmic [1]. However, HYPK could also be present in the nucleus [1]. The nuclear localization of HYPK was also indirectly proven in another study which showed the HYPK interaction with several nuclear proteins [2]. However, the nuclear localization signal (NLS) of HYPK was not known. Hence, we investigated to understand if HYPK-ΔN was also present in nucleus, or it was distinctively present only in the cytoplasm. While the full-length HYPK protein was capable of showing the nuclear presence, the HYPK-ΔN protein could not translocate to the nucleus. This was evident in confocal microscopy images (Figure 6(a–c)) and in western blot studies (Figure 6d). The confocal microscopy images showed complete lack of HYPK-ΔN presence in the nucleus. On the other hand, HYPK was present in the nucleus of significantly high number of cells. Similarly, in the immunoblotting experiments, we did not find any trace of HYPK-ΔN in the nucleus.

Many of the cellular proteins are transported to the nucleus by a mechanism that involved the protein's interaction with importin-α/β proteins. We investigated if HYPK was transported to the nucleus by binding to the importin-α, and whether HYPK-ΔN lacked the capacity of binding to the importin-α. Indeed, we found that full-length HYPK protein could strongly bind to the importin-α (Figure 6e), whereas the HYPK-ΔN protein did not show binding to the importin-α (Figure 6e). This had confirmed that the truncated isoform of

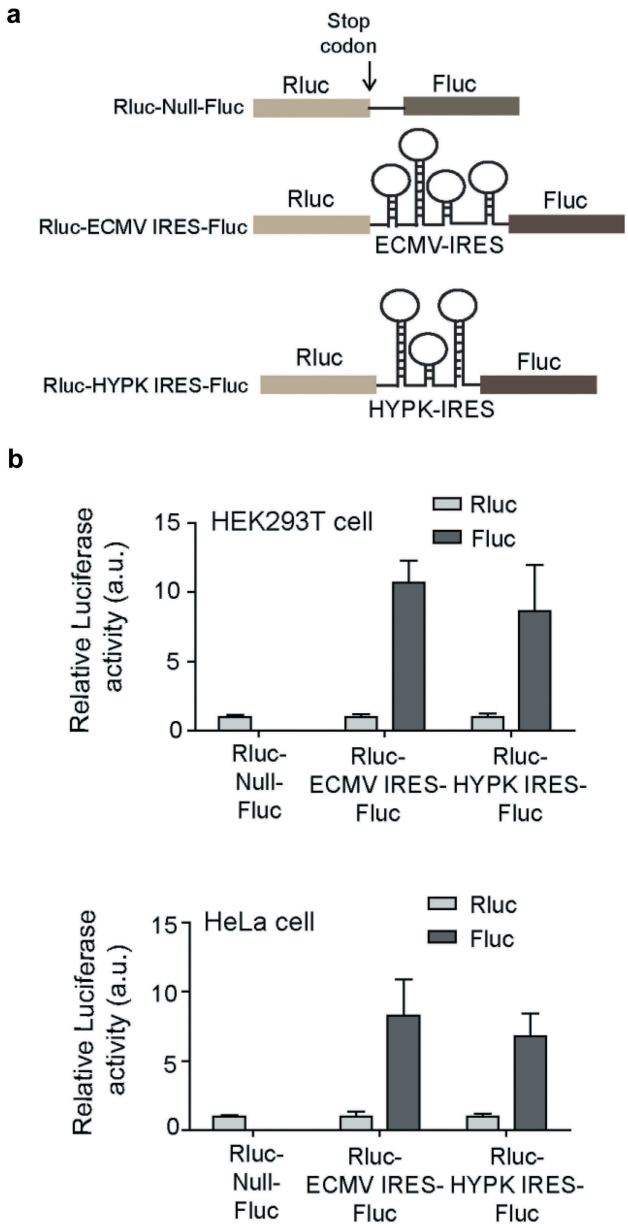

Figure 4. HYPK IRES initiates the translation of the downstream coding sequence.

(a) Schematic representation of different constructs. [See results for the details of the constructs.] (b) In HEK293T and HeLa cells, HYPK IRES can initiate the translation of the downstream Fluc open reading frame from the Rluc-HYPK IRES-Fluc bicistronic construct. Quantitative expression of Fluc is many fold increased in the Rluc-HYPK IRES-Fluc and the Fluc-ECMV IRES-Rluc transfected cells compared to the Rluc-null-Fluc transfected cells.

HYPK, i.e. HYPK-ΔN, was a cytosolic protein and it lacked the nuclear localization capacity. On the contrary, full-length HYPK was a nucleo-cytoplasmic protein, which could shuttle from cytosol to nucleus by binding to the importin-α protein.

In the next phase of our study, we analyzed the nuclear localization signal (NLS) of HYPK that triggered the localization of HYPK in the nucleus. Since the HYPK-ΔN could not translocate to the nucleus, it was evident that the NLS resided in the N-terminal eight amino acid region (MRRRGEID) of HYPK. Various previous reports had suggested that a patch of positively charged amino acids (lysine and/or arginine) could act as nuclear localization signal [25]. In the N-terminus of HYPK, we found a similar positive charge-rich patch that contained a repeat of three arginine residues (R2, R3 and R4) (Figure 7a). We suspected that this arginine repeat could act as an NLS in the full-length HYPK. We generated different alanine mutations of the arginine repeat (Figure 7a) to understand the NLS activity of 'RRR' sequence and pinpoint the effectiveness of each arginine in the cytoplasm-to-nuclear shuttling process. We found that the tri-arginine sequence was necessary and sufficient for transport of the full-length HYPK to the nucleus (Figure 7 (b,c)). HYPK's nuclear transport was prevented when all the three arginine amino acids were mutated to alanine (HYPK-RRR/AAA) (Figure 7b second row, 7c). We found that the central arginine (R3) was indispensable for nuclear transport of HYPK. Mutation of this residue to alanine (R3A in HYPK-RRR/RAR) had completely abolished the nuclear

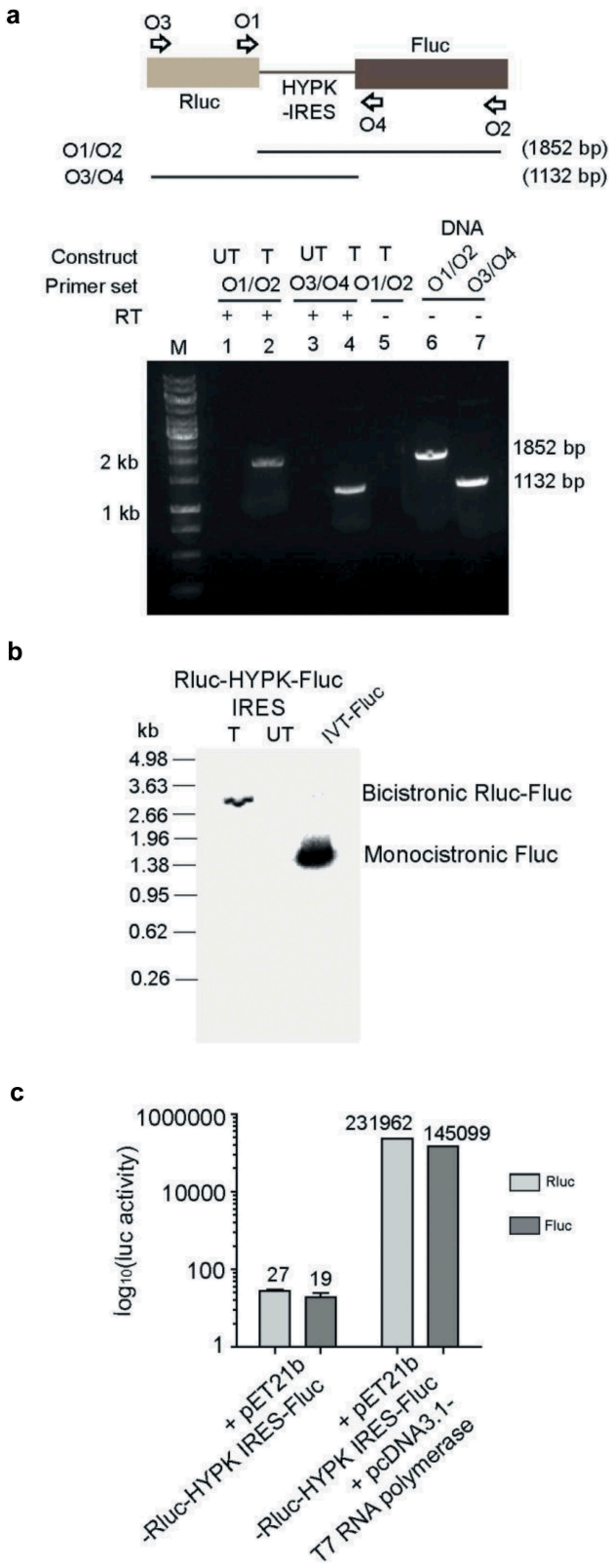

Figure 5. The Rluc-HYPK IRES-Fluc bicistronic construct does not contain any splice site or cryptic promoter.

(a) RT-PCR analysis of the RNA of Rluc-HYPK IRES-Fluc transfected HEK293T cells show no splice variant arising from the bicistronic Rluc-HYPK IRES-Fluc mRNA. The experiment is conducted with two sets of primers which are represented as, O1: oligo-1, O2: oligo-2, O3: oligo-3, O4: oligo-4. RT is reverse transcriptase, UT is RNA from untransfected cells, T is RNA from Rluc-HYPK IRES-Fluc transfected cells. The lane 5 is RT untreated/RT negative control reaction. Lanes 6 and 7 are positive control reactions that used Rluc-HYPK IRES-Fluc DNA instead of Rluc-HYPK IRES-Fluc mRNA. (b) Northern blot analysis for identification of the Fluc-containing mRNA from the total isolated RNA of the Rluc-HYPK IRES-Fluc transfected HEK293T cells. The nucleotide length of the identified Fluc-containing mRNA correlates to the nucleotide length of bicistronic Rluc-HYPK IRES-Fluc mRNA, representing no smaller splice variant mRNA arising from the Rluc-HYPK IRES-Fluc mRNA. (c) Rluc and Fluc activity in the HEK293T cells that were transfected with pET21b-Rluc-HYPK IRES-Fluc construct. The expression/activity of Fluc and Rluc is very low in the transfected cell due to the lack of production of bicistronic mRNA (lack of eukaryotic promoter in pET21b). It represents that no cryptic promoter is present in the Rluc-HYPK IRES-Fluc construct. Fluc and Rluc expression/activity rise to a very high level in the cells which are co-transfected with pET21b-Rluc-HYPK IRES-Fluc-pET21b and pcDNA3.1-T7 RNA polymerase.

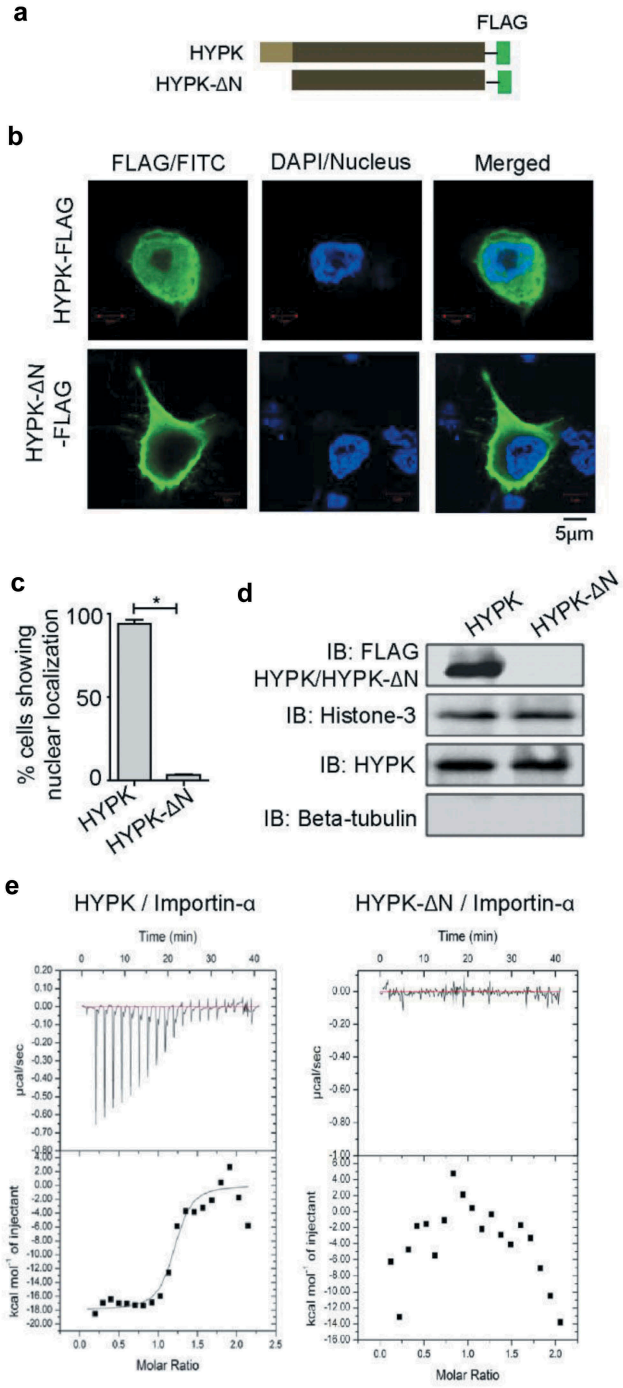

Figure 6. HYPK, but not the HYPK-ΔN, is transported to nucleus.

(a) Schematic presentation of HYPK and HYPK-ΔN. (b) Confocal microscopy images show that the full-length HYPK can be transported to the nucleus of HEK293T cells, whereas HYPK-ΔN can not translocate to the nucleus of HEK293T cells. (c) Nuclear localization of full-length HYPK is significantly higher than HYPK-ΔN in HEK293T cells [* nuclear localization of HYPK vs HYPK-ΔN: t-test, n = 106, df = 104, $p < 0.001$]. (d) Western blot of nuclear contents of HEK293T cells shows that HYPK, but not the HYPK-ΔN, is present in nucleus. The purity of nuclear fraction is indicated by the absence of cytosolic protein beta-tubulin. (e) Isothermal titration calorimetric assays for the protein–protein interaction studies show that HYPK can bind to importin-α. Dissociation constant [K_d] of HYPK-importin-α binding is 1.71μM. HYPK-ΔN can not bind to importin-α.

presence of HYPK (Figure 7b fourth row, 7c). The upstream and downstream arginine amino acids (R2 and R4) were required to strengthen the nuclear transport phenomenon. Individual mutations of these arginine residues (R2A in HYPK-RRR/ARR and R4A in HYPK-RRR/RRA) resulted in a meager nuclear presence of HYPK (Figure 7b third and fifth rows, 7c). The quantitative amounts of nuclear HYPK were less for these two arginine mutants (Figure 7c). These results identified the N-terminal arginine repeat as the nuclear localization signal in HYPK. The central arginine in this repeat had played the most significant role in the nuclear shuttling process.

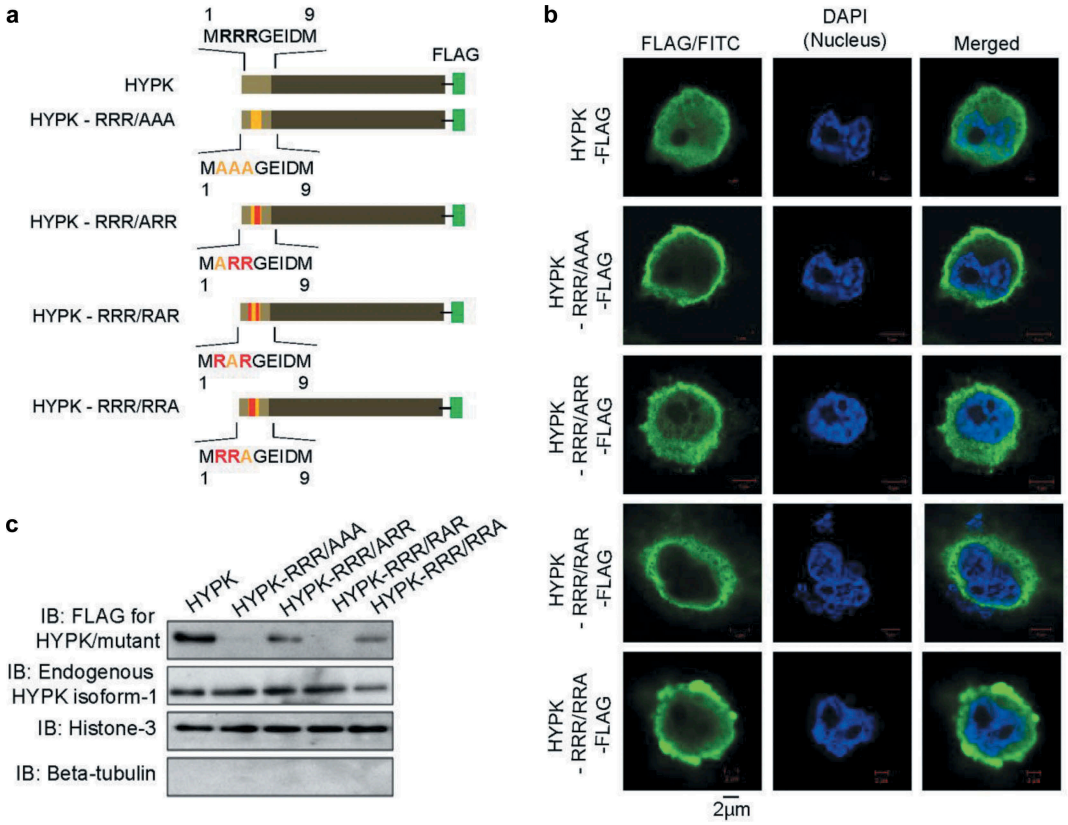

Figure 7. The N-terminal arginine repeat in HYPK functions as the nuclear localization signal.

(a) HYPK and its different arginine-to-alanine mutant constructs that are used to decipher the nuclear localization signal of HYPK. [See results for the details of the constructs.] (b) [First row] The N-terminal consecutive three arginine amino acids function as the nuclear localization signal of HYPK. The wild-type HYPK protein is capable of being transported to the nucleus. [Second row] Mutation of all the three N-terminal arginine residues to alanine [HYPK-RRR/AAA] renders HYPK unable to be localized in the nucleus. [Third row] Mutation of the first arginine-to-alanine of the N-terminal three arginines [HYPK-RRR/ARR] reduces the localization of HYPK in nucleus. [Fourth row] Mutation of the central arginine-to-alanine in the N-terminal arginine repeat [HYPK-RRR/RAR] causes complete loss of HYPK's nuclear localization capacity. [Fifth row] Mutation of the third arginine-to-alanine in the N-terminal arginine repeat [HYP-RRR/RRA] causes partial reduction of HYPK's localization to nucleus. (c) Immunoblotting of the nuclear fractions of the HEK293T cells that were transfected with HYPK or its different arginine-to-alanine mutants. The blot shows that the RRR repeat is necessary for HYPK's nuclear presence. The central arginine [R3] is necessary and sufficient to drive the translocation of HYPK to the nucleus. The two neighbouring arginine residues [R2 and R4] only enhance the nuclear localization ability of HYPK. Western blot is carried out with pure nuclear fraction, and without any contamination of cytosolic content. This is represented by the absence of beta-tubulin.

HYPK is ubiquitously expressed in different organisms. We asked the question if all the HYPK's in different organisms had retained the NLS in the evolution process. A comprehensive alignment and phylogenetic analysis of the HYPK sequences of 30 different organisms (heterogeneous mixture of lower and higher eukaryotic plants and animals) had shown that the presence of the NLS was not evolutionarily conserved in the HYPK protein of different organisms (Figure 8(a,b)). The NLS was only present in the HYPK protein of higher eukaryotic animals, like human, mice, etc. Most of the other organisms did not contain an NLS in their HYPK protein. This showed that the NLS was not phylogenetically conserved in different organisms. If we could extrapolate this observation, it would mean that HYPK function was probably not indispensable in the nucleus, and its nuclear presence in higher eukaryotic animals could only serve for advantageous functions.

HYPK, but not the HYPK-ΔN, can function as an aggregation preventing protein in the nucleus

Having found the exclusive presence of HYPK, but not the HYPK-ΔN, in the nucleus, we were interested to see what function(s) HYPK could exert in the nucleus. We had previously reported that HYPK could prevent/reduce the aggregation of different toxic aggregation-prone proteins [1,7]. HYPK formed annular-shaped macromolecular structures, termed as H-granules, that sequestered the aggregates of poly-glutamine expanded Huntingtin-exon1 (Htt97Q-exon1), α-synuclein-A53T and superoxide dismutase 1-G93A [1]. Another study had also pointed out the chaperone-like and aggregation preventive properties of HYPK [4]. We argued that HYPK could perform similar aggregation preventing function in the nucleus. We chose a mutant of p53 as the model protein that could form nuclear aggregates. An arginine-to-glutamine mutation at the 248th position (R248Q) of p53 protein was known to impart prion-like aggregation properties in p53 [31,32]. We tested the abilities of HYPK and HYPK-ΔN in terms of preventing the nuclear aggregates of p53-R248Q. The nuclear presence of full-length HYPK prevented the aggregation of p53-R248Q (Figure 9(a,b)). In the condition of the ectopically induced over-expression of HYPK, p53-R248Q showed diffused nuclear distribution. On the contrary, p53-R248Q formed significantly high number of nuclear aggregates during the ectopic expression of HYPK-ΔN (Figure 9

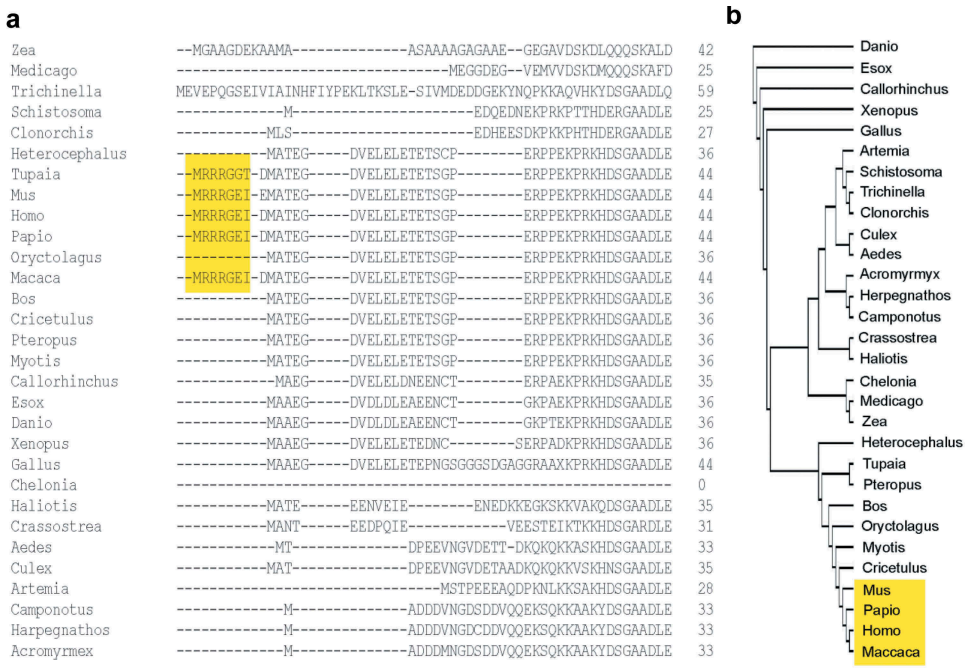

Figure 8. HYPK's NLS is not conserved in different organisms.

(a) Alignment of HYPK protein sequences of different organisms. (b) Phylogenetic analysis derived cladogram of the HYPK protein sequences of different organisms. The cladogram shows that NLS is only present in higher eukaryotic animals [yellow highlighted region].

(a,b)). Since HYPK-ΔN could not penetrate to the nucleus, it showed no effect in preventing the aggregation of p53-R248Q protein. HYPK reduced the aggregation of p53-R248Q by effectively co-associating with the p53-R248Q. In the confocal microscopy images, HYPK showed high colocalization with the p53-R248Q protein (Figure 9(a,c)). HYPK-ΔN showed significantly less colocalization with p53-R248Q (Figure 9(a,c)). The direct interaction of HYPK, but not the HYPK-ΔN, with p53-R248Q was also observed in the immunoprecipitation/immunoblotting (IP/IB) experiments (Figure 9d). A similar observation was also reported in a previous study in which it was shown that HYPK could physically interact with the p53 protein [2].

We had previously shown that not only HYPK could prevent the aggregation of aggregation-prone proteins, but it could also enhance the degradation of the aggregates/aggregation-prone proteins in the cells [1]. In the present study with p53-R248Q, we noticed a similar phenomenon. Other than reducing the nuclear aggregates of p53-R248Q (Figure 9e), HYPK could actually reduce the quantitative amount of cellular p53-R248Q (Figure 9e). Together, these findings indicated that the full-length HYPK had nuclear functions inside the cell. NLS containing full-length HYPK could cope up the challenges of nuclear protein aggregates. The N-terminal truncated isoform of HYPK, HYPK-ΔN, probably functions as a cytoplasmic protein that prevented the aggregation of other cellular proteins. Owing to its inability to be transported to the nucleus, HYPK-ΔN had no function in the nucleus.

HYPK-ΔN lacks the cell cycle regulatory activity

HYPK has cell cycle regulatory activity [2]. Regulation of cell cycle by HYPK had been attributed to the fact that HYPK was able to interact to a diverse range of nuclear and cytoplasmic proteins. We checked the effects of HYPK and HYPK-ΔN on the restoration of cell cycle during the p53-R248Q mediated proteotoxic stress. The expression of p53-R248Q had detrimental effects upon the cell cycle and cell survival [33]. We evaluated the effects of HYPK and HYPK-ΔN in restoring the normalcy of cell cycle and reducing the cell death that were associated with the toxic aggregates of p53-R248Q. While the expression of p53-R248Q had disturbed the normal cell cycle (Figure 10a upper panel), the ectopic co-expression of HYPK with p53-R248Q had substantially reduced the toxic effects of p53-R248Q. This was manifested by the reversion of the cell cycle to normal state and reduction of the cell deaths (Figure 10a middle panel). Though p53 had been known to induce apoptotic cell death [34], its apoptotic activity was thwarted by the activity of HYPK. This was in-line with the previous study that explained how HYPK could help in cell survival and proliferation [2]. The HYPK-ΔN failed to show any positive effect in restoring the p53-R248Q mediated cell cycle abnormalities and cell death. Deregulated cell cycle and sufficiently high number of cell deaths by p53-R248Q was observed even in the presence of HYPK-ΔN (Figure 10a lower panel). This had proven that HYPK, but not the HYPK-ΔN, had the ability to positively regulate the cell cycle and survival of human cells.

We presume that the cell cycle regulatory activity of HYPK was intrinsically associated with its ability of nuclear translocation and interaction with the nuclear proteins, like p53, Cdk(s) etc. Among all the HYPK interacting proteins, we found that a sufficiently high number of proteins were nuclear

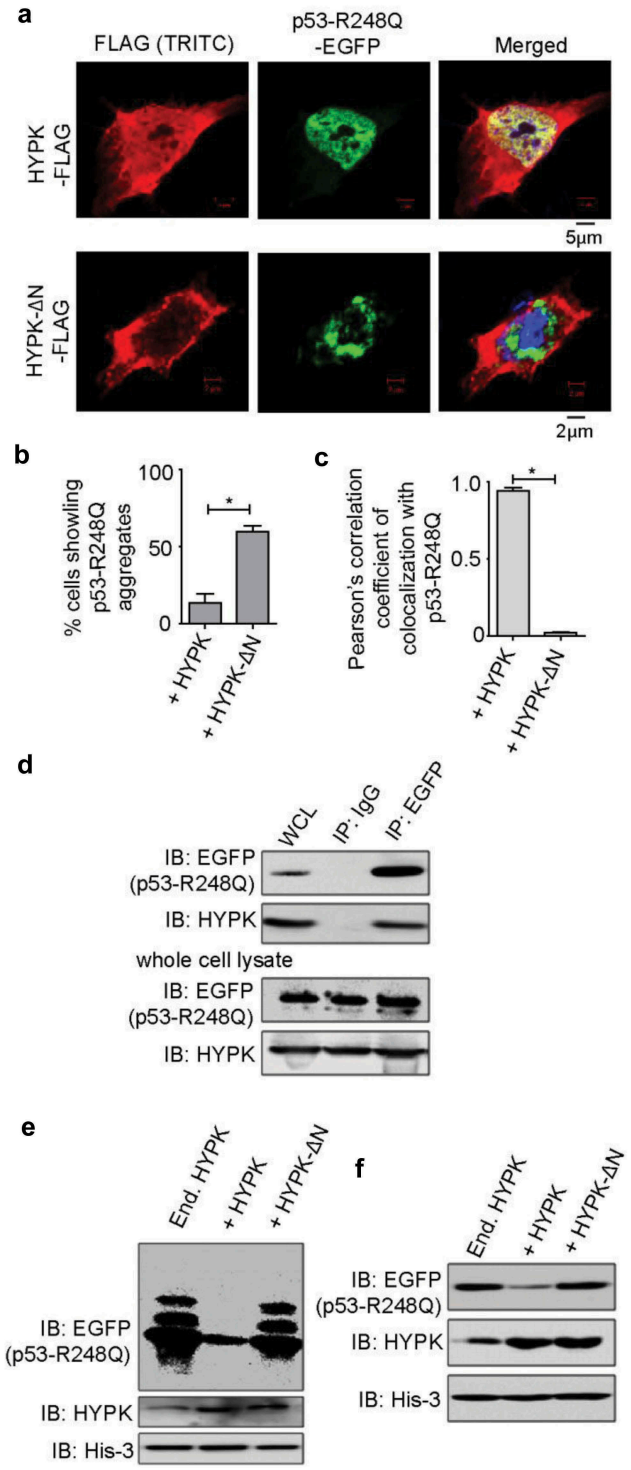

Figure 9. HYPK, but not the HYPK-ΔN, has the proteostastic effects in nucleus.

(a) Presence of HYPK in nucleus reduces the aggregation of p53-R248Q protein in the nucleus of HEK293T cells. HYPK-ΔN is unable of being transported to the nucleus. HYPK-ΔN can not reduce the nuclear aggregation of p53-R248Q. (b) HYPK mediated reduction of p53-R248Q's nuclear aggregation is significantly higher than HYPK-ΔN's activity towards the reduction of nuclear p53-R248Q aggregation. [Nuclear p53-R248Q aggregation in presence of HYPK vs. HYPK-ΔN: t-test, n = 85, df = 83, $p < 0.005$]. (c) HYPK shows significantly higher nuclear colocalization with p53-R248Q than the nuclear colocalization of HYPK-ΔN with p53-R248Q. [Pearson's correlation coefficient of p53-R248Q colocalization with HYPK vs. HYPK-ΔN: t-test, n = 15, df = 13, $p < 0.005$]. (d) Immunoprecipitation and immunoblotting experiments show that HYPK is a true interacting partner of p53-R248Q. (e) Immunoblots show that the over-expression of HYPK prevents nuclear aggregation of p53-R248Q in HEK293T cells. HYPK-ΔN has no such effect. (F) HYPK reduces the quantitative content of cellular p53-R248Q protein. HYPK-ΔN does not show any effect in reducing the total protein content of p53-R248Q.

proteins (Figure 10b). Many of them were directly involved in cell cycle progression, cell cycle inhibition and apoptosis. It was, therefore, reasonable to assume that HYPK mediated regulation of cell cycle/cell survival was dependent upon its ability to interact with other nuclear cell cycle regulatory proteins.

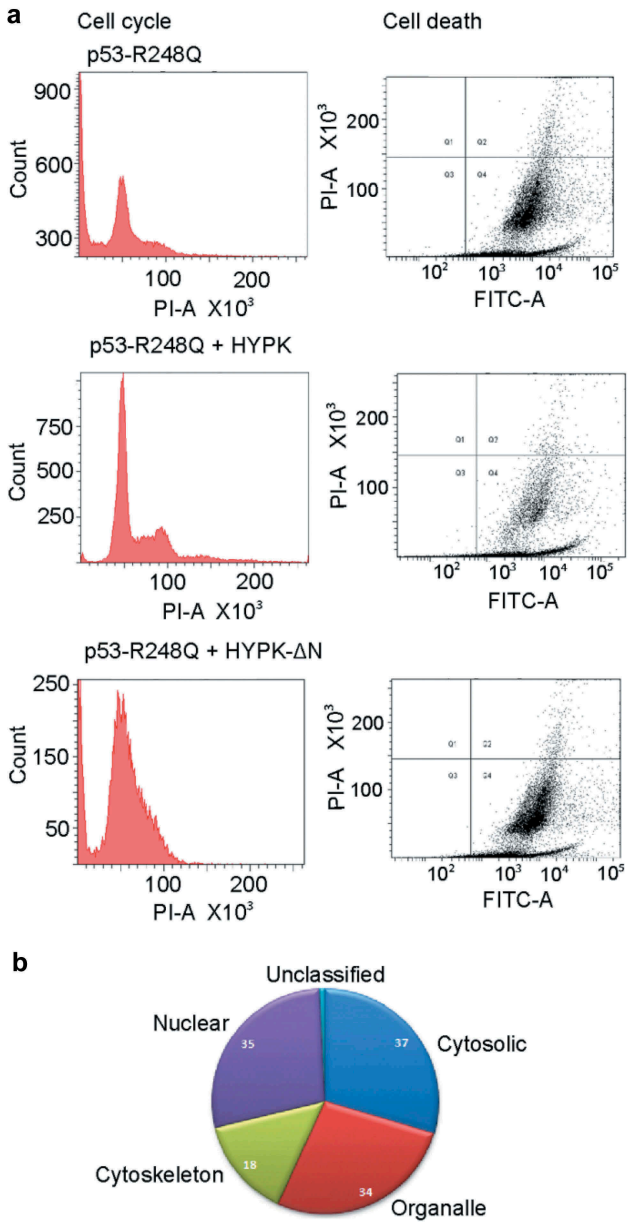

Figure 10. HYPK can restore the perturbed cell cycle. HYPK can also reduce the cell deaths that are caused by the proteotoxic effects of p53-R248Q.
(a) Over-expression of p53-R248Q disturbs the cell cycle in HEK293T cells. It also increases the death of HEK293T cells. Ectopic co-expression of HYPK with p53-R248Q leads to restoration of the normal pattern of cell cycle along with a reduction of cell death. HYPK-ΔN had no effect in modulating the disturbed cell cycle and increased cell deaths that are caused by the proteotoxicity of p53-R248Q. (b) Gene ontology analysis of cellular compartments of HYPK interacting proteins shows that a high number of nuclear proteins can interact with HYPK. Many of these proteins are cell-cycle regulatory proteins.

Discussion

Internal ribosome entry site mediated translation initiation is an alternative mechanism to start the translation process from some of the (m)RNA(s) in the physiological or stressed conditions [35]. While several of the viral RNAs employ IRES-dependent translation as the primary mechanism of translation initiation, some of the eukaryotic mRNAs also exploit this strategy to start the synthesis of polypeptides. For example, growth factor mRNAs, like VEGF mRNA, FGF mRNA, can undergo IRES-dependent translation initiation during cellular stresses [36,37]. Previous reports suggest that an N-terminus truncated isoform of HYPK, i.e. HSPC136 (or HYPK-ΔN), is expressed in human tissues. Our study finds that the translation of the HYPK-ΔN isoform happens by an IRES-mediated translation initiation of HYPK mRNA. The 5′UTR of the HYPK mRNA consists of an IRES element that is predicted to assume a three-dimensional tertiary structure. Such a structure in the 5′UTR of the HYPK mRNA is conducive for the cap-independent and IRES-dependent translation of the HYPK mRNA, resulting in the start of the translation from an alternative (second) initiation codon of HYPK mRNA.

While we have deciphered the IRES-dependent translation initiation of HYPK mRNA from the second AUG codon, we

do not discard the possibility of other mechanisms of translation initiation from the second AUG codon of HYPK mRNA. In fact, the choice of the second AUG codon may be determined by its better nucleotide context (purine bases at positions −3 and +4) compared to the less favourable context of the upstream initiating triplet, wherein the same positions are pyrimidine bases. Due to the presence of pyrimidine bases at the −3 and +4 positions of the first AUG codon of HYPK mRNA, it is possible that the scanning ribosome bypasses this AUG by leaky scanning. The ribosome may start the translation from the second AUG codon of the HYPK mRNA due to the presence of suitable purine bases at the −3 and +4 positions. In addition, the choice of second AUG codon and leaky scanning can be strongly influenced by the activities of some initiating factors, in particular eIF1 and eIF5. These activities can vary as the cell growth conditions changes. In the context of cellular perspectives, ribosomal shunting is another mechanism that can account for the differential translation of the HYPK mRNA from the second AUG codon. It is possible that the tertiary structure of the 5′UTR region of HYPK mRNA disrupts the ribosome scanning in such a way that the second AUG codon is used preferentially over the first AUG codon for translation initiation under certain circumstances. The ribosome dissociates from the HYPK mRNA upstream of the tertiary structure of 5′UTR, followed by landing of the ribosome at a different site after the tertiary structure of HYPK mRNA to continue the scanning for the subsequent AUG codon. Since the predicted structure of the 5′UTR of HYPK mRNA shows that the first AUG codon, but not the second AUG codon, resides in a highly structured region, it is reasonable to test if the ribosomal shunting mechanism can manifest the translation initiation from the second AUG codon of HYPK mRNA. Some cellular mRNAs undergo translation reinitiation after the translation of an upstream ORF [38]. Such a translation reinitiation from the second AUG codon of HYPK mRNA is theoretically possible after the translation is completed from the first AUG codon of HYPK mRNA. Since no previous study has identified any of the above mechanisms as an alternative mechanism of translation initiation from the second AUG codon of HYPK mRNA, it remains to be validated if such mechanisms really occur as non-canonical mode of translation initiation of HYPK mRNA.

HYPK is a member protein of the proteostasis network. It prevents the aggregation of different toxic aggregation-prone proteins [1]. It is well reported that cap-dependent translation process level attenuates and IRES-dependent translation of several mRNAs increase during the proteotoxic stress [39]. Since HYPK's function is critically required during the burden of protein aggregates, it is indispensable that the translation of HYPK mRNA continues in the condition of proteotoxic stress. When the global rate of cap-dependent translation is reduced, the IRES-mediated translation initiation of HYPK mRNA can help in continuous synthesis of this protein. IRES-dependent translation of HYPK mRNA maintains the necessary quantity of cellular HYPK during the imbalanced proteostasis. This can also be true for HYPK expression in thermal stress. Previous studies have reported that the transcriptional expression of HYPK is upregulated during the thermal stress, leading to cellular accumulation of HYPK protein [11,12]. Since normal/cap-dependent translation rate is suppressed during thermal stress, we believe that the continuous initiation of the HYPK mRNA translation occurs through IRES-dependent process during heat shock. Given the fact that heat shock responsive IRES-mediated translation is also reported for other chaperone proteins [40], there is the possibility of IRES-dependent translation of HYPK during heat stress. However, this notion is completely speculative and it requires further investigations for definitive conclusions.

The IRES-dependent translation initiation of HYPK mRNA is associated with the generation of the translation product (HYPK-ΔN) that has limited abilities of being translocated to the nucleus. The full-length HYPK protein has an N-terminal nuclear localization signal. The NLS is a small repeat of arginine amino acids. Since the IRES translated HYPK-ΔN lacks the first nine amino acids of the N-terminal region and the NLS, it is incapable of being transported to the nucleus. Thus, it can not manifest its activity in the nucleus. Unlike this truncated isoform, the full-length HYPK shows activity in preventing the self-association of aggregation-prone nuclear proteins, like mutant p53 (p53-R248Q). Similar to many cellular proteins that are transported to nucleus due to the presence of positively charged patch(es) in their NLS, the full-length HYPK also shuttles to the nucleus due to the activity of the positively charged arginine repeat in its NLS. Nuclear transport of HYPK is mediated by the binding of HYPK to the importin-α protein. We have previously shown that the N-terminus of HYPK is a disordered nanostructure [6]. The existence of the NLS in this disordered region helps HYPK to easily bind to importin-α. The central arginine of the tri-arginine appears to be essential for HYPK transport to nucleus. Mutation of this arginine residue to alanine renders the HYPK unable of being transported to nucleus. The neighbouring two arginine amino acids enhance the efficiency of nuclear transport of HYPK. Individual mutation of these two arginine residues to alanine does not compromise the nuclear transport of HYPK, albeit HYPK's transport to nucleus becomes quantitatively less. The NLS of HYPK is not very strong. HYPK is mostly a cytosolic protein, with a very less presence in nucleus. Since the positive charge-rich patch in the NLS of HYPK is very short, the nuclear localization of HYPK is a not heavily favoured. Though the nuclear import mechanism of HYPK is evident from our study, we are unable to comment if HYPK also shuttles out from nucleus to cytoplasm by the activity of any existing nuclear export signal in it.

Not only the nuclear localization of the full-length HYPK allows this protein to balance the nuclear proteostasis, but it also helps in maintaining the cell cycle. Though HYPK is known to be a cell cycle regulatory protein [2], its mode of action in cell cycle regulation is not known. Because the HYPK-ΔN potentially lacks any activity in restoring the abnormal cell cycle, we believe that the nuclear localization of HYPK is essential for its cell cycle regulatory activity. This is also apparent from the fact that HYPK interacts with several cell cycle regulatory proteins, like p53, etc., in the nucleus. We understand that the cell cycle regulation by HYPK is an indirect effect that is caused by the differential spatio-

temporal interaction of HYPK with nuclear cell cycle regulatory proteins.

The NLS of HYPK is not evolutionary conserved in different organisms. In fact, the NLS is only present in the HYPK protein of higher eukaryotic animals. This implies that the primary localization and function of HYPK is cytosolic, and not nuclear. The addition of the NLS in the N-terminus of HYPK of higher eukaryotic animals has provided an additional functional advantage to HYPK in terms of monitoring the nuclear proteostasis. It is for this reason that the NLS is positively selected and retained in the N-terminus of HYPK of some organisms.

It is important to understand that the primary function of HYPK deals with proteostasis, which is mainly a cytosolic event. Hence, the NLS can be regarded as a less essential part of HYPK sequence. In the stressful conditions, maintaining the cellular homeostasis is more essential than the progression of the cell cycle. So, HYPK-ΔN, which is the native HYPK protein in many of the organisms, appears to be functionally relevant irrespective of its inability to translocate to nucleus and control the cell cycle. In human and other mammals, the full-length HYPK has the NLS and it has the extra ability to control some of the nuclear events as well. In the physiological condition, the full-length HYPK is involved in both proteostasis and cell cycle regulation in cytoplasm and nucleus, whereas stress-dependent IRES-mediated translation of HYPK-ΔN can only be involved in the cytosolic proteostasis.

Methods

Cloning and constructs

mCherry was cloned in the multiple cloning site (MCS) of pEGFPN1 vector to generate the mCherry-null-EGFP construct. The *Renilla* Luciferase (Rluc) and firefly Luciferase (Fluc) open reading frame were PCR-amplified from the pCMV-IRES-Renilla Luciferase-IRES-Gateway-Firefly Luciferase (pIRIGF) plasmid which was obtained from Addgene (plasmid number: 101139). Rluc- and Fluc-specific primers were used to PCR-amplify the Rluc and Fluc DNA. The Rluc and Fluc were consecutively cloned in the multiple cloning site (MCS) of pCDNA3.1(+) to generate the Rluc-null -Fluc construct. To obtain the DNA sequence corresponding to the 5′UTR DNA of HYPK mRNA, total RNA from HeLa cell was isolated, followed by RT-PCR of the HYPK mRNA using the 5′UTR specific primers. The encephalomyocarditis virus (ECMV) IRES element was PCR-amplified from the pIRES plasmid using ECMV-IRES specific primers. The 5′UTR/IRES of HYPK and ECMV IRES were separately inserted in the intercistronic region (part of the MCS) of mCherry-null-EGFP and Rluc-null-Fluc to generate the following constructs: mCherry-HYPK IRES-EGFP, mCherry-ECMV IRES-EGFP, Rluc-HYPK IRES-Fluc and Rluc-ECMV IRES-Fluc. The Rluc-HYPK IRES-Fluc segment was cut from the pcDNA3.1 and re-inserted in the MCS of pET21b vector to generate the pET21b-Rluc-HYPK IRES-Fluc construct. T7 polymerase gene was PCR-amplified from the genome of BL21DE strain of *Escherichia coli*. The T7 polymerase gene was cloned in the pcDNA3.1(+) vector. The DNA of the full-length coding sequence of HYPK and the N-terminal deletion constructs of HYPK (HYPK-ΔN), importin-α were generated by RT-PCR of the HYPK mRNA and importin-α mRNA. The HYPK, HYPK-ΔN and importin-α were separately cloned in pET21b vector. The HYPK and HYPK-ΔN were also cloned in pcDNA3.1(+) with the in-frame FLAG peptide sequence at the C-terminus. The FLAG sequence was incorporated in the reverse primer while cloning the HYPK and HYPK-ΔN in the pcDNA3.1(+). The site-directed mutagenesis of the arginine residues in the N-terminus of the full-length HYPK were done by changing the arginine codon(s) to alanine codon(s) in the forward primer(s) that were used for the PCR of HYPK. The following arginine(R)-to-alanine(A) mutant constructs were generated for HYPK: HYPK-AAA, HYPK-ARR, HYPK-RAR, HYPK-RRA (see results for the details of the constructs). All of these HYPK arginine-to-alanine mutants had the C-terminal in-frame FLAG sequence attached to them. The p53-EGFP clone was kindly gifted by Dr. Murali Daran Bashyam of the laboratory of molecular oncology of CDFD. The R248Q mutant of p53 was generated by overlapping PCR-based method using four primers. All the cloning processes involved the same method that was described by us in our earlier studies [1,41]. Briefly, the DNA constructs were PCR amplified, followed by restriction digestion of the PCR products and plasmid vectors. Restriction digested PCR products were ligated with the cognate vectors and the ligation products were transformed into the ultra-competent cells of DH5α strain of *Escherichia coli*. Viable colonies were selected in the antibiotic-resistant media and the positive clones were identified by colony PCR. All clones were sequenced at the sophisticated equipment facility of the research support service group of CDFD. Details of all constructs, clones and primer sequences are available upon request.

Recombinant protein production

Recombinant HYPK, HYPK-ΔN and importin-α proteins were produced in a process that was similar to what was described in our earlier studies [42,43]. Briefly, the bacterial expression clones of HYPK, HYPK-ΔN and importin-α were separately transformed into the BL21DE3 strain of *Escherichia coli*, followed by initiating of the protein production by addition of 1 mM (final concentration) of IPTG in the LB-ampicillin medium. After sufficient time (12–16 h) of protein production, bacterial cells were lysed in lysis buffer [50 mM Tris-Cl (pH: 8.0), 300 mM NaCl, 10 mM imidazole and 1 mM PMSF]. Cleared cell lysate was run through the Ni-NTA beads containing column to allow the binding of 6X-Histidine containing recombinant proteins to the nickel beads. This was followed by washing of the beads with wash buffer [50 mM Tris-Cl (pH: 8.0), 300 mM NaCl, 40 mM imidazole] and eluting the proteins in elution buffer [50 mM Tris-Cl (pH: 8.0), 300 mM NaCl, 300 mM imidazole]. Proteins were dialyzed in 20 mM Tris-Cl (pH: 8.0), 20 mM NaCl dialysis buffer.

Isothermal titration calorimetry

Isothermal titration calorimetry studies were done to analyze the qualitative and quantitative binding properties of HYPK and HYPK-ΔN to importin-α. The cell proteins (HYPK or

HYPK-ΔN) concentrations were 20 μM, whereas the syringe protein (importin-α) concentration was 200μM. All binding experiments were conducted at 25°C in the MicroCal iTC200 isothermal titration calorimeter. Temperature control was done by a thermostat which was attached to the instrument. The instrument was subjected to the following parameters: total number of binding events: 20, reference power of the instrument was 10 μcal/s, first binding event started after 60 s of experiment initiation, 120 s of regular interval was maintained between two binding events, the rotation speed of syringe was maintained at 300rpm. Most of the instrument parameters were similar to one of our earlier studies [43].

Cell culture

The HeLa and HEK293T cells were obtained from the National Centre for Cell Sciences (NCCS, India). Cells were maintained in DMEM medium that was supplemented with 10% fetal bovine serum, 2 mM l-glutamine and 1X penicillin-streptomycin solution. Cells were grown in a moistened incubator at 37°C and 5% CO_2 level. Transfection of DNA was done by Lipofectamine2000 (Invitrogen).

Luciferase assay

The Luciferase assays were conducted by following the protocol that was given for the use of Promega Luciferase assay system (E1483). After the cultured cells were detached from the culture dishes by the application of trypsin-EDTA, they were washed with PBS buffer, followed by lysis of the cells by CCLR lysis buffer. Following the lysate clearance by centrifugation, Luciferase reagent was added to the defined volume of cell lysate supernatant. The emitted luminescent light was detected in a luminometer for 10 s.

In vitro transcription

For the *in vitro* transcription of the Fluc gene, a T7 promoter was added upstream of the start codon of Fluc gene. The T7 promoter sequence was kept in the forward primer, and the PCR using this primer had added the T7 promoter sequence upstream of the start codon of the Fluc gene. The linear PCR product was *in vitro* transcribed using the protocol and reagents provided in the MEGAscript T7 transcription kit of Promega.

Reverse transcription-polymerase chain reaction

Using the Ambion RNA miniprep kit, the total RNA was isolated from the HEK293T cells that were untransfected or transfected with Rluc-HYPK IRES-Fluc (in pcDNA3.1+) construct. The RNA pool was treated with DNase-I to remove any contaminating DNA, followed by RT-PCR of the desired RNA (using Superscript III reverse transcriptase of Thermo Fischer Scientific) with defined combinations (see results and Figure 4) of the following primers: forward primer O1 – corresponds to the 3′region of Rluc gene, reverse primer O2 – corresponds to the 3′region of Fluc gene, forward primer O3 – corresponds to the 5′region of Rluc gene, reverse primer O4 – corresponds to the 5′region of Fluc gene.

Northern blotting

The total RNA was isolated from the HEK293T cells that were transfected with Rluc-HYPK IRES-Fluc (in pcDNA3.1+) construct. This RNA pool and the *in vitro* transcribed Fluc mRNA were separated on a 1% agarose gel, followed by blotting the RNAs on a nylon membrane. To identify the Fluc containing mRNA, the RNA containing nylon membrane was flooded with a hot (^{32}P) riboprobe that hybridized with the Fluc mRNA.

Nucleus isolation

The HEK293T cells were detached from the culture dishes by mild treatment of trypsin-EDTA, followed by washing of the cells with PBS buffer. Cells were lysed with 0.1% NP40-PBS lysis buffer by repeated trituration in ice-cold condition. The cell lysate was centrifuged at 600 g for 10 s, followed by re-suspending the pellet in 0.1% NP40-PBS buffer. The suspension was again centrifuged at 600 g for 30 s, and the pellet was collected. This pellet contained the nuclei.

Immunoblotting and immunoprecipitation

Whole cells or the nuclei fraction of the cells were lysed in RIPA buffer [20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM EGTA 1% NP-40, 1% sodium deoxycholate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, 1 μg/mL leupeptin], followed by clearance of the lysate by centrifugation. 40-60 μg of protein in cleared lysate was loaded onto 12% SDS-PAGE gel for electrophoretic separation. Separated proteins were transferred to PVDF membrane, followed by sequential application of primary and secondary antibodies with intermittent washing with TBST buffer. Following primary antibodies were used – anti HYPK antibody (Sigma Aldrich, HPA055252; dilution – 1:2000), anti-FLAG antibody (Sigma Aldrich, F3165; dilution – 1:5000), anti-histone H3 antibody (Sigma Aldrich, 06–755; dilution: 1:2000), anti-beta tubulin antibody (Sigma Aldrich, T8328; dilution – 1:4000), anti-GFP antibody (Sigma Aldrich, G1544; dilution – 1:2000), anti-mCherry antibody (Sigma Aldrich, MAB131873; dilution – 1:2500). Secondary antibodies: anti-mouse IgG (Fab specific)–Peroxidase antibody produced in goat (A9917 Sigma Aldrich, dilution: 1:5000), anti-rabbit IgG (whole molecule)–Peroxidase antibody produced in goat (A0545 Sigma Aldrich, dilution: 1:5000).

The Immunoprecipitation experiments were conducted by following the protocol that was given in the Pierce Crosslink Magnetic IP/Co-IP Kit (Thermo Fisher Scientific). The anti-GFP antibody (Sigma Aldrich, G1544) was used to pull-down the p53-R248Q-EGFP protein.

Fluorescence microscopy

Adherent cells were washed with PBS, followed by fixation of the cells with 4% paraformaldehyde (in PBS, pH: 7.4) and subsequent permeabilization of cells with 0.2% triton

X-100 (in PBS, pH: 7.4). This was followed by consecutive applications of primary and secondary antibodies. Following primary antibody was used: anti-FLAG antibody (Sigma Aldrich, F3165; dilution – 1:400). Following secondary antibodies were used: anti-mouse IgG-FITC (F0257 Sigma Aldrich, dilution: 1:400), anti-mouse IgG–TRITC (T5393 Sigma Aldrich, dilution: 1:400). Intermittent washing was done by PBS buffer, and the blocking step was performed by 1% BSA solution (in PBS, pH: 7.4). DAPI containing fluoroshield (Sigma Aldrich) mounting medium was used.

EGFP and mCherry expressing cells were directly visualized after the paraformaldehyde fixation and PBS washing steps.

Imaging of the cells were done in LSM-700 confocal microscope (Zeiss) by the 63X Plan-Apo (with DIC) objective. Image processing was done in Zen-lite (black) software.

Fluorescence associated cell sorting

Fluorescent associated cell sorting (FACS) was done to analyze the cell cycle status of p53-R248Q and HYPK/HYPK-ΔN expressing cells. Transfected cells were detached from the culture dishes by application of trypsin-EDTA, followed by washing of the cells with PBS. For cell cycle analysis, cells were re-suspended in 500 µL of staining solution that was prepared by adding 20µl of 0.1% tritonX-100, 50 µL of 100 mg/mL ribonulcease-A and 300 µL of 50 µg/mL propidium iodide to 500µl of PBS. Cells were kept in the staining solution for 10 min. Apoptotic cell detection was done by using Annexin V-FITC apoptosis detection kit (Sigma Aldrich, Cat. No. APOAF) using manufacturer's protocol. Briefly, detached cells were washed with PBS and incubated with appropriate volume of Annexin V-FITC and propidium iodide solutions followed by incubation of cells at room temperature for 15 min. FACS experiments were done in BD-FACSARIA-III (Becton Dickenson Biosciences) instrument and the analyses were conducted in BD-FACSDIVA software.

Computational studies

Sequence alignment: Alignment of HYPK protein sequences of different organisms were done in Clustal Omega [44].

Phylogenetic analysis: Phylogenetic analysis of HYPK protein sequences of different organisms were done by following the same process that was described by us in one of our earlier studies with p97/VCP [41]. Aligned sequences were saved in phylip format (.phy) which was subjected to analysis in the Phylip3.695 software package. The aligned sequences were sequentially run in the seqboot, protdist, neighbour and consense programs. The outfile of each program was used as the input file of the next program. The details of the each program's parameter and specifications are available upon request. The cladogram was prepared in Treeview.

RNA secondary structure prediction: The predicted secondary structure of the 5′UTR of HYPK mRNA was generated by using the algorithm of MFOLD.

Three-dimensional structure prediction of RNA: The three-dimensional (3D) structure of the 5′UTR of HYPK mRNA was constructed based on the secondary structure of the 5′UTR of HYPK mRNA. The 3D structure was generated in RosettaCommons.

Statistical test

Comparisons of the means of groups were done by unpaired t-test.

Graphical drawing

The schematic representations of different constructs were generated in Adobe Illustrator.

Acknowledgments

The authors are thankful to Dr. Murali Dharan Bashyam (laboratory of molecular oncology of CDFD) for providing the p53-EGFP clone. The authors also thank the DBT-IPLS facility of the University of Calcutta for providing the ITC facilities. The authors heartily appreciate the members of CFG laboratory for their inputs in different experiments and manuscript preparation. The personnel of the sophisticated equipment facility of the research support of service group of CDFD are acknowledged for their helps in operating different instruments. Research in CFG lab is supported by the core grants of CDFD. DKG is a recipient of graduate study fellowship from the council for scientific and industrial research (CSIR, Government of India).

Authors contributions

AR and DKG identified the scope and rationale of the study. DKG performed all the experiments. AR was responsible for mentoring the experiments, data analysis, resource acquisition and arrangement of funds. AR and DKG wrote the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Ghosh DK, Roy A, Ranjan A. Aggregation-prone regions in HYPK help it to form sequestration complex for toxic protein aggregates. J Mol Biol. 2018;430:963–986.
[2] Choudhury KR, Raychaudhuri S, Bhattacharyya NP. Identification of HYPK-interacting proteins reveals involvement of HYPK in regulating cell growth, cell cycle, unfolded protein response and cell death. PLoS One. 2012;7:e51415.
[3] Otto H, Conz C, Maier P, et al. The chaperones MPP11 and Hsp70L1 form the mammalian ribosome-associated complex. Proc Natl Acad Sci U S A. 2005;102:10064–10069.
[4] Raychaudhuri S, Sinha M, Mukhopadhyay D, et al. HYPK, a Huntingtin interacting protein, reduces aggregates and apoptosis induced by N-terminal Huntingtin with 40 glutamines in Neuro2a cells and exhibits chaperone-like activity. Hum Mol Genet. 2008;17:240–255.
[5] Ghosh DK, Kumar A, Ranjan A. Metastable states of HYPK-UBA domain's seeds drive the dynamics of its own aggregation. Biochim Biophys Acta Gen Subj. 2018;1862:2846–2861.
[6] Ghosh DK, Roy A, Ranjan A. Disordered nanostructure in huntingtin interacting protein K acts as a stabilizing switch to prevent protein aggregation. Biochemistry. 2018;57:2009–2023.
[7] Arnesen T, Starheim KK, Van Damme P, et al. The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of Huntingtin aggregation. Mol Cell Biol. 2010;30:1898–1909.

[8] Weyer FA, Gumiero A, Lapouge K, et al. Structural basis of HypK regulating N-terminal acetylation by the NatA complex. Nat Commun. 2017;8:15726.
[9] Ayyadevara S, Balasubramaniam M, Gao Y, et al. Proteins in aggregates functionally impact multiple neurodegenerative disease models by forming proteasome-blocking complexes. Aging Cell. 2015;14:35–48.
[10] Bell R, Hubbard A, Chettier R, et al. A human protein interaction network shows conservation of aging processes between human and invertebrate species. PLoS Genet. 2009;5:e1000414.
[11] Sakurai H, Sawai M, Ishikawa Y, et al. Heat shock transcription factor HSF1 regulates the expression of the huntingtin-interacting protein HYPK. Biochim Biophys Acta. 2014;1840:1181–1187.
[12] Das S, Bhattacharyya NP. Transcription regulation of HYPK by heat shock factor 1. PLoS One. 2014;9:e85552.
[13] Komar AA, Hatzoglou M. Internal ribosome entry sites in cellular mRNAs: mystery of their existence. J Biol Chem. 2005;280:23425–23428.
[14] Lewis SM, Holcik M. IRES in distress: translational regulation of the inhibitor of apoptosis proteins XIAP and HIAP2 during cell stress. Cell Death Differ. 2005;12:547–553.
[15] Lee KM, Chen CJ, Shih SR. Regulation mechanisms of viral IRES-driven translation. Trends Microbiol. 2017;25:546–561.
[16] Audigier S, Guiramand J, Prado-Lourenco L, et al. Potent activation of FGF-2 IRES-dependent mechanism of translation during brain development. RNA. 2008;14:1852–1864.
[17] Shatsky IN, Dmitriev SE, Terenin IM, et al. Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. Mol Cells. 2010;30:285–293.
[18] Martinez-Salas E, Pineiro D, Fernandez N. Alternative mechanisms to initiate translation in eukaryotic mRNAs. Comp Funct Genomics. 2012;2012:391546.
[19] Costantino D, Kieft JS. A preformed compact ribosome-binding domain in the cricket paralysis-like virus IRES RNAs. RNA. 2005;11:332–343.
[20] Pestova TV, Lomakin IB, Lee JH, et al. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to initiate coding during internal translation initiation of hepatitis C and classical swine fever virus. Genes Dev. 1998;11:67–83.
[21] Lukavsky PJ. Structure and function of HCV IRES domains. Virus Res. 2009;139:166–171.
[22] Belsham GJ. Divergent picornavirus IRES elements. Virus Res. 2009;139:183–192.
[23] Kaufman RJ. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. Genes Dev. 1999;13:1211–1233.
[24] Marfori M, Mynott A, Ellis JJ, et al. Molecular basis for specificity of nuclear import and prediction of nuclear localization. Biochim Biophys Acta. 2011;1813:1562–1577.
[25] Freitas N, Cunha C. Mechanisms and signals for the nuclear import of proteins. Curr Genomics. 2009;10:550–557.
[26] Lui K, Huang Y. RanGTPase: A key regulator of nucleocytoplasmic trafficking. *Mol Cell Pharmacol*. 2009;1:148–156.
[27] Miyamoto Y, Yamada K, Yoneda Y. Importin alpha: a key molecule in nuclear transport and non-transport functions. J Biochem. 2016;160:69–75.
[28] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31:3406–3415.
[29] Jaafar ZA, Kieft JS. Viral RNA structure-based strategies to manipulate translation. Nat Rev Microbiol. 2019;17:110–123.
[30] Hellen CU, Sarnow P. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 2001;15:1593–1612.
[31] Silva JL, Rangel LP, Costa DC, et al. Expanding the prion concept to cancer biology: dominant-negative effect of aggregates of mutant p53 tumour suppressor. Biosci Rep. 2013 33.
[32] Ano Bom AP, Rangel LP, Costa DC, et al. Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. J Biol Chem. 2012;287:28152–28162.
[33] Ng JW, Lama D, Lukman S, et al. R248Q mutation–beyond p53-DNA binding. Proteins. 2015;83:2240–2250.
[34] Haupt S, Berger M, Goldberg Z, et al. Apoptosis - the p53 network. J Cell Sci. 2003;116:4077–4085.
[35] Grainger L, Cicchini L, Rak M, et al. Stress-inducible alternative translation initiation of human cytomegalovirus latency protein pUL138. J Virol. 2010;84:9472–9486.
[36] Bornes S, Prado-Lourenco L, Bastide A, et al. Translational induction of VEGF internal ribosome entry site elements during the early response to ischemic stress. Circ Res. 2007;100:305–308.
[37] Creancier L, Morello D, Mercier P, et al. Fibroblast growth factor 2 internal ribosome entry site (IRES) activity ex vivo and in transgenic mice reveals a stringent tissue-specific regulation. J Cell Biol. 2000;150:275–281.
[38] Morris DR, Geballe AP. Upstream open reading frames as regulators of regulators of mRNA translation. Mol Cell Biol. 2000;20:8635–8642.
[39] Liu B, Qian SB. Translational reprogramming in cellular stress response. Wiley Interdiscip Rev RNA. 2014;5:301–315.
[40] Pacheco A, Twiss JL. Localized IRES-dependent translation of ER chaperone protein mRNA in sensory axons. PLoS One. 2012;7:e40788.
[41] Ghosh DK, Roy A, Ranjan A. The ATPase VCP/p97 functions as a disaggregase against toxic Huntingtin-exon1 aggregates. FEBS Lett. 2018;592:2680–2692.
[42] Roy A, Reddi R, Sawhney B, et al. Expression, functional characterization and X-ray analysis of HosA, A member of marr family of transcription regulator from uropathogenic escherichia coli. Protein J. 2016;35:269–282.
[43] Kumar A, Ghosh DK, Ali J, et al. Characterization of lipid binding properties of plasmodium falciparum acyl-coenzyme a binding proteins and their competitive inhibition by mefloquine. ACS Chem Biol. 2019;14:901–915.
[44] Sievers F, Higgins DG. Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol. 2014;1079:105–116.