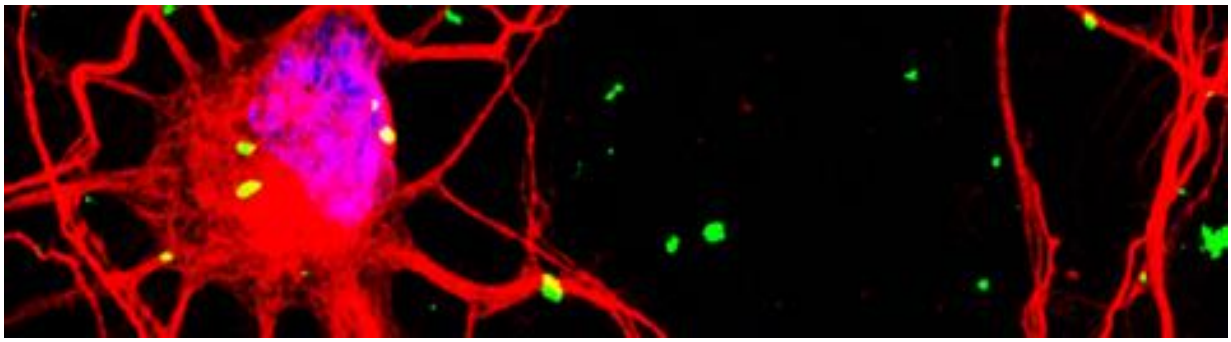


Corso di Dottorato in Neuroscienze
Curriculum Neuroscienze e Neurotecnologie
Ciclo XXIX

Modulation of RE1-Silencing Transcription Factor (REST) *via* DNA and RNA strategies

Autore: **Stefania Criscuolo**

Supervisor: **Fabio Benfenati, MD**
Fabrizia Cesca, PhD

ISTITUTO ITALIANO DI TECNOLOGIA

UNIVERSITÀ DEGLI STUDI DI GENOVA

«Alla fine sono arrivata a credere in una ricerca che io chiamo “La Fisica dell’Anima”, una forza della natura governata da leggi reali quanto la legge di gravità. La regola di questo principio funziona più o meno così: se sei abbastanza coraggiosa da lasciarti indietro tutto ciò che ti è familiare e confortevole e che può essere qualunque cosa, dalla tua casa a vecchi rancori, e partire per un viaggio alla ricerca della verità, sia esterna che interna, se sei veramente intenzionata a considerare tutto quello che ti capita durante questo viaggio come un indizio, se accetti tutti quelli che incontri strada facendo come insegnanti, e se sei preparata soprattutto ad affrontare e perdonare alcune realtà di te stessa veramente scomode, allora la verità non ti sarà preclusa.»

Eat, Pray, Love

E. Gilbert

Table of Contents

Chapters

1 Introduction

1.1 The RE1-Silencing Transcription Factor (REST)

The acquisition of our identity depends on extrinsic and intrinsic factors that, if harmoniously orchestrated, allow us to BE our proper phenotype. Likewise, neuronal maturation occurs through an intricate network of signalling molecules and epigenetic modifications that leads to the acquisition and maintenance of neuronal identity.

A key actor in this regulation process is the repressor element 1 (RE1)-silencing transcription factor (REST), also known as Neuron-Restrictive Silencer Factor (NRSF), and its co-repressors that play a particularly important role in neuronal maturation and function. REST is a member of the Kruppel-type zinc finger transcription factor family. It represses transcription by binding a DNA sequence of 21–23 bp, called neuron-restrictive silencer element (NRSE, also known as RE1)[1, 2]. REST was discovered in 1995. Originally believed to silence neuronal gene expression in non-neuronal cells, several works revealed its diverse functions as a transcriptional and epigenetic regulator, whose function is highly context-dependent in neural and non-neural cells, where it can act as both silent repressor and active protector [3].

The human REST gene spans 24 kb of genomic DNA, which are transcribed into a full-length messenger RNA of 1097 amino acids, encoding a glycoprotein of 210 kDa. REST harbors three functional domains: a DNA binding domain containing nine zinc-finger motifs that bind to the RE1 motif, and two independent repressor domains. The repressor domain at the amino terminal interacts with mSin3, a corepressor that recruits histone deacetylases (HDAC-1, -2), while the domain at the carboxy-terminal interacts with the CoREST complex, which contains HDAC1/2, the ATP-dependent chromatin-remodelling enzyme BRG1, the H3K4 demethylase LSD1 and the H3K9 methylase G9a. The recruitment of REST and its associated corepressors onto target genes leads to the removal of several modifications that are associated with active gene transcription, thus altering the chromatin landscape to achieve gene silencing **(Fig. 1.1)** [4].

REST mRNA is composed of three alternative 5' non-coding exons associated with different promoters, three coding exons and an internal alternative exon that is spliced into some neuron- and disease-associated transcripts. Five splice variants of REST have been identified [5], four of which

(REST1, sNRSF, REST-N62 and REST4) encode for truncated proteins, whereas REST5 lacks the second exon. REST4 retains the N-terminal repressor domain and five of the nine DNA-binding zinc fingers. This neuron specific isoform is conserved in human, mouse and rat [6] and is formed by an insertion of 16 nucleotides, followed by an in-frame stop codon, regulated during alternative splicing by neural-specific Ser/Arg repeat-related protein of 100 kDa (nSR100/SRRM4). Thus REST4 itself does not bind to RE1 sites but has a dominant negative function by inhibiting the REST-DNA contact [7, 8]. Conversely, REST directly represses nSR100 in non-neural cells to prevent the activation of neural-specific splicing events.

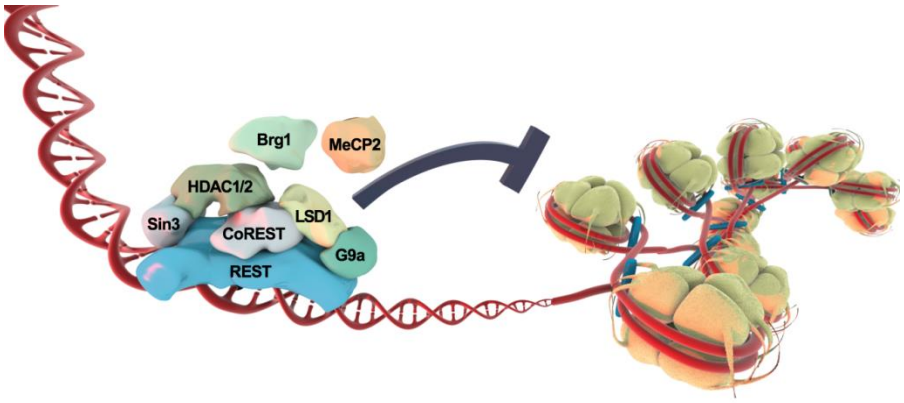


Figure 1.1. Representation of REST and associated corepressors when bound on a target sequence. The initial step in REST-mediated gene repression is the recruitment of the mSin3a and CoREST co-repressors to the RE1 *cis*-sites on target gene promoters. mSin3a and CoREST in turn recruit the chromatin-remodelling enzymes HDAC1, HDAC2, G9a and LSD1. The REST-DNA interaction is stabilized by the ATP-dependent enzyme BRG1. When bound to chromatin, the BRG1 complex repositions nucleosomes with respect to DNA, presumably allowing REST to form a more stable interaction with the DNA. The REST complex creates a negative chromatin landscape, thus preventing gene transcription. Image taken from Hermanson [9]

1.1.1 REST: a neuroepigenetic factor

Nervous system development requires the expression of negative and positive factors that regulate the precise spatial and temporal acquisition of the different neural lineages. This orderly acquisition of the various neural fates is mediated by specific networks of transcriptional activators and repressors in response to environmental and intrinsic cues. A number of studies have demonstrated that REST is a key determinant of neuron-specific gene expression, and an important regulator of neuronal gene expression during early embryogenesis. Indeed the full deletion of REST in mice is lethal at embryonic day 11.5 [10]. REST knockout (KO) embryos develop normally until embryonic day 9.5, after which they undergo massive apoptotic cell death that results in

malformations of the developing nervous system and insufficient growth. Interestingly, the absence of REST per se does not activate REST target genes, indicating that the expression of neuronal differentiation genes requires both the absence of REST and the presence of other tissue-specific positive transcription factors [10]. REST is predicted to repress more than 2000 neuron-specific genes, including genes encoding synaptic vesicle proteins, ion channels, neurotransmitter receptors, and miRNAs, which have been analyzed in non-neuronal tissues and undifferentiated neural precursors [3, 11, 12]. Its expression decreases during neuronal development allowing the correct differentiation of neurons.

Recent works have revealed different mechanisms through which REST orchestrates the epigenetic regulation of target genes, suggesting that it can also activate gene transcription [13]. For example, REST binds the short form of TET3, guiding it to the promoters of RE1-containing genes. TET3 is the major methylcytosine dioxygenase expressed in neurons, which catalyses the conversion of 5 methylcytosine to 5 hydroxymethylcytosine. This modification is labile and can rapidly regenerate unmethylated cytosines, which in turn promote gene activation [14]. In addition, the REST-TET3 complex interacts with the histone-lysine N methyltransferease and two other H3K36 methyltransferases, to add a trimethylation moiety to the core histone protein H3, thus generating H3K36me3, which is a strong mark of gene activation [14]. REST cooperates also with other remodelling epigenetic factors, such as the methyltransferase enhancer of Zeste homologue (EZH2), the catalytically active component of Polycomb repressor complex 2 (PRC2), which act synergistically to silence target genes with REST [15]. Altogether, these observations reveal that the mechanisms by which REST controls the expression of neuronal genes are more complex than initially appreciated.

1.1.2 REST in diseases

The control of REST levels is so important that its dysregulation, both increment or decrement, gives rise to several neuropathologies. In brain, high REST expression levels have been associated to medulloblastoma and glioblastoma tumors [16, 17]. REST knockdown in human medulloblastoma cells abrogated their tumorigenic potential in mouse orthotopic models, whereas its constitutive expression in Myc-immortalized neural progenitors promoted tumor formation in vivo. REST levels are also related to the overall survival rate in patients affected by medulloblastoma, indeed patients with high REST levels have the worst overall survival compared with patients with REST-negative or low expressing tumors [18, 19]. REST is involved in the pathology of human glioblastoma multiforme (GBM), as it regulates the self-renewal, survival and

differentiation of malignant GBM cells. Thus, REST-regulated molecular circuitries may represent sensible targets for therapeutic intervention in GBM [17].

Epileptic seizures and ischemia are also associated to an increase of REST levels [20, 21]. In rats, ischemia induces an increase in REST expression in the CA1 region of the hippocampus, repressing the transcription of the μ-opioid receptor 1 (Opmr1) gene in inhibitory interneurons [22]. Inhibition of Opmr1 in turn induces an increased release of γ-aminobutyric acid (GABA), which lowers neuronal activity and probably has a neuro-protective effect [22]. REST plays also a role in downregulation of 4 aminopyridine (4AP)-induced neuronal hyperexcitability, re-establishing a physiological spiking activity in the neuronal network. As shown in the work of *Pozzi et al*, the long-term treatment of neural cultures with 4AP enhances the expression of REST, which results in down-regulation of Nav1.2 channels, thus decreasing the neuronal Na^+ current density. This work showed that REST is involved in the molecular mechanisms driving the homeostatic response induced by neuronal hyperactivity [23]. Quite a different result was obtained in the pentylenetetrazol (PTZ) model of acute seizures, using animals in which the conditional deletion of REST was achieved through mating with mice expressing the Cre recombinase under the neuron-specific enolase (NSE) promoter, which ablates the REST gene from most if not all neurons. Although the initial clonic convulsions caused by PTZ were not different between REST cKO and control mice, tonic convulsions and death required a higher PTZ dose [24] These findings would suggest that REST might contribute to seizure initiation or generalization. Therefore the role of REST in epigenetic mechanisms related to epilepsy is still debated.

Recently, a new role of REST as protective factor in neurodegenerative diseases has been unveiled. In the work of *Lu et al*, the role of REST in stress resistance, ageing and Alzheimer's disease (AD) has been investigated. The authors measured the protein and mRNA levels of REST in the prefrontal cortex (PFC) of three groups of individuals: young adults (20-35), aged (73-106), and patients with AD. An increment of REST levels in the ageing human PFC was observed, compared to the young adult group, thus suggesting that REST expression is increased in ageing neurons. ChIP-seq analysis showed that REST represses genes involved in cell death, such as p38 MAP kinase, BAX, and PUMA, which are associated with AD and dementia. Reduced REST binding to these genes, and a consequent higher mRNA expression, were characteristic features of AD brains. Because of the observed REST-mediated repression of pro-apoptotic genes, in order to assess the role of REST in neuroprotection, REST deficient neurons were treated with hydrogen peroxide and their vulnerability to oxidative stress was assessed. REST-deficient neurons showed a markedly increased degeneration and cell death compared to controls, consistent with a neuroprotective role

of REST [25]. REST activity has been also implicated in the pathogenesis of Huntington's disease (HD). Wild-type huntingtin binds REST, sequestering it in the cytosolic department, while the mutated, pathogenic form loses the ability to bind REST thus causing its accumulation in the cell nuclei. This causes the consequent reduction of the transcription of REST-dependent neuronal genes including BDNF, which is a fundamental trophic factor required for the correct activity of cortico-striatal synapses and for the survival of the GABAergic medium-sized spiny striatal neurons that die in HD [26, 27].

The neuroprotective role of REST has been explored also in the pathogenesis of prion disease. In the work of *Song et al*, the authors treated primary cultured cortical neurons with the neurotoxic prion protein fragment 106-126 (PrP106-126), a widely used model for the in vitro study of prion-associated pathological damage. REST protein expression was examined at various time points (up to 48 h) by immunofluorescence and western blot analysis after 24 h of treatment with PrP106-126. REST expression was 1.6 to 1.9 fold higher than in non treated samples between 6 and 24 h. Thereafter, its expression gradually declined reaching control levels at 48 h, which was probably due to proteasomal degradation. The nucleo-cytoplasmic shuttling of REST was also analysed, revealing decreased protein levels in the cytoplasm and a corresponding increase in the nucleus, compared with the untreated control group. In addition, overexpression of REST rescues the physiological alterations induced by PrP106-126, decreasing apoptosis by stabilizing the level of the pro-survival protein FOX1. The authors also explored the relationship between REST and LRP6, a key component of the LRP5/LRP6/Frizzled co-receptor group that is involved in the canonical Wnt pathway. They observed that the activation of Wnt-β-catenin signaling induced by LRP6 receptor after stress stimulus in turn activates transiently REST and β-catenin, suggesting a critical role of REST and Wnt-β-catenin signalling in maintaining cellular integrity and protecting neurons from death [28].

REST is also significantly depleted in frontotemporal dementia and dementia with Lewi bodies. In samples from patients affected by these pathologies, REST is lost from the nucleus and it appears in authophagosomes with pathological misfolded proteins, including Aβ, phosphorylated tau and alpha synuclein, suggesting that epigenetic regulation of chromatin may modulate the cognitive outcome of a variety of pathologic states [25].

In non-neural cells, including epithelia, REST expression is high and plays the role of tumor suppressor. Indeed, reducing REST function through RNAi or expression of dominant-negative REST peptides, promotes transformation of human epithelial cells. Conversely, reconstitution of REST expression elicits a dramatic proliferation defect in colon cancer cells that have lost endogenous REST function [29, 30]. More evidence in support to the tumor suppressor role of

REST comes from works demonstrating that REST expression was absent in a subset of small cell lung cancers, and in prostate cancer [31, 32].

1.1.3 REST and REST-dependent pathways as targets of molecular therapy

Given the relevance of REST dysregulation in a number of neuronal pathologies, various molecular strategies have been developed to modulate its action. In order to decrease the activity of REST in Huntington disease, a oligonucleotide (ODN)-based approach has been used, consisting in a short double-stranded-DNA that mimics the RE1 cis-site to sequester REST and reduce its activity. Despite the promising results, the ODN decoy strategy faces great challenges when target gene expression has to be inhibited *in vivo* only in a single organ or tissue type. In fact, systemic delivery of ODNs may result in a widespread uptake and potential nonspecific side effects. Moreover, a successful use of ODN decoys depends on the efficient delivery of the synthetic DNA to target cells [33]. The reduction of REST activity *in vivo* has been achieved through the use of adeno-associated viral vectors (AAVs) expressing a dominant negative REST protein (DN:REST). Such protein comprised only the Zn finger domain region, without the N- and C-terminus, and was able to compete with the wild-type protein for the binding to target promoters, thus maintaining an open chromatin conformation [34]. However, the main drawback of this approach is that the long-term expression of DN:REST driven by the AAV promoter produced excessive levels of the protein. To achieve a more selective competition of REST a small peptide that blocks the mSin3-REST interaction has been used, which inhibited the formation of the multi- protein complex on gene promoters. This approach is promising, however small peptides still present major challenges *in vivo*, such as limited tissue penetration and possible immune reaction that can cause toxicity and neutralize their therapeutic activity [35].

1.2 Light-sensitive Proteins and Optogenetics

Optogenetic techniques recently emerged as a powerful and promising tool, particularly in the field of neuroscience. The optogenetic approach is based on the artificial introduction of photo-sensitive proteins into or on the surface of a target cell, with the aim of making specific cellular processes light-sensitive. The earliest photoactivatable systems were light-controlled caged molecules, such as neurotransmitters and ion channels [36]. The first optogenetic tool used in primary neurons was channelrhodopsin2 (ChR2), which enables the light-driven depolarization of neuronal membranes, and halorhodopsin (NpHR), [37-39] which induces the light-tunable hyperpolarization of membranes to inhibit action potential. These molecules are still widely employed, but have been flanked by new and more refined optogenetic tools, which provide a more robust and precise cell manipulation, enabling more complex experimental designs. Another interesting class of optogenetic protein is represented by OptoXRs opsin–receptor chimaeras in which the intracellular loops of rhodopsin were replaced with the intracellular loops from other G protein-coupled receptors (such as adrenergic receptors) to obtain a light-induced activation of specific G protein-coupled signalling pathways in targeted neurons in vitro and in freely moving mammals. The use of OptoXRs allows to alter excitability in a given neuronal population on a longer time scale, bypassing the need to continuously evoke action potential firing (**Fig. 1.2**).

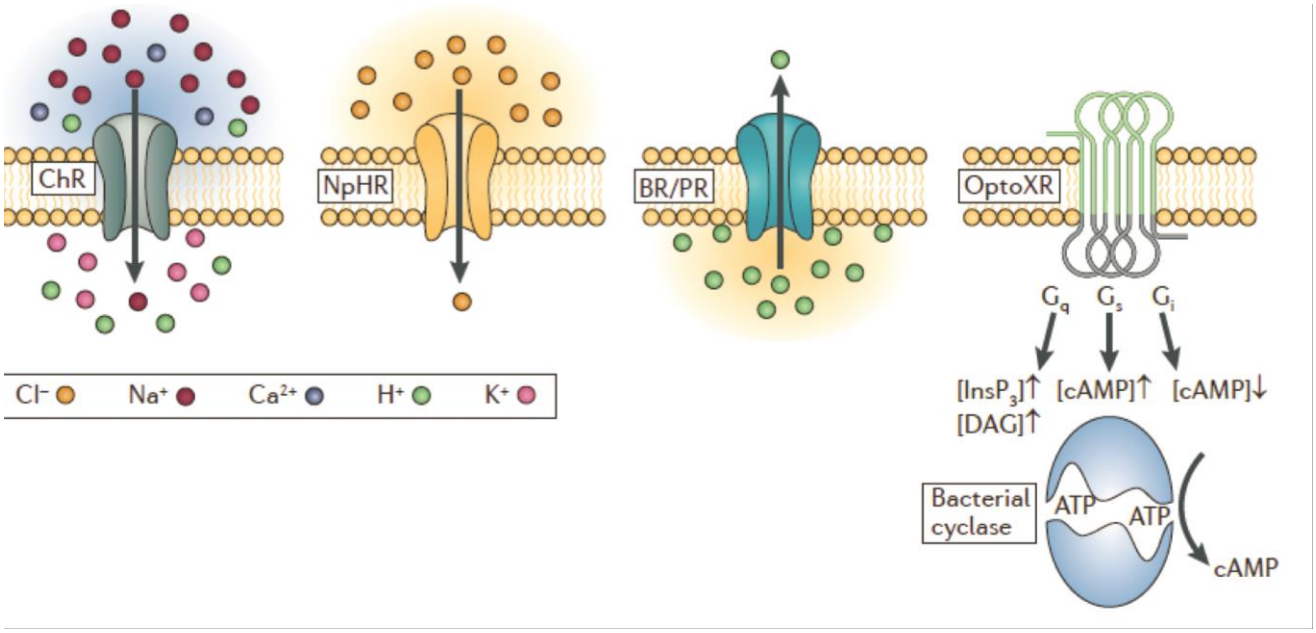


Figure1.2. Membrane-associated optogenetic tools. Optogenetic molecules include cation permeable channels for membrane depolarization (channelrhodopsin, ChR), chloride pumps (halorhodopsin, NpHR) and proton pumps (bacteriorhodopsin or proteorhodopsin (BR/PR)) for membrane hyperpolarization. The light-activated membrane-bound G protein-coupled (OptoXR) or soluble (bacterial cyclase) receptors can be employed to trigger the activation of various signalling cascades. Taken from Tye *et al* [40].

Besides the described membrane-associated proteins, other classes of intracellular optogenetic molecules are used. These proteins mainly belong to three categories, i.e. phytochromes, cryptochromes and light-oxygen voltage domains (detailed in the next section). Phytochromes are expressed in plants and cyanobacteria and contain a chromophore that consists of a covalently linked tetrapyrrole, Upon red light exposure, the isomerization of the tetrapyrrole cofactor induces a conformational change of the protein structure, which switches from a red-absorbing (Pr) to a far-red-absorbing (Pfr) state. In the dark, the Pfr state spontaneously reverts to the Pr state with slow kinetics, alternatively, a far-red light stimulus can induce a quicker conversion.[41] (**Fig. 1.3a**). Cryptochromes are flavin adenine dinucleotide (FAD)-binding proteins that regulate growth in plants and circadian clocks in animals. In plants, blue light induces FAD reduction and a consequent conformational change in the cryptochromes, which spontaneously reverses in the dark. In cryptochrome 2 (CRY2) of *A. thaliana*, the light-induced conformational change facilitates the interaction with the calcium and integrin-binding protein 1 (CIB1). This event has been used by *Kennedy et al.* to recruit proteins to the cell membrane or to induce the binding between a transcriptional activation domain and a DNA-binding domain [42](**Fig. 1.3b**).

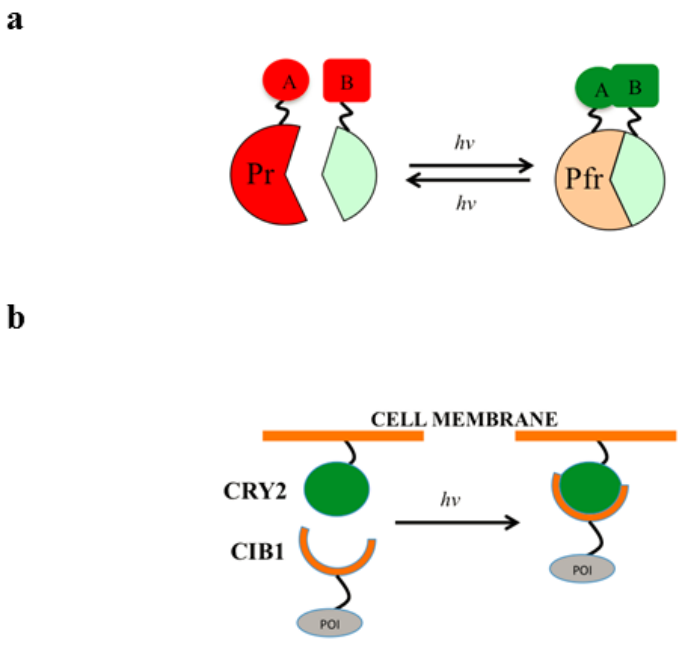


Figure 1.3. Intracellular optogenetic tools. a. Phytocrome-based strategy. Upon red light absorption, Pr interacts with a target domain, inducing the activation of the fused proteins, Pr, red light-absorbing molecule; Pfr, far red light absorbing molecule. **b.** Examples of cryptochrome-associated strategies. CRY2-CIB1 interaction recruits a selected protein to cell membrane.

1.3 Light Oxygen Voltage (LOV) Domains

Light, oxygen, or voltage (LOV) domains were first identified as protein sensors involved in blue-light absorption in plants. Phototropins LOV domains are included in key photoreceptor kinases that modulate light-dependent processes such as phototropism, light-directed chloroplast movement and signalling processes across all life kingdoms, particularly in plants, fungi and bacteria [43, 44]. Structurally, LOV domains belong to the large and versatile Per-ARNT-Sim (PAS) domain superfamily. Their molecular conformation reflects the canonical PAS domain fold, with a central structure containing a five-stranded antiparallel β sheet and several α helices (**Fig. 1.4a**). The inner portion of this α/β core is non-covalently associated to the flavin mononucleotide (FMN) cofactor, the principal form of riboflavin found in cells and tissues. The small FMN cofactor absorbs blue light with a maximum peak ranging from 450 to 470 nm. Upon photon absorption, FMN enters in a photo-excited state enabling the formation of a covalent thioether adduct between the 4a carbon (C4a) of its isoalloxazine ring and a conserved cysteine residue (C450) within the LOV domain [45] (**Fig. 1.4b**). On the microsecond time scale, this induces a larger conformational rearrangement of the inner β sheet, which propagates to the outer structures. The external portion of the LOV domain is formed by one amino- and one carboxy-terminal α helix, termed A′α and Jα, respectively [46, 47], which are associated to the central sheet, opposite to the FMN-binding pocket. Following the photo-induced structural changes, both terminal α helices are displaced and unfolded. Giving that the carboxy-terminal Jα helix function as an interaction hub, as in other PAS domains, the rearrangement of this region mediates signal transduction to the effector domains.

LOV domain-containing proteins can be subdivided into five functional categories, depending on their function: phototropins, proteins regulating circadian rhythms, LOV histidine kinases, LOV-STAS proteins, and LOV phosphodiesterases. Amongst the various LOV-containing proteins, phototropins are particularly interesting as they contain two LOV domains (LOV1 and LOV2) instead of a single LOV sequence. LOV2 exhibits a higher efficiency of the light-dependent switching and a slower rate of dark recovery, compared to LOV1. Considering the properties described, the phototropin LOV2 domain has been preferentially used in the design and construction of engineered photoreceptors. In this context, the *Avena sativa* LOV2 (AsLOV2) domain has been widely employed. Strickland et al. [48] created a chimeric protein fusing the

AsLov2 Jα helix to the N-terminus of the Trp repressor protein (TrpR) obtaining a longer shared linking helix (**Fig. 1.4c**). The resulting fusion protein, named LOV-TAP, displayed light-regulated activity, showing a weak DNA affinity in the dark state and a ~5-fold higher affinity upon illumination. In this model, the shared linker helix was folded against the LOV core in the dark obstructing the TrpR domain, while upon blue-light stimulus the linker helix relaxed, thus restoring the structure and function of TrpR. With a different approach, Wu et al. used the AsLOV2 domain to modulate the activity of the small GTPase Rac1 [49]. The rationale underlying this strategy was to block the active site of Rac1 in a light-dependent manner. In the chimeric protein, named PA-Rac1, LOV2 was fused via its C-terminal Jα helix to the N-terminus of Rac1 with the intent to obstruct the active site in the dark (**Fig. 1.4c**). Following blue-light irradiation, the Jα helix unfolded, releasing the Rac1 domain and exposing its active site. The open PA-Rac1 was able to bind its downstream effector protein PAK with approximately the same affinity as wild-type Rac1, whereas in the dark the affinity was 10-fold lower. Moreover, the expression of PA-Rac1 allowed to control the motility of cultured fibroblasts when the appropriate light stimulus was provided. In conclusion LOV photosensor domains, and particularly the AsLOV2 domain, can efficiently be coupled to effector proteins by domain fusion to obtain light-induced changes in the biological activity of cells and organisms.

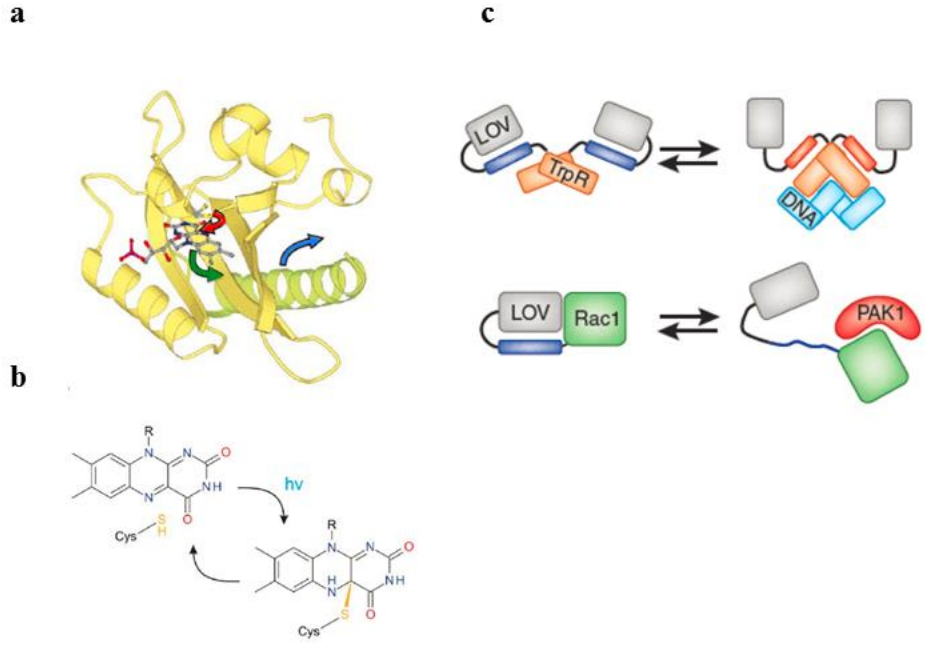


Figure 1.4. Light, oxygen, voltage (LOV) domain and some of its applications. a. AsLOV2 is folded in the canonical PAS conformation and binds FMN. Blue-light absorption induces the formation of a covalent bond between the flavin ring and the conserved cysteine residue C450 (red arrow). Light signals are propagated to and through the β sheet (green arrow). Here, they promote unfolding of the C-terminal Jα helix (in green; blue arrow). **b.** The LOV photocycle depends on the blue-light induced formation of a thioether bond between C450 residue and atom C(4a) of the flavin cofactor. The lit state reverts thermally to the ground state (adapted from *Moglich et al*) [41]. **c.** In LOV-TAP, the Trp repressor has a weak affinity for

DNA in the dark, whereas the DNA binding domain is exposed upon illumination, thus promoting Trp-DNA binding. In PA-Rac, the enzyme Rac1 is fused to the C-terminus of AsLOV2. Upon blue light stimulus, the Jα helix unfolds allowing the interaction between Rac1 and PAK1 (adapted from *Hahn KM et al*) [50]

1.4 Target RNA to Play with Cells!

RNA molecules play an active role within cells by catalyzing a number of fundamental biological reactions. RNA controls the expression of the encoded protein by providing a template for translation, and regulates translation through alternative splicing or RNA editing. Several types of non coding RNA (ncRNA), such as micro- or short interfering RNAs (miRNA, siRNA) can downregulate gene expression through specific association to a complementary mRNA sequence [51, 52]. Moreover localization of mRNAs to specific destinations within a cell or an embryo is important for local control of protein expression. Thus, because of the high number of processes in which RNA is involved, it represents an optimal candidate to manipulate these functions in living cells. Many progresses have been made in designing DNA binding proteins able to activate, silence or modify double-stranded DNA with chosen specificity [53]. Promising systems for DNA manipulation are now available such as Zinc Finger (ZF) domains, the transcription activator-like effector (TALE), and the guide RNA-based CRISPR-associated (CAS9) system, a powerful tool in the field of genome editing and recently also in RNA knowdown [54]. However equivalent programmable RNA binding modules to enable manipulation of a specific RNA are not widely used [55].

The most common approach to manipulate RNA is through RNA interference (RNAi), which uses sequence-specific siRNAs, synthetic miRNA and miRNA inhibitors that regulate transcriptional or post-transcriptional gene silencing [56]. However the RNAi technology has some limitations, such as the instability of siRNAs, their off target effects and the immune response elicited by their injection *in vivo*. Thus, the development of tools able to safely edit RNA with minimal side effects is highly needed and desirable.

1.4.1 RNA binding proteins

RNA-binding proteins (RBPs) play an essential role in every aspects of RNA biology. Their dynamic association with RNA determines the cellular localization, lifetime, processing and translational rate of coding and non-coding RNAs. In the last two decades researchers have focused on understanding the structural basis for RNA recognition, in the interest of designing RNA-binding proteins with tailored sequence specificity for therapeutic applications. It is now established that RBPs are composed of modular repeats of just few domains that, combined in various arrangements, can generate versatile recombinant proteins capable to bind RNA with higher affinity and specificity than the individual domains [52].

Several efforts have been made to alter gene expression by targeting RNA through RNA-binding proteins. The most known and used RBDs are the spliceosomal protein U1A, the MS2 coat protein that link proteins to RNA with the purpose of tracking mRNA localization, and the iron responsive element (IRE) RNA-binding protein IRP-1 fused with the C-terminal region of the eIF4G human translation initiation factor, used to drive translation of reporter genes. However all these RBPs require the insertion of the protein-binding sequences into the target RNAs, since they are unable to operate on endogenous mRNAs, thus requiring the engineering of an artificial reporter bearing the specific recognition site. The RNA Recognition Motifs (RRMs) have a size of 80-90 amino acids and are characterized by α helices packed against a four stranded anti parallel β-sheet. The centre of the β-sheet allows the recognition of two nucleotides on one side and of two additional nucleotides on the other side. This conserved platform is not sufficient for sequence specificity and many dynucleotide combinations are recognized by this motif; moreover the topology of the bound RNA varies in each complex, depending on its secondary structure. Because of all these reasons, it has been difficult to engineer RNA binding proteins using RRMs [57].

Another class of RNA binding domains is represented by repeats that generate extended structures with a large surface suitable for protein-protein and protein-RNA interactions. Many of these proteins have been identified, including Pumilio and FBF protein (PUF) (**Fig. 1.5a**, described in detail in the next section), pentatricopeptide repeats (PPR) and trp RNA-binding attenuation proteins. In all these proteins, each repeat interacts with one base in a modular fashion, thus dictating sequence specificity in a predictable manner (**Fig. 1.5**). PPRs, less characterized than PUF proteins, localize primarily in mitochondria and chloroplasts where they influence various aspects of RNA metabolism [58]. PPRs contain 2-30 tandem repeats each one composed by 35 amino acids forming a super-helical binding surface (**Fig. 1.5b**). The RNA is bound in an anti-parallel

orientation with respect to the PUF protein. Each motif binds one nucleic acid; the amino acid in charge of recognition is at position 6, distinguishing purines from pyrimidines. The next most important amino acid is at position 1, which distinguishes between C/A and G/U. The amino acid at position 3 may also affect binding and specificity directly or indirectly, providing hydrophobic interactions with the RNA bases. PPR proteins do not discriminate well between uridine and cytidine and because of this feature show some limitations in sequence specificity with appreciable off target effects. Another limitation is represented by the physiological localization of these proteins to organelles, and their widespread diffusion in plants and groups of protists and metazoans, which pose difficulties when transferring these tools into mammalian cells.

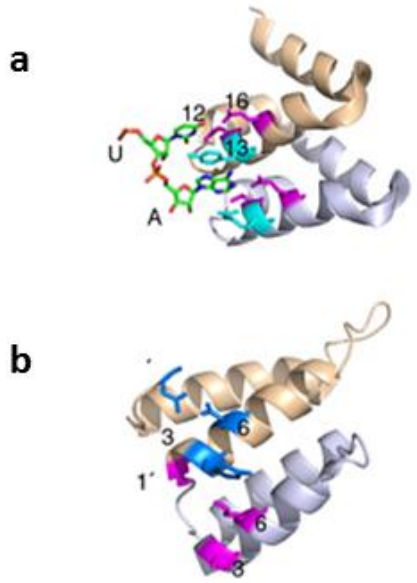


Figure 1.5. Conformational structure of PUF and PPR protein with RNA. a. Crystal structure of PUM1 (residues 996-1067; PDB ID 1M8W). The RNA base contacts residues 12 and 16 (magenta) in the PUF repeats. Residue 13 (cyan) stacks between adjacent nucleotides. **b.** Crystal structure of two PPR motifs from human mitochondrial RNaseP (residues 96-175; PDB ID 4G23). Amino acids that determine RNA binding specificity are shown in magenta and blue for repeat 1 (light gray ribbon) and repeat 2 (light brown ribbon), respectively. Adapted from *Yagi et al* [58].

1.4.2 PUF proteins

The PUF family of proteins is named after the founding members Pumilio in *Drosophila melanogaster* (DmPum) and fem-3mRNA-binding factor (FBF) in *Caenorhabditis elegans*. PUF homology proteins have been found in all eukaryotic organisms [59], where they regulate diverse processes including stem cell maintenance, organelle biogenesis, oogenesis, neuron function, and memory formation. At the molecular level, PUF proteins promote translational regulation of target mRNAs by first interacting with conserved cis-elements in the 3' untranslated region (UTR)[60]. PUF proteins are typically composed of 8 imperfectly repeated 36 amino acid motifs (PUF repeats) flanked by conserved sequences, which form a sequence-specific single-stranded (ss) RNA-binding domain. Each repeat is characterized by three helices that pack together to form a crescent-shaped right-handed super helix with a continuous hydrophobic core [61]. The RNA binds to the concave surface of the protein, where each repeat interacts with a single RNA base. The N-terminal repeat interacts with the 3’ end of the mRNA sequence, in an antiparallel configuration (**Fig. 1.6**).

The crystal structure of a PUF domain from the human Pumilio 1 (HsPum1) protein in complex with RNA revealed its modular recognition code [62]. In each repetition, the amino acids at position 12 and 16 interact with the Watson-Crick edge of RNA bases, while amino acid at position 13 forms stacking interactions between aromatic rings of adjacent RNA bases. These three RNA-binding residues in HsPum1 are contained within a five-residue motif represented as (12-13-X-X-16) (X being any hydrophobic residue). In this code, uracil is recognized by asparagine, glutamine and tyrosine; guanine by serine, glutamate and asparagine; adenine by cysteine, glutamine and arginine, and cytosine, identified by yeast three-hybrid selection [62], by serine, arginine and tyrosine (**Fig. 1.6**).

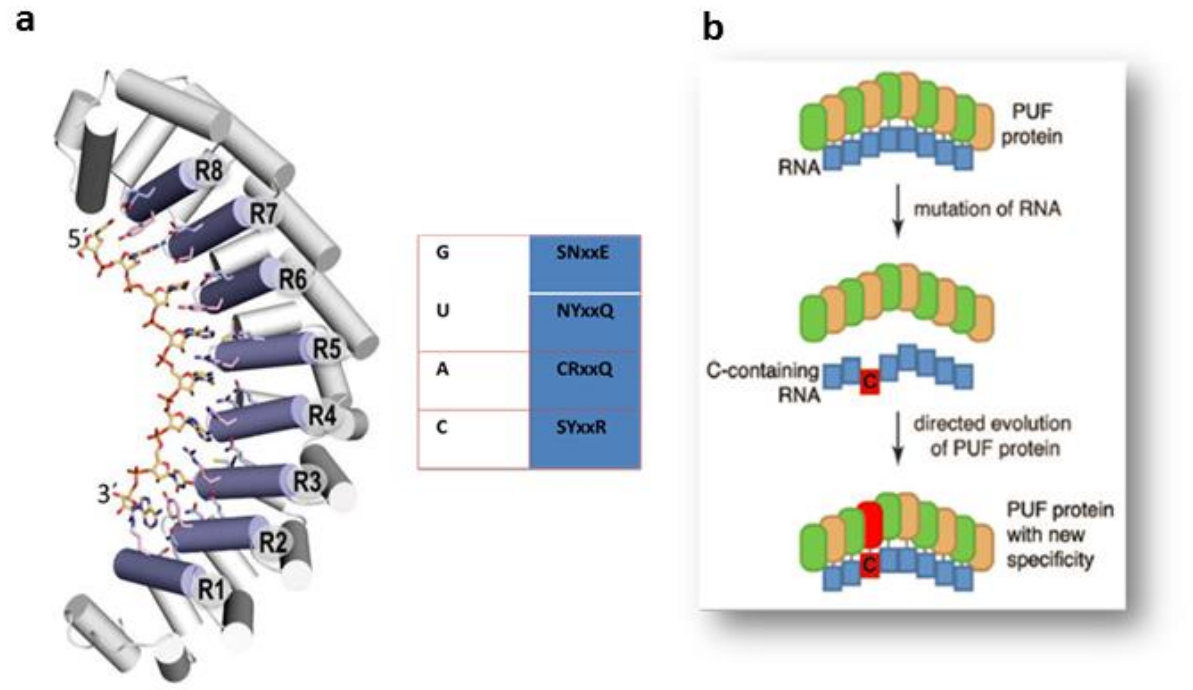


Figure 1.6. PUF protein and its modular recognition code used to engineered artificial PUFs. a. *Left panel* the crystal structure of HsPum1 binds to RNA in antiparallel manner by an arc-like shape disposal. Each repeats bind to single ribonucleotide following the recognition code reported in table (*Right panel*) In each repeat three RNA-binding residues, contained within a five-residue motif represented as (12-13-X-X-16) (X being any hydrophobic residue), are involved in RNA base recognition. **b.** The code of PUF proteins allows to design artificial RNA binding proteins specific for a selected sequence. Rearranged from *Filipovska et al* [62]

PUF proteins are versatile regulators, which employ multiple mechanisms to regulate their mRNA targets, repressing/ activating translation or contributing to mRNA localization [63]. Yeast PUF5 acts as a post-transcriptional repressor recruiting the Ccr4-Pop2-NOT deadenylase complex on its specific recognition sequence in the 3'UTR of the target mRNA, shortening the poly(A) tail and consequently influencing both RNA stability and translation [64]. In addition to the deadenylase, the repressive complex recruited by PUF5 includes the decapping factors Dhh1 and Dcp1, which are associated with the Ccr4-Pop2-NOT complex and cause mRNA repression by affecting the hydrolysis of the 5' cap. This mechanism of repression is conserved; however other repression mechanisms have also been investigated. For instance *D. melanogaster* PUM can recruit d4EHP on mRNA via its cofactor Brat, inhibiting translation by competing with the translational initiator factor eIF4E [65]. Two Pumilio proteins, Pum1 and Pum2, have been found in humans. They recruit Nanos, which is also involved in RNA localization, and Brat, which binds eIF4 inhibiting

transcription initiation. In addition, Pumilio has cap-binding activity itself and so it may compete with eIF4 to block initiation [66]. In some cases Pumilio could positively regulate mRNA stability and translation. In *X. laevis* cyclin B mRNA, PUM- and CPEB-binding sites cooperate to trigger translational activation, probably by stabilizing CPEB on the transcripts [67]. In *C. elegans*, PUF (named FBF) physically interacts with and activates the CCF-1/Pop2p deadenylase and also physically interacts with the GLD-2 poly(A) polymerase, enhancing its activity and ultimately affecting mRNA polyadenilation and translation *in vitro* [68]. One other mechanism of transcriptional activation is through competitive binding with repressor. Ribonucleoprotein-immunoprecipitation microarrays and genome-wide analyis revealed that PUF motifs are enriched around predicted miRNA binding sites and that high-confidence miRNA binding sites are significantly enriched in the 3′UTRs of experimentally determined human PUM targets, strongly suggesting an interaction with the miRNA regulatory system [69]. Last but not least, experiments in yeast suggested an involvement of PUF proteins also in mRNA localization, indeed, PUF3 localizes mRNA to mitochondria, and PUF6 contributes to the asymmetric localization of ASH1 during transport to the yeast bud [70].

1.4.3 Applications of PUF proteins

The elucidation of the RNA recognition code of PUF proteins allowed engineering of artificial RBPs for targeting endogenous mRNAs, permitting the manipulation of the transcriptome. These artificial constructs are composed of a RNA recognition module that binds a specific RNA target, and by a catalytic domain. Such constructs can be implemented in the regulation of a broad range of biological processes [71]. For instance, PUF constructs have been used for mRNA tracking and localization of endogenous mitochondrial RNA. *Ozawa et al* developed a tracking system to target mtRNA to the mitochondrial matrix; this system is composed by an enhanced GFP or Venus split in two parts, where each fragment is fused to a human PUMILIO1 specific for the NADH dehydrogenase subunit 6 (ND6) mtRNA. The addition of a mitochondrial targeting signal (MTS) to the PUF proteins allowed the correct localization to the mitochondrial matrix, permitting at the same time the real-time imaging of ND6 mtRNA (**Fig. 1.7 a**)[72]. One similar tool has been developed by *Cooke et al* to regulate translation in *X. laevis* oocytes through the stabilization of mRNA poly(A) tail. In this work, FBF-2 from *C. elegans* was fused to the poly(A) PAP polymerase GLD-2. The resulting fusion protein was able to guide polyadenylation of luciferase mRNA containing a specific FBF recognition site. In the same work FBF-2 was also fused to CAF1b, which could remove poly(A) from a radiolabeled RNA reporter (**Fig. 1.7 b-c**) [73]. In another study

the PUF domain was designed to target the 5' UTR of mRNA and fused to the eukaryotic translation initiation factor 4E (eIF4E) to enhance translation of a reporter luciferase mRNA. The light-inducible heterodimerization of PUF and eIF4E through light-sensitive protein partners was also demonstrated, which allowed light-inducible translation activation [74]. In another approach the manipulation of mRNA splicing has been explored, whereby the PUF domain was used as a scaffold for engineering artificial splicing factors with designed sequence specificities and activities. These engineered splicing factors (ESFs) were constructed from a wild type or modified PUF domain of HsPUM1 fused to a glycine-rich domain of hnRNP A1 (Gly-PUF) or the arginine–serine-rich domain of ASF/SF2 (RS-PUF). The designer ESF was able to shift the splicing of the endogenous pre-mRNA of Bcl-x towards the short Bcl-xS isoform, thereby promoting apoptosis, as well as promote splicing towards the anti-angiogenic isoform b of endogenous VEGFA gene in cultured cancer cells (**Fig. 1.7 d-e**) [75]. PUF-based fusion proteins with RNA cleavage activity have also been created in order to cleave (CUG)n repeats-containing transcripts in patient-derived DM1 (myotonic dystrophy type 1) cells. This recombinant protein could target the nuclear accumulation of the pathogenic RNA, thus rescuing the phenotype (**Fig. 1.7 f**) [76].

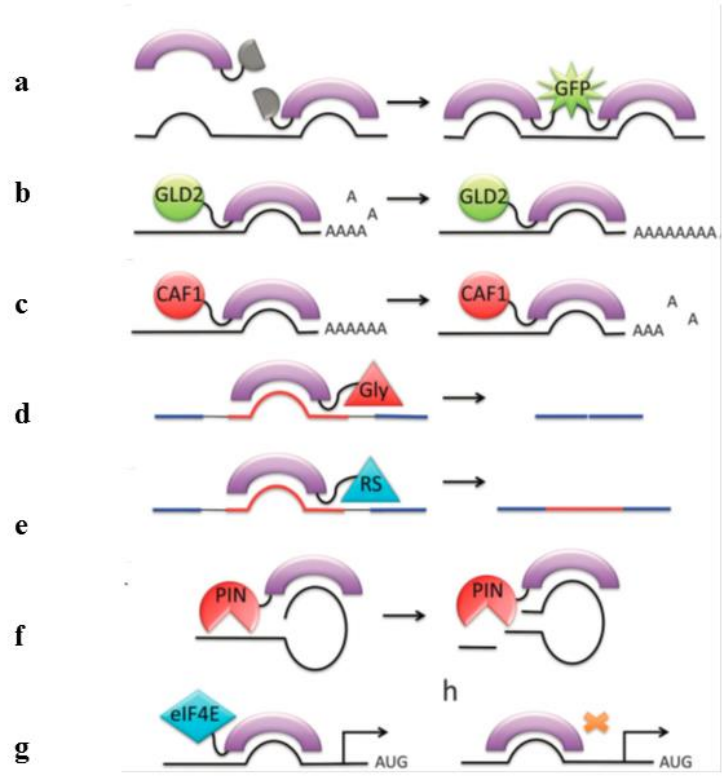


Figure 1.7. PUF applications. a. RNA tracking and localization using two PUF domains fused to a split Venus protein. **b-c.** PUF proteins engineered with a translational activator (GLD2) and a transcriptional repressor (CAF1) that regulate poly(A) tail length. **d-e.** PUF proteins able to influence alternative splicing. **f.** An RNA endonuclease fused to PUF. **g.** PUF designed to bind to the 5' UTR region drives the initiation factor eIF4E on mRNA to activate translation. Taken from *Abil et al* [71].

1.5 Aim of the study

To interfere with REST dysregulation in neuronal pathologies, various molecular strategies have been developed with the purpose to restore the correct levels of expression of REST target genes. However, these approaches suffer from a number of limitations, including their transient activity, and the possibility to generate important side effects when used in long-term applications. In order to recover the expression of REST activity we choose two different approaches. In the first approach we decreased REST activity based on optogenetic tools able to interfere with the transcriptional factor in a spatiotemporal precise manner. We employed the *A. sativa* AsLOV2 domain to engineer two recombinant proteins (AsLOV2-fused peptides) that can (i) prevent the interaction between REST and its cofactor mSin3a, or (ii) inhibit REST binding to target gene promoters decreasing indirectly REST activity in a light-sensitive way.

In the second approach we attempted to increment REST expression, and explored the possibility to target directly its mRNA. We took advantage of the RNA binding proteins PUF, and engineered six PUF constructs designed to bind a specific sequence in the 3' UTR region of murine REST mRNA. Three constructs are designed to recognize a specific eight-ribonucleotide sequence, and the other three are designed to bind to a sixteen-ribonucleotide sequence in the same region of the transcript. We verified the binding and the affinity of the six proteins for their target sequences, and identified two constructs with the highest affinity for endogenous REST mRNA. We confirmed our results through a molecular modelling analysis, in collaboration with Drs. Gatti and Maragliano (IIT-NSYN, Genova), who simulated the binding of our engineered proteins for their target RNA sequences. In addition we checked the function of our PUF constructs, demonstrating that they do not have any intrinsic effect on REST mRNA translational regulation. Thus, the PUF-based constructs described in this work are configured as excellent target-specific platforms. In the future they will be linked to functional proteins capable of modulating the stability of the REST transcript.

2 Materials and methods

2.1 Materials

All biochemical reagents and drugs were from Sigma-Aldrich (Milano, IT) and Promega (Milano, IT) unless otherwise specified. Tissue culture reagents and media were from Gibco (Life Technologies Corp, Monza, Italy) or Sigma-Aldrich. List of antibodies used: rabbit polyclonal anti-GAPDH (#SAB3500247;Sigma-Aldrich), mouse monoclonal anti-His-tag (HIS.H8, #sc-57598; Santa Cruz Biotechnology D.B:A, Milano IT), rabbit polyclonal anti-GFP (#A-11122; Life Technologies), rabbit polyclonal anti-BDNF (H-117) (#sc-20981; Santa Cruz Biotechnology), rabbit monoclonal anti-FLAG (#F7425 Sigma), rabbit polyclonal anti-calnexin (#10286 Abcam, Milano, IT). Horseradish peroxidase (HRP)-conjugated secondary antibodies for western blot analysis were: stabilized goat anti-mouse IgG (H+L), peroxidase conjugated (#32430; Thermo Scientific, Monza, IT) and stabilized goat anti-rabbit IgG (H+L), peroxidase conjugated (#32460; Thermo Scientific).

2.2 Mammalian Cell Culture and Light Stimulation Experiments

Murine Neuro2a (N2a) neuroblastoma cells and human HEK293T cells were cultured in DMEM (#11965-092, Gibco) supplemented with 10% (vol/vol) FBS, glutamine (2 mM), and antibiotics, in a humidified 5% CO2 atmosphere at 37 °C. For light stimulation experiments, 750,000 cells were plated in 35 mm dishes, and the day after were transfected with 2.5 μg of the indicated vectors. Twenty-four hours after transfection, cells were subjected to light stimulation. Stimulation parameters were as follows: 0.34 mW/cm2, 470 nm, and 50% duty cycle (1-s light pulses) for the indicated time. After the stimulation period, RNA was extracted. For the control dark points, cells were maintained in the incubator wrapped in aluminum foil, to ensure complete dark conditions. Expression vectors were transiently transfected into cultured cells using Lipofectamine 2000 (Life Technologies) following standard procedures.

2.3 Protein Extraction and Western Blotting.

Total protein lysates were obtained from cells lysed on ice for 20 min in RIPA buffer [20 mM Tris·HCl, pH 7.4; 1% Triton X-100; 10% (vol/vol) glycerol; 150 mM NaCl; 1% PMSF; protease inhibitor mixture: Complete protease inhibitor mixture tablets; #04693116001; Roche Applied Science (Monza, IT)]. The final protein concentration was quantified by using the BCA protein assay kit (#23225; Pierce Biotech, Monza, IT). Nuclear extracts were prepared as follows: the pellet from 5×10^6 cells was resuspended in 400 μL cytoplasmic buffer (10 mM Hepes, pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT, protease inhibitor mixture) and incubated 20 min on ice with constant shaking. After the addition of 25 μL of 10% (vol/vol) IGEPAL, the solution was briefly vortexed and centrifuged 2 min at 8,000 × g at 4 °C. The supernatant containing the cytoplasmic extracts was collected, the pelleted nuclei were suspended in 50 μL nuclear buffer [20 mM Hepes, pH 7.9, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 25% (vol/vol) glycerol, protease inhibitor mixture] and incubated 20 min on ice with constant shaking. Nuclei were centrifuged 25 min at 10,000 × g at 4 °C, and the supernatant (nuclear extract) was collected. The final protein concentration was quantified by using the Bradford protein assay (#23200; Pierce Biotech.). Proteins were separated using precast 10% NuPAGE Novex Bis-Tris Gels (Life Technologies) and transferred to a nitrocellulose membrane. After incubation with primary antibodies, membranes were incubated with HRP-conjugated secondary antibodies and revealed by autoradiography using the SuperSignal West Pico Chemiluminescent Substrate (#34077; Thermo Scientific).

2.4 Immunofluorescence and Confocal Microscopy

Cells were fixed with 4% (wt/vol) paraformaldehyde and 20% (wt/vol) sucrose in PBS for 15 min at room temperature (RT) and permeabilized with 0.1%. Triton X-100 in PBS for 5 min at RT. Samples were blocked for 30 min in immunofluorescence buffer [2% (wt/vol) BSA, 10% (vol/vol) goat serum in PBS]. Primary and secondary antibodies were diluted in immunofluorescence buffer and incubated for 45 min at RT. Coverslips were mounted using ProLong antifade (#P36931; Life Technologies) and imaged by confocal microscopy. Confocal fluorescent images were obtained using a Leica SP5 confocal microscope with a 40x and 20x objective and analyzed with the Leica LAS AF software.

2.5 RNA Preparation and qRT-PCR

Total cellular RNA was extracted using RNeasy Mini Kit (#74104 Qiagen), and isolated RNA was subjected to DNase I (# 9PIM610, Promega) treatment. cDNA was synthesized starting from 0.5 μg treated RNA according to the High-Capacity cDNA Reverse Transcription Kit manual (#4368814; Applied Biosystems) and used for qRT-PCR. qPCR was performed using the Power SYBR Green PCR Master Mix (#4309155 Applied Biosystems). The primers used in qRT-PCR experiments are the following:

Syn1 Fw 5'-AGCTCAACAAATCCCAGTCTCT-3'

Syn1 Rv 5'-CGGATGGTCTCAGCTTTCAC-3'

BDNF Fw 5'-ATTACCTGGATGCCGCAAA -3'

BDNF Rv 5'-TAATACTGTCACACACGCTCA-3'

BDNF-II Fw 5'-GCCATCCACACGTGACAAAAC-3'

BDNF-II Rv 5'- TGCTGAATGGACTCTGCTCTC-3'

SNAP25 Fw 5'-CCTAGTAGGTCTTGCACATACAC-3'

SNAP25 Rv 5'-GACAGAGCACACAGGACATTT-3'

NAV1.2 Fw 5'-GGCTCTGCTGTCATTGTTGGTA-3'

NAV1.2 Rv 5'-GAAGGCTAGGTGAGTACATCCC-3'

HPRT1 Fw 5'-TCAGTCAACGGGGGACATAAA-3'

HPRT1 Rv 5'-GGGGCTGTACTGCTTAACCAG-3'

GAPDH Fw 5'-AGGTCGGTGTGAACGGATTTG-3'

GAPDH Rv 5'-TGTAGACCATGTAGTTGAGGTCA-3'

RPS9 Fw 5'-CTGGACGAGGGCAAGATGAAGC-3'

RPS9 Rv 5'-TGACGTTGGCGGATGAGCACA-3'

2.6 List of Plasmids (LOV-based probes)

pCS2+MTmSin3AN205 was a kind gift from T. Kouzarides, Gurdon Institute, Cambridge, UK; and pCS2+PRIKLE1 was a kind gift from A. Bassuk, University of Iowa, Ames, IA. pcDNA3.1His-AsLOV2 was cloned as follows: the AsLOV2 sequence was amplified from a bacterial plasmid encoding for AsLOV2 (gift from T. R. Sosnick, University of Chicago, Chicago) and inserted in the pcDNA3.1V5/His vector (Life Technologies) between the HindIII and AgeI sites. pcDNA3.1His/AsLOV2-PAH1 and pcDNA3.1His/AsLOV2-RILPshort were obtained as follows: the PAH1 sequence was amplified from pCS2+MTmSin3AN205 and cloned in pcDNA3.1His-AsLOV2 at the AgeI site; the RILP N313 sequence was amplified from pCS2+PRIKLE1 and cloned in pcDNA3.1-His-AsLOV2 at the AgeI site.

2.7 Lentivirus Production and Infection Procedures

The lentiviral pCCL.sin.cPPT.PGK.GFP.WPRE bidirectional expression vector and packaging plasmids were a kind gift from L. Naldini, TIGET, San Raffaele Sci. Institute, Milan, Italy. For the coordinated expression of AsLOV2-PAH1b and GFP or of scrambled/REST shRNA and mCherry, the low-affinity nerve growth factor (NGF) receptor in the MA1 construct was replaced with REST cDNA, leading to GFP expression from the miniCMV promoter and REST expression from the phosphoglycerate kinase (PGK) promoter. Third-generation lentiviruses were produced by transient four-plasmid co-transfection into HEK293T cells using the calcium phosphate transfection method. Supernatants were collected, passed through a 0.45 μm filter, and purified by ultracentrifugation as previously described. Viral vectors were titrated at concentrations ranging from 1×10^8 to 5×10^9 transducing units (TU)/mL and used at a multiplicity of infection (MOI) of 1–10. The efficiency of infection was estimated to range between 70% and 90% by counting neurons expressing GFP protein with respect to the total number of cells stained with DAPI. Primary cortical neurons were infected at 7 DIV. After 24 h, half of the medium was replaced with fresh medium. Experiments were performed 5–7 d after infection (between 12 and 14 DIV).

2.8 Electrophysiological Recordings

Primary cortical cultures were prepared from mouse C57BL/6J (E17–E18) embryos. All experiments were carried out in accordance with the guidelines established by the European Communities Council (Directive 2010/63/EU of 22 September 2010) and were approved by the Italian Ministry of Health. Lentivirus production and infection were performed following standard

procedures. Neurons were infected with lentiviral vectors encoding GFP, AsLOV2-PAH1b/GFP, scrambled/mCherry or REST shRNA/mCherry and subjected to 24 h light stimulation or kept in the dark. One day after the end of the stimulation protocol, excitability and Na^+ current density were studied by patch-clamp recordings in current-clamp and voltage-clamp configurations. All experiments were performed on transduced neurons identified by green or red fluorescence using a Multiclamp 700B amplifier (Axon Instruments, Molecular Devices) and an upright bx51WI microscope (Olympus) equipped with Nomarski optics. Current clamp recordings were performed at a holding potential of −70 mV, and action potential firing was induced by injecting current steps of 25 pA lasting 500 ms. Sodium currents were obtained by stepping from a holding potential of −90 to −70 to +50 mV with 10 mV steps.

2.9 Photostimulation Hardware and arduino sketch

In vitro light stimulation experiments were performed using a custom built LED photostimulation device, which was fabricated with a precision-machined aluminum mounting plate and a 5 W high efficiency Blue LED array, with peak emission at 470 nm. LED intensity was regulated by a separate, 12 channel low-noise, linear power driver; on off operation (and duty cycle) was modulated by an ARDUINO 2000 microcontroller (the system can accept every TTL or digital positive signal to enable channels output or can be used in stand-alone mode). Radiation output was measured from a distance of 1.5 cm above the array using a Thorlabs PM100 power meter and an S121b circular sensor probe.

ARDUINO sketch.

```
// Assign channels to pins on the Arduino board
const int pin1 = 2;
const int pin2 = 4;
const int pin3 = 6;
const int pin4 = 8;
const int pin5 = 10;
const int pin6 = 12;

// Define initial output states of pins.  For LED illumination, “LOW” means the LED is off, and “HIGH” //means it’s on
int ledState1 = LOW;
int ledState2 = LOW;
int ledState3 = LOW;
```

```
int ledState4 = LOW;
int ledState5 = LOW;
int ledState6 = LOW;

// Set reference timers for each channel to zero
long prevMillis1 = 0;
long prevMillis2 = 0;
long prevMillis3 = 0;
long prevMillis4 = 0;
long prevMillis5 = 0;
long prevMillis6 = 0;

// Define pulse intervals for each channel
long interval1 = 2000;
long interval2 = 2000;
long interval3 = 2000;
long interval4 = 2000;
long interval5 = 2000;
long interval6 = 2000;

// Define pulsewidth for each channel
long pulse1 = 1000;
long pulse2 = 1000;
long pulse3 = 1000;
long pulse4 = 1000;
long pulse5 = 1000;
long pulse6 = 1000;

void setup(){
  // Define each pin as an output pin
  pinMode(pin1,OUTPUT);
  pinMode(pin2,OUTPUT);
  pinMode(pin3,OUTPUT);
  pinMode(pin4,OUTPUT);
```

```
  pinMode(pin5,OUTPUT);
  pinMode(pin6,OUTPUT);

}

void loop(){
  // Assign  current time to time variables for each channel
  unsigned long currentMillis1 = millis();
  unsigned long currentMillis2 = millis();
  unsigned long currentMillis3 = millis();
  unsigned long currentMillis4 = millis();
  unsigned long currentMillis5 = millis();
  unsigned long currentMillis6 = millis();
```

/* Code to determine on or off state of channel 1. If the current time is greater than the previously updated reference time by more than the defined interval, the LED turns on ("HIGH") and the reference time gets updated to the present time (the time at which the LED turned on). If the LED had previously been triggered and the pulsewidth time has passed, the LED will turn off ("LOW"). The same script is used below to control each individual channel*/

```
 if (currentMillis1-prevMillis1 > interval1)
   {

    if (ledState1 ==LOW){
      ledState1 = HIGH;
      prevMillis1 = currentMillis1;
    }
   }
    else {
      if (currentMillis1-prevMillis1 > pulse1)
      ledState1 = LOW;
    }
```

```
// Control of channel 2
if (currentMillis2-prevMillis2 > interval2)
  {

   if (ledState2 ==LOW){
     ledState2 = HIGH;
     prevMillis2 = currentMillis2;
   }
  }
   else {
     if (currentMillis2-prevMillis2 > pulse2)
     ledState2 = LOW;
   }


 // Control of channel 3
 if (currentMillis3-prevMillis3 > interval3)
  {

   if (ledState3 ==LOW){
     ledState3 = HIGH;
     prevMillis3 = currentMillis3;
   }
  }
   else {
     if (currentMillis3-prevMillis3 > pulse3)
     ledState3 = LOW;
   }


 // Control of channel 4
 if (currentMillis4-prevMillis4 > interval4)
  {
```

```
  if (ledState4 ==LOW){
    ledState4 = HIGH;
    prevMillis4 = currentMillis4;
  }
 }
  else {
    if (currentMillis4-prevMillis4 > pulse4)
    ledState4 = LOW;
  }


// Control of channel 5
if (currentMillis5-prevMillis5 > interval5)
 {

  if (ledState5 ==LOW){
    ledState5 = HIGH;
    prevMillis5 = currentMillis5;
  }
 }
  else {
    if (currentMillis5-prevMillis5 > pulse5)
    ledState5 = LOW;
  }


// Control of channel 6
if (currentMillis6-prevMillis6 > interval6)
 {

  if (ledState6 ==LOW){
    ledState6 = HIGH;
    prevMillis6 = currentMillis6;
```

```
  }
 }
  else {
   if (currentMillis6-prevMillis6 > pulse6)
   ledState6 = LOW;
  }


 // Output the state of each LED to the pins to physically regulate the LEDs
 digitalWrite(pin1,ledState1);
 digitalWrite(pin2,ledState2);
 digitalWrite(pin3,ledState3);
 digitalWrite(pin4,ledState4);
 digitalWrite(pin5,ledState5);
 digitalWrite(pin6,ledState6);

}
```

2.10 PUF vectors and cloning strategy

PUF8wt and PUF16wt were kindly provided by Dr. Rackham (The University of Western Australia, Perth, Australia). PUFrest8, PUFrest8STACK(Y), PUFrest8STACK(H) and PUF16rest have been obtained starting from PCR amplification of PUF8wt and PUF16wt using Pfu DNA polymerase (#M7745; Promega) with the following mutagenesis primers:

Fw PUF8rest-ns (SE-NQ) 5’-CAAATTTGCAAatAATGTTGTGcAGAAGTGTGTTACTC-3’

Rv PUF8rest-ns (SE-NQ) 5’-GAGTAACACACTTCTgCACAACATTatTTGCAAATTTG-3’

Fw PUF8rest-s 5’- GCATAAATTTGCCAATtACGTGGTTCAAAAATGTG-3’

Rv PUF8rest-s 5’-CACATTTTTGAACCACGTAATTGGCAAATTTATGC-3’

Fw PUF8rest-sH 5’- GCATAAATTTGCCAATCACGTGGTTCAAAAATGTG-3’

Rv PUF8rest-sH 5’-CACATTTTTGAACCACGTgATTGGCAAATTTATGC-3’

For PUF16rest-ns:

Fw Rep3(C-N) 5’- GGCACTGCAAATGTATGGTAATCGTGTTATTCAGAAAGCCCTGG-3’

Rv Rep3(C-N) 5’- CCAGGGCTTTCTGAATAACACGATTACCATACATTTGCAGTGCC-3’

Fw Rep4(N-C) 5’-GTGTGAAAGATCAGAATGGCTGTCATGTTGTGCAGAAATG-3’

Rv Rep4(N-C) 5’-CATTTCTGCACAACATGACAGCCATTCTGATCTTTCACAC-3’

Fw Rep14(NQ-SR) 5’-GTATGGAAGCTATGTGATTCGTCATGTTCTGGAACATG-3’

Rv Rep14(NQ-SR) 5’-CATGTTCCAGAACATGACGAATCACATAGCTTCCATAC-3’

Fw Rep12(SE-NQ) 5’-CAAATTTGCAAatAATGTTGTGcAGAAGTGTGTTACTC-3’

Rv Rep12(SE-NQ) 5’-GAGTAACACACTTCTgCACAACATTatTTGCAAATTTG-3’

PCR conditions were as follows: 95 °C, 5 min; 95 °C, 30 s; 55 °C, 30 s; 72 °C, 17, min; 30 cycles; 72 °C 5 min. PCR products were digested using the DpnI enzyme (Promega) and transformed into DH5α cells. Positive colonies were verified by DNA sequencing. To avoid recombination events in PUF16rest-ns the mutations in REP3-4-6 were performed on PUF8wt while mutation in REP7 on PUF16wt. Mutated PUF16wt and PUF8wt were then digested with the SacI enzyme (Promega) and ligated with T4 DNA Ligase (Promega). PUF16rest-s and PUF16rest-2.0 were produced by BIOMATIK CORPORATION (Cambridge, ON, Canada). Mutated constructs were then amplified using Pfu DNA polymerase (#M7745, Promega) with the following primers, flanked by restriction sites for NotI and BamHI:

Fw PUF8NotI 5’-GCATAAATTTGCCAATAACGTGGTTCAAAAATGTG-3’

Rv PUF8BamHI 5’-CACATTTTTGAACCACGTTATTGGCAAATTTATGC-3’

Fw:PUF16NotI 5’-CATAGCGGCCGCACCATGGGTCGTAGCCGTCTG-3’

Rv:PUF16BamHI 5’-CATAGGATCCGCCCAGGTCCACGCCATTTTTC-3’

and then cloned in the NotI /BamHI (Promega) digested CMV_3Xflag vector, kindly provided by Dr D. di Bernardo (Telethon Institute of Genetics and Medicine, TIGEM, Naples). Reporter plasmids for luciferase assays have been produced starting from annealed oligos: 1µg sense and 1 µg antisense oligos were resuspended in the following solution: 10 mM Tris HCl (pH7.5), 0.1M

NaCl, 1mM EDTA, and incubated in a (Biorad T100 Thermocycler) as follows: 95°C 4min, 70°C 10min, decrease temperature to 4° (0.1°C/min).

RESTRNA8

Sense 5'-TCGAGTTTATATAGC-3'

Antisense 5'-CGCCGCTATATAAA3-'

NRE

Sense 5'-TCGAGTGTATATAGC-3'

Antisense 5'-CGCCGCTATATACAC-3'

RESTRNA16

Sense 5'-TCGAGTGCTTTATATAAATTAGC-3'

Antisense 5'-CGCCGTAATTTATATAAAGCAC-3'

2XNRE

Sense 5'-TCGAGTGTTGTATATAATATAGC-3'

Antisense 5'-CGCCGTATATTATATACAACAC-3'

Annealed oligos were inserted via standard ligation procedures in the digested XhoI /NotI psiCHECK Vector 2.0, provided by Dr A. Contestabile (Istituto Italiano di Tecnologia, IIT Genoa).

2.11 Purification of Flag-tagged Proteins

Cytosolic protein lysate (300 µg) was purified with 30 µl of Anti-FLAG® M2 Magnetic Beads (Sigma-Aldrich) pre washed with PBS/tween 0.1%, and incubated in the following solution: 50 mM Tris HCl,150 mM NaCl, pH 7.4; in a total volume of 200 µl for 2 h RT. After incubation, beads were washed five times with PBS/tween 0,1%, and bound proteins were eluted by incubation with 3X FLAG peptides (# F4799, Sigma) (150 ng/µl final concentration) for 45 min at 4 °C. 6% of the total volume elution was resolved on a denaturing 10% polyacrylamide gel and stained with Pierce™ Silver Stain Kit (#24612 Thermo Fisher Scientific) for purity rate verification. The final protein concentration was quantified by using the BCA protein assay kit (#23225; Pierce Biotech.).

2.12 RNA Electrophoretic mobility shift assay (EMSA)

3'biotinylated RNA sequences **(NRE**: CCUGUAUAUAAGU; **RESTRNA8**: CCUUUAUAUAAGU; **2XNRE**: CCUGUUGUAUAUAAUAU**A**AGU; **RESTRNA16**: CCUGCUUUAUAUAAAUUAAGU; **RESTRNA16-2.0** CCAUUGGCUUAGUAAAUUAGU)

were chemically synthesized (Sigma Aldrich). To perform EMSAs, the LightShift® Chemiluminescent RNA EMSA Kit (#20158, Thermo Fisher Scientific) was used. For each gel-shift reaction, a total of 10 nM biotin-labeled probe was dissolved in the Binding Buffer provided, supplemented with glycerol 50%, 2 µg tRNA, BSA 4 µg, KCl 50mM, DTT 2 mM, Tween 0.02%, EDTA 1 mM, and incubated with 1 µM -10 µM of purified proteins. The reaction mixture was pre-incubated 40 min at RT and resolved on a non-denaturing 10% polyacrylamide gel 0.5X TBE (1h 4°C) that had been pre-electrophoresed for 60 min. RNA-protein complexes were transferred to nylon membranes and then crosslinked for 13 min using a commercial UV-light crosslinking instrument (254 nm bulbs). To obtain titration curves, known concentrations of biotin-labeled RNA (2.5-5-10-20-40-80-160 pg) were used to create a standard curve from which the complexed RNA concentration was interpolated. Data were then fitted using nonlinear regression analysis with the following formula Y=Bmax*X^h/(Kd^h + X^h), thus obtaining Kd values.

2.13 Cross-linking RNA Immunoprecipitation (CLIP)

Transfected N2a cells (1x10^6) were subjected to UV crosslink at 4000x100 µJ/cm^2. Cells were subsequently harvested in buffer A [20 mM Tris HCl (pH 8), 10 mM NaCl, 3mM $MgCl_2$, 0.1% NP40, 10% Glycerol, 0.2 mM EDTA, 1 mM DTT, protease inhibitor mixture tablets (#04693116001, Roche), RNase inhibitor RiboLock (#EO0381 Thermo Scientific)] and centrifuged at 2500 x g at 4°C for 5 min. Agarose beads protein A salmon sperm (#16-157. Millipore) were pre-cleaned with CLIP Buffer (1XPBS, 0.1% SDS 0.5% NP-40) and incubated for 1 h with the lysates plus RIPA buffer (10 mM Tris·HCl, pH 7.5, 1 mM EDTA, 140 mM NaCl, 0.5 mM EGTA, 1% TritonX-100, 0.1% SDS, 0.1% Na-deoxycolate) to a final volume of 1mL. After centrifugation at 2500 x g for 5 min, the supernatant was divided in two equal parts, to incubate with anti-flag Rabbit (#F7425 SIGMA) (5 µg) and anti-rabbit IgG antibodies (PP54B; Millipore) (2.5 µg) at 4°C overnight, (keeping 50 µl for Input RNA Real time detection and 20µl for Input protein WB detection). Pre-cleaned beads were incubated with the antibody-protein-RNA complex at 4 °C for 2 h. Beads were washed with CLIP Buffer and High Salt wash buffer [5XPBS without Mg^{2+}/Ca^{2+}(#D1408, Sigma-Aldrich), 0.1% SDS, 0.5% NP-40], and resuspended in RIPA buffer. Proteinase K (#P6556, Sigma-Aldrich) was added to the samples and incubated 45 min at 45 °C before proceeding to RNA extraction by using the miRNeasy kit (#217004, Qiagen). 50% of the total RNA extracted volume was used for cDNA synthesis and qRT-PCR. List of primers used in qRT-PCR:

FwREST3’UTR: 5’-GCATAAATCTTAGCAAATCCTCGGGAG-3’

RvREST3’UTR: 5’-GGCAGACAAGGCAAGTGGTGTG-3’

FwGapdh3’UTR: 5’-GAAACCCTGGACCACCCAC-3’

RvGapdh3’UTR: 5’-GTGGGTGCAGCGAACTTTATTG-3’

FwRESTCDS: 5’-TTCACATTTATACGGGCGTTC-3’

RvRESTCDS: 5’-CCTGCAGCAAGTGCAACTAC-3’

FwGapdhCDS: 5’-AGGTCGGTGTGAACGGATTTG-3’

RvGapdhCDS: 5’-TGTAGACCATGTAGTTGAGGTCA. -3’

2.14 Molecular Dynamics Simulations

The starting conformation of the PUF8wt system was obtained from the PDB structure (3Q0P) and not from PDB (1M8Y), (same amino acid sequence) since this structure was more correct for simulations. PUF16wt was generated by hand as described in par. 3.2.2. All required protein and RNA mutations were performed using the VMD program by deleting the original sidechain atoms and letting VMD generate the positions of the new atoms. When overlaps were observed, the structure was optimized by hand based on local alignment with similar crystal structures. All systems were solvated in explicit water using the TIP3P water model, and the total charge was neutralized by adding Na+ and Cl–ions at physiological concentration. The total number of atoms, including water and ions, was 130,000 for PUF8wt and 290,000 for PUF16wt. Periodic Boundary Conditions were used to avoid finite-size effects, and long range electrostatic interactions were treated with the Particle Mesh Ewald method. For each system, the internal energy was minimized via 1000 steps of steepest descent and then 2 ns of equilibration were performed at constant pressure and temperature (NPT ensemble) with pressure 1 atm and temperature 300K. A timestep of 2 fs was used. Finally, production runs were carried out in the constant volume and temperature NVT ensemble, for a total of 200 ns per system. All simulations were performed with the NAMD program and using the Amber force field. All analyses were carried out with VMD, and distances and energy contributions were calculated using home-written Tcl scripts.

2.15 Statistical Analysis

Results are expressed as means ± S.E.M. throughout. Data were analyzed by either the unpaired Student's t test or one-way ANOVA followed by Tukey's multiple comparison test.

3 Results

3.1 Transcriptional modulation of REST activity

3.1.1 Design of AsLOV2-based strategies to inhibit REST activity.

To modulate REST activity, we engineered chimeric proteins formed by the light-sensitive AsLOV2 protein fused to two REST-interfering domains. AsLOV2 would sterically block the effector domain fused to its C terminus in the dark, whereas the unfolding of the Jα-helix upon illumination would cause the reversible exposure of the effector sequence (**Fig. 3.1a**). To select REST-interfering molecules, we focused on domains able to alter two key events in REST activity, namely the recruitment of the cofactor mSin3a and the REST-RE1 site interaction. To impair mSin3a recruitment, we used the N-terminal portion of mSin3a (mSin3a-N205), which was identified as the minimal REST binding region of mSin3a [77] (AsLOV2-PAH1). To interfere with RE1 recognition, we chose the REST-interacting LIM domain protein (RILP), an endogenous REST interactor that, by binding to the Zn-finger domains of the repressor, displaces it from the target chromatin [78]. We used the first N-terminal portion of RILP (RILP-N313), which contains the minimal REST binding domain (AsLOV2-RILP N313) (**Fig. 3.1b**).

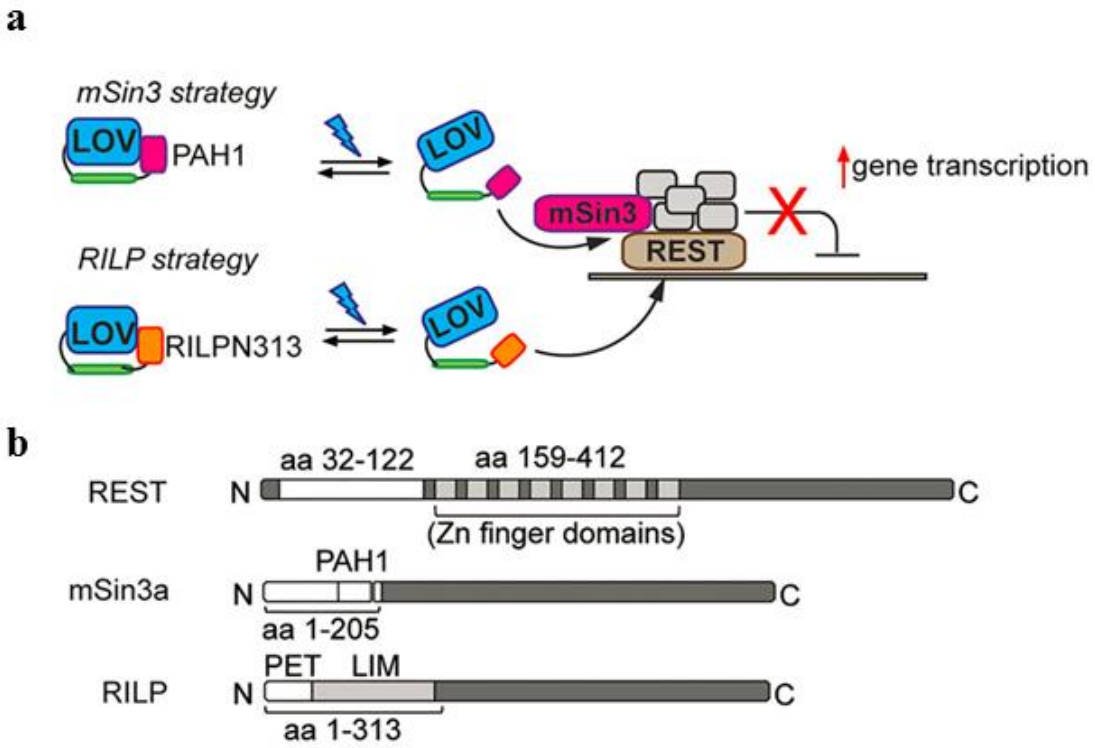


Figure 3.1 Strategies for REST inhibition. a Schematic cartoon of the inhibition strategy. mSin3a (Upper) or RILP (Lower) interfering domains are directly fused to AsLOV2. In the dark, AsLOV2 is in a closed conformation, masking the REST-binding sites. Upon blue light illumination (470 nm), AsLOV2 unfolds thus freeing the C-terminal domains to interact with REST, displacing the endogenous mSin3a or the entire REST complex from target DNA. This would result in the increased transcription of REST-target genes. **b** Schematic representation of REST, mSin3a, and RILP protein sequences. The interacting portions between REST and mSin3a/RILP are highlighted. PAH1, paired amphipathic helix 1; PET, Prickle Espinas Testin; LIM, Lin11, Isl-1, andMec-3.

3.1.2 Light-Driven Modulation of REST Activity

Once the expression and the REST interfering activity of the two chimeras (AsLOV2-PAH1 and AsLOV2-RILP N313) had been validated, we evaluated the activity of the AsLOV2 constructs in response to specific patterns of illumination. N2a cells were transfected with AsLOV2-PAH1 and AsLOV2-RILP N313 and after 24 h, cells were transferred to the illumination apparatus and subjected to the light stimulation protocol (470-nm light illumination 0.5 Hz) (**Fig. 3.2a**). A significant increase in the transcription of the REST targets BDNF, BDNFII, SYN1, SNAP25, and NAV1.2 mRNAs was observed upon illumination in cells expressing either chimera. No increase in gene transcription was observed in parallel samples kept in the dark or in control cells transfected with AsLOV2 alone and exposed to the same illumination pattern (**Fig. 3.2a**). Next, we wanted to define the time course of the light-induced transcriptional changes from the beginning of the light stimulation protocol up to a maximum time of 48 h, using BDNF as a target gene. We found that, with both AsLOV2 chimeras, BDNF expression levels significantly increased already after 12 h of illumination, reached a plateau of ∼1.5- to 2-fold increase after 24 h and returned to baseline levels in ∼12 h after switching off the light (**Fig. 3.2b**). To evaluate whether the observed changes in mRNA levels were followed by a parallel change in protein levels, BDNF protein was measured in cell extracts by western blotting after 48 h of light stimulation. As expected, a significant increase in BDNF protein was observed only in N2a cells expressing the active probes and exposed to light (**Fig. 3.2c**). Taken together, these data demonstrate that both AsLOV2-PAH1 and AsLOV2-RILP N313 effectively act as light modulated inhibitors of REST activity on target genes, resulting in increased levels of transcription and translation.

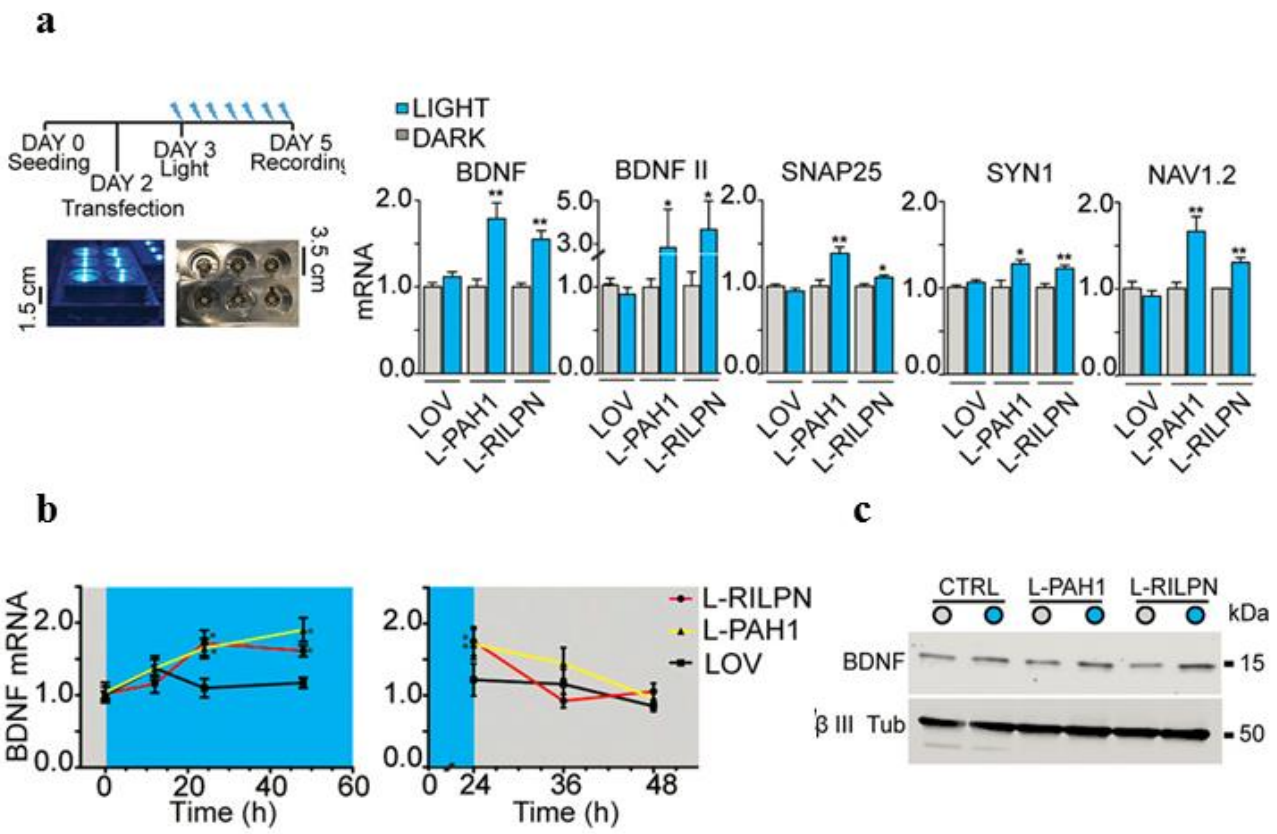


Figure 3.2 Photostimulation of AsLOV2-PAH1 and AsLOV2-RILP N313 constructs. a Illustration of the illumination protocol and LED stimulation apparatus. Undifferentiated N2a cells were transfected with the indicated constructs. Cells were subjected to 470-nm light illumination (0.5 Hz) or kept in the dark, as indicated. After 48 h, the mRNA levels of the indicated genes were quantified via qRT-PCR. GAPDH and HPRT1 were used as control housekeeping genes (*P < 0.05; **P < 0.01; Student t test; n = 3 independent experiments). **b** Time-course of light-dependent BDNF up-regulation and post illumination recovery (*P < 0.05; Student t test vs. respective dark condition; n = 3 independent experiments). **c** Western blot analysis of BDNF in undifferentiated N2a cells transfected with the indicated constructs.

3.1.3 Transduction of AsLOV2-PAH1 in Primary Neurons Increases Intrinsic Excitability.

To study the effects of the light-driven modulation of REST in primary neurons, we engineered a bidirectional lentiviral vector expressing AsLOV2-PAH1 and EGFP. Once transduced into primary neurons, the probe showed a clear enrichment in the nuclear compartment as we expected (**Fig. 3.3a**). Primary cortical neurons were infected at 7 d in vitro (DIV) and then subjected to 470-nm light stimulation at 12 DIV for 24 h. Starting from the illumination parameters used for N2a cells, we optimized our protocol for primary neurons, by changing both the illumination intensity and the duty cycle. We found that the optimal conditions for primary neurons, which maximized cell viability while inducing appreciable changes in gene transcription, were 0.34 mW/cm2; 250-ms pulses; 0.25 Hz. The analysis of gene transcription reported a significant increase in both NAV1.2 and BDNF mRNAs in AsLOV2-PAH1 infected neurons exposed to light compared with the same neurons kept in the dark or to neurons infected with a control vector expressing only EGFP (**Fig. 3.3b**), indicating that the lentiviral probe was effective in inhibiting REST activity in primary neurons. Next, we analyzed whether the optogenetic suppression of REST activity was associated with changes in intrinsic neural excitability. Transduced neurons, held at a membrane potential of –70 mV in whole-cell current-clamp mode were subjected to current injection of 500-ms duration and stepwise increasing amplitude. Interestingly, neurons expressing AsLOV2-PAH1 and subjected to light stimulation displayed a strongly increased firing frequency compared with the same transduced neurons kept in the dark or to parallel cultures infected with the control vector (**Fig. 3.3c**). Interestingly, the increase in the firing frequency obtained by optogenetic inhibition of REST activity with either probe was fully comparable with that obtained by silencing REST with a lentiviral vector expressing a REST shRNA (**Fig. 3.3d**). These observations demonstrate that the AsLOV2-PAH1 probe is able to influence REST activity in primary neurons by inducing light-evoked transcriptional regulation with functional consequences on neuronal activity.

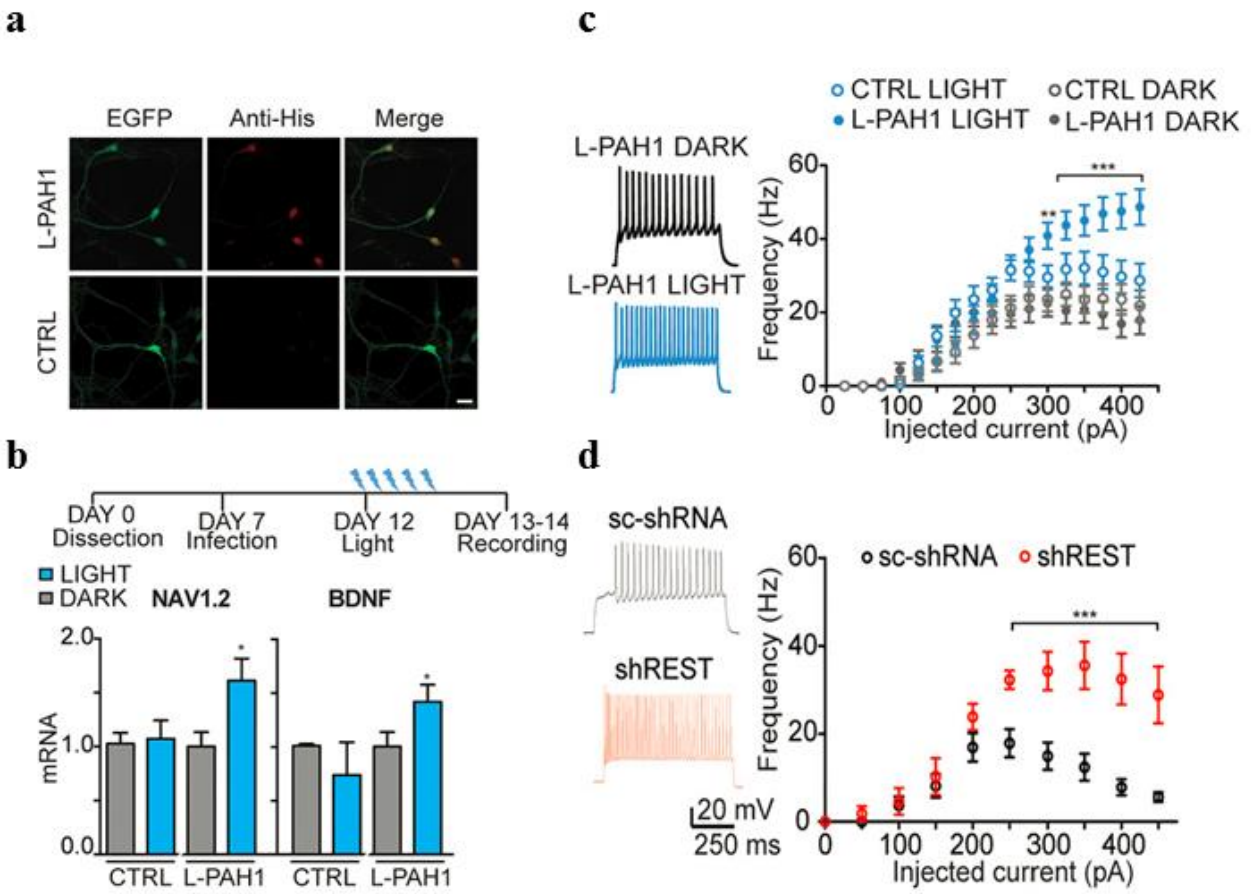


Figure 3.3 AsLOV2-PAH1 inhibits REST activity in primary neurons. a Confocal images of primary cortical neurons infected with a bidirectional lentiviral vector expressing either EGFP (CTRL) or AsLOV2-PAH1/EGFP. Neurons were stained using anti-Histidine tag antibodies (red) to detect the AsLOV2 construct, and anti-EGFP (green) antibodies to detect EGFP Scale bar, 10 μm.**b** (Upper) Schematic representation of the time course of the experiment: neurons were infected at day 7, the illumination protocol was started 5 d after infection (day 12) for 24 h. (Lower) qRT-PCR for NAV1.2 and BDNF mRNA in EGFP (CTRL) or AsLOV2-PAH1–expressing neurons kept under dark (gray bars) or lit (blue bars) conditions for 24 h. GAPDH and HPRT1 were used as control housekeeping genes (*P < 0.05; one-way ANOVA followed by the Tukey's multiple comparison test vs. the respective dark control; n = 3 independent experiments).**c.** Plot of the mean (±SEM) firing frequency vs. injected current (F/I curve) in cortical neurons infected with either AsLOV2-PAH1 (closed symbols) or EGFP alone (CTRL, open symbols) and subjected to either 24 h light (blue symbols) or 24 h dark (gray symbols) (AsLOV2-PAH1/dark, n = 22; AsLOV2-PAH1/light, n = 21; EGFP/dark, n = 19; EGFP/light, n = 20; from 3 independent experiments). Representative current-clamp recordings of spike trains evoked by injection of 225 pA for 500 ms in neurons infected with AsLOV2-PAH1 and kept for 24 h under dark (black) or lit (blue) conditions are shown on the left. **d** The same experiment was performed in neurons infected with either a scrambled shRNA or shRNA against REST (n= 11 for each condition).

.

3.2 Targeting REST at the transcriptional level

3.2.1 Design of synthetic PUF proteins specific for REST mRNA

The above-described results show that we were able to modulate REST activity by using two light-sensitive chimeras that interfered with the interaction between REST and its cofactor mSin3a, or inhibited REST binding to target gene promoters, thus decreasing indirectly REST activity in a light-sensitive way. In order to increase REST expression, we decided to modulate directly REST mRNA by engineering a REST RNA-specific binding protein. To this aim, we used the PUF protein PUMILIO-Homology Domain from Human Pumilio1 (HsPUM-HD), whose amino acidic sequence and secondary structure have been well characterized, since the crystallographic structure of the protein in complex with the RNA sequence is availabile (PDB code 1M8Y). The chosen PUF recognizes an eight ribonucleotide RNA sequence (5'**UGUAUAUA**3') called Nanos Response Element (*NRE*)[79]. PUF binds every RNA base through a single repeat for a total of eight repetitions. In this work, the wild type PUF protein is referred to as PUF8wt. We analyzed the 3'UTR of mouse REST mRNA (713 bp) to select a PUF target sequence that minimized the number of mutations required to achieve specific binding, while conserving the original folding (**Fig. 3.4a**). In addition we selected a sequence that is not overlapping with, or in close proximity to, miRNA sites. Eventually, we identified a sequence of eight ribonucleotides that diverges from the original PUF target for only one base, i.e. 5'**UUUAUAUA**3', which we call *RESTRNA8*. Naturally occurring PUF proteins typically contain eight RNA-binding repeats that allow them to regulate specific developmental processes, often by binding multiple RNAs [80]. With the aim of minimizing the off-targeting and increasing the sequence specificity of our REST-specific construct, we took advantage of a PUF protein composed of 16 RNA-binding repeats, referred to as PUF16wt, kindly provided by Dr. Rackham (The University of Western Australia, Perth, Australia). This extended PUF is composed of two HsPUM-HD, and binds to a double *NRE* sequence (5'**UGUUGUAUAUAAUAUA**3', *2XNRE*)[62].

The REST-specific PUF16 protein target was built on a sequence spanning *RESTRNA8*, three additional ribonucleotides at 5' and other five ribonucleotides at the 3', obtaining a 16 bp sequence (5'**UGCUUUAUAUAAAUUA**3', *RESTRNA8* sequence in red, additional nucleotides in black) that differs from *2XNRE* in four base pairs. In addition, to evaluate the feasibility of engineering PUF proteins for any RNA sequence irrespective of the similarity to the original *NRE* site, we picked also a 16 bp sequence in REST 3'UTR, 5'**AUUGGCUUAGUAAAUU**3' that we

defined *RESTRNA16-2.0*, characterized by minimal cross-talk across all mouse mRNAs (**Fig. 3.4a**). Since there are no structural data about the extended PUF domain, we coupled biological experiments with computational molecular modeling in order to support our future experiments. Computational analyses over both PUF8 and PUF16 systems (detailed in the next sections) were performed in collaboration with Drs. M. Gatti and L. Maragliano, (IIT-NSYN, Genova) by engineering a 3D model of PUF16wt together with the PUF8wt model, whose crystallographic structure is available (PDB code 1M8Y) (**Fig 3.4b**).

a

REST mRNA 3'UTR

```
  1  CUGAGCCUCGGCAGAAGCACCGUGCAGACUUUGUGAGCAUGCAAUUUUAA  50
 51  UUUGUAGACAAACGCAAGCUUGCUUUAAUUAGUCUCCAAGGCUGAGUUUU 100
101  CAGUAACAUUCUUUUUCUUAGGACUGUACAUCUAUUUAGUGUUUGUUGCA 150
151  UAAAUCUUAGCAAAUCCUCGGGAGUUAAUGUAAGAGGACAGAUAUGUAAC 200
201  UAGCUCGUGCAGGCAGGUGCAAGGAGAAGGGUAAGAUGGUGGAACACACC 250
251  ACUUGCCUUGUCUGCCUACAACCUGUUGGGUUUUCUUUUCACGGUAGUUC 300
301  CUAAUUUUUAGUUACUUGUUUAGAUCGAUAAAAAUUGGCUUAGUAAAUUA 350
351  CUUGAAGAAUUUGCCUGCUUUAUAUAAAUUAAGUUAGCACUUUACAGUUU 400
401  CUUUAGAGAUGAAAAAAAAGAGAUUUUAAUUGGAGAGAAAUUCUCAACAU 450
451  UGGACAUUGUAUCUGUCCAGGUAAUUGCUUCCUAACUUGCUAUCAAUAUU 500
501  UUGUGUUUAUAUGUUAAUCGUUAUAAAAAGUGAUUUUUGUUUUUUGGGUA 550
551  UUUUUUAUUUUGGUGCUUUUCUGGCUUAAGAUGUUGCACAUGGUUCUUGU 600
601  UUUUGUUUCUUUAACCUAUGCAGUUAAUCUCCCUUCCCCCUGAAACAGCGU 650
651  UGUGUUAAAUAGUAACACUAUACAGAUAUAUGCAUGGUUUUUUUUUUUGU 700
701  UUGUUUGUUUGUU                                        713
```

b

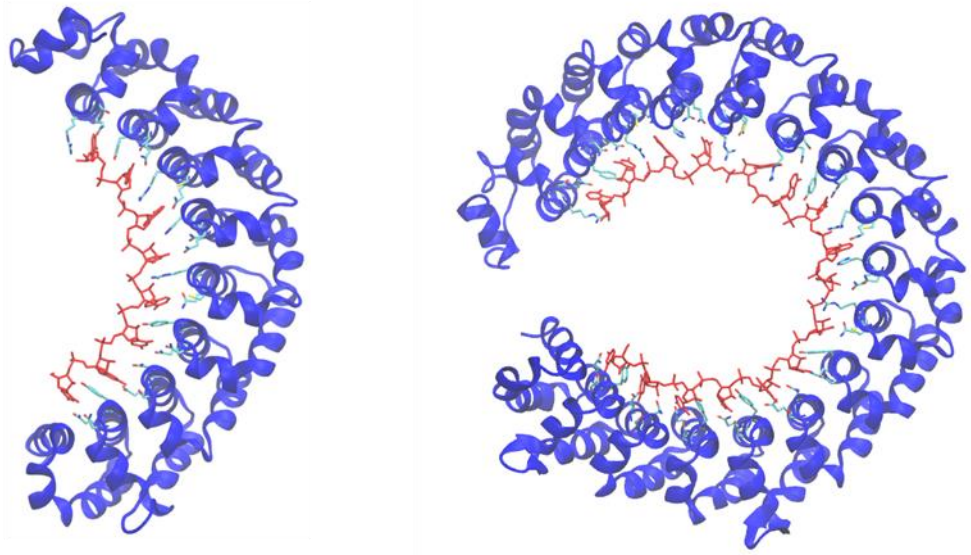


Figure 3.4 Sequence of REST mRNA 3'UTR and 3D model of PUF8wt and PUF16wt. a. REST 3'UTR RNA sequence (713bp, NCBI #ENSMUSG00000029249). *RESTRNA8* is highlighted in red, *RESTRNA16*, spanning *RESTRNA8* and flanking ribonucleotides, is marked by the black rectangle. The sequence of *RESTRNA16-2.0* is in blue. **b.** *Left*: PUF8wt crystal structure bound to the *NRE* sequence (PDB code 1M8Y). *Right*: 3D structural model of the PUF16wt, built from 2 PUF8wt domains, in complex with *2XNRE*. The protein backbone is represented in blue and the ribonucleotide sequence in red; protein residues involved in the recognition code are highlighted in cyan (kindly provided by Drs M. Gatti and L. Maragliano, IIT-NSYN, Genova).

We then mutagenized PUF8wt and PUF16wt through site-directed mutagenesis to obtain proteins able to bind the REST selected sequences. According to *Cheong et al* [81], a single PUF repeat (R) binds its specific ribonucleotide using two amino acid side chains, at the position 12 and 16 (the first and last residues in the five-residue code shown in Fig. 1.6a), which make hydrogen bonds or van der Waals contacts with the Watson–Crick edge of the RNA base; a third amino acid side chain, at position 13 (the second residue in the five-residue code shown in Fig. 1.6a), makes stacking interactions with the aromatic rings of the RNA base. In many studies, artificial PUFs are designed changing only the hydrogen-bonding amino acid residues, which is sufficient to achieve specificity for a new sequence [82, 83]. However the residue mediating the stacking interactions with the aromatic rings can also affect the affinity for the target RNA [84]. In order to elucidate the importance of the stacking residues in the PUF/RNA recognition mechanism, we engineered two variants of synthetic REST-PUF proteins, one mutated in hydrogen bonds or van der Waals residues, which we referred as "ns" (not stacking) and another one mutated in hydrogen-bonding and stacking amino acid residues referred as "s" (stacking). Therefore, based on the RNA-recognition code of PUF proteins (**Fig. 1.6**), we replaced the serine and the glutamic acid of PUF8wt in repeat R7 with an asparagine and a glutamine at position 12 and 16 (SE to NQ), obtaining PUF8rest-ns (**Fig. 3.5a**). We then replaced the asparagine at position 13 with a tyrosine (N-Y) obtaining PUF8rest-s, and according to the expanded recognition code in artificial PUF scaffolds (5) we also replaced the same asparagine with a histidine (N-H), which is the second most frequent amino acid after tyrosine in repeat R7, thus obtaining PUF8rest-sH (**Fig. 3.5a**).

Following the same approach used for PUF8wt, we mutated the hydrogen-bonding amino acid residues at position 12 and 16 of repeats R14, R12, R4 and R3 of PUF16wt, obtaining PUF16rest-ns (**Fig. 3.5b**). Specifically, in repeat R14 we replaced asparagine and glutamine with serine and arginine (NQ to SR); in repeat R12 serine and glutamic acid with asparagine and glutamine (SE to NQ), in repeat R4 asparagine with cysteine (N to C), in repeat R3 cysteine with asparagine (C to N) (**Fig. 3.5b**). We also modified the stacking residues, creating the PUF16rest-s protein, in which we replaced asparagine with tyrosine (N to Y), in repeat R12, tyrosine with arginine (Y to R) in repeat R4, and arginine with tyrosine (R to Y) in repeat R3 (**Fig. 3.5b**). We also engineered a PUFrest16-2.0 specific for binding the *RESTRNA16-2.0* sequence. In this case, we changed the amino acids at position 12, 13 and 16, according to the code, to achieve the specificity for the new sequence (**Fig. 3.5c**).

In summary, we have engineered a total of six REST-specific PUF proteins of which three are designed to bind *RESTRNA8*, i.e. PUF8rest-ns, PUF8rest-s, PUF8rest-sH, two to bind

RESTRNA16, i.e. PUF16rest-ns and PUF16rest-s, and one for *RESTRNA16-2.0*, i.e. PUFrest16 2.0. PUF8wt and PUF16wt with their cognate sequences have been used as controls in all the experiments described in the next sections.

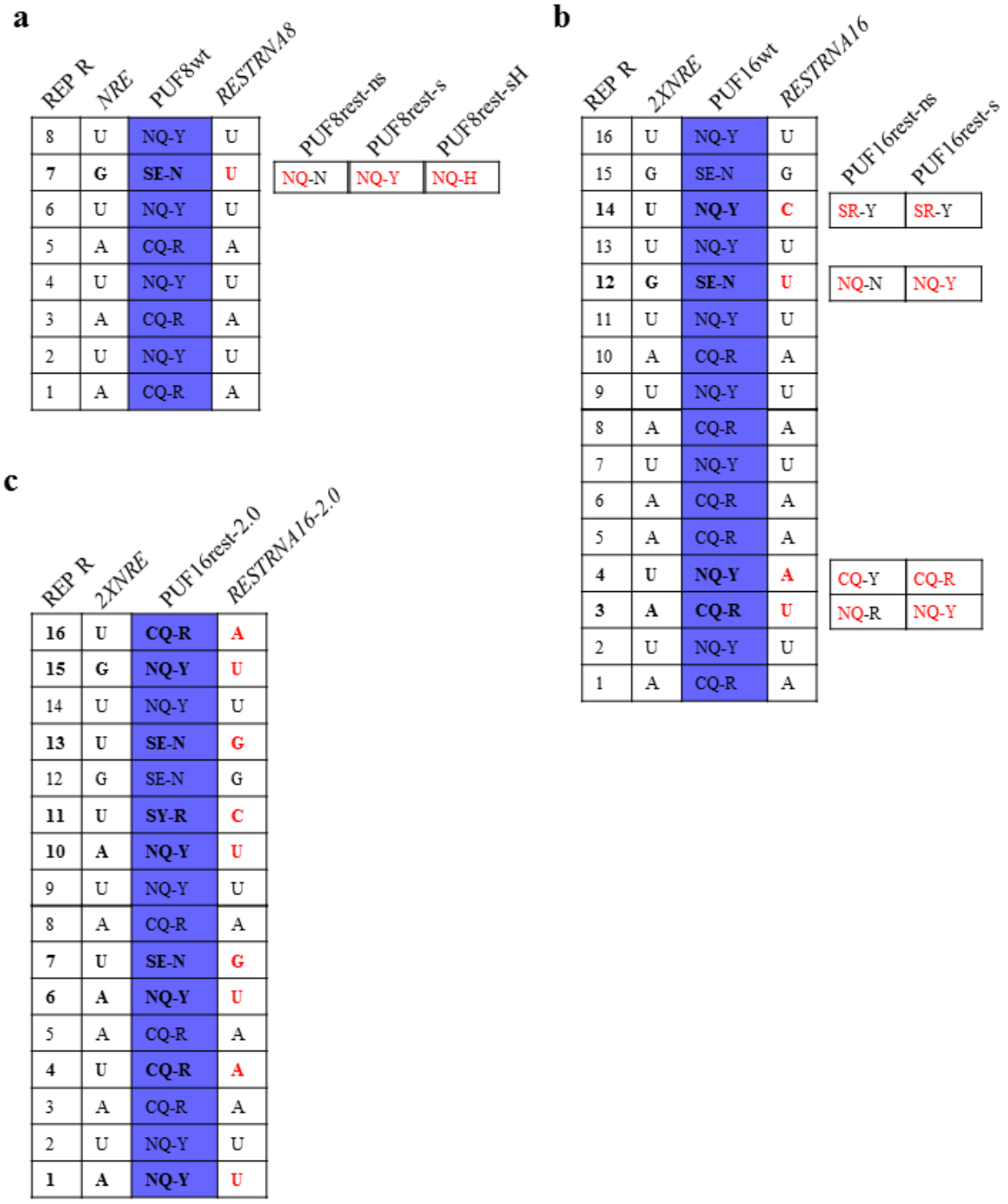


Figure 3.5 Schematic representation of the PUF constructs. a. *NRE* target sequence and the amino acids at position 12, 16 and 13 (stacking) within each of the eight repetitions of PUF8wt. *RESTRNA8* differs from *NRE* for one ribonucleotide (uracil) corresponding to repeat R7. The mutations performed to obtain the three REST-specific PUF8 proteins are represented in red: amino acid 12 and 16 for PUF8rest-ns, amino acids 12, 13 and 16 for PUF8rest-s and PUF8rest-sH. **b**. *2XNRE* target sequence and the amino acids at position 12, 16 and 13 (stacking) within each of the sixteen repeats of PUF16wt. *RESTRNA16* differs from *2XNRE* in four ribonucleotides corresponding to repetitions R14, 12, 4 and 3 (in red). The mutations performed to obtain the two REST-specific PUF16 proteins (PUF16rest-ns and PUF16rest-s) are highlighted in red. **c.** PUF16rest-2.0 with its own target sequence *RESTRNA16-2.0.*

3.2.2 Structural modeling of PUF16wt in complex with its cognate *2XNRE* sequence

A 3D model of the extended PUF domain (PUF16wt), was built starting from two identical PUF proteins and joining them by creating new chemical bonds using the PSFgen Plugin of the VMD program [85], employing the same strategy described in Filipovska *et al* [62]. The first five repeats (R1 to R5) were selected and cut at the end of the R5 repeat and the same procedure was applied for the last three (R6-R8). The proper number of residues to cut was carefully determined by using the Uniprot structure domain section of the PUF8wt (http://www.uniprot.org/uniprot/Q14671#structure) as reference. The second PUF protein was kept as a whole, without the flanking regions, in such way to allow the newly cut parts from the first PUF protein to be fused and, at the same time, preserving the modular structure (**Fig. 3.6**). The same procedure was applied for the RNA molecule.

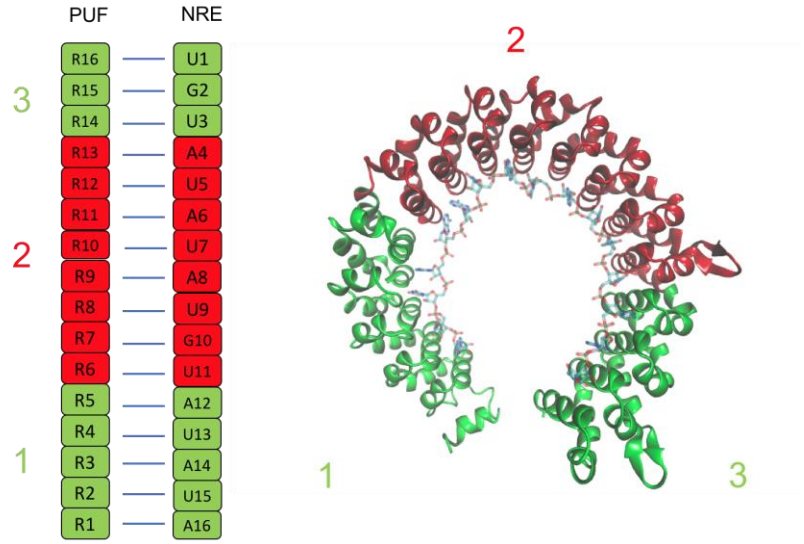


Figure 3.6: *Left*: Schematic representation of the extended PUF16wt structural model built according to Filipovska et al.[62]. *Right*: 3D structure of PUF16wt coupled with the 16 bp *2XNRE* sequence, built using the VMD software (images kindly provided by Drs M. Gatti and L. Maragliano, IIT-NSYN, Genova).

3.2.3 The REST-designed PUF proteins are expressed and localize in the cytosol

As a first step in the validation of the engineered PUF constructs, we checked whether they are correctly expressed in mammalian cells. To this aim, we transfected the various constructs in HEK293T cells and analyzed the expression of the PUF proteins by western blotting analysis using anti-flag antibodies (expected MWs: eight-repeat PUF constructs, 47 kDa; sixteen-repeat PUF constructs, 75kDa). We detected good expression levels of all the proteins (**Fig. 3.7a**). We subsequently purified the proteins starting from mammalian cells, since it has been demonstrated that some mutations may affect the protein solubility and the efficiency of production in bacteria [82] Proteins isolated from transfected HEK293T cells were analyzed by denaturing polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by silver staining, showing a satisfactory purification rate (**Fig. 3.7b**). Finally, to verify the intracellular localization of the new constructs, we transfected HEK293T cells and performed immunostaining and confocal analysis using anti-flag antibodies to mark the PUF proteins. All the proteins localized in the cytosol, as expected (**Fig. 3.8**).

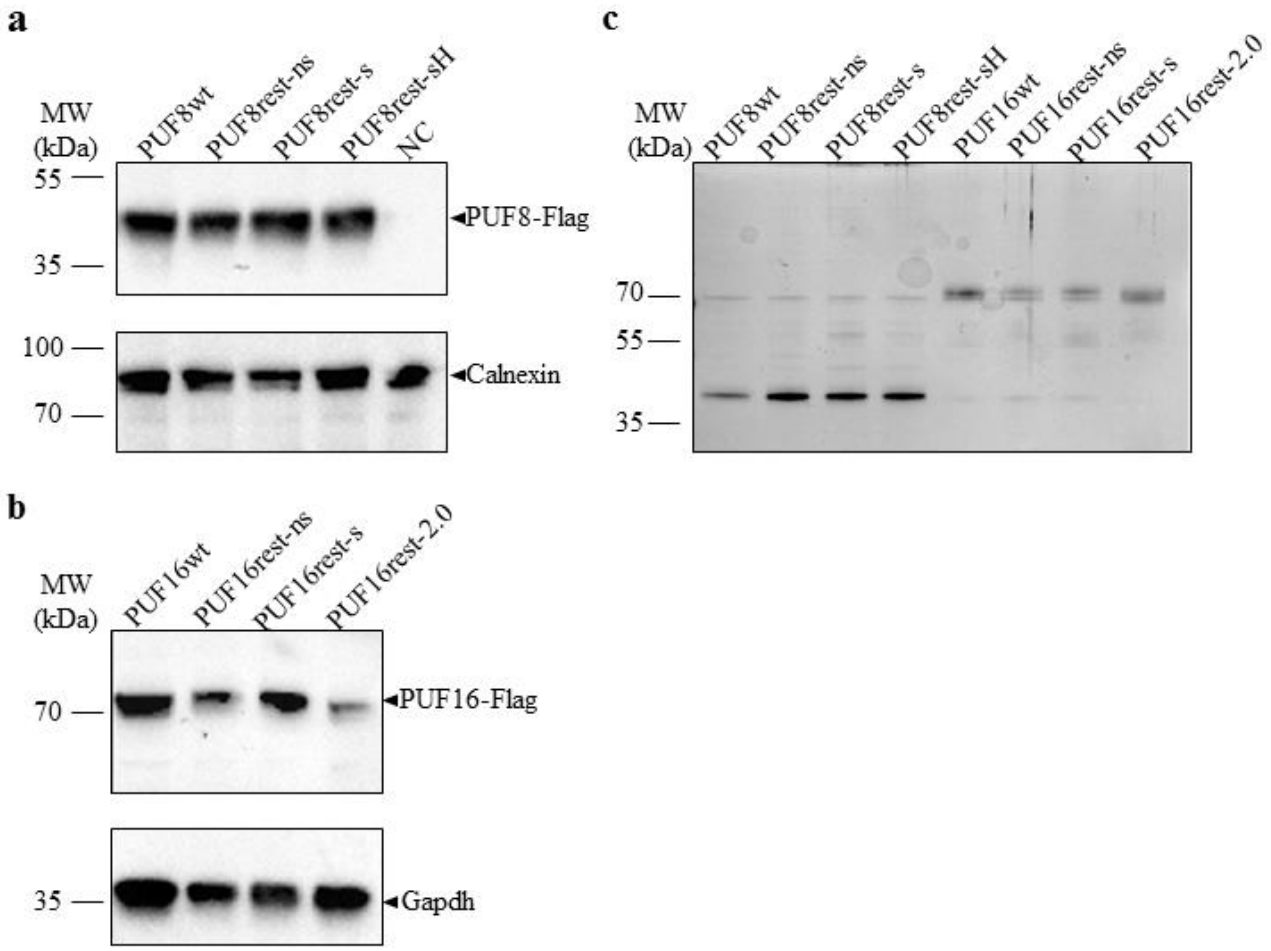
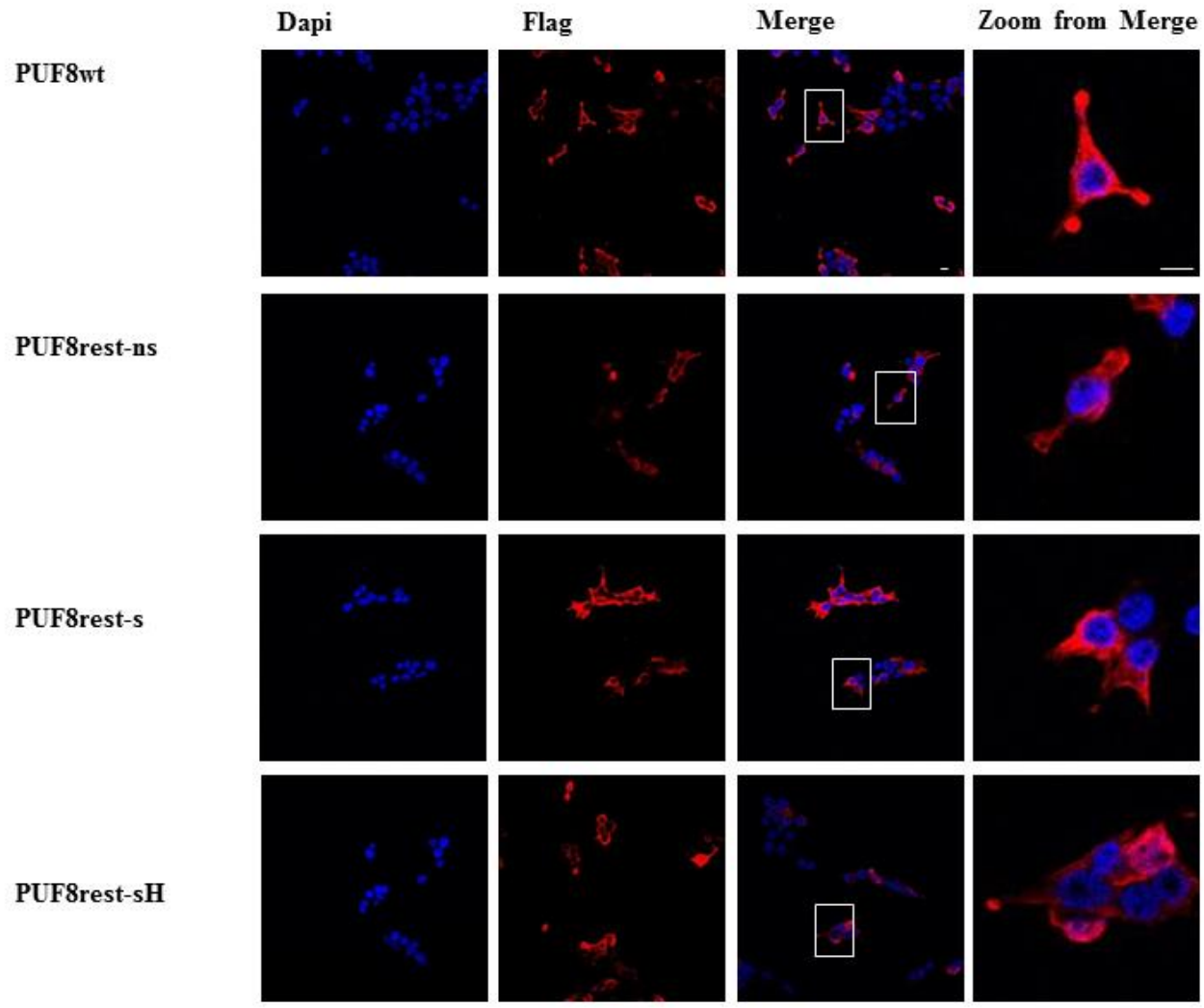


Figure 3.7. REST-designed PUF proteins expression and purification. a. HEK293T cells were transfected with the indicated PUF constructs. NC = negative control (cells transfected with the empty Flag vector). Protein expression was analyzed by western blotting using anti-Flag antibodies and antibodies for the housekeeping genes calnexin and gapdh. **b.** Silver-stained gel of the PUF constructs after immunoprecipitation with anti-flag beads. Cytosolic extracts were incubated with magnetic beads conjugated to anti-Flag antibodies, eluted with 3x flag peptides and analyzed on a 10% polyacrylamide gel, subsequently silver-stained.

a


Dapi
Flag
Merge
Zoom from Merge
PUF8wt
PUF8rest-ns
PUF8rest-s
PUF8rest-sH

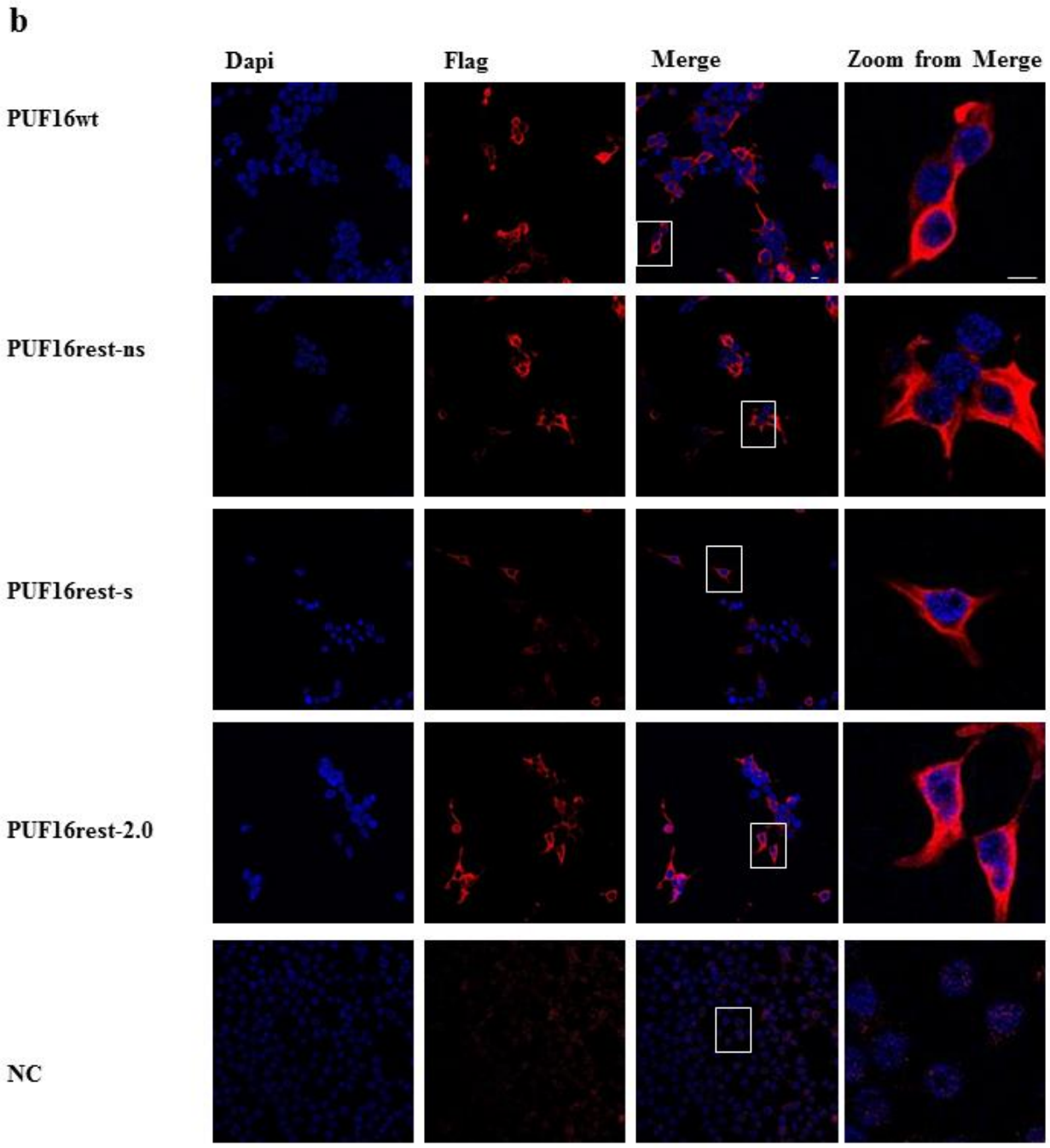


Figure 3.8. REST-designed PUF proteins intracellular localization. Confocal images of HEK293T cells transfected with the indicated constructs were processed for indirect immunofluorescence using anti-Flag antibodies (red) to detect PUF constructs and Hoechst (blue) to visualize cell nuclei. The overlay images (merge) reveal the cytosolic localization of the constructs**.** Scale bars: 10 μm in the lower magnification panel, 10 μm in the higher magnification panels.

3.2.4 Eight- and sixteen-repeat PUF constructs mutated in the stacking residues bind to REST sequences

To verify whether the designed PUF constructs were able to bind specifically their cognate sequences, we performed electrophoretic mobility shift assay (EMSA) analysis using a constant concentration of purified PUF proteins (1 µM) with biotinylated RNA targets. For the eight-repeat constructs (PUF8rest-ns, PUF8rest-s and PUF8rest-sH), the target RNA sequences were *NRE* and *RESTRNA8*. For the sixteen-repeat constructs, the target RNA sequences were *2XNRE* and *RESTRNA16* for PUF16wt, PUF16rest-ns, PUF16rest-s and *RESTRNA16-2.0* for PUFrest16-2.0.

We first checked the binding of every PUF construct for its cognate and non-cognate sequences (**Fig. 3.9**). For what concerns the 8-repeat constructs, we observed the formation of a specific RNA-protein complex between PUF8rest-s and PUF8rest-sH with *RESTRNA8,* while we did not detect any binding to *NRE*. Vice versa, we observed the formation of a specific complex of PUF8wt with *NRE* and not with *RESTRNA8.* However, we did not observe any RNA-protein complex for PUF8rest-ns with any of the target sequences, even when the concentration of PUF8rest-ns used in the assay was raised from 1 to 10 µM (**Fig. 3.9a-b**). Similar results were obtained when EMSA was performed with the 16-repeat constructs. In this case, we observed the formation of a specific complex of PUF16rest-s with *RESTRNA16*, and of PUF16wt with 2X*NRE*, and not with their non-cognate sequences. No RNA-protein complexes were detected for PUF16rest-ns and PUFrest16-2.0 with any of the sequences (**Fig. 3.9c**), not even when the PUF concentration used in the assay was increased to 10 µM (**Fig. 3.9d**).

Altogether, this first series of experiments shows that mutations in the stacking residues are indeed crucial to confer binding specificity to REST-PUF constructs, compared to the same constructs mutated only in the backbone residues. Therefore all the subsequent experiments were restricted to these two proteins.

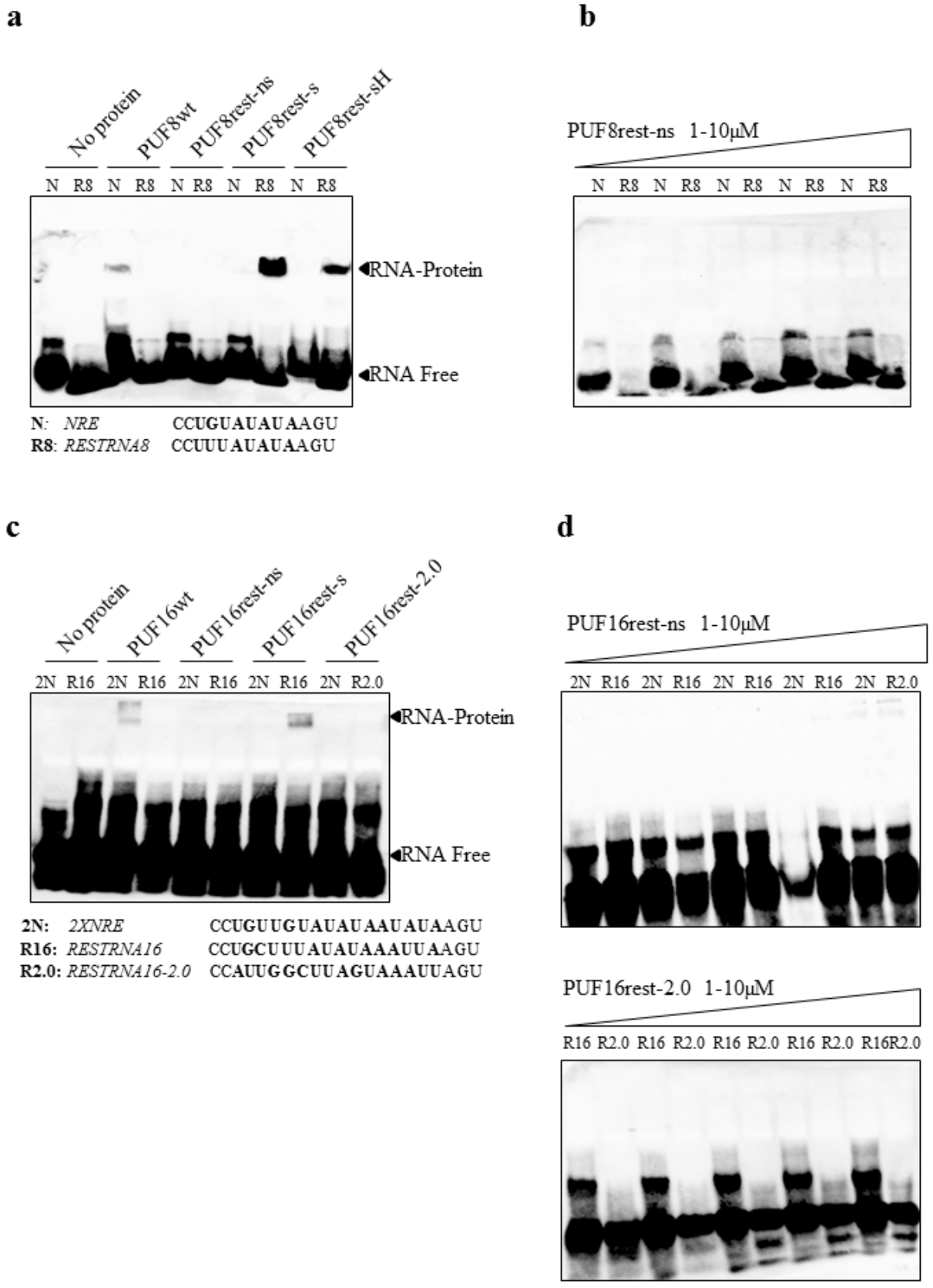


Figure 3.9 Evaluation of PUF proteins specificity by EMSA analysis. a. The indicated 8-repeat PUF proteins were incubated with biotinylated *NRE* (*N*) and *RESTRNA8* (*R8*) ribonucleotide sequences. Protein-RNA complexes and unbound RNA are indicated. **b.** Increasing concentrations of PUF8rest-ns were incubated with *N* and *R8* sequences, as indicated. **c.** The indicated 16-repeat PUF proteins were incubated

with biotinylated *2XNRE* (*2N*), *RESTRNA16* (*R16*) and *RESTRNA16-2.0* (*R2.0*) ribonucleotide sequences. Protein-RNA complexes and unbound RNA are indicated. **d.** Increasing concentrations of PUF16rest-ns were incubated with *2N* and *R16* sequences (upper panel), and PUF16rest-2.0 with *R16* and *R2.0*, as indicated. Biotinylated sequences used in the assays are indicated below the EMSA gels in (a) and (c). highlighted in bold black.

3.2.5 Eight- and sixteen-repeat PUF proteins mutated in the stacking residues bind to REST RNA with high affinity

We next determined the binding affinity of the REST-PUF constructs that showed the highest affinity for their target sequences, i.e. PUF8rest-s and PUF16rest-s, and compared them to the respective wt constructs, i.e. PUF8wt and PUF16wt. We performed EMSA experiments by increasing the concentration of PUF proteins up to 10 µM, in the presence of molar excess of target RNAs. The amount of RNA bound to proteins was quantified by densitometric analysis and absolute amounts (ng) were extrapolated by running in the same blot an RNA titration curve. The amount of bound RNA (ng) was subsequently plotted as a function of PUF protein concentration (µM), and data fitted by non-linear regression (see Materials and Methods and legends to Fig. 3.10 and Fig. 3.11). We measured the equilibrium dissociation constant for all the analyzed systems. Our results indicate that PUF8wt binds *NRE* with the highest affinity (Kd = 1.3 µM) (**Fig. 3.10a**), while PUF8rest-s has a lower affinity for *RESTRNA8* (Kd = 3.4 µM) despite the fact that it binds more RNA (Bmax = 0.8 ng vs 0.020 ng for PUF8wt) (**Fig. 3.10b**). For what concerns the 16-repeat constructs, instead, we obtained a Kd of 4.1 µM for PUF16wt and of 2.6 µM for PUF16rest-s. Thus, in this case, the REST-specific construct bound its target sequence with a higher affinity than its wt counterpart (**Fig. 3.11**).

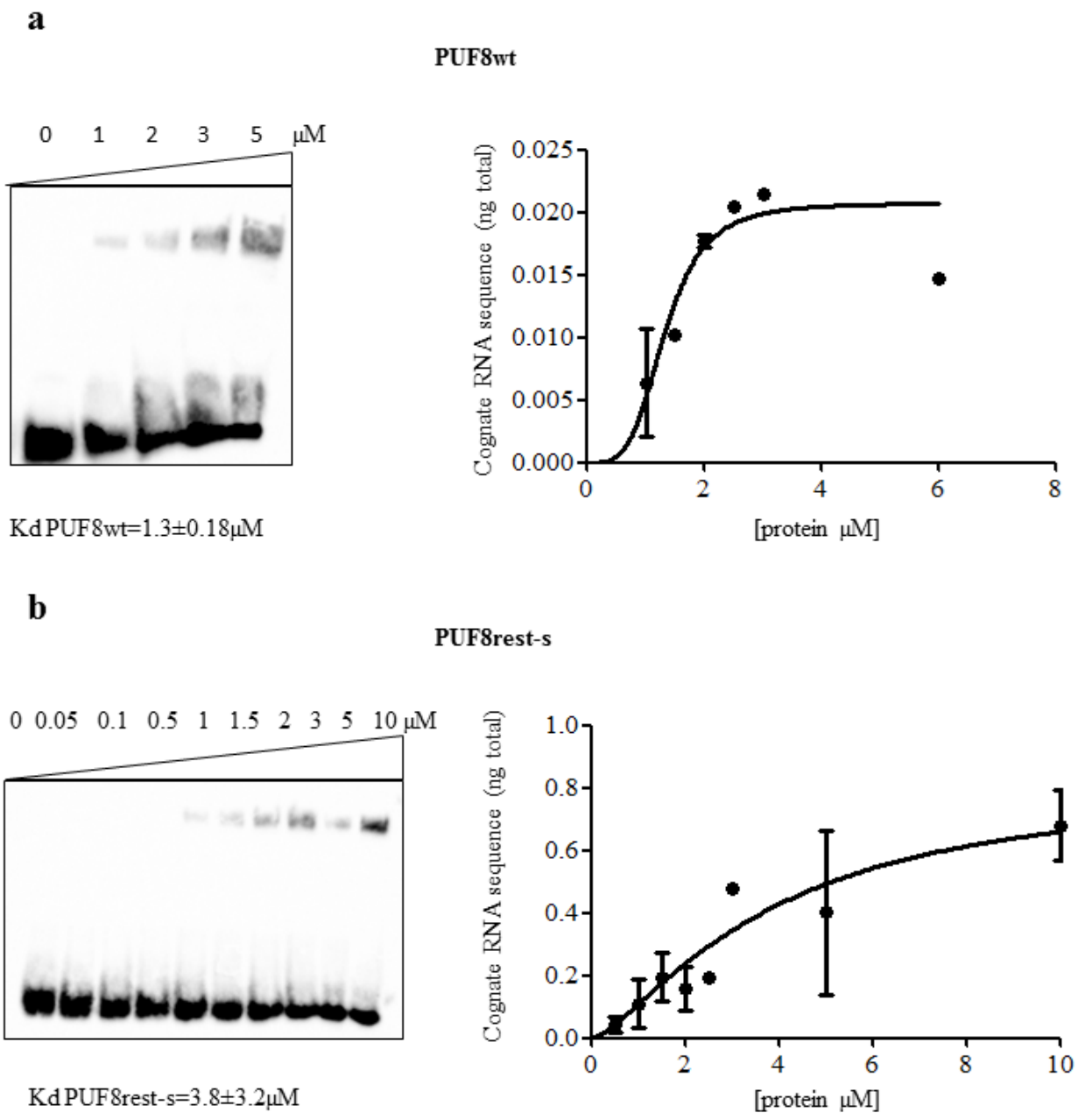


Figure 3.10 Binding isotherms of eight-repeat PUFs to their target RNA sequences. a *Left panel:* Binding of PUF8wt induces a mobility shift in biotinylated *NRE*. *Right panel:* The amounts of PUF-bound cognate RNA, plotted as a function of the purified PUF8wt protein concentration, were quantified in relation to known quantity of biotinylated RNA. Data, shown as means ±SD of two independent experiments, were fitted by nonlinear regression using the one site specific binding with hill slope model Y=Bmax*X^h/(Kd^h + X^h). **b** The same analysis was performed for PUF8rest-s with biotinylated *RNAREST8*.

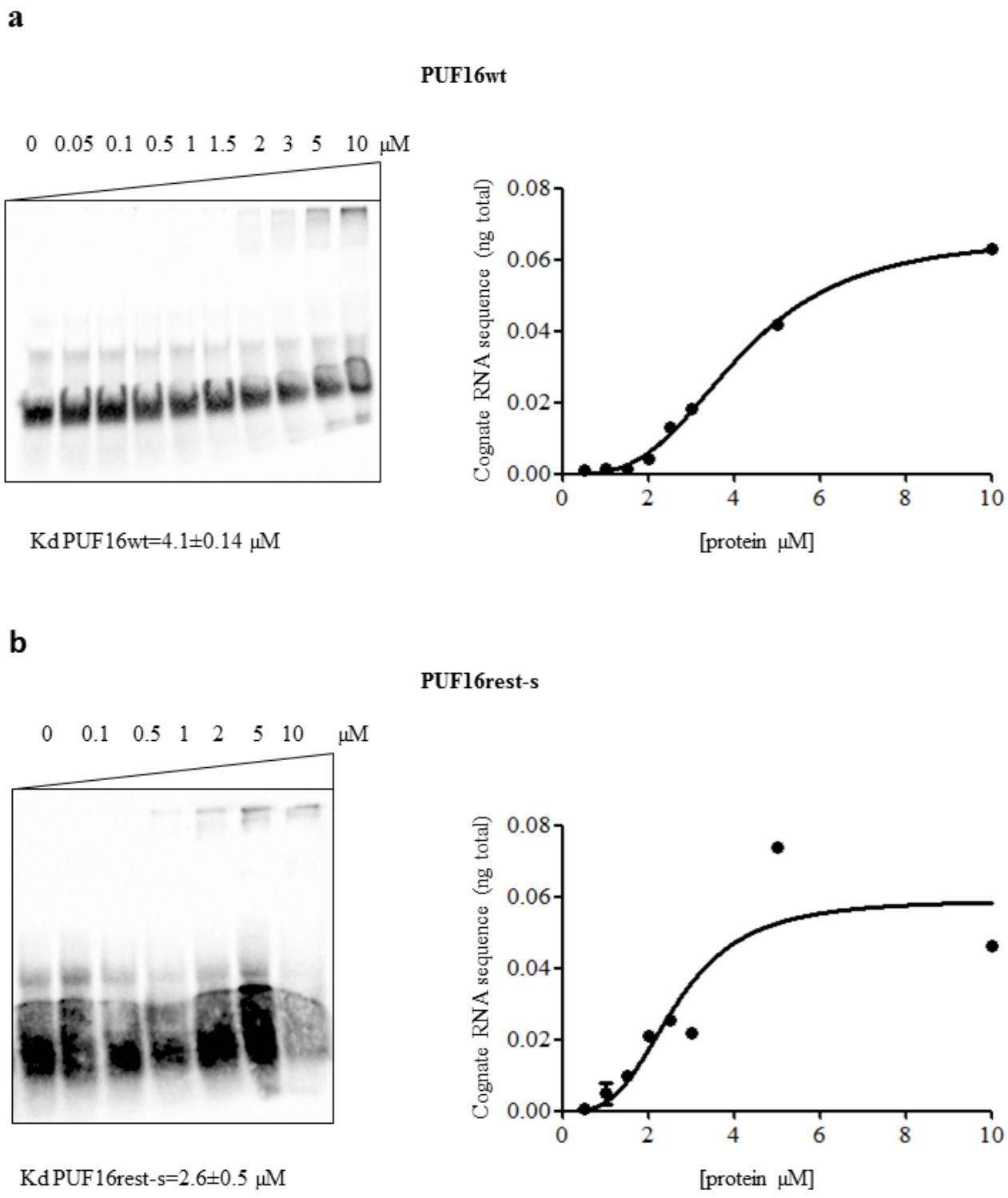


Figure 3.11 Binding isotherms of sixteen-repeat PUFs to their target RNA sequences. A. *Left panel:* Binding of PUF16wt induces a mobility shift in biotinylated *2XNRE*. *Right panel:* The amounts of PUF-bound cognate RNA, plotted as a function of the purified PUF16wt protein concentration (1-10 µM), were quantified in relation to known quantity of biotinylated RNA. Data, shown as means ± SD of two independent experiments, were fitted by nonlinear regression using the one site specific binding with hill slope model Y=Bmax*X^h/(Kd^h + X^h) **b.** The same analysis was performed for PUF16rest-s with biotinylated *RNAREST16.*

3.2.6 Molecular dynamics simulations of PUF-RNA interaction

In order to investigate in detail the structural determinants of the PUF/RNA interactions, structural modeling and Molecular Dynamics (MD) simulations of the PUF/RNA complexes were performed in collaboration with Drs. M. Gatti and L. Maragliano, (IIT-NSYN, Genova). Using MD, it is possible to reconstruct the time evolution of a system by employing mathematical and physical models based on Newtonian dynamics. These simulations are usually performed starting from the 3D atomic structure of a system obtained from X-Ray crystallography, and allow gaining insights on the atomistic details at the basis of the molecular interactions. For the 8-repeat PUF systems the crystal structure of the PUMILIO-homology domain from Human PUMILIO1 with hunchback *NRE* (PDBcode: 3Q0P) was used, while the 16-repeat structures were modeled starting from two copies of the same crystal that were fused as described above. All systems were simulated using the Amber force field [86] and the NAMD [87] software packages. The VMD program [85] was used to visualize the 3D structures and analyze the trajectories. Here we show simulations only for the PUF constructs whose RNA specificity was described in the previous section. Thus, three simulations were produced for the 8-repeat constructs: PUF8wt-*NRE*, PUF8wt-*RESTRNA8*, PUF8rest-s-*RESTRNA8*, and three for the 16-repeat constructs: PUF16wt-*2XNRE*, PUF16wt-*RESTRNA16*, PUF16rest-s-*RESTRNA16*. PUF8wt-*RESTRNA8* and PUF16wt-*RESTRNA16* were considered as negative controls, and designed to investigate the structural consequences of not respecting the recognition code. To this aim, the PUF proteins were left unaltered, while the RNA molecules were mutated to match REST RNA.

To analyze the simulations, we measured the distances and the non-bonded interaction energy between centers of mass (COMs) of RNA nucleotides and facing PUF residues. The COMs were defined using the ribonucleotide's nucleobase and the side chains of the two edge-recognizing residues and the residue involved in the stacking interaction. As for the interaction energy calculation, electrostatic and van der Waals energy contributions were measured between the same groups. Fig. 3.12 reports the results of the distances and energy analysis for the groups that, once mutated, showed the most pronounced effect. Specifically, for the 8-repeat systems we show results for R7 and the facing 2nd ribonucleotide N2 (**Fig. 3.12a-b**), while for the 16-repeat systems we show results for R4 and the corresponding ribonucleotide N13 (**Fig. 3.12c-d**). In detail, these interaction groups are SE-N with guanosine in PUF8wt-*NRE*, NQ-Y with uracil in PUF8rest-s-*RESTRNA8*, and SE-N with uracil in PUF8wt-*RESTRNA8*; similarly, they were NQ-H with uracil in

PUF16wt-*2XNRE*, CQ-R with adenine in PUF16rest-s-*RESTRNA16*, and NQ-H with adenine in PUF16wt-*RESTRNA16*.

The PUF8wt-*RESTRNA8* system (green curve in **Fig. 3.12a-b**) shows the highest values and fluctuations for both distance and energy. Energy values are positive, indicating a destabilizing interaction. Between 0 and 50 ns a sudden increase in distance values is observed, which is due to a base-flipping event of the ribonucleotide. The corresponding conformation is shown in the snapshot at the bottom left of **Fig 3.12**. It can be observed that the nucleobase points out of the protein-RNA interaction surface, away from the corresponding amino acids. Conversely, both PUF8wt-*NRE* (orange curves in **Fig. 3.12a-b**) and PUF8rest-s-*RESTRNA8* (cyan curve) show stable distance profiles throughout the simulation at about 5 Å, corresponding to stable non-bonded interaction energy values around -30/-20 kcal/mol. For what concerns the 16-repeat constructs, the PUF16wt-*RESTRNA16* system (violet curves in **Fig 3.12c-d**) shows also high energy and very high distance values over the whole simulated trajectory, corresponding again to a base-flipping event. The associated conformation is shown in the snapshot at the bottom left of **Fig. 3.12**, where it can again be observed that the nucleobase points away from the protein-RNA interaction surface. PUF16wt-*2XNRE* (red curve) and PUF16rest-s-*RESTRNA16* (blue curve) show a similar behavior as their 8-repeat counterparts. In particular, the PUF16rest-s-*RESTRNA16* system displays even lower distances and energies than the PUF16wt-*2XNRE*, with plateaus at around 4.5 Å and -20/-30 kcal/mol, respectively.

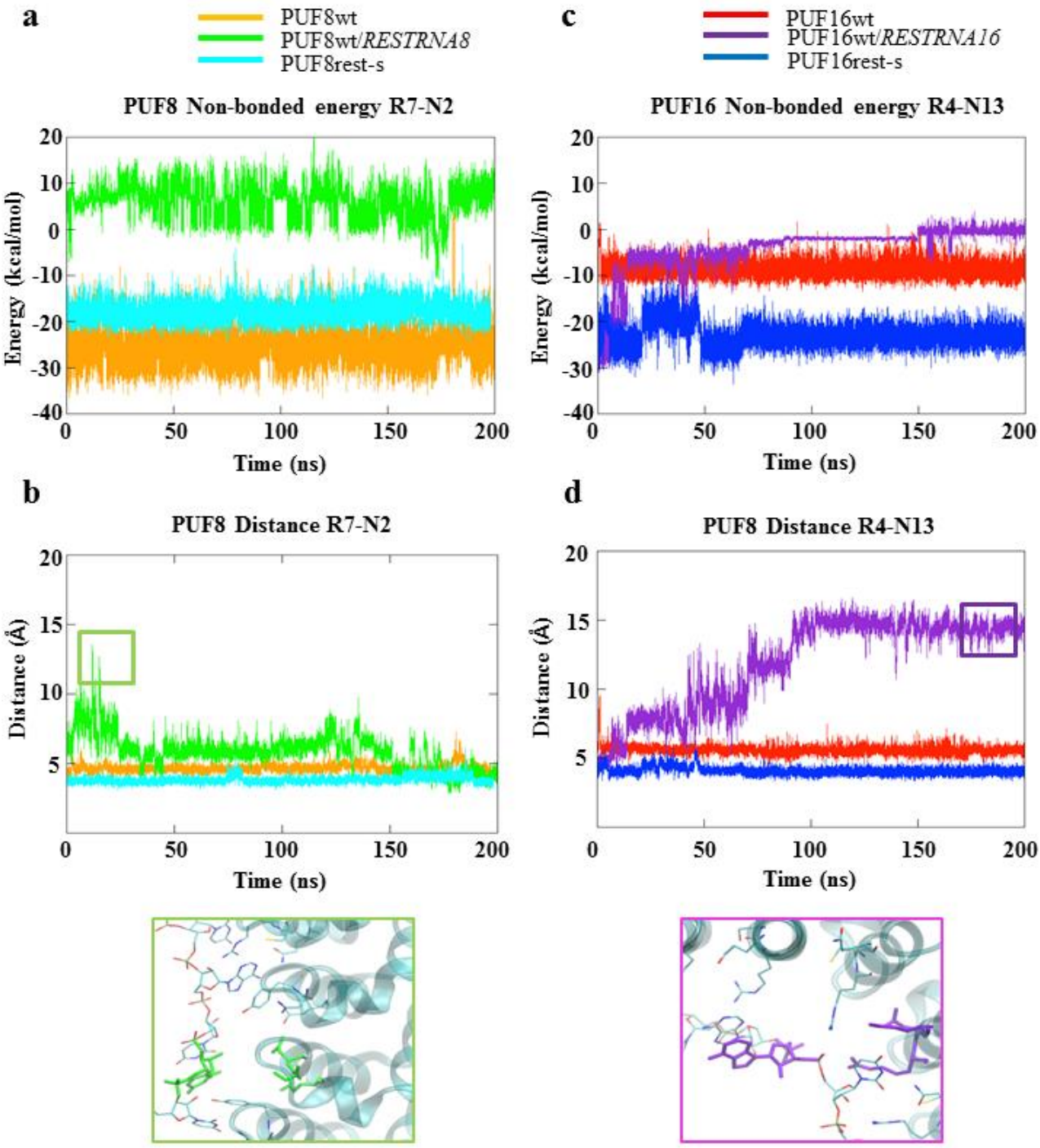


Figure 3.12 Distance and energy plots for 8-repeat (left) and 16-repeat (right) PUF/RNA systems. On the x-axis the simulated time is represented in nanoseconds (ns), while on the y-axis energy and distance are reported in kcal/mol and Å, respectively. **a.** Non-bonded energy plot for the R7-N2 interaction in PUF8wt-*NRE* (orange), PUF8wt-*RESTRNA8* (green) and PUF8rest-s-*RESTRNA8* (cyan). **b.** Distance between COMs of the R7-N2 repeat for PUF8wt-*NRE* (orange), PUF8wt-*RESTRNA8* (green) and PUF8rest-s *RESTRNA8* (cyan). **c.** Non-bonded energy plot of the R4-N13 interaction in PUF16wt-*2XNRE* (red), PUF16wt-*RESTRNA16* (violet) and PUF16rest-s-*RESTRNA16* (blue). **d.** Distance between COMs on the R4-N13 interaction over the PUF16wt-*2XNRE* (red), PUF16wt-*RESTRNA16* (violet) and PUF16rest-s-*RESTRNA16* (blue). Indicated by the colored boxes at the bottom of the figures are snapshots from the simulations showing the 3D structures of the flipped bases for PUF8wt-*RESTRNA8* (green box) and PUF16wt-*RESTRNA16* (violet box) (images kindly provided by Dr M. Gatti and L. Maragliano, IIT-NSYN, Genova).

3.2.7 Eight- and sixteen-repeat PUF proteins mutated in the stacking residues selectively tether endogenous REST mRNA without altering its expression

Having demonstrated the ability of REST-PUF constructs to bind their specific sequences, confirming the selectivity and the affinity of the designed proteins, we decided to investigate whether they are able to bind endogenous REST mRNA. To address this point, we performed cross-linking RNA immunoprecipitation (CLIP) followed by a quantitative analysis with qPCR. N2a cells were transfected with PUF8rest-s, PUF16rest-s, PUF8wt and PUF16wt; after 48 h cells were lysed and immunoprecipitated with anti-flag antibodies or IgGs as control, and agarose A beads. Immunoprecipitated RNA was subsequently analyzed by qRT-PCR with primers specific for REST 3'UTR and Gapdh as control. Our results indicate that both REST-specific proteins (PUF8rest-s, PUF16rest-s) bind endogenous REST mRNA more strongly than their wild type counterparts (PUF8wt, PUF16wt). In particular, PUF16rest-s shows a very high binding affinity, as it precipitates nine times more REST mRNA than PUF16wt (**Fig. 3.13a**). We do not detect any specific binding of any of the proteins to Gapdh mRNA, confirming the specificity of our results, assessed also by the immunoblot of protein fractions (**Fig. 3.13b**). These data confirm our previous results, and establish that eight- and sixteen-repeat PUF proteins mutated in the stacking residues specifically bind endogenous REST mRNA.

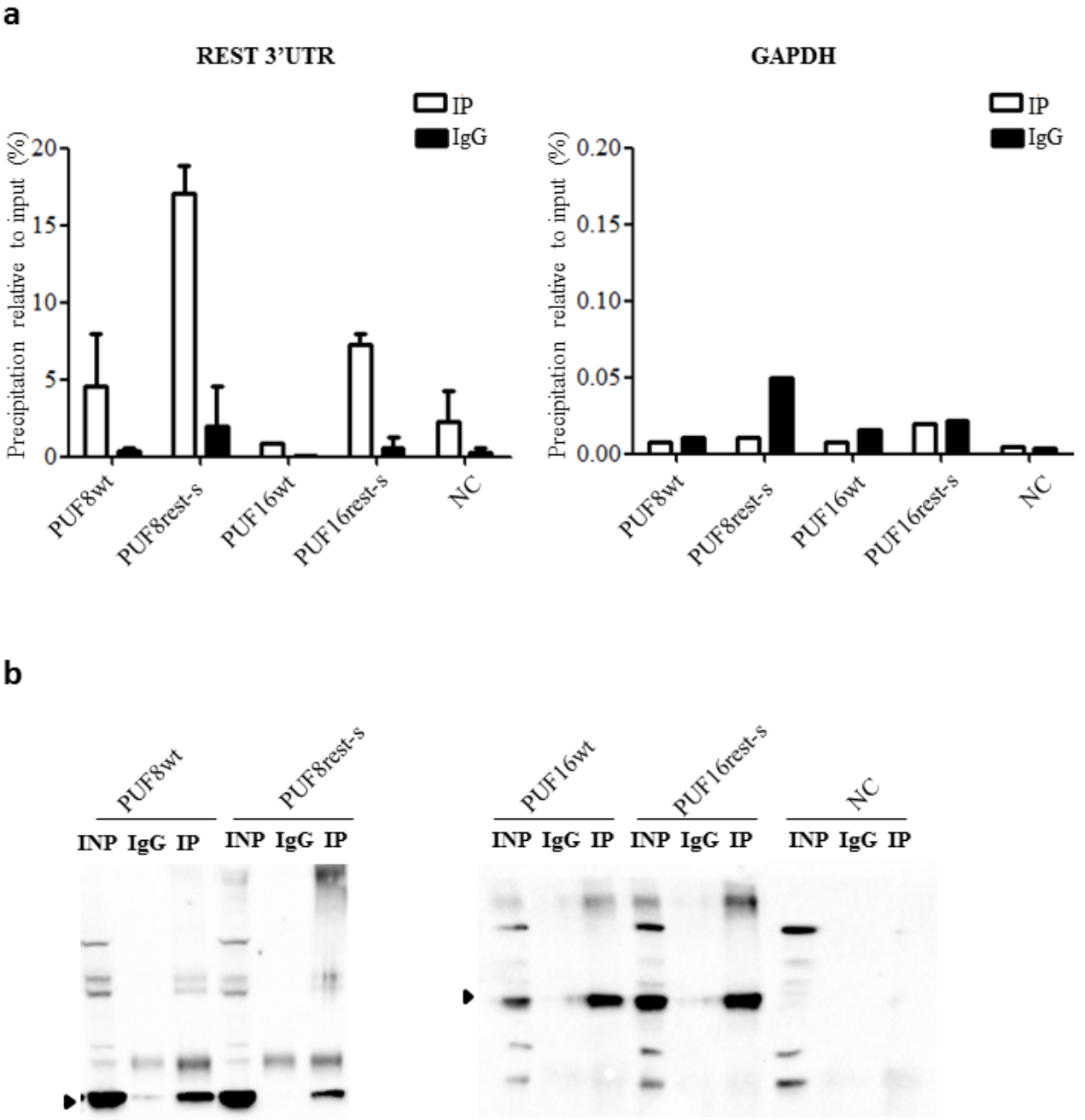


Figure 3.13. Crosslink RNA immunoprecipitation (CLIP) of the REST-PUF stacking constructs for endogenous REST mRNA. a. CLIP was performed with agarose A beads, anti-flag and IgGs antibodies on N2a cell lysates transfected with the indicated constructs. REST (left panel) and Gapdh (right panel) values were normalized against the input value and plotted as percentage of specific precipitation. Data are shown as means ± SD of two independent experiments. **b.** Immunoblot of immunoprecipitated complexes revealed with anti-flag antibodies. Input (INP) represents 20% of the cell extract before immunoprecipitation for all the indicated constructs. NC, negative control: cells transfected with the empty 'flag' vector.

PUF proteins mediate the translational regulation of their target mRNAs through various mechanisms. We asked whether our PUF constructs have translational repression or activation activity when they bind to their target sequence. To answer this question, we cloned the PUF target sequences downstream of the Renilla coding sequence before the polyA tail, on a plasmid carrying also the luciferase gene (psicheck-2) (**Fig. 3.14a**) We co-transfected HEK293T cells with PUF8rest-s, PUF16rest-s, PUF8wt and PUF16wt together with the corresponding target sequences into the psicheck-2 vector and performed luciferase assays. We measured the ratio of Renilla over Luciferase expression (**Fig. 3.14b**) and compared it to control samples co-transfected with the psichek-2 vector carrying the same target sequence and the flag vector in the absence of PUF proteins. We did not observe any change in Renilla expression for all the PUF constructs, demonstrating they do not have any intrinsic effect on the translational activity of the target gene. Furthermore, we evaluated the effect of PUF constructs on endogenous REST mRNA. We transfected N2a cells with PUF8rest-s, PUF16rest-s, PUF8wt and PUF16wt and performed qRT-PCR analysis to measure REST mRNA levels. As shown in **Fig. 3.14c**, we did not detect any variation in REST mRNA levels, confirming that our PUF constructs do not affect the stability of REST mRNA.

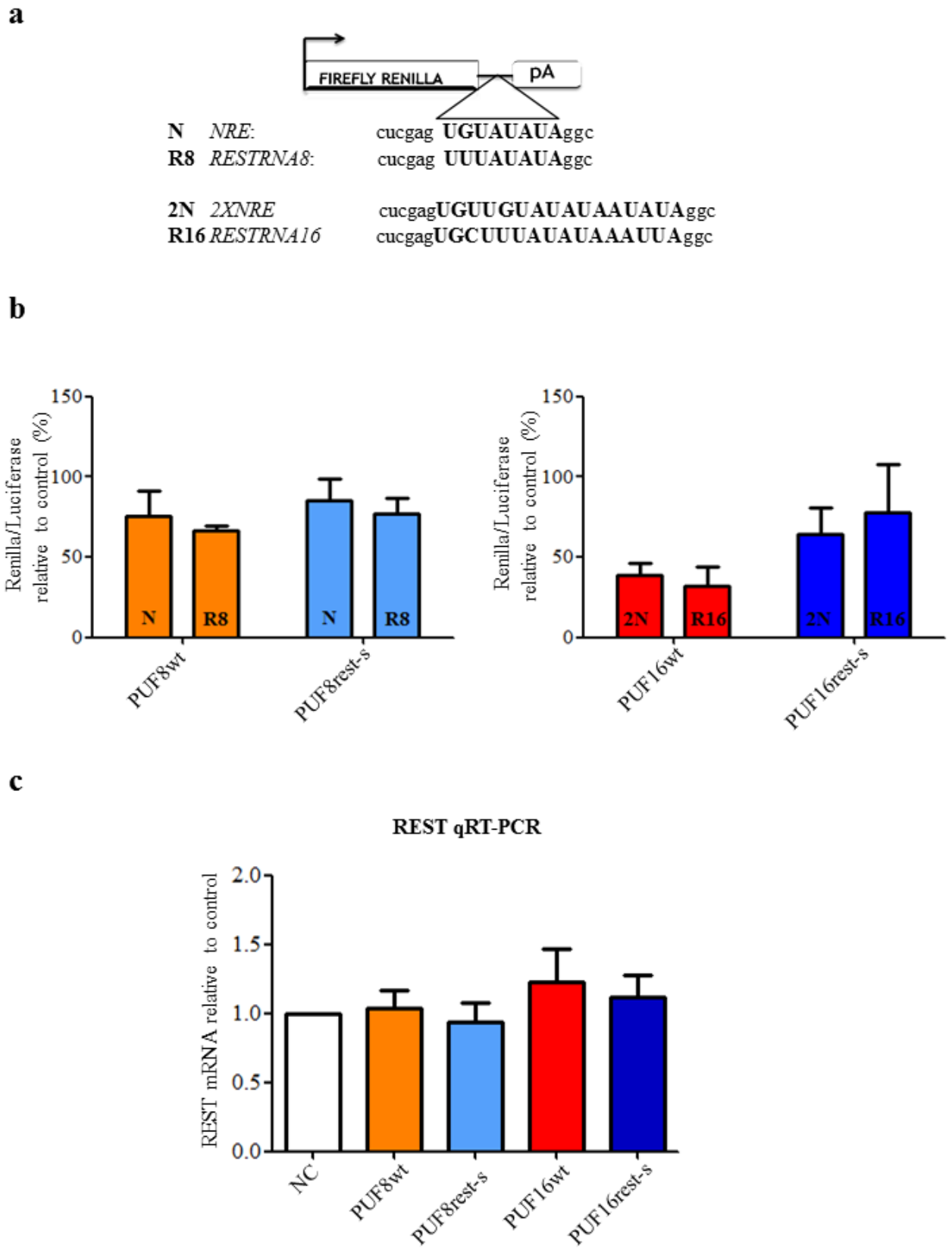


Figure 3.14. Evaluation of the translational regulation capability of PUF constructs. a. Schematic representation of the PUF target sequences, cloned downstream of the Renilla coding sequence, before the polyA tail, on the psicheck-2 plasmid. **b.** Luciferase assay of HEK293T cells co-transfected with the indicated constructs and the corresponding target sequences. Luciferase activity was measured 48 h after transfection. The Renilla / Luciferase ratio was first calculated for every sample, and data (means ± sem)

were subsequently normalized to the activity of the negative control, set to 100 % Renilla (n=3 independent experiments). **c.** N2a cells were transfected with the indicated constructs. After 48 h, REST mRNA levels were quantified via qRT-PCR analysis. Gapdh was used as control housekeeping gene. Data (means ± sem) of n = 7 independent experiments.

4 Discussion

LOV-based probes to inhibit REST activity. The ability to regulate the activity of transcription factors represents an essential step for a better understanding of the molecular and epigenetic mechanisms that constitute the basis of cell physiology, and a promising approach in the treatment of pathologies linked to dysregulation of specific gene clusters. Epigenetic modifications are crucial to rapidly adapt the neuronal transcriptional response to developmental and environmental hints. In this context, REST plays a central role in the determination of the neuronal fate [88], as well as in the modulation of neuronal activity and plasticity [23]. The role of REST in the onset of pathologies is complex, acting under some circumstances as an oncogene and neuroprotector [25], and under other conditions as a promoter of insult-induced neuronal death (brain ischemia, HD) [20, 27] or dysfunctions like epilepsy [89]. Due to its crucial role in several neuropathologies, various molecular strategies have been developed to target REST activity, including decoy oligodeoxynucleotides, interfering peptides, and stable expression of dominant-negative/constitutively active forms of REST [19, 33, 90]. However, all these approaches are burdened by intrinsic limitations such as the rapid degradation of the probes or the constitutive and non-tunable activation of REST-target genes, which may lead to unwanted side effects. The optogenetic tools described here are endowed with several features that circumvent these limitations, such as (i) their effect is reversible and can be finely modulated by specific illumination protocols; (ii) they can be used for long-term applications, being genetically encoded; and (iii) they counteract REST dysregulation without affecting REST expression. In this work, we describe an optogenetic approach aimed at modulating gene transcription in a dynamic and reversible way, by acting on the activity of the master transcriptional repressor REST. We engineered chimerical proteins based on the blue light absorbing AsLOV2 domain and exploited the inhibitory potential of small protein domains belonging to the endogenous REST interactors mSin3a and RILP/PRICKLE1. Following this strategy, we were able to target two key events in REST physiology: (i) the assembly of the repressor complex on RE1 sites and (ii) REST binding to DNA. Optogenetic activation of our LOV-based opto-probes selectively induced transcription of REST target genes in neuroblastoma cells and in primary neurons. By applying specific illumination protocols, we were able to drive and modulate important physiological processes, such as the neural differentiation of N2a cells and the firing properties of cortical neurons. These data open the possibility to devise novel therapeutic strategies for brain diseases based on the optogenetic control of the neuronal epigenome.

PUF-based probes: anchors to bind REST mRNA. With the purpose of achieving the bi-directional control of REST activity we followed a different strategy, based on the direct regulation of REST expression acting on its mRNA. In particular we employed PUF domains, a class of RNA binding proteins (RBPs) that typically bind to the 3'UTR of mRNA transcripts, acting as post-transcriptional regulators. The PUF-RNA code has been identified and used in other works to engineer PUF proteins endowed with novel sequence specificity. When fused to functional domains, the engineered PUF-based proteins were able to modulate the stability of the targeted transcripts [71]. Following the published code, we selected a sequence in the 3'UTR of mouse REST mRNA that differed in only one base from the wild type PUF target sequence *NRE*. We developed three eight-repeat REST-specific PUFs, one mutated only in the edge-on residues (PUF8rest-ns) and the other two mutated in the edge-on and stacking residues (PUF8rest-s and PUF8rest-sH). Although the PUF domains are very specific for their targets, a sequence of eight ribonucleotides is present at many off target sites in the transcriptome. To overcome this issue, we took advantage of a sixteen-repeat PUF domain, which recognizes a 16 bp sequence [62], and created three sixteen-repeat REST-specific PUFs, one mutated only in the edge-on residues (PUF16rest-ns), one mutated in the edge-on and stacking residues (PUF16rest-s), and another one mutated in edge-on and stacking residues, recognizing a different sequence in REST 3'UTR (PUF16rest-2.0). All the constructs were correctly expressed in immortalized cell lines (HEK293T, N2a), and localized to the cytosol, as expected. Moreover, all the recombinant proteins could be purified with similar efficiency, which allow to perform the subsequent analysis of in vitro specificity. Our EMSA experiments demonstrated that only the PUF constructs mutated in edge-on and stacking residues are able to interact with the selected REST sequence, and not with the *NRE* site. We then determined the binding affinity of PUF8rest-s and PUF16rest-s to their cognate RNA, obtaining similar values (Kd 3.8 and 2.6 μM, respectively), which differ from the values obtained for the wild type PUF proteins (PUF8wt, 1.3 μM and PUF16wt, 4.1 μM) of 4.9 and 0.6 folds, respectively. Thus, the affinity of the mutant proteins for their target RNA is similar but still not as good as the one of wild type PUF proteins. This result is in line with the work of *Cheong et al* [82] where the PUF mutants, although specific for their cognate sequences, do not present the same Kd values as reported for wild type (for example: mutant7-2(G2U), 6.0 nM, MUT6-2/7-2 (GU23UG), 18 nM vs wild type, 0.48 nM)[82].

These data together with structural modeling and molecular dynamics simulations reveal the importance of respecting the recognition code in order to obtain binding between PUF proteins and cognate RNA molecules. Indeed, when the code is not maintained, the simulated systems sample unstable structures with highly perturbed protein/RNA interaction surfaces. This is reflected in

weaker local interactions between the ribonucleotides and the protein residues involved in the recognition repeat, which in turn might imply weakly bound complexes. Conversely, when the new interfaces for the REST transcript are designed following the recognition code, the simulated systems behave locally as the wild-type, or even show an increase in stability as for the PUF16rest-s-*RESTRNA16* system. Although in many works [74, 76, 91] it is proved that by changing only the amino acids at positions 12 and 16 PUF reaches the desired specificity, our results, in line with the work of *Koh YY et al* [84], demonstrate that PUF proteins mutated in the residues at positions 12-16 and 13 are the only mutant proteins that specifically bind REST mRNA, underlying the importance to change also the stacking residues in the mutation design since they contribute to the protein specificity for RNA sequences, as well as to its affinity. In a recent work *Adamala et al* [83] reported the use of a single repeated module from Pumilio (Pumby) to support protein generation, analogous to the TALE design for DNA sequences, with different lengths from 6 to 18 modules, whose performances they argue are equivalent to the original Pumilio protein, underlying the fact that is possible to engineer PUF proteins specific for any RNA sequence. However these results are in contrast with our EMSA results for PUF16rest-2.0, since even if we use the same PUF scaffold, changing 9 repeats according to the code (stacking residues included), we are not able to detect any specificity. Therefore if on one hand it is fundamental to respect the code in creating new PUF constructs, on the other side there are other factors that definitely need to be evaluated in the creation of these structures, such as the secondary and tertiary structure of the PUF scaffold that determine the correct three-dimensional curvature for RNA binding, ultimately determining binding specificity.

Binding of PUF proteins to their target mRNA could in principle degrade it, stabilize it or leave it unaffected [92-94], for this reason we evaluated the affinity of PUF proteins for the endogenous REST transcript. In agreement with our in vitro data, our CLIP experiments clearly showed that the eight- and sixteen-repeat PUF proteins mutated in the stacking residues bound strongly the endogenous REST mRNA. Thus, we provide the first evidence that also the sixteen repeat PUF protein can be mutated according to the code in order to specifically bind endogenous sequences. Our functional studies demonstrated that PUF8rest-s and PUF16rest-s do not alter the translation of a reporter gene, and therefore they do not have any intrinsic effect on REST mRNA stability. This is relevant as these constructs will be used as an anchor to target REST mRNA and will be fused to effector proteins in order to modulate REST mRNA expression.

5 Future Perspectives

The objective of this thesis was to achieve the bi-directional regulation of the transcriptional repressor REST. For what concerns the LOV-based optogenetic strategy, our next step will be to test the efficacy of our probes in vivo. In particular, we are interested in using such tools to investigate the role of REST a number of plasticity phenomena, such as visual cortex plasticity in adult animals. Moreover, we also want to employ these tools to address the role of REST in pharmacological and genetic mouse models of epilepsy, in which the function of REST is still debated.

The described PUF-based probes, instead, will serve as a platform to anchor specific effector proteins to REST mRNA. As first strategy we will fuse our constructs to GLD-2, a member of the PAP family of PolyA Polymerases that add adenines to the polyA tail of mRNA transcripts, making them more stable [95-97]. The PUF-GLD2 fusion proteins should therefore increase REST mRNA stability, thus increasing the rate of protein translation. As a final step we will make our probes light-sensitive by fusing them to appropriate light-sensitive domains. In this regard, we have envisaged two possible strategies. In the first approach we will fuse the LOV domain directly to the PUF-GLD-2 chimera. In the resulting construct, illumination should expose the RNA binding site allowing PUF to bind REST mRNA, and allowing GLD-2 to add adenines to the polyA tail, increasing REST mRNA stability and consequently its translation rate. In the dark state instead the Jα helix would mask the RNA recognition site of PUF preventing the binding of PUF to REST mRNA and blocking the action of GLD-2 (**Fig. 5.1**). To identify the best way to link the various components of the chimera to each other (i.e. by varying the length and / or the sequence of the linker regions), we will perform molecular modeling simulations based on the known crystal structure of PUF alone and of the PUF-GLD2 chimera (in collaboration with Drs. L. Maragliano and M. Gatti, IIT-NSYN, Genova). As a second strategy we will plan to use the CRY2/CIB cryptochrome system, described in par 1.2, fusing CRY2 to PUF and CIB to GLD-2. The light-induced dimerization of the two fusion proteins would anchor GLD2 to REST mRNA thus allowing its action, ultimately stabilizing REST mRNA.

Altogether, the approaches here described will allow us to control REST activity bi-directionally and on-demand, and will therefore represent a powerful tool to interrogate the function of this important transcription factor in health and disease.

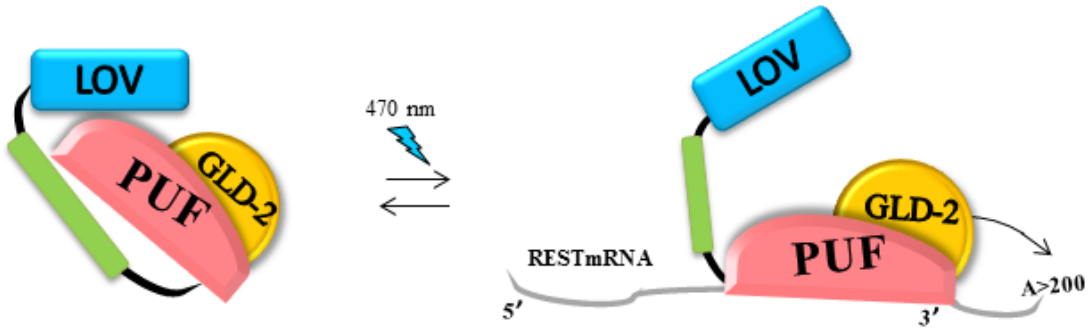


Figure 5.1. a. *Schematic representation of the ASLOV2-PUF-GLD-2 chimera.* PUF is fused to GLD-2, which adds adenines to the polyA, stabilizing the target mRNA. The fusion of this complex to AsLOV2 makes it light-inducible. In the dark (left) the Jα helix of AsLOV2 masks the PUF RNA recognition site for REST mRNA. Upon illumination (right), a change in the Jα conformation allows the binding of PUF on REST 3'UTR region and the action of GLD-2, thus increasing REST mRNA stability.

6 Bibliography

1. Schoenherr, C.J. and D.J. Anderson, *The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes.* Science, 1995. **267**(5202): p. 1360-3.
2. Chong, J.A., et al., *REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons.* Cell, 1995. **80**(6): p. 949-57.
3. Zhao, Y., et al., *Brain REST/NRSF Is Not Only a Silent Repressor but Also an Active Protector.* Mol Neurobiol, 2017. **54**(1): p. 541-550.
4. Battaglioli, E., et al., *REST repression of neuronal genes requires components of the hSWI.SNF complex.* J Biol Chem, 2002. **277**(43): p. 41038-45.
5. Palm, K., et al., *Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene.* J Neurosci, 1998. **18**(4): p. 1280-96.
6. Coulson, J.M., et al., *A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker.* Cancer Res, 2000. **60**(7): p. 1840-4.
7. Lee, J.H., Y.G. Chai, and L.B. Hersh, *Expression patterns of mouse repressor element-1 silencing transcription factor 4 (REST4) and its possible function in neuroblastoma.* J Mol Neurosci, 2000. **15**(3): p. 205-14.
8. Raj, B., et al., *Cross-regulation between an alternative splicing activator and a transcription repressor controls neurogenesis.* Mol Cell, 2011. **43**(5): p. 843-50.
9. Hermanson, O., *Stem cells have different needs for REST.* PLoS Biol, 2008. **6**(10): p. e271.
10. Chen, Z.F., A.J. Paquette, and D.J. Anderson, *NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis.* Nat Genet, 1998. **20**(2): p. 136-42.
11. Ballas, N. and G. Mandel, *The many faces of REST oversee epigenetic programming of neuronal genes.* Curr Opin Neurobiol, 2005. **15**(5): p. 500-6.
12. Ooi, L. and I.C. Wood, *Chromatin crosstalk in development and disease: lessons from REST.* Nat Rev Genet, 2007. **8**(7): p. 544-54.
13. Hwang, J.Y., K.A. Aromolaran, and R.S. Zukin, *The emerging field of epigenetics in neurodegeneration and neuroprotection.* Nat Rev Neurosci, 2017. **18**(6): p. 347-361.
14. Perera, A., et al., *TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression.* Cell Rep, 2015. **11**(2): p. 283-94.
15. Tsai, M.C., et al., *Long noncoding RNA as modular scaffold of histone modification complexes.* Science, 2010. **329**(5992): p. 689-93.
16. Taylor, P., et al., *REST Is a Novel Prognostic Factor and Therapeutic Target for Medulloblastoma.* Molecular Cancer Therapeutics, 2012. **11**(8): p. 1713-1723.
17. Conti, L., et al., *REST Controls Self-Renewal and Tumorigenic Competence of Human Glioblastoma Cells.* Plos One, 2012. **7**(6).
18. Fuller, G.N., et al., *Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST) neuron-restrictive silencer factor, which can be functionally countered by REST-VP16.* Molecular Cancer Therapeutics, 2005. **4**(3): p. 343-349.
19. Lawinger, P., et al., *The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells.* Nat Med, 2000. **6**(7): p. 826-31.
20. Calderone, A., et al., *Ischemic insults derepress the gene silencer REST in neurons destined to die.* J Neurosci, 2003. **23**(6): p. 2112-21.
21. Garriga-Canut, M., et al., *2-deoxy-D-glucose reduces epilepsy progression by NRSF-CtBP-dependent metabolic regulation of chromatin structure.* Nature Neuroscience, 2006. **9**(11): p. 1382-1387.

22. Formisano, L., et al., *Ischemic insults promote epigenetic reprogramming of mu opioid receptor expression in hippocampal neurons.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 4170-5.
23. Pozzi, D., et al., *REST/NRSF-mediated intrinsic homeostasis protects neuronal networks from hyperexcitability.* Embo Journal, 2013. **32**(22): p. 2994-3007.
24. Liu, M., et al., *Neuronal conditional knockout of NRSF decreases vulnerability to seizures induced by pentylenetetrazol in mice.* Acta Biochim Biophys Sin (Shanghai), 2012. **44**(6): p. 476-82.
25. Lu, T., et al., *REST and stress resistance in ageing and Alzheimer's disease.* Nature, 2014. **507**(7493): p. 448-54.
26. Zuccato, C. and E. Cattaneo, *Role of brain-derived neurotrophic factor in Huntington's disease.* Prog Neurobiol, 2007. **81**(5-6): p. 294-330.
27. Zuccato, C., et al., *Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes.* Nature Genetics, 2003. **35**(1): p. 76-83.
28. Song, Z., et al., *REST alleviates neurotoxic prion peptide-induced synaptic abnormalities, neurofibrillary degeneration and neuronal death partially via LRP6-mediated Wnt-beta-catenin signaling.* Oncotarget, 2016. **7**(11): p. 12035-52.
29. Westbrook, T.F., et al., *A genetic screen for candidate tumor suppressors identifies REST.* Cell, 2005. **121**(6): p. 837-48.
30. Neumann, S.B., et al., *Relaxation of glycine receptor and onconeural gene transcription control in NRSF deficient small cell lung cancer cell lines.* Brain Res Mol Brain Res, 2004. **120**(2): p. 173-81.
31. Shimojo, M., et al., *The small cell lung cancer-specific isoform of RE1-silencing transcription factor (REST) is regulated by neural-specific Ser/Arg repeat-related protein of 100 kDa (nSR100).* Mol Cancer Res, 2013. **11**(10): p. 1258-68.
32. Liang, H., et al., *Down-regulation of RE-1 silencing transcription factor (REST) in advanced prostate cancer by hypoxia-induced miR-106b~25.* Exp Cell Res, 2014. **320**(2): p. 188-99.
33. Soldati, C., et al., *Rescue of gene expression by modified REST decoy oligonucleotides in a cellular model of Huntington's disease.* J Neurochem, 2011. **116**(3): p. 415-25.
34. Greenway, D.J., et al., *RE1 Silencing transcription factor maintains a repressive chromatin environment in embryonic hippocampal neural stem cells.* Stem Cells, 2007. **25**(2): p. 354-63.
35. Conforti, P., et al., *Binding of the repressor complex REST-mSIN3b by small molecules restores neuronal gene transcription in Huntington's disease models.* J Neurochem, 2013. **127**(1): p. 22-35.
36. Deisseroth, K., *Optogenetics: 10 years of microbial opsins in neuroscience.* Nat Neurosci, 2015. **18**(9): p. 1213-25.
37. Gradinaru, V., et al., *Molecular and cellular approaches for diversifying and extending optogenetics.* Cell, 2010. **141**(1): p. 154-165.
38. Chow, B.Y., et al., *High-performance genetically targetable optical neural silencing by light-driven proton pumps.* Nature, 2010. **463**(7277): p. 98-102.
39. Mattis, J., et al., *Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins.* Nat Methods, 2011. **9**(2): p. 159-72.
40. Tye, K.M. and K. Deisseroth, *Optogenetic investigation of neural circuits underlying brain disease in animal models.* Nat Rev Neurosci, 2012. **13**(4): p. 251-66.
41. Moglich, A., et al., *Structure and function of plant photoreceptors.* Annu Rev Plant Biol, 2010. **61**: p. 21-47.
42. Kennedy, M.J., et al., *Rapid blue-light-mediated induction of protein interactions in living cells.* Nat Methods, 2010. **7**(12): p. 973-5.
43. Crosson, S., S. Rajagopal, and K. Moffat, *The LOV domain family: photoresponsive signaling modules coupled to diverse output domains.* Biochemistry, 2003. **42**(1): p. 2-10.
44. Pudasaini, A., K.K. El-Arab, and B.D. Zoltowski, *LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling.* Front Mol Biosci, 2015. **2**: p. 18.
45. Salomon, M., et al., *An optomechanical transducer in the blue light receptor phototropin from Avena sativa.* Proc Natl Acad Sci U S A, 2001. **98**(22): p. 12357-61.
46. Herrou, J. and S. Crosson, *Function, structure and mechanism of bacterial photosensory LOV proteins.* Nat Rev Microbiol, 2011. **9**(10): p. 713-23.

47. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch.* Science, 2003. **301**(5639): p. 1541-4.
48. Strickland, D., K. Moffat, and T.R. Sosnick, *Light-activated DNA binding in a designed allosteric protein.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10709-14.
49. Wu, Y.I., et al., *A genetically encoded photoactivatable Rac controls the motility of living cells.* Nature, 2009. **461**(7260): p. 104-8.
50. Hahn, K.M. and B. Kuhlman, *Hold me tightly LOV.* Nat Methods, 2010. **7**(8): p. 595, 597.
51. Huang, Y., et al., *Biological functions of microRNAs: a review.* J Physiol Biochem, 2011. **67**(1): p. 129-39.
52. Chen, Y. and G. Varani, *Engineering RNA-binding proteins for biology.* Febs Journal, 2013. **280**(16): p. 3734-3754.
53. Perez-Pinera, P., D.G. Ousterout, and C.A. Gersbach, *Advances in targeted genome editing.* Curr Opin Chem Biol, 2012. **16**(3-4): p. 268-77.
54. Abudayyeh, O.O., et al., *RNA targeting with CRISPR-Cas13.* Nature, 2017. **550**(7675): p. 280-284.
55. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.
56. Czech, M.P., M. Aouadi, and G.J. Tesz, *RNAi-based therapeutic strategies for metabolic disease.* Nat Rev Endocrinol, 2011. **7**(8): p. 473-84.
57. Auweter, S.D., F.C. Oberstrass, and F.H. Allain, *Sequence-specific binding of single-stranded RNA: is there a code for recognition?* Nucleic Acids Res, 2006. **34**(17): p. 4943-59.
58. Yagi, Y., T. Nakamura, and I. Small, *The potential for manipulating RNA with pentatricopeptide repeat proteins.* Plant J, 2014. **78**(5): p. 772-82.
59. Wickens, M., et al., *A PUF family portrait: 3'UTR regulation as a way of life.* Trends Genet, 2002. **18**(3): p. 150-7.
60. Kiani, S.J., et al., *PUF Proteins: Cellular Functions and Potential Applications.* Curr Protein Pept Sci, 2017. **18**(3): p. 250-261.
61. Edwards, T.A., et al., *Structure of Pumilio reveals similarity between RNA and peptide binding motifs.* Cell, 2001. **105**(2): p. 281-9.
62. Filipovska, A., et al., *A universal code for RNA recognition by PUF proteins.* Nat Chem Biol, 2011. **7**(7): p. 425-7.
63. Quenault, T., T. Lithgow, and A. Traven, *PUF proteins: repression, activation and mRNA localization.* Trends Cell Biol, 2011. **21**(2): p. 104-12.
64. Goldstrohm, A.C., et al., *PUF proteins bind Pop2p to regulate messenger RNAs.* Nat Struct Mol Biol, 2006. **13**(6): p. 533-9.
65. Cho, P.F., et al., *A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP.* Cell, 2005. **121**(3): p. 411-23.
66. Arvola, R.M., et al., *Combinatorial control of messenger RNAs by Pumilio, Nanos and Brain Tumor Proteins.* RNA Biol, 2017: p. 1-12.
67. Pique, M., et al., *A combinatorial code for CPE-mediated translational control.* Cell, 2008. **132**(3): p. 434-48.
68. Suh, N., et al., *FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline.* Genetics, 2009. **181**(4): p. 1249-60.
69. Galgano, A., et al., *Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system.* PLoS One, 2008. **3**(9): p. e3164.
70. Gu, W., et al., *A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization.* Genes Dev, 2004. **18**(12): p. 1452-65.
71. Abil, Z. and H. Zhao, *Engineering reprogrammable RNA-binding proteins for study and manipulation of the transcriptome.* Mol Biosyst, 2015. **11**(10): p. 2658-65.
72. Ozawa, T., et al., *Imaging dynamics of endogenous mitochondrial RNA in single living cells.* Nat Methods, 2007. **4**(5): p. 413-9.

73. Cooke, A., et al., *Targeted translational regulation using the PUF protein family scaffold.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(38): p. 15870-15875.
74. Cao, J., et al., *Bidirectional regulation of mRNA translation in mammalian cells by using PUF domains.* Angew Chem Int Ed Engl, 2014. **53**(19): p. 4900-4.
75. Wang, Y., et al., *Engineering splicing factors with designed specificities.* Nat Methods, 2009. **6**(11): p. 825-30.
76. Zhang, W., et al., *Treatment of type 1 myotonic dystrophy by engineering site-specific RNA endonucleases that target (CUG)(n) repeats.* Mol Ther, 2014. **22**(2): p. 312-320.
77. Grimes, J.A., et al., *The co-repressor mSin3A is a functional component of the REST-CoREST repressor complex.* J Biol Chem, 2000. **275**(13): p. 9461-7.
78. Shimojo, M. and L.B. Hersh, *Characterization of the REST/NRSF-interacting LIM domain protein (RILP): localization and interaction with REST/NRSF.* J Neurochem, 2006. **96**(4): p. 1130-8.
79. Wang, X., et al., *Modular recognition of RNA by a human pumilio-homology domain.* Cell, 2002. **110**(4): p. 501-12.
80. Gerber, A.P., D. Herschlag, and P.O. Brown, *Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast.* PLoS Biol, 2004. **2**(3): p. E79.
81. Hall, T.M., *Expanding the RNA-recognition code of PUF proteins.* Nat Struct Mol Biol, 2014. **21**(8): p. 653-5.
82. Cheong, C.G. and T.M. Hall, *Engineering RNA sequence specificity of Pumilio repeats.* Proc Natl Acad Sci U S A, 2006. **103**(37): p. 13635-9.
83. Adamala, K.P., D.A. Martin-Alarcon, and E.S. Boyden, *Programmable RNA-binding protein composed of repeats of a single modular unit.* Proc Natl Acad Sci U S A, 2016. **113**(19): p. E2579-88.
84. Koh, Y.Y., et al., *Stacking interactions in PUF-RNA complexes.* RNA, 2011. **17**(4): p. 718-27.
85. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
86. Maier, J.A., et al., *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.* J Chem Theory Comput, 2015. **11**(8): p. 3696-713.
87. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* J Comput Chem, 2005. **26**(16): p. 1781-802.
88. Mandel, G., et al., *Repressor element 1 silencing transcription factor (REST) controls radial migration and temporal neuronal specification during neocortical development.* Proc Natl Acad Sci U S A, 2011. **108**(40): p. 16789-94.
89. McClelland, S., et al., *The transcription factor NRSF contributes to epileptogenesis by selective repression of a subset of target genes.* Elife, 2014. **3**: p. e01267.
90. Conforti, P., et al., *In vivo delivery of DN:REST improves transcriptional changes of REST-regulated genes in HD mice.* Gene Ther, 2013. **20**(6): p. 678-85.
91. Opperman, L., et al., *A single spacer nucleotide determines the specificities of two mRNA regulatory proteins.* Nat Struct Mol Biol, 2005. **12**(11): p. 945-51.
92. Chritton, J.J. and M. Wickens, *Translational repression by PUF proteins in vitro.* RNA, 2010. **16**(6): p. 1217-25.
93. Miller, M.A. and W.M. Olivas, *Roles of Puf proteins in mRNA degradation and translation.* Wiley Interdiscip Rev RNA, 2011. **2**(4): p. 471-92.
94. Friend, K., et al., *A conserved PUF-Ago-eEF1A complex attenuates translation elongation.* Nat Struct Mol Biol, 2012. **19**(2): p. 176-83.
95. Kashiwabara, S., et al., *Non-canonical poly(A) polymerase in mammalian gametogenesis.* Biochim Biophys Acta, 2008. **1779**(4): p. 230-8.
96. Kwak, J.E., et al., *Mammalian GLD-2 homologs are poly(A) polymerases.* Proc Natl Acad Sci U S A, 2004. **101**(13): p. 4407-12.
97. Eckmann, C.R., C. Rammelt, and E. Wahle, *Control of poly(A) tail length.* Wiley Interdiscip Rev RNA, 2011. **2**(3): p. 348-61.

Appendix

Article published by Stefania Criscuolo during the PhD course

Regulation of neural gene transcription by optogenetic inhibition of the RE1-silencing transcription factor.

***Paonessa F**, **Criscuolo S**, Sacchetti S, Amoroso D, Scarongella H, Pecoraro Bisogni F, Carminati E, Pruzzo G, Maragliano L, Cesca F, Benfenati F,* PNAS 2016

Abstract

Optogenetics provides new ways to activate gene transcription; however, no attempts have been made as yet to modulate mammalian transcription factors. We report the light-mediated regulation of the repressor element 1 (RE1)-silencing transcription factor (REST), a master regulator of neural genes. To tune REST activity, we selected two protein domains that impair REST-DNA binding or recruitment of the cofactor mSin3a. Computational modeling guided the fusion of the inhibitory domains to the light-sensitive Avena sativa light-oxygen-voltage-sensing (LOV) 2-phototrophin 1 (AsLOV2). By expressing AsLOV2 chimeras in Neuro2a cells, we achieved light-dependent modulation of REST target genes that was associated with an improved neural differentiation. In primary neurons, light-mediated REST inhibition increased Na(+)-channel 1.2 and brain-derived neurotrophic factor transcription and boosted Na(+) currents and neuronal firing. This optogenetic approach allows the coordinated expression of a cluster of genes impinging on neuronal activity, providing a tool for studying neuronal physiology and correcting gene expression changes taking place in brain diseases.