International Ph.D. Program in Biomolecular Science
Department of Cellular, Computational and Integrative Biology
CIBIO

XXXIII cycle

CHARACTERIZATION OF A NEW PUTATIVE ELAV-LIKE BINDING PROTEIN IN *ACINETOBACTER BAUMANNII*

Tutor: Prof. ALESSANDRO PROVENZANI
Department of Cellular, Computational
and Integrative Biology – CIBIO
Laboratory of Genomic Screening

Ph.D. Thesis of
CATERINA CIANI
Department of Cellular, Computational and Integrative Biology – CIBIO
Laboratory of Genomic Screening
Academic Year 2019-2020

Original authorship

Declaration:

I, Caterina Ciani, confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

Post-transcriptional regulations (PTRs) have always been considered features of organisms with higher complexity. However recently, the interest toward the post-transcriptional mechanisms in prokaryotes increased.

The bacterial proteome is much more complex compared to the genome size, suggesting a tight and articulate regulation of proteins production, extremely important for the bacterial adaptation to an always changing environment. Bacterial PTRs are responsible of modulation of mRNA stability and decay, translation initiation and elongation, modulation of the access of ribosome to the ribosome binding site and control of termination of the transcript. The main actors in the PTRs are small non-coding RNA (responsible of the inhibition of the transcription) and RNA binding proteins (RBPs), which modulate the translation and half-life of the mRNA.

RBPs, are particularly of my interest since I wanted to find a possible orthologous of the eukaryotic Elav-like (Elavl) family of proteins in *Acinetobacter baumannii*. Elav-like proteins are present in all metazoans and are characterized by two highly conserved sequences: RNP-1 (a quite well conserved hexamer) and RNP-2 (a really well conserved octamer) that are responsible of binding to the mRNA. Each species has a different number of Elavl paralogous that is totally independent from the complexity of the organisms, suggesting a more ancient origin. In particular, I focused on the human paralog HuR (human antigen R). HuR is characterized by three RNA Recognition motif (RRM) -domains, is ubiquitously expressed and is mainly localized into the nucleus (where it is responsible of maturation of the mRNA), but under stress stimuli, can shuttle into the cytoplasm where protect the target mRNA from degradation, by binding AU/U rich sequences (ARE sequences). Its high concentration into the cytoplasm can lead to the overexpression of oncogenes and pro-tumorigenic factors.

The choice of *Acinetobacter baumannii* comes from the increasing worldwide concern toward this pathogen that is becoming multidrug resistant. Indeed, in Italy, more the 50% of nosocomial infections are caused by *A. baumannii*. I found a putative protein (*AB*-Elavl), composed by a single RRM domain endowed with similar features of the eukaryotic RRM domain as the presence of a quite well conserved RNP-2 and a less conserved RNP-1. I expressed this protein with recombinant tools and confirmed the production of the protein in the host by western blot and mass spectrometry. I evaluated the binding activity of *AB*-Elavl testing the EC_{50} and the Kd with different biochemical assays (EMSA, AlphaScreen and HTRF-

FRET) toward three different RNA sequences, in order to test the specificity. By X-RAY and NMR, I confirmed the folded structure that can be overlapped to the HuR's one and the interaction with the probes tested, highlighting the presence of binding, but with different specificity. I also tested some small molecules developed for interfering in the binding of HuR with the target sequence and found a possible compound able to interact with *AB*-Elavl, by disrupting the binding with the target probe.

All these results suggest an ancient origin of the metazoans' Elavl family of proteins that probably share a common ancestor with *AB*-Elavl. More studies should be performed to better understand the role of *AB*-Elavl in *A. baumannii* as well as in other bacteria. In fact, I found the presence of other ARE sequence-binding proteins also in *Pseudomonas aeruginosa*. Interesting would be to check the presence of this protein in all the multidrug resistant ESKAPE bacteria.

Table of Content

Abbreviation and Acronyms

5'UTR	5' Untraslated region
3'UTR	3' Untraslated region
A. baumannii	*Acinetobacter baumannii*
AB-Elavl	*Acinetobacter baumannii* embryonic lethal abnormal vision-like
ARE	AU/U rich elements
ASCH	Activating signal cointegrator homology
Bi-ARE	Biotinylated ARE sequences
BSA	Bovine serum albumin
CSD	cold shock domain
dsRBD	Double strand RNA binding protein
EC_{50}	Half maximal effective concentration
E. coli	*Escherichia coli*
E. faecium	*Enterococcus faecium*
ELAVL	Embryonic Lethal Abnormal Visual system - like
ESKAPE	*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacter* species
FNE	Found in neurons
hnRNP	Heterogeneous ribonucleoprotein particle
HuR	Human antigen R
KH	K homology domain
K. Pneumoniae	*Klebsiella pneumoniae*
eIF	Initiation factor
IP	Immunoprecipitation
LUCA	Last universal common ancestor
MDR	Multidrug resistant
MW	Molecular weight
mRNA	messenger Ribonucleic Acid
miRNA	micro–Ribonucleic Acid
nELAVL	neuronal Embryonic lethal abnormal vision-like

OB	Oligonucleotide/oligosaccharide binding superfamily
OmpA	Outer membrane protein A
P. aeruginosa	Pseudomonas aeruginosa
PEI	Polyethyleneimine
PLD	Phospholipase D
PTR	Post-transcriptional regulation
rAB-Elavl	recombinant *Acinetobacter baumannii* embryonic lethal abnormal vision-like
RBP	RNA binding protein
RBS	Ribosome binding site
RNP1-RNP2	Ribonucleoprotein sequence 1 and 2
RRM	RNA Recognition Motif
RBP9	RNA binding protein 9
S. aereus	*Staphylococcus aureus*
SM	Small molecule
sRNA	Small non-coding RNA
TNFα	Tumor Necrosis Factor Alpha
WHO	World Health Organization

Introduction

1. Post transcriptional mechanisms

1.1. Evolution of post-transcriptional mechanisms

The most ancient system that connect all the organisms from bacteria to mammals, is the RNA metabolism, seen as all the processes that involve RNA: from transcription to degradation, passing through the post-transcriptional mechanisms responsible of maturation and regulation of the mRNA (Anantharaman et al., 2002).

Post-transcriptional regulation (PTR) was for long believed to be a distinctive characteristic belonging to organisms with a high complexity, but a relatively small genome size. Indeed, PTR can differently modulate protein production, according to the environmental inputs (in mammals ≈60-80% of RNA can be differently recruited by ribosomes, according to the different stimuli) (Halbeisen et al., 2008; Van Assche et al., 2015). RNA metabolism is tightly controlled by two main factors: *cis*-elements and *trans*-elements. While *cis*-elements are sequences located in the proximity of the controlled sequence, regulating RNA translation and decay, the *trans*-element are RNA binding proteins (RBPs) responsible of recognizing and binding the *cis*-elements, affecting the RNA abundance and stability, the translational efficiency and protein abundancy (Halbeisen et al., 2008; Hör et al., 2018; Schaefke et al., 2018).

Proteins interacting with RNA are 3-11% of all the proteome belonging to the three domains: bacteria, archaea and eukarya. The fact that there are several orthologous shared by organisms so distant in evolution, suggest the presence of a common ancestor. The Last Universal Common Ancestor (LUCA) (Anantharaman et al., 2002; Fox, 2010; Schaefke et al., 2018) is a reconstructed organism, or a population of organisms, connecting the "primordial RNA world" (where the RNA used to have the dominant role), to the modern protein-dominated cellular system. Indeed, 80 different orthologous groups of proteins can be already found in LUCA, suggesting the presence of a distinct functional RNA metabolism composed of ribosomes, accessory apparatus of protein synthesis and RNA binding proteins responsible of its maturation and degradation (Anantharaman et al., 2002).

It is truthful to assume that in the pre-LUCA world, there was a limited set of simple proteins with generic functions, while the specificity of the binding was totally related to the RNA. With evolution, the number of proteins increased thanks to

multiple episodes of genomic duplication. New proteins were able to occupy different functional niches, increasing in the specificity of their function and becoming the main controller of RNA metabolism.

In the post-LUCA phase, although with the advent of eukaryotes there was a burst of innovation, RNA kept conserved many ancient components (Anantharaman et al., 2002; Fox, 2010). This difference is still observable when comparing the RBPs functions between bacteria and eukaryotes. Despite in eukarya it is possible to find RBPs' paralogous with well distinct functions and targets, in bacteria the number of proteins able to interact with RNA are more limited, and less specific.

The evolution of post-transcriptional mechanisms led to the passage from unicellular forms to multicellular organisms without increasing the genomic material. This was possible thanks to the alternative splicing, increased number of proteins responsible for the maturation and control of RNA and reduction of noise in the gene expression performed by non-coding RNAs as miRNAs (Schaefke et al., 2018). Evolution led also to an adjusted RNA decay: RNA became more stable according to the complexity of the organism, passing from a half-life of minutes in bacteria, to days in mammals (Halbeisen et al., 2008; Schaefke et al., 2018).

1.2. Post-transcriptional mechanisms in Prokaryota and Eukaryota

The phenotype of an individual is determined by its unique proteome, from here the importance of mRNA control (Keene & Lager, 2005). Evolution kept more attention on conserving proteins more than genome (Schaefke et al., 2018). The most ultra-conserved genes in humans are the sequences encoding for RNA binding proteins that derive from ancestral RBPs with domains that are still present both in eukaryotes and prokaryotes as: oligonucleotide/oligosaccharide binding (OB) domains, K homology and double strand RNA binding domains (Holmqvist & Vogel, 2018). Until recently it was believed the post-transcriptional mechanisms were a characteristic of eukaryotes, but it was observed how the bacterial proteome is much more complicated compared to the genome size, leading to the study of post-transcriptional mechanisms that, even if less complex, are anyway present also in prokaryotes. PTRs include all the modifications that lead from a pre-mRNA to a mature RNA (as capping of the 5'UTR, polyadenylation of the tail at the 3' UTR, splicing and several possible chemical modifications of the internal sequence, as m^6A), as well as all the processes that control the stability of mRNA (its half-life

and decay) and the regulation and control of translation with several differences between bacteria and Animalia (Day & Tuite, 1998; Picard et al., 2009; Revel & Groner, 1978; Zhao et al., 2016).

A mention should be done for Archaea that are a special dominium: they are considered as a sister lineage to eukaryotes in the Woese's "universal tree" (Woese et al., 1990), but they share some characteristics of prokaryotes, as the presence of a proteoglycan wall or the ability to metabolise nitrogen. Interestingly, the post-transcriptional mechanisms in archaea are more like eukaryotes for the complexity of RNA machinery and spliceosome (Spang et al., 2015).

1.2.1. Eukarya

Post-transcriptional mechanisms start in the nucleus from the pre-mRNA. Indeed, the mRNA has to pass through several maturation steps before it can move in the cytoplasm. The extremities of the mRNA are protected against the exonucleases' activities by adding a 7-methylguanosine cap to the 5'UTR (m^{7G}), and a poli(A) tail to the 3'UTR (Day & Tuite, 1998). mRNA sequence is also modified by splicing events, in which the non-coding sequences are cut (Revel & Groner, 1978) and by chemical modifications (Zhao et al., 2016). Once the mRNA moves out from the nucleus, other post-transcriptional mechanisms control the mRNA stability and decay, as well as the translational control. The half-life of mRNA is controlled by the presence of internal sequences, mainly located at the 3'UTR, that can destabilize the transcript, and by the length of the poli(A) tail (Day & Tuite, 1998). One of the most studied destabilizing sequences is the ARE sequence (an AU/U rich region) which promotes the binding of endonucleases that can cleave the mRNA. While the intrinsic half-life of mRNA in the eukaryotic cell is from minutes to days, the general control of the transcriptome is performed by small non-coding RNAs and RBPs. RBPs can recognize and bind the destabilizing sequences, protecting them from endonucleases. It is hypothesized that RBPs bind specifically genes involved in the same processes in order to produce functionally related proteins (Keene & Lager, 2005; Nishtala et al., 2016). RBPs can also help in the control of translation initiation, by avoiding the binding of the 43S translation complex, responsible for binding to the starting codon and scanning the mRNA sequence for protein production (Day & Tuite, 1998). mRNA stability and translation are coupled in time and space in order to have a strict control on protein production as an answer to external inputs.

1.2.2. Prokarya

In prokaryotes, the half-life of mRNA is extremely shorter compared to the eukaryotic one (the average is around 7 minutes, according to several studies performed at the stationary phase of bacteria) (Picard et al., 2009). This is important because the bacterial cell has to adapt in a fast way to an always changing environment (Martínez & Vadyvaloo, 2014). While in eukaryotes the mRNA is first matured and then moved into the cytoplasm where it can be translated, in bacteria the translation starts when the transcription is still ongoing. There is a tight control of the gene expression thanks to small non-coding RNAs (sRNAs) and RBPs. sRNAs can bind sequences close to the ribosome's binding sites (RBS) with limited complementarity, regulating translational efficiency and RNA stability by avoiding binding of proteins. RBPs are able to change the secondary structure of the mRNA, "hiding" the starting sequences, and to shield the RBSs competing actively with ribosomes and RNAses (Holmqvist & Vogel, 2018).

Our interest is mainly in RNA binding proteins, in particular those that are able to target the ARE sequences.

2. RNA binding proteins and the RNA recognition motif

As previously mentioned, RBPs are extremely important for stabilization and destabilization of mRNA, in order to answer to the cell needs. Evolutionarily derived from a common ancestor, RBPs show similarities and differences in eukaryotes and prokaryotes.

2.1. Eukarya

RBPs in eukaryotes are in general constituted by multiple RNA-binding domains. The best characterized are (C. Oliveira et al., 2017):

- Heterogeneous ribonucleoprotein particle (hnRNP) K-homology (KH domain) formed by three α-helices around the surface of a central antiparallel β-sheet.
- RGG motif, an evolutionary conserved sequence, composed of a repetition of arginine and glycine residues

- C_2H_2 zinc finger proteins, composed of two cysteine and two histidine that coordinate a zinc ion. The conformation adopted by these proteins is small and characterized by two β-strands and one α-helix.
- Double strand RNA binding domain (dsRBD), a conserved protein domain which bind to double stranded or highly structured RNAs
- PUF domain, containing multiple tandem repeats of 35-39 residues recognizing specific RNA bases. It is found in Pumilio family proteins, important in regulation of translation of mRNA target and interaction with mRNA regulatory system.
- PAZ domain, an interesting motif, since it binds with a low affinity single strand RNAs in a sequence independent manner, probably by recognition of the 3'-ends of the target RNA
- RNA recognition motif (RRM) domain, the most abundant domain and the most studied. Present in all form of life, is composed of four β-strands and two α-helices

In our laboratory, we focus on human antigen R (HuR) protein, belonging to the RRM family domain. Hereafter, I will discuss and get deep into details about this group of proteins.

The RRM domain was discovered in the 1980s when it was demonstrated that pre-mRNA is bound to RNA binding proteins. It is composed of mainly 90 amino acids with some peptides highly conserved among all the life kingdoms: RNP1 and RNP2. RNP1 is an octapeptide highly conserved, while RNP2 is an hexapeptide that remains less conserved (fig.1). These two sequences are responsible for binding RNA with a wide range of specificity that often could be in the nM scale. The affinity of the binding is not always related only to the mRNA sequence, but also to the secondary structure of the mRNA. It was demonstrated that the same sequence can be bound with much less efficiency if linearized. In eukaryotes the RRM domain can associate with other different domains, increasing its biological functions. Indeed, this protein domain is involved in lots of different roles in the eukaryotic cell: pre-mRNA processing, splicing, alternative splicing, mRNA stability and editing, mRNA export and translational regulation, degradation. RRM domain folds into an αβ sandwich: $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ structure, with RNP1 and RNP2 localized in the internal strains β_1 and β_3. The loops between the two different secondary structures can have different lengths, while the α-helix1 seems to be responsible for the interaction with other proteins (Maris et al., 2005). HuR protein is composed by a combination of three RRM domains distributed in tandem (fig 2). The first and second domain are

responsible for binding with a high affinity the target RNA, while the third domain is responsible for the interaction with a second HuR protein for dimerization and stabilization of the binding on the RNA (Assoni et al, 2022). HuR characterization and description will be explained in more details in the next chapter.

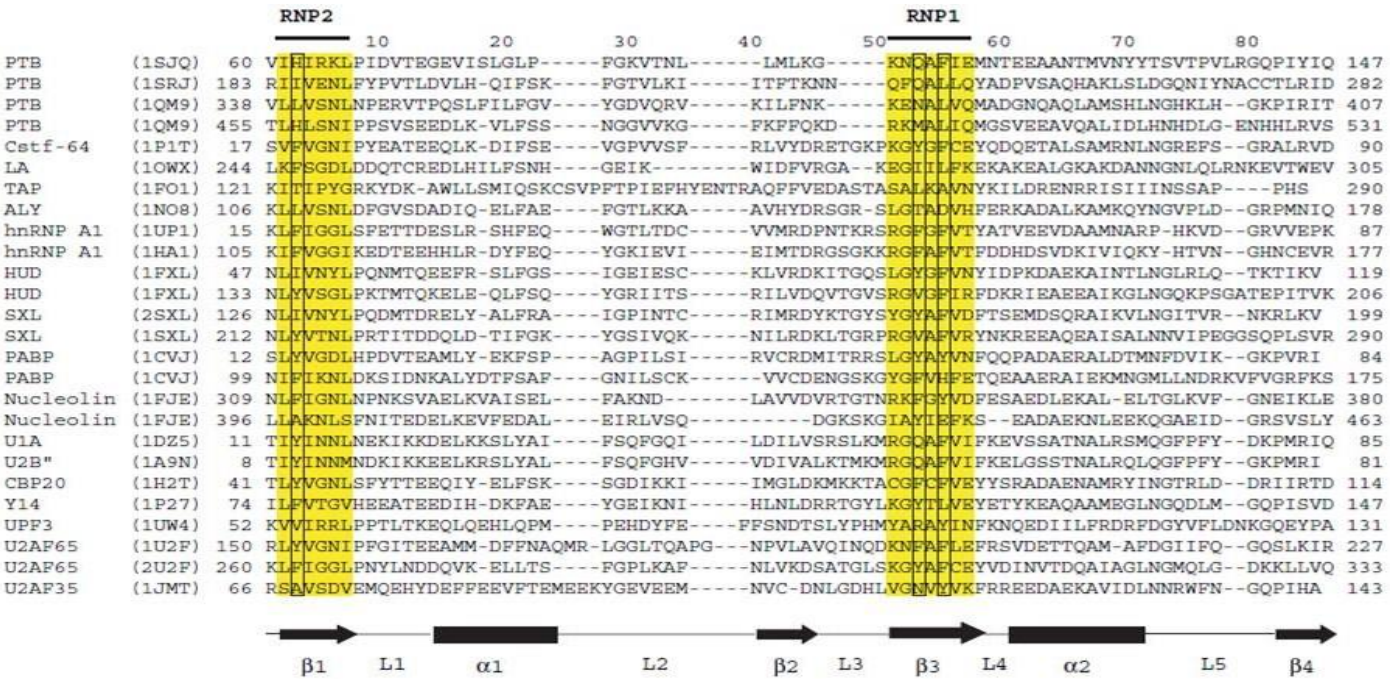
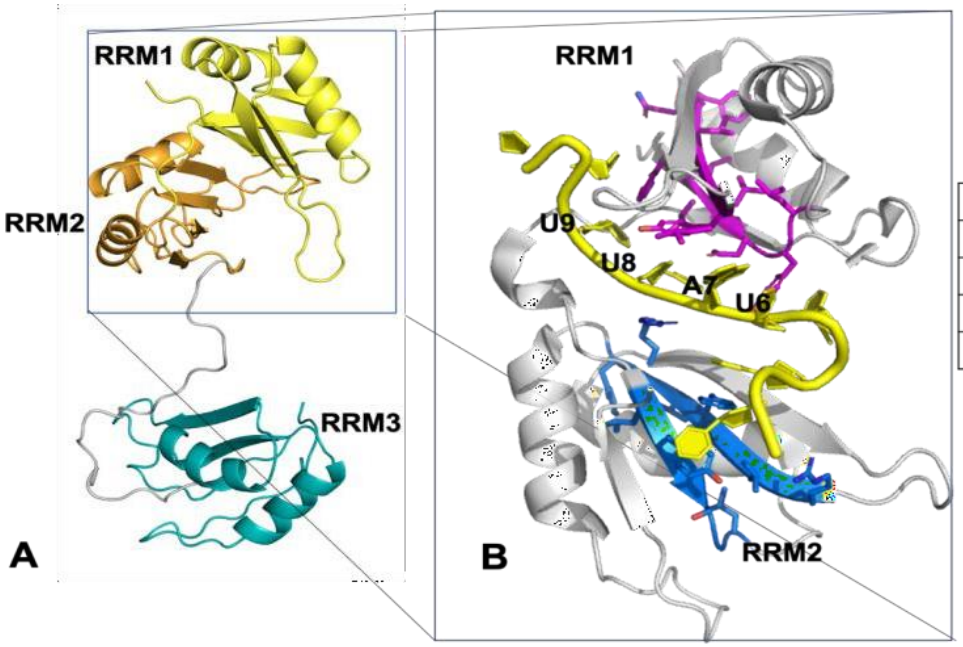


Figure 1. Alignment of a selection of human RRM domains (from Maris et al. 2005). In yellow the sequences RNP1 and RNP2, the amino acid in boxes are the aromatic residues important for RNA binding. In brackets the PDB codes.



Domain	Sequence
RNP1_RRM1	$_{61}$LGYGFVNY$_{68}$
RNP2_RRM1	$_{22}$LIVNYL$_{27}$
RNP1_RRM2	$_{147}$RGVAFIRF$_{154}$
RNP2_RRM2	$_{108}$LYISGL$_{113}$

Figure 2. (From Assoni et al. 2022) A) Representative 3D model of full length HuR (Chimera model obtained with Prime-Schrodinger software); the protein is shown in ribbons with RRM1, RRM2 and RRM3 displayed in yellow, orange and cyan, respectively, and the long basic linker between RRM2 and RRM3 in gray. B) Co-crystal structure of the tandem RRM1 and RRM2 HuR-mRNA complex (pdb code 4ED5); the protein backbone is shown in gray, the RNP1 and RNP2 sequences of the RRM1 domain in magenta, the RNP1 and RNP2 sequences of the RRM2 domain in blue, mRNA in yellow.

2.2. Prokarya

There are several similarities between bacterial and eukaryotic RBPs: from the role (structural and regulatory role involved in synthesis, modification, translation, processing and degradation of the mRNA) to the structures of the domains:

- S1 domain and cold-shock domain (CSD) of the oligonucleotide/oligosaccharide binding (OB) superfamily are conserved from bacteria to human: globular structure of around 70 amino acids and with two highly conserved sequences (Heinemann, 2021).
- Sm and Sm-like domain, also present both in eu- and pro-karyotes are mainly composed of homohexamer subunits (Murina & Nikulin, 2011)
- RNA recognition motif (RRM) domain characterized by conserved RNP1 and RNP2, derived from a common ancestor with eukaryotes
- K homology (KH) domain, arose early in the evolution, is the only case of protein that conserve the sequence, but fold in different way in eukaryota and prokaryota (Grishin, 2001)
- Double stranded RNA-binding domain (dsRBD), 65–70 amino acids, interesting for its ability to recognize secondary structures more than sequences of the target RNA.
- PAZ and PIWI domain, found in two families of proteins that are involved in post-transcriptional gene silencing

Differences are also widely present: while eukaryotic RBPs are mainly composed of more than one domain in tandem leading to an exponential increase in the affinity for the target RNA (Maris et al., 2005), in bacteria there is most often just one single domain, with a less specialized binding selectively and lower affinity. Bacterial RBPs can form big complexes that are not present in eukaryotes, like the degradosome (a complex of RBPs able to drive endo- and exo-nucleases to the site of cleavage). RNA binding proteins in bacteria are not as well-known as in eukaryotes: many proteins are still unknown or with unknown functions. Recently it was discovered some of them can also show a bifunctionality: they are not only involved in RNA metabolism, but also in other totally unrelated mechanisms (e.g., Aconitase: an iron containing tricarboxylic acid cycle protein that in *Bacillus subtilis* is responsible for post-transcriptional regulation as well as sporulation) (Holmqvist & Vogel, 2018).

I am mainly interested in the RRM domain in bacteria. Between all the domains, this one seems to be not well characterized in prokaryotes. Interestingly, most Cas

proteins of the system CRISPR-Cas (the main prokaryotic adaptive immunity system) are RRM domain-containing proteins (Sharma et al., 2018). Another example of RRM domain in bacteria is in the DEAD-box family of proteins, and in particular in the DbpA subfamily, responsible for structural rearrangement of RNA (Christopoulou & Granneman, 2021). This seems to suggest an important role of the RRM domain in the modification and modulation of the target RNA.

3. ELAV-like family

The *Elav-like* (*Elavl*) family is a group of genes encoding for proteins able to bind ARE regions localized at the 3'UTR of target mRNA with a high affinity. They were discovered in the late 80s in *Drosophila melanogaster* in a study for neuronal development. The first gene found was the *embryonic lethal abnormal visual system* (*Elav*), an essential gene for the nervous system formation and maintenance (Robinow & White, 1991). Subsequent studies showed Elav has other two paralogous genes: Rbp9 and Fne. Rbp9 (RNA binding protein 9) is the second identified member of the Elav-like gene family and, contrarily to *Elav* gene is not essential in the knockout mutant, but show motor deficit and female sterility phenotype. Fne (found in neuron) has different molecular function that are not overlapping to those of its paralogous proteins, since is localized in the cytoplasm (while Elav and Rbp9 are mainly nuclear), with vital null mutants (Colombrita et al., 2013). In humans, ELAVL family are characterized by the presence of three RRM domains: RRM1 and RRM2 toward the N-terminal, and RRM3 separated from the first two by a hinge region that allow the passage of the ELAVL through the nuclear membrane.

3.1. Elav-like family phylogenies

The Elav-like proteins are widely spread across all metazoans and are characterized by a high degree of conservation between different species (identity score: >45%) (fig. 4) (Samson, 2008). Indeed, they are composed by three distinct RRM domains, which individually contribute to the mRNA binding, with the most conserved sequences at the level of RNP1 and RNP2 (Good, 1995; Zucal et al., 2015). Interestingly, the number of paralogous genes in the different species vary from one to four without any direct link with the complexity, size, development and brain structure of the organism. This suggests that these genes could have been developed

before the evolution of metazoans (Samson, 2008; Tang et al., 2012). Mammals are characterized by four paralogous genes: HuR (or HuA), Hel-N1 (or HuB), HuC, and HuD, with different roles and cellular localization. Hel-N1, HuC and HuD are specifically located in the neural tissue (Hel-N1 is also present in the testis and ovary) and are mainly localized in the cytoplasm, even though they can easily shuttle into the nucleus under certain stimuli. HuR, conversely, is ubiquitous expressed in all human tissues and is mainly localized in the nuclear compartment, but, as its paralogous genes, can shuttle, moving to the cytoplasm as response to stress stimuli. The neural ELAVL (nELAVL) shares redundant roles into the cell and their expression is tightly regulated in tissue- and cell-specific manner throughout the development. In fact, the knockout of the HuD gene in mice determines only a mild phenotype due to a transient delay of cranial nerve growth in the early stages of development. This delay leads to motor defects and reproductive failure in adulthood (Colombrita et al., 2013), while HuR knockout causes embryonic lethality in mice. This huge difference in phenotype could be due to the fact that HuD role is in some way counterbalanced by the other two nELAVL, while HuR's activity is involved in key processes with high relevance for all the organism (Akamatsu et al., 2005; Assoni et al., 2022; Good, 1995; Kasashima et al., 1999; Katsanou et al., 2009; Lang et al., 2017; Samson, 2008).

I am mainly interested in one of the human paralogous: HuR

3.1.1. HuR

HuR is mainly localized in the nucleus where it is responsible for promotion of splicing of pre-mRNA and of alternative polyadenylation (Bakheet et al., 2018; Izquierdo, 2008). In presence of stress stimuli HuR acts as a shuttle helping the export of mature target mRNA into the cytoplasm, where it acts as a stabilization factor and translation regulator. HuR regulates the fate of thousands of coding and noncoding RNAs containing ARE sequences (fig. 3). HuR is able to self-regulate its expression, by binding its own mRNA, avoiding degradation by miRNAs (Guo et al., 2009; Marasa et al., 2010). Its activity is also modulated by multiple post-translational modifications as phosphorylation and methylation, affecting its affinity for the targets and its localization (Noh et al., 2016). Since it is crucial for several cellular mechanisms (from spermatogenesis to lipolysis and for the proper development during embryogenesis), its depletion results in development of many

different diseases and embryonic lethality within 10 days (Assoni et al., 2022). HuR high concentration in cytoplasm is correlated also with overexpression of oncogenes and pro-tumorigenic factors leading to cancer onset and progression as renal, urothelial and esophageal carcinomas, and small-lung cancer (Assoni et al., 2022; Cha et al., 2013).

Given its physiological importance, I envisioned that an investigation about its phylogeny would have revealed further details about its pleiotropic roles within the cell. Indeed, nothing is known about HuR's orthologous in the prokaryotic world and, in particular, in bacteria.

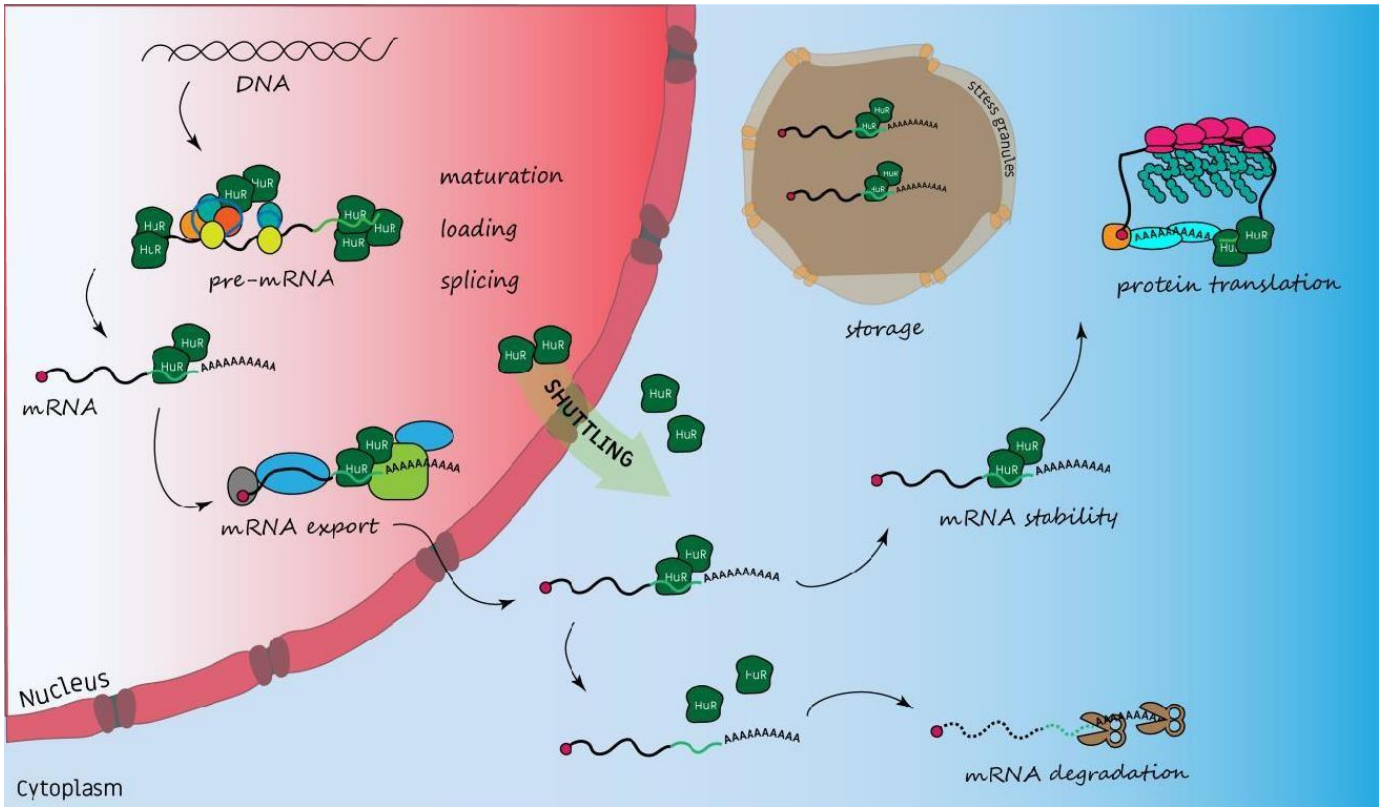


Figure 3. Schematic representation of HuR functions within the cell (from Assoni et al., 2022). Inside the nucleus, HuR (green) interacts with pre-mRNA introns (light blue thin lines) and untranslated regions (light green lines), promoting mRNA maturation events, and with splicing factors (round colored dots), guiding splicing and alternative splicing events. HuR helps the mRNA export in the cytoplasm by interacting with transportation factors (colored shapes). In the cytoplasm HuR is responsible of mRNA stability and storage and promotion of target translation.

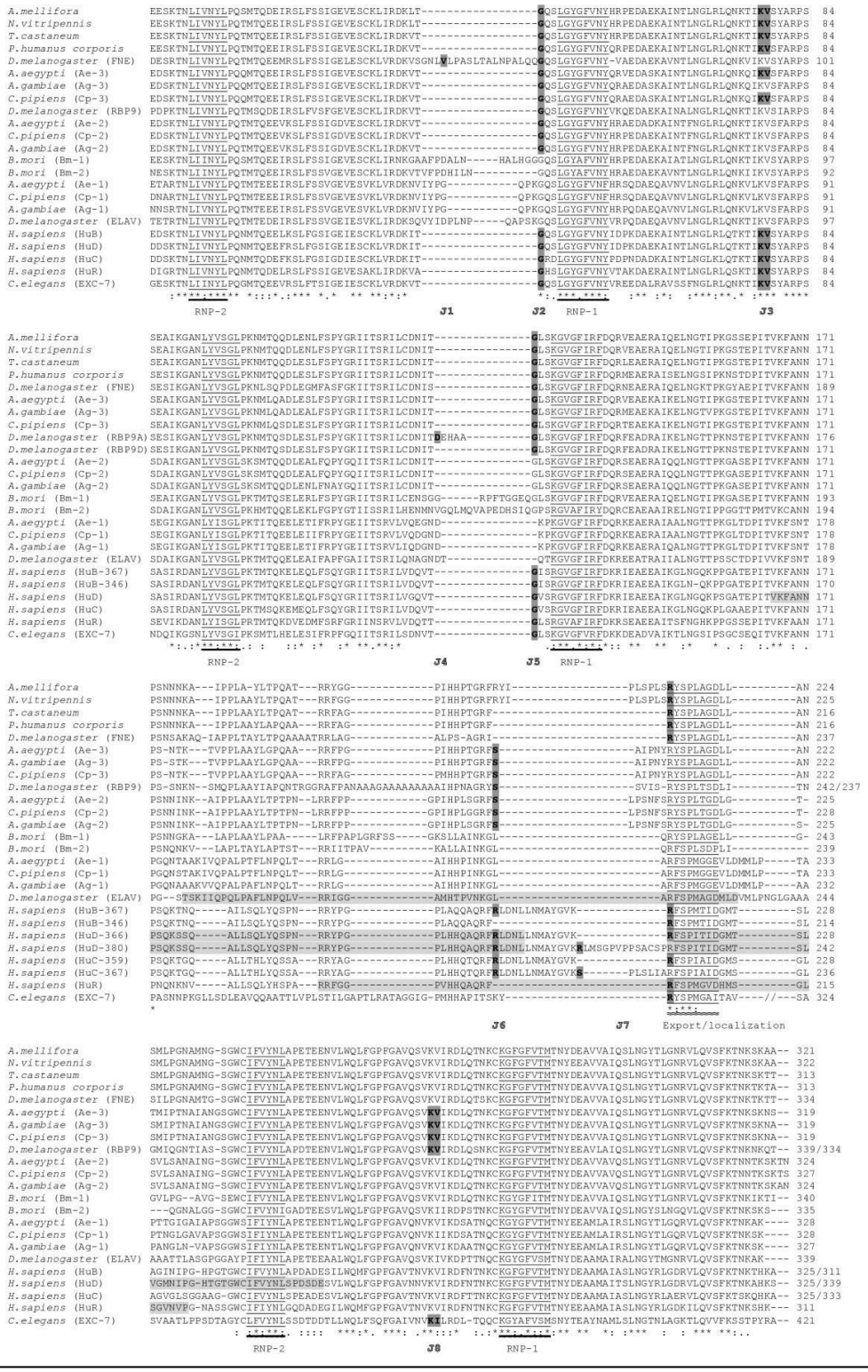


Figure 4. Protein sequence comparison among 27 Elav-like proteins forms (from Samson, 2008). Alternative protein forms are included, specifically for Drosophila Rbp9 (A and D) and three of the human proteins (HuB, HuC and HuD, where HuX-n refers to the n amino acid long form of the HuX protein). "*****" indicate that amino acids are identical in all 27 sequences, "**:**" and "**.**" respectively indicate conserved and semi-conserved substitutions. The octamer RNP-1, the hexamer RNP-2, and the octamer of the hinge region are underlined.

4. Antibiotic resistance: a threat to health systems across the globe

4.1. ESKAPE bacteria

In the 2020, reports estimated a number of infections in the United States of more than 2 million, with a death toll of 29,000 victims per year and a cost of more than $4.7 billion, while it is not possible to have a worldwide estimation since there is a lack of reports (Benkő et al., 2020). Persistent use of antibiotics since the turn of the '90s, self-medication and exposure of infections in hospitals have promoted the emergence of multidrug resistant (MDR) bacteria (resistant to at least three antibiotic classes) (Benkő et al., 2020; Mulani et al., 2019; D. M. P. De Oliveira et al., 2020). In the ESKAPE group are classified all the pathogens able to "escape" the action of most antimicrobial agents currently available: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species. The mechanisms of resistance can be classified into four main classes: (i) inactivation or alteration of the antimicrobial molecules by producing enzymes able to destroy or neutralize the antibiotic agents (as β-lactamases), (ii) modification of the antibiotic target site leading to a reduction of the affinity or even to the prevention of the binding, (iii) reduction of antibiotic penetration and accumulation, by downregulating or losing the outer membrane protein channels (porins) responsible of the penetration of hydrophilic agents (as most of the antimicrobial molecules), and overexpression of the efflux pumps (able to extrude drugs out of the cell), (iv) biofilm production which protect colonies from antibiotic activity (although this mechanism differs from the previous ones since it is a phenotypical change rather than genotypical) . There are also other strategies that can be applied, such as a reversible arrest of the growth state when the colony enters in contact with the drugs or the internalization and survival of the bacteria into the host cells forming a latent infection. These mechanisms are mainly acquired by mutation and acquisition of mobile genetic elements (as transposons and plasmids) (D. M. P. De Oliveira et al., 2020). Every year the number of antibiotics effective against ESKAPE is declining, opening to a new era in which there is not going to be any limit to infections of more and more aggressive bacteria (Mulani et al., 2019). In 2017 the World Health Organization (WHO) published a list of pathogens for which there is the urgent need of development of new antimicrobial agents. They described three categories of critical, high and medium priority according to the drug resistance of the pathogen. Carbapenem resistant *A. baumannii*, *P. aeruginosa*, *K. pneumoniae* and *Enterobacter* species are listed in the critical priority class of pathogens; whereas vancomycin resistant *E. faecium* and methicillin and vancomycin resistant *S. aureus* are in the high priority group ("WHO Publishes List of Bacteria for Which New

Antibiotics Are Urgently Needed," 2017). Development of novel therapeutics to treat drug resistant infections, especially those caused by ESKAPE pathogens is the need of the hour.

I focused on one bacterium of the ESKAPE group: *Acinetobacter baumannii*

4.1.1. Acinetobacter baumannii

Acinetobacter species contain a great diversity of bacteria with some being just environmental organisms, while others are human pathogens with high resistance to antibiotics responsible of outbreaks (in 2015, it was reported that in several European countries, more than 50% of nosocomial infections are caused by *Acinetobacter* species (fig. 5) (Peleg et al., 2012; "WHO Publishes List of Bacteria for Which New Antibiotics Are Urgently Needed," 2017).

Between all the species, *Acinetobacter baumannii* is a gram-negative opportunistic bacterium that causes infections with serious morbidity and mortality (Gallagher et al., 2015). Its high resistance to antibiotics developed recently, as the first report of *A. baumannii* in an intensive care unit goes back to the 1960s. The genus *Acinetobacter* was first isolated from the soil in 1911, but it became a serious hospital threat in diverse geographical locations worldwide in few years. In the 1970s it was easily controlled using β-lactams and sulphonamides. Unfortunately, since then, antibiotic resistance has just increased, with few developments of new drugs that could control this infection (Gonzalez-Villoria & Valverde-Garduno, 2016). In the 1990s it was found an association between ventilator use and increased mortality with strains resistant to more than three antibiotics, bringing *Acinetobacter baumannii* between the MDR strains. In 1985 the development of carbapenems as a therapeutic option gave the hope to be able to control the outbreak of this pathogen, but already in the same year there were reports of resistant strains. The more and more urgency to control the spread pushed to new studies for understanding the mechanism of infection of *A. baumannii*. It was found there is a small genomic region of 86 kilobases containing 45 resistance genes.

It is still not totally elucidated how this bacterium is able to enter in the host cells (Asensio, 2021); it was demonstrated that the interaction pathogen-cell is receptor-dependent through key factors as the Outer Membrane Protein A (OmpA) (that is fundamental for cell invasion since knockout for this protein are unable to infect), Phospholipase D (PLD) and autotransporter adhesion Ata, but the general comprehension of the human-pathogen interaction is made more complicated by *A.*

baumannii's ability to create latent infections, remaining dormant into vesicles in the host cells, protected from antibiotic and immune-response exposures (Asensio, 2021).
Asensio et al, showed that *baumannii* can survive and replicate in the human epithelial cells, leading to a rearrangement of gene expression as cytoskeletal remodelling, activation of apoptosis and stress response and down-regulation of chemokines.

In this context, the comprehension of the physiology of *Acinetobacter baumannii* is extremely important. Indeed, it is predicted that in 2050, 10 million people per year will die from *A. baumannii* infections if there are no improvements in the treatments (Regulation, 2017). RBPs were proven to be overexpressed in resistant strains such as enolase, RNAse E and NusA, all involved in mRNA processing and gene expression modulation (Quendera et al., 2020).
Among the most studied RBPs in *A. baumannii*, there are Hfq and Csr(A). Hfq (an Sm-like domain containing-protein) is one of the main proteins of the post-transcriptional machine shared by gram-negative bacteria. Hfq is an RNA chaperone that modulates gene expression by helping the bind of sRNAs to their target and has an important role as a virulence factor since its knockout leads to reduced growth rate and stress tolerance (Moll et al., 2003; Sharma et al., 2018). Csr(A) is able to bind as a symmetrical homodimer to the 5'UTR of mRNA, regulating the amino acids metabolism. Its deletion leads to bacterial inability to use amino acids for carbon uptake and, consequently, it avoids the growth and adaptation to environments as the urinal one, full of peptides and alcohol. These results are important to delineate the main role played by Csr(A) in urinary tract infections and, mainly, to underline the importance of RBPs in pathogenicity (Gallagher et al., 2020; P. Wang et al., 2021).

Taken together, all these data show the importance of more studies of this pathogen that is creating a great concern worldwide.

Figure 5. Multidrug resistant *Acinetobacter baumannii* species in the European region in 2015 (from the Central Asian andEastern European Surveillance of Antimicrobial Resistance (CAESAR, @WHO 2016)).

Hypothesis and Aim

The family of Elav-like proteins is widely spread in all metazoans, usually composed of several members with redundant roles, suggesting they derive from a common single ancestor gene. HuR, specifically, differs from its paralogous playing a fundamental role into the human organism, since its knockout leads to embryonic death, and being expressed in all cell types.

It is widely believed that the Elavl family developed "recently" with metazoans. However, the RRM domain is widespread also in prokarya and I wondered whether an RRM containing protein, like HuR, existed in *Acinetobacter baumannii* and which role the hypothetical orthologue may play. I started from *A. baumannii* since this bacterium is characterized by an easy development of resistance to treatments and good capabilities to survive in harsh conditions.

Therefore, I aimed at identifying RRM containing, ELAV-like, genes in *Acinetobacter* and investigating their function. Indeed, I reasoned that if HuR is endowed with stress response function, bacterial homologue may play similar functions with similar mechanisms. I therefore searched for a gene of interest and elucidated its biochemical function. The physiological characterization is currently ongoing.

Results

1. Homology search of the HuR protein within the *Acinetobacter baumannii* genomes identifies the homologous protein *AB*-Elavl.

In collaboration with doc. Alfonso Esposito from lab of Microbial Genomic (CIBIO), We searched using tblastn the human HuR protein on the RefSeq Genome Database, limiting the search to the species *Acinetobacter baumannii*, and found 25 hits. All the subject sequences were the same orthologous protein from different *Acinetobacter baumannii* genomes. In 11 cases, the region of homology on HuR query sequence was limited to the RRM1 (amino acids 20-98 on the human protein); the remaining 14 hits instead covered the RRM3 (the amino acids 244-322, fig 1A). The percentage of protein identity ranged 32.90 - 35.48% (on average 34.08 ± 1.00%) for the alignment covering the RRM1 region, and it ranged 31.94 - 46.00% (on average 35.70 ± 3.67%) for the alignment covering on the RRM3 (fig 1B). The amino acids matching as similar, ranged 50.67 - 60.26% when the subject was RRM1 (on average 54.11 ± 3.15%), and 51.39 - 60.00% when the subject was RRM3 (on average 53.84 ± 2.66%) (Table 1). The orthologous protein (ELAV-like protein, here named *AB*-Elavl) was present on nearly all the *A. baumannii* genomes (4946 out of 4972 available), this suggests that it belongs to the core genome of this species. Its genomic context included an activating signal cointegrator homology (ASCH) domain-containing protein with RNA-binding properties, starting 8 bp downstream respect to *AB*-Elavl, and an ATP-dependent helicase, ending 74 bp upstream the putative human ELAV homologous (fig 1C).

We searched for homologous proteins to *AB*-Elavl in all bacterial genomes contained in OrthologueDB (a comprehensive database of bacterial and archaeal orthologs) (Whiteside et al., 2013). There were 207 hits in genomes spanning 12 bacterial phyla suggesting that the protein is frequently found in the prokaryotic kingdom. The sequence lengths of the bacterial Elavl proteins ranged 78-241 aa with an average length of 106.68 ± 30.06, multiple sequence alignment consisted of a 375aa alignment. The bacterial proteins homologous to HuR shared an identity score ranging 9.1% (*Dyadobacter fermentans* versus *Shewanella pealeana*) – 99.3% (between two *Shewanella* spp.), sharing an average id of 31.1 ± 10.8%.

The visual evaluation of the multiple sequence alignment suggested that there were conserved regions within the bacterial homologous of HuR (fig 1D), so we ran the web tool MEME for motif discovery. We found that there were two motifs which were significantly conserved across all sequences (fig 1E), one had the pattern (I/L)(Y/F/L)YGNL (p-value $3.0e^{-1314}$), the second (K/R)GF(G/A)FVEM (p-value

$3.0e^{-1407}$). Those two patterns match the locations and the order of the ribonucleoprotein motifs RNP-2 and RNP-1 in each of the RRMs in the gene of HuR (Samson, 2008), thus suggesting that *AB*-Elavl protein is endowed with RNA binding ability.

Table 1: Most significant tblastn hits using as query HuR protein sequence on the *Acinetobacter baumannii* genomes, the search was performed as restricted to the reference prokaryotic representative genomes database.

Query	Subject	Description	% identity	% positives	Alignment length	Mismatches	Gap opens	e-value	Bit score
Query_26761	NZ_VMIG01000771.1	Acinetobacter baumannii strain 20821 NODE_805_length_314_cov_0.491979 , whole genome shotgun sequence	37.975	54.43	79	49	0	8.45E-06	51.6
Query_26761	NZ_VMJH01000115.1	Acinetobacter baumannii strain 20858 NODE_138_length_443_cov_0.215190 , whole genome shotgun sequence	38.889	56.94	72	38	2	1.33E-05	52.4
Query_26761	NZ_VMJF01000450.1	Acinetobacter baumannii strain 20856 NODE_475_length_291_cov_0.280488 , whole genome shotgun sequence	39.062	57.81	64	39	0	1.53E-05	50.8
Query_26761	NZ_RFCE01000004.1	Acinetobacter baumannii strain TG22158 Acinetobacter-spp-TG22158_56, whole genome shotgun sequence	34.615	60.26	78	51	0	1.05E-04	52.8

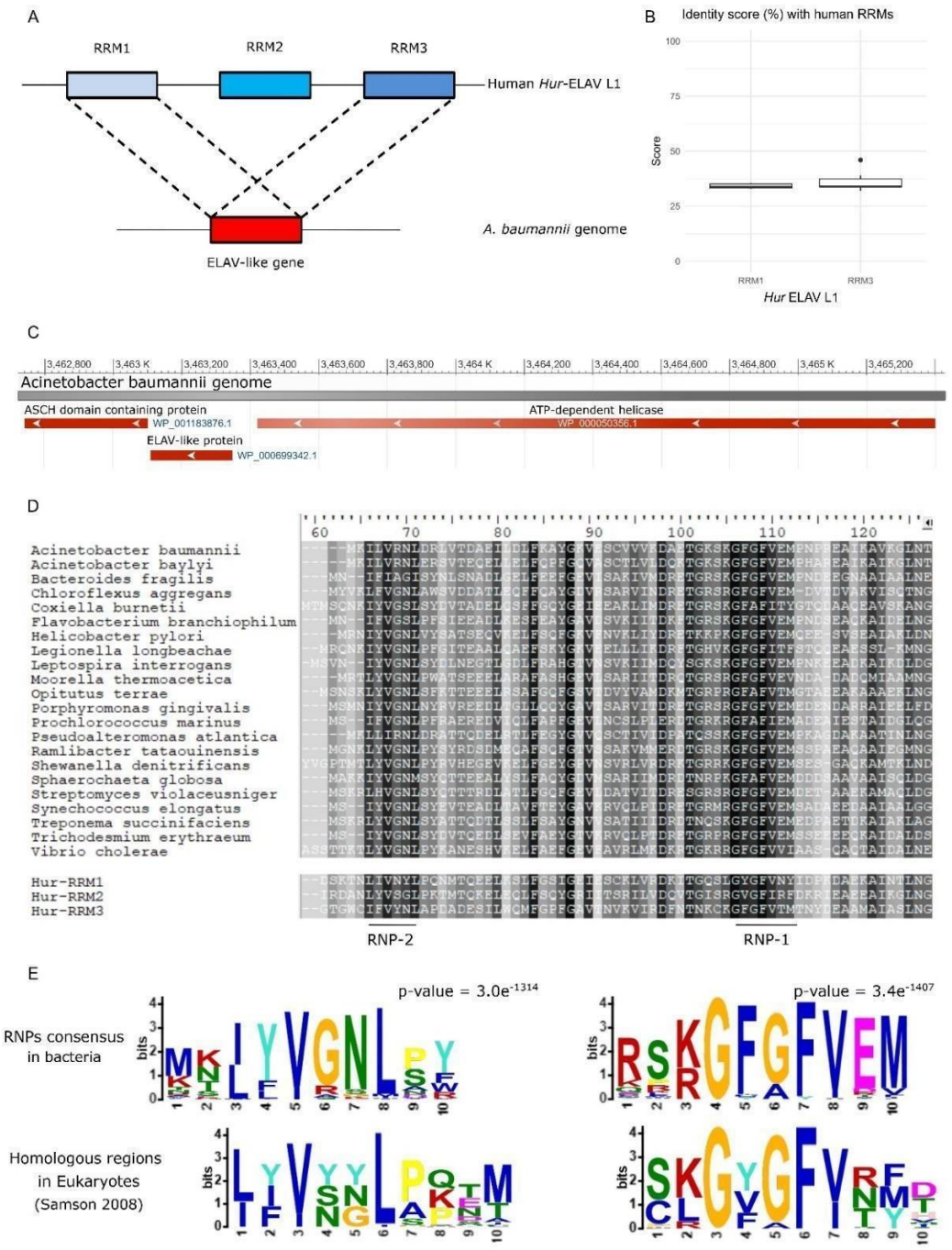


Figure 1. *In silico* analysis. A) The tblastn search, using as query HuR and restricting the search to *A. baumannii* genomes, gave 25 outputs (25 hits), all hits corresponded to the same orthologous protein which found a homology in both RRM1 and RRM3; B)Boxplots showing the percentages of identity. C) Genomic context of *AB*-Elavl. It is shown that the three genes, namely the ATP-dependent helicase, *AB*-Elavl and the ASCH domain containing protein are organized in tandem. D) Alignment of the bacterial homologues of human ELAVL in selected bacterial species with clinical or environmental relevance spanning seven phyla. The background colors denote the level of conservation in that position, darker background mean more conserved residue in that position. E) Motifs significantly conserved across all the bacterial sequences matching the ribonucleoprotein motifs RNP-2 and RNP-1 in each of the RRMs in the gene of HuR.

2. Validation of the expression of *AB*-Elavl protein in *Acinetobacter baumannii*.

I initially checked whether the RNA transcript corresponding to *Acinetobacter baumannii* Elav-like (*AB*-Elavl) was expressed in the bacterium. Using a Reverse Transcriptase enzyme, I generated a complementary DNA (cDNA) from the total transcriptome of the *Acinetobacter baumannii* reference strain (ATCC 19606) and amplified the surroundings of the gene of interest by PCR (fig 2A) using three different pairs of primers. The sequences of the amplicons were confirmed by Sanger sequence analysis and matched the DNA deposited sequence. I observed that our gene of interest is expressed and embedded into a longer mRNA of at least 764 bp, suggesting that this genomic region expresses a polycistronic mRNA. Indeed, as already stated before, bioinformatics analysis shows our gene is localized between two other genes: downstream, a gene encoding an ASCH domain containing protein and, upstream, a gene encoding an ATP-dependent helicase (fig 1C).

To evaluate whether the polycistronic mRNA is translated into a protein containing the domain of interest, in collaboration with the Mass Spectrometry facility of CIBIO, we performed proteome analyses by mass spectrometry on the protein lysate of *Acinetobacter baumannii* ATCC 19606. Protein cell lysate was separated into a polyacrylamide gel, a band (6-14 kDa) comprising the MW of the predicted protein was cut, trypsin digested and submitted to LC-MS/MS analysis (fig 2C). Among the detected fragments, we obtained 40% coverage of the entire protein with complete matching of the experimental amino acid sequence with the predicted sequences (fig 2C, below). Interestingly, the detected protein fragments contained the region of the highly conserved heptapeptide KGFGFVT which we found conserved in the protozoans and that corresponds to the octapeptide RGFGFVTF, the ribonucleoprotein motif 1 (RNP-1) in metazoans (fig 2D).

We performed an LC-MS/MS analysis also on the total proteome of *Acinetobacter baumannii*, but we could not find the peptides of our interest. This suggests that our protein of interest could be expressed at low concentration.

Taken together, these data indicate that the DNA encoding the hypothetical *AB*-Elavl protein is effectively transcribed and translated into a protein that contains the conserved RNP-1 amino acid sequences responsible for nucleic acid binding.

3. Expression of the recombinant *Acinetobacter baumannii* Elav- like (r*AB*-Elavl) protein.

In order to express the recombinant protein, I cloned the sequence of *AB*-Elavl into the expression plasmid pET30a(+) in frame with a 6XHis tag in the C terminal region, and expressed the recombinant *Acinetobacter baumannii* Elav-like (r*AB*-Elavl) protein in *E.coli* Rosetta BL21. I purified the recombinant protein (predicted molecular weight (MW): 11885.05 Da) from Rosetta BL21's cell lysate by affinity purification: I incubated the cell lysate with Nickel NTA agarose beads, in order to bind r*AB*-Elavl to the beads. The blocked protein was washed and eluted by increasing imidazole concentration. The final concentration of the eluted protein was measured using a UV-VIS spectrophotometer and I observed a high quantity of nucleotides in the solution, probably bound to the protein. I treated the eluted protein with Polyethyleneimine 5% (PEI 5%) to eliminate the nucleotides and obtain a pure solution. The PEI treatment reduced the yield of the recombinant protein from 1.4 mg/ml (117.8µM) to 0.5 mg/ml (42µM), but the 260 absorbance of nucleotides dropped to nearly zero. The purity of the r*AB*-Elavl protein was evaluated by Coomassie staining of protein polyacrylamide gel of each purification step performed (fig 2B, arrows). The yield of protein production was always around 12.5mg/L (1µM). The purified r*AB*-Elavl protein was subjected to mass spectrometry analysis after trypsin and chymotrypsin digestion. I obtained an 82.45% coverage of the entire recombinant sequence, and the detected peptides perfectly matched the encoded amino acid sequence (fig 2C, above).

To get more insight on the presence and on the MW of the hypothetical *AB*-Elavl protein translated in the bacterium, I developed a custom antibody against the recombinant protein with the collaboration of "Davids Biotechnologie". Rabbits were immunized with the denatured r*AB*-Elavl protein, and after two cycles of immunization, the serum was collected, and the IgG titre quantified. Specificity of the IgG in recognizing the protein of interest was investigated by performing western blot against the r*AB*-Elavl protein, the recombinant human HuR and MCF7 cell lysate (human breast cancer cell line expressing HuR protein) (fig 2E). I confirmed the immunized serum can recognize the r*AB*-Elavl but not the recombinant human homologous, neither any protein from the human cell lysate. Comparing the signal from the total protein lysate of *Acinetobacter baumannii* and the one from the recombinant protein, it is possible to observe a band at a slightly heavier MW

compared to the hypothesized one (around 10kDa *versus* around 12kDa). This suggests there could be either post-translational modification on the protein sequence, the native protein is longer than the predicted one or the recombinant one is digested in any part during the protein production into *E. coli* (fig 2E).

Using the online Expasy tool I was able to calculate the isoelectric point (pI: 9.33) and the molecular weight (10.8 kDa) of the hypothesized protein, while for the recombinant protein, in which there is the presence of a 6*His tag, the isoelectric point remains equal to 9.33 but the molecular weight increases to 11.8 kDa.

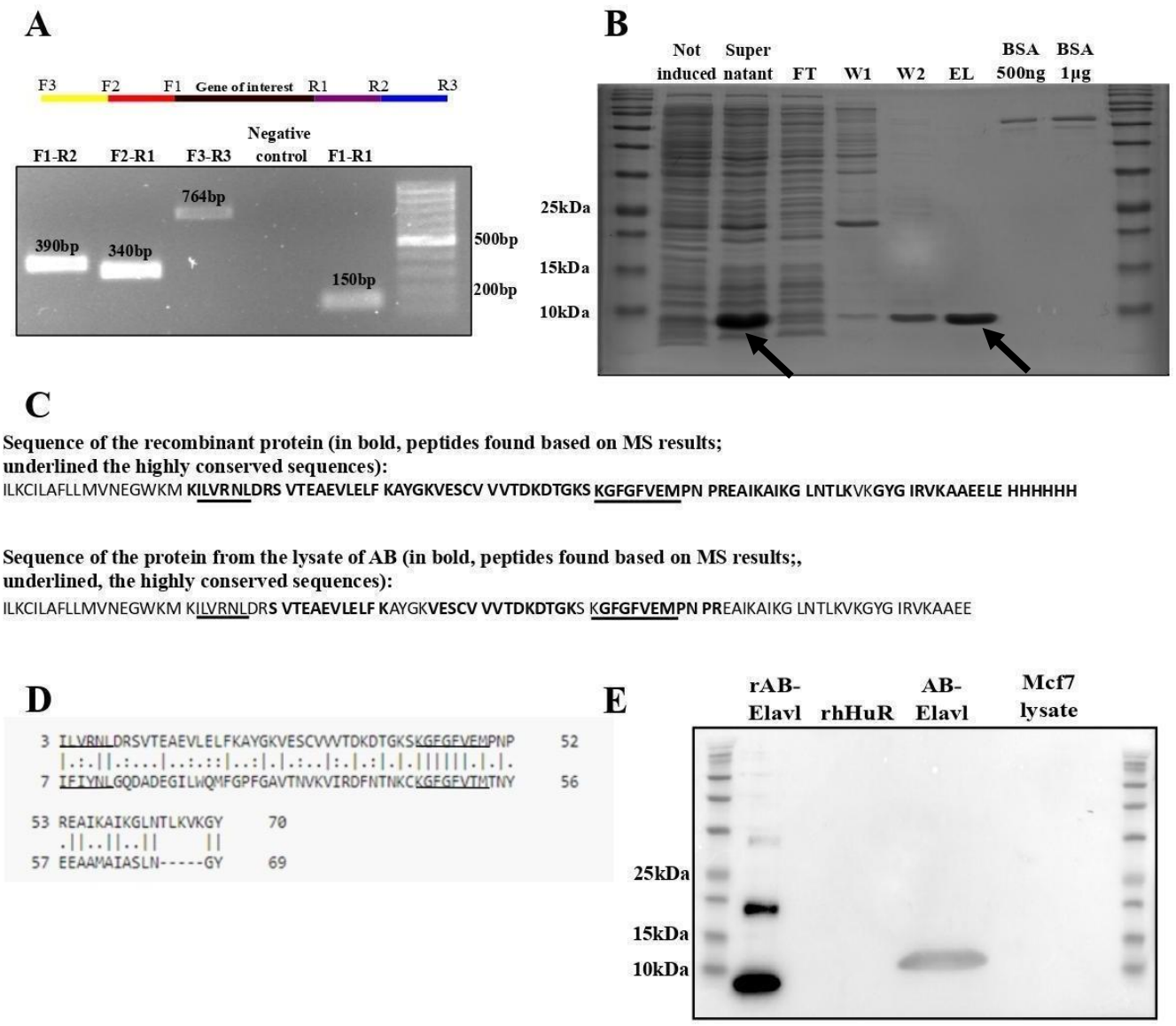


Figure 2. Protein identification and purification. A) PCR amplification of the transcription of the polycistronic mRNA containing the sequence of interest. Different pairs of primers were designed, as shown in the upper part of the fig. The amplicons produced are 390bp for F1-R2, 340bp for F2-R1, 764bp for F3-R3 and 150bp for F1-R1 (this amplicon was also used as positive control). B) Purification of the recombinant protein. FT: flow through, W: wash, EL: elution, Mk: marker. Arrows indicate the recombinant protein. C) Mass spectrometry analysis. The recombinant protein was analyzed as first in orderto confirm the sequence. It was then used as a reference for the analysis of *Acinetobacter baumannii* proteome. In bold found peptides with high confidence, underlined the conserved peptides. The predicted molecular weight is 11.8KDa for the recombinant protein and 10.8 for the protein from *A. baumannii*. The predicted isoelectric point is 9.33 for both the proteins. D) Alignment of *AB*-Elavl (above) and the RRM3 domain of HuR (below). "|" means that the residues in column are identical.; ":" means that the amino acid in column has been substituted by one with similar characteristics; "." means that semi-conserved substitutions are observed. E) Western blot analysis to confirm the presence of the protein of interest in the protein lysate of *Acinetobacter baumannii* and in the r*AB*-Elavl sample, as well as the absence of signal in the sample of the recombinant proteins HuR and in the Mcf7 lysate.

4. Determination of the *AB*-Elavl protein structure

X-RAY and NMR analysis were performed by our collaborator prof. Marco Fragai in Centro Risonanze Magnetiche (CERM) in Florence. X-RAY analysis showed the typical domain of an RRM domain: β_1-α_1-β_2-β_3-α_2-β_4 (Tang et al., 2012) (fig 3C). This structure is also typical in the ELAV-like family. The two conserved sequences, the hexamer (ILVNRL) and the heptamer (KGFGFVE), are localized at the level of the two internal strands of the β-sheet: β_2-β_3 (fig 3C). These amino acids are believed to be responsible for the binding to the RNA (Samson, 2008).

4.1. Crystal structure analysis

The structure was solved using the molecular replacement method; the model showing the highest sequence identity (about 40%) was HuD (a human paralog of HuR that share with it the ability to bind AU rich element of the C-FOS RNA). MODELLER was used based on this structure to generate the model with the correct sequence for molecular replacement. The solved structure shows the absence of the first 18 residues in the electron density with respect to the cloned sequence. It is not trivial to tell whether this is due to their high mobility or rather to their loss due to some protein degradation before/during crystallization. Figure 3A shows the superposition between HuR (red) and HuD (green). It appears quite clear that the fold of HuD is very similar to that of the model used for structure solution and, in turn, similar to the typical RNA binding motif (fig 3C). The greatest discrepancy between the two structures is in the region involving residues from 50 to 58, just before RNP-1. In our case electron density is missing for those residues confirming thus a very high mobility. This mobility is not present in the case of HuD because this region interacts with RNA.

4.2. NMR resonance assignment.

The 2D ^{1}H–^{15}N HSQC spectrum of the *AB*-Elavl, shows 79 well-dispersed and resolved signals in agreement with a small, uniform and well-folded protein structure (fig 3B). The backbone assignment of the protein was obtained from the analysis of triple resonance spectra. All the residues from Lys-22 to Glu-101 were assigned in the spectra, while the first 21 N-terminal residues were missing. TALOS+ (Schen et al. 2009) program was used to obtain secondary structure prediction. Chemical shifts are, indeed, reporters of the protein structure. The same

residue, when located on a secondary structure element (α-helix or β-strand) or on random coil, experiences different values of chemical shifts for proton, carbon and nitrogen. Talos+ algorithm estimates the differences between the provided experimental chemical shifts (HN, Cα, Cβ, C', N) for each residue with respect to their corresponding random coil values. In particular, the program uses chemical shift and sequence information and matches the experimental chemical shift values with those reported for known structures available in the database. Secondary structure probability is, thus, provided for each residue, along with an estimation of the backbone torsion angles. The predictions of secondary structure obtain from the assigned chemical shift or RRM domain of *AB*-Elav1 indicate that this domain is constituted according to TALOS+ predictions, by two α-helices and four β-strands, in agreement with the currently resolved crystal structure and with the reported structures of the RRM domains (fig 3A and 3C).

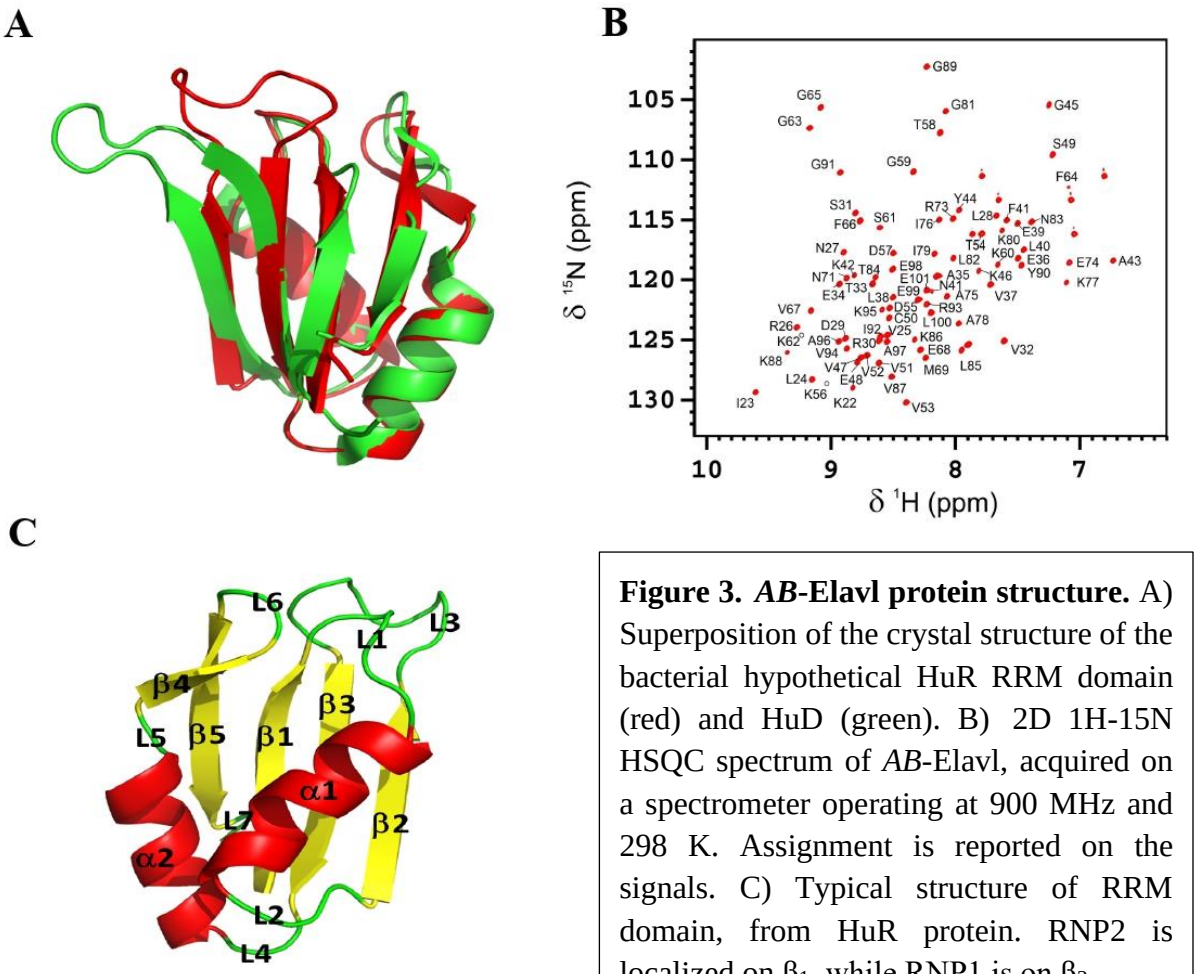


Figure 3. *AB*-Elavl protein structure. A) Superposition of the crystal structure of the bacterial hypothetical HuR RRM domain (red) and HuD (green). B) 2D 1H-15N HSQC spectrum of *AB*-Elavl, acquired on a spectrometer operating at 900 MHz and 298 K. Assignment is reported on the signals. C) Typical structure of RRM domain, from HuR protein. RNP2 is localized on β_1, while RNP1 is on β_2.

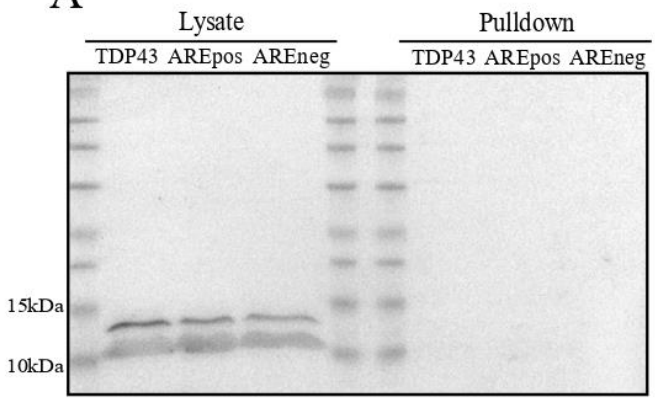
5. Molecular characterization of *AB*-Elavl

The limited abundance of *AB*-Elavl makes difficult to capture it using molecular approaches from lysates. In the pulldown assay I incubated the total lysate of *A. baumannii* with three different 5'-biotinylated single strand RNA (ssRNA) probes: the biotinylated ARE pos (Bi-ARE pos), the biotinylated ARE negative (Bi-ARE neg) and an unrelated RNA sequence taken from the consensus sequence of the RNA binding protein TDP43 (Bi-TDP43). The complex RNA-protein was then pulled down using magnetic streptavidin beads. There was not an enrichment of *AB*-Elavl that resulted to remain all in the lysate (fig 4A). In order to check that the pulldown was technically well performed, I tested the recombinant protein: incubating r*AB*-Elavl with the three probes: Bi-ARE pos, Bi-ARE neg and Bi-TDP43. I found an enrichment of the complex RNA-r*AB*-Elavl that was pulled down, sustaining the hypothesis that *AB*-Elavl has a limited concentration in the bacterial lysate (Fig 4B).

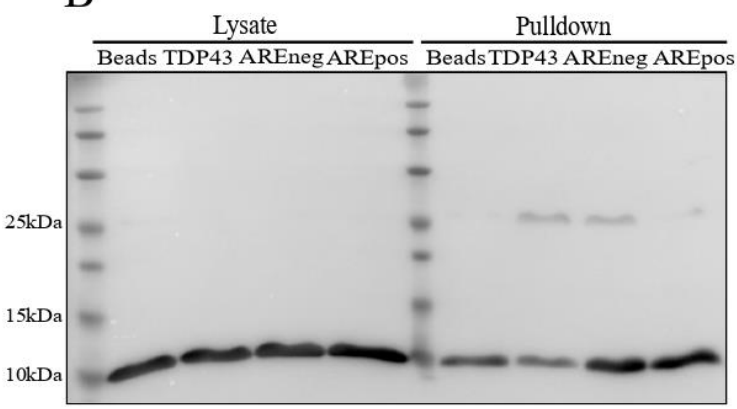
Using the antibody against *AB*-Elavl, I performed a protein immunoprecipitation from the total lysate of *Acinetobacter baumannii*: the complex antibody-*AB*-Elavl was precipitated by conjugating the antibody to magnetic beads. Proteins precipitated were run on polyacrylamide gel but again, no bands were detected, likely due to the already hypothesized low concentration of the protein of interest into the lysate (antibody against IgG was used as a control of the well performed assay) (fig 4C). Therefore, in collaboration with the Mass Spectrometry facility from CIBIO, I performed an immunoprecipitation followed by mass spectrometry (IP-MS), to investigate which proteins were enriched, with respect to rabbit IgG, used as control. About 5675 proteins were enriched into the immunoprecipitated material and, among the most enriched ones, we found three entries in the Uniprot database related to hypothetical RNA binding proteins of *Acinetobacter baumannii* (fig 4D and 4E) that are extremely similar to our protein of interest as the percentage of homology spans from 81% to 54.5%. The three entries are D0CAL6, 86 aa, predicted MW 9560.22 Da, A0A009GG82, 79 aa, predicted MW 8715.23 Da and A0A4R5S8D9, 58 aa, predicted MW 6445.52 Da (fig 4G). All of them showed a predicted MW lower than the recombinant protein. Notably, in addition to the previously identified protein fragments in the protein lysate, we found eight more amino acids that completed the retrieval of the hexapeptide conserved sequence (ILVRNL) in protozoans homologous to the RNP-2 involved in the recognition of RNA in metazoans (fig 4F).

In sum, molecular approaches are difficult to use, since this protein is extremely small and not abundant into the bacterium. An important achievement is the proof of the presence of the conserved RNP-2, responsible, together with RNP-1 for recognition and binding of the target nucleotides. The limited availability did not permit the RNA immunoprecipitation assay, which remains a goal still to be achieved.

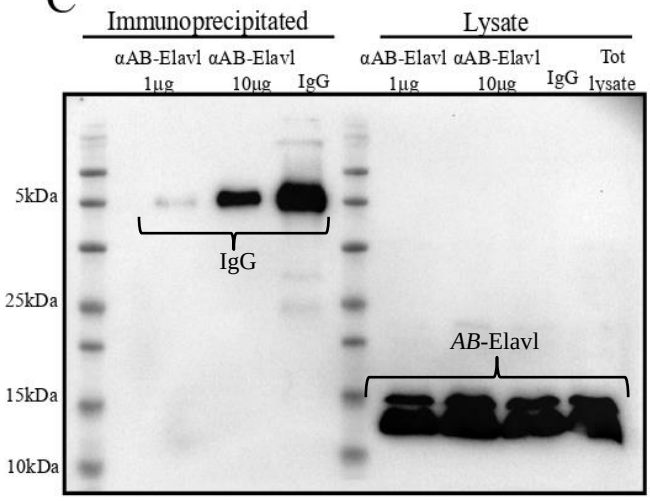
A



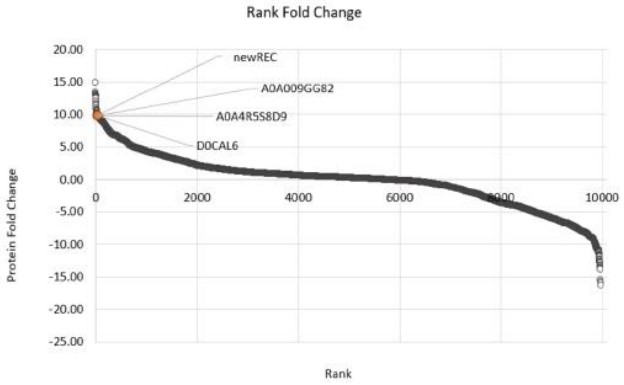
B



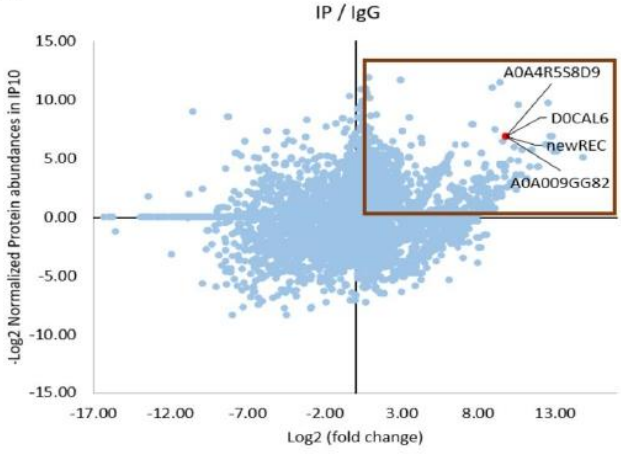
C



D



E



F

Sequence of the protein
(in bold, peptides found based on MS results;
underlined the high confidence peptides):

ILKCILAF LLMVNEGWKM KILVRNLDRS
VTEAEVLELF KAYGKVESCV
VVTDKDTGKS KGFGFVEMPN PREAIKAIKG
LNTLKVKGYG IRVKAAEELE

G

r*AB*-Elavl: HMILKCILAF LLMVNEGWKM KILVR NLD RS VTEAEVLELF KAYGKVESCV VVTDKDTGKS KGFGFVEM PN PR EAIKAIKG LNTLKVKGYG IRVK AAEELE HHHHHH
D0CAL6: MVNEGWKM KILVR NLD RS VTEAEVLELF KAYGKVESCV VVTDKDTGKS KGFGFVEM PN PR EAIKAIKG LNTLKVKGYG IRVKAAEE
A0A009GG82: M KILVR NLD RS VTEAEVLELF KAYGKVESCV VVTDKDTGKS KGFGFVEM PN PR EAIKAIKG LNTLKVKGYG IRVKAAEE
A0A4R5S8D9: M KILVR NLD RS VTEAEVLELF KAYGKVESCV VVTDKDTGKS KGFGFVEM PN PR EAIKA

Figure 4. Molecular characterization of *AB*-Elavl. A) Pulldown assay on the total protein lysate of *A. baumannii*. The protein was not enriched in the precipitated samples, since it remained all in the inputs. B) Pulldown assay on the recombinant protein in order to check if effectively the protein can recognize and bind the biotinylated probes. C) Immunoprecipitation assay on the total protein lysate of *A. baumannii*. IgG was used as a control. No enrichment of the protein was visible by western blot analysis in the immunoprecipitated samples. D) and E) Mass spectrometry analysis on IP samples shows an enrichment of *AB*-Elavl in the top ten proteins precipitated in the assay. newRec is the sequence of r*AB*-Elavl F) Peptides found as enriched in the IP assay belonging to the sequence of our interest. In bold the peptides found by mass spectrometry, among which underlined the high confidence ones. G) Entries from the Uniprot database related to the peptides found in the IP-MS analysis.

6. Biochemical characterization of r*AB*-Elavl

Given the high structural similarity to the mammalian ELAVL and the presence of the conserved RNP-1 and RNP-2, I investigated whether the r*AB*-Elavl protein had RNA binding ability as its mammalian counterparts. I initially evaluated whether in the proteome of *Acinetobacter baumannii* there were proteins able to recognize and bind the ARE sequence as the one from the 3'UTR of *TNFα*, a known target of HuR, using non-denaturing and non-cross linked RNA Electro Mobility Shift assay (REMSA) (D'Agostino et al., 2013). The incubation of proteins with their RNA target leads to the formation of complexes that migrate slowly in a native acrylamide gel, compared to the free nucleotides. The single strand RNA positive ARE probe was bound with an infra-red dye DY681 (IR-AREpos). By mixing higher concentrations of the protein lysate with a fixed 2.5 nM concentration of the IR-AREpos probe I observed a decreased quantity of free RNA probe and the formation of a protein-RNA complex. This indicates the presence of one or more proteins able to bind the IR-AREpos (fig 5A). I then evaluated if the r*AB*-Elavl was able to bind to ARE positive probes or also to probes that did not contain the HuR consensus sequence (ARE negative RNA probes), by REMSA. I mixed increasing amounts (2μM and 4μM) of protein with 2.5 nM IR-AREpos and 2.5 nM IR-AREneg. As shown in the mobility shift assay, r*AB*-Elavl clearly caused the RNA probe electrophoretic retardation detectable as one prominent band, with both probes, however showing a binding preference towards the AREpos probe in this biochemical condition (fig 5B, arrows). Using the antibody against r*AB*-Elavl, I assessed if the formation of the heavier complex antibody-protein-RNA could create a super shift due to a major electrophoretic retardation compared to the one composed of just r*AB*-Elavl-RNA complex. I observed a decrease of the signal at the level of the transient complex and an increase in the free RNA (fig 5C, arrows), but I did not detect the antibody-protein-RNA complex super-shift, likely because the antibody developed is not able to recognize the folded structure of the protein (fig 5C).

To quantitatively characterize the binding activity of the r*AB*-Elavl to different RNA probes, I used AlphaScreen technology, a bead-based assay that uses luminescence of acceptor beads excited by singlet oxygen to detect the proximity of donor beads. I used the three different 5'-biotinylated ssRNA probes: Bi-ARE pos, Bi-ARE neg and Bi-TDP43. I optimized the assay to identify the best molar ratio between the two interacting partners coupled with anti-His-Acceptor and Streptavidin-Donor beads; the best signal (before the hook point, see arrows in fig 5E) was calculated with the three different probes at 50 nM concentration and using a 2 μM concentration of r*AB*-Elavl. I also tested different concentrations of Bi-AREpos in order to find the best ratio protein/RNA (fig 5D).

The recombinant protein shows a high affinity for Bi-ARE Pos (EC_{50}: 42.5 nM) while it has low affinity for the Bi-ARE neg probe (EC_{50}: 257.1nM) and, surprisingly, a high affinity for Bi-TDP43 (EC_{50}: 12.4nM) (Fig 5F).

I then evaluated the minimal ARE sequence length required for binding. I observed that r*AB*-Elavl, as its human orthologous, has a higher affinity for longer ARE sequences then for shorter ones (EC_{50} ARE pos 29 nt: 35.62, ARE pos 19 nt: 64.76, ARE pos 11 nt: not converged) mimicking the behaviour of HuR as it is reported in literature (fig 5G) (Ma et al., 1996).

The affinity evaluation was confirmed using the HTRF-FRET assay. This technology is similar to the AlphaScreen assay since it is based on the distance-dependent interaction between the acceptor beads europium-labeled anti-6X His-Antibody and donor beads XL665–conjugated for biotin detection. This assay as well, was optimized in order to identify the best molar ratio between the two interacting partners: the optimal protein concentration, before the hooking effect, was observed at around 200nM for all the probes (fig. 6A). I calculated the EC_{50} of the different 5'-biotinylated ssRNA probes (Bi-ARE pos, Bi-ARE neg and Bi-TDP43) that resulted to be really similar to those calculated with the AlphaScreen: ARE pos- EC_{50}: 24.27 nM, ARE neg- EC_{50}: Ambiguous, TDP43- EC_{50}:8.6 nM (fig 6B).

To further define the binding affinity between the r*AB*-Elavl protein and ARE sequence, I performed a time course experiment in which different concentrations of the Bi-AREpos were mixed to the protein with different incubation time. The experiment shows that the binding of r*AB*-Elavl to Bi-AREpos probe was both time and dose dependent (Figure 6C). Data were globally fitted using the association kinetic model of multiple ligand concentration: derived association (kon of 2.035 M-1 min-1) and dissociation (koff of 0.02687 min-1) rates indicated a very high affinity of the r*AB*-Elavl protein towards this RNA substrate and a low dissociation rate. According to the law of mass action, the equilibrium binding constant Kd calculated as koff/kon was obtained as Kd value of 13,2 nM. I performed the same type of assay for Bi-AREneg (fig 6D), for which the binding resulted as ambiguous, and TDP43 for which the kon is 762585 M-1 min-1 and a koff is 0.01528 min-1, for a final Kd of 20nM (fig 6E).

I tried the Isothermal titration calorimetry (ITC) as another method to further characterize the interaction between r*AB*-Elavl and the three probes. ITC is able to directly measure the heat released or consumed in a reaction. As control, I measured the heat released from injection of RNA into the pure dialysis buffer (fig

7A) used to elute the recombinant protein. I then tested two different concentrations of protein (5μM, fig 7B and 10 μM, fig 7C) in the chamber, and a unique concentration of RNA (50 μM) in the syringe. Unfortunately, I did not detect any difference of heat released or absorbed during the molecular binding event between the protein and the RNA in both cases.

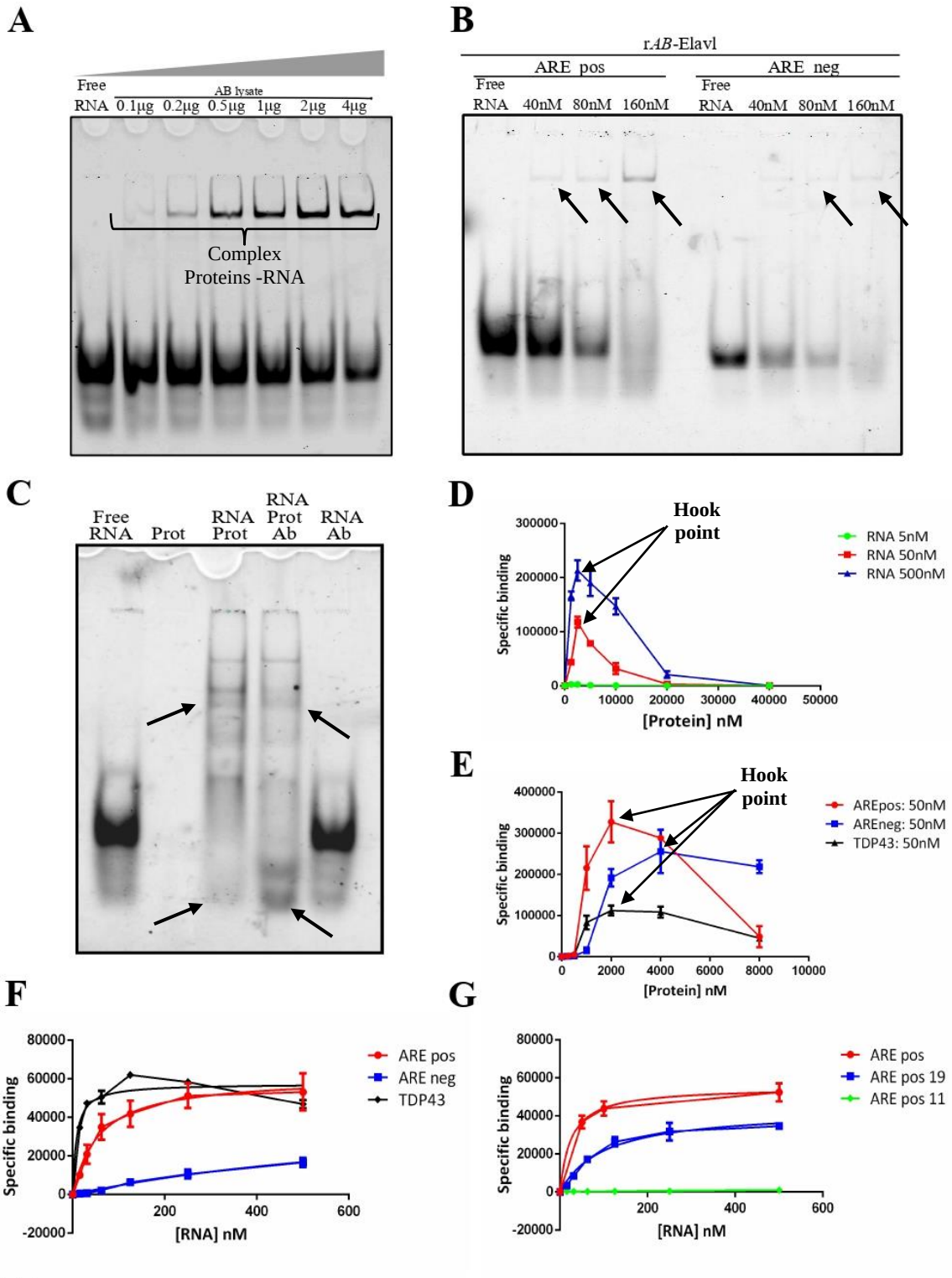
I also tried the Grating Coupled Interferometry (GCI), with the collaboration of the "2bind" company, in which the kinetical interaction between r*AB*-Elavl and Bi-AREpos, Bi-AREneg and Bi-TDP43 was investigated. They captured the protein on a microchip thanks to the 6XHis tag and incubated it with the probe in a multicycle kinetics experiment by injecting the analyte (AREpos, AREneg or TDP43) in increasing concentrations over the surfaces where is bound the ligand (AREpos fig 7D, AREneg fig 7E and TDP43 fig 7F). No binding was detected with any of the probes tested. This made us hypothesize that probably the binding affinity between our protein and the RNA is increased by the formation of a protein complex on a consensus sequence, i.e., upon multimerization of the protein onto the target RNA. In order, to evaluate this hypothesis, I performed the Dynamic light scattering assay (DLS). This assay studies the diffusion behaviour of macromolecules in solution, giving information on size and shape of the molecules. Again, the maximum concentration of my protein (1.4μg/μl, 118.6μM) resulted to be lower than the sensitivity of the Instrument.

Finally, prof. Fragai also titrated the r*AB*-Elavl with the available probes by NMR. The binding of different RNA probes to the RRM domain of *AB*-Elavl was investigated by solution NMR. The 15N isotopically enriched protein was titrated with increasing aliquots of each RNA strand. Solution NMR allows to obtain residue specific information about the interaction of a protein with a partner. In particular, the value of chemical shifts of the protein atoms depends on their chemical environment. The interaction with a partner (i.e., ligand, nucleic acid, other protein, etc.) causes environmental changes on the protein surface, that is reflected by a chemical shift perturbation and/or line broadening that can be measured as a change in the signal intensity. Residues, whose corresponding signals are mainly affected by these variations, can be mapped on the protein surface, and are presumably located at the binding interface.

Chemical shift perturbations (CSPs) and changes in signal intensity were analysed in the 2D 1H 15N BEST-TROSY NMR spectra recorded during the NMR titration. After the addition of TDP43 to the protein solution a general decrease of the protein resonances is observed (fig 8C), with some residues (Ile23, Leu24, Asn27, Leu28, Asp29, Leu38, Glu48, Val53, Gly65, Phe66, Ile79, Thr84, Ile92) experiencing a

larger effect (fig 8A-top and 9A). These residues are mainly located on the β-platform of the domain. The generalized decrease of signal intensity could be explained with protein aggregation on the RNA fragment. Interestingly, few residues, located on the N-terminal β-strand (Lys22, Val25) and on the C- terminal loop (Glu98, Glu99, Leu100, Glu101), that are close to the β-platform, experience a significant chemical shift perturbation (CSP) (figure 8B-top and 9A). The interactions of the *AB*-Elavl with AREpos and AREneg were also investigated through solution NMR. The effect of the decrease in signal intensity and CSP with both the positive and the negative probes is by far reduced with respect to what is observed in the presence of TDP43. At the protein/RNA molar ratio of 1:0.5 a global decrease of signal intensity is not observed. However, in the presence of AREpos, some residues (Ile23, Val25, Asp29, Ser31, Gly63, Phe64, Phe66, Val67, Ile92, Glu98), located in the same region interacting with TDP43, experience a decrease in signal intensity (fig 8A-middle and 9B). Some of these residues (or the neighbouring ones) are also affected by CSP (Val25, Arg30, Gly63, Phe64, Ile92, Glu98, Glu99, Leu100, Figure 8B-middle and 9B). A similar behaviour is observed in the presence of AREneg with some residues of the β-platform experiencing a decrease in signal intensity (Val25, Asn27, Leu38, Leu40, Phe41, Lys62, Phe64, Ile76, Glu101; Fig 8A-bottom and 9C) and/or a CSP (Lys22, Arg30, Lys56, Phe64, Phe66, Ile92, Glu99,Leu100, fig 8B-bottom and 9C).

Passing through several techniques and using the recombinant r*AB*-Elavl, I was able to establish *AB*-Elavl interacts with the probes tested, even if with different affinity. Compared to the human counterpart, this protein is not so selective for AU-rich regions only. This is not surprising, since in evolution RNA binding proteins resulted to be more specialized and increased in number compared to their ancestors.



H

ARE pos: AUUAUUUAUUAUUUAUUUAUUAUUUA
ARE pos 19: AUUAUUUAUUAUUUAUUUA
ARE pos 11: AUUAUUUAUUA
ARE neg: ACCACCCACCACCCACCCACCACCCA
TDP43: CCGGGGCCGGGGCCGGGGCCGGGG

Figure 5. Analysis of the protein binding abilities. A) EMSA assay on the total protein lysate of *Acinetobacter baumannii* and a probe mimicking the AU rich sequence of TNFα (ARE pos) with an infrared tag (IR-AREpos). B) EMSA assay on the recombinant protein incubated with different probes with an infrared tag: IR-AREpos and IR-AREneg. AREpos is bound with a high affinity, while AREneg show a lower affinity. C) EMSA assay for detection of the super-shift in presence of the antibody against *AB*-Elavl. In the presence of the antibody, the super-shift is not detectable, even though it is visible a reduction of the transient complex (arrows) and an increase of the free RNA compared to sample without antibody. D) AlphaScreen for detection of the hook point on the recombinant protein and three different concentrations of AREpos probe. I chose to use a concentration of protein minor then 250nM and a concentration of RNA equal to 50nM. E) AlphaScreen on the recombinant protein and different probes but with similar length: ARE pos, ARE neg and TDP43. It is evident that the hook point is similar for all the three probes and that AREpos has the higher affinity to the recombinant protein. F) AlphaScreen saturation assay for detection of the minimal probe length for binding of the protein. The probe are AREpos with 3' deletions: ARE pos: ARE sequence full length, ARE pos 19: ARE sequence with 19 nucleotides, ARE pos 11: ARE sequence with 11 nucleotides. The minimal number of nucleotides in order to obtain the binding is 19, but longer sequences have a higher affinity. G) AlphaScreen saturation experiment for Kd calculation between the recombinant protein and AREpos, AREneg and TDP43. AREpos and TDP43 show a high affinity while AREneg is not well bound. H) Sequences of the probes used in the different assays.

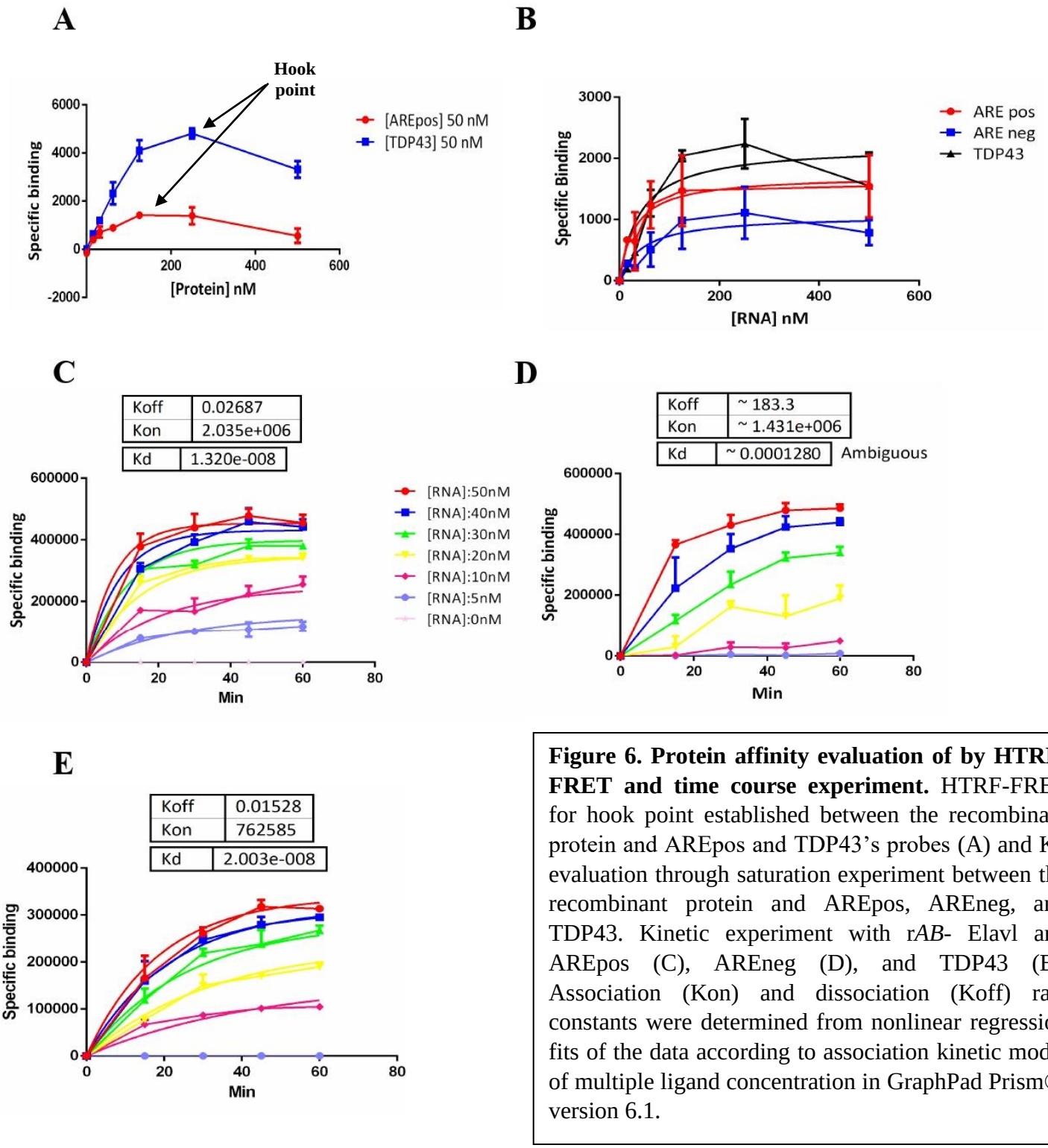


Figure 6. Protein affinity evaluation of by HTRF-FRET and time course experiment. HTRF-FRET for hook point established between the recombinant protein and AREpos and TDP43's probes (A) and Kd evaluation through saturation experiment between the recombinant protein and AREpos, AREneg, and TDP43. Kinetic experiment with r*AB*- Elavl and AREpos (C), AREneg (D), and TDP43 (E). Association (Kon) and dissociation (Koff) rate constants were determined from nonlinear regression fits of the data according to association kinetic model of multiple ligand concentration in GraphPad Prism®, version 6.1.

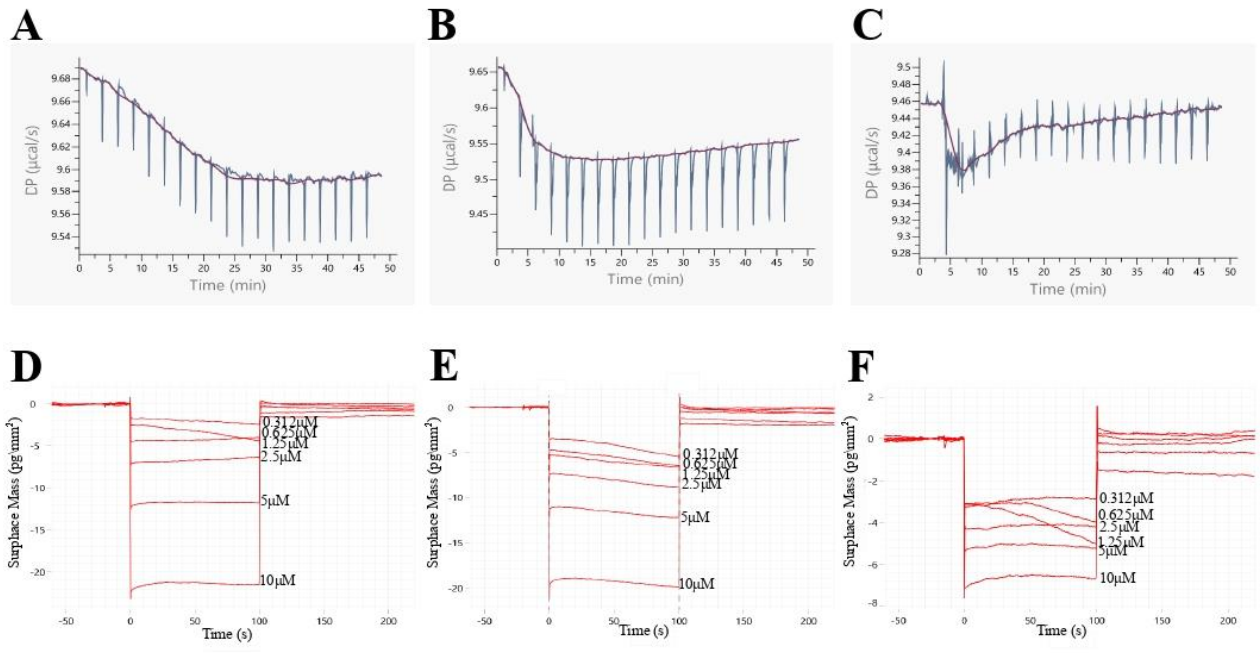


Figure 7. Confirmation of binding abilities of r*AB*-Elavl by ITC, and GCI assays. A) ITC experiment: measurement of the heat released by injecting AREpos into the protein's buffer, as control. It was not measured a relevant change in heat due to the reaction. B) ITC experiment for measuring the heat of binding released in the formation of the complex protein-RNA (AREpos). [Protein]: 5µM, [RNA]: 50µM. It was not measured a relevant change in heat due to the reaction C) ITC experiment for measuring the heat of binding release in the formation of the complex protein-RNA (AREpos). [Protein]: 10µM, [RNA]: 50µM. It was not measured a relevant change in heat due to the reaction. D) GCI experiment in which r*AB*-Elavl was bound on a microchip surface and exposed to different concentration of AREpos; no bound was detected. The experiment was performed by 2bind GmbH. E) GCI experiment in which r*AB*-Elavl was bound on a microchip surface and exposed to different concentration of AREneg; no bound was detected. The experiment was performed by 2bind GmbH. F) GCI experiment in which r*AB*-Elavl was bound on a microchip surface and exposed to different concentration of TDP43; no bound was detected. The experiment was performed by 2bind GmbH.

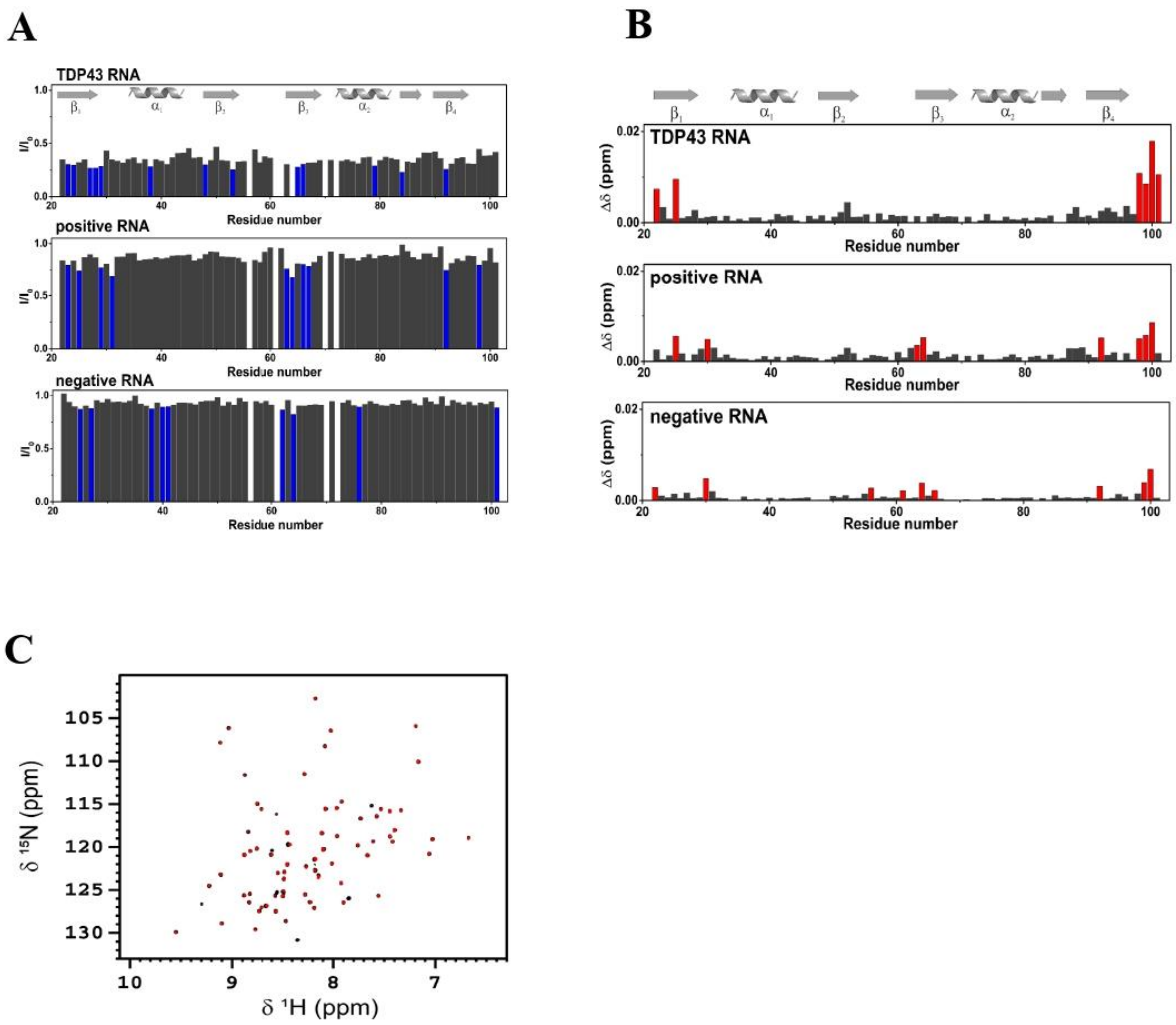


Figure 8. NMR assay. A) Intensity decreases plots of free r*AB*-Elavl RRM domain (70 µmol dm-3) with respect to the protein in the presence of 35 µmol dm-3 TDP43 (top), 35 µmol dm-3 positive RNA (middle), and 35 µmol dm-3 negative RNA (bottom). The residues experiencing the largest decreases have been highlighted in blue. B) Chemical shift perturbation (CSP) of r*AB*-Elavl RRM domain (70 µmol dm-3) with respect to the protein in the presence of 35 µmol dm-3 TDP43 RNA (top), 35 µmol dm-3 positive RNA (middle), and 35 µmol dm-3 negative RNA (bottom). The residues experiencing the largest decreases have been highlighted in red. C) Superimposed 2D 1H-15N HSQC spectra of free *AB*-Elavl RRM domain (70 µmol dm-3, black) and in the presence of 35 µmol dm-3 of TDP43 RNA (red). The spectra were acquired on a spectrometer operating at 950 MHz and 298 K.

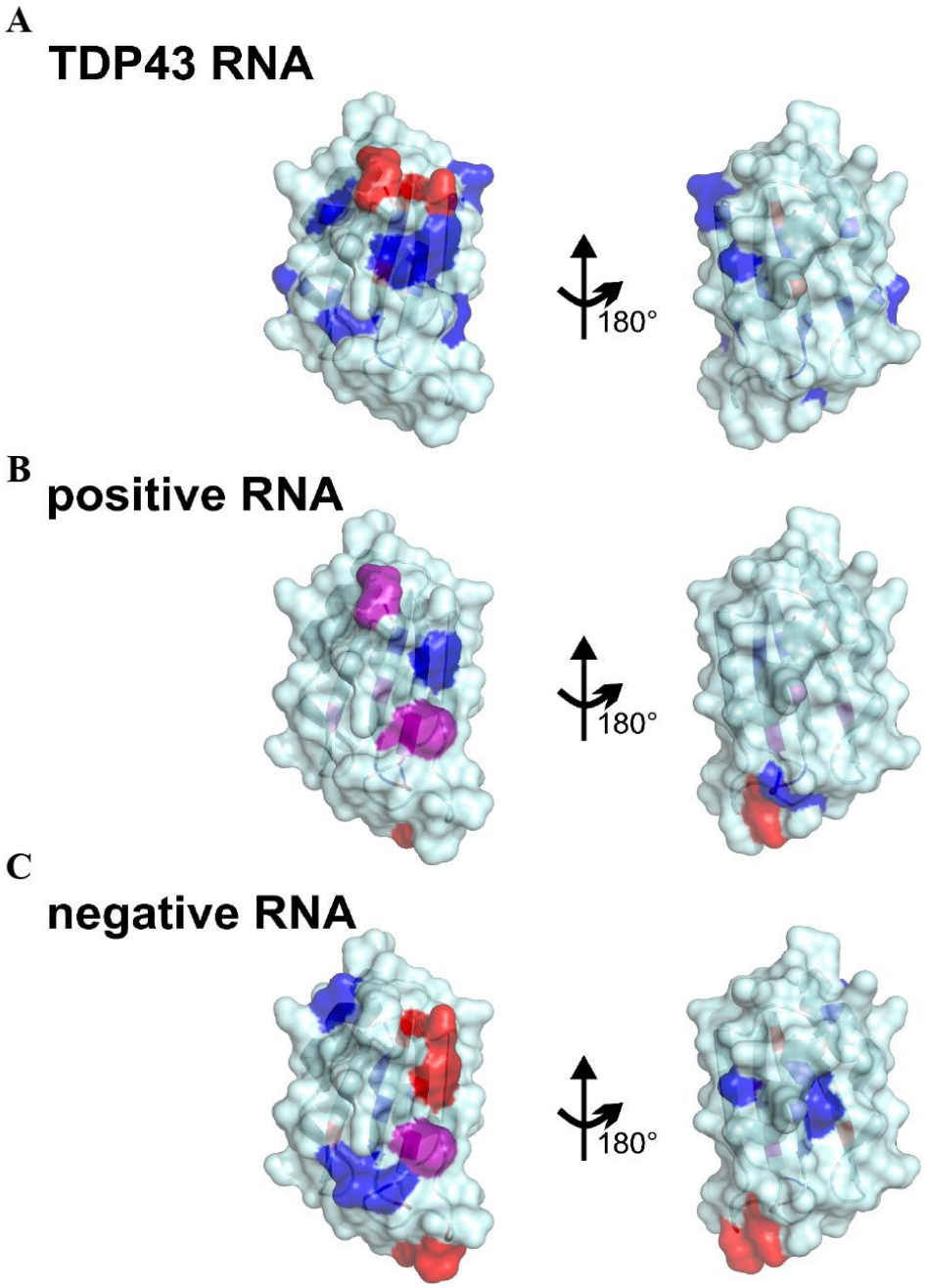


Figure 9. Surface representation of r*AB*-Elavl. Highlighted in blue the residues experiencing the largest intensity decreases, inred the residues experiencing the largest CSP, and in violet the residues experiencing the largest intensity decreases and CSP, in the presence of 35 µmol dm^{-3} TDP43 RNA (A), 35 µmol dm^{-3} positive RNA (B), and 35 µmol dm^{-3} negative RNA (C).

7. *AB*-Elavl and HuR inhibitors

As I proved before, *AB*-Elavl is characterized by two highly conserved peptides (RNP-1 and RNP-2), localized into its binding site. RNP-1 and RNP-2 are also characteristic of HuR. My lab developed small molecules able to dislocate the target RNA into the binding site of HuR (D'Agostino et al., 2013; Manzoni et al., 2018), and I wondered if those small molecules (SM) could also interact with *AB*-Elavl.

I screened the disrupting ability of our HuR inhibitors, named MFM48, GA03, GA13, MFM48, GD95, MB39 and VB35, against r*AB*-Elavl by EMSA assay at a final concentration of 10μM. I compared the activity of the SMs both on HuR (fig 10A) and *AB*-Elavl (fig 10B). I chose to focus my attention on those compounds which showed an effect on the bacterial protein but not on the human one. Particularly interesting is MFM48 (fig 10E), a monophenol able to reduce the formation of the complex bacterial protein-RNA and increase the free RNA signal but is inactive on the human one. As a control for all the following experiments, I used MFM49 (fig 10F) that is the corresponding ketone version of MFM48. Surprisingly MFM49 can reduce the binding signal for HuR, while is inactive on r*AB*-Elavl (fig 10A and 10B, arrows). GA03 and GD95 did not show any activity, while GA13, MB39 and VB 35 showed a mild activity on both proteins.

To confirm the activity of MFM48 and MFM49, I repeated the EMSA assay, testing three different concentrations of the SMs on *AB*-Elavl (20μM, 10 μM and 5 μM). The results showed again that while MFM49 is not able to disrupt the RNA binding ability of r*AB*-Elavl (fig 10C), MFM48 is instead able to increase the free RNA signal and reduce the binding (fig 10D, arrows). I then evaluated which is the minimal concentration of MFM48 able to inhibit the formation of the complex protein-RNA. I tested 7 different concentrations of the SMs ranging from 6.25μM to 500μM (fig 11A). I found the free RNA signal increases in intensity already at 12.5μM, but the binding signal remains visible also at the highest concentration.

I next investigated whether our SMs had a toxic effect on the *Acinetobacter baumannii* performing Minimal Inhibitory Concentration (MIC) experiments (we exposed *A. baumannii* directly to the SMs to evaluate the inhibition of the bacterial growth). Being a gram-negative bacterium, *Acinetobacter* has a double lipid membrane and a middle proteoglycan membrane that limit the passage of molecules. I could not find any significant results in the change of the growth ability using both MFM49 (fig 11B), and MFM48 (fig 11C).

I demonstrated that there could be the possibility that some SMs used to inhibit HuR can be also adapted for targeting the *AB*-Elavl. More studies on this topic are required since it is important to better understand which are the interactions that could lead to RNA displacement and which part of the protein is interested in the interaction with the SMs, and which mechanisms are used to let the molecules pass the proteoglycan wall.

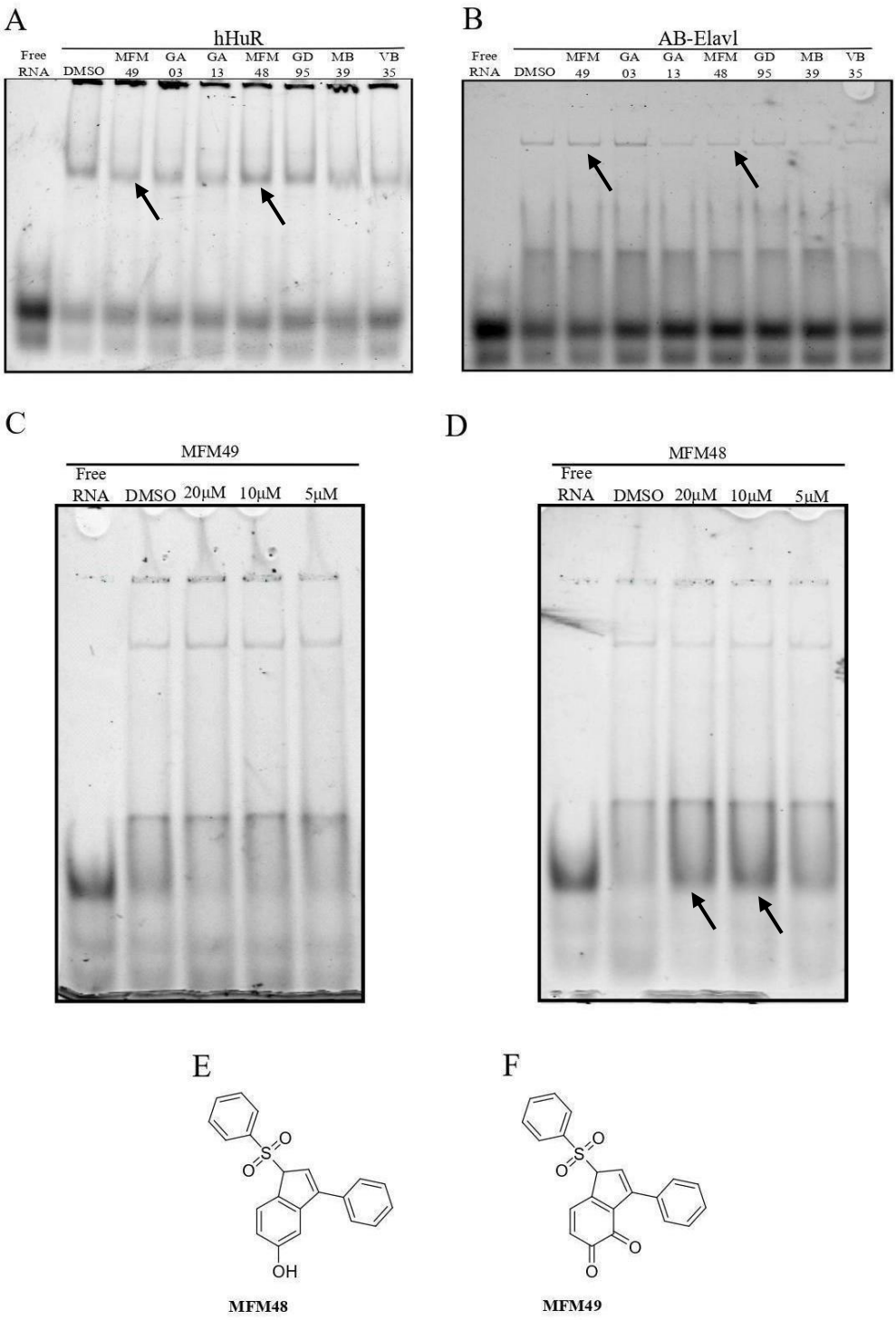


Figure 10. Small molecules activity on r*AB*-Elavl. A) EMSA assay for testing the ability of the small molecule to dislocate the formation of the complex rHuR-AREpos. MFM49 show a lower intensity in the shifted line, while MFM48 show a clear darker shifted band (see arrows). B) EMSA assay for testing the ability of the small molecule to displace the formation of the complex r*AB*-Elavl-AREpos. Unlike with rHuR, MFM49 shows a darker shifted band compared to MFM48 (see arrows). C) EMSA assay for evaluating if MFM49, is unable to interfere with the complex formation even at increasing concentration (5µM, 10 µM, 20 µM). D) EMSA assay for evaluating the ability of MFM48, at different concentrations, to dislocate the complex r*AB*- Elavl-ARE pos. E) Structure of MFM48. F) Structure of MFM49.

8. Pseudomonas aeruginosa

Acinetobacter baumannii belongs to the ESKAPE group. It is interesting to know if also the other bacteria of this group express orthologous RNA binding protein.

As it was done for *Acinetobacter*, I evaluated if the proteome of *Pseudomonas aeruginosa* (PA14) could recognize and bind the AREpos sequence, performing an EMSA in which I incubated the total proteome of PA14 with IR-AREpos. As for *A. baumannii*, I found a dose dependent increase of the signal at the level of the complex protein-RNA (fig 11E). The formation of a transient complex that reduces in intensity increasing in lysate concentration is clearly observed, evolving into a second heavier band (fig 11E, arrows). This suggests that the probe is bound by one or more proteins that can multimerize on the RNA.

I was eager to know if MFM48 was active also on another powerful pathogen. As *A. baumannii*, *P. aeruginosa* is a gram-negative bacterium, therefore characterized by a limited passage of molecules due to the double lipid membrane and the middle proteoglycan membrane. I tested MFM48 inhibitory activity on *P. aeruginosa*, by EMSA on the total proteome of *P. aeruginosa* and MIC assays. EMSA did not show any activity of MFM48 in the disruption of the complex protein-RNA. This is not surprising since I used the total protein lysate, suggesting there is more than one protein able to bind the ARE sequences (fig 11F).
I tested the minimal concentration needed for MFM48 in order to inhibit the growth of the bacterium (fig 11D): *P. aeruginosa* showed a similar trend to that of *A. baumannii*, and, also in this case, the inhibition is not statistically relevant.

These data suggest more studies should be done on *Pseudomonas aeruginosa*, repeating all the steps that were performed on *Acinetobacter baumannii*, since (as shown in fig 1D) proteins orthologous to *AB*-Elavl are present in several other bacteria.

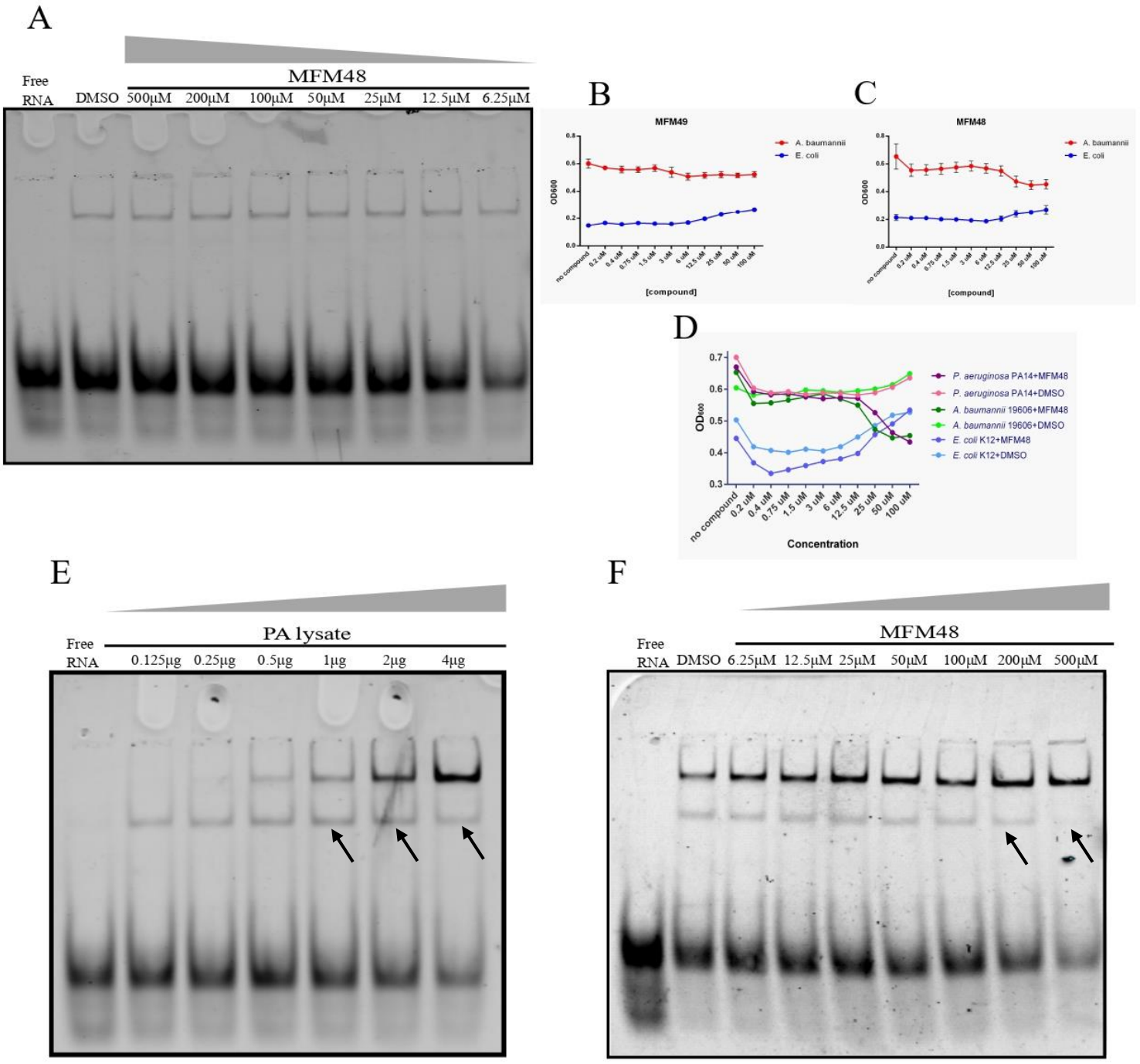


Figure 11. Small molecule activity on *A. baumannii* and *P. aeruginosa*. A) EMSA assay for testing the minimal concentration of MFM48 able to decrease the concentration of the complex r*AB*-Elavl-AREpos. B) MIC assays for evaluation of the activity of MFM49 on bacterial growth on *A. baumannii* (*E. coli* used as control). C) MIC assays for evaluation of the activity of MFM48 on bacterial growth on *A. baumannii* (*E. coli* used as control). D) MIC assay for evaluation of MFM48 on *A. baumannii*, *E. coli* and *P. aeruginosa*.
E) EMSA assay on the total proteins lysate of *P. aeruginosa* and the probe ARE pos. Arrows indicate the band of the transient complexes that evolve into heavier bands. F) EMSA assay with different concentrations of MFM48 to check the possible activity of MFM48 on *P. aeruginosa*.

Discussion

Not much information is available about the proteins responsible of mRNA regulation in bacteria other than *Escherichia coli* (Van Asche et al., 2015). Efforts for a better understanding of the transcriptional and post-transcriptional mechanisms are necessary, since several studies showed how *Acinetobacter*'s infectivity is linked to protein-protein interaction between pathogen and host cells (Asensio, 2021; "WHO Publishes List of Bacteria for Which New Antibiotics Are Urgently Needed," 2017). Indeed, *Acinetobacter baumannii* is able to evade the immune response and the action of antibacterial drugs by interacting with the cytoskeleton of the hosting cells, entering into the cell as a phagosome, or by creating biofilms able to protect the dormant bacteria against harsh environments, showing how the tight control of gene expression is fundamental for the survival and resistance of this pathogen (Gonzalez-Villoria & Valverde-Garduno, 2016; Malaka De Silva et al., 2020; Peleg et al., 2012). Furthermore, Hubloher et al. (Hubloher et al. 2021) showed that the silencing of RBPs can have dramatic effects on the bacterium. Indeed, the silencing of global post-transcriptional regulator Csr(A), impairs the growing abilities of *A. baumannii* in media rich in amino acids, alcohols and aromatic compounds. In fact, *Acinetobacter* is not able to use glucose as a carbon and energy source, thus the uptake of nutrients relies on other mechanisms like the one controlled by Csr(A). These results are important to delineate the main role played by Csr(A) in urinary tract infections, but also to underline the importance of RBPs in pathogenicity. Another evidence is Hfq (an RNA chaperone that modulates gene expression by helping the bind of sRNAs to their target) that, as *AB*-Elavl, is interesting from an evolutionary point of view, since it shares a common ancestor with the eukaryotic Sm-like proteins. In bacteria, Hfq has an important role as a virulence factor since its knockout leads to reduced growth rate and stress tolerance (Moll et al., 2003; Sharma et al., 2018).

For a better comprehension of the physiology of *A. baumannii*, I identified a new RNA binding protein which shares characteristics with one of the most studied human RNA binding proteins: HuR (Ma et al., 1996).

At the moment, less than 90 RRM domain-containing proteins are known in bacteria (mainly in Gram-negative bacteria), and all of them show just one single RRM domain per protein; in eukaryotes more than 6000 proteins were identified, often characterized from two to six RRM domains per protein (Koonin & Makarova, 2013; Mulani et al., 2019). The similarity score between HuR's RRMs and the *AB*-Elavl protein are in the same range as the similarity scores among the three human RRMs

(RRM1-RRM2: 32% id and 57% positives, RRM1-RRM3: 36% id and 53% positives, RRM2-RRM3: 30% id and 48% positives). This supports the hypothesis that the region found in the *Acinetobacter baumannii* genome encodes for a homologous protein (or at least its functional domain).

Furthermore, *AB*-Elavl displays significant homology with the RRM1 and RRM3 of HuR, suggesting that in animals, those two regions are under stronger purifying selective pressure, as they kept higher similarity with its putative ancient homologous gene (Marris et al., 2005).

I first investigated whether an mRNA was transcribed from the identified genomic locus, suggesting the eventual production of the encoded protein. I found that the sequence of our interest is inserted in a polycistronic mRNA. From the *A. baumannii*'s genome, I know *AB*-Elavl is positioned upstream to an ASCH domain containing protein and downstream to an ATP-dependent helicase (fig 1). PCR analyses confirmed the position of the gene within the polycistronic transcript. The presence of the protein into the proteome lysate was confirmed by western blot detection and mass spectrometry; more in detail I found the typical conserved sequences which characterize the RRM domains also in eukaryotes: RNP-1 and RNP-2, considered responsible of the binding with the RNA (Maruyama et al., 1999), (fig 2).

I cloned, expressed and purified the recombinant protein to test its biochemical properties. The r*AB*-Elavl was proved to be smaller compared to the wild one, as it was visible by western blot detection using the antibody developed directly on the denatured recombinant protein. Although there is a small difference in size, I believe that the most important part of the protein, responsible for the interaction with the RNA, are expressed anyway, as confirmed by mass spectrometry analysis. The r*AB*-Elavl in fact shows both the conserved domains in mass analysis, while just the starting peptides are missed (fig 2).

Based on the phylogenetic studies we performed, I tried to get more into the possible relationship with the ELAV-like family widely shared by eukaryotes (Samson, 2008). By NMR analysis we characterized the crystal structure of r*AB*-Elavl, which retrace the common structure of the classical RRM domain: β_1-α_1-β_2-β_3-α_2-β_4, in which RNP-1 and RNP-2 are in the internal strands of the beta-sheet (β_1-β_3) (fig 3) Holmqvist & Vogel, 2018; Maris et al., 2005). While in the proteins belonging to the Elav-like family the RRM domain is present twice or thrice, in *Acinetobacter*, the *AB*-Elavl is composed of just a single domain. Indeed, evolution led to an increase in the number and specificity of eukaryotic RBPs: they are often

characterized by a repetition of domains (that could be different or equals) that collaborate for a better affinity to the target RNA. On the contrary, bacteria tend to be more streamline, with simpler RBPs composed by just one single domain but with wider functions, since they are less specific for their targets, as it is the case for Csr(A), Hfq and several other RNA binding proteins in bacteria (Asensio et al., 2021; Iii et al., 2020; Kaur et al., 2021; Popovitchenko 2020).

In order to understand the role of *AB*-Elavl, I tried different biochemical assays: I tested the binding affinities of r*AB*-Elavl for RNA probes which mimic sequences known to be targeted by HuR. I first tested the presence in the *Acinetobacter*'s proteome of possible proteins able to recognize and bind a sequence characterized by a repetition of adenine and uracil by performing a REMSA assay. Several proteins in *Acinetobacter baumannii* were proved to recognize AU/U rich regions: Hfq recognizes the ARE sequence thanks to its Sm-like domain, protecting them from endonucleases' attach; RNAse E, one of the key factors for adaptation of *A. baumannii* to harsh environments, is a big endoribonuclease protein, composed by several subdomains and is able to recognize and cleave ARE sequences thanks to its highly conserved S1 subdomain (D. M. P. De Oliveira et al., 2020; Kuo et al., 2017; Stetefeld et al., 2016). As I was expecting, I found that at least one protein can target in a concentration dependent manner the probe, leading us to the next step (fig 5).

Since HuR is reported to specifically bind the AREpos but not the AREneg probe (D'Agostino et al., 2013), I evaluated the specificity of binding of r*AB*-Elavl. Effectively, the protein can target both AREpos and AREneg sequences but with one order higher affinity for the AREpos. After having developed an antibody for the r*AB*-Elavl, I tried to create a super shift in EMSA assay, to evaluate if the antibody can recognize the protein and drag the complex protein-RNA. I did not find a major electrophoretic retardation of the signal, possibly due to the fact that the antibody was developed on the denatured protein. Even the EMSA assay showed a preference of the protein toward the AU rich sequence compared to the AC rich peptides (fig 5).

As already stated, bacterial RBPs are able to exploit different functions and to bind different sequences (even if with more or lower affinity). Hfq has a wide substrate selection, underling the different roles covered by this protein: from RNA chaperon to ribosome biogenesis, DNA compaction, protein-protein interactions, and involvement in RNA degradation machinery. In order to check the specificity of *AB*-

Elavl, I also tested the affinity for TDP43's targeted sequence (a repetition of cytosine and guanine), which resulted to be comparable to the AREpos sequence (EC_{50}-TDP43: 8nM *versus* EC_{50}-AREpos: 24.27nM; Kd-TDP43: 20 nM *versus* Kd-AREpos: 13.2 nM). I measured the minimal length of the consensus sequence in AREpos's probe necessary to be bound by r*AB*-Elavl, that is equal to a minimum of 19 nucleotides, exactly as it was already shown for HuR (Ma et al., 1996) (fig 5). I further confirmed the results obtained with a time course kinetic assay: increasing the RNA concentration at different timepoints, there is an increase of the complex protein-RNA that also can be compared to the kinetic measured between HuR and the ARE sequences (Gallagher et al., 2015; Martínez & Vadyvaloo, 2014; Moll et al., 2003; Nishtala et al., 2016; C. Oliveira et al., 2017; Van Assche et al., 2015; H. Wang et al., 2013). To confirm this result, I evaluated the EC_{50} by FRET assay, which confirmed the same affinity I found with saturation, and time-course experiment by AlphaScreen (fig 6).

I could not measure the enthalpy of the formation of the complex protein-RNA by using the isothermal titration calorimetry, as well as I could not understand if the protein is able to dimerize or to form complex on the RNA, using the Dynamic Light Scattering assay. Indeed, I could not reach a sufficient concentration of the protein for an accurate measure since the maximum concentration I obtained was 12.5mg/L. However, I believe *AB*-Elavl, can dimerize on the RNA, since the Grating Coupled Interferometry analysis performed by binding *AB*-Elavl on a chip and incubating it with the AREpos probe did not give any signal of binding. I suggest this can happen because there is not formation of dimers of the protein on the RNA, lowering the affinity for the target sequence, as it is demonstrated for the single domains of HuR: cutting one of the RRM domain present in HuR, the binding abilities decrease (Wang et al., 2013) (fig 7). Dimerization of bacterial single domain-RNA binding protein on the target RNA is reported in literature: Csr(A) creates homodimers, and each homodimer can bind two strands of RNA (Christopoulou & Granneman, 2021; Iii et al., 2020).

The limited availability of the protein and the low affinity of the antibody for the bacterial folded protein made it difficult to study it with molecular approaches (fig 4). In fact, performing a pulldown of the total *A. baumannii*'s protein lysate, I could not detect any enrichment of the protein that clearly remained in the lysate. There is not any technical issue in the approach as I am able to pull down the recombinant protein. I was not able to obtain a visible enrichment of the immunoprecipitated complex by western blot. However, we could confirm the presence of *AB*-Elavl by mass spectrometry, again suggesting the low production

of the protein at basal condition into the bacteria. MS analysis was able to identify three possible candidates from the Uniprot database that share a high percentage of similarity with our protein. It is interesting that all of them are characterized by a lower molecular weight compared to our recombinant protein, that is also smaller compared to the bacterial one (fig 4).

Taking together these last results, about the molecular weight and the inability of the antibody to bind properly the folded wild type protein, I suggest that either the bacterial protein is longer than what expected, is characterized by a post-transcriptional modification or is partially degraded during its purification.

RBPs were found upregulated in MDR *Acinetobacter baumannii* (e.g.: regulators of transcription, RNAseE, enolase), suggesting an involvement in the resistance mechanisms (Wang et al. 2019). Since *AB*-Elavl and HuR share a high conservation of RNP-1 and RNP-2, and, since our lab is currently working into the characterization of small molecules able to target HuR and inhibit the binding with the target RNA, I tested some molecules on r*AB*-Elavl and on the bacterial susceptibility. In particular, I was looking for molecules that were not active on the human protein but that could exert some inhibition of the binding of r*AB*-Elavl with the RNA target. I found an interesting derivative of Dihydrotanshinone-I: MFM48. Dihydrotanshinone-I is an anti-inflammatory agent, proven to target HuR and to interfere with its binding to the target RNA (Lal et al., 2017). MFM48 is able to reduce the electrophoretic retardation of the signal and to increase the signal of the free RNA when incubated with r*AB*-Elavl but is not active on HuR. MFM48 is the monophenol version of another tanshinone mimic, MFM49, that instead behaves in the opposite way: it reduces the binding signal with HuR, but not with r*AB*-Elavl. I confirmed the activity of both molecules making a titration of different concentrations in EMSA assay, obtaining a dose-response curve (fig 10). In the MIC assay, I could not find a significant decrease of the bacterial growth using MFM48. One possibility is that the small molecules were not able to enter the cells, but other options are clearly possible, as the fact they may not be toxic for the bacteria. However, there is the need of further studies in order to clearly understand the mechanism of action and the concentration range needed to obtain a proper minimalinhibitory concentration (fig 11).

A. baumannii belongs to the ESKAPE organisms (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter*). These are bacteria able to survive to all the treatments available at the moment by avoiding the action of the

antibacterial molecules (Asensio, 2021). I was also interested in other bacteria from this group in order to compare them to *A. baumannii*. Using PA14 strain for *Pseudomonas aeruginosa*, I evaluated the presence of proteins able to recognize the ARE sequence using EMSA assay. The results were interesting since, not only I found a dose- dependent response, but it is also visible the formation of a progressive heavier electrophoretic retardation due to an increased number of proteins bound to the probe (fig 11). This is in line with what is reported in literature, since *P. aeruginosa* expresses an ortholog of Hfq that, as well, is reported to bind AU/U rich sequences (Sharma et al., 2018). Inspired by these results I also tried to evaluate if MFM48 could have in *P. aeruginosa* the same effects I found in *A. baumannii*. I could observe a diminished signal of the lighter complex in EMSA assay at higher MFM48's concentration (fig 11).

Taken together, these results suggest *Acinetobacter baumannii* express an RNA binding protein that shares characteristics similar to the mammalian Elav-like family proteins, or, more broadly with RRM containing RBPs. However, the RNA binding ability appears to be less specific than the human counterpart. Even more interestingly, I confirmed the presence of highly conserved sequences and structure, which suggest a probable common ancestor with the Elav-like family. I evaluated the ability of this protein to be inhibited by small molecules developed to target HuR and I tried also to include in this analysis *Pseudomonas,* belonging it as well to the ESKAPE organisms.

More studies are needed in order to understand the role of this protein into the bacteria, not only as an RNA binding protein but also in terms of essentiality to the bacterial life.

Conclusions and future perspectives

Through this thesis, I have elucidated the biochemical characteristic of an RNA binding protein expressed by *Acinetobacter baumannii*.

I evaluated if it could be possible to find a bacterial orthologous of one of the most studied human RNA binding proteins: HuR. I was able to confirm the production of *AB*-Elavl by finding the polycistronic mRNA responsible for the protein transcription as well as several peptides belonging to the protein. I also confirmed its ability to bind sequences rich in adenine and uracil (mimicking HuR behaviour) even though *AB*-Elavl is less specific compared to the human counterpart since it can also bind other targets such as TDP43's consensus sequence and ARE negative probe. This is not surprising, since in bacteria, RBPs show less specific binding properties.

This study has opened new lights on the ancient origin of the ELAV-like family of gene, deriving possibly from LUCA. I also increased the knowledge on the proteome and of *Acinetobacter baumannii*, sadly known for its multidrug resistance.

It will be of great interest to study the phenotype of mutated *Acinetobacter baumannii* for this gene. Indeed, Gallagher et. al (2015) were able to annotate the genome sequence of *Acinetobacter baumannii*'s strain AB5075 (isolated in 2008 from a combatant wound infection, showing a high virulence and multiple antibiotic resistances in animal models). They created a transposon mutant library of the strain freely available on internet. In this database I found two different transposon insertions into my gene of interest. This suggests that *AB*-Elavl is not an essential gene. In the future, it will be extremely interesting to study the phenotype of *A. baumannii* mutant, in which the gene producing the protein of interest has been inactivated. In particular, I would like to evaluate if there are changes at the level of drug resistance, stress response and viability, as well as finally characterize the role of *AB*-Elavl, understanding which target sequences are recognized and bound and finally, to exploit its role on the destiny of the bound RNA.

Materials and methods

1. Bioinformatics analysis to search homologous proteins to ELAV-like protein in *Acinetobacter baumannii*.

The human HuR protein sequence was used as a query to search for the most similar protein in the *Acinetobacter baumannii* genome, using tblastn on the NCBI web server. The search was restricted to the species *A. baumannii* within the RefSeq Genome Database. The best scoring hit (i.e. the *A. baumannii* protein displaying highest similarity with the human HuR) was searched in all *A. baumannii* genomes available using tblastn, and in all other bacterial genomes using the Ortholuge database (Whiteside et al. 2013).

The sequences of the bacterial ELAV-like proteins, found as described above, were submitted to the MEME-suite tool MEME v5.3.3 (Bailey et al., 2009), to find conserved motifs. We performed the search using the following parameters: -mod zoops -nmotifs 50 -minw 6 -maxw 10 (that means search for at least 50 motifs occurring zero or one time per sequence and spanning 6-10 aa in length).

2. Preparation and detection of recombinant orthologous human ELAV-like protein in Acinetobacter baumannii.

The mRNA of the orthologous of the ELAV-like family in *Acinetobacter baumannii* was retro-transcribed into cDNA and the sequence was amplified and inserted into the pET30a(+) vector (GenScript) by using the forward (5'-CGGC CATATG ATACTCAAATGTATA-3') and reverse (5'-ATAT CTCGAG CTCTTCAGCTGCCTT-3') primers containing the NdeI and the XhoI restriction sites, respectively. Frame and sequence of the full-length ORF, with the His tag-encoding sequence located at the 3'-end, was confirmed by Sanger sequencing. The recombinant vector pET30a(+)-Elavl was amplified in competent *E.coli* Top10 and the recombinant protein has been expressed into *E. coli* Rosetta BL21. Overnight cultures of *E. coli* BL21 were diluted at 1:50 with the LB medium. At A_{600} of 0.5, cultures were induced with isopropyl β-D-thiogalactoside (IPTG) at 0.2 mM, and grown overnight at 18°C. Cells were spun down and lysed in buffer containing 20mM Hepes pH 7.5, 300mM NaCl, 3mM $MgCl_2$, and 0.5 mg/ml Proteases Inhibitor Cocktail (Leupeptin, Aprotinin and Pepstatin from Sigma Aldrich) and then centrifuged at 16.000xg for 30 min at 4°C. The supernatant was incubated with Ni-

NTA Agarose beads (Ni-NTA Agarose, Qiagen GmbH) for 2 h at 4°C. After washing the beads with buffer A (20mM Hepes, pH 7.5, 150mM NaCl, 3mM $MgCl_2$ and 20mM imidazole), buffer B (as buffer A but containing 50mM imidazole) and buffer C (as buffer A but containing 100mM imidazole), protein was eluted with buffer D (as buffer A but containing 250mM imidazole). The eluted protein was dialyzed against storage buffer (20mM Hepes, pH 7.5, 150mM NaCl, 3mM $MgCl_2$, 5% glycerol) and stored at -80°C (Bhattacharyya & Filipowicz, 2012). Recovered recombinant protein was analyzed by Coomassie staining on 12%-SDS PAGE. The relative protein concentration was determined in three different ways: using bovine serum albumin (BSA) standards and densitometry quantification (ImageJ 1.4 software, NIH) of corresponding bands on acrylamide gels, using the Bradford assay and by UV-vis spectrometry using the molar extinction coefficient.

3. Polyethyleneimine (PEI) treatment.

In order to eliminate the high concentration of nucleic acids from the purified protein, I treated the dialyzed protein with PEI, a cationic polyelectrolyte able to disrupt the interaction of protein-nucleotides by displacing the protein. PEI was added to a final concentration of 0,8% and incubated for 30 min at 4°C. Soluble proteins were harvested by centrifugation at 8000rpm for 30 min. Supernatant was collected and washed twice with ammonium sulphate at a final saturation of 70% (calculated with the ENCOR calculator) for 30 min at 4°C. Proteins were precipitated by centrifugation at 8000rpm for 30 min. The pellet was finally resuspended in the dialysis buffer.

4. Expression and purification of r*AB*-Elavl for X-ray and NMR analysis.

r*AB*-Elavl in plasmid pET-3a(+) was overexpressed in BL21(DE3) GOLD cells. Cells were grown in LB or M9 minimal media supplemented with 15NH4Cl or 15NH4Cl and 13C-glucose at 37 ºC until optical density (OD_{600}) reached 0.6-0.8. Subsequently, protein production was induced with 0.2 mmol dm-3 of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 18 ºC overnight and harvested by centrifugation at 4 ºC, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer (50 mmol dm-3 HEPES, pH 6.8, 300 mmol dm-3 NaCl, 3 mmol dm-3 $MgCl_2$, 1 mmol dm-3 inhibitor proteases), ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was collected and treatment with 5% PEI solution was performed in order to remove DNA/RNA attached to the protein. Re-suspension of the protein was performed with the lysis buffer. Soluble protein was filtered with a 0.22 µm membrane and purified by a Ni^{2+}-

affinity chromatography step using a His-Trap HP 5 cm3 column previously equilibrated in 50 mmol dm-3 HEPES pH 6.8, 300 mmol dm-3 NaCl, 3 mmol dm-3 MgCl2, 1 mmol dm-3 inhibitor proteases. r*AB*-Elavl was eluted with increasing concentration of imidazole (20 - 50 - 100 - 250 mmol dm-3) in the buffer and subsequently dialyzed overnight against 4 dm3 of 20 mmol dm-3 HEPES buffer at pH 6.8, containing 150 mmol dm-3 NaCl and 3 mmol dm-3 $MgCl_2$. The protein was filtered and further purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column that was previously equilibrated in 20 mmol dm-3 HEPES pH 6.8, 150 mmol dm-3 NaCl, 3 mmol dm-3 $MgCl_2$ and 1 mmol dm-3 proteases inhibitors.

5. Crystallization of r*AB*-Elavl

r*AB*-Elavl was concentrated to 6 mg cm-3 in 20 mmol dm-3 HEPES buffer pH 6.8, containing 150 mmol dm-3 NaCl, 3 mmol dm-3 $MgCl_2$ and 1 mmol dm-3 proteases inhibitors. 1.6 Å crystals were obtained by sitting drop vapor diffusion at 293 K, in which 5 mm3 of protein solution was mixed with 5 mm3 of reservoir solution and suspended over 600 mm3 of the same reservoir solution. Crystals were obtained in the reservoir solution consisting of 0.1 mol dm-3 sodium acetate trihydrate pH 4.5, 3 mol dm-3 sodium chloride and optimized with 15% of cryo-protector ethylene glycol.

6. X-ray data collection

The dataset was collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K; the crystal used for data collection were cryo-cooled using 25% ethylene glycol in the mother liquor. The crystals diffracted up to 1.6 Å resolution: they belong to space group I41 with one molecule in the asymmetric unit, a solvent content of about 50%, and a mosaicity of 0.3°. The data were processed using the program XDS, reduced and scaled using XSCALE and amplitudes were calculated using XDSCONV. The structure was solved using the molecular replacement technique; the model used was obtained through MODELLER. The successful orientation hand translation of the molecule within the crystallographic unit cell was determined with MOLREP. The refinement was carried out using PHENIX, applying TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT. The quality of the refined structures was assessed using the program MOLPROBITY.

7. NMR measurements and protein assignment.

Experiments for backbone assignment were performed on samples of the 13C, 15N isotopically enriched RRM domain of r*AB*-Elavl at protein concentration of 300 µmol dm-3 in buffer solution (20 mmol dm-3 HEPES, pH 6.8, 150 mmol dm-3 NaCl, 3 mmol dm-3 MgCl2, 1 mmol dm-3 protease inhibitors). NMR spectra were recorded at 298 K on a Bruker AVANCE NEO 900 spectrometer, equipped with a triple-resonance Cryo-Probe. Spectra were processed with the Bruker TOPSPIN software packages and analyzed with CARA (Computer Aided Resonance Assignment, ETH Zurich). The backbone resonance assignment of RRM domain was obtained by the analysis of 3D HN(CO)CA, 3D HNCA, 3D HNCO, 3D HN(CA)CO, 3D CBCA(CO)NH and 3D HNCACB spectra. Secondary structure prediction was performed with TALOS+(8) by using the chemical shifts of HN, N, C', Cα, and Cβ as input data.

8. Titration of r*AB*-Elavl with RNA.

The effect of three different types of RNA (TDP43-M, AREpos and AREneg) on the 15N-isotopically enriched RRM domain of *AB*-Elavl (70 mmol dm-3) was evaluated in the following experimental conditions: 20 mmol dm-3 HEPES, pH 6.8, 150 mmol dm-3 NaCl, 3 mmol dm-3 MgCl2, 1 mmol dm-3 protease inhibitors. 2D 1H 15N BEST-TROSY. NMR spectra were acquired at 298 K on a Bruker Avance III 950 MHz NMR spectrometer to monitor the effect of increasing amounts (17.5, 35, 52.5, 70 µmol dm-3) of each RNA added to the protein solution.

9. RNA-Electrophoresis Mobility Shift Assay (REMSA).

rELAV-like protein (at indicated concentrations) and RNA probes with DY681 infra-red tag (at a concentration of 2.5nM) were incubated, in REMSA buffer (20 mM HEPES pH 7.5, 50 mM KCl, 0.5 µg BSA, 0.25% Glycerol) in a final volume of 20µl at room temperature. The reaction mix was then loaded onto 6% native polyacrylamide gel containing 0.5% Glycerol. Run was performed in 0.5X TBE buffer at 80 V for 40 min and then 100V for 20 min, at 4° C. Free and complexed RNA probe were detected with Odyssey infrared Imaging System (LI-COR Odissey Infrared Imager Biosciences) using filters for red light emission detection. (D'Agostino et al., 2019; Manzoni et al., 2018)

10. Amplified Luminescent Proximity Homogeneous Assay (ALPHA Screen).

AlphaScreen assays have been performed using histidine (nickel) chelate detection kit (Histidine detection kit Nickel Chelate 6760619C, PerkinElmer) in white 384 Optiplates. AlphaScreen assay was applied to study the interaction between rELAV-like protein and the different biotinylated single-stranded probes: ARE pos (5'-Bi-AUUAUUUAUUAUUUAUUUAUUAUUUA-3'), ARE pos 19 (5'-Bi-AUUAUUUAUUAUUUAUUUA-3'), ARE pos 11 (5'-Bi-AUUAUUUAUUA-3'), ARE neg (5'-Bi-ACCACCCACCACCCACCCACCACCCA-3') and TDP43 (CCG GGG CCG GGG CCG GGG CCG GGG) (Eurofins Genomics). All reagents were reacted in ALPHA buffer (25 mM HEPES pH 7.4, 100 mM NaCl, 0.1% BSA). For the optimization of the assay, it was check which is the best protein:RNA ratio (hook point): a series of concentrations of the recombinant protein (0 – 40μM) were incubated with different concentrations of ARE pos probe (0 – 500 nM). Once found the best RNA's concentration, it was calculated the hook point of the protein also in incubation with the ARE neg and TDP43's probe at a final concentration of 50nM. For the Kd calculation 500 nM of the r*AB*-Elavl protein was incubated with a series of concentrations of probes (0 – 500nM) for 15 min at room temperature, then anti-His-Acceptor beads (20 μg/ml final concentration) and Streptavidin-Donor beads (20 μg/ml final concentration) were added and reaction was incubated in the dark at room temperature for 60 min to reach equilibrium. Fluorescence signals were detected on Enspire plate reader instrument (PerkinElmer; 2300 Multilabel Reader). Non-specific interference with the assay has been evaluated by reacting the same amount of acceptor and donor beads (20 μg/mL/well) without the probe and with just the protein buffer in the same experimental conditions. The half maximal effective concentration (EC_{50}) was calculated with GraphPad Prism software v6.1. (D'Agostino et al., 2019; Manzoni et al., 2018)

11. Time course experiments kinetic

Time course experiments were carried out incubating in a final volume of 20μL, a series of concentrations (0–50 nM) of the RNA probe (Bi-ARE pos, Bi-AREneg or Bi-TDP43) with a constant concentration of r*AB*-Elavl protein (500 nM) and anti-His-Acceptor beads (20 μg/ml) and Streptavidin-Donor beads (20μg/ml) in ALPHA buffer, as described above. Assays were performed in triplicate. The wells were all seeded with a cocktail containing Alpha buffer and beads, while r*AB*-Elavl protein and probes were added in a second moment, according to the time checkpoints. The

signals of the whole 384-well plate were detected at the end of the time course. Association and dissociation rate constants were determined from nonlinear regression fits of the data according to association kinetic model of multiple ligand concentration in GraphPad Prism®, version 6.1. The resulting Kd values obtained by koff/kon ratio were compared with the Kd of classical AlphaScreen assay (D'Agostino et al., 2019).

12. Protein lysate of *Acinetobacter baumannii*

Acinetobacter baumannii strain ATCC 19606 was grown in Luria-Bertani (LB) medium, in the incubator shaker at 200 rpm at 37°C. Inoculum was grown overnight and the next day it was diluted to a final concentration of 0.05 OD. The bacteria were let grow to a final OD of 0.5 (they were measured at the spectrophotometer at a λ: 600nm) and then spinned at 4000rpm for 20 min at 4°C. The pellet was incubated for 30 min in ice with lysis buffer (50mM tris HCl pH7.5, 100mM NaCl, 10% glycerol, 0.1% triton, 1mM DTT, 1mM EDTA, Leupetine, Aprotinin, Lysozime, 2.5U/μL) to a final volume equal to 1/20 of the initial culture, and then sonicated.

13. Western blot

Human cells were rinsed with PBS and lysed in ice-cold RIPA buffer, while bacteria were lysed in a bacterial lysis buffer (20mM Tris Hcl pH8, 150mM KCl, 1mM MgCl2, 1mM DTT, DNAsi, Proteinase inhibitors and RNAses inhibitors). Proteins were boiled in SDS gel sample buffer, separated by SDS-PAGE and immunoblotted onto a polyvinylidene difluoride membrane. Primary antibody against *AB*-Elavl was developed by Davids biotechnologie in rabbit, while the antibody against HuR is HuR (6A97) from Santa Cruz Biotechnologies. Bands were visualized with anti-rabbit or anti-mouse HRP-conjugated secondary antibodies and scanned on Biorad Chemidoc.

14. Time resolved fluorescence resonance energy transfer (HTRF-FRET)

All assays were performed in 20μL in 96 wells low-volume white plates. All assays were performed in triplicate. First it was found the best protein:RNA ratio by testing different concentrations of protein with the probes Bi-AREpos and Bi-TDP43 at a final concentration of 50nM (the one that was also used for AlfaScreen). Once found

the hook point, it was measured the saturation ability of r*AB*-Elavl by calculating the Kd testing increasing concentrations of RNA. The experiments were performed by incubating the protein with the RNA for few min before to add the mix composed of beads (Acceptors beads europium-labeled anti-6X His-Antibody and donors beads XL665–conjugated for biotin detection at a final concentration of 35nM), Kf buffer and FRET reaction buffer 1x provided by the kit. The plate was spin at 1000rpm for 1 min and incubated for 1 hour at 4°C. The single signals of acceptors and donors were detected using Tecan Spark and the final results were calculated using the following equation: Acceptors/Donors*10000.

15. Biotinylated RNA Pull Down assay (PD)

For each PD sample it was used 1mg of total protein lysate from *Acinetobacter baumannii*, lysed in RIP buffer (20mM Tris Hcl pH8, 150mM KCl, 1mM MgCl2, 1mM DTT, DNAsi, Proteinase inhibitors and RNAses inhibitors). It was followed the protocol adapted from Panda et al., 2016. Beads were rinsed in TENT buffer (10mM Tris HCl pH8, 1mM EDTA pH8, 250mM NaCl, 0,5% Triton). Lysate was pre-cleared with a 10µL of Streptavidin Magnetic Beads (Life technologies, 11205D) and then incubated with 2.5µM of the probe for 1 hours at 4°C in TENT buffer. Solutions were incubated for further 2 hours with 20µL/sample of streptavidin magnetic beads. 10% of the total lysate for each sample has been stored as Input. Analysis of the results were performed by western blot assay.

16. Immunoprecipitation (IP) assay.

The protocol was adapted from the one published by Keene et al., 2006, taking just the part related to immunoprecipitation of the protein. For each single IP it was used 2.5mg of total protein lysate from *Acinetobacter baumannii*, lysed in RIP lysis buffer (20mM Tris Hcl pH8, 150mM KCl, 1mM $MgCl_2$, 1mM DTT, DNAsi, Proteinases inhibitors and RNAses inhibitors). Lysate was incubated with Pierce A/G beads (Thermo Scientific Pierce 88847-88848) for pre-clearing steps 2 hour at 4°C; in parallel 50% A and 50% G beads have been incubated either with 10µG of α-r*AB*-Elavl or 10µG of IgG antibodies for ab-coating step for 2 hours at RT. At the end of the 2 hours of incubation, the protein lysate was incubated with antibodies and beads overnight at 4°C. Finally, samples have been washed (5 times, 5 minutes each wash)

with NT2 buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1mM $MgCl_2$, 0.05% NP40). The pellet was then analysed by western blot or mass spectrometry assay.

17. Minimal Inhibitory Concentration (MIC)

The activity of the different small molecules was tested by determining the MIC in liquid media. The MIC was defined as the lowest concentration of small molecules that prevents visible growth of the bacteria.

Serial dilutions, ranging from 0 µM to 100 µM of each small molecule, were made in 96 well-plates using the Muhler Hinton medium. *Acinetobacter baumannii* was inoculated into each well to an approximate starting OD_{600} of 0.0001. For endpoint MIC readings, plates were read at 600 nm for the initial OD (t = 0) and then incubated in an incubator at 37 °C. After 24h, the plates were read again at 600 nm.

References

1. Holmqvist E, Vogel J. RNA-binding proteins in bacteria. *Nat Rev Microbiol*. 2018;16(10):601-615. doi:10.1038/s41579-018-0049-5

2. Sharma A, Dubey V, Sharma R, et al. The unusual glycine-rich C terminus of the Acinetobacter baumannii RNA chaperone Hfq plays an important role in bacterial physiology. *J Biol Chem*. 2018;293(35):13377-13388. doi:10.1074/jbc.RA118.002921

3. Christopoulou N, Granneman S. The role of RNA-binding proteins in mediating adaptive responses in Gram-positive bacteria. *FEBS J*. Published online 2021. doi:10.1111/febs.15810

4. Regulation R. crossm. 2017;199(11):1-15.

5. Quendera AP, Seixas AF, dos Santos RF, et al. RNA-Binding Proteins Driving the Regulatory Activity of Small Non-coding RNAs in Bacteria. *Front Mol Biosci*. 2020;7(May):1-9. doi:10.3389/fmolb.2020.00078

6. Wang P, Li RQ, Wang L, Yang WT, Zou QH, Xiao D. Proteomic Analyses of Acinetobacter baumannii Clinical Isolates to Identify Drug Resistant Mechanism. *Front Cell Infect Microbiol*. 2021;11(February):1-15. doi:10.3389/fcimb.2021.625430

7. Gallagher LA, Bailey J, Manoil C. Ranking essential bacterial processes by speed of mutant death. *Proc Natl Acad Sci U S A*. 2020;117(30):18010-18017. doi:10.1073/pnas.2001507117

8. Gonzalez-Villoria AM, Valverde-Garduno V. Antibiotic-Resistant Acinetobacter baumannii Increasing Success Remains a Challenge as a Nosocomial Pathogen . *J Pathog*. 2016;2016:1-10. doi:10.1155/2016/7318075

9. Malaka De Silva P, Patidar R, Graham CI, Brassinga AKC, Kumar A. A response regulator protein with antar domain, avnr, in acinetobacter baumannii ATCC 17978 impacts its virulence and amino acid metabolism. *Microbiol (United Kingdom)*. 2020;166(6):554-566. doi:10.1099/mic.0.000913

10. Peleg AY, de Breij A, Adams MD, et al. The Success of Acinetobacter Species; Genetic, Metabolic and Virulence Attributes. *PLoS One*. 2012;7(10). doi:10.1371/journal.pone.0046984

11. Koonin E V, Makarova KS. Koonin and Makarova 2013. 2013;(May):679-686.

12. Mulani MS, Kamble EE, Kumkar SN, Tawre MS, Pardesi KR. Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review. *Front Microbiol*. 2019;10(APR). doi:10.3389/fmicb.2019.00539

13. Iii JMF, Wells G, Palethorpe S, Adams MD, Pesci EC. crossm Associated Growth of Acinetobacter baumannii. 2020;(May):1-19.

14. Kaur H, Kalia M, Taneja N. Identification of novel non-homologous drug targets against Acinetobacter baumannii using subtractive genomics and comparative metabolic pathway analysis. *Microb Pathog*. 2021;152(November 2020):104608. doi:10.1016/j.micpath.2020.104608

15. Asensio NC, Rendón JM, Burgas MT. Time-resolved transcriptional profiling of epithelial cells infected by intracellular acinetobacter baumannii. *Microorganisms*. 2021;9(2):1-14. doi:10.3390/microorganisms9020354

16. Tatiana Popovitchenko ALL RIGHTS RESERVED ©2020. Published online 2020.

17. De Oliveira DMP, Forde BM, Kidd TJ, et al. Antimicrobial resistance in ESKAPE pathogens. *Clin Microbiol Rev*. 2020;33(3):1-49. doi:10.1128/CMR.00181-19

18. Stetefeld J, McKenna SA, Patel TR. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys Rev*. 2016;8(4):409-427. doi:10.1007/s12551-016-0218-6

19. Kuo HY, Chao HH, Liao PC, et al. Functional characterization of Acinetobacter baumannii Lacking the RNA chaperone Hfq. *Front Microbiol*. 2017;8(OCT):1-12. doi:10.3389/fmicb.2017.02068

20. Nishtala S, Neelamraju Y, Janga SC. Dissecting the expression relationships between RNA-binding proteins and their cognate targets in eukaryotic post-transcriptional regulatory networks. *Sci Rep*. 2016;6(April):1-14. doi:10.1038/srep25711

21. Oliveira C, Faoro H, Alves LR, Goldenberg S. RNA-binding proteins and their role in the regulation of gene expression in trypanosoma cruzi and saccharomyces cerevisiae. *Genet Mol Biol*. 2017;40(1):22-30. doi:10.1590/1678-4685-GMB-2016-0258

22. Martínez LC, Vadyvaloo V. Mechanisms of post-transcriptional gene regulation in bacterial biofilms. *Front Cell Infect Microbiol*. 2014;5(MAR):1-15. doi:10.3389/fcimb.2014.00038

23. Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, Bläsi U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *Rna*. 2003;9(11):1308-1314. doi:10.1261/rna.5850703

24. Gallagher LA, Ramage E, Weiss EJ, et al. Resources for genetic and genomic analysis of emerging pathogen Acinetobacter baumannii. *J Bacteriol*. 2015;197(12):2027-2035. doi:10.1128/JB.00131-15

25. Van Assche E, Van Puyvelde S, Vanderleyden J, Steenackers HP. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol*. 2015;6(MAR):1-16. doi:10.3389/fmicb.2015.00141

26. Wang H, Zeng F, Liu Q, et al. The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr Sect D Biol Crystallogr*. 2013;69(3):373-380. doi:10.1107/S0907444912047828

27. Robinow S, White K. Characterization and spatial distribution of the ELAV protein during Drosophila melanogaster development. *J Neurobiol*. 1991;22(5):443-461. doi:10.1002/neu.480220503

28. Takada H, Roghanian M, Murina V, et al. The C-Terminal RRM/ACT Domain Is Crucial for Fine-Tuning the Activation of 'Long' RelA-SpoT Homolog Enzymes by Ribosomal Complexes. *Front Microbiol*. 2020;11(February):1-16. doi:10.3389/fmicb.2020.00277

29. Romeo T, Babitzke P. Global Regulation by CsrA and Its RNA Antagonists. *Microbiol Spectr*. 2018;6(2):1-24. doi:10.1128/microbiolspec.rwr-0009-2017

30. Wang S, Overgaard MT, Hu YX, McKay DB. The Bacillus subtilis RNA helicase YxiN is distended in solution. *Biophys J*. 2008;94(1):1-3. doi:10.1529/biophysj.107.120709

31. Karginov F V., Caruthers JM, Hu YX, McKay DB, Uhlenbeck OC. YxiN is a modular protein combining a DExD/H core and a specific RNA-binding domain. *J Biol Chem*. 2005;280(42):35499-35505. doi:10.1074/jbc.M506815200

32. Mardle CE, Shakespeare TJ, Butt LE, et al. A structural and biochemical comparison of Ribonuclease E homologues from pathogenic bacteria highlights species-specific properties. *Sci Rep*. 2019;9(1):1-11. doi:10.1038/s41598-019-44385-y

33. Wiegand I, Hilpert K, Hancock REW. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc*. 2008;3(2):163-175. doi:10.1038/nprot.2007.521

34. Marenchino M, Armbruster DW, Hennig M. Rapid and efficient purification of RNA-binding proteins: Application to HIV-1 Rev. *Protein Expr Purif*. 2009;63(2):112-119. doi:10.1016/j.pep.2008.09.010

35. Bailey TL, Boden M, Buske FA, et al. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37(SUPPL. 2):202-208. doi:10.1093/nar/gkp335

36. Whiteside MD, Winsor GL, Laird MR, Brinkman FSL. OrthologeDB: A bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res*. 2013;41(D1):366-376. doi:10.1093/nar/gks1241

37. Hubloher JJ, Schabacker K, Müller V. CsrA Coordinates Compatible Solute Synthesis in Acinetobacter baumannii and Facilitates Growth in Human Urine. 2021

38. Chihara K, Bischler T, Barquist L, et al. Conditional Hfq Association with Small Noncoding RNAs in Pseudomonas aeruginosa Revealed through Comparative UV Cross-Linking Immunoprecipitation Followed by High-Throughput Sequencing. *mSystems*. 2019;4(6). doi:10.1128/msystems.00590-19

39. Pérez-ràfols A. Advanced Drug Delivery Reviews HuR-targeted agents : an insight into medicinal chemistry , biophysical , computational studies and pharmacological effects on cancer models .

40. Leoncini E, Ricciardi W, Cadoni G, et al. Adult height and head and neck cancer: A pooled analysis within the INHANCE Consortium. *Head Neck*. 2014;36(10):1391. doi:10.1002/HED

41. Benkő R, Gajdács M, Matuz M, et al. Prevalence and antibiotic resistance of eskape pathogens isolated in the emergency department of a tertiary care teaching hospital in hungary: A 5-year retrospective survey. *Antibiotics*. 2020;9(9):1-17. doi:10.3390/antibiotics9090624

42. Ma WJ, Cheng S, Campbell C, Wright A, Furneaux H. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J Biol Chem*. 1996;271(14):8144-8151. doi:10.1074/jbc.271.14.8144

43. Marasa BS, Srikantan S, Martindale JL, et al. MicroRNA profiling in human diploid fibroblasts uncovers miR-519 role in replicative senescence. *Aging (Albany NY)*. 2010;2(6):333-343. doi:10.18632/aging.100159

44. Bakheet T, Hitti E, Al-Saif M, Moghrabi WN, Khabar KSA. The AU-rich element landscape across human transcriptome reveals a large proportion in introns and regulation by ELAVL1/HuR. *Biochim Biophys Acta - Gene Regul Mech*. 2018;1861(2):167-177. doi:10.1016/j.bbagrm.2017.12.006

45. Guo X, Wu Y, Hartley RS. MicroRNA-125a represses cell growth by targeting HuR in breast cancer. *RNA Biol*. 2009;6(5):575-583. doi:10.4161/rna.6.5.10079

46. Noh JH, Kim KM, Abdelmohsen K, et al. HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. *Genes Dev*. 2016;30(10):1224-1239. doi:10.1101/gad.276022.115

47. Zucal C, D'Agostino V, Loffredo R, et al. Targeting the Multifaceted HuR Protein, Benefits and Caveats. *Curr Drug Targets*. 2015;16(5):499-515. doi:10.2174/1389450116666150223163632

48. Izquierdo JM. Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition. *J Biol Chem*. 2008;283(27):19077-19084. doi:10.1074/jbc.M800017200

49. Heinemann U. Cold-Shock Domains—Abundance, Structure, Properties, and Nucleic-Acid Binding. Published online 2021.

50. Grishin N V. KH domain: One motif, two folds. *Nucleic Acids Res*. 2001;29(3):638-643. doi:10.1093/nar/29.3.638

51. Murina VN, Nikulin AD. RNA-binding Sm-like proteins of bacteria and archaea. Similarity and difference in structure and function. *Biochem*. 2011;76(13):1434-1449. doi:10.1134/S0006297911130050

52. Picard F, Dressaire C, Girbal L, Cocaign-Bousquet M. Examination of post-transcriptional regulations in prokaryotes by integrative biology. *Comptes Rendus - Biol*. 2009;332(11):958-973. doi:10.1016/j.crvi.2009.09.005

53. Zhao BS, Roundtree IA, He C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol*. 2016;18(1):31-42. doi:10.1038/nrm.2016.132

54. Anantharaman V, Koonin E V., Aravind L, et al. Post-transcriptional and translational controls of gene expression in eukaryotes. *Mol Cell*. 2018;9(5):1079-1126. doi:10.1016/j.molcel.2017.12.023

55. Day DA, Tuite MF. Post-transcriptional gene regulatory mechanisms in eukaryotes: An overview. *J Endocrinol*. 1998;157(3):361-371. doi:10.1677/joe.0.1570361

56. Keene JD, Lager PJ. Post-transcriptional operons and regulons co-ordinating gene expression. *Chromosome Res*. 2005;13(3):327-337. doi:10.1007/s10577-005-0848-1

57. Anantharaman V, Koonin E V., Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res*. 2002;30(7):1427-1464. doi:10.1093/nar/30.7.1427

58. Roberts E, Sethi A, Montoya J, Woese CR, Luthey-Schulten Z. Molecular signatures of ribosomal evolution. *Proc Natl Acad Sci U S A*. 2008;105(37):13953-13958. doi:10.1073/pnas.0804861105

59. Hör J, Gorski SA, Vogel J. Bacterial RNA Biology on a Genome Scale. *Mol Cell*. 2018;70(5):785-799. doi:10.1016/j.molcel.2017.12.023

60. Schaefke B, Sun W, Li YS, Fang L, Chen W. The evolution of posttranscriptional regulation. *Wiley Interdiscip Rev RNA*. 2018;9(5):1-20. doi:10.1002/wrna.1485

61. Halbeisen RE, Galgano A, Scherrer T, Gerber AP. Post-transcriptional gene regulation: From genome-wide studies to principles. *Cell Mol Life Sci*. 2008;65(5):798-813. doi:10.1007/s00018-007-7447-6

62. Parker SJ, Meyerowitz J, James JL, et al. Neurochemistry International Endogenous TDP-43 localized to stress granules can subsequently form protein aggregates. *Neurochem Int*. 2012;60(4):415-424. doi:10.1016/j.neuint.2012.01.019

63. Konopka A, Whelan DR, Jamali S, et al. Impaired NHEJ repair in amyotrophic lateral sclerosis is associated with TDP-43 mutations. 2020;4:1-28.

64. Fox GE. Origin and Evolution of the Ribosome. Published online 2010:1-18.

65. بررسی فلور و پراکنش جغرافیایی گیاهان در ارتباط با اقلیم در مرتع منطقه ایرانشهر *No Title* بارانی سنعسوح. Vol 148.

66. Adam SA, Nakagawa T, Swanson MS, Woodruff TK, Dreyfuss G. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol Cell Biol*. 1986;6(8):2932-2943. doi:10.1128/mcb.6.8.2932-2943.1986

67. Maris C, Dominguez C, Allain FHT. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J*. 2005;272(9):2118-2131. doi:10.1111/j.1742-4658.2005.04653.x

68. D'Agostino VG, Adami V, Provenzani A. A Novel High Throughput Biochemical Assay to Evaluate the HuR Protein-RNA Complex Formation. *PLoS One*. 2013;8(8):1-9. doi:10.1371/journal.pone.0072426

69. Colombrita C, Silani V, Ratti A. ELAV proteins along evolution: Back to the nucleus? *Mol Cell Neurosci*. 2013;56:447-455. doi:10.1016/j.mcn.2013.02.003

70. Maruyama K, Sato N, Ohta N. Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: Probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins. *Nucleic Acids Res*. 1999;27(9):2029-2036. doi:10.1093/nar/27.9.2029

71. Akamatsu W, Fujihara H, Mitsuhashi T, et al. The RNA-binding protein HuD regulates neuronal cell identity and maturation. *Proc Natl Acad Sci U S A*. 2005;102(12):4625-4630. doi:10.1073/pnas.0407523102

72. Katsanou V, Milatos S, Yiakouvaki A, et al. The RNA-Binding Protein Elavl1/HuR Is Essential for Placental Branching Morphogenesis and Embryonic Development. *Mol Cell Biol*. 2009;29(10):2762-2776. doi:10.1128/mcb.01393-08

73. Okano HJ, Darnell RB. A hierarchy of Hu RNA binding proteins in developing and adult neurons. *J Neurosci*. 1997;17(9):3024-3037. doi:10.1523/jneurosci.17-09-03024.1997

74. Lang M, Berry D, Passecker K, et al. HuR small-molecule inhibitor elicits differential effects in adenomatosis polyposis and colorectal carcinogenesis. *Cancer Res*. 2017;77(9):2424-2438. doi:10.1158/0008-5472.CAN-15-1726

75. Lal P, Cerofolini L, D'Agostino VG, et al. Regulation of HuR structure and function by dihydrotanshinone-I. *Nucleic Acids Res*. 2017;45(16):9514-9527. doi:10.1093/nar/gkx623

76. Tang YH, Han SP, Kassahn KS, Skarshewski A, Rothnagel JA, Smith R. Complex evolutionary relationships among four classes of modular RNA-binding splicing regulators in eukaryotes: The hnRNP, SR, ELAV-Like and CELF proteins. *J Mol Evol*. 2012;75(5-6):214-228. doi:10.1007/s00239-012-9533-0

77. Good PJ. A conserved family of elav-like genes in vertebrates. *Proc Natl Acad Sci U S A*. 1995;92(10):4557-4561. doi:10.1073/pnas.92.10.4557

78. Samson ML. Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins. *BMC Genomics*. 2008;9:1-11. doi:10.1186/1471-2164-9-392