



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA
INTERNATIONAL DOCTORATE IN STRUCTURAL BIOLOGY

CICLO XXXV

COORDINATOR Prof. Lucia Banci

**Design of new RNA binding proteins and new analytical
approaches for biopharmaceutical formulations**

Settore Scientifico Disciplinare CHIM/03

PhD student

Dott. Anna Pérez i Ràfols

Anna Pérez i Ràfols

Tutor

Prof. Marco Fragai and Dott. Tommaso Martelli

Marco Fragai Tommaso Martelli

Coordinator

Prof. Lucia Banci

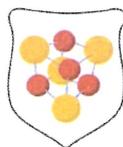
Lucia Banci

November 2019 – October 2022

***This thesis has been approved by the University of Florence,
the University of Frankfurt and the Utrecht University***



Universiteit Utrecht



This project was supported by the European Union through the project “RNAct” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITNAnalytical 2018 (contract n° 813239).

Index

1. Introduction	1
1.1 RNA BINDING PROTEINS (RBPs):.....	3
1.2 RNA RECOGNITION MOTIF (RRM):	4
1.2.1 RRMs biological and pathological roles:	6
1.3 RNAct: ENABLING PROTEINS WITH RNA RECOGNITION MOTIFS FOR SYNTHETIC BIOLOGY AND BIO-ANALYTICS:	8
1.4 MUSASHI (MSI) FAMILY:.....	9
1.4.1 Musashi-1 structure and RNA interaction:.....	9
1.4.2 Musashi-1 expression, localization, and function:	12
1.4.3 Musashi-1 targeting as a therapeutic strategy:	16
1.5 HUMAN ANTIGEN (HU-ELAVL) FAMILY:.....	16
1.5.1 HuR expression, structure, and overall function:	17
1.5.2 HuR role in diseases and therapeutic strategy:.....	18
1.6 BIOPHYSICAL APPROACHES TO STUDY PROTEIN-RNA AND PROTEIN-LIGAND INTERACTIONS:.....	18
1.7 ARTICLE 1:.....	21
1.7.1 HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models.....	21
2. Aims and topic of this Thesis	51
3. Methodological aspects	55
3.1 PROTEIN EXPRESSION AND PURIFICATION:	57
3.1.1 Site directed mutagenesis:	59
3.1.2 UV-Visible spectroscopy (UV-Vis):	59
3.1.3 Poly (ethyleneimine) (PEI) treatment:.....	59
3.1.4 Preparation of RNase free water:	60
3.1.5 AB-Elavl protein expression and purification protocol.....	60
3.1.6 Crystallization of rAB-Elavl	61
3.1.7 Human Musashi-1 (MSI-1) (1-200) RRM ₁₋₂ expression and purification protocol	61
3.1.8 Human MSI-1 RRM-1 (1-103) domain expression and purification protocol	62
3.1.9 Human MSI-1 RRM-2 (104-200) domain expression and purification protocol	62
3.1.10 Human MSI-1 mutants' expression and purification protocol	63
3.1.11 L-Asparaginase II (ANSII) expression and purification protocol.	63
3.1.12 Transthyretin (TTR) expression and purification protocol.....	64
3.2 PROTEIN CHARACTERIZATION:.....	64
3.2.1 Nuclear Magnetic Resonance (NMR) spectroscopy:	64
3.2.2 X-ray Crystallography:.....	66

3.3	PROTEIN – RNA INTERACTION:	66
3.3.1	NMR spectroscopy to study of protein -RNA interactions:	66
3.3.2	Size Exclusion Chromatography with Multi Angle Light Scattering (SEC-MALS):	67
3.3.3	Ligand Tracer for monitoring protein-RNA binding in vivo in real time:	67
3.3.4	Surface Plasmon Resonance (SPR)	68
4.	Results	69
4.1	ARTICLE 2:	71
4.1.1	Identification and Characterization of an RRM-Containing, RNA Binding Protein in <i>Acinetobacter Baumannii</i>	71
4.1.2	Supplementary Information	95
4.2	ARTICLE 3:	103
4.2.1	The interaction between Musashi-1 and RNA: a multiplayer game	103
4.2.2	Supplementary Information	143
4.3	ARTICLE 4:	153
4.3.1	Deciphering the RNA recognition by Musashi-1 to develop new artificial biomolecules for in vitro and in vivo applications	153
4.3.2	Supplementary Information	171
4.4	ARTICLE 5:	179
4.4.1	Monitoring RNA-protein interactions in living bacteria in real-time	179
4.5	ARTICLE 6:	199
4.5.1	Evaluation of the Higher Order Structure of Biotherapeutics Embedded in Hydrogels for Bioprinting and Drug Release	199
4.5.2	Supplementary Information	211
4.6	ARTICLE 7:	221
4.6.1	Transthyretin shows new possibilities for the structure-based molecular design of protein-drug conjugates	221
4.6.2	Supplementary Information	251
5.	Conclusions and perspectives	261
6.	References	265
7.	Acknowledgements	273

1. Introduction

1.1 RNA BINDING PROTEINS (RBPs):

RNA binding proteins (RBPs) are proteins capable to recognize and bind RNA targets and form ribonucleoprotein particles (RNPs). RBPs is one of the largest group of proteins in the cell; they play important roles in the regulation of gene expression and participate in every step of the RNA metabolism. In particular, RBPs are involved in transcription, splicing, modification, intracellular trafficking, translation, and decay (Figure 1).^{1,2} RBPs are highly conserved throughout evolution, widely distributed across tissues, and may act on RNA targets or with other regulatory partners that display tissue-specific expression. They can bind RNA sequences with a wide range of affinity and specificity, interacting with local sequences or structures of the RNA and can be modulated by post-transcriptional modifications, leading to the formation of dynamic and cell type specific regulatory complexes. Hence, even though RBPs can bind a wide variety of RNA targets, some may be regulated only under particular cellular conditions.²

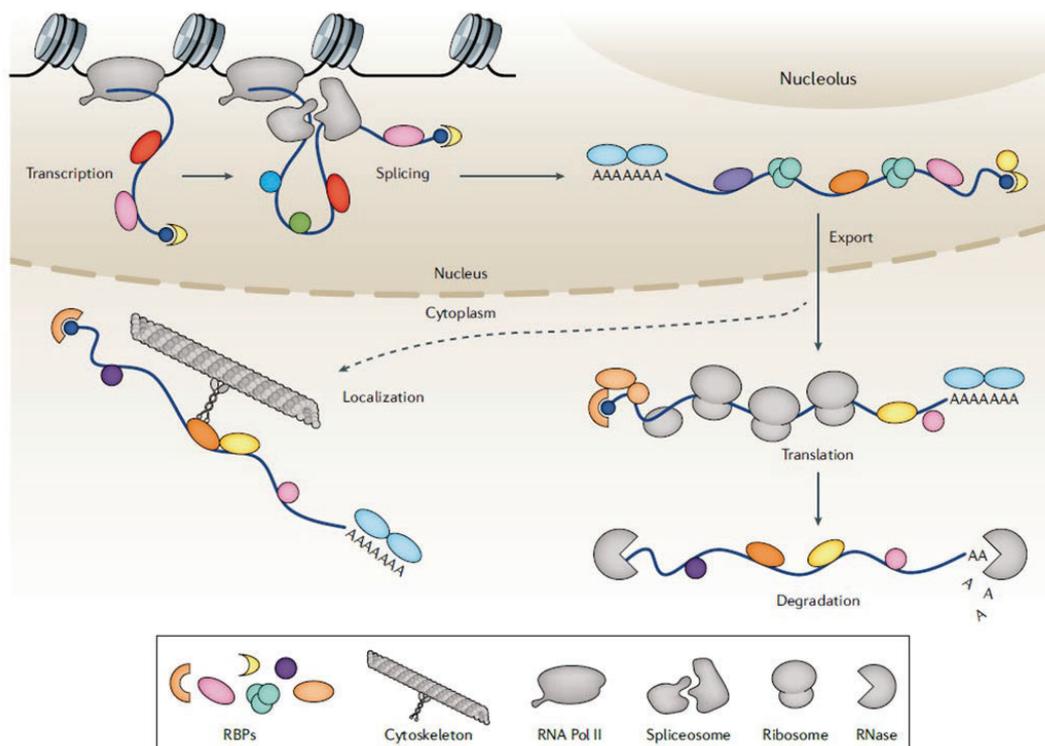


Figure 1. RNA-binding proteins involvement in several steps of RNA life both in the nucleus and cytoplasm. Nuclear functions involve transcription, splicing, capping and polyadenylation. Cytoplasmic functions involve transport (from nucleus to cytoplasm), localization, translation, and degradation of mRNA targets.²

RBPs can have one or multiple domains responsible for recognizing and binding the RNA. The presence of multi-domains can often help the coordination and enhance the binding to RNA; in fact, the combination of several RNA binding regions give to RBPs the ability to recognize with high specificity particular regions in RNAs. The most prevalent RNA binding domains (RBDs) include RNA Recognition Motif (RRM), K-homology (KH), zinc-finger domains and DEAD/DEAH helicase. Within

all of these, RRM domains are the most abundant RNA-binding domain in higher vertebrates and the most studied domain, both in terms of structure and biochemistry.^{3,4}

1.2 RNA RECOGNITION MOTIF (RRM):

RRMs are found in prokaryotes, eukaryotes, and viruses. In prokaryotes, they are normally found in proteins containing a single RRM domain and therefore are rather small. Conversely, when RRM domains are found in eukaryotes, they are often found as multiple copies within protein or together with other domains.⁵ The fact that RNA Recognition Motifs are found in such abundance through species and in particular, in higher vertebrates, highlights the importance of their biological function and their association in many functions in the cell. Although it is known that eukaryotic RRM proteins participate in all post-transcriptional events involving RNA (including pre-mRNA processing, and splicing, mRNA stability, RNA editing, pre-rRNA complex formation, translation regulation and degradation), their roles in bacteria are still unknown.⁶

A typical RRM domain comprises approximately 90 amino acids and consists in four-stranded antiparallel β -sheets packed against two alpha helices, adopting the canonical $\alpha\beta$ structure with a typical $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology (Figure 2A)^{4,7}. The loops between the secondary structure elements are often disordered when the protein is in its free form and can vary in length. Only in the loop5 there is often the formation of a secondary structure such as a small two-stranded β -sheet.⁵ The two central β -strands β_3 and β_1 of the domain (Figure 2A), contain two conserved sequences of eight and six amino-acids, that define the ribonucleoprotein domain (RNP) called RNP1 and RNP2 respectively (Figure 2B). They play important roles in nucleic acid recognition, making the beta sheet surface the binding area of the RRM domains. On top of this, both α -helices interact with each other keeping the β -sheet surface completely free and ready for RNA binding.¹ Both consensus sequences of RNP in RRM domains, contain 6-8 residues that are mainly aromatic and positively charged. The consensus sequence for RNP1 is [Arg/Lys]-Gly-[Phe/Tyr]-[Val/Ile/Leu]-X-[Phe/Tyr] and for RNP2 is [Ile/Leu/Val]-[Phe/Tyr]-[Ile/Leu/Val]-X-Asn-Leu where X can be any amino acid.^{4,5} These residues are involved in either stacking interactions with RNA bases or inserted between two sugar rings.⁸

Although they share a really conserved RNP domains, RRM-containing proteins can recognize a wide range of different RNA sequences. Indeed, many efforts have been made to shed light on this RNA recognition mechanism, and despite the identification of RNA sequences for several RRM domains, many RRM domains still have no consensus sequences identified yet. The different binding modes and differences between the RRM protein families, make the identification of a general code for RRM-RNA a current challenge.⁹ In general, each RRM can recognize between four and eight nucleotides; several

studies emphasize that the two central β -strands of the β -sheet surface are crucial for nucleic recognition, but that also the loops connecting β -strands and α -helices (Figure 2A) are involved in the interaction. Loop1 connecting β 1 and α 1, loop3 connecting β 2 and β 3 and loop 5, connecting α 2 and β 4, are in fact often involved in the formation of the main binding interface.^{4,5}

When multiple RRM domains are present in a protein, it is thought to increase its binding affinity and specificity for RNA targets; the combination of two or more RRMs in fact allow the protein to recognize a larger RNA sequence which often results in an increase of affinity. Furthermore, when two RRMs are located close to each other in tandem and they are connected by a short inter-domain linker, they can accommodate the RNA molecule and achieve a stronger interaction. When this happens, both RRMs and the interdomain linker cooperatively bind the RNA target.^{4,5,10} In other cases, RRMs do not interact with each other but rather seem to bind RNA independently, increasing the chance for the protein to encounter their specific RNA binding sequence. This is the case of PTB protein where both RRMs are separated by longer flexible linkers.¹¹ Finally, there are also cases where, independently by the length of the linker, the two domains interact in the free form of the protein. This is the case of hnRNPA1 for example where, in presence of a single strand RNA target, those two domains can affect the RNA topology inducing the formation of RNA loops between the two binding sites of each RRM.¹¹

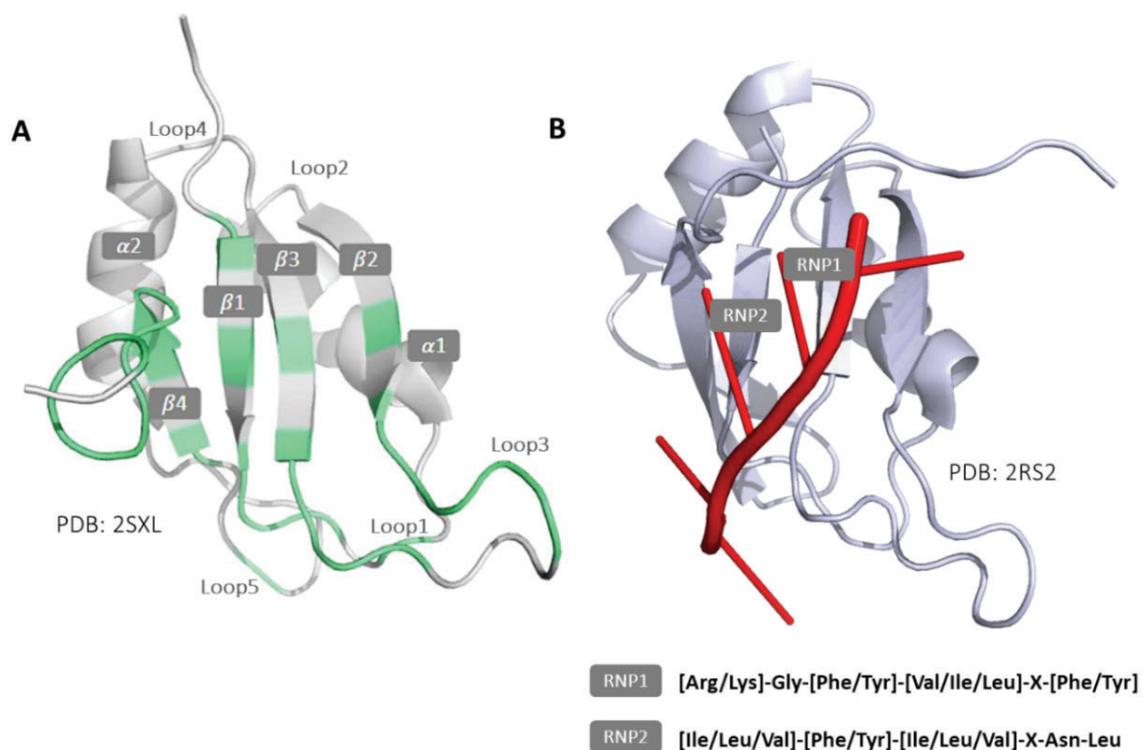


Figure 2. A) RNA Recognition Motif canonical $\alpha\beta$ structure with a typical β 1 α 1 β 2 β 3 α 2 β 4 topology. B) RRM binding interface comprehending the β -sheet surface and the conserved RNP1 and RNP2 regions located in β 3 and β 1 respectively.

On top of that, having multiple RRM proteins with similar affinity for the same RNA targets, but using a different mode of interaction, is likely to be frequent in RNA biology. This can create a competition effect which can make the molecular mechanisms of post-transcriptional gene regulation difficult to deconvolute or model.^{11,12}

Over the last few years, in addition to the known capability of multiple RRM domains to cooperate with each other to enhance the affinity for target RNA, several studies have shown that RRMs are also involved in protein-protein interaction. The RRM-protein interactions seem so important that sometimes other protein acts as a cofactor, enabling the interaction of the RRM with its target RNA.

It has also been demonstrated that other RRMs don't bind RNA at all but are however involved in RNA-related function through the interaction with other proteins; this usually happens when the beta sheet surface of the RRM is involved in the interaction with the protein.^{5,13}

From this brief glimpse on RRM interactions it is quite clear that to understand the interplay of multiple RRMs, it is important to properly characterize their relative orientation and hence, their interaction. An exhaustive characterization, that also includes the shape of the protein surface, the secondary structure of the RNA and the network of non-covalent interactions such as Van der Waals, hydrogen bonds and ionic bonds between the protein and the RNA, can be in fact the key to understand their mode of interaction and define the exact RNA binding surface.⁸

1.2.1 RRMs biological and pathological roles:

As we said before, RRMs are involved in elaborate networks of RNA-protein and protein-protein interactions that control RNA metabolism; an alteration in their binding capability or specificity can hence have an impact on many different pathways.^{14,15} For this reason, any change in their expression, any dysfunction, or any aggregation of RRMs can lead to major human disease including neurological disorders, muscular atrophies, and cancer. Moreover, aggregation of RRMs, i.e. induced by mutations, can play an important role in the context of unfolded protein diseases since RRMs can promote their aggregation or alter their function, which provides a direct mechanism for overactive Stress Granules (SGs) formation.¹⁵

Altogether, these alterations of the physiological role of RRM-containing proteins can lead to a loss of RRM function due to protein inactivation, or toxic RNA gain of function, usually observed by a dysregulation on their level of expression that tends to lead to an overexpression.^{14,16}

Therefore, altered expression, post-translational modifications, or altered localization of RRMs can contribute to tumorigenesis and diseases. Altered RNA metabolism due to a malfunction in RRM-containing proteins can affect cell growth, proliferation, invasion, and death. It is in fact common that

altered expression of these proteins has a role during the development and progression of cancer. Thus, over the past decades, RRM-containing proteins have gained more attention and a special potential for targeting them for cancer therapeutics has risen.^{17,18}

Here some examples between the numerous cases reported in literature throughout the years involving RRM-containing proteins: spinal muscular atrophy (SMA), a common autosomal recessive neuromuscular disorder, is associated with mutations or deletions in the survival of motoneuron (SMN-1) protein,^{15,16} mutations in FUS, Ataxin-2 (ATXN2), heterogeneous nuclear RNA proteins (hnRNPs) and TDP-43 all cause neurodegenerative diseases including Amyotrophic Lateral Sclerosis (ALS);^{15,19,20} altered expression of Musashi (MSI), RNA binding protein 38 (RBM38), or TET family, proteins that participate in cell growth control, are involved in several cancers;^{21,22} and HuR protein, of the Hu/ELAV-like family, has been implicated in promoting inflammation and inflammatory diseases.¹⁰

During these past years, various strategies have been used to target RRM-containing proteins in cancer and other human diseases. The most common approach to target them is via small molecules.¹⁸ These might be used to inhibit the protein function in different ways including disruption of the RNA-protein interaction by binding the RRM-containing protein^{23,24}, induce RRM protein degradation²⁵ or prevent functional RRM modifications by targeting the activity of the enzymes involved²⁶.

On the other hand, other approaches that have been gaining attention over the past years¹⁸ involve the use of i) antisense oligonucleotides (ASOs), designed to bind to the RNA to perturb protein production by enhancing degradation of the targeted RNA and therefore alter RNA metabolism or up-regulate gene expression²⁷; ii) target delivery of small interfering RNA (siRNAs) nanoparticles for cancer treatment applied for example in HuR protein²⁸, iii) nucleic acid or peptide based aptamers that fold into sequence specific three-dimensional structure that can recognize their unique target²⁹; iv) peptides for targeted drug delivery³⁰ or v) circular RNAs (circRNAs) with several RRM recognition sites to regulate their function by acting as a decoy or sponge³¹.

However, to this date, many RRM-containing proteins still remain considered as “undruggable” and only a few other have been targeted for therapy. This is due to the many difficulties and challenges that these proteins entail. Some of these, include the fact that the complex biological function of most RRM proteins in diseases still needs to be understood, the unknown RNA target or the specific mode of interaction of the RRM-containing protein, the selective targeting of cancer cells as most of the RRM proteins are required for both normal and cancer cells, or the identification of critical targets of an RRM that can mediate its function in cancers.

To overcome these many difficulties one might encounter when trying to target these proteins, research in many areas and focus on many RRM-containing proteins is needed. Understanding the structural bases of the interaction with RNA, discover and understanding the biological function of the protein in

cells, detecting the implication that each of these proteins have in diseases, finding selective targets, and developing or improving technology or targeted therapies are just a few of the many paths that research needs to focus on to continue making progress in the therapeutical field of RRM-containing proteins. In this regard, the European project “RNAct: Enabling proteins with RNA Recognition Motifs for synthetic biology and bio-analytics” has emerged as one of the multi-disciplinary and networking projects that aims to study and improve the general knowledge in RRM-RNA interaction to try to overcome some of these challenges.

1.3 RNAct: ENABLING PROTEINS WITH RNA RECOGNITION MOTIFS FOR SYNTHETIC BIOLOGY AND BIO-ANALYTICS:

RNAct is Marie Skłodowska-Curie Innovative Training Network (ITN) project that focuses its research on proteins containing RNA Recognition Motifs (RRM) to post-transcriptionally regulate gene expression and detect specific RNAs. The ability to manipulate these RRM proteins offers a wide application potential in synthetic biology, in bio-analytics and for the creation of synthetic pathways in cells, such as enabling their activation on triggering their inhibition via small molecules.

As already mentioned in *section 1.2*, RRMs are extremely versatile in their RNA recognition capability and are also able to modify RNA conformation, enabling its function or recognition by other proteins. Because of the complexity of RRM-RNA binding and the intrinsic dynamics present in both RRMs and RNA targets, we lack understanding of how the isolated domains recognize and bind specific RNA fragments and how multiple RRMs act to achieve complex functions in multidomain RRM proteins.

RNAct research is organized throughout a design cycle. This cycle starts with the computational analysis of sequences and structures of both proteins and RNA, in order to identify the key amino acid positions that drive the interaction. With this analysis achieved, a computational approach is used to try to manipulate the interaction between protein and RNA changing the primary structure of both sequences. The cycle continues with an analysis at atomic level using integrated structural biology approaches and the final step is the synthetic biology study of the post-transcriptionally regulation of fatty acid processing via RRMs, and within the bio-analytics field, the determination of RNA *in-cell* and the design of RNA-based chips.

These design cycle steps address three general aims: i) modify the RNA specificity of single-domain RRMs by modulating their interaction, ii) allosterically control RRM-RNA binding via a small ligand and iii) design multi-domain RRM protein switches where RNA binding changes the linker conformation or where allosteric changes in the domain linker change RNA specificity.

This PhD thesis falls within the RNAct consortium and dives mainly in the first two general aims of the project. In particular, it focuses on the expression, purification, and characterization of single- and multi-

domain RRM containing proteins using biophysical and structural biology techniques in order to determine how those proteins bind to RNA targets. Furthermore, it is also studied the development of a new application to monitor time-resolved RNA-RRM interaction analysis on living bacteria cells and its modification in presence of an allosteric inhibitor.

From all the RRM-containing proteins that still need to be studied, this thesis focuses on two really relevant RNA binding protein families that are implicated in several diseases and have a high interest within the research community from a therapeutic point of view. This thesis focuses its research in two RRM-containing families, the Musashi (MSI) family, with regards to the Musashi-1, and the Elavl-like family, with the analysis of the bacterial form of the Human Antigen R (HuR) protein.

1.4 MUSASHI (MSI) FAMILY:

The Musashi (MSI) family is a widespread and highly conserved protein family, first identified in *Drosophila* due to its role in the regulation of the sensory organ precursor cells division. This protein family comprises two Musashi homolog proteins: Musashi-1 (MSI-1) and Musashi-2 (MSI-2), that function as translational regulators of mRNA targets and play critical roles in stem cell maintenance and self-renewal. MSI-1 and MSI-2 share 69% of sequence identity and contain two RRM domains that recognize and bind target mRNA.^{18,32}

1.4.1 Musashi-1 structure and RNA interaction:

Musashi-1 is a multi-domain RRM protein that consist of 362 amino acid residues. Its structure, Figure 4A, contains two highly conserved RRMs, RRM-1 (comprising residues Cys20-Val94) and RRM-2 (residues Lys109-Ser191) in its N-terminal region connected by a short flexible interdomain linker. The tandem domain is followed by a highly disordered region in its C-terminal portion, that serves as a putative nuclear export signal and mediates protein-protein interactions. As it can be seen in Figure 4B and 4C, residues Arg61-Phe68 and Arg150-Phe157, and residues Met22-Leu27 and Ala111-Leu116 correspond to the RNP1 and RNP2 sequences of the two domains, respectively.

Both RRMs contain the classical conformation of four beta sheets packed against two alpha helices adopting the $\beta 1\alpha 1 \beta 2\beta 3\alpha 2\beta 4$ topology. In RRM-1 (Figure 4B), in addition to the canonical structure, two short antiparallel beta sheets have been identified in loop5 between $\alpha 2$ and $\beta 4$ (Glu84-Leu85 and Lys88-Thr89, respectively), and a β -bulge structure, common in most RRM-containing proteins, has been found in the $\beta 2$ region comprising the residues Val46-Glu48. Furthermore, residues Phe23, Phe63 and Phe65 form a hydrophobic patch exposed to the solvent.^{32,33} In RRM-2 (Figure 4C), in addition to the main structure, there is a short beta-strand located between the residues Phe171-Ile174, and, a β -

bulge, has been identified in the Val135-Asp137 region, again in the $\beta 2$. A hydrophobic patch is present also in this second RRM, and involves residues Phe112, Phe152 and Phe154.^{33–35}

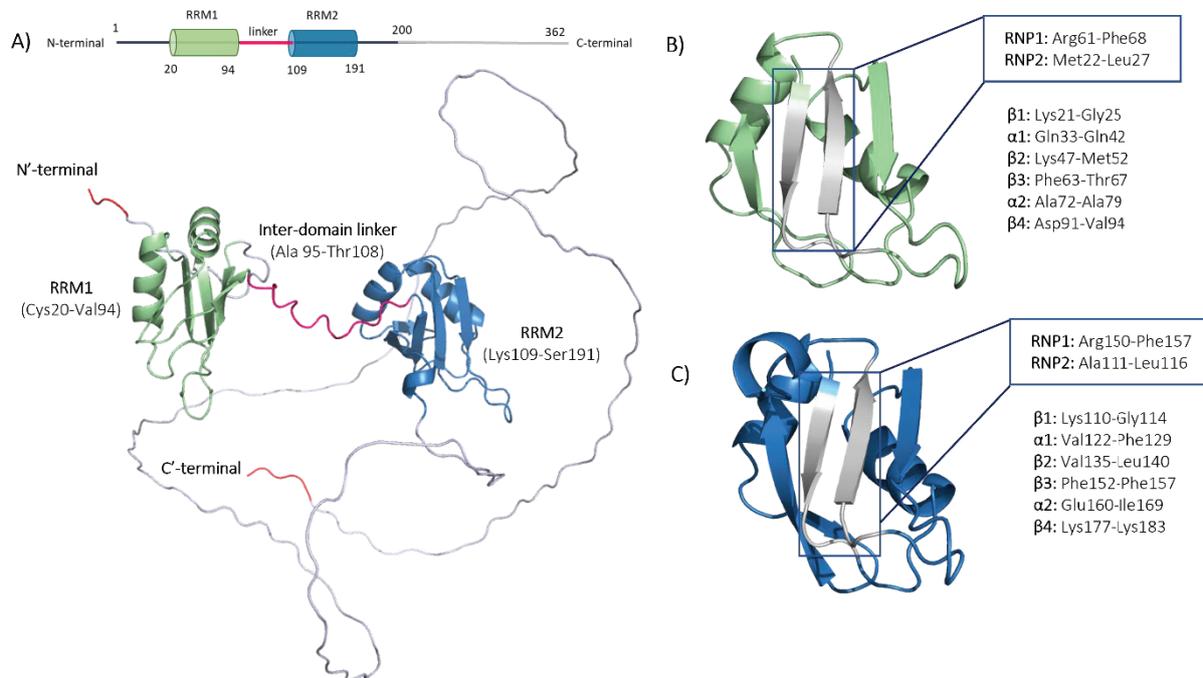


Figure 4. Structural scheme of human Musashi-1 protein. Structures from Alpha fold: AF-O43347-F1. A) General overview of MSI-1 full-length structure. RRM-1 (in green) and RRM-2 (in blue) are located in tandem in the N'-terminal region and connected by an inter-domain linker (pink). N'- and C'-terminal regions are highlighted in red. B) Structural information of RRM-1. Highlighted in grey the residues comprising RNP1 and RNP2 regions. C) Structural information of RRM-2. Highlighted in grey the residues comprising RNP1 and RNP2 regions.

Due to its role in the neural precursor cells and its implication in many diseases, experiments throughout the years with the mouse homolog of the protein (m-MSI-1, 98.9% of sequence identity with the human MSI-1) have been performed as an attempt to deconvolute the interaction mode with the target RNA. First binding experiments were performed by Northwestern blot assays using RNA homopolymers.^{36,37} Those experiments detected a strong binding of the m-MSI-1 full length protein with poly(G) and a significantly weaker binding to poly(U). When analysing the isolated domains, RRM-1 showed similar behaviour as the full-length protein, while no binding was detected with RRM-2.

In 1997 gel-retardation experiments with a proposed specific sequence (r(GUUAGUUAGUUAGUU)), also indicated that RRM-1 alone was able to bind the RNA oligomer while again no binding was detected for RRM-2.³⁵ Further experiments with solution NMR were carried out to determine the three-dimensional structure of RRM-2 (Lys109-Ser191) in its free form.

In 2001, an in vitro selection of high-affinity (SELEX) RNA ligands for m-MSI-1, allowed the identification of the consensus RNA sequence for Musashi-1, and demonstrated putative translational

repression of the mammalian numb (m-numb) gene.³⁸ In this SELEX study, the authors identified sequence-specific uridine-rich RNAs fragments in mRNAs that bind MSI-1, and all the selected fragments were containing the (G/A)U₁₋₃AGU sequence motif that frequently appeared in two or three tandem repeats.

Using the SELEX motif as a starting point, Miyanoiri's group³³ performed RNA binding experiments with solution NMR with the isolated domains and the same RNA sequence used in gel-retardation experiments mentioned above. Contrary to gel retardation experiments, where binding to RRM-2 was not detected, NMR experiments were able to successfully detect the binding of RRM-1 and RRM-2 with the RNA target.

Solution NMR³² and Fluorescence Polarization³⁹ experiments, using different length RNA fragments and the isolated domains of MSI-1, were performed to find the optimal Musashi binding determinant and define the core motif for either recognition motif. The minimal consensus sequences identified for the binding of RRM-1 and RRM-2 were r(GUAG) and r(UAG), respectively. These consensus motifs are suggested to mainly be located in single stranded or partially bulged hairpins and within the 3' untranslated region (3'UTR) region or mRNA targets.^{40,41}

Subsequently, structure of the RRM-1³² and of the RRM-2³⁴ bound to r(GUAGU) (referred from this point as G1-U2-A3-G4-U5) were determined by solution NMR.

Among the residues in the β -sheet surface of both RRMs involved in the RNA binding, RRM-1 residues Lys21, Phe23, Arg61, Phe63, Phe65, Asp91 and Lys93 are conserved in the second recognition motif (RRM-2), being Lys110, Phe112, Arg150, Phe152, Phe154, Glu180 and Lys182, respectively. These residues share function and mode of interaction upon the binding (Figure 4, panel A and B).

In RRM-1 (Figure 4, panel C), the Trp29 located in the loop between β 1 α 1 recognizes G1 of the RNA sequence throughout aromatic-aromatic stacking, and a hydrogen bond is formed between the same G1 and the Lys88. In the RRM-2 the Trp29 is not maintained and the Val118 is present instead; since the second domain cannot undergo aromatic-aromatic stacking interaction and therefore is not able to bind G1.^{32,34}

To date, no proper RNA binding analysis has been performed on the tandem domain and no structural biology experiments has been done on the human homologue protein. Furthermore, no deep kinetic analysis has been done, and values of K_d obtained throughout the years came from fluorescence polarization or densitometric analysis of gel retardation experiments.

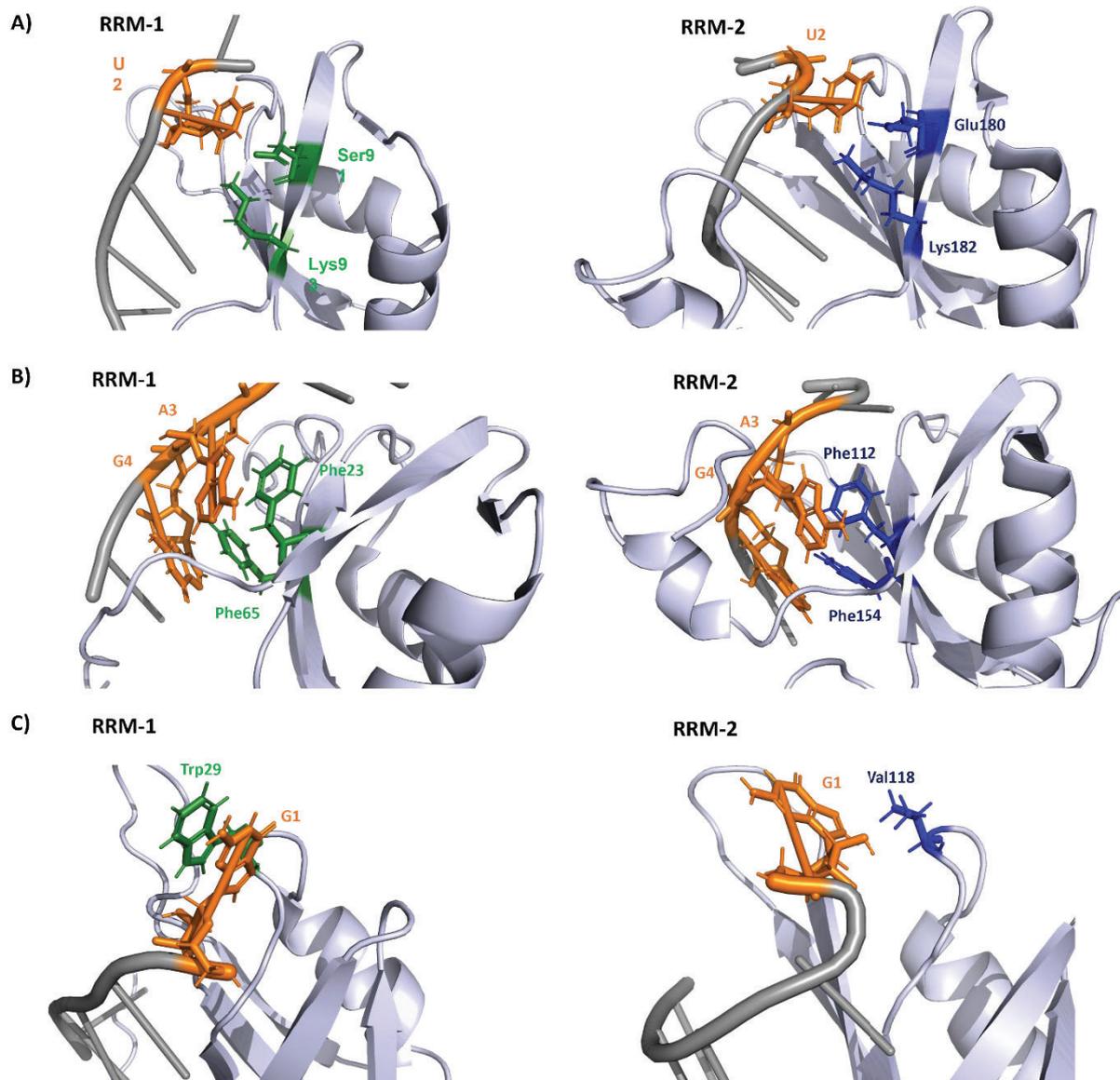


Figure 4. Structure of the residues involved in the RNA binding of RRM-1 (PDB:2rs2) and of the RRM-2 (PDB: 5x3z) bound to r(GUAGU) (referred as G1-U2-A3-G4-U5). A) Asp91, Lys93 and Glu180, Lys182 formation of hydrogen bonds with U2. B) phenylalanine residues located in the RNP2 (Phe23 and Phe112) and RNP1 (Phe65 and Phe154) interaction through aromatic-aromatic stacking interactions with A3 and G4 respectively and salt bridges are present between G4 and Lys21 and Lys110 to support the stacking interactions. C) In the RRM-1 the Trp29, recognizes G1 of the RNA sequence, throughout aromatic-aromatic stacking whereas in RRM-2 Trp29 is not maintained and the Val118 is present instead. Val118 cannot undergo aromatic-aromatic stacking interaction and therefore is not able to bind G1. ^{32,34}

1.4.2 Musashi-1 expression, localization, and function:

Musashi-1 is considered a specific RBP, because it interacts with small subset of mRNAs, and has a restrictive pattern of expression such as being expressed only in certain tissues or during certain specific developmental stages.⁴⁰ To be precise, MSI-1 protein is selectively expressed in neural precursors cells,

including NSCs, and in central neuron system (CNS) of progenitor cells during the embryonic stage. It plays important role in brain development, maintenance of stem-cell state and cell differentiation.⁴²

Neural RNA binding proteins in mammalian, play important roles in post-transcriptional regulation of gene expression and neural development.³³ Indeed, Musashi expression is gradually down regulated during neural differentiation; it associates with ribosome and is involved in maintaining the undifferentiated state of neural stem cells through the posttranscriptional control of downstream genes.^{38,43}

Once Musashi is expressed and activated in its specific tissues and developmental stages, it is expected to regulate hundreds of different mRNA and thus, to be involved in regulation of cell cycle, cell proliferation, cell differentiation, apoptosis, and post-translational modifications.⁴⁰ In physiological conditions, MSI-1 expression in mammary epithelial cells activates Notch and Wnt pathways and promotes proliferation of mammary stem cells,⁴³ any dysregulation of the expression or activation of the protein might lead to a disruption of these signalling pathways. The Wnt signalling pathway plays important role in the maintenance of pluripotency and in the process of somatic cell reprogramming. p21^{WAF-1} mRNA encodes for a cyclin-dependent kinase inhibitor, which negatively regulates Wnt gene expression. Musashi binds to the 3'-UTR region of the p21^{WAF-1} mRNA and represses its translation resulting in an upregulation of the cell cycle via the Wnt signalling, maintaining cell proliferation and multipotency (Figure 5).³²

Separately, numb mRNA encodes for the Numb protein that promotes the differentiation of neural stem cells (NSCs) via the suppression of the Notch signalling pathway (Figure 5).^{32,40} MSI-1 has been shown to upregulate numb expression in gastric tissue.³⁹ NSCs are undifferentiated tissue stem cells present in the CNS that possess multipotency, ability to self-renew and to repair damaged tissue. MSI-1 has been found to bind the 3'-UTR region of Numb mRNA and inhibit its translation. On the other hand, Numb protein binds the intracellular domain of Notch-1 protein and inhibits activation of the Notch signalling pathway. Hence, MSI-1 positively regulates Notch signalling to maintain stem-cell NSCs self-renewal.⁴² Thus, Musashi controls cell proliferation of cancer stem cells by modulating Notch and Wnt signalling pathways, and promotes G0/G1 to S phase transition of cell proliferation by inhibiting translation of checkpoint regulators through the direct binding to p21, p27 and p53 mRNAs.⁴³

An additional in vitro affinity selection study of E12 embryonic mouse brain cells was performed to identify mRNA and proteins that bind specifically to MSI-1 with the aim to understand and comprehend its role in NSCs.^{32,44} Among the mRNA identified as Musashi binders, this study focused on the doublecortin (dxc) mRNA, a protein related to the migration of new-born neurons and neural development (Figure 5). They selected it because of its structural and functional features: i) the mRNA sequence includes the MSI-1 consensus sequence for both RRMs, ii) these motifs are located near or

within an hairpin, and iii) are located on the 3'-UTR region of the mRNA; iv) doublecortin (dxc) is expressed only in neural precursors from NSCs, v) exclusive and co-expression of both Musashi and dxc was observed in human brain, and vi) knock-out experiments on MSI genes reduced the number of neurospheres isolated from embryonic mouse brains while on dxc prevented their migration and maintained their structure. In this study they confirmed binding both in vitro and in cultured cells, as well as demonstrated binding of Musashi to dxc mRNA and a subsequently inhibition of its translation throughout reporter assay systems.

The C-terminal region of Musashi-1 is involved in protein-protein interaction which allows other proteins to both induce and repress protein translation (Figure 5). For instance, Musashi-1 is co-expressed complementarily with HuR protein localized in differentiated neurons in the central nervous system. HuR recognizes and directly binds the 3'-UTR region of Musashi and positively regulates MSI-1 expression in glioblastoma.^{33,37} In this regard, in the mammalian system, MSI-1 was identified as a marker for neuronal stem cells (NSCs) together with other RBPs from the ELAV family.⁴¹ Separately, the interaction with a poly-A-binding protein (PABP), with a specific domain located after RRM-2 of MSI-1 disrupts binding between PABP and the eukaryotic translation factor 4 gamma (eIF4G) and thus, represses initiation of translation of a subset of Musashi targets. In addition, in mammalian cells, Musashi binds 3'-UTR region of the numb mRNA and represses m-numb 5'-cap dependent translational by competing with eIF4G for PAB binding, thereby maintaining the stem cell state.^{32,40,45}

In the C-terminal region of MSI-1, a Lin28-binding motif has been identified; hence a cooperation within the two proteins can control the subcellular localization of proteins involved in microRNA (miRNA) biogenesis and regulate the post-transcriptional miRNA biogenesis in the nucleus^{32,41}, influencing stem cell maintenance and differentiation. In line with miRNA, MSI-1 was demonstrated to impair CD44 mRNA by limiting miRNA-mediated decay.

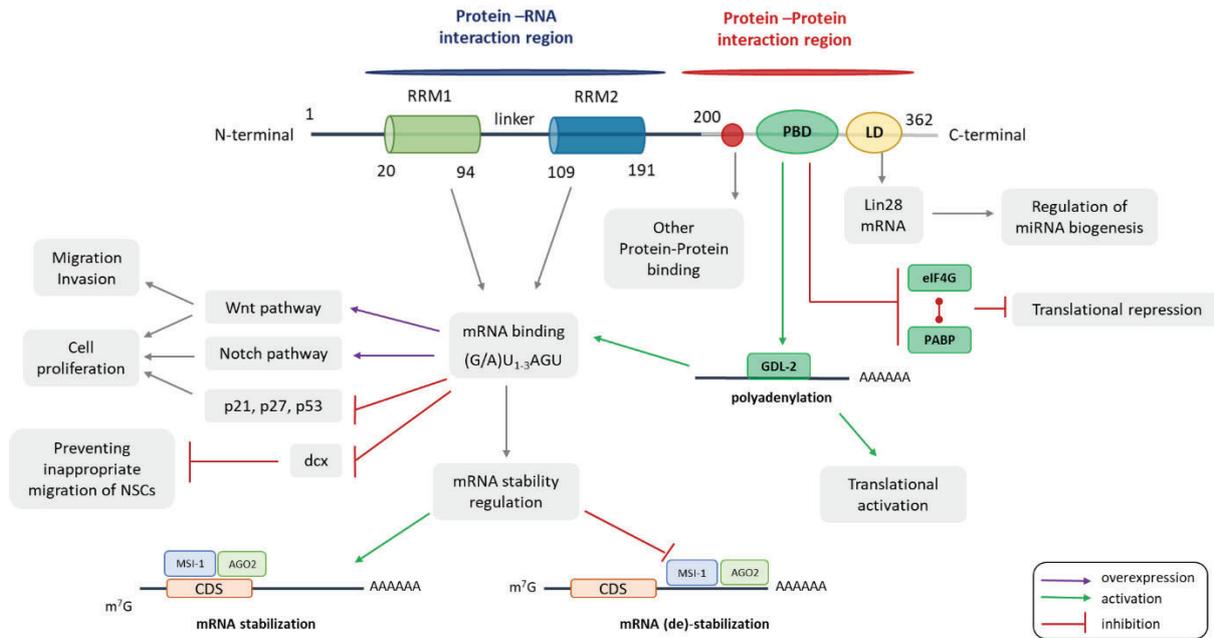


Figure 5. General scheme of Musashi role and function in cells.

MSI-1 is localized in both the nucleus and cytoplasm, and shuttling was reported under stressed conditions such as hypoxia or platinum-based therapy (chemodrugs). In response to the stress condition, MSI-1 translocates to the cytosol where it recruits Argonaute 2 (AGO2) protein, another RBP that play a central role in RNA silencing processes by inhibiting the translation of their targets and mediating their decay (Figure 5). Once AGO2 is recruited by Musashi, it post-transcriptionally regulates the expression of target mRNA. Their interaction MSI-1/AGO2 to mRNA targets within the 3'-UTR region enhances their degradation whereas binding to CDS region (coding sequence) prevents their rapid degradation. By coordinating these mechanisms under stress conditions, MSI-1/AGO2 complex enhances tumour proliferation and ensures cancer cell survival.^{41,46} In addition, MSI is also recruited by stress granules (SGs) like many RBPs and affects SG formation. However, whether SG-recruitment is linked to its roles in controlling cytoplasmatic mRNA remains unknown.⁴¹

In 2001 Musashi was first linked to cancer where studies showed overexpression of MSI-1 in medulloblastomas, gliomas and hepatomas^{41,45}. Over the years, elevated expression⁴⁵ has been also identified in many types of solid tumours (like pancreatic, lung, colon, brain, or breast cancer). In addition, overexpression has also been found in different types so leukaemia⁴⁷. In line with this, overexpression of MSI-1 has been correlated with poor prognosis in several cancers, i.e., in breast cancer⁴⁸. Because of the strong oncofetal expression of Musashi, its association with a poor prognosis and its pro-oncogenic properties suggest MSI-1 as a potent marker and promising therapeutic target in various solid cancers. It is considered a potent regulator of various cancer hallmark pathways such as proliferation, apoptosis and self-renewal, migration, and tumour growth *in vivo*.^{18,41,45}

1.4.3 Musashi-1 targeting as a therapeutic strategy:

The importance of this protein in various cancers motivated researchers to develop small molecules inhibitors of these proteins. However, as many other RBPs, MSI-1 was considered hard to target by small molecules due to the lack of well-defined binding pockets and structural information. Moreover, the strong electrostatic binding between the negatively charged RNA and the positively charged protein, makes it even harder to obtain small molecules inhibitors. However, attempts have been made and to date some small molecules (oleic acid and other fatty acids, (-)-gossipol and luteolin), have been identified and tested.⁴¹

Oleic acid was proven to bind the first recognition motif of Musashi and thus, disrupt the RNA binding. It was also shown to inhibit proliferation of a cell line expressing Musashi, although the mechanism of action remains unestablished.²³

Luteolin disrupts MSI-1 binding to the consensus sequence via a direct interaction with the protein, and diminishes Musashi's positive impact on the pro-oncogenic target genes' expression, resulting in reduced proliferation, migration, and invasion of glioblastoma cells.⁴¹

(-)-gossipol, a natural product extracted from cottonseed, inhibits RNA binding via a direct interaction with RRM-1. It has been shown that it also reduces Notch/Wnt signalling pathways in colon cancer cell lines, and suppress tumour growth in a mouse xenograft model. However, on the other hand, it has shown only minimal activity against prostate cancer.⁴⁵

Since the situation described above, extensive research to find proper inhibitors of Musashi is needed. These inhibitors may be valuable in cancer therapy or as regulators of responses to treatments like chemotherapy or target therapy.

Also, MSI-1 is induced as part of the damage response processes in arthritis, overexpressed in atherosclerotic plaques and critical for the functionality and survival of photoreceptors in the retina.⁴⁵ Therefore, more work is needed to define Musashi's role in other tissues with the aim to lighten pathologic conditions.

1.5 HUMAN ANTIGEN (HU-ELAVL) FAMILY:

The Hu family comprises four mammalian neuronal RBPs that codify for the ELAV gene (embryonic lethal abnormal visual system), a gene that encodes a nuclear product in neurons and that is required for the maintenance and differentiation of postmitotic neurons. All members of the Hu or ELAV family are essentially expressed within the nervous system as in neurons, neuroblasts and glial cells, and in less frequency in other tissue.^{49,50}

The ELAV-like family proteins are well conserved throughout species and the four human Hu proteins (HuR alternatively called HuA, HuB, HuC and HuD) themselves, share >71% of sequence identity. Within the four proteins, HuR is the only one ubiquitously expressed, whereas the expression of the remaining

three are limited to neuronal tissues.^{49,51} These RRM-containing proteins are both located in the nucleus and cytoplasm and thus perform various roles at different stages of gene expression including splicing, nucleo-cytoplasmic transport, translation, and degradation of mRNA.⁵²

The Hu family is known to bind mRNAs by targeting uridine- or adenine/uridine-rich (AU-rich) elements (ARE) in the 3'-UTR region, and play important roles in the regulation of these mRNA stability and translational efficiency. AREs are specific regulatory sequences that determine whether mRNA decay is delayed or facilitated and can be grouped in three categories. Class I AREs contain several copies of the AUUUA motif dispersed within U-rich regions, Class II AREs contain two or more overlapping UUAUUUA(U/A)(U/A) monomers and Class III AREs are U-rich regions that do not contain the AUUUA motif.^{52,53}

1.5.1 HuR expression, structure, and overall function:

The HuR protein is an ubiquitously expressed protein mainly localized in the nucleus, but can shuttle to the cytoplasm under proliferative and stress conditions including heat shock, UV radiation or nutrient and energy depletion. In the nucleus it is engaged in the regulation of pre-mRNA splicing and polyadenylation, while, in the cytoplasm, it is involved in the stabilization of various mRNAs ARE-containing encoding for cell-cycle regulators, growth factors, tumour suppressors, proto-oncogenes, apoptosis and various inflammatory enzymes. HuR can also increase or repress the translation of mRNA targets.⁵²⁻⁵⁴

HuR is composed by three RNA Recognition Motifs (RRMs) that contribute to the RNA binding. RRM-1 and RRM-2 are the most conserved across the Hu family, and are located in tandem at the N-terminal region after an initial intrinsically disordered region (about 19 residues) and are connected via a short flexible linker. These two domains are known to provide the main surface of interaction for ARE binding throughout a cooperative assembly. The third RRM domain is located at the C-terminal regions and separated from the others by a much longer linker (about 60 residues) that includes the HuR Nucleocytoplasmic Shuttling Sequence (HNS), which is responsible for the nuclear/cytoplasmic shuttling. RRM-3 interacts with the poly(A) tails of mRNA and preferably binds U-rich stretches rather than AUUA motifs. In addition, both RRM-3 and the RRM2-RRM3 linker have been shown to also contribute to the stabilization of the HuR-AREs complexes.^{51,54,55}

Both RRM-1 and RRM-3 domains have been demonstrated to self-associate and, therefore, dimerize in solution. For RRM-1, the dimerization is established by the α -helices of the domain but without engaging the β 2- β 3 loop.⁵⁵ For RRM-3, several conformations with different orientations have been

proposed, however, all agree that the Trp261 residue is critical for the stabilization of the dimer conformation.^{54,56-58}

Studies with RNA binding and small inhibitors like dihydritanshione-I (DHTS), have revealed an equilibrium between an open-close conformation of the RRM1-RRM2 tandem domain in the absence of mRNA while adopting a close globular conformation upon the binding. The conformational change induced subsequent contacts throughout the inter-domain linker and RRM-2 that increased the binding affinity.^{24,53}

A detailed information on the expression, localization, structure, function, and RNA binding can be found in Article 1, section 1.7.

1.5.2 HuR role in diseases and therapeutic strategy:

As many HuR targets encode for proteins that play critical roles in the establishment and development of tumours, high or abnormal expression of HuR, and its nuclear-cytoplasmic localization, have been found in several types of cancers, including lung, colon, breast, and ovarian tumours. Hence, HuR protein is considered as a potential therapeutic target and a prognostic marker of cancer. HuR has been proven to modulate cancer traits including enhanced cell proliferation, increased angiogenesis, anti-apoptotic resistance, reduced immunosurveillance, invasion and metastasis. Moreover, HuR has been implicated in promoting inflammation and inflammatory diseases and has been shown to be involved in other pathologies like cardiovascular, neurological, and muscular disorders.^{55,59}

Given its functions, HuR has been proposed as a useful therapeutic target. Approaches to inhibit HuR function can be beneficial to inhibit or reduce tumorigenesis or useful for treating chronic inflammatory diseases.⁵⁹ A detailed information about the implication of HuR in diseases and information on therapeutical approaches can be found in Article 1, section 1.7.

1.6 BIOPHYSICAL APPROACHES TO STUDY PROTEIN-RNA AND PROTEIN-LIGAND INTERACTIONS:

Biophysical approaches to study protein-ligand interactions aim to identify the molecular components of the complex and quantify their equilibrium populations and kinetics (association and dissociation rates).⁶⁰ Any signal that affects the amount of complex formed can be used to monitor the binding and determine the thermodynamic and kinetic characteristics of the interaction. The choice of the methodology to investigate a particular protein system depends on many factors, including specificity

of the ligand for the protein, quantities of protein and ligand available and presence or absence of the complex formation.⁶⁰

Protein complex formation can be monitored both *in vitro* and *in-cell*. *In vitro*, particular buffer and salt solutions are used to maintain each component stable and soluble whereas *in-cell* proteins are found in physiological conditions in the cellular cytoplasm or blood plasma. A lot of different techniques exist to study, in *in vitro* or *in-cell*, protein-ligands interaction and all together they give complementary information one to the other leading to a broader and more complete understanding of the system.

These experimental strategies⁶⁰ aim to detect the formation of complex via monitoring a change in the ligand or protein properties, or in their structure, through detection and quantification of the population of the free components and detect the location or biological implications resulting upon the complex formation.

Starting from *in vitro* studies, many experimental strategies rely on changes in spectral properties of either the protein or the ligand, for example, UV or visible light absorption, circular dichroism (CD) or fluorescence detection. Other techniques such as nuclear magnetic resonance (NMR) and X-ray crystallography, are used instead for the study of structural features of complexes and are widely used for screening applications.⁶⁰

A wide range of methods that have also been developed rely on the immobilization of the complex components and the detection of their changes upon the binding. Examples of these methods are Surface Plasmon Resonance (SPR) and SwitchSENSE[®] technology.

Continuing with *in-cell* studies, some of the wide range of approaches that can be used to study protein-ligand interactions are focused on monitoring the interaction in real-time and estimate the affinity and kinetic constants within living cells like the LigandTracer technology.⁶¹ Other techniques like confocal microscopy, instead, are broadly used to resolve the localization or the structure of specific objects within the cell.

To understand the mechanism of interaction between a protein and its ligand, the usual approach is the study of multiple ligands (structurally related to each other) and the study of the interactions of mutants of the protein or the ligand.⁶⁰ Focusing on any changes that appear upon small changes in the structure, it is possible to retrieve important information that could lead to a structural explanation of the binding affinity or specificity of a particular ligand, or any change on the cellular function of the protein.

Using all the techniques mentioned above, among others, the complex formation can be studied using different approaches. Experiments such as direct binding assays, competition experiments, binding

kinetics, or enzyme activity, for example, are the most used. Thus, combining techniques both *in vitro* or *in-cell* and employing different kind of experiments, it is possible to obtain a complete understanding of the system collecting information on the complex formation from different and complementary points of view. Exploiting these techniques and developing new methodologies helps improve the study of protein-ligand interactions of all kinds.

In this PhD thesis several of these techniques were exploited for the study of proteins-ligand interaction with a major focus on RNA binding proteins and their interaction with RNA targets.

1.7 ARTICLE 1:

1.7.1 HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models

Giulia Assoni ^{a,c,1}, Valeria La Pietra ^{b,1}, Rosangela Digilio ^{c,1}, Caterina Ciani ^c, Nausicaa Valentina Licata ^c, Mariachiara Micaelli ^c, Elisa Facen ^c, Weronika Tomaszewska ^c, Linda Cerofolini ^d, Anna Pérez-Ràfols ^e, Marta Varela Rey ^f, Marco Fragai ^g, Ashwin Woodhoo ^{f,h,i,j,k,2}, Luciana Marinelli ^{b,2}, Daniela Arosio ^{l,2}, Isabelle Bonomo ^{c,2}, Alessandro Provenzani ^{c,2,*}, Pierfausto Seneci ^{a,2,*}

^b Department of Pharmacy, University of Napoli Federico II, Via D. Montesano 49, 80131 Napoli, Italy

^c Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, via Sommarive 9, 38123 Trento, Italy

^d Magnetic Resonance Center (CERM), University of Florence and Interuniversity Consortium for Magnetic Resonance of Metalloproteins (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy

^e Giotto Biotech S.R.L., Via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy

^f Gene Regulatory Control in Disease Group, Center for Research in Molecular Medicine and Chronic Diseases (CIMUS), Health Research Institute of Santiago de Compostela (IDIS), University of Santiago de Compostela, 15706 Santiago de Compostela, Spain

^g Magnetic Resonance Center (CERM), University of Florence and Interuniversity Consortium for Magnetic Resonance of Metalloproteins (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy

^h Department of Functional Biology, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain

ⁱ Galician Agency of Innovation (GAIN), Xunta de Galicia, Santiago de Compostela, Spain

^j Center for Cooperative Research in Biosciences (CIC bioGUNE, Basque Research and Technology Alliance (BRTA), Bizkaia Technology Park, Building 801A, 48160 Derio, Spain

^k IKERBASQUE, Basque Foundation for Science, Bilbao 48013, Spain

^l Istituto di Scienze e Tecnologie Chimiche “G. Natta” (SCITEC), National Research Council (CNR), Via C. Golgi 19, I-20133 Milan, Italy

Advanced Drug Delivery Reviews 181 (2022) 114088

DOI: <https://doi.org/10.1016/j.addr.2021.114088>



Contents lists available at ScienceDirect

Advanced Drug Delivery Reviews

journal homepage: www.elsevier.com/locate/adr

HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models



Giulia Assoni^{a,c,1}, Valeria La Pietra^{b,1}, Rosangela Digilio^{c,1}, Caterina Ciani^c, Nausicaa Valentina Licata^c, Mariachiara Micaelli^c, Elisa Facen^c, Weronika Tomaszewska^c, Linda Cerofolini^d, Anna Pérez-Ràfols^e, Marta Varela Rey^f, Marco Fragai^g, Ashwin Woodhoo^{f,h,i,j,k,2}, Luciana Marinelli^{b,2}, Daniela Arosio^{l,2}, Isabelle Bonomo^{c,2}, Alessandro Provenzani^{c,2,*}, Pierfausto Seneci^{a,2,*}

^a Chemistry Department, University of Milan, Via Golgi 19, I-20133 Milan, Italy

^b Department of Pharmacy, University of Napoli Federico II, Via D. Montesano 49, 80131 Napoli, Italy

^c Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, via Sommarive 9, 38123 Trento, Italy

^d Magnetic Resonance Center (CERM), University of Florence and Interuniversity Consortium for Magnetic Resonance of Metalloproteins (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy

^e Giotto Biotech S.R.L., Via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy

^f Gene Regulatory Control in Disease Group, Center for Research in Molecular Medicine and Chronic Diseases (CIMUS), Health Research Institute of Santiago de Compostela (IDIS), University of Santiago de Compostela, 15706 Santiago de Compostela, Spain

^g Magnetic Resonance Center (CERM), University of Florence and Interuniversity Consortium for Magnetic Resonance of Metalloproteins (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy

^h Department of Functional Biology, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain

ⁱ Galician Agency of Innovation (GAIN), Xunta de Galicia, Santiago de Compostela, Spain

^j Center for Cooperative Research in Biosciences (CIC bioGUNE, Basque Research and Technology Alliance (BRTA), Bizkaia Technology Park, Building 801A, 48160 Derio, Spain

^k IKERBASQUE, Basque Foundation for Science, Bilbao 48013, Spain

^l Istituto di Scienze e Tecnologie Chimiche "G. Natta" (SCITEC), National Research Council (CNR), Via C. Golgi 19, I-20133 Milan, Italy

ARTICLE INFO

Article history:

Received 6 April 2021

Revised 7 October 2021

Accepted 16 December 2021

Available online 20 December 2021

Keywords:

HuR modulators

Hallmarks of cancer

HuR role in immunity

Chemical features of HuR modulators

Structural interactions of HuR modulators

ABSTRACT

The Human antigen R (HuR) protein is an RNA-binding protein, ubiquitously expressed in human tissues, that orchestrates target RNA maturation and processing both in the nucleus and in the cytoplasm. A survey of known modulators of the RNA-HuR interactions is followed by a description of its structure and molecular mechanism of action – RRM domains, interactions with RNA, dimerization, binding modes with naturally occurring and synthetic HuR inhibitors. Then, the review focuses on HuR as a validated molecular target in oncology and briefly describes its role in inflammation. Namely, we show ample evidence for the involvement of HuR in the hallmarks and enabling characteristics of cancer, reporting findings from *in vitro* and *in vivo* studies; and we provide abundant experimental proofs of a beneficial role for the inhibition of HuR-mRNA interactions through silencing (CRISPR, siRNA) or pharmacological inhibition (small molecule HuR inhibitors).

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

1.1. ELAVL family conservation through evolution

Human antigen R (ELAVL1, HuR or HuA) belongs to the ELAVL (Embryonic Lethal Abnormal Vision-like) protein family [1] that

includes also HuB (or Hel-N1), HuC and HuD [2]. A characteristic feature of these proteins is the tripartite architecture of three functionally distinct RNA Recognition Motif (RRM) domains, which individually contribute to mRNA binding [3,4]. Elav-like orthologous genes are present in all metazoans with a high identity score (>45%), but the number of paralogous genes in the different species varies from one to four, without a clear relationship between the complexity level of the organism and its size, development, and brain structure [5]. These observations suggest that the diversification of ELAVL members could have occurred before or at the first stages of metazoan evolution [5–7]. In mammals, the four ELAV-like proteins show a 70–85% identity. The most conserved

* Corresponding authors.

E-mail addresses: alessandro.provenzani@unitn.it (A. Provenzani), pierfausto.seneci@unimi.it (P. Seneci).¹ First authors.² Last authors.

sequences are contained in the RRM domains and include a conserved octapeptide termed ribonucleoprotein motif (RNP-1) and a conserved hexapeptide termed RNP-2, both responsible for the binding with nucleotides of their RNA targets. The four human ELAVL paralogous genes have different roles and different cellular localizations. HuB, HuC and HuD are mainly neural (HuC and HuD are strictly neural-specific, HuB targets are neural-specific RNA species although several transcripts are also expressed in ovary and testis), and are localized mainly in the cytoplasm, although they can translocate into the nucleus. HuR, conversely, is ubiquitously expressed in all human tissues [8–10], is localized mainly in the nuclear compartment, but shuttles to the cytoplasm under certain stimuli. The roles of neuronal ELAVL (nELAVL) proteins are overlapping, undergo strict tissue- and cell-specific modulation, and their expression is time regulated throughout development [5,8,9,11]. HuR knock-out (KO) leads to embryonic lethality in mice, due to extra-embryonic defects of placenta, showing abnormalities in skeleton and spleen development in the surviving embryos. These findings suggest that HuR is involved in regulating the fate of mRNAs encoding proteins involved in key processes, such as organ development and tissue homeostasis, and highlights its relevance for the entire organism [12,13].

1.2. HuR expression, localization and molecular functions

In the nucleus, HuR binds target pre-mRNA introns, promoting splicing and alternative polyadenylation events [14–17]. Upon intrinsic (such as DNA damage) [18] or extrinsic stimuli (such as UV irradiation), HuR acts as a shuttle, exporting associated mature mRNAs to the cytoplasm, where it mainly stabilizes and promotes the translation of such mRNAs. In so doing, HuR regulates the fate of thousands of coding and noncoding RNAs containing AU/UU-rich elements (AREs) sequences primarily located in their 3' untranslated regions (UTR) (Fig. 1).

The expression of HuR is finely regulated at multiple levels. HuR expression depends on the transcription factor Nuclear Factor

kappa-light-chain-enhancer of activated B cells (NF- κ B) [19] and on Smad family proteins [20], but the mechanism of transcriptional regulation of HuR still remains unclear. It was shown that HuR binds and stabilizes its own mRNA [21], and that *HuR* mRNA shows alternative polyadenylation variants to protect itself from degradation, decay and nucleus-cytoplasmic export promotion [22,23]. HuR translation can be negatively controlled by several microRNAs such as miR-519 and miR-125a, as observed in human MCF7, WI-38 fibroblast and Hela cells [24,25]. Furthermore, HuR activity is regulated by multiple post-translational modifications (PTMs), such as phosphorylation and methylation. In particular, HuR can be phosphorylated by the serine/threonine-kinase Chk2, the mitogen-activated protein kinase (MAPK), p38, the protein kinase C alpha (PKC- α) [26], the protein kinase C delta (PKC- δ), and the I κ B kinase α (IKK α) [18,27–30], can be methylated by coactivator associated methyltransferase (CARM1) [31] and neddylated by murine double minute 2 (MDM2) [32]. Moreover, HuR activity can be regulated by PARylation through the poly(ADP-ribose) polymerase 1 (PARP-1) [33,34] and by sulfhydration through the cystathionine δ -lyase (CSE) [35]. PTMs on HuR can occur at its RRMs, at the hinge region, and at the HuR nucleocytoplasmic shuttling domain (HNS), affecting either its binding ability and its localization [36]. Furthermore, HuR can be degraded via the proteasome after ubiquitination [37] or cleaved by caspases during apoptosis [38].

As HuR is involved in the regulation of post-transcriptional gene expression for thousands of targets (approximately 7% of the human protein-coding gene transcripts), it is not surprising that an altered expression or localization of HuR leads to the emergence of multiple pathologic phenotypes.

1.3. HuR in healthy tissues and human pathologies

HuR is crucial in promoting a proper differentiation of different cellular lineages, including spermatocytes, myocytes, and adipocytes [39–42]. In testes, HuR is essential for spermatogenesis by

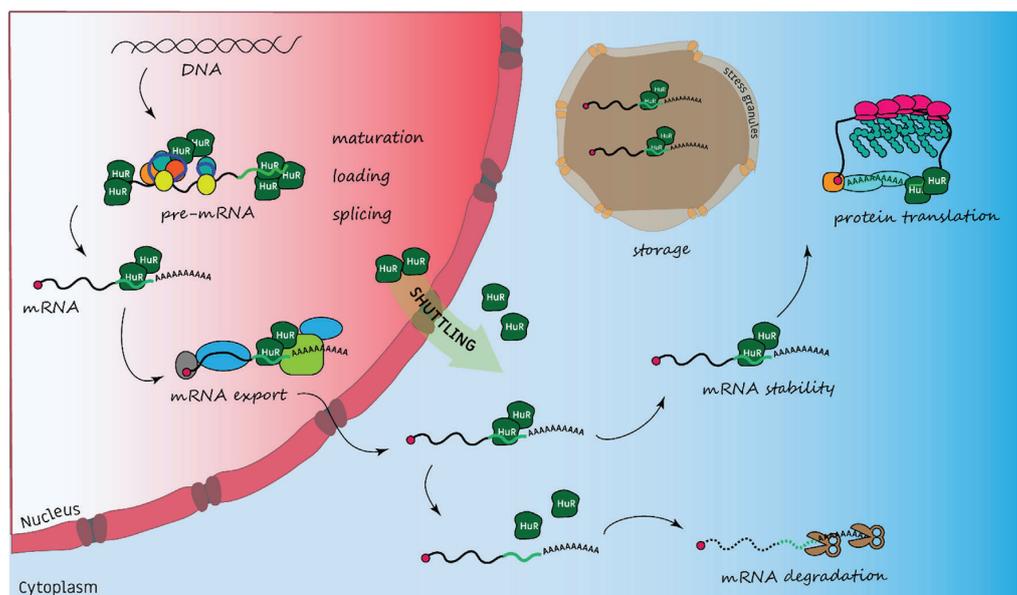


Fig. 1. Schematic representation of HuR functions within the cell. Inside the nucleus HuR (green) binds pre-mRNA introns (light blue thin lines) and untranslated regions (light green lines), promoting nuclear processing and mRNA maturation events. HuR cooperates with splicing factors (round colored dots), guiding splicing and alternative splicing events, and favors mRNA export to the cytoplasm by interacting with transportation factors (colored shapes). In the cytoplasm HuR promotes mRNAs stability, helps its storage (as in stress granules), and modulates target translation.

specifically binding the heat shock protein A2 (HspA2), and while its ablation causes sterility, HuR overexpression causes the failure of spermatid differentiation [43]. In adipose tissue, HuR positively regulates the expression of adipose triglyceride lipase (ATGL), thus promoting lipolysis, while HuR ablation increases the risk of obesity [44]. Furthermore, in muscle, HuR plays a critical role in myogenesis by positively regulating myogenic factors such as *MyoD*, *myogenin*, and *p21* [39], although during muscle wasting, HuR is involved in the repression of myogenic differentiation [45].

In vivo studies targeting HuR functions clearly demonstrated its key functions in development, differentiation and control of tissue homeostasis, and its importance in the pathogenesis of various disorders. HuR is essential for mouse embryogenesis, as genetic ablation of HuR in the germ line using Deleter-Cre lines [12] or Hrpt-Cre lines [46] both lead to embryonic lethality. Global ablation of HuR using tamoxifen-inducible Cre recombinase-mediated gene excision in adult mice also resulted in lethality within 10 days, showing its non-redundant function in postnatal life. In these mice, a severe depletion in progenitor cell populations in hematopoietic and intestinal systems was observed [46].

Subsequent studies over the years, targeting HuR and using cell-type-specific Cre lines in various cells ranging from immune to neuronal cells, from hepatocytes to adipose cells have further shown the exquisite role of HuR in regulating many aspects of cellular differentiation and functions in various cell types (Table 1).

These are just some examples of the physiological role of HuR in regulating tissues homeostasis. Indeed, physiological functions of HuR were also reported in other organs. Namely, de-regulation of HuR was shown in cardiovascular [76,77] and retinal diseases [78,79], nephropathies [80] and neurological disorders [81].

In vivo models have been pivotal to better understand the role of HuR in diseases. In some cases, HuR has a protective function during disease initiation and progression, since its ablation leads to disease aggravation, as for intestinal epithelial cells [59,60,62]. In other circumstances, HuR knockdown leads to beneficial effects, most notably in experimental models of multiple sclerosis [63,64]. Such a complex functional profile for HuR is further demonstrated by surprisingly opposite effects in the same cell type, depending on the insult. In cardiomyocytes, for example, HuR ablation aggravates the effect of isopentol-induced myocardial hypertrophy and cardiac fibrosis [54]; conversely, in a transverse aortic constriction (TAC) model of pressure overload-induced hypertrophy HuR deletion reduces left ventricular hypertrophy, dilation, and fibrosis while preserving the cardiac function [55]. Importantly, being one of the major regulators of gene expression, a de-regulation of HuR has been also associated in the development of a variety of cancers. Indeed, the majority of its mRNA targets encodes for extremely relevant proteins in oncogenesis and tumor progression, such as p21 [82], c-FOS [83], the vascular endothelial growth factor (VEGF) [84], SIRT1 [85], tumor necrosis factor alpha (TNF- α) [86], B-cell lymphoma 2 (Bcl-2) [87], cyclooxygenase 2 (COX-2) [88] and p53 [89]. An aberrant overexpression of HuR and a disturbance of its nuclear/cytoplasmic localization have been associated with malignant transformations [90] in a significant number of human cancers, including breast [91], colon [92], ovarian [93,94], prostate [95], pancreatic [96] and oral cancer [97] among others.

1.4. HuR nucleus to cytoplasm transport and association with diseases

The functions of HuR in the nucleus are still to be fully elucidated, only recently being targeted in multiple studies. Conversely, the role of HuR in the cytoplasm and the mechanisms driving its shuttling have been clarified. Indeed, several PTMs regulate HuR subcellular localization, influencing its interactions with several transporters, and thus impacting on its regulatory feedback [36,98–103]. A correct and balanced subcellular localization for

HuR, either in the nucleus or in the cytoplasm, is pivotal for normal development and organism homeostasis. As an example, a proper cytoplasm to nucleus redistribution influences the 3' UTR driven alternative splicing of different mRNAs, such as focal adhesion kinase (FAK) guiding adult neurodevelopment in mice neural stem cells [104]. As to bone marrow derived macrophages (BMDMs), HuR retention in the nucleus caused by its interaction with macrophage associated atherosclerosis lncRNA sequences (*MAARS*) increases macrophage apoptosis, leading to a decrease in their efferocytosis and, consequently, exacerbating the pathophysiology of atherosclerosis [105]. When liberated from the *MAARS* sponging effect and after translocation to the cytoplasm, HuR can then appropriately regulate mRNAs (e.g. *p53*, *BCL-2*) that drive beneficial apoptotic processes. Furthermore, nuclear HuR regulates the alternative splicing of 3' UTRs in different oncogenic mRNAs (e.g. *CENPN*), so that treating MCF7 cells with doxorubicin (DOXO) prevents HuR nuclear localization, impairing cell cycle progression and thus cancer development [16]. Therefore, abnormal HuR subcellular localization and accumulation correlates with multiple diseases. Many studies on histological and clinical samples demonstrated a positive correlation between cytoplasmic HuR accumulation and *VEGFA*, *VEGFC*, *COX2* and *IL-8* levels in human tumor samples [88,106–109]. HuR is also correlated with a high-density of blood microvessels, and its presence in the cytoplasm is associated with large different tumors [53]. HuR cytoplasmic accumulation is mostly associated with cancer onset and progression, along with worse prognosis and poor outcomes in renal, urothelial and esophageal carcinomas, and small-lung cancer [90,110–114]. It also correlates with an overexpression of oncogenes and pro-tumorigenic factors [115,116]. Conversely, HuR nuclear accumulation is reported as a causative factor in the onset of gallbladder carcinoma [117].

1.5. HuR targeting as a therapeutic strategy

Considering the earlier described pathogenic functions of HuR, its inhibition *via* either small-interfering RNAs or small molecules has emerged as a putative therapeutic approach to ameliorate the outcome of multiple diseases. Indeed, many reports focused on the identification and characterization of HuR targeting agents, each exhibiting their advantages and disadvantages have been published. The modulation of either the expression, the translocation and the PTMs profile of HuR, and its silencing were found to be effective in a context-dependent manner, and rely on the activation of specific factors [111,118]. Conversely, inhibitors (mostly small molecules) impairing the HuR-mRNA interaction are less context-dependent, although their potency may depend on the PTM profile of HuR [33,119]. Their use could lead to unpredictable side effects due to the ubiquitous expression of HuR, its pleiotropic and controversial functions, and the lethal phenotype connected with its complete ablation. Indeed, the *in vivo* efficacy and tolerability of these HuR-mRNA modulators must still be completely determined [12,13]. Moreover, HuR conditional KO animal models (Table 1) suggest that its presence is essential for tissue homeostasis (i.e. hepatocytes, hematopoietic stem cells and epithelia), therefore a strong inhibition of HuR function by small molecules may cause potentially severe side effects. On the other side, pharmacological modulation of HuR only partially recapitulates a complete HuR KO condition and may be less impactful. In addition, considering the structural similarity of HuR with other ELAVL-like proteins, specificity remains a largely unknown issue. The aim of this work is to critically review the HuR modulators discovered so far, focusing mainly on naturally occurring and synthetic small molecules, describing their physico-chemical and structural properties, and commenting on their bioavailability. Their *in vitro* and *in vivo* anti-tumorigenic activities is recapitulated in association with

Table 1
Cellular studies showing cell-specific HuR functions, major observations and genes regulated by HuR.

Targeted cell type	Cre line used	Major observations	Regulated genes	Refs
Ubiquitous	<i>Morpholinos (Zebrafish)</i>	Knockdown of Elavl1a using specific morpholinos results in a striking loss of primitive embryonic erythropoiesis.	<i>Gata1</i>	[47]
All tissues (global)	<i>Hrpt-Cre; Rosa26Cre/ERT2</i>	Postnatal global deletion of HuR1 induces atrophy of hematopoietic organs, extensive loss of intestinal villi, obstructive enterocolitis, and lethality within 10 days.	<i>Mdm2</i>	[46]
Germ line, epiblast cells; endothelial cells	<i>Deleter-Cre; Sox-Cre; Tie 1-Cre</i>	Targeted HuR ablation leads to defects in placental labyrinth morphogenesis, skeletal specification patterns, and splenic ontogeny.	<i>Fgf10, Tbx4, Ets2, Hoxd13, Hoxb9</i>	[12]
Ubiquitous	<i>HuR-transgenic</i>	Transgenic overexpression of HuR, prevents the production of fully functional gametes.	–	[40]
Reproductive tissue	<i>Sycp1-Cre; Vav-Cre; Nestin-Cre; Vasa-Cre; HuR-transgenic</i>	Targeted deletion of HuR specifically in germ cells leads to male but not female sterility. Mutant males are azoospermic because of the extensive death of spermatocytes at meiotic divisions and failure of spermatid elongation. The latter defect is also observed upon HuR overexpression.	<i>Hspa2</i>	[48]
Epiblast cells; epithelial compartment of the lung endoderm	<i>Sox-Cre; Sftpc-Cre</i>	The loss of HuR blocks lung branching morphogenesis in the mouse by controlling mesenchymal but not epithelial responses during branching.	<i>Fgf10, Tbx4,</i>	[49]
Hepatocytes	<i>Albumin-cre</i>	HuR knockout in hepatocytes reduces liver lipid transport and ATP synthesis, and aggravates high fat diet (HFD)-induced NAFLD.	<i>Apob, Uqcrb, Ndufb6</i>	[50]
Adipose tissue; brown adipose tissue	<i>Adipoq-Cre; UCP1-Cre</i>	Fat-specific knockout of HuR significantly enhances adipogenic gene program in adipose tissues, accompanied by a systemic glucose intolerance and insulin resistance. HuR knockout also results in depot-specific phenotypes: it can repress myogenesis program in brown fat, enhance inflammation program in epididymal white fat and induce browning program in inguinal white fat.	<i>Insig1</i>	[51]
Adipose tissue	<i>Adipoq-Cre</i>	Mice lacking HuR in adipose tissue show obesity when induced with a high-fat diet, along with insulin resistance, glucose intolerance, hypercholesterolemia and increased inflammation in adipose tissue.	<i>adipose triglyceride lipase (ATGL)</i>	[44]
Vascular smooth muscle cells	<i>alpha-smooth muscle actin (α-SMA)-Cre</i>	Mice lacking expression of HuR in vascular smooth muscle cell show hypertension and cardiac hypertrophy.	<i>RGS (regulator of G-protein signaling) protein(s) RGS2, RGS4, and RGS5</i>	[52]
Endothelial cells	<i>VE-cadherin-Cre</i>	Endothelial-specific HuR knockout mice exhibit reduced revascularization after hind limb ischemia and tumor angiogenesis in oncogene-induced mammary cancer, resulting in attenuated blood flow and tumor growth, respectively.	<i>Ej4enif1</i>	[53]
Cardiomyocytes	<i>α-MHC-Cre</i>	Deletion of HuR in cardiomyocytes aggravates the effect of isopentol-induced myocardial hypertrophy and cardiac fibrosis.	<i>PLB; β1-AR</i>	[54]
Cardiomyocytes	<i>αMHC-mER-Cre-mER</i>	HuR deletion reduces left ventricular hypertrophy, dilation and fibrosis while preserving cardiac function in a transverse aortic constriction (TAC) model of pressure overload-induced hypertrophy	<i>Tgfb</i>	[55]
Muscle cells	<i>MyoD-Cre</i>	Muscle-specific HuR knockout mice have high exercise endurance that is associated with enhanced oxygen consumption and carbon dioxide production. These mice exhibit a significant increase in the proportion of oxidative type I fibers in several skeletal muscles.	<i>KSRP</i>	[56]
Muscle cells	<i>MyoD-Cre</i>	Genetic ablation of HuR in muscle cells protects mice from tumor-induced muscle loss (cachexia)	<i>STAT3</i>	[57]
Skeletal muscle	<i>Myf1-Cre</i>	Male, but not female, mice lacking HuR in skeletal muscle exhibit metabolic inflexibility, with mild obesity, impaired glucose tolerance, impaired fat oxidation and decreased <i>in vitro</i> palmitate oxidation.	–	[58]
Intestinal - epithelial cells	<i>Villin-Cre</i>	Mice lacking intestinal expression of HuR display reduced cell proliferation in the small intestine and increased sensitivity to DOXO-induced acute intestinal injury. Intestinal deletion of HuR also decreases tumor burden in genetic and pharmacological models.	–	[59]
Intestinal epithelial cells	<i>Villin-Cre</i>	Intestinal epithelium-specific HuR knockout inhibits repair of damaged mucosa induced by mesenteric ischemia/reperfusion in the small intestine and by dextran sulfate sodium in the colon.	<i>cdc42</i>	[60]
Intestinal epithelial cells	<i>Villin-Cre</i>	Intestinal tissues from intestinal epithelium-specific HuR knockout mice have reduced numbers of Paneth cells, and Paneth cells have fewer lysozyme granules per cell.	<i>Cnpy3</i>	[61]
Intestinal epithelial cells	<i>Villin-Cre</i>	Targeted deletion of HuR in intestinal epithelial cells disrupts mucosal regeneration and delays repair after injury.	<i>Nucleophosmin (NPM)</i>	[62]
TH17 cells	<i>OX40-Cre</i>	Knockout of HuR reduces the number of pathogenic IL-17 + IFN- γ + CD4 + T cells in the spleen during experimental autoimmune encephalomyelitis, impairs splenic Th17 cell migration to the CNS and abolishes the disease.	<i>Irf4, Runx1, il12rb1</i>	[63]
TH17 cells	<i>OX40-Cre</i>	Targeted deletion of HuR in Th17 cells delays initiation and reduces disease severity in the onset of experimental autoimmune encephalomyelitis.	<i>IL-17</i>	[64]
CD4 + T cell	<i>distal lck-Cre</i>	Conditional HuR knockout in CD4 + T cells results in loss of IL-2 homeostasis and defects in JAK-STAT signaling, Th2 differentiation, and cytokine production.	<i>Il2ra</i>	[65]
B lineage cells	<i>Mb1-Cre</i>	Mice lacking HuR have reduced numbers of immature bone marrow and mature splenic B cells.	–	[66]

Table 1 (continued)

Targeted cell type	Cre line used	Major observations	Regulated genes	Refs
B lineage cells	<i>Mb1-Cre</i>	In the absence of HuR, defective mitochondrial metabolism resulted in large amounts of reactive oxygen species and B cell death, showing that HuR controls the balance of energy metabolism required for the proliferation and differentiation of B cells.	<i>Dlst</i>	[67]
Myeloid cells	<i>LysM-Cre (Mice): Morpholinbos (zebrafish)</i>	Tumor growth, angiogenesis, vascular sprouting, branching, and permeability are significantly attenuated in HuR-knockout mice, suggesting that HuR-regulated myeloid-derived factors modulate tumor angiogenesis. Zebrafish embryos injected with an <i>elav1</i> morpholino show angiogenesis defects in the subintestinal vein plexus.	<i>Vegf-a</i>	[68]
Myeloid cells	<i>LysM-Cre (Mice): HuR-transgenic</i>	Mice lacking HuR in myeloid-lineage cells, which include many cell types of the innate immune system, display enhanced sensitivity to endotoxemia, rapid progression of chemical-induced colitis, and severe susceptibility to colitis-associated cancer. Conversely, myeloid overexpression of HuR reduces inflammatory profiles, and protects mice from colitis and cancer.	<i>Tnf, Il10, Ccl2, Ccl7</i>	[69]
Myeloid cells	<i>HuR-transgenic</i>	HuR overexpression in murine innate compartments suppresses inflammatory responses <i>in vivo</i> .	<i>Tnf, Cox2</i>	[70]
Myeloid cells	<i>LysM-Cre</i>	Macrophage-specific HuR knockout mice display a markedly diminished microvascular angiogenic response to an inflammatory stimulus, and blood flow recovery and ischemic muscle neovascularization after femoral artery ligation.	<i>Vegf, MMP-9</i>	[40]
Thymocytes	<i>Lck-Cre</i>	Mice with genetic deletion of HuR in thymocytes possess enlarged thymi but display a substantial loss of peripheral T cells.	<i>Tnf, Dr3</i>	[71]
Microglia/macrophages	<i>Cx3cr1 -Cre</i>	Targeted deletion of HuR in microglia/macrophages reduces tumor growth and proliferation associated with prolonged survival in a murine model of glioblastoma.	-	[72]
Excitatory neurons; pyramidal neurons of the cerebral cortex and hippocampus	<i>AAV-CaMKIIamCherry-Cre; Nex-Cre</i>	HuR is a critical modulator of stress-induced synaptic plasticity. In adult mice, AAV-Cre-mediated knockout of HuR prevents anxiety-like and depression-like behaviors induced by chronic stress, and is required for stress-induced dendritic spine loss and synaptic transmission deficits. Genetic inactivation of HuR during embryonic development leads to enhanced synaptic functions.	<i>RhoA</i>	[73]
Radial glia neural progenitors; neuroepithelial cells	<i>Emx1-Cre ; Foxg1-Cre</i>	Deletion of HuR before embryonic day 10 disrupts both neocortical lamination and formation of the main neocortical commissure, the corpus callosum.		[74]
Neurons	<i>Thy1Cre-ERT2-EYFP</i>	Inducible, neuron-specific HuR-deficient mice develop a phenotype consisting of poor balance, decreased movement, and decreased strength.	<i>Immediate Early Response 2 (IER2)</i>	[75]

pivotal hallmarks of Cancer and enabling characteristics [120,121], providing evidence for a modulating role of HuR targeting agents on all of them. In addition, a summary guide on the molecular and cellular tools used so far for the identification and characterization of authentic HuR-mRNA modulators is presented (Fig. 2).

2. HuR modulators from various sources: Chemical classes, structural and mechanistic features

This Section focuses on small molecule HuR inhibitors, divided in three classes depending on their source. *Naturally occurring compounds* (Paragraph 2.1, Table 2 entries 1n-13n) were the first to be reported in literature; they include a number of heavily oxygenated chemotypes with varying potency on HuR and limited bioavailability. *Synthetic compounds* (Paragraph 2.3, Table 3 entries 1s-17s) resulted either from HTS campaigns on large compound collections, or from rational drug design efforts using computational tools. Finally, *nature-inspired synthetic compounds* (Paragraph 2.2, Table 4 entries 1ns-7ns) were discovered either by testing semi-synthetic derivatives of biologically active natural compounds, or by rationally designing and synthesizing analogues inspired by known, naturally occurring HuR inhibitors.

We decided to include in our review both direct/HuR-binding inhibitors, and indirect modulators, acting to reduce HuR functional activity. Conversely, we only included a brief Paragraph 2.4 to HuR inhibitors other than small molecules (e.g., siRNA or miRNA sequences, antisense nucleotides, nano-objects containing or decorated with HuR-interacting moieties). Some recent reviews [122,123] can be accessed to cover in details these areas.

2.1. Natural compounds

A first study, aimed at demonstrating the druggability of HuR with small organic molecules, was initially reported [124]. A HTS campaign, taking advantage of a confocal fluctuation spectroscopy homogeneous assay format, was run on ~50,000 microbial, mycological and plant extracts tested on a shortened variant of recombinant HuR (HuR₁₂). RP-HPLC-fractionation of 13 most active extracts led to the identification of chrysantone-like MS-444, dehydromutactin and okicenone as HuR-RNA binding inhibitors (Table 2, respectively entries 1n-3n). Namely, cell-free inhibition constants in the low-medium micromolar range were determined for the three hits; HuR-dependent cellular effects were confirmed for MS-444 and dehydromutactin. Through RNA electrophoretic mobility assay (EMSA) screening of 179 chemicals representing a subset from a Korean compound library, quercetin (Table 2, entry 4n), a flavonoid extracted from more than 20 plant varieties, was identified together with the synthetic compounds b-40 and b-41 (Table 4, entries 1s and 2s respectively) as inhibitors of the binding of HuR to the ARE domain of TNF- α mRNA [125]. A 1.4 μ M IC₅₀ was subsequently determined by measuring their effect on HuR-mRNA formation in a filter binding assay; a good specificity vs. other RNA-binding proteins was also observed.

Additional, naturally occurring flavonoids and coumarins were reported as HuR-mRNA interaction modulators. Clorobioicin/C11 [126] (Table 2, entry 5n) – a natural aminocoumarin extracted from *Streptomyces roseochromogenes*, known to inhibit DNA gyrase [127] – was identified together with semi-synthetic daunomycin 3-oxime/C10 (Table 3, entry 2ns) through a high throughput

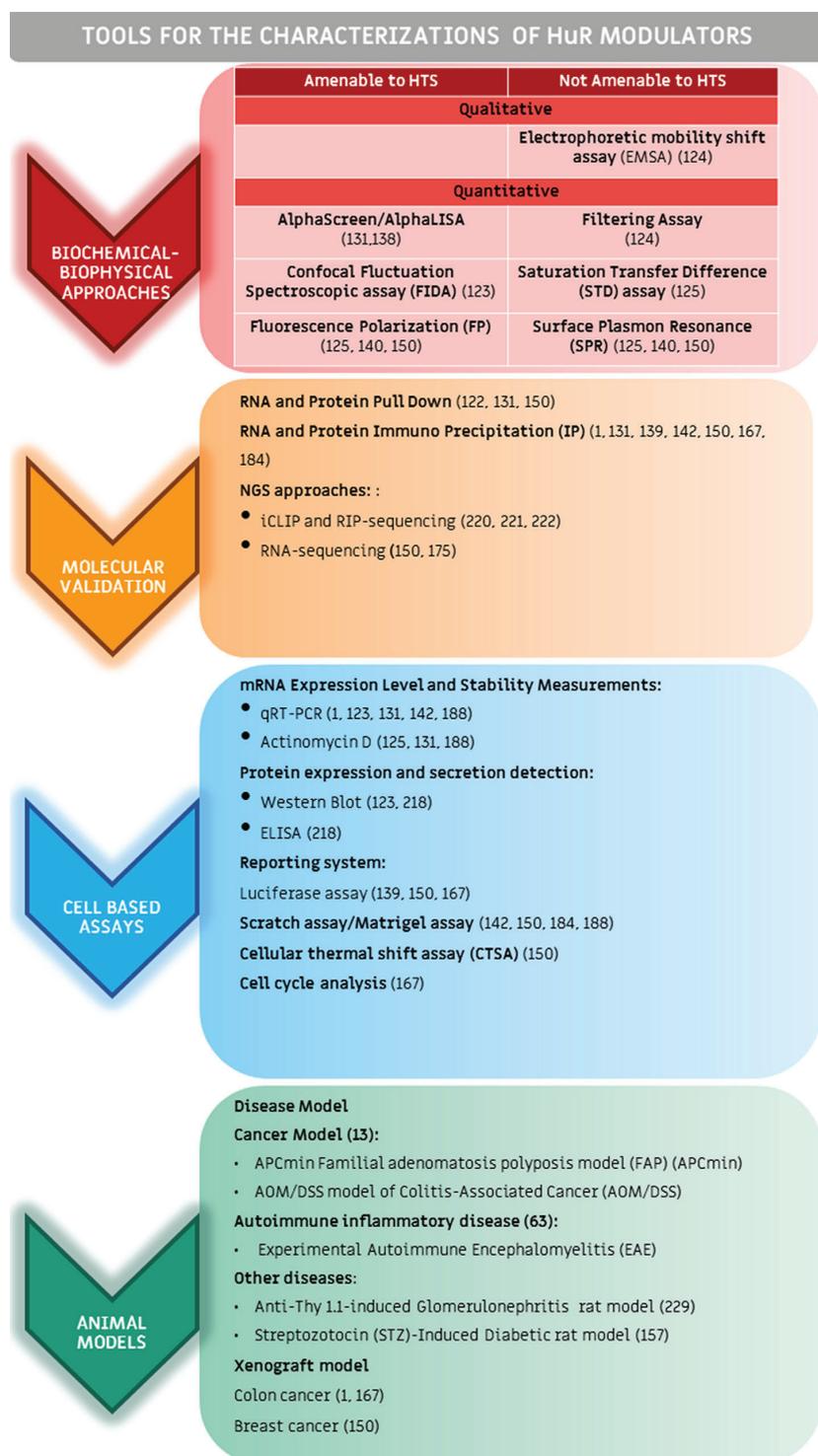
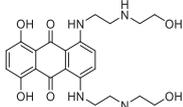
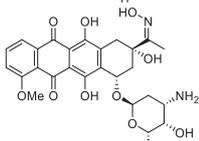
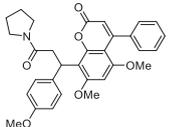
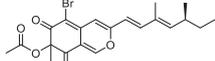
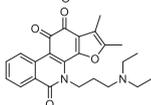
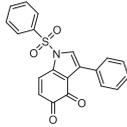
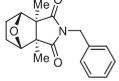


Fig. 2. Tools for the discovery, development and characterization of HuR modulators. This pipeline lists techniques, strategies, cellular and *in vivo* models mostly applied for the identification and characterization of HuR inhibitors. A number of biochemical and biophysical approaches (red boxes) were used for the identification (e.g. high-throughput screening-HTS methods) and *in vitro* validation of HuR modulators (e.g., EMSA assays). Molecular validation assays (orange boxes) were then used to elucidate the molecular mechanisms leading to HuR inhibition (e.g., RNA immunoprecipitation). Finally, cell-based assays (e.g., scratch assays) (blue boxes) and *in vivo* cancer models (green boxes) used to preclinically evaluate HuR hits and leads as anti-cancer and inflammatory candidates are listed.

Table 2
Naturally occurring HuR inhibitors: chemical structures, identification process. P = proven HuR binders, O = others.

	Entry	Molecule	Structure	Origin	P/O	Ref
Natural compounds	1n	MS-444		HTS of ≈50,000 microbial, mycological and plant extracts using a confocal fluctuation spectroscopic assay	P	[124]
	2n	Dehydromutactin		As 1n	P	[124]
	3n	Okicenone		As 1n	P	[124]
	4n	Quercetin		EMSA screening of 179 molecules from a chemical library at the Korea Research Institute of Chemical Technology (KRICT).	P	[125]
	5n	C11/Clorobiocin		Fluorescence polarization (FP)-assay followed by STD-NMR validation, using the NCI diversity set V.	P	[126]
	6n	Rutin		STD-NMR and molecular modelling.	P	[130]
	7n	Novobiocin		As 6n	P	[130]
	8n	DHTS I		AlphaScreen assay of 107 commercially available anti-inflammatory compounds	P	[1]
	9n	Cryptotanshinone		Known DHTS analogue	P	[132]
	10n	Triptolide		Known anti-inflammatory and anticancer activities	O	[134]
	11n	Kalopanax saponin A (KPS-A)		In vitro biological profiling of KPS-A	O	[135]
	12n	Ursolic Acid		In vitro biological profiling of ursolic acid	O	[136]
	13n	Latrunculin A		Known antimitotic effects of latrunculin A	O	[137]

Table 3
HuR inhibitors inspired by natural products: chemical structures, identification process. P = proven HuR binders, O = others.

	Entry	Molecule	Structure	Origin	P/O	Ref
Nature-inspired synthetic compounds	1ns	Mitoxantrone		HTS of ≈ 2000 molecules and secondary verification with RNA-EMSA.	P	[138]
	2ns	C10/Daunomycin 3-oxime		As 5n – Table 2	P	[126]
	3ns	CMLD2		FP-based HTS of ≈ 6000 compounds, validation by an AlphaLISA assay, SPR, RNP IP, and luciferase reporter functional studies	P	[139]
	4ns	AZA-9		FP-based HTS of ≈ 2000 compounds from the NCI library plus in-house compounds, validation by SPR, NMR, and computational modeling	P	[140]
	5ns	22h		DHTS I-inspired modifications	P	[141]
	6ns	AzaTanshinone 6a		DHTS I-inspired modifications	P	[142]
	7ns	N-Benzyl-cantharidinamide		Known antitumoral effects of cantharidine	O	[143]

fluorescence polarization (HT-FP) assay run on the National Cancer Institute (NCI) diversity set V ($\approx 1,600$ compounds). Twelve selected hits were validated using saturation transfer difference (STD)-NMR [128] and analytical ultracentrifugation [129], which confirmed a direct HuR-hit interaction for clorobiocin and daunomycin 3-oxime and a medium micromolar inhibition constant for both ($41.9 \mu\text{M}$ and $21.7 \mu\text{M}$ respectively).

A computational- and NMR-driven effort on a small subset of 28 validated, naturally occurring HuR inhibitors run a first stability-solubility test for STD-NMR studies, selecting 13 drug-like hits; STD-NMR confirmed the interaction between HuR and twelve hits, and rutin and novobiocin (Table 2, entries 6n and 7n respectively) were selected as the most prospective validated hits. Inhibition constants were not reported [130].

A validated amplified luminescent proximity homogeneous assay (AlphaScreen) format, measuring the inhibition of an interaction between human HuR and the ARE domain of TNF- α mRNA, was used to screen a set of 107 commercially available anti-inflammatory compounds [131]. Out of eight hits, after further validation with an RNA EMSA assay, dihydrotanshinone I (DHTS I) (Table 2, entry 8n) was identified as a nM inhibitor of recombinant HuR (rHuR)-mRNA complex formation ($0.149 \mu\text{M}$ in REMSA, $0.068 \mu\text{M}$ in AlphaScreen assay), endowed with cytotoxic cellular activity [1,131]. A few other, naturally occurring tanshinones (i.e., cryptotanshinone – Table 2, entry 9n) resulted to be slightly less potent [131,132].

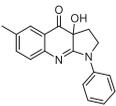
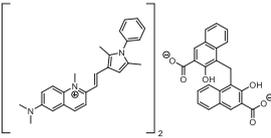
In addition to validated HuR interactors, several compounds of natural origin affected HuR functions; although no direct HuR-compound interaction was proven for any of them, they are described here. Several naturally occurring terpenoids, well known in traditional Chinese medicine as anti-inflammatory and anti-cancer compounds [133], were submitted to mechanistic studies in order to identify their molecular target; they were found to interfere with the cytoplasm-nucleus translocation of HuR, reducing stability and expression of various HuR-interacting, tumor-related mRNAs. Triptolide (Table 2, entry 10n) was characterized as a TNF- α -dependent COX-2 expression inhibitor [134], due to a reduction of cytoplasmic HuR in A549 cells. Similarly, kalopanax saponin A (KPS-A) (Table 2, entry 11n), used in traditional medicine against rheumatoid arthritis and diabetes, inhibited phorbol 12-myristate 13-acetate (PMA)-induced cytoplasmic translocation of HuR, and reduced HuR-dependent matrix metalloproteinase 9 (MMP-9) mRNA stability and expression [135]. Ursolic acid (UA) (Table 2, entry 12n), isolated from the leaves of many plants, fruits and flowers, was also characterized as being able to reduce adriamycin resistance by promoting HuR cytoplasm-nucleus translocation, and by decreasing the mRNA stability of the HuR interactor multidrug resistance gene 1 (MDR1), thus reducing MDR1 expression [136]. A few modulators of actomyosin cytoskeleton assembly such as latrunculin A (Table 2, entry 13n), an actin polymerization inhibitor toxin extracted from the Red Sea sponge *Negombata magnifica*, and blebbistatin (Table 4, entry 16s), a synthetic highly

Table 4
Fully synthetic HuR inhibitors: chemical structures, identification process. P = proven HuR binders, O = others.

	Entry	Molecule	Structure	Origin	P/O	Ref
Synthetic compounds	1s	b-40		As 4n - Table 2	P	[125]
	2s	b-41		As 4n - Table 2	P	[125]
	3s	H1N		HTS of ≈89,000 compounds by CONA (confocal nanoscanning/bead picking)	P	[144]
	4s	5		AlphaScreen assay	P	[145]
	5s	Suramin		DSF screening of an FDA approved library	P	[147]
	6s	1		Rational design by computational methods	P	[148]
	7s	3		Rational design by computational methods	P	[148]
	8s	4		Rational design by computational methods	P	[148]
	9s	2		Biophysical fragment-screening	P	[149]
	10s	3		Biophysical fragment-screening	P	[149]
	11s	KH-3		FP-based HTS of ≈2000 compounds from the NCI library plus in-house compounds, validation with an AlphaLISA assay and SPR	P	[150]
	12s	VP12/14		Rational design by computational methods	P	[151]
	13s	VP12/110		Rational design by computational methods	P	[151]
	14s	SRI-42127		Structural optimization of a HTS hit	P	[152]
	15s	MPT0B098		Known anticancer properties of MP70B098	O	[153]

(continued on next page)

Table 4 (continued)

Entry	Molecule	Structure	Origin	P/O	Ref
16s	Blebbistatin		Known antimetabolic effects of blebbistatin	0	[137]
17s	Pyrvinium pamoate		HTS of FDA-approved drugs for inhibition of HuR expression after UVC irradiation	0	[111]

selective inhibitor of non-muscle myosin II ATPase activity, were able to reduce the translocation of HuR from nucleus to cytoplasm in HepG2 and Huh7 hepatocarcinoma (HCC) cells [137].

2.2. Nature-inspired synthetic compounds

In addition to previously mentioned daunomycin 3-oxime/C10 (Table 3, entry 2ns) [126], several nature-inspired synthetic compounds have been identified as HuR inhibitors.

Using the experimental setup described earlier (AlphaScreen homogeneous assay, followed by an RNA EMSA assay) a library of 2,000 small molecules, including clinically tested candidates and natural products, was screened [138]. The anthraquinone mitoxantrone (Table 3, entry 1ns) resulted to be able to interfere with rHuR – TNF α mRNA complex formation.

Two years later, the coumarin-like derivative CMLD-2 (Table 3, entry 3ns) was identified through a FP-based HTS on a 6,000-membered library containing FDA-approved drugs and in house made compounds [139]. CMLD-2 was then validated through a rich profiling cascade including a biochemical AlphaLISA assay, surface plasmon resonance (SPR), ribonucleotide immunoprecipitation (RNP IP), and luciferase reporter functional studies, displaying a dose-dependent effect on HuR. Another library of \approx 2,000 molecules, including compounds from the NCI Diversity Set II, a small set of natural products, FDA-approved oncology drugs and a few in-house compounds was screened using the same FP-based biochemical competition assay [140]. Azaphilone 9 (AZA-9) (Table 3, entry 4ns), a semi-synthetic derivative built on the fungal natural asperbenzaldehyde scaffold, was characterized as the most potent hit through SPR and NMR (1.2 μ M for full length HuR, 7.4 μ M for HuR RRM1/2); computational studies were carried out to better pinpoint the molecular interaction between HuR and azaphilone 9.

Using previously mentioned 8n (Table 2) as structural guidance, two research groups successfully obtained synthetic analogues endowed with higher potency and better bioavailability. A small set of lactam tetracyclic compounds inspired by 8n was synthesized [141], out of which compound 22 h (Table 3, entry 5ns) was selected as a bioavailable early lead. Rational design and medicinal chemistry led us to a small array of bicyclic indolesulfonamide tanshinone mimics, out of which azatanshinone 6a (Table 3, entry 6ns) was selected as an early lead [142]. Namely, a profiling cascade entailing biochemical Alpha-Screen and RNA EMSA assays, followed by cytotoxicity assays in breast MCF7 and MDA-MB-231 cancer cell lines, and on pancreatic ductal adenocarcinoma PANC-1 cell lines, established the potent cellular antitumor activity of compound 6a.

As to inhibitors of HuR functions without a proven direct interaction, *N*-benzyl-cantharidinamide (Table 3, entry 7ns), a synthetic analogue of the topic, naturally occurring cantharidine drug, was reported to reduce MMP-9 expression and the invasive potentials

of hepatoma Hep3B cells by inhibiting cytosolic translocation of HuR, thus reducing HuR-mediated MMP-9 mRNA stability [143].

2.3. Synthetic compounds

In addition to previously mentioned b-40 [125], b-41 [125] and blebbistatin [137] (Table 4, entries 1s, 2s and 16s respectively), a few fully synthetic compounds have been characterized as effective HuR inhibitors.

At first, a representative diversity set of 89,000 compounds from a large, 2.2 M bead-based library was screened by confocal nanoscanning/bead picking (CONA) [144]; after decoding by MS the structures of 46 most recurring hits, they were re-synthesized and validated as single compounds with full-length HuR and HuR₁₂. Out of six confirmed hits, polyamidoamine H1N (Table 4, entry 3s) was identified as the first RRM3-targeted, low molecular weight HuR inhibitor, while the vast majority of reported HuR inhibitors bind to HuR RRM1 and RRM2.

In a glioma-targeted project, an AlphaScreen assay measuring hinge phosphorylation and subsequent oligomerization of HuR was used for a HTS campaign [145]; no details were given about size and nature of the screened collection. Benzimidazoleamide compound 5 (Table 4, entry 4s) was identified as a putative disruptor of HuR multimerization. Suramin (Table 4, entry 5s), a polysulfonated naphthylurea known for its antitrypanosomal, anticancer, and antiviral properties [146], was identified through a differential scanning fluorimetry (DSF)-based screening of a 1570-membered library of FDA-approved compounds; 55 initial hits were further profiled in more demanding DSF conditions, and suramin resulted as a single, fully validated hit. A low affinity HuR-suramin interaction ($K_d = 0.24 \mu$ M) was confirmed by SPR [147].

The first structure-based rational design of HuR ligands, based on a virtual screening (vHTS) campaign using a platform specifically set to identify novel scaffolds/molecules as inhibitors of macromolecular interactions, was recently reported [148]. Ranking among 200 virtual chemotypes led to the selection of 17 specific hits, which were docked into the HuR binding site and more extensively studied. As a result four compounds were synthesized, three (Table 4, entries 6s, 7s and 8s) were confirmed as HuR interactors using a combination of STD-NMR and *in silico* studies, identifying contacts with the RNP regions of RRM1 and RRM2 HuR domains, and polyphenol benzamide 4 (Table 4, entry 8s) was qualitatively suggested being the most potent based on the intensity of the STD signal.

Aiming to improve the activity of polyphenol benzamide 4 (8s), compounds 2 and 3 (Table 4, entries 9s and 10s respectively) were designed on the basis of a SPR-fragment screening [149]. Their interaction with HuR was then also evaluated by STD-NMR, and their interference with the HuR-RNA complex was quantitated with an FP assay ($IC_{50} = 105 \mu$ M and 92μ M for compound 2 and 3 respectively).

Following the earlier described discovery of CMLD-2 [139] and Aza-9 [140], using the same FP-based HTS – AlphaLISA and SPR profiling cascade on a ≈ 2000 -membered compound collection, benzothioephene hydroxamate KH-3 (Table 4, entry 11s) was identified as a HuR-ARE^{Msi1} (Musashi RNA-binding protein 1) disruptor with low micromolar potency (3.5 μM in an FP assay, 2.3 μM in an AlphaLISA assay) and functional HuR effects [150].

In a study directed towards the identification of novel HuR-mRNA binding inhibitors, 28 derivatives based on indole and caffeic acid phenethyl ester (CAPE) scaffolds were designed and synthesized [151]. Among them, indole-based compounds VP12/14 and VP12/110 (Table 4, entries 12s and 13s respectively) were confirmed as HuR interactors in an *in vitro* assay.

Recently, bisheteroaryl compound SRI-4217 (Table 4, entry 14s) was found to bind and inhibit HuR dimerisation in primary patient-derived glioblastoma xenografts (PDGx) with a 1.2 μM IC₅₀. An interaction with the RRM1 and RRM2 domains was proposed through computational studies [152].

Few other synthetic HuR inhibitors were identified, such as Indoline sulfonamide MPTOB098 (Table 4, entry 15s) [153]. In addition to its antimetabolic activity through microtubule binding, MPTOB098 significantly decreased HuR translocation from the nucleus to the cytoplasm in A549, HONE-1 and PC3 tumor cells, subsequently reducing hypoxia-inducible factor 1- α (HIF-1- α) protein expression.

Finally, pyrvinium pamoate (Table 4, entry 17s), an FDA-approved anthelmintic drug, was found to dose-dependently inhibit cytoplasmic accumulation of HuR by activating the AMP-activated kinase/importin $\alpha 1$ cascade and inhibiting the checkpoint kinase1/cyclin-dependent kinase 1 pathway [111].

2.4. Nanoparticle-encapsulated HuR siRNA as therapeutic agents

HuR was successfully targeted also through nanoparticle (NP) -based HuR-specific small interfering RNA (HuR siRNA) delivery. The efficacy of folate receptor- α (FRA)-targeted DOTAP:Cholesterol lipid NPs carrying HuR siRNA (HuR-FNPs) was tested against human lung cancer cells [154]. A folic acid (FA)-based FA-PEG-DSPE construct, 1,2-dioleoyl-3-trimethylammonium-propane chloride salt (DOTAP), and cholesterol were used to build liposomes with a particle size of 303 nm and a surface charge of + 4.3 mV. Folate-conjugated liposomes showed efficient internalization in lung cancer cells *via* folate receptor endocytosis, and serum stability and gel retardation assays revealed that such liposomes protected HuR siRNA from rapid degradation [154]. A method to deliver siRNA against HuR *in vitro* and *in vivo* was set up and optimized by conjugating FA to 3DNA nanocarriers in ovarian cancer models [155]. A transferrin receptor-targeted, liposomal NP-encapsulated HuR siRNA (siRNA-Tf-NP) was investigated as a therapeutic agent against HuR by employing sulfhydryl reactive crosslinking chemistry to synthesize Tf-PEG-DSPE [156]. HuR siRNA was administered intraocularly as nano-based lipidic systems in streptozotocin (STZ)-induced diabetic rat models. Such lipoplexes caused an efficient decrease of pathologically elevated HuR and VEGF retinal levels. Finally, nanocarrier-transported siRNA showed an amelioration of the retinal damage caused by STZ, increasing retina thickness and the number of retinal ganglion cells (RGCs) up to homeostatic levels, compared to their reduction observed in diabetic rats alone and when receiving naked siRNAs [157].

In conclusion, the large number of reported small molecules demonstrates the druggability of HuR. HTS campaigns were at first employed to identify small molecules able to either directly or indirectly interfere with HuR activity, due to the lack of structural information about the HuR-mRNA interaction at that time. Although these structurally heterogeneous compounds, being

either natural, synthetic or natural compound-inspired, often consist of complex molecular structures and/or show sub-optimal pharmaceutical properties, they were useful starting points to develop more specific HuR modulators. More recently, a detailed knowledge of HuR structure and binding modes to mRNA and small molecule modulators has enabled the rational design, synthesis and characterization of new ligands specifically designed through computational methods, as extensively described in the next Chapter. We do believe that both HTS campaigns / access to unpredictable structures of HuR inhibitors, and rational drug discovery / further exploitation of structural information on HuR will be exploited in future to enrich the panel of existing HuR modulators and their properties. Their potential as putative clinical candidates against oncological diseases will be commented upon in details in Chapter 4.

3. HuR structural features and interactions with low molecular weight inhibitors

Full-length HuR is a multi-domain protein constituted by three RRM domains of about 90 amino acids long. Namely, two conserved RRM domains near the N-terminus are named RRM1 (20–98) and RRM2 (106–186), linked by a 12 amino acid linker and preceded by an intrinsically disordered region of 20 amino acids; and a third recognition motif, named RRM3 (244–322), is located at the C-terminal region. RRM3 is separated from RRM2 by a longer basic linker that involves the 60 amino acid-long HNS, which is mainly responsible for the nuclear-cytoplasmic shuttling of HuR to stabilize and/or enhance the target mRNA translation efficiency (see Fig. 3A) [3,4,130,158].

In the last decade ten 3D structures were deposited in the PDB databank (nine resolved by X-ray diffraction, one by solution NMR). Among them, three include only the RRM1 domain (PDB codes: 4FXV, 3H19, and 5SZW for the NMR structure), two span the tandem RRM1-RMM2 domains (PDB codes: 4ED5 and 4EGL), and five relate to the RRM3 motif (PDB codes: 6GD2, 6GD3, 6G2K, 6GD1 and 6GC5); none of them include any complexed inhibitor in the crystal structure. Although structural guidance for HuR modulation is now available, most HuR inhibitors discovered so far (see Chapter 2) have been identified through experimental HTS.

3.1. Structural insights on HuR

X-ray structures collected on the isolated domains and on the RRM1-RMM2 tandem domain of HuR reveal the typical architecture of the RNA recognition motifs in eukaryotic RNA binding proteins. They present four-stranded antiparallel β -sheets packed against alpha helices, adopting the canonical $\alpha\beta$ structure with a $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$ topology [159,160]. Also, high structural similarity has been detected between the first two RRM domains of HuR, with the exception of a different conformation involving a β -hairpin located at the $\alpha 2-\beta 4$ loop, which in RRM1 adopts a β -turn- β conformation that is not present in RRM2 [158,159].

Characterization of the HuR-mRNA-binding activity allowed the understanding of the role for each RRM domain. The interactions between mRNAs and HuR are generally affected by any modification of the residues within the RRM domains. Conversely, modifications on the residues within the HNS sequence alter the sub-cellular localization of HuR [10].

Information on the mRNA-bound structures of HuR have been obtained from the complexes of the protein with short AREs [3,4,158,160,161]. HuR RRM1 and RRM2 control the recognition of any target mRNAs. In particular, the RRM1-RMM2 tandem construct has been reported to bind a 11-base AU-rich strand with high nanomolar affinity, resulting from adding the micromolar

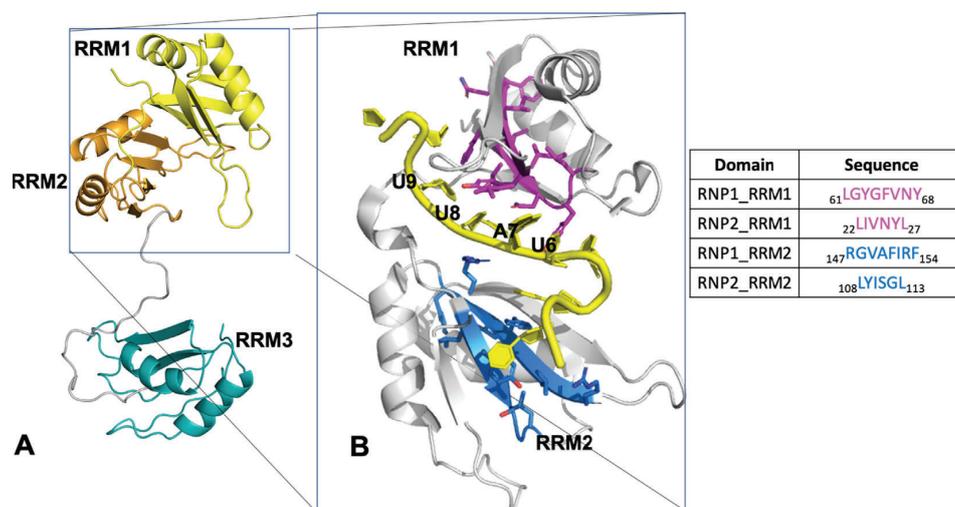


Fig. 3. A) Representative 3D model of full length HuR (Chimera model obtained with Prime-Schrodinger software); the protein is shown in ribbons with RRM1, RRM2 and RRM3 displayed in yellow, orange, and cyan, respectively, and the long basic linker between RRM2 and RRM3 in grey. B) Co-crystal structure of the tandem RRM1 and RRM2 HuR-mRNA^{c-fos} complex (pdb code 4ED5); the protein backbone is shown in grey, the RNP1 and RNP2 sequences of the RRM1 domain in magenta, the RNP1 and RNP2 sequences of the RRM2 domain in blue, mRNA^{c-fos} in yellow.

affinity of RRM1 to the weak affinity of RRM2 and to further contacts established by the short interconnecting linker between them. The same tandem construct recognizes and binds U-rich sequences (U-rich RNA and U-rich DNA), with even higher affinity and with a preference for U-rich RNA sequences [162].

A detailed analysis of the experimental structure of HuR complexed with a 11-mer RNA oligonucleotide derived from c-fos mRNA (PDB code 4ED5) shows that the binding site is mostly hydrophobic. The residues involved in the interaction are placed at the two canonical ribonucleoprotein sequences named RNP1 and RNP2 (see Fig. 3B) [163], located at the center of the β -platform ($\beta 1$ and $\beta 3$) in the RRM domain.

Additional residues located around the β -strands and in the RRM1-RRM2 linker contribute to the binding with mRNA^{c-fos} by establishing weaker interactions. Residues N24 and F65 located respectively at the $\beta 1$ and $\beta 3$ strands [158,159] are reported to play a binding role, while residues located at the $\alpha 2$ helix, $\beta 2$ - $\beta 3$ and $\alpha 2$ - $\beta 4$ loops and belonging to the inter-domain linker experience conformational change upon RNA binding.

As to RNA, structural data show that the RRM1 domain recognizes up to 5 consecutive uracils, while the inter-domain linker and RRM2 bind only to one or two nucleotides [1,3]. A mechanism for the HuR-mRNA binding has been recently proposed [3]. It entails a first interaction with the RRM1 domain, followed by conformational changes involving the inter-domain linker and RRM2, inducing them both to bind the RNA strand, and thus stabilizing the HuR-mRNA complex.

As to the third recognition motif, despite previous studies suggested its negligible contribution to the interaction between HuR and mRNA strands, more recently an isolated RRM3 domain was also shown to recognize ARE sequences, contributing to the interaction with the target mRNAs by binding their poly-A tail, as well as being necessary for the cooperative assembly of HuR oligomers. In fact, disruptions at the dimerization interface result in a decreased binding affinity between HuR and its mRNA targets [10,160,164,165].

The RRM3 domain can bind both to AUUUA motifs and poly(U) sequences, with a preference for the U-rich strands [160]. Similarly, to other RRMs, the binding interface in RRM3 is located in RNP1 ($\beta 3$), involving residues K285-M292, and in RNP2 ($\beta 1$)

involving residues I246-L251. Residues F287 and M292 from RNP1 and F247-L251 from RNP2 are crucial for the interaction. Later studies [4,161] confirmed both the binding site and preference for U-rich sequences and determined that HuR-mRNA complexes are stabilized by a combination of base stacking and polar and hydrogen bond interactions. Although there are conflicting reports on the number of binding pockets on the RRM3 domain, different studies agree that RRM3 can recognize the UUU/A motif. RNP2 is responsible for the direct interaction with uracil and RNP1 for other nucleotides such as adenines. (A/U)UU(A/U) has been proposed as the RNA binding recognition sequence [161], while mutations at the binding interface of isolated RRM3 domains as well as within full-length HuR [4] confirmed the key role of F247 and Y249 in the HuR-mRNA interaction, and the minor contribution provided by F287 and F288 residues. Notably, ATP was identified as a natural RRM3 ligand responsible of a surprising, RNA-modifying, terminal adenosyl transferase activity of HuR, suggesting an additional role for HuR in the maturation and metabolism of mRNA targets [144].

As to the dimerization process of HuR, studies based on EMSA assays have also shown the propensity of isolated RRM1 domains to dimerize [166]. Interestingly, an equilibrium between multimeric forms of RRM1 has been described, with their predominance on monomeric and dimeric forms in the absence of mRNA sequences. The RRM1 domain region responsible for the dimerization process comprises the β -hairpin and α -helices. Besides, mutagenesis experiments reveal the involvement of a disulfide bridge at Cys13, which may have functional significance in redox modulation of HuR activity in response to oxidative stress [159]. Of note, the dimerization region is placed on the opposite site with respect to the mRNA binding surface. However, the interaction with mRNA sequences involves a slightly overlapped protein region, thus promoting the dissociation of the RRM1 dimeric form and shifting the equilibrium to the monomeric state [158,159].

Another important recognition motif for the dimerization of HuR is in the RRM3 domain. Recent studies on the C-terminal region of the protein proved the existence of a dynamic equilibrium between dimeric and monomeric forms of RRM3 even in the absence of RNA sequences. The extent of dimerization appears to be dependent on the concentration of RRM3 domains, and the dimeric form is stabilized by interaction with mRNA sequences.

It is interesting to note that both sequence and length of the mRNA strand play a role in the stabilization of a HuR dimer. In particular, in presence of short mRNA constructs, the dimeric state of HuR might be disfavoured, while both RRM3 domains forming a dimer can bind to the same, long mRNA strand (more than 14 nucleotides) [4,160,161,164]. X-ray and NMR studies on the RRM3 domain, together with Molecular Dynamic (MD) calculations allowed the identification of a dimerization surface constituted by the α 1 helix and the loop between α 1- β 2. Moreover, the W261 residue plays a crucial role, since any mutation involving this amino acid leads to an increase of HuR monomers [4,160,161,164].

The internal dynamics of HuR have been characterized in detail by NMR. Measurements of the longitudinal (R1) and transverse (R2) relaxation rates of the backbone amide nitrogens recorded on the RRM1-RRM2 tandem domains prove that this construct does not behave as a rigid body, but rather displays inter-domain flexibility with the two domains that can reorient with respect to one another [1].

Interestingly, the same relaxation data suggest that RRM1-RRM2 dimers are not present in solution as stable complexes. Another important finding from NMR studies concerns the RRM3 domain, that does not seem to interact with the other two recognition motifs in the absence of mRNA [160].

A comprehensive analysis of the experimental structures obtained by X-ray crystallography and NMR data indicates that in the absence of mRNA strands the three domains in full-length HuR, as well as RRM1 and RRM2 in the tandem domain construct, move independently and maintain an open/flexible conformation (Fig. 4).

After binding to a target mRNA, the protein adopts a closed globular conformation, forming a positively charged cleft where additional contacts between RRM1 and RRM2 can be detected. Namely, RNP1 and RNP2 sequences come in close proximity, especially around the U6, A7, U8 and U9 nucleotides, providing in this region a narrower binding pocket, possibly amenable to the rational design of organic HuR/RNA disruptors (see Fig. 3B).

3.2. Computational and NMR studies to elucidate the interaction between HuR and small molecule inhibitors

Many HuR inhibitors were discovered and tested, especially *in vitro*, in the last decade. However, for most of them the molecular mechanism of HuR inhibition has been poorly investigated and characterized at the atomic level, and for a few it is completely unknown. It has been proposed, solely relying on docking studies, that small organic ligands should bind HuR at the cleft between the RRM1 and RRM2 domains [149,150,152,163,167]. This theory has been recently validated through combined NMR and computational studies [1,130,140,142].

For instance, using both NMR titration and docking calculations, the interaction of **4ns** (see Table 3) with HuR was elucidated. Specifically, NMR-titration experiments allowed the identification of a pool of residues interacting with the molecule (I103, L138, Y26, R97, I103, Y109 and R153) which delimit the mRNA binding cavity. In line with NMR results, docking of **4ns** into HuR suggests Y26, K55, R97, and R153 as residues surrounding **4ns**, establishing electrostatic, hydrogen bond, hydrophobic, and pi-stacking interactions with the small molecule (Fig. 5A). Thus, **4ns** appears to disrupt the HuR-mRNA interaction by competitively binding to the RRM1-RRM2 interdomain cleft of HuR [140].

Similar results were subsequently obtained with twelve naturally occurring HuR inhibitors including flavones, flavonols and coumarins, which were studied through STD-NMR experiments and docking to explore the ligand-protein interaction mode [130]. These studies revealed that all compounds interacted with

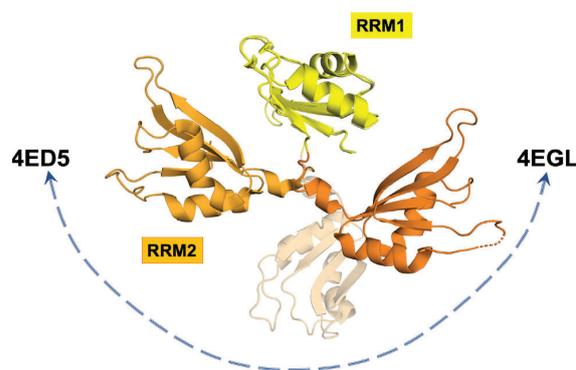


Fig. 4. A superposition of NMR- (PDB code: 4EGL) and X-ray-derived (PDB code: 4ED5) HuR structures shows that RRM1 (yellow cartoon) and RRM2 (different shades of orange cartoon) domains experience conformational freedom with respect to each other [3].

HuR in STD-NMR with different strengths (see for example **6n** and **7n** in Table 3); that all bind into the deep pocket between the RRM1 and the RRM2 domains, stabilizing a “closed” conformation of the protein; and that some interactions are conserved among flavones, flavanols and coumarins, including those with Y63, R97, N25 and R153 side chains that have been described as important also for **4ns** and **6s** [148].

We reported structural details on the interaction mode of **8n** and **6ns** (see respectively Table 2 and Fig. 5B, and Table 3 and Fig. 5C) with HuR [1,142]. A combined approach using NMR, MD simulations and mutagenesis coherently indicated that these small molecules bind to the region between the interconnecting linker and the β -platform of both RRM1 and RRM2, altering the conformational freedom and the reciprocal reorientation of RRM1 and RRM2. This stabilizes an unproductive, “closed” conformation that hampers target mRNA binding (Fig. 5D).

Noteworthy, a different mechanism of action has been described for compound **3s** [144]. Confocal nanoscaning-bead picking experiments showed that this molecule binds to the RRM3 domain. *In vitro* studies indicate that **3s** interferes with ATP as well as RNA binding within the RRM3 domain. Docking calculations of **3s** into the homology modeling structure of RRM3 (no RRM3 domain was crystallized at that time) showed that the relatively large ligand occupies both a conserved DxD motif (D₂₅₄, D₂₅₆), commonly used for recognition of ATP, and an adjacent shallow and positively charged cleft, very likely binding the target mRNA, thus hampering the accommodation of both RNA and ATP in line with experimental data.

3.3. Computer-aided hit discovery

In 2019, a first example of virtual screening has been reported [148] through which three structurally diverse ligands were found to inhibit HuR using the NucleoQuery application within the free Web-server platform AnchorQuery. Specifically, the authors selected the mRNA U8 and U9 nucleotides as anchor points and the NucleoQuery application detected all the possible interactions with the protein interface, allowing the final selection of several pharmacophore points. Such pharmacophoric query identified a library of 800 structurally diverse, synthetically accessible molecules as putative HuR inhibitors, all possessing a specific aromatic moiety (superimposable with the U8 of mRNA) as an anchor point. Four representative compounds were synthesized, three of which (Table 4, **6s-8s**) were characterized as HuR binders in an STD-NMR assay. Just recently, **8s** has been successfully optimized

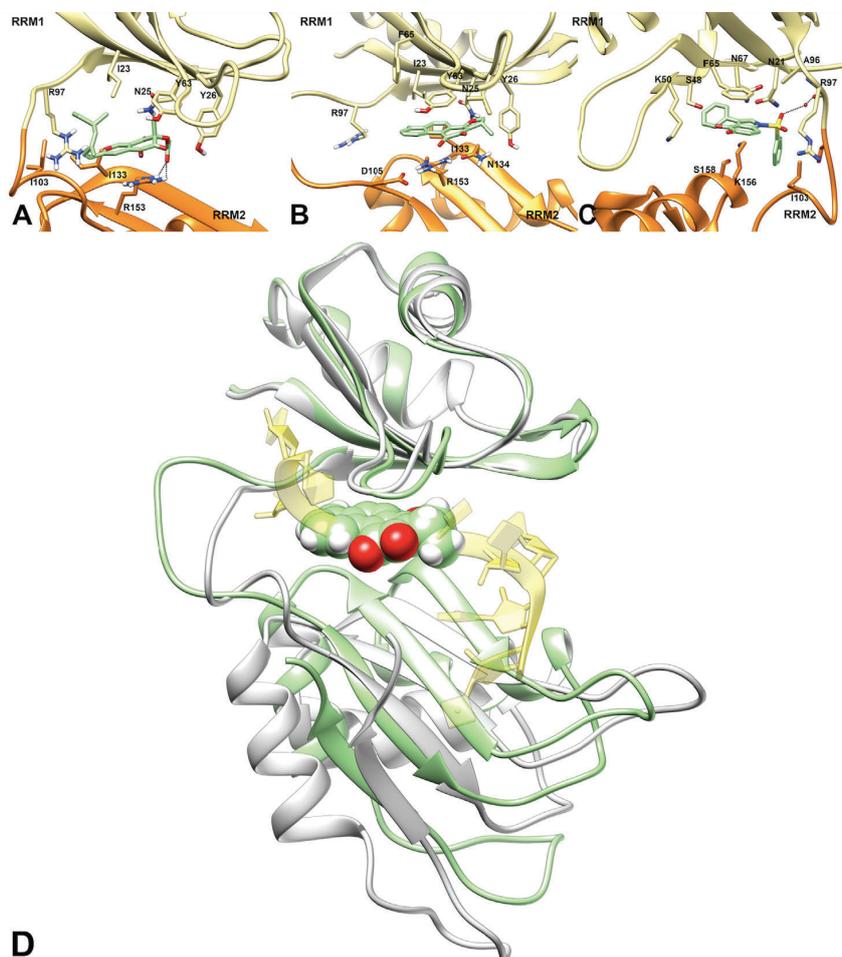


Fig. 5. A), B), C) Binding modes of **4ns** (A), **8n** (B) and **6ns** (C) (green sticks) into HuR (yellow and orange cartoons), as determined by computational studies. HuR residues involved in binding interactions are displayed as sticks, and H-bonds are depicted as dashed black lines. D) 3D superposition of the HuR conformation (green cartoons) induced by **8n** (green spheres) upon binding, and the mRNA-bound HuR conformation (PDB code 4ED5, grey and transparent yellow ribbons for HuR and mRNA, respectively).

through SPR and STD-NMR fragment screening [149]. Docking experiments confirm that also optimized compounds **9s** and **10s** bind in the same region occupied by the U8 mRNA base. Most importantly, both studies demonstrate that computer based HuR ligand discovery is achievable, and that the molecular interaction between U8 in mRNA and Y63 in HuR represents an important anchor point to design small molecules suitable for HuR recognition and binding [148,149].

Another successful, recent example of computer-aided hit discovery on HuR consisted of a shape similarity screening using **8n** (Table 2) as a template and a proprietary database of 182 drug-like molecules [151]. As a result, 28 putative HuR ligands mostly built on indole or caffeic acid phenethyl ester scaffolds were found and clustered on the basis of their structural interaction fingerprints (SIFts). Two putative ligands (**12s** and **13s**, Table 4) were found in the same cluster of potent **1n** and **8n**, and thus were synthesized. Although a direct binding assay with HuR was not reported, cellular assays demonstrated that **12s** and **13s** modulate HuR expression and decrease VEGF and TNF- α release, similarly to **8n**. Noteworthy, MD simulations of **12s**, **13s** and **8n** into the binding site of HuR demonstrated that **13s** and **8n** should compete with mRNA to bind to HuR, possibly inhibiting the functional effects of HuR on mRNAs.

NMR and modeling data reported so far for HuR inhibitors surely indicate, with the single exception of **3s**, their binding in the mRNA U6-U9 binding site of the RRM1-RRM2 construct, stabilizing a closed, unproductive conformation of HuR that in turn hampers mRNA accommodation in its binding site. Although HuR inhibitors are chemically diverse, and no pharmacophore hypothesis has been proposed up to now, some general, common features could be detected. Namely, one or more aromatic rings (rarely saturated rings) in the same scaffold, which confer structural rigidity and a hydrophobic character to HuR inhibitors, then fitting well in the mRNA U6-U9 binding region of the HuR protein; and functional groups such as carbonyls, hydroxyls, amines and carboxylic acid to increase polarity and interact with the hydrophilic residues in the binding cleft. For instance, at least one polar contact with the side chains of R97 or R153 residues should be established, as compounds **6n**, **6ns** and **6s** possess a carbonyl group (a sulfonyl for **6ns**) that establishes a hydrogen bond with R97, while compounds **7n**, **4ns** and **6s** are endowed with a hydroxyl or a carbonyl moiety establishing one or two hydrogen bonds with R153. Other residues frequently interacting with active molecules are S99 and S100 on the interdomain linker, and Y63 and N25 in RRM1.

Taken together, the reported studies have laid the foundation for the rational finding / design of novel HuR inhibitors through a

Table 5
In vivo studies showing the effects of HuR targeting in different cancers.

Cancer type	<i>In vivo</i> model	HuR modulation	Major observations	Refs
Malignant peripheral nerve sheath tumors (MPNST)	– Tumor Xenograft – Lung metastasis model	Genetic (Constitutive and inducible shRNAs) Pharmacological (1n)	– HuR knockdown prevents formation of xenograft tumors and induces regression of already formed tumors. – Genetic ablation of HuR prevents formation and growth of metastatic lung nodules. – Pharmacological inhibition of HuR blunts growth of xenograft tumors and metastatic nodules.	[171]
Gastric cancer	– Tumor Xenograft	Genetic (Overexpressing)	Overexpression of HuR promoted gastric cancer cell growth <i>in vivo</i> .	[172]
Fibrosarcoma	– Tumor Xenograft	Genetic (CRISPR/Cas9)	HuR deletion markedly diminished tumor growth on its own whereas AGI-5198 (a mutant IDH1 inhibitor) treatment combined with HuR deletion had the greatest impact on tumor growth.	[173]
Pancreatic ductal adenocarcinoma (PDAC)	– Tumor Xenograft	Genetic (DOXO-inducible shRNAs)	HuR inhibition enhances PARP inhibitor olaparib-mediated reduction of PDAC tumor growth <i>in vivo</i> .	[174]
Colon carcinoma	– Tumor Xenograft	Pharmacological (8n)	Compound 8n induces significant anti-tumor effects, with approximately a 4-fold reduction in tumor size.	[1]
Pancreatic ductal adenocarcinoma (PDA), colorectal cancer	– Tumor Xenograft	Genetic (CRISPR/Cas9)	HuR-deficient PDA cells were unable to engraft tumors <i>in vivo</i> compared with control cells, whereas HuR-deficient colon cancer cells showed significantly reduced <i>in vivo</i> tumor growth.	[175]
Lung cancer	– Tumor Xenograft – Lung metastasis model	Transferrin receptor-targeted liposomal nanoparticle-based HuR siRNA (HuR-TJNP) Pharmacological (17s)	HuR-TJNP treatment suppresses lung tumor growth <i>n vivo</i> and suppresses experimental lung metastasis	[154]
Bladder cancer	– Tumor Xenograft	Pharmacological (17s)	Combining compound 17s with chemotherapeutic agents (e.g. cisplatin, DOXO, vincristine and oxaliplatin) synergistically suppressed the growth of patient-derived bladder tumor xenografts in mice	[111]
Colorectal cancer (CRC)	– Tumor Xenograft	Pharmacological (1n)	Compound 1n led to an approximately 1.7-fold reduction in tumor size	[167]
Melano	– Tumor Xenograft	Pharmacological (9n)	Compound 9n effectively inhibited tumor growth and angiogenesis, decreasing the cytoplasm translocation of HuR.	[132]
Small intestinal and colon cancer	– Apcmin/– mice, a TG model of intestinal tumorigenesis. – inflammatory colon carcinogenesis protocol - AOM-DSS administration	Conditional intestine-specific HuR deleter mice: <i>Hurff villin-Cre-ERT2 mice</i>	Intestinal deletion of the HuR gene caused a three-fold decrease in tumor burden in Apcmin/– mice characterized by reduced proliferation and increased apoptosis, and a two-fold decrease in tumor burden in mice subjected to an inflammatory colon carcinogenesis protocol	[59]
Lung adenocarcinoma	– Tumor Xenograft	Pharmacological (15s)	Mice receiving compound 15s significantly delayed tumor growth in a concentration-dependent manner	[153]
Glioblastoma (GBM)	– Intracranial injections	Genetic (shRNA)	HuR knockdown induces a significant attenuation of tumor growth and invasiveness	[176]
Conventional renal cell carcinoma (CRCC)	– Tumor Xenograft	Genetic (siRNAs)	HuR knockdown inhibits human CRCC tumor growth in nude mice <i>in vivo</i> .	[177]
Colon carcinoma	– Tumor Xenograft	Genetic (siRNA & Overexpressing)	HuR-overexpressing cancer cells produced significantly larger tumors; conversely, cells expressing reduced HuR developed significantly more slowly.	[92]

range of diverse methodologies. We expect that, now that diverse HuR inhibitors are known, a few pharmacophore models will be developed and used in virtual screening (vHTS) campaigns. Besides, taking advantage of recently released RRM3 X-ray structures, novel HuR ligands targeting the RRM3 domain should be soon discovered with the aid of receptor-based finding techniques.

4. HuR inhibitors and the hallmarks of cancer

HuR is often overexpressed in many human cancers, with increased levels associated with tumor aggressiveness and worse prognosis. HuR plays a critical role in controlling almost all key cancer-associated traits, including proliferation, survival and dissemination amongst others, by regulating the expression of a plethora of genes [106,122,168–170]. Due to this fundamental role in cancer progression and metastasis, HuR has received considerable attention as a therapeutic target. Remarkably, over the years hundreds of *in vitro* and *in vivo* studies (Table 5) have consistently shown that targeting HuR is a promising strategy for a variety of cancers. In addition, genetic silencing in several cancer models

has shown benefits in terms of tumor regression, but, while these studies are promising, advancing siRNA-based therapy to the clinic remains a challenge. Thus, the importance of developing a potent, bioavailable and safe small molecule inhibitor directed against HuR to be tested and proposed for clinical studies cannot be understated.

We decided to build our detailed description of biological and pharmacological studies on HuR and on its modulators by taking inspiration from the Hallmarks and Enabling Characteristics of Cancer [120,121]. A surprisingly long list of molecular targets impacted by HuR modulation through biologicals and/or small molecules, and referred to cancer hallmarks, is graphically depicted in Fig. 6.

4.1. Genomic instability

Genome instability embraces all the mutations which affect the genome of a cellular lineage, from changes of a single base in the nucleic acid sequence to chromosomal rearrangements or aneuploidy. Malfunctions of the DNA repair machinery are among the

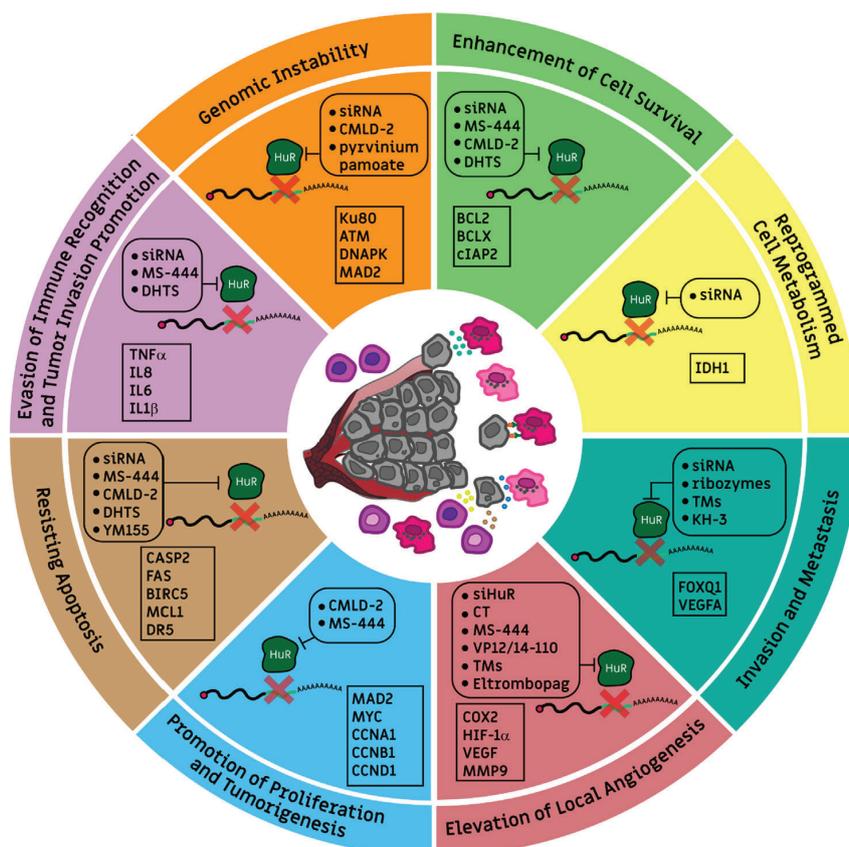


Fig. 6. HuR modulators antagonize the cancer traits. Each feature is represented by a colored slice in a circle. Each slice contains a list of HuR modulators (e.g. siRNA, small molecules, confined in rounded boxes), and a list of molecular targets affected by HuR inhibition (cellular or *in vivo* models, confined in square boxes). In the middle of the figure, a schematic representation of the tumor progression, characterized by vascular tissues (in red) cancer cells (grey) and immune cells (pink and violet) is shown.

most common causes of genomic instability. Due to its consequences, genome instability plays a central role in the development of cancer.

Regulators of centrosome stability and integrity during mitosis, such as growth factors (i.e., EGF and fibroblast bFGF), induce tyrosine phosphorylation on HuR and association with centrosomes. A proper centrosomal duplication represents a key point during mitosis. HuR dysregulations could therefore lead to strong chromosomal rearrangements [178]. The effect of HuR silencing on genomic instability has been studied through various approaches. HuR siRNA were used on triple negative breast cancer cells to study HuR knockdown in radiosensitization, and the consequent involvement of double strand break (DSB) repair [179]. HuR silencing correlates with a reduced expression of proteins involved in DNA repair, such as Ku80, ataxia telangiectasia mutated (ATM) and DNA-PK. However, the consequent positive or negative effects of HuR on cancer therapies triggering DNA damage is not straightforward. Indeed, HuR silencing combined with radiotherapy appears to improve the radio sensitization of tumor cells by decreasing DNA repair and enhancing radiation-induced ROS production. During DNA damage response (DDR) HuR translocates from the nucleus to the cytoplasm to regulate DNA repair genes upon chemotherapeutic stress, improving cancer cells' resistance to chemotherapy [103,111,180]. However, in pancreatic cancer cells, HuR overexpression favors the expression of deoxycytidine kinase (dCK) and the consequent metabolic activation of gemcitabine; therefore, HuR becomes a key determinant for gemcitabine activity [96].

Similarly, in breast cancer cells, HuR mediates doxorubicin efficacy by post-transcriptionally regulating topoisomerase II α [181] and its downregulation can lead to doxorubicin resistance in *in vitro* cell models [182]. In ovarian cancer patients, HuR nuclear localization during chemotherapy correlates with a good outcome, while its cytoplasmic localization increases paclitaxel resistance [183].

The action of HuR inhibitors on genomic instability has also been tested. The coumarin-like HuR inhibitor **3ns** (Table 3) showed reduction of cell viability in thyroid cancer cells, impairing HuR binding with the microtubule regulator MAD2 [184]. Quinolinium salt **17s** (Table 4) has been observed to improve chemotherapy efficacy in urothelial bladder carcinoma (UCB) by affecting HuR's ability to translocate into the cytoplasm [111]. To sum up, any possible beneficial effect of HuR silencing on genomic instability for cancer therapy needs to be carefully evaluated within the cancer context, and the possible combinations with other treatments, due to the multifaceted profile of HuR.

4.2. Enhancement of cell survival

Tumor cells are characterized by enhanced cell survival thanks to resisting recognition and attack by immune cells and apoptosis [185]. HuR is a potential coordinator of this pathologic behavior, by controlling antiapoptotic genes such as *Bcl2*, *Bcl-xL*, and *cIAP2* [3]. Even if HuR effects on cell survival were most likely both tissue- and cancer type-dependent, silencing HuR expression utilizing

HuR siRNA or shRNA significantly reduced tumor growth and inhibited tumor cells' survival *in vitro* and *in vivo* [137,184,186].

The therapeutic potential of targeting HuR for the treatment of ovarian [156] and pancreatic cancer [186] was evaluated in models stably expressing shHuR through a lentiviral gene transducing system. Ovarian OVCAR3 and MiaPaCa2 pancreatic ductal adenocarcinoma cancer cell lines stably expressed DOXO-inducible HuR-targeted shRNAs. The suppression of HuR expression reduced proliferation, anchorage-independent growth, and invasion of ovarian and pancreatic cancer cells *in vitro* [156,186]. MiaPaCa2 cells were also stably transfected with a tetracycline-responsive plasmid to overexpress HuR in response to DOXO treatment, without affecting cell proliferation [186].

Several small molecules showed inhibitory activity on HuR and cell survival, as observed with genetic silencing. Chrysanthone-like **1n** (Table 2) interferes with the multimerization of HuR and formation of HuR-mRNA complexes, leading to a reduction in cancer cell proliferation and survival in glioblastoma, PDA and colorectal cancer [145,167,187,188]. Treatment with **1n** in inflammatory bowel disease (IBD) and inflammatory colon cancer mice models seems to increase tumor size and invasiveness by counteracting eosinophil recruitment into tumors, probably leading to enhanced cell survival [13]; notably, **1n** did not cause any relevant toxicity in mice models. Compound **3ns** (Table 3) disrupts the interaction between HuR and ARE oligomers in pancreatic, colon, lung and thyroid cancer cells [184,189]. At genome wide level, compound **8n** (Table 2) dysregulates HuR by switching it to preferential binding to longer, ARE-enriched mRNAs at the 3' UTRs, including mRNAs that encode apoptotic and cell-cycle regulatory proteins. Interestingly, it inhibits colon cancer cell growth *in vivo* [1], and also reduces or disrupts HuR dimers in the cytoplasmic fraction of glioblastoma cells [145]; therefore, the exact, and possibly multifactorial mechanism of action of **8n** against cancer is still to be fully clarified. Microfilament inhibitors, such as macrocycle **13n** (Table 2) and fused pyridone **16s** (Table 4), mimic HuR silencing by interfering with HuR-dependent mRNA-stabilization and attenuating cytoplasmic HuR levels in hepatocellular carcinoma cells [137]. Thus, in conclusion, tumor cells may be dependent from the survival signals provided by HuR-regulated oncoproteins, thanks to HuR activity in the cytoplasm and its multimerization. This could be an adaptive survival mechanism of cancer cells under stress conditions, that can be interfered with by the ability of small molecules to inhibit the translocation of HuR in the cytoplasm, or to inhibit its dimerization.

4.3. Reprogrammed cell metabolism

So far, there is no clear evidence about any effect of small molecule HuR inhibitors on the regulation of cell metabolism. However, pancreatic tumor HuR-deficient cells were shown to be less resistant to glucose deprivation when compared to isogenic controls [190]. Changes in the cellular microenvironment, and in particular glucose deprivation, are a specific tumor signature of acute metabolic stress. HuR may activate a pro-survival signaling pathway in response to acute metabolic stress after translocation from the nucleus to the cytoplasm, where it regulates the expression of specific metabolic genes [190]. Three mRNAs that were selected for further validation (*GPI*, *IDH1*, and *PRPS2*) were observed as downregulated in HuR depleted pancreatic tumor cells.

In a similar manner, HuR was identified as the main regulator of the *IDH1* gene by deleting HuR in PDAC cells deprived of nutrients and performing RNA deep sequencing [191]. PDAC cells reprogrammed their metabolism in response to nutrient deprivation, and caused the development of resistance to chemotherapy, that induced an increase in oxidative stress. Namely, HuR regulation of *IDH1*, the NADPH-generating enzyme, activated a rapid antioxi-

dant response that enhanced PDAC cell survival, thus resulting in a potential therapeutic target for this cancer type.

HuR shows an important role in the regulation of metabolic pathways, and most importantly its inhibition could impair metabolic flexibility, which is the difference in the respiratory exchange ratio in fed and starved states. The conditional KO of HuR in skeletal muscle caused mild obesity in female but not male mice, due to impaired glucose tolerance, fat oxidation and palmitate oxidation. Thus, HuR is involved in the regulation of important skeletal muscle metabolism-related genes [58].

Although there is no evidence about the impact of small molecule HuR inhibition on the regulation of cell metabolism, targeting HuR could also affect cancer cell metabolism, and in particular their pro-survival response to glucose deprivation, as observed in HuR-deficient pancreatic cells.

4.4. Invasion and metastasis

HuR extensively influences invasion and metastasis, two key aspects of cancer progression. As a representative example for the relevance of HuR in the progression of these hallmarks, HuR was observed to stabilize both *Snail* mRNA, whose protein is involved in the epithelial mesenchymal transition (EMT), and metalloproteases, involved in the degradation of the extracellular matrix [192]. HuR is involved in peroxide-induced cellular migration of mammary carcinoma cells, and HuR KO reduces this migration ability. HuR is overexpressed in pre-neoplastic lesions of invasive breast cancer, such as atypical ductal hyperplasia (ADH) and ductal carcinoma *in situ* (DCIS), when compared to healthy controls, confirming its role in the progression of invasion [193]. In addition, high levels of cytoplasmic HuR significantly correlate with metastasis in bladder cancer [194]. In clinical studies, HuR is also associated with lymph node metastasis in non-small lung carcinoma [109].

Due to its strong influence on invasion and metastatic ability of cancer cells, HuR activity was impaired through different approaches, using anti-HuR siRNA and small molecule HuR inhibitors. MCF7 and MDA-MB-231 cells were transfected with a ribozyme transgene consisting of hammerhead ribozyme and HuR-specific antisense. The effect of HuR knockdown on human breast cancer cells reduced their growth and invasion and affected the expression of Cyclin D1 and MMP-9 [195]. Anti HuR siRNA conjugated with folate NPs (FNP) were designed and tested on human non-small-cell lung carcinoma (NSCLC), showing impairment in tumor cell migration [154]. *In vitro* scratch assays revealed the impairment of cellular migration also in an ovarian cancer cell model with altered HuR expression (OVCAR5-shHuRc257) [156].

Benzothioephene hydroxamate HuR inhibitor **11s** (Table 4) strongly reduced the invasiveness of breast cancer cells by inhibiting HuR-FOXQ1 interaction [150]. FOXQ1 is a transcription factor involved in breast cancer EMT, and its interaction with HuR contributes to cell invasion ability [150]. Finally, recently identified tanshinone mimic HuR inhibitors have been observed, through scratch assays, to remarkably reduce the migration ability of MDA and PANC-1 cells [142]. Tanshinone mimic **6ns** (Table 3), in addition to interfering with HuR-RNA interactions, efficiently blocked PANC-1 and MDA-MB-231 cell migration [142]. Tongue carcinoma HSC-3 and SAS cell lines were treated with polysulfonate **5s** (Table 4), and their ability to grow without adherence to the extracellular matrix and to neighbouring cells was investigated. Upon treatment with **5s**, attenuated motile and invasive properties were detected through *in vitro* wound healing and Matrigel invasion assays; the same happened in HuR knockdown HSC-3 cells [147].

Recently the role of HuR in malignant peripheral nerve sheath tumors (MPNST) was examined by us, finding a striking dependence of these cancer cells on HuR for their proliferation, survival and dissemination [171]. HuR was overexpressed in these highly aggressive sarcomas that originate in the peripheral nervous system, regulating the expression of numerous cancer-associated transcripts in human MPNST samples. A constitutive genetic inhibition of HuR in MPNST cells was sufficient to completely prevent tumor formation using xenograft models, whereas inducible ablation of HuR in pre-formed tumors led to their shrinkage. Remarkably, we found that HuR inhibition also prevented the formation and growth of metastatic nodules in lungs using a surrogate model of lung metastasis. This was particularly relevant since MPNSTs have a high metastatic potential, and up to 50% of patients develop metastatic disease, usually to the lung, which worsens the 5-year survival rates of patients [171]. Notably, pharmacological inhibition of HuR by chrysanthone-like **1n** (Table 2) could reduce tumor growth in xenograft models as well as metastatic growth in lungs, highlighting the relevance of HuR as a potential therapeutic target for MPNSTs. Finally, using experimental and computational approaches, it was ascertained that HuR exerts such profound effects on these cancer cells because of its capacity to simultaneously regulate multiple essential oncogenic pathways, which converged on key transcriptional networks [171].

4.5. Elevation of local angiogenesis

Tumor cells can promote vascular growth or angiogenesis through different mechanisms. Angiogenesis subsequently contributes to tumor growth and helps cancer cells enter the peripheral circulation [196].

Many clinical investigations have shown that HuR plays a critical role in promoting angiogenesis. In fact, cytoplasmic HuR accumulation stabilizes angiogenic factors, such as *VEGF-A*, interleukin-8 (*IL-8*), *HIF- α* , and *COX-2* [196–198]. Moreover, HuR acts by augmenting the hypoxic or inflammatory signal and promoting the proliferation and migration of endothelial cells [196,197].

In mesangial cells (hMCs) the vasoconstricting peptide angiotensin II (AngII) increases the capability of HuR to bind and stabilize *COX-2* mRNA. *COX-2* mRNA is a specific target of cytoplasmic HuR, and AngII stimulation is necessary for their interaction. Indeed, attenuated HuR expression mediated by siRNA affected the expression and function of *COX-2* in cells after AngII treatment [118].

HIF-1 is a master regulator of tumor neovascularization. The heat shock transcription factor 1 (Hsf1) is overexpressed in a variety of tumors and plays a critical role in tumor progression by regulating numerous genes, including HuR. Upon Hsf1 knockdown, HuR levels decreased by about 70% in MCF7 and Hs578 cells, MCF7 xenografts and in Hsf1-KO mice, consequently affecting HIF-1 expression [199].

In myeloid cells HuR stabilizes different genes, including angiogenic regulators bearing ARE or U-rich sequences at their 3' UTR region. *VEGF-A* is a key angiogenic growth factor directly regulated by HuR and miR-200b in an antagonistic manner. After myeloid-specific deletion of HuR in mice (Elavl1M \emptyset KO), bone marrow-derived macrophages presented a robust *VEGF-A* [68], *VEGF* and *MMP-9* [200] downregulation. Furthermore, tumor slices from Lewis lung carcinoma (LLC) cells injected into Elavl1M \emptyset KO mice displayed reduced vascular density and permeability, sprouting and branching, and contained vessels with dilated lumens. The same phenotype also occurred in zebrafish embryos through the injection of morpholino (MO)-targeted HuR, resulting in severe defects in subintestinal vein (SIV) vascular development [68]. Therefore, macrophages and monocytes are the primary sources

of the inflammatory angiogenic process, and macrophage-specific HuR inhibition impairs the posttranscriptional mechanisms of angiogenic regulator genes [68,200].

Among HuR small molecule inhibitors, natural tanshinone **9n** (Table 2) is endowed with anti-tumor and anti-angiogenic activity both *in vitro* and *in vivo*. *TNF α* , besides being an important mediator of inflammation, also has proangiogenic effects. Compound **9n** has been observed to markedly reduce lipopolysaccharide (LPS)-induced *TNF α* mRNA expression in a dose dependent manner by increasing nuclear HuR localization, thus impairing HuR cytoplasmic translocation and, consequently, *TNF α* mRNA stabilization [132].

Tanshinone mimic **6ns** (Table 3) was tested for its capability to interfere with HuR-RNA binding in MCF7 cells. Compound **6ns** treatment revealed a decrease of *VEGF* mRNA copies and a reduction in its expression levels [142].

As previously mentioned, HuR interacts with *COX-2* mRNA, stabilizing its expression [118,196]. Moreover, chrysanthone-like **1n** (Table 2) was found to interfere with HuR cytoplasmic localization and to impair its binding activity [123]. In Cajal-Retzius (CR) cells and in mice bearing HCT116 and HCA7 cell xenografts, **1n** inhibited HuR and decreased *COX-2* expression in a dose dependent manner. Finally, HuR targeting by **1n** altered *COX-2* expression and decreased angiogenesis both *in vitro* and *in vivo* [167].

HuR is associated with diabetic retinopathy (DR) [78], and is abundant in human retinal endothelial cells (HRECs) [151]. The aetiology of diabetic retinopathy is the onset of a diabetic macular oedema that causes retinal detachment and visual loss. Indoles **12s** and **13s** (Table 4) were selected as HuR inhibitors, presenting anti-inflammatory and anti-angiogenic properties. On the basis of MD experiments, these small molecules blocked HuR functions by competing with its mRNA binding site and modifying the structure of the HuR binding cleft. In addition to HuR impairment, *TNF α* and *VEGFA* expression were significantly decreased in HRECs treated with **12s** and **13s**, in high glucose media cultured conditions [151].

A matrigel tubule formation assay revealed Indoline sulfonamide **15s** (Table 4), and indoles **12s** and **13s** (Table 4) [151] as antiangiogenic HuR inhibitors, because their effects inhibited VEGF-induced migration and tube formation ability of human umbilical vein endothelial cells (HUVEC) [153] and of HRECs [151].

Eltrombopag is an approved oral drug used in clinics for the treatment of thrombocytopenia, severe aplastic anaemia as an agonist of the thrombopoietin receptor [201]. It is effective in disrupting the interaction between HuR and the ARE sequence of *VEGFA* mRNA. Furthermore, eltrombopag is a good anti-angiogenic drug, as demonstrated by its *in vivo* reduction of microvessels in tumor tissue [202].

4.6. Promotion of cell proliferation and tumorigenesis

The CRISPR/Cas9 technology was used to define the role of HuR in PDA and colorectal cancer cells (CRCs) [175]. MiaPaca2, Hs766T, HCT116 and HCA7 cells lacking HuR displayed attenuated growth compared to the control and the HuR overexpression in HuR-null MiaPaca2 cells restored a tumor growth phenotype. According to HuR function in promoting tumor cell growth, MIA.HuR-KO^(-/-) and HsT.HuR-KO^(-/-) cells were unable to grow into 3D cultures. Namely, cells with CRISPR/Cas9-mediated deletion of HuR only formed single cell sheets, and were unable to form substantial spheroids even after four weeks of plating, while positive control cells easily formed spheroids in 3D cultures within a passage, continuing to grow and doubling in size [175].

As mentioned earlier, HuR was knocked down in four MPNST cell lines by lentiviral delivery of shRNAs [171]. HuR is highly aggressive in sarcomas originated from Schwann cells, and its downregulation produced a dramatic decrease of percentage in

BrdU positive cells and in the ability of single cells to form colonies. After HuR silencing in MPNST cells, an analysis of β -galactosidase staining revealed a marked increase in cellular senescence. Therefore, HuR genetic inhibition has cytotoxic and cytostatic effects on tumor growth [171].

In normal intestinal epithelium, HuR has a predominant nuclear localization, while in adenocarcinomas it is mostly relocalized in the cytoplasm, where it is aberrantly regulating the stability of key pro-oncogenes responsible for cell proliferation, such as *cyclinA*, *cyclinB1* and *cFos*.

HuR-siRNA encapsulated in lipid NPs was delivered in human melanoma cell lines and by the Trypan blue exclusion assay method, cellular viability was assessed. HuR-NP treatment in combination with the kinase inhibitor UO126, generated a significant inhibitory effect in a human melanoma cell line overexpressing the microphthalmia-associated transcription factor (MITF) [203].

Several colorectal cancer cell lines overexpressing HuR (HCT116, HCA-7, RKO, HT-29, and SW480 cells) showed growth inhibition after treatment with chrysantone-like **1n** (Table 2) at different concentrations (1–100 μ M), whereas only a weak effect was observed on normal epithelial cells [167]. Further studies highlighted the impact of **1n** on HuR-mediated colorectal tumorigenesis, discriminating among cells from different patient subtypes. Targeting HuR was used to prevent development of cancer in high-risk groups, such as those with familial adenomatous polyposis (FAP) or IBD. The level of HuR expression and localization was different in each condition, and to better address the question of how HuR expression is involved in each stage of tumor progression, azoxymethane/dextran sodium sulfate (AOM/DSS) and adenomatous polyposis coli (APC)^{Mim} mice models were used [13]. The *c-myc* gene, a known HuR target, was chosen as a biomarker to verify the functional inhibition of HuR upon **1n** treatment in intestinal tissue, due to its involvement in colorectal cancer development. As expected, both *c-myc* expression and the rate of tumorigenesis were reduced, as confirmed by the decrease of Ki67 positive cells among small intestinal crypt cells derived from the aforementioned mouse models [13].

Chrysantone-like **1n** was also tested against malignant glioma cells [188]. Treatment with **1n** (20 μ M and 30 μ M) strongly impaired the invasion rate of JX12 cells, and their CD33⁺ subpopulation was even more sensible to **1n**, as shown by an increased inhibition of their invasivity at 10 μ M. To assess the functional phenotype of brain tumor initiating cells, the formation of neurospheres was quantified by a limiting dilution assay. At 2.5 μ M, **1n** attenuated the initiating stem cell frequency in forming neurospheres [188].

Compound **3ns** (Table 3) was tested for its potential antitumor activity in human lung cancer cells [189]. Several NSCLC cell lines (H1299, A549, HCC827, H1975) were treated for 24 and 48 h with **3ns** (20 μ M and 30 μ M), and a consistent, dose dependent inhibition of tumor proliferation and induction of G1 cell cycle arrest was consistently observed. Conversely, limited or no effects were observed on normal epithelial cells.

The antitumor efficacy of **3ns** was evaluated in four thyroid cancer cell lines (SW1736, 8505C, BCPAP and K1) [184]. Through a scratch assay, treatment with **3ns** (35 μ M) reduced their invasion ability and impaired colony formation using an anchorage-independent assay in soft agar. This effect on colony formation and proliferation was ascribed to downregulation of mitochondrial associated protein MAD2, a HuR target overexpressed in thyroid cancer cells. MAD2 is involved in the regulation of cell division, in particular in the metaphase to anaphase transition, and treatment with **3ns** downregulated the expression of this gene via HuR inhibition, thus blocking tumorigenesis. This was confirmed by siRNA silencing of MAD2 and through a rescue experiment in which MAD2 was overexpressed after treatment with **3ns** [184].

Compound **3ns** was also tested in combination with the YAP inhibitor verteporfin (VP), and the CA3 and CDK4/6 inhibitor abemaciclib in PDAC cells. The combination of abemaciclib with **3ns** decreased the number of PDAC colonies compared with both monotherapies, most likely due to their shared regulatory role of the cyclinD1 pathway [204].

Finally, HCT116 and HCT116 HuR-KO colorectal carcinoma cells were grown under anchorage- and serum-independent conditions. Cells lacking HuR were unable to form tumor spheroids and, after treatment with natural tanshinone **8n** (Table 2), HCT116 spheroids were significantly reduced [175].

According to these evidence, HuR modulators seem to be able to decrease cancer cell tumorigenesis at early stages, as evaluated by colony formation or through spheroids assays. Nevertheless, further data are needed to fully validate this hypothesis.

4.7. Resisting apoptosis

Resistance to apoptosis is a key event in tumor development. HuR finely regulates the balance between cell survival and cell death by caspase-mediated apoptosis in response to lethal stress. Indeed, in normal conditions, HuR promotes cell survival by stabilizing and increasing the translation of mRNAs coding for antiapoptotic factors [18], whereas under such lethal stress HuR promotes apoptosis by increasing the expression of proapoptotic proteins [205].

Several studies have demonstrated that inhibition of HuR, by using either genetic approaches (gene deletion and siRNA) or small molecule inhibitors, promotes apoptosis [59,66,67]. For example, HuR acts in B cells as a key factor for a proper metabolic switch and cell growth during B cell maturation. In a B-cell precursor lineage, HuR deletion led to the induction of apoptosis [66,67]. Moreover, annexin V staining of siHuR MiaPaCa2 cells exposed to death receptor 5 (DR5) showed an increase in the apoptotic signal [206].

Among several small molecule HuR inhibitors, natural tanshinone **8n** (Table 2) showed anti-tumor effects in different human breast [207] and colon cancer cell lines [126] by inducing apoptosis. Induction of apoptosis was also observed with compound **1n** (Table 2) in colon cancer [167], in glioblastoma cells [188] and in MPNST cells [171]. Finally, apoptosis induction was observed in thyroid cancer cells treated with coumarin-like **3ns** (Table 3) [184].

Interestingly, microtubule inhibitor **15s** (Table 4) led to a concentration-dependent increment of TUNEL-positive A549 cells population, by inhibiting the expression of *HIF-1 α* through reduced translocation of HuR to the cytoplasm [153]. Finally, in UCB, pyruvium pamoate **17s** (Table 4), in combination with genotoxic agents, increased apoptosis by triggering DNA damage in 5637 cells [111].

4.8. Evasion of immune recognition and tumor invasion promotion

Cancer onset and progression are strongly determined by tumor capacity of evading the immune response, with subsequent promotion of inflammation [121]. The tumor microenvironment is preserved by either inflammatory, stroma and tumor cells, often exploiting signalling molecules, such as cytokines, chemokines, to promote invasion, migration and metastasis [91]. Considering that \approx 90% of mRNAs coding cytokines and chemokines contain repeated ARE sequences in their 3' UTR structure, HuR probably regulates these unstable transcripts in competition with stabilizing and destabilizing *trans*-factors, such as other RBPs like tristetraprolin (TTP) and T cell-restricted intracellular antigen 1 (TIA1). This gives rise to a rapid degradation and turnover of these mRNAs in response to changes in cells' or tissues' environments. Consequently, HuR plays an important role in innate, adaptive immunity and inflammatory pathways, in physiological and pathological

condition [67,69,169,208,209]. Indeed, cancer-driven immune escape still represents an issue for anticancer therapy [210]. In this context, the role of HuR remains controversial, as indeed its deletion or inhibition leads to different outcomes. Co-culture of MCF-7 tumor-spheroids with primary human CD14⁺ monocytes attracted and retained macrophages in the 3D tumor spheroids [211]. Tumor progression is strongly affected by infiltrating immune cells, and environmental changes may affect the activity of HuR. In fact, when HuR was knocked down by shRNAs, the expression level of C-C motif chemokine ligand 5 (CCL5) increased concomitantly to the infiltration of macrophages in the tumor spheroid. Moreover, a model of 3D breast cancer showed a decrease in size upon HuR depletion, supporting a role for HuR in enhancing cancer proliferation [211].

Recently, in glioblastoma, HuR deletion has been associated with a decrease of tumor growth and proliferation, as in particular a reduction in the number of tumor associated macrophages (TAMs), showing M1-like increased polarization. In fact, HuR KO mice were characterized by substantial changes for key parameters determining cancer progression, such as migratory and chemoattractive capabilities, with substantial rearrangements of chemokine and cytokine production, modifying microenvironment conditions and reducing tumor growth [72]. This demonstrated the importance of HuR as a valuable target for therapy in this field. To this purpose, chrysanthone-like **1n** (Table 2), whose activity has been previously described to downregulate different mRNA cytokines in cellular contexts such as GBM xeno-lines [72,123,167], has been tested *in vivo* in AOM/DSS mice as earlier described. However, in this model the treatment with **1n** induced an attenuation of eosinophils associated with tumor and a decrease of pro-inflammatory molecules causing an exacerbation of tumor development and invasiveness, therefore worsening the outcome of the treatment. On the other side, using **1n** in an APC^{Min} model of FAP and colon cancer partially ameliorated their carcinogenic conditions [13]. These results are in line with genetic models in which the ablation of HuR in the colon tissue protects from tumor burden [59], while myeloid cell-specific deletion of HuR exacerbated chemically induced colitis [64].

In other words, small molecule HuR inhibitors are validated agents for immune-restoring therapy, although the complexity of the HuR regulatory functions presented so far must be kept in mind.

In conclusion, after having substantiated the claim of HuR influencing the whole Hallmarks of Cancer panel, we must also say that its emerging role in inflammatory processes and diseases is of significant relevance as well, also in terms of its pharmacological-small molecule targeting. Thus, in the next Chapter we provide a somewhat detailed overview of this fast growing field, commenting upon multiple and sometimes contrasting opinions.

5. A controversial role for HuR in immunity

5.1. HuR in immunity

HuR determines the development of a pro-inflammatory response to agents such as LPS, since it prevents degradation of toll-like receptor 4 (TLR4) mRNA, giving rise to an upregulation of inflammation processes in models of vascular inflammation and atherogenesis [212]. Furthermore, HuR stabilizes several inducible transcripts, including *interferon- γ* (*IFN- γ*), *TNF- α* , *IL-6*; *IL-8*, *IL-3*, *IL-1 β* and the *urokinase-type plasminogen activator* (*uPA*), that are key mediators of the inflammatory and immune responses [166,213–219]. CX3CL1/fractalkine is a chemokine ligand specific for natural killer (NK) cells and monocytes maturation. It contains ARE elements in its 3' UTR and can be post-transcriptionally regu-

lated by HuR. This suggests that HuR can modulate the development of two among the major players in the innate immunity system [220]. Recently, sequential photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) experiments on normal and HuR KO bone-marrow derived macrophages (BMDMs) showed that the expression of lineage specific genes involved in vascular development and angiogenesis is post-transcriptionally determined by an intricate interplay between HuR and various mRNAs [221]. Moreover, similar PAR-CLIP experiments were carried out in primary macrophages under LPS stimuli, proving the existence of a complex post-transcriptional landscape driven by balancing activities of several RBPs, in particular HuR and TTP [222]. In fact, groups of transcripts bound and modulated exclusively by TTP or HuR were identified; another group, including mRNAs for *TNF α* or *CXCL2*, can interact and bind simultaneously to both TTP and HuR, establishing a competition between stabilizing and destabilizing effects that is reflected in the insurgence of a vulnerable and tightly regulated post-transcriptional pattern influenced by both TTP and HuR [222,223]. Moreover, another molecular mechanism has been proposed to explain the HuR-mediated increase of different cytokines and chemokines (e.g. *CXCL2*) in macrophages, due to stabilization of their mRNAs. Such mechanism entails a PTM carried out by poly(ADP-ribose)polymerase 1 (PARP1) to HuR, known as a PAR-ylation. This modification of HuR influences its shuttling to the cytoplasm, and increases its interaction with different pro-inflammatory mRNAs [120]. Moreover, in mammalian macrophages, HuR behaves as a mRNAs sponge and derepresses inflammatory agents by counteracting the induction of anti-inflammatory response driven by infection of pathogens, such as *Leishmania donovani* [224].

Despite these evidences, the role of HuR in immunity remains controversial. For example, HuR deficiency in macrophages derived from mice lacking HuR in the myeloid-lineage cells, has been linked with an increase of pro-inflammatory cytokines. HuR KO cells presented an increased CCR2-mediated chemotaxis and enhancements in the expression of inflammatory mRNAs (including *Tnf*, *Tgfb*, *Il10*, *Ccr2*, and *Ccl2*), due to mis-regulation in their translational and stability levels. This caused an increased susceptibility of these mice to colitis-associated cancer [69]. Furthermore, experiments performed in a co-culture model of primary human macrophages and MCF7 breast cancer cell lines, demonstrated that HuR suppresses the production of the leukocyte attracting chemokine *CCL5*, reducing the infiltration of pathogenic macrophages in the tumor site, and preventing the exacerbation of inflammation processes [211]. This confirms for HuR a regulatory role for both pro-inflammatory and anti-inflammatory agents, with contrasting, strongly cell lineage-dependent effects whose overall results are difficult to be precisely predicted and determined. Accordingly, exploiting intestinal inflammation mice models in which HuR has been deleted in Intestinal epithelial cells (IECs) and in myeloid-derived immune compartments, HuR KO increased in cell death, confirming that HuR has a role in the maintenance of the intestinal barrier integrity and homeostasis. In parallel, though, the lack of HuR in myeloid lineages fueled the inflammation process, thus exacerbating a pathological condition [225]. Conversely, a conditional HuR KO mice model confirmed that HuR is pivotal for the maintenance of hematopoietic stem cells during hematopoiesis, but also its essential role for B cell development and a homeostatic balance between T and B cells [46,67,71,226]. In particular, using a Cre-LoxP system, HuR deletion in thymocyte development resulted in a loss of peripheral T cells, highlighting its importance in controlling thymocyte maturation and trafficking [71]. Nevertheless, in T cells HuR modulates maturation and polarization of Th2 and Th17 cell lineages, enhancing the stability of IL-2 and IL-17 by binding to their 3' UTRs [64,65].

5.2. HuR targeting in inflammatory diseases

In a pathologic context, HuR deletion in distal lung epithelium decreased neutrophilia and pulmonary inflammation levels induced by IL-17, through increasing mRNA decay of chemokines such as CXCL1 and CXCL5 [227]. Th17 cells are the major mediators of the generation of inflammatory infiltrations in the central nervous system (CNS) during neuroinflammation. Functional studies and HuR conditional KO in CD4⁺ T cells of a mouse strain, determined that in an experimental autoimmune encephalomyelitis (EAE) model, HuR binds directly to the 3' UTR of the C-C chemokine receptor 6 (CCR6). CCR6 is a surface receptor on Th17 cells, and is responsive to migration stimuli exerted by the release of chemokine ligand 20 (CCL20), which is constitutively secreted by choroid-plexus epithelial cells at the site of inflammation. A complete ablation of HuR significantly decreases the production of CCR6, thus lowering the number of migrating Th17 cells, ameliorating the pathogenic neuroinflammation processes in the CNS, and the EAE outcome in mice [228].

Consequently, there is emerging evidence concerning the targeting of HuR with small molecule inhibitors in order to ameliorate disease onset and reduce autoimmune inflammation [63]. In fact, administration of natural tanshinone **8n** (Table 2) in EAE mice models led to a reduction of the aggressiveness, and to a delayed onset of the MS-like pathology. Compound **8n**, injected at 10 mg/kg every 48 h from day 5 to day 15 of disease induction, caused a decreasing number of infiltrates in CNS and lower demyelination. Furthermore, through flow cytometry assays the number of CD4⁺ T cells producing IFN- γ and IL17 in mice spleen was significantly reduced, as was the number of CD11b⁺ myeloid cells present in their spinal cords. Moreover, signals of pro-inflammatory mRNAs (e.g. *IL17*, *IFN- γ*) and adhesion factors (*Vla-4*) were lowered in CNS after treatment with **8n**. Nevertheless, no alteration of the balance between CD4, CD8 T cells, B cells and macrophages was observed in the spleen of treated mice compared with control, suggesting that compound **8n** did not cause systemic toxicity *in vivo* [63]. Lastly, a HuR targeting strategy has been applied to reduce inflammatory contributions during the progression of chronic kidney disease (CKD), that is characterized by sustained inflammation and fibrosis development, leading to final end-stage renal disease (ESRD). In particular, benzothioephene hydroxamate HuR inhibitor **11s** (Table 4) was tested at 50 mg/kg daily for 5 days in an experimental anti-Thy 1.1 nephritis rat model, characterized by high level of glomerular HuR. Injections of **11s** showed no side effects (e.g. peritonitis insurgence), but caused the lowering of urea levels in the serum and of protein content in urines when compared with untreated rats. Moreover, the analysis of histological renal sections staining showed a compound **11s**-dependent reduction of glomerulosclerosis, followed by a reduction in mRNA expression and protein production of profibrotic markers such as TGF β 1, plasminogen activation inhibitor 1 (PAI-1) and fibronectin. Finally, **11s** treatments decreased the number of monocytes and macrophages invading glomerular sites, and stimulated M2 macrophage activation and renal production of NF- κ B-p65, that promotes HuR transcription during tumor development [229] and induces glomerular HuR transcription and shuttling to the cytoplasm [230].

6. Conclusion

In this review, we focused mostly on the therapeutical potential of HuR inhibition in cancer. Considering that overexpression of HuR, or its mis-localization along with accumulation either in the nucleus or in the cytoplasm correlates with tumor development and progression, HuR modulators – in particular small molecule inhibitors – have been repeatedly used in multiple *in vitro* and

in vivo models, providing promising preliminary results against cancerogenic traits also known as the hallmarks and enabling characteristics of cancer (Fig. 6). Thus, they represent a suitable and prospective option in cancer therapy.

Although their effects in multiple cancer cellular contexts were described, their *in vivo* evaluation remains limited, and should significantly increase in the next years. Being HuR ubiquitously expressed, and regulating a variety of different key mRNAs, the development of suitable, HuR-centered *in vivo* models is still controversial. Indeed, HuR KO in complex systems usually leads to the exacerbation of the disease [224], and similar results were obtained when treating with chrysantone-like **1n** (Table 2) in similar models [13], suggesting that targeting the pleiotropic functions of HuR could arise some undesired effects.

Another concern, mentioned in Paragraph 1.5, regards the specificity of these compounds versus other members of the ELAVL family of proteins, with shared structural properties with respect to HuR. Indeed, available data in this regard are not yet exhaustive for HuR inhibitors tested either *in vitro* or *in vivo*. Nevertheless, the scenario described in this review is strongly supporting the rational design, synthesis and structural optimization of synthetic HuR inhibitors, to overcome the limitations of current, mostly naturally occurring HuR modulators in terms of bioavailability (e.g. solubility) and specificity. When properly optimized for HuR selectivity, safety and *in vivo* efficacy, such small molecule HuR inhibitors could on one side minimize the possible insurgence of side effects, and on the other side become potent and bioavailable enough to foresee the identification of one or more HuR-targeted clinical candidate as anticancer agents in the next years.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Associazione Italiana Ricerca sul Cancro (IG grant # 21548 to A.P.), by AEI/FEDER, EU (RTI2018-097503-B-I00, 2020-119486RB-100 and SAF2015-62588-ERC to A.W.), Severo Ochoa Excellence Accreditation (SEV-2016-0644 to A.W.), Fundación BBVA (Becas Leonardo to A.W. and M.V.R.), Gilead Sciences International Research Scholars Program in Liver Disease to M.V.R., the European Community's H2020 Framework Program ERC Consolidator Grant (865157-MYERIBO to A.W.), "RNAct" Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITN2018 (contract N. 813239), MIUR-PRIN 2017 (Grant PHRC8X to L.M. and V.L.P.) and Regione Campania-POR Campania FESR 2014/2020 (Project N. B61G18000470007 to L.M. and V.L.P.).

References

- [1] P. Lal, L. Cerofolini, V.G. D'Agostino, C. Zucal, C. Fuccio, I. Bonomo, E. Dassi, S. Giuntini, D. Di Maio, V. Vishwakarma, R. Preet, S.N. Williams, M.S. Fairlamb, R. Munk, E. Lehrmann, K. Abdelmohsen, S.R. Elezgarai, C. Luchinat, E. Novellino, A. Quattrone, E. Biasini, L. Manzoni, M. Gorospe, D.A. Dixon, P. Seneci, L. Marinelli, M. Fragai, A. Provenzani, Regulation of HuR structure and function by dihydrotanshinone-I, *Nucleic Acids Res.* 45 (2017) 9514–9527.
- [2] W.-J. Ma, S. Cheng, C. Campbell, A. Wright, H. Furneaux, Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein, *J. Biol. Chem.* 271 (14) (1996) 8144–8151.
- [3] H. Wang, F. Zeng, Q. Liu, H. Liu, Z. Liu, L. Niu, M. Teng, X. Li, The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding, *Acta Crystall Section D*, D 69 (2013) 373–380.
- [4] M. Pabis, G.M. Popowicz, R. Stehle, D. Fernandez-Ramos, S. Asami, L. Warner, S.M. Garcia-Mauriño, A. Schlundt, M.L. Martinez-Chantar, I. Diaz-Moreno, M.

- Sattler, HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs, *Nucleic Acids Res.* 47 (2019) 1011–1029.
- [5] M.-L. Samson, Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins, *BMC Genomics* 9 (2008) 392.
- [6] H.J. Okano, R.B. Darnell, A hierarchy of Hu RNA binding proteins in developing and adult neurons, *J. Neurosci.* 17 (9) (1997) 3024–3037.
- [7] Y.H. Tang, S.P. Han, K.S. Kassahn, A. Skarszewski, J.A. Rothnagel, R. Smith, 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.* 75 (2012) 214–228.
- [8] P.J. Good, A conserved family of elav-like genes in vertebrates, *PNAS* 92 (10) (1995) 4557–4561.
- [9] K. Kasashima, K. Terashima, K. Yamamoto, E. Sakashita, H. Sakamoto, Cytoplasmic localization is required for the mammalian ELAV-like protein HuD to induce neuronal differentiation, *Gen. Cells* 4 (11) (1999) 667–683.
- [10] C. Zucal, P. Lal, V.G. D'Agostino, R. Loffredo, B. Mantelli, N. Thongon, P. Lal, E. Latorre, A. Provenzani, Targeting the multifaceted HuR protein, benefits and caveats, *Curr. Drug Targets* 16 (2015) 1–17.
- [11] W. Akamatsu, H. Fujihara, T. Mitsushashi, M. Yano, S. Shibata, Y. Hayakawa, H. J. Okano, S.-i. Sakakibara, H. Takano, T. Takano, T. Takahashi, T. Noda, H. Okano, The RNA-binding protein HuD regulates neuronal cell identity and maturation, *PNAS* 102 (12) (2005) 4625–4630.
- [12] V. Katsanou, S. Milatos, A. Yiakouvakli, N. Sgantzis, A. Kotsoni, M. Alexiou, V. Harokopos, V. Aidinis, M. Hemberger, D.L. Kontoyiannis, The RNA-binding protein Elavl1/HuR is essential for placental branching morphogenesis and embryonic development, *Mol. Cell. Biol.* 29 (10) (2009) 2762–2776.
- [13] M. Lang, D. Berry, K. Passecker, I. Mesteri, S. Bhujii, F. Ebner, V. Sedlyarov, R. Evstatiev, K. Dammann, A. Loy, O. Kuzyk, P. Kovarik, V. Khare, M. Beibel, G. Roma, N. Meisner-Kober, C. Gasche, HuR small-molecule inhibitor elicits differential effects in adenomatous polyposis and colorectal carcinogenesis, *Cancer Res.* 77 (9) (2017) 2424–2438.
- [14] J.M. Izquierdo, Hu Antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition, *J. Biol. Chem.* 283 (27) (2008) 19077–19084.
- [15] S. Lebedeva, M. Jens, K. Theil, B. Schwanhäusser, M. Selbach, M. Landthaler, N. Rajewsky, Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR, *Mol. Cell* 43 (3) (2011) 340–352.
- [16] M. Dutertre, F.Z. Chakrama, E. Combe, F.-O. Desmet, H. Mortada, M. Polay Espinoza, L. Grataudou, D. Auboeuf, A recently evolved class of alternative 3'-terminal exons involved in cell cycle regulation by topoisomerase inhibitors, *Nat. Commun.* 5 (2014) 3395.
- [17] T. Bakheet, E. Hitti, M. Al-Saif, W.N. Moghrabi, K.S.A. Khabar, The AU-rich element landscape across human transcriptome reveals a large proportion in introns and regulation by ELAVL1/HuR, *Biochim. Biophys. Acta* 1861 (2) (2018) 167–177.
- [18] K. Abdelmohsen, R. Pullmann, A. Lal, H.H. Kim, S. Galban, X. Yang, J.D. Blethrow, M. Walker, J. Shubert, D.A. Gillespie, H. Furneaux, M. Gorospe, Phosphorylation of HuR by Chk2 regulates SIRT1 expression, *Mol. Cell* 25 (4) (2007) 543–557.
- [19] M.-J. Kang, B.-K. Ryu, M.-G. Lee, J. Han, J.-H. Lee, T.-K. Ha, D.-S. Byun, K.-S. Chae, B.-H. Lee, H. S. Chun, K.Y. Lee, H.-J. Kim, S.-G. Chi, NF- κ B activates transcription of the RNA-binding factor HuR, via PI3K-AKT signaling, to promote gastric tumorigenesis, *Gastroenterology* 135 (2008) 2030–2042.
- [20] Selvi C. Jeyaraj, Mamata Singh, Dina A. Ayupova, Suman Govindaraju, Beth S. Lee, Transcriptional control of Human antigen R by bone morphogenetic protein, *J. Biol. Chem.* 285 (7) (2010) 4432–4440.
- [21] Rudolf Pullmann, Hyeon Ho Kim, Kotb Abdelmohsen, Ashish Lal, Jennifer L. Martindale, Xiaoling Yang, Myriam Gorospe, Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs, *Mol. Cell. Biol.* 27 (18) (2007) 6265–6278.
- [22] W. Al-Ahmadi, M. Al-Ghamdi, L. Al-Haj, M. Al-Saif, K.S.A. Khabar, Alternative polyadenylation variants of the RNA binding protein, HuR: abundance, role of AU-rich elements and auto-regulation, *Nucleic Acids Res.* 37 (2009) 3612–3624.
- [23] Na Chang, Jie Yi, Gaier Guo, Xinwen Liu, Yongfeng Shang, Tanjun Tong, Qinghua Cui, Ming Zhan, Myriam Gorospe, Wengong Wang, HuR uses AUF1 as a cofactor to promote p16INK4 mRNA decay, *Mol. Cell. Biol.* 30 (15) (2010) 3875–3886.
- [24] Xun Guo, Yuehan Wu, Rebecca Hartley, MicroRNA-125a represses cell growth by targeting HuR in breast cancer, *RNA Biol.* 6 (5) (2009) 575–583.
- [25] Bernard S. Marasa, Subramanya Srikantan, Jennifer L. Martindale, Mihee M. Kim, Eun Kyung Lee, Myriam Gorospe, Kotb Abdelmohsen, MicroRNA profiling in human diploid fibroblasts uncovers miR-519 role in replicative senescence, *Aging* 2 (6) (2010) 333–343.
- [26] A. Doller, A. Huwiler, R. Müller, H.H. Radeke, J. Pfeilschifter, W. Eberhardt, Protein kinase C alpha-dependent phosphorylation of the mRNA-stabilizing factor HuR: implications for posttranscriptional regulation of cyclooxygenase-2, *Mol. Biol. Cell* 18 (2007) 2137–2148.
- [27] Po-Chen Chu, Hsiao-Ching Chuang, Samuel K. Kulp, Ching-Shih Chen, The mRNA-stabilizing factor HuR protein is targeted by β -TrCP protein for degradation in response to glycolysis inhibition, *J. Biol. Chem.* 287 (52) (2012) 43639–43650.
- [28] T.-X. Yu, P.-Y. Wang, J.N. Rao, T. Zou, L. Liu, L. Xiao, M. Gorospe, J.-Y. Wang, Chk2-dependent HuR phosphorylation regulates occludin mRNA translation and epithelial barrier function, *Nucleic Acids Res.* 39 (2011) 8472–8487.
- [29] Sebastian Schulz, Anke Doller, Nicole R. Pardini, Jacqueline A. Wilce, Josef Pfeilschifter, Wolfgang Eberhardt, Domain-specific phosphomimetic mutation allows dissection of different protein kinase C (PKC) isotype-triggered activities of the RNA binding protein HuR, *Cell. Signalling* 25 (12) (2013) 2485–2495.
- [30] Y. Akaike, K. Masuda, Y. Kuwano, K. Nishida, K. Kajita, K. Kurokawa, Y. Satake, K. Shoda, I. Imoto, K. Rokutana, HuR regulates alternative splicing of the TRA2 gene in human colon cancer cells under oxidative stress, *Mol. Cell. Biol.* 34 (2014) 2857–2873.
- [31] L. Pang, H. Tian, N. Chang, J. Yi, L. Xue, B. Jiang, M. Gorospe, X. Zhang, W. Wang, Loss of CARM1 is linked to reduced HuR function in replicative senescence, *BMC Mol. Biol.* 14 (2013) 15.
- [32] Nieves Embade, David Fernández-Ramos, Marta Varela-Rey, Naiara Beraza, Marcella Sini, Virginia Gutiérrez de Juan, Ashwin Woodhoo, Nuria Martínez-López, Begoña Rodríguez-Iruretagoyena, Francisco Javier Bustamante, Ana Belén de la Hoz, Arkaitz Carracedo, Dimitris P. Xirodimas, Manuel S. Rodríguez, Shelly C. Lu, José M. Mato, María L. Martínez-Chantar, Murine Double Minute 2 Regulates Hu antigen R stability in human liver and colon cancer through NEDDylation, *Hepatology* 55 (4) (2012) 1237–1248.
- [33] Y. Ke, Y. Han, X. Guo, J. Wen, K. Wang, X. Jiang, X. Tian, X. Ba, I. Boldogh, X. Zeng, PARP1 promotes gene expression at the post-transcriptional level by modulating the RNA-binding protein HuR, *Nat. Commun.* 8 (2017) 14632.
- [34] Yuehuang Ke, Xueping Lv, Xingyue Fu, Jing Zhang, Ameer Ali Bohio, Xianlu Zeng, Wenjing Hao, Ruoxi Wang, Istvan Boldogh, Xueqing Ba, Poly(ADP-ribose)ylation enhances HuR oligomerization and contributes to pro-inflammatory gene mRNA stabilization, *Cell. Mol. Life Sci.* 78 (4) (2021) 1817–1835.
- [35] Sofia-Iris Bibli, Jiong Hu, Fragiska Sigala, Ilka Wittig, Juliana Heidler, Sven Zukunft, Diamantis I. Tsilimigras, Voahanginirina Randriamboavonjy, Janina Wittig, Baktybek Kojonazarov, Christoph Schürmann, Mauro Siragusa, Daniel Siuda, Bert Luck, Randa Abdel Malik, Konstantinos A. Filis, George Zografos, Chen Chen, Dao Wen Wang, Josef Pfeilschifter, Ralf P. Brandes, Csaba Szabo, Andreas Papapetropoulos, Ingrid Fleming, Cystathionine γ lyase sulfhydrylates the RNA binding protein human antigen R to preserve endothelial cell function and delay atherosclerosis, *Circulation* 139 (11) (2019) 101–114.
- [36] I. Grammatikas, K. Abdelmohsen, M. Gorospe, Posttranslational control of HuR function, *Wiley Interdiscip. Rev. RNA* 8 (2017) e1372.
- [37] Kotb Abdelmohsen, Subramanya Srikantan, Xiaoling Yang, Ashish Lal, Hyeon Ho Kim, Yuki Kuwano, Stefanie Galban, Kevin G. Becker, Davida Kamara, Rafael de Cabo, Myriam Gorospe, Ubiquitin-mediated proteolysis of HuR by heat shock, *EMBO J.* 28 (9) (2009) 1271–1282.
- [38] R. Mazroui, S. Di Marco, E. Clair, C. von Roretz, S.A. Tenenbaum, J.D. Keene, M. Saleh, I.-Eddine Gallouzi, Caspase-mediated cleavage of HuR in the cytoplasm contributes to pp32/PHAP-I regulation of apoptosis, *J. Cell. Biol.* 180 (2008) 113–127.
- [39] Angélica Figueroa, Ana Cuadrado, Jinshui Fan, Ulus Atasoy, George E. Muscat, Pura Mun-oz-Canoves, Myriam Gorospe, Alberto Mun-oz, Role of HuR in skeletal myogenesis through coordinate regulation of muscle differentiation genes, *Mol. Cell. Biol.* 23 (14) (2003) 4991–5004.
- [40] Marilynne Levadoux-Martin, Agnès Gouble, Bernard Jégou, Virginie Vallet-Erdtmann, Jacques Auriol, Pascale Mercier, Dominique Morello, Impaired gametogenesis in mice that overexpress the RNA-binding protein HuR, *EMBO Rep.* 4 (4) (2003) 394–399.
- [41] Kate van der Giessen, Sergio Di-Marco, Eveline Clair, Imed Eddine Gallouzi, Eddine Gallouzi, RNAi-mediated HuR depletion leads to the inhibition of muscle cell differentiation, *J. Biol. Chem.* 278 (47) (2003) 47119–47128.
- [42] J. Cherry, V. Karschner, H. Jones, P.H. Pekala, HuR, an RNA-binding protein, involved in the control of cellular differentiation, *In vivo* 20 (2006) 17–24.
- [43] M. Nguyen-Chi, D. Morello, RNA-binding proteins, RNA granules, and gametes: is unity strength?, *Reproduction* 142 (2011) 803–817.
- [44] J. Li, L. Gong, S. Liu, Y. Zhang, C. Zhang, M. Tian, H. Lu, P. Bu, J. Yang, C. Ouyang, X. Jiang, J. Wu, Y. Zhang, Q. Min, C. Zhang, W. Zhang, Adipose HuR protects against diet-induced obesity and insulin resistance, *Nat. Commun.* 10 (2019) 2375.
- [45] S. Di Marco, R. Mazroui, P. Dallaire, S. Chittur, S.A. Tenenbaum, D. Radzioch, A. Marette, I.-Eddine Gallouzi, NF- κ B-mediated MyoD decay during muscle wasting requires nitric oxide synthase mRNA stabilization, HuR protein, and nitric oxide release, *Mol. Cell. Biol.* 25 (2005) 6533–6545.
- [46] Mallika Ghosh, Hector Leonardo Aguila, Jason Michaud, Youxi Ai, Ming-Tao Wu, Annabrita Hemmes, Ari Ristimaki, Caiying Guo, Henry Furneaux, Timothy Hla, Essential role of the RNA-binding protein HuR in progenitor cell survival in mice, *J. Clin. Invest.* 119 (12) (2009) 3530–3543.
- [47] X. Li, Y.-C. Lu, K. Dai, I. Torregroza, T. Hla, T. Evans, Elavl1a regulates zebrafish erythropoiesis via posttranscriptional control of gata1, *Blood* 123 (2014) 1384–1392.
- [48] Mai Nguyen Chi, Jacques Auriol, Bernard Jégou, Dimitris L. Kontoyiannis, James M.A. Turner, Dirk G. de Rooij, Dominique Morello, A. Gregory Matera, The RNA-binding protein ELAVL1/HuR is essential for mouse spermatogenesis, acting both at meiotic and postmeiotic stages, *Mol. Cell. Biol.* 22 (16) (2011) 2875–2885.
- [49] Nikos Sgantzis, Anthie Yiakouvakli, Eumorphia Remboutsika, Dimitris L. Kontoyiannis, HuR controls lung branching morphogenesis and mesenchymal FGF networks, *Dev. Biol.* 354 (2) (2011) 267–279.
- [50] Z. Zhang, C. Zong, M. Jiang, H. Hu, X. Cheng, J. Ni, X. Yi, B. Jiang, F. Tian, M.W. Chang, W. Su, L. Zhu, J. Li, X. Xiang, C. Miao, M. Gorospe, R. de Cabo, Y. Dou, Z.

- Ju, J., Yang, C., Jiang, Z., Yang, W., Wang, Hepatic HuR modulates lipid homeostasis in response to high-fat diet, *Nat Comm.* 11 (2020) 3067.
- [51] D.T.C. Siang, Y.C. Lim, A.M.M. Kyaw, K.N. Win, S.Y. Chia, U. Degirmenci, X. Hu, B.C. Tan, A.C.E. Walet, L. Sun, D. Xu, The RNA-binding protein HuR is a negative regulator in adipogenesis, *Nat Comm.* 11 (2020) 213.
- [52] Shanshan Liu, Xiuxin Jiang, Hanlin Lu, Mengdan Xing, Yanning Qiao, Cheng Zhang, Wencheng Zhang, HuR (Human antigen R) regulates the contraction of vascular smooth muscle and maintains blood pressure, *Arterioscler. Thromb. Vasc. Biol.* 40 (4) (2020) 943–957.
- [53] Sung-Hee Chang, Olivier Elemento, Jiasheng Zhang, Zhen W. Zhuang, Michael Simons, Timothy Hla, ELAVL1 regulates alternative splicing of eIF4E transporter to promote postnatal angiogenesis, *PNAS* 111 (51) (2014) 18309–18314.
- [54] H. Hu, M. Jiang, Y. Cao, Z. Zhang, B. Jiang, F. Tian, J. Feng, Y. Dou, M. Gorospe, M. Zheng, L. Zheng, Z. Yang, W. Wang, HuR regulates phospholamban expression in isoproterenol-induced cardiac remodeling, *Cardiovasc. Res.* 116 (2020) 944–955.
- [55] L.C. Green, S.R. Anthony, S. Slone, L. Lanzillotta, M.L. Nieman, X. Wu, N. Robbins, S.M. Jones, S. Roy, A.P. Owens III, J. Aube, L. Xu, J.N. Lorenz, B.C. Blaxall, J. Rubinstein, J.B. Benoit, M. Tranter, Human antigen R as a therapeutic target in pathological cardiac hypertrophy, *JCI Insight* 4 (2019) e121541.
- [56] B.J. Sánchez, A.-M.K. Tremblay, J.-P. Leduc-Gaudet, D.T. Hall, E. Kovacs, J.F. Ma, S. Mubaid, P.L. Hallauer, B.L. Phillips, K.E. Vest, A.H. Corbett, D.L. Kontoyiannis, S.N.A. Hussain, K.E.M. Hastings, S. Di Marco, I.-Eddine Gallouzi, Depletion of HuR in murine skeletal muscle enhances exercise endurance and prevents cancer-induced muscle atrophy, *Nat. Commun.* 10 (2019) 4171.
- [57] Souad Mubaid, Jennifer F. Ma, Amr Omer, Kholoud Ashour, Xian J. Lian, Brenda J. Sanchez, Samantha Robinson, Anne Cammas, Virginia Dormoy-Raclet, Sergio Di Marco, Sridar V. Chittur, Scott A. Tenenbaum, Imed-Eddine Gallouzi, I.-Eddine Gallouzi, HuR counteracts miR-330 to promote STAT3 translation during inflammation-induced muscle wasting, *PNAS* 116 (35) (2019) 17261–17270.
- [58] Randall L. Mynatt, Robert C. Noland, Carrie M. Elks, Bolormaa Vandanmagsar, David S. Bayless, Allison C. Stone, Sujoy Ghosh, Eric Ravussin, Jaycob D. Warfel, The RNA binding protein HuR influences skeletal muscle metabolic flexibility in rodents and humans, *Metabolism* 97 (2019) 40–49.
- [59] Antonina Giammanco, Valerie Blanc, Grace Montenegro, Coen Klos, Yan Xie, Susan Kennedy, Jianyang Luo, Sung-Hee Chang, Timothy Hla, İlke Nalbantoglu, Sekhar Dharmarajan, Nicholas O. Davidson, Intestinal epithelial HuR modulates distinct pathways of proliferation and apoptosis and attenuates small intestinal and colonic tumor development, *Cancer Res.* 74 (18) (2014) 5322–5335.
- [60] L. Liu, R. Zhuang, L. Xiao, H.K. Chung, J. Luo, D.J. Turner, J.N. Rao, M. Gorospe, J.-Y. Wang, HuR enhances early restitution of the intestinal epithelium by increasing Cdc42 translation, *Mol. Cell. Biol.* 37 (2017) e00574–16.
- [61] Lan Xiao, Xiao-Xue Li, Hee Kyoung Chung, Sudhakar Kalakonda, Jia-Zhong Cai, Shan Cao, Ning Chen, Yulan Liu, Jaladanki N. Rao, Hong-Ying Wang, Myriam Gorospe, Jian-Ying Wang, RNA-binding protein HuR regulates Paneth cell function by altering membrane localization of TLR2 via posttranscriptional control of CNPY3, *Gastroenterology* 157 (3) (2019) 731–743.
- [62] Lan Liu, Lan Xiao, Hee K. Chung, Min S. Kwon, Xiao-Xue Li, Na Wu, Jaladanki N. Rao, Jian-Ying Wang, RNA-binding protein HuR regulates Rac1 nucleocytoplasmic shuttling through nucleophosmin in the intestinal epithelium, *Cell. Mol. Gastroent. Hepatol.* 8 (3) (2019) 475–486.
- [63] Jing Chen, Jennifer L. Martindale, Kotb Abdelmohsen, Gaurav Kumar, Paolo M. Fortina, Myriam Gorospe, Abdolmohamad Rostami, Shiguang Yu, RNA-binding protein HuR promotes Th17 cell differentiation and can be targeted to reduce autoimmune neuroinflammation, *J. Immunol.* 204 (8) (2020) 2076–2087.
- [64] Jing Chen, Jason Cascio, Joseph D. Magee, Patsharaporn Techasintana, Matthew M. Gubin, Garrett M. Dahm, Robert Calaluce, Shiguang Yu, Ulus Atasoy, Posttranscriptional gene regulation of IL-17 by the RNA-binding protein HuR is required for initiation of experimental autoimmune encephalomyelitis, *J. Immunol.* 191 (11) (2013) 5441–5450.
- [65] Patsharaporn Techasintana, Jason S. Ellis, Jacqueline Glascock, Matthew M. Gubin, Suzanne E. Ridenhour, Joseph D. Magee, Marcia L. Hart, Peng Yao, Hao Zhou, Maryln S. Whitney, Craig L. Franklin, Jennifer L. Martindale, Myriam Gorospe, Wade J. Davis, Paul L. Fox, Xiaoxia Li, Ulus Atasoy, The RNA-binding protein HuR posttranscriptionally regulates IL-2 homeostasis and CD4+ Th2 differentiation, *ImmunoHorizons* 1 (6) (2017) 109–123.
- [66] Amy DeMicco, Martin S. Naradikian, Vishal J. Sindhava, Je-Hyun Yoon, Myriam Gorospe, Gerald B. Wertheim, Michael P. Cancro, Craig H. Bassing, B cell-intrinsic expression of the HuR RNA-binding protein is required for the T cell-dependent immune response in vivo, *J. Immunol.* 195 (7) (2015) 3449–3462.
- [67] M.D. Diaz-Muñoz, S.E. Bell, K. Fairfax, E. Monzon-Casanova, A.F. Cunningham, M. Gonzalez-Porta, S.R. Andrews, V.I. Bunik, K. Zarnack, T. Curk, W.A. Heggermont, S. Heymans, G.E. Gibson, D.L. Kontoyiannis, J. Ule, M. Turner, The RNA-binding protein HuR (Elavl1) is essential for the B cell antibody response, *Nat. Immunol.* 16 (2015) 415–425.
- [68] Sung-Hee Chang, Yi-Chien Lu, Xi Li, Wan-Ying Hsieh, Yuquan Xiong, Mallika Ghosh, Todd Evans, Olivier Elemento, Timothy Hla, Antagonistic function of the RNA-binding protein HuR and miR-200b in post-transcriptional regulation of vascular endothelial growth factor-A expression and angiogenesis, *J. Biol. Chem.* 288 (7) (2013) 4908–4921.
- [69] Anthie Yiakouvakis, Marios Dimitriou, Ioannis Karakasiliotis, Christina Eftychi, Stamatis Theocharis, Dimitris L. Kontoyiannis, Myeloid cell expression of the RNA-binding protein HuR protects mice from pathologic inflammation and colorectal carcinogenesis, *J. Clin. Invest.* 122 (1) (2012) 48–61.
- [70] Vicky Katsanou, Olympia Papadaki, Stavros Milatos, Perry J. Blackshear, Paul Anderson, George Kollias, Dimitris L. Kontoyiannis, HuR as a negative posttranscriptional modulator in inflammation, *Mol. Cell* 19 (6) (2005) 777–789.
- [71] Olympia Papadaki, Stavros Milatos, Sofia Grammenoudi, Neelanjan Mukherjee, Jack D. Keene, Dimitris L. Kontoyiannis, Control of thymic T cell maturation, deletion and egress by the RNA-binding protein HuR1, *J. Immunol.* 182 (11) (2009) 6779–6788.
- [72] Jiping Wang, Jianmei W. Leavenworth, Anita B. Hjelmeland, Reed Smith, Neha Patel, Ben Borg, Ying Si, Peter H. King, Deletion of the RNA regulator HuR in tumor associated microglia and macrophages stimulates antitumor immunity and attenuates glioma growth, *Glia* 67 (12) (2019) 2424–2439.
- [73] Z.-X. He, H.-F. Song, T.-Y. Liu, J. Ma, Z.-K. Xing, Y.-Y. Yin, L. Liu, Y.-N. Zhang, Y.-F. Zhao, H.-L. Yu, X.-X. He, W.-X. Guo, X.-J. Zhu, HuR in the medial prefrontal cortex is critical for stress-induced synaptic dysfunction and depressive-like symptoms in mice, *Cereb. Cortex* 29 (2019) 1–11.
- [74] M.L. Kraushar, K. Thompson, H.R.S. Wijeratne, B. Viljetic, K. Sakers, J.W. Marson, D.L. Kontoyiannis, S. Buyske, R.P. Hart, M.-R. Rasin, Temporally defined neocortical translation and polysome assembly are determined by the RNA-binding protein Hu antigen R, *PNAS* 111 (36) (2014) E3815–E3824.
- [75] Kevin Sun, Xiao Li, Xing Chen, Ying Bai, Gao Zhou, Olga N. Kokiko-Cochran, Bruce Lamb, Thomas A. Hamilton, Ching-Yi Lin, Yu-Shang Lee, Tomasz Herjan, Neuron-specific HuR-deficient mice spontaneously develop motor neuron disease, *J. Immunol.* 201 (1) (2018) 157–166.
- [76] Lei Gu, Huiwen Wang, Jun Wang, Yuting Guo, Yinglong Tang, Yang Mao, Lijuan Chen, Hua Lou, Guangju Ji, Reconstitution of HuR-inhibited CUGBP1 Expression Protects Cardiomyocytes from Acute Myocardial Infarction-Induced Injury, *Antiox. Redox. Signal.* 27 (14) (2017) 1013–1026.
- [77] Meng-yuan Zhou, Jin-Ming Yang, Xing-dong Xiong, The emerging landscape of circular RNA in cardiovascular diseases, *J. Mol. Cell. Cardiol.* 122 (2018) 134–139.
- [78] M. Amadio, C. Bucolo, G.M. Leggio, F. Drago, S. Govoni, A. Pascale, The PKC β /HuR/VEGF pathway in diabetic retinopathy, *Biochem. Pharmacol.* 80 (2010) 1230–1237.
- [79] Johanna Viiri, Marialaura Amadio, Nicoletta Marchesi, Juha M.T. Hyttinen, Niko Kivinen, Reijo Sironen, Kirsi Rilla, Saeed Akhtar, Alessandro Provenzani, Vito Giuseppe D'Agostino, Stefano Govoni, Alessia Pascale, Hansjurgen Agostini, Goran Petrovski, Antero Salminen, Kai Kaarimäntä, Michael E. Boulton, Autophagy activation clears ELAVL1/HuR-mediated accumulation of SQSTM1/p62 during proteasomal inhibition in human retinal pigment epithelial cells, *PLoS ONE* 8 (7) (2013) e69563.
- [80] C. Yu, W. Xin, J. Zhen, Y. Liu, A. Javed, R. Wang, Q. Wan, Human antigen R mediated post-transcriptional regulation of epithelial-mesenchymal transition related genes in diabetic nephropathy, *J. Diabetes* 7 (2015) 562–572.
- [81] A. Skliris, O. Papadaki, P. Kafasla, I. Karakasiliotis, O. Hazapis, M. Reczko, S. Grammenoudi, J. Bauer, D.L. Kontoyiannis, Neuroprotection requires the functions of the RNA-binding protein HuR, *Cell Death Differ.* 22 (5) (2015) 703–718.
- [82] Wengong Wang, Henry Furneaux, Huiming Cheng, M. Craig Caldwell, Dorothy Hutter, Yusen Liu, Nikki Holbrook, Myriam Gorospe, HuR regulates p21 mRNA stabilization by UV light, *Mol. Cell. Biol.* 20 (3) (2000) 760–769.
- [83] S.S.-Y. Peng, C.-Y.A. Chen, N. Xu, Ann-Bin Shyu, RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein, *EMBO J.* 17 (1998) 3461–3470.
- [84] Nina S. Levy, Sangmi Chung, Henry Furneaux, Andrew P. Levy, Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR, *J. Biol. Chem.* 273 (11) (1998) 6417–6423.
- [85] Wenhui Zhao, Jinfeng Zhao, Miaomiao Hou, Yue Wang, Yang Zhang, Xin Zhao, Ce Zhang, Dawei Guo, HuR and TIA1/TIAL1 are involved in regulation of alternative splicing of SIRT1 pre-mRNA, *Int. J. Mol. Sci.* 15 (2) (2014) 2946–2958.
- [86] D. Cao, J. Bian, Z.-C. Hua, L. Ma, W. Chen, X. Zhang, R. Zhou, S. Cheng, W.-Z. Sun, Q.-C. Jiao, W. Yin, Modulation of TNF- α mRNA stability by human antigen R and miR181s in sepsis-induced immunoparalysis, *EMBO Mol. Med.* 7 (2015) 140–157.
- [87] Daniella Ishimaru, Sivakumar Ramalingam, Tapas K. Sengupta, Sumita Bandyopadhyay, Stephanie Dellis, Baby G. Tholanikunnel, Daniel J. Fernandes, Eleanor K. Spicer, Regulation of Bcl-2 expression by HuR in HL60 leukemia cells and A431 carcinoma cells, *Mol. Cancer Res.* 7 (8) (2009) 1354–1366.
- [88] Dan A. Dixon, Neal D. Tolley, Peter H. King, L. Burt Nabors, Thomas M. McIntyre, Guy A. Zimmerman, Stephen M. Prescott, Altered expression of the mRNA stability factor HuR promotes cyclooxygenase-2 expression in colon cancer cells, *J. Clin. Invest.* 108 (11) (2001) 1657–1665.
- [89] Krystyna Mazan-Mamczarz, Stefania Galbán, Isabel López de Silanes, Jennifer L. Martindale, Ulus Atasoy, Jack D. Keene, Myriam Gorospe, RNA-binding protein HuR enhances p53 translation in response to ultraviolet light irradiation, *PNAS* 100 (14) (2003) 8354–8359.
- [90] Hanna Ronkainen, Markku H. Vaarala, Pasi Hirvikoski, Ari Ristimäki, HuR expression is a marker of poor prognosis in renal cell carcinoma, *Tumor Biol.* 32 (3) (2011) 481–487.
- [91] Zonglin Zhang, Aimei Huang, Aihong Zhang, Chenxia Zhou, HuR promotes breast cancer cell proliferation and survival via binding to CDK3 mRNA, *Biomed. Pharmacother.* 91 (2017) 788–795.

- [92] Isabel López de Silanes, Jinshui Fan, Xiaoling Yang, Alan B. Zonderman, Olga Potapova, Ellen S. Pizer, Myriam Gorospe, Role of the RNA-binding protein HuR in colon carcinogenesis, *Oncogene* 22 (46) (2003) 7146–7154.
- [93] X. Yi, Y. Zhou, W. Zheng, S.K. Chambers, HuR expression in the nucleus correlates with high histological grade and poor disease-free survival in ovarian cancer, *Austr. New Zeal. J. Obstet. Gynecol.* 49 (2009) 43–48.
- [94] Xiaohui Yu, Yujiao Li, Yumei Ding, Hong Zhang, Ning Ding, Ming Lu, HuR promotes ovarian cancer cell proliferation by regulating TIMM44 mRNA stability, *Cell Biochem. Biophys.* 78 (4) (2020) 447–453.
- [95] S. Niesporek, G. Kristiansen, A. Thoma, W. Weichert, A. Noske, A.-C. Buckendahl, K. Jung, C. Stephan, M. Dietel, C. Denkert, Expression of the ELAV-like protein HuR in human prostate carcinoma is an indicator of disease relapse and linked to COX-2 expression, *Int. J. Oncol.* 32 (2008) 341–347.
- [96] C.L. Costantino, A.K. Witkiewicz, Y. Kuwano, J.A. Cozzitorto, E.P. Kennedy, A. Dasgupta, J.C. Keen, C.J. Yeo, M. Gorospe, J.R. Brody, HuR's role in gemcitabine efficacy in pancreatic cancer: HuR upregulates the expression of the gemcitabine metabolizing enzyme, deoxycytidine kinase, *Cancer Res.* 69 (2009) 4567–4572.
- [97] H. Hasegawa, W. Kakuguchi, T. Kuroshima, T. Kitamura, S. Tanaka, Y. Kitagawa, Y. Totsuka, M. Shindoh, F. Higashino, HuR is exported to the cytoplasm in oral cancer cells in a different manner from that of normal cells, *Br. J. Cancer* 100 (12) (2009) 1943–1948.
- [98] Ana Rebane, Alar Aab, Joan A. Steitz, Transportins 1 and 2 are redundant nuclear import factors for hnRNP A1 and HuR, *RNA* 10 (4) (2004) 590–599.
- [99] W. Wang, X. Yang, T. Kawai, I. Lopez de Silanes, K. Mazan-Mamczarz, P. Chen, Y.M. Chook, C. Quensel, M. Kohler, M. Gorospe, AMP-activated protein kinase-regulated phosphorylation and acetylation of importin 1, *J. Biol. Chem.* 279 (2004) 48376–48388.
- [100] Barbara Fries, Jochen Heukeshoven, Ilona Hauber, Cordula Grüttner, Carol Stocking, Ralph H. Kehlenbach, Joachim Hauber, Jan Chemnitz, Analysis of nucleocytoplasmic trafficking of the HuR ligand APRIL and its influence on CD83 expression, *J. Biol. Chem.* 282 (7) (2007) 4504–4515.
- [101] Imed-Eddine Gallouzi, Christopher M. Brennan, Joan A. Steitz, Protein ligands mediate the CRM1-dependent export of HuR in response to heat shock, *RNA* 7 (9) (2001) 1348–1361.
- [102] G.L. Gravina, M. Tortoreto, A. Mancini, A. Addis, E. Di Cesare, A. Lenzi, Y. Landesman, D. McCauley, M. Kauffman, S. Schacham, N. Zaffaroni, C. Festuccia, XP01/CRM1-selective inhibitors of nuclear export (SINE) reduce tumor spreading and improve overall survival in preclinical models of prostate cancer (PCa), *J. Hematol. Oncol.* 7 (2014) 46.
- [103] Timothy K. Williams, Christina L. Costantino, Nikolai A. Bildzukewicz, Nathan G. Richards, David W. Rittenhouse, Lisa Einstein, Joseph A. Cozzitorto, Judith C. Keen, Abhijit Dasgupta, Myriam Gorospe, Gregory E. Gonye, Charles J. Yeo, Agnieszka K. Witkiewicz, Jonathan R. Brody, Janine Santos, pp32 (ANP32A) Expression inhibits pancreatic cancer cell growth and induces gemcitabine resistance by disrupting HuR binding to mRNAs, *PLoS ONE* 5 (11) (2010) e15455.
- [104] Yuhuan Wang, Ye Guo, Changyong Tang, Xiu Han, Mingyue Xu, Jing Sun, Yifei Zhao, Yiwen Zhang, Min Wang, Xiaofeng Cao, Xiaojuan Zhu, Weixiang Guo, Developmental cytoplasmic-to-nuclear translocation of RNA-binding protein HuR is required for adult neurogenesis, *Cell Rep.* 29 (10) (2019) 3101–3117. e7.
- [105] V. Simion, H. Zhou, S. Haemmig, J.B. Pierce, S. Mendes, Y. Tesmenitsky, D. Pérez-Cremades, J.F. Lee, A.F. Chen, N. Ronda, B. Papotti, J.A. Marto, M.W. Feinberg, A macrophage-specific lncRNA regulates apoptosis and atherosclerosis by tethering HuR in the nucleus, *Nat. Commun.* 11 (2020) 6135.
- [106] L.B. Nabors, G.Y. Gillespie, L. Harkins, P.H. King, HuR, a RNA stability factor, is expressed in malignant brain tumors and binds to adenine- and uridine-rich elements within the 3' untranslated regions of cytokine and angiogenic factor mRNAs1, *Cancer Res.* 61 (2001) 2154–2161.
- [107] Takahiro Sakuma, Takao Nakagawa, Kazunori Ido, Hiroaki Takeuchi, Kazufumi Sato, Toshihiko Kubota, Expression of vascular endothelial growth factor-A and mRNA stability factor HuR in human meningiomas, *J. Neurooncol.* 88 (2) (2008) 143–155.
- [108] T.-L. Erkinheimo, H. Lassus, A. Sivula, S. Sengupta, H. Furneaux, T. Hla, C. Haglund, R. Butzow, A. Ristimäki, Cytoplasmic HuR expression correlates with poor outcome and with cyclooxygenase 2 expression in serous ovarian carcinoma, *Cancer Res.* 63 (2003) 7591–7594.
- [109] Jun Wang, Weipeng Zhao, Yan Guo, Bicheng Zhang, Qichao Xie, Debing Xiang, Jianfei Gao, Baocheng Wang, Zhengtang Chen, The expression of RNA-binding protein HuR in non-small cell lung cancer correlates with vascular endothelial growth factor-C expression and lymph node metastasis, *Oncology* 76 (6) (2009) 420–429.
- [110] Łukasz Piotr Fus, Paweł Pihowicz, Łukasz Koperski, Janina Maja Marczevska, Barbara Górnicka, High cytoplasmic HuR expression is associated with advanced pT stage, high grade and increased microvessel density in urothelial bladder carcinoma, *Ann. Diagn. Pathol.* 33 (2018) 40–44.
- [111] Jiawei Guo, Jing Lv, Siyu Chang, Zhi Chen, Weiqiang Lu, Chuanliang Xu, Mingyao Liu, Xiufeng Pang, Inhibiting cytoplasmic accumulation of HuR synergizes genotoxic agents in urothelial carcinoma of the bladder, *Oncotarget* 7 (29) (2016) 45249–45262.
- [112] P.-I. Liang, W.-M. Li, Y.-H. Wang, T.-F. Wu, W.-R. Wu, A.C. Liao, K.-H. Shen, Y.-C. Wei, C.-H. Hsing, Y.-L. Shiue, H.-Y. Huang, H.-P. Hsu, L.-T. Chen, h.-Y. Lin, C. Tai, C.-M. Lin, C.-F. Li, HuR cytoplasmic expression is associated with increased cyclin A expression and poor outcome with upper urinary tract urothelial carcinoma, *BMC Cancer* 12 (2012) 611.
- [113] Cong Zhang, Guoliang Xue, Jingwang Bi, Ming Geng, Huili Chu, Yaping Guan, Jun Wang, Baocheng Wang, Cytoplasmic expression of the ELAV-like protein HuR as a potential prognostic marker in esophageal squamous cell carcinoma, *Tumor Biol.* 35 (1) (2014) 73–80.
- [114] Jun Wang, Baocheng Wang, Jingwang Bi, Cong Zhang, Cytoplasmic HuR expression correlates with angiogenesis, lymphangiogenesis, and poor outcome in lung cancer, *Med. Oncol.* 28 (S1) (2011) 577–585.
- [115] Jeong-Dan Cha, Shengjin Li, In-Ho Cha, Association between expression of embryonic lethal abnormal vision-like protein HuR and cyclooxygenase-2 in oral squamous cell carcinoma, *Head Neck* 33 (5) (2011) 627–637.
- [116] J.-D. Cha, H.-K. Kim, I.-H. Cha, Cytoplasmic HuR expression: Correlation with cellular inhibitors of apoptosis protein-2 expression and clinicopathologic factors in oral squamous cell carcinoma cells, *Head Neck* 36 (2014) 1168–1175.
- [117] D.-P. Sun, C.-Y. Lin, Y.-F. Tian, L.-T. Chen, L.-C. Lin, S.-W. Lee, C.-H. Hsing, H.-H. Lee, Y.-L. Shiue, H.-Y. Huang, C.-F. Li, P.-I. Liang, Clinicopathological significance of HuR expression in gallbladder carcinoma: with special emphasis on the implications of its nuclear and cytoplasmic expression, *Tumor Biol.* 34 (2013) 3059–3069.
- [118] A. Doller, E.-S. Akool, A. Huwiler, R. Muller, H.H. Radeke, J. Pfeilschifter, W. Eberhardt, Posttranslational modification of the AU-rich element binding protein HuR by protein kinase C elicits angiostatin II-induced stabilization and nuclear export of cyclooxygenase 2 mRNA, *Mol. Cell. Biol.* 28 (2008) 2608–2625.
- [119] Hyeon Ho Kim, Yuki Kuwano, Subramanya Srikantan, Eun Kyung Lee, Jennifer L. Martindale, Myriam Gorospe, Gorospe1, HuR recruits let-7/RISC to repress C-Myc expression, *Genes Dev.* 23 (15) (2009) 1743–1748.
- [120] Douglas Hanahan, Robert A. Weinberg, The Hallmarks of Cancer, *Cell* 100 (1) (2000) 57–70.
- [121] Douglas Hanahan, Robert A. Weinberg, Hallmarks of Cancer: the next generation, *Cell* 144 (5) (2011) 646–674.
- [122] Christopher W. Schultz, Ranjan Preet, Teena Dhir, Dan A. Dixon, Jonathan R. Brody, Understanding and targeting the disease-related RNA binding protein human antigen R (HuR), *Wiley Int. Rev. RNA* 11 (3) (2020), <https://doi.org/10.1002/wrna.v11.310.1002/wrna.1581>.
- [123] S. Zhu, S. Rooney, G. Michlewski, RNA-targeted therapies and high-throughput screening methods, *Int. J. Mol. Sci* 21 (2020) 2996.
- [124] Nicole-Claudia Meisner, Martin Hintersteiner, Kurt Mueller, Roman Bauer, Jan-Marcus Seifert, Hans-Ulrich Naegeli, Johannes Ottl, Lukas Oberer, Christian Guenat, Serge Moss, Nathalie Harter, Maximilian Woisetschlaeger, Christof Buehler, Volker Uhl, Manfred Auer, Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR, *Nat. Chem. Biol.* 3 (8) (2007) 508–515.
- [125] Min-Ju Chae, Hye Youn Sung, Eun-Hye Kim, Mira Lee, Hojoong Kwak, Chong Hak Chae, Sunwoo Kim, Woong-Yang Park, Chemical inhibitors destabilize HuR binding to the AU-rich element of TNF- α mRNA, *Exp. Mol. Med.* 41 (11) (2009) 824, <https://doi.org/10.3858/emmm.2009.41.11.088>.
- [126] Zhonghua Wang, Akash Bhattacharya, Dmitri N. Ivanov, Yong-Bin Yan, Identification of small-molecule inhibitors of the HuR/RNA interaction using a fluorescence polarization screening assay followed by NMR validation, *PLoS ONE* 10 (9) (2015) e0138780.
- [127] D.C. Hooper, J.S. Wolfson, G.L. McHugh, M.B. Winters, M.N. Swartz, Effects of novobiocin, coumermycin A1, clorobiocin, and their analogs on Escherichia coli DNA gyrase and bacterial growth, *Antimicrob. Ag. Chemother.* 22 (4) (1982) 662–671.
- [128] S. Walpole, S. Monaco, R. Nepravishta, J. Angulo, STD NMR as a technique for ligand screening and structural studies, in: J. Wand (Ed.), *Methods in Enzymology* 615, Elsevier Inc., Amsterdam, 2019, pp. 423–451.
- [129] J.L. Cole, J.W. Lary, T.P. Moody, T.M. Laue, Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium, *Methods Cell Biol.* 84 (2010) 143–179.
- [130] F. Vasile, S. della Volpe, F.A. Ambrosio, G. Costa, M.Y. Unver, C. Zucal, D. Rossi, E. Martino, A. Provenzani, A.K.H. Hirsch, S. Alcaro, D. Potenza, S. Collina, Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD-NMR and Molecular Modelling approach, *Sci. Rep.* 8 (2018) 13780.
- [131] V.G. D'Agostino, P. Lal, B. Mantelli, C. Tiedje, C. Zucal, N. Thongon, M. Gaestel, E. Latorre, L. Marinelli, P. Seneci, M. Amadio, A. Provenzani, Dihydrotanshinone-I interferes with the RNA-binding activity of HuR affecting its post-transcriptional function, *Sc. Rep.* 5 (2015) 16478.
- [132] Zhijie Zhu, Yang Zhao, Junbo Li, Li Tao, Peiliang Shi, Zhonghong Wei, Xiaobo Sheng, Dandan Shen, Zhaoguo Liu, Liang Zhou, Chao Tian, Fangtian Fan, Cunsi Shen, Pingting Zhu, Aiyun Wang, Wenxing Chen, Qingshun Zhao, Yin Lu, Cryptotanshinone, a novel tumor angiogenesis inhibitor, destabilizes tumor necrosis factor- α mRNA via decreasing nuclear-cytoplasmic translocation of RNA-binding protein HuR, *Mol. Carcinogen.* 55 (10) (2016) 1399–1410.
- [133] D. Cox-Georgian, N. Ramadoss, C. Dona, C. Basu, Therapeutic and medicinal uses of terpenes, in: N. Joshee, S. Dhekney, P. Parajuli (Eds.), *Medicinal Plants*, Springer International Publishing, Cham, 2019, pp. 333–359.
- [134] Lixin Sun, Shuang Zhang, Zhenzhou Jiang, Xin Huang, Tao Wang, Xiao Huang, Han Li, Luyong Zhang, Triptolide inhibits COX-2 expression by regulating mRNA stability in TNF- α -treated A549 cells, *Biochem. Biophys. Res. Commun.* 416 (1–2) (2011) 99–105.
- [135] Young Sun Hwang, Kwang-Kyun Park, Won-Yoon Chung, Kalopanaxsaponin A inhibits the invasion of human oral squamous cell carcinoma by reducing

- metalloproteinase-9 mRNA stability and protein trafficking, *Biol. Pharm. Bull.* 35 (3) (2012) 289–300.
- [136] W. Li, L. Luo, W. Shi, Y. Yin, S. Gao, Ursolic acid reduces Adriamycin resistance of human ovarian cancer cells through promoting the HuR translocation from cytoplasm to nucleus, *Environ. Toxicol.* (2020) 1–9.
- [137] Anke Doller, Amel Badawi, Tobias Schmid, Thilo Brauß, Thomas Pleli, Dagmar Meyer zu Heringdorf, Albrecht Piiper, Josef Pfeilschifter, Wolfgang Eberhardt, The cytoskeletal inhibitors latrunculin A and blebbistatin exert antitumorigenic properties in human hepatocellular carcinoma cells by interfering with intracellular HuR trafficking, *Exp. Cell Res.* 330 (1) (2015) 66–80.
- [138] Vito G. D'Agostino, Valentina Adami, Alessandro Provenzani, Albert Jeltsch, A novel high throughput biochemical assay to evaluate the HuR protein-RNA complex formation, *PLoS ONE* 8 (8) (2013) e72426.
- [139] Xiaoqing Wu, Lan Lan, David Michael Wilson, Rebecca T. Marquez, Weichung Tsao, Philip Gao, Anuradha Roy, Benjamin Andrew Turner, Peter McDonald, Jon A. Tunge, Steven A. Rogers, Dan A. Dixon, Jeffrey Aubé, Liang Xu, Identification and Validation of Novel Small Molecule Disruptors of HuR-mRNA Interaction, *ACS Chem. Biol.* 10 (6) (2015) 1476–1484.
- [140] Kawaljit Kaur, Xiaoqing Wu, James K. Fields, David K. Johnson, Lan Lan, Miranda Pratt, Amber D. Somoza, Clay C.C. Wang, John Karanicolas, Berl R. Oakley, Liang Xu, Roberto N. De Guzman, Dimitris L. Kontoyiannis, The fungal natural product azaphilone-9 binds to HuR and inhibits HuR-RNA interaction in vitro, *PLoS ONE* 12 (4) (2017) e0175471.
- [141] Chunyong Ding, Qianting Tian, Jie Li, Mingkun Jiao, Shanshan Song, Yingqing Wang, Zehong Miao, Ao Zhang, Structural modification of natural product tanshinone I leading to discovery of novel nitrogen-enriched derivatives with enhanced anticancer profile and improved drug-like properties, *J. Med. Chem.* 61 (3) (2018) 760–776.
- [142] L. Manzoni, C. Zucal, D. di Maio, V.G. D'Agostino, N. Thongon, I. Bonomo, P. Lal, M. Miceli, V. Baj, M. Brambilla, L. Cerofolini, S. Elezgarai, E. Biasini, C. Luchinat, E. Novellino, M. Fragai, L. Marinelli, A. Provenzani, P. Seneci, Interfering with HuR-RNA interaction: design, synthesis and biological characterization of tanshinone mimics as novel, effective HuR inhibitors, *J. Med. Chem.* 61 (2018) 1483–1498.
- [143] Ji-Yeon Lee, Tae-Wook Chung, Hee-Jung Choi, Chang Hyun Lee, Jae Soon Eun, Young Taek Han, Jun-Yong Choi, So-Yeon Kim, Chang-Woo Han, Han-Sol Jeong, Ki-Tae Ha, A novel cantharidin analog N-benzylcantharidinamide reduces the expression of MMP-9 and invasive potentials of Hep3B via inhibiting cytosolic translocation of HuR, *Biochem. Biophys. Res. Commun.* 447 (2) (2014) 371–377.
- [144] Nicole-Claudia Meisner, Martin Hintersteiner, Jan-Marcus Seifert, Roman Bauer, Roger Marc Benoit, Armin Widmer, Torsten Schindler, Volker Uhl, Michaela Lang, Hubert Gstach, Manfred Auer, Terminal adenosyl transferase activity of posttranscriptional regulator HuR revealed by confocal on-bead screening, *J. Mol. Biol.* 386 (2) (2009) 435–450.
- [145] Natalia Filippova, Xiuhua Yang, Subramaniam Ananthan, Anastasia Sorochinsky, James R. Hackney, Zachery Gentry, Sejong Bae, Peter King, L. Burt Nabors, Hu antigen R (HuR) multimerization contributes to glioma disease progression, *J. Biol. Chem.* 292 (41) (2017) 16999–17010.
- [146] N. Wiedemar, D.A. Hauser, P. Mäser, 100 years of suramin, *Antimicrob. Ag. Chemother.* 64 (2019) e01168–19.
- [147] Wataru Kakuguchi, Takao Nomura, Tetsuya Kitamura, Satoko Otsuguro, Kazuhiro Matsushita, Masahiro Sakaitani, Katsumi Maenaka, Kanchu Tei, Suramin, screened from an approved drug library, inhibits HuR functions and attenuates malignant phenotype of oral cancer cells, *Cancer Med.* 7 (12) (2018) 6269–6280.
- [148] S. della Volpe, R. Nasti, M. Queirolo, M.Y. Unver, V.K. Jumde, A. Dömling, F. Vasile, D. Potenza, F.A. Ambrosio, G. Costa, S. Alcaro, C. Zucal, A. Provenzani, M. di Giacomo, D. Rossi, A.K.H. Hirsch, S. Collina, Novel compounds targeting the RNA-binding protein HuR. Structure-based design, synthesis, and interaction studies, *ACS Med. Chem. Lett.* 10 (2019) 615–620.
- [149] S. Della Volpe, P. Linciano, R. Listro, E. Tumminelli, M. Amadio, I. Bonomo, W. A.M. Elgaher, S. Adam, A.K.H. Hirsch, F.M. Boeckler, F. Vasile, D. Rossi, S. Collina, Identification of N, N-arylalkyl-picolinamide derivatives targeting the RNA-binding protein HuR, by combining biophysical fragment-screening and molecular hybridization, *Bioorg. Chem.* 116 (2021) 105305, <https://doi.org/10.1016/j.bioorg.2021.105305>.
- [150] X. Wu, G. Gardashova, L. Lan, S. Han, C. Zhong, R.T. Marquez, L. Wei, S. Wood, S. Roy, R. Gowthaman, J. Karanicolas, F.P. Gao, D.A. Dixon, D.R. Welch, L. Li, M. Ji, J. Aubé, L. Xu, Targeting the interaction between RNA-binding protein HuR and FOXQ1 suppresses breast cancer invasion and metastasis, *Commun. Biol.* 3 (2020) 193.
- [151] C.B.M. Platania, V. Pittalà, A. Pascale, N. Marchesi, C.D. Anfuso, G. Lupo, M. Cristaldi, M. Olivieri, F. Lazzara, L. di Paola, F. Drago, C. Bucolo, Novel indole derivatives targeting HuR-mRNA complex to counteract high glucose damage in retinal endothelial cells, *Biochem. Pharmacol.* 175 (2020) 113908.
- [152] N. Filippova, X. Yang, S. Ananthan, J. Calano, V. Pathak, L. Bratton, R.H. Vekariya, S. Zhang, E. Ofori, E.N. Hayward, D. Namkoong, D.K. Crossman, M.R. Crowley, P.H. King, J. Mobley, L.B. Nabors, Targeting the HuR Oncogenic Role with a New Class of Cytoplasmic Dimerization Inhibitors, *Cancer Res.* 81 (2021) 2220–2233.
- [153] Yun-Ching Cheng, Jing-Ping Liou, Ching-Chuan Kuo, Wen-Yang Lai, Kuang-Hsing Shih, Chi-Yen Chang, Wen-Yu Pan, Joseph T. Tseng, Jang-Yang Chang, MPT0B098, a novel microtubule inhibitor that destabilizes the hypoxia-inducible factor-1 α mRNA through decreasing nuclear-cytoplasmic translocation of RNA-binding protein HuR, *Mol. Cancer Ther.* 12 (7) (2013) 1202–1212.
- [154] R. Muralidharan, A. Babu, N. Amreddy, K. Basalingappa, M. Mehta, A. Chen, Y. D. Zhao, U.B. Kompella, A. Munshi, R. Ramesh, Folate receptor-targeted nanoparticle delivery of HuR-RNAi suppresses lung cancer cell proliferation and migration, *J. Nanobiotechnol.* 14 (2016) 47.
- [155] Ranganayaki Muralidharan, Anish Babu, Narsireddy Amreddy, Akhil Srivastava, Allshine Chen, Yan Daniel Zhao, Uday B. Kompella, Anupama Munshi, Rajagopal Ramesh, Tumor-targeted nanoparticle delivery of HuR siRNA inhibits lung tumor growth in vitro and in vivo by disrupting the oncogenic activity of the RNA-binding protein HuR, *Mol. Cancer Ther.* 16 (8) (2017) 1470–1486.
- [156] Yu-Hung Huang, Weidan Peng, Narumi Furuuchi, Jacquelyn Gerhart, Kelly Rhodes, Neelanjan Mukherjee, Masaya Jimbo, Gregory E. Gonye, Jonathan R. Brody, Robert C. Getts, Janet A. Sawicki, Delivery of therapeutics targeting the mRNA-binding protein HuR using 3DNA nanocarriers suppresses ovarian tumor growth, *Cancer Res.* 76 (6) (2016) 1549–1559.
- [157] Marialaura Amadio, Alessia Pascale, Sarha Cupri, Rosario Pignatello, Cecilia Osera, Velia D'Agata, Agata Grazia D'Amico, Gian Marco Leggio, Barbara Ruozzi, Stefano Govoni, Filippo Drago, Claudio Bucolo, Nanosystems based on siRNA silencing HuR expression counteract diabetic retinopathy in rat, *Pharmacol. Res.* 111 (2016) 713–720.
- [158] Carolina Lixa, Amanda Mujo, Mariana T.Q. de Magalhães, Fabio C.L. Almeida, Luis Mauricio T.R. Lima, Anderson S. Pinheiro, Oligomeric transition and dynamics of RNA binding by the HuR RRM1 domain in solution, *J. Biomol. NMR* 72 (3–4) (2018) 179–192.
- [159] R.M. Benoit, N.C. Meisner, J. Kallen, P. Graff, R. Hemmig, R. Cèbe, C. Ostermeier, H. Widmer, M. Auer, The X-ray crystal structure of the first RNA recognition motif and site-directed mutagenesis suggest a possible hur redox sensing mechanism, *J. Mol. Biol.* 397 (2010) 1231–1244.
- [160] R.M. Scheiba, A.I. De Opakua, A. Díaz-Quintana, I. Cruz-Gallardo, L.A. Martínez-Cruz, M.L. Martínez-Chantar, F.J. Blanco, I. Díaz-Moreno, The C-terminal RNA binding motif of HuR is a multi-functional domain leading to HuR oligomerization and binding to U-rich RNA targets, *RNA Biol.* 11 (2014) 1250–1261.
- [161] Nina Ripin, Julien Boudet, Malgorzata M. Duszczyk, Alexandra Hinniger, Michael Fallner, Miroslav Krepl, Abhilash Gadi, Robert J. Schneider, Jiří Šponer, Nicole C. Meisner-Kober, Frédéric H.-T. Allain, Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM, *PNAS* 116 (8) (2019) 2935–2944.
- [162] H.S. Kim, M.C.J. Wilce, Y.M.K. Yoga, N.R. Pendini, M.J. Gunzburg, N.P. Cowieson, G.M. Wilson, B.R.G. Williams, M. Gorospe, J.A. Wilce, Different modes of interaction by TIAR and HuR with target RNA and DNA, *Nucleic Acids Res.* 39 (2011) 1117–1130.
- [163] R. Nasti, D. Rossi, M. Amadio, A. Pascale, M.Y. Unver, A.K.H. Hirsch, S. Collina, Compounds interfering with embryonic lethal abnormal vision (ELAV) protein-RNA complexes: an avenue for discovering new drugs, *J. Med. Chem.* 60 (2017) 8257–8267.
- [164] A. Díaz-Quintana, S.M. García-Mauriño, I. Díaz-Moreno, Dimerization model of the C-terminal RNA recognition motif of HuR, *FEBS Lett.* 589 (2015) 1059–1066.
- [165] Elizabeth J. Fialcowitz-White, Brandy Y. Brewer, Jeff D. Ballin, Chris D. Willis, Eric A. Toth, Gerald M. Wilson, Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences, *J. Biol. Chem.* 282 (29) (2007) 20948–20959.
- [166] Jonathan L.E. Dean, Robin Wait, Kamal R. Mahtani, Gareth Sully, Andrew R. Clark, Jeremy Saklatvala, The 3' untranslated region of tumor necrosis factor alpha mRNA is a target of the mRNA-stabilizing factor HuR, *Mol. Cell. Biol.* 21 (3) (2001) 721–730.
- [167] Fernando F. Blanco, Ranjan Preet, Andrea Aguado, Vikalp Vishwakarma, Laura E. Stevens, Alok Vyas, Subhash Padhye, Liang Xu, Scott J. Weir, Shrikant Anant, Nicole Meisner-Kober, Jonathan R. Brody, Dan A. Dixon, Impact of HuR inhibition by the small molecule MS-444 on colorectal cancer cell tumorigenesis, *Oncotarget* 7 (45) (2016) 74043–74058.
- [168] Z. Zhu, B. Wang, J. Bi, C. Zhang, Y. Guo, H. Chu, X. Liang, C. Zhong, J. Wang, Cytoplasmic HuR expression correlates with P-gp, HER-2 positivity, and poor outcome in breast cancer, *Tumor Biology* 34 (2013) 2299–2308.
- [169] X.C. Fan, J.A. Steitz, HNS, a nuclear-cytoplasmic shuttling sequence in HuR, *PNAS* 95 (26) (1998) 15293–15298.
- [170] K. Mazan-Mamczarz, P.R. Hagner, S. Corl, S. Srikantan, W.H. Wood, K.G. Becker, M. Gorospe, J.D. Keene, A.S. Levenson, R.B. Gartenhaus, Post-transcriptional gene regulation by HuR promotes a more tumorigenic phenotype, *Oncogene* 27 (47) (2008) 6151–6163.
- [171] M. Palomo-Irigoyen, E. Pérez-Andrés, M. Iruarrizaga-Lejarreta, A. Barreira-Manrique, M. Tamayo-Caro, L. Vila-Vecilla, L. Moreno-Cugnon, N. Beitia, D. Medrano, D. Fernández-Ramos, J.J. Lozano, S. Okawa, J.L. Lavín, N. Martín-Martín, J.D. Sutherland, V. Gutiérrez de Juan, M. Gonzalez-Lopez, N. Macías-Cámara, D. Mosén-Ansorena, L. Laraba, C.O. Hanemann, E. Ercolano, D.B. Parkinson, C.W. Schultz, M.J. Araújo-Bravo, A.M. Ascensión, D. Gerovska, H. Iribar, A. Izeta, P. Pytel, P. Krastel, A. Provenzani, P. Seneci, R.D. Carrasco, A. Del Sol, M.L. Martínez-Chantar, R. Barrio, E. Serra, C. Lázaro, A.M. Flanagan, M. Gorospe, N. Ratner, A.M. Aransay, A. Carracedo, M. Varela-Rey, A. Woodhoo, HuR/ELAVL1 drives malignant peripheral nerve sheath tumor growth and metastasis, *J. Clin. Invest.* 130 (2020) 3848–3864.
- [172] H. Wang, Y. Chen, J. Guo, T. Shan, K. Deng, J. Chen, L. Cai, H. Zhou, Q. Zhao, S. Jin, J. Xia, Dysregulation of tristetraprolin and human antigen R promotes

- gastric cancer progressions partly by upregulation of the high-mobility group box 1, *Sci. Rep.* 8 (2018) 7080.
- [173] Mahsa Zarei, Shruti Lal, Ali Vaziri-Gohar, Kevin O'Hayer, Venugopal Gunda, Pankaj K. Singh, Jonathan R. Brody, Jordan M. Winter, RNA-binding protein HuR regulates both mutant and wild-type IDH1 in IDH1-mutated cancer, *Mol. Cancer Ther.* 17 (2) (2019) 508–520.
- [174] Saswati N. Chand, Mahsa Zarei, Matthew J. Schiewer, Akshay R. Kamath, Carmella Romeo, Shruti Lal, Joseph A. Cozzitorto, Avinoam Nevler, Laura Scolari, Eric Londin, Wei Jiang, Nicole Meisner-Kober, Michael J. Pishvaian, Karen E. Knudsen, Charles J. Yeo, John M. Pascal, Jordan M. Winter, Jonathan R. Brody, Posttranscriptional regulation of PARG mRNA by HuR facilitates DNA repair and resistance to PARP inhibitors, *Cancer Res.* 77 (18) (2017) 5011–5025.
- [175] Shruti Lal, Edwin C. Cheung, Mahsa Zarei, Ranjan Preet, Saswati N. Chand, Nicole C. Mambelli-Lisboa, Carmella Romeo, Matthew C. Stout, Eric Londin, Austin Goetz, Cinthya Y. Lowder, Avinoam Nevler, Charles J. Yeo, Paul M. Campbell, Jordan M. Winter, Dan A. Dixon, Jonathan R. Brody, CRISPR knockout of the HuR gene causes a xenograft lethal phenotype, *Mol. Cancer Ther.* 15 (6) (2017) 696–707.
- [176] Natalia Filippova, Xiuhua Yang, Yimin Wang, G. Yancey Gillespie, Cathy Langford, Peter H. King, Crystal Wheeler, L. Burt Nabors, The RNA-binding protein HuR promotes glioma growth and treatment resistance, *Mol. Cancer Res.* 9 (5) (2011) 648–659.
- [177] S. Danilin, C. Sourbier, L. Thomas, V. Lindner, S. Rothhut, V. Dormoy, J.-J. Helwig, D. Jacqmin, H. Lang, T. Massfelder, Role of the RNA-binding protein HuR in human renal cell carcinoma, *Carcinogenesis* 31 (2010) 1018–1026.
- [178] Natalia Filippova, Xiuhua Yang, Louis Nabors, Growth factor dependent regulation of centrosome function and genomic instability by HuR, *Biomolecules* 5 (1) (2015) 263–281.
- [179] M. Mehta, K. Basalingappa, J.N. Griffith, D. Andrade, A. Babu, N. Amreddy, R. Muralidharan, M. Gorospe, T. Herman, W.-Q. Ding, R. Ramesh, A. Munshi, HuR silencing elicits oxidative stress and DNA damage and sensitizes human triple-negative breast cancer cells to radiotherapy, *Oncotarget* 7 (2016) 64820–54835.
- [180] S. Lal, R.A. Burkhart, N. Beeharry, V. Bhattacharjee, E.R. Londin, J.A. Cozzitorto, C. Romeo, M. Jimbo, Z.A. Norris, C.J. Yeo, J.A. Sawicki, J.M. Winter, I. Rigoutsos, T.J. Yen, J.R. Brody, HuR posttranscriptionally regulates WEE1: Implications for the DNA damage response in pancreatic cancer cells, *Cancer Res.* 74 (2014) 1128–1140.
- [181] Subramanya Srikantan, Kotb Abdelmohsen, Eun Kyung Lee, Kumiko Tominaga, Sarah S. Subaran, Yuki Kuwano, Ritu Kulshrestha, Rohit Panchakshari, Hyeon Ho Kim, Xiaoling Yang, Jennifer L. Martindale, Bernard S. Marasa, Mihee M. Kim, Robert P. Wersto, Fred E. Indig, Dipanjan Chowdhury, Myriam Gorospe, Translational control of TOP2A influences doxorubicin efficacy, *Mol. Cell. Biol.* 31 (18) (2011) 3790–3801.
- [182] E. Latorre, T. Tebaldi, G. Viero, A.M. Spartà, A. Quattrone, A. Provenzani, Downregulation of HuR as a new mechanism of doxorubicin resistance in breast cancer cells, *Mol. Cancer* 11 (2012) 13.
- [183] S. Prislei, E. Martinelli, M. Mariani, G. Raspaglio, S. Sieber, G. Ferrandina, S. Shahabi, G. Scambia, C. Ferlini, MiR-200c and HuR in ovarian cancer, *BMC Cancer* 13 (2013) 72.
- [184] L. Allegri, F. Baldan, S. Roy, J. Aubé, D. Russo, S. Filetti, G. Damante, The HuR CMLD-2 inhibitor exhibits antitumor effects via MAD2 downregulation in thyroid cancer cells, *Sci. Rep.* 9 (2019) 7374.
- [185] H. Zhu, Z. Berkova, R. Mathur, L. Sehgal, T. Khashab, R.-H. Tao, X. Ao, L. Feng, A. L. Sabichi, B. Blechacz, A. Rashid, F. Samaniego, HuR suppresses Fas expression and correlates with patient outcome in liver cancer, *Mol. Cancer Res.* 13 (2015) 809–818.
- [186] Masaya Jimbo, Fernando F. Blanco, Yu-Hung Huang, Aristeidis G. Telonis, Brad A. Screnci, Gabriela L. Cosma, Vitali Alexeev, Gregory E. Gonye, Charles J. Yeo, Janet A. Sawicki, Jordan M. Winter, Jonathan R. Brody, Targeting the mRNA-binding protein HuR impairs malignant characteristics of pancreatic ductal adenocarcinoma cells, *Oncotarget* 6 (29) (2015) 27312–27331.
- [187] Carmella Romeo, Matthew C. Weber, Mahsa Zarei, Danielle DeCicco, Saswati N. Chand, Angie D. Lobo, Jordan M. Winter, Janet A. Sawicki, Jonathan N. Sachs, Nicole Meisner-Kober, Charles J. Yeo, Rajanikanth Vadigepalli, Mark L. Tykocinski, Jonathan R. Brody, HuR contributes to TRAIL resistance by restricting death receptor 4 expression in pancreatic cancer cells, *Mol. Cancer Res.* 14 (7) (2016) 599–611.
- [188] Jiping Wang, Anita B. Hjelmeland, L. Burt Nabors, Peter H. King, Anti-cancer effects of the HuR inhibitor, MS-444, in malignant glioma cells, *Cancer. Biol. Ther.* 20 (7) (2019) 979–988.
- [189] R. Muralidharan, M. Mehta, R. Ahmed, S. Roy, L. Xu, J. Aubé, A. Chen, Y.D. Zhao, T. Herman, R. Ramesh, A. Munshi, HuR-targeted small molecule inhibitor exhibits cytotoxicity towards human lung cancer cells, *Sci. Rep.* 7 (2017) 9694.
- [190] R.A. Burkhart, D.M. Pineda, S.N. Chand, C. Romeo, E.R. Londin, E.D. Karoly, J.A. Cozzitorto, I. Rigoutsos, C.J. Yeo, J.R. Brody, J.M. Winter, HuR is a post-transcriptional regulator of core metabolic enzymes in pancreatic cancer, *RNA Biol.* 10 (2013) 1312–1323.
- [191] M. Zarei, S. Lal, S.J. Parker, A. Nevler, A. Vaziri-Gohar, K. Dukleska, N.C. Mambelli-Lisboa, C. Moffat, F.F. Blanco, S.N. Chand, M. Jimbo, J.A. Cozzitorto, W. Jiang, C.J. Yeo, E.R. Londin, E.L. Seifert, C.M. Metallo, J.R. Brody, J.M. Winter, Posttranscriptional upregulation of IDH1 by Hu establishes a powerful survival phenotype in pancreatic cancer cells, *PNAS* 95 (1998) 15293–15298.
- [192] Rui Dong, Jian-Guo Lu, Qing Wang, Xian-Li He, Yan-Kui Chu, Qing-Jiu Ma, Stabilization of Snail by HuR in the process of hydrogen peroxide induced cell migration, *Biochim. Biophys. Res. Commun.* 356 (1) (2007) 318–321.
- [193] Mira Heinonen, Annabrita Hemmes, Kaisa Salmenkivi, Kotb Abdelmohsen, Sivi-Tuuli Vilén, Marko Laakso, Marjut Leidenius, Tuula Salo, Sampa Hautaniemi, Myriam Gorospe, Päivi Heikkilä, Caj Haglund, Ari Ristimäki, Role of RNA binding protein HuR in ductal carcinoma in situ of the breast, *J. Pathol.* 224 (4) (2011) 529–539.
- [194] Yasuyoshi Miyata, Shin-ichi Watanabe, Yuji Sagara, Kensuke Mitsunari, Tomohiro Matsuo, Kojiro Ohba, Hideki Sakai, Kaustubh Datta, High expression of HuR in cytoplasm, but not nuclei, is associated with malignant aggressiveness and prognosis in bladder cancer, *PLoS ONE* 8 (3) (2013) e59095.
- [195] Z. Yuan, A.J. Sanders, L. Ye, Y. Wang, W.G. Jiang, Knockdown of human antigen R reduces the growth and invasion of breast cancer cells in vitro and affects expression of cyclin D1 and MMP-9, *Oncology Rep.* 26 (2011) 237–245.
- [196] Jun Wang, Yan Guo, Huihui Chu, Yaping Guan, Jingwang Bi, Baocheng Wang, Multiple functions of the RNA-binding protein HuR in cancer progression, treatment responses and prognosis, *Int. J. Mol. Sci.* 14 (5) (2013) 10015–10041.
- [197] R. Dong, G.-D. Yang, N.-A. Luo, Y.-Q. Qu, HuR: a promising therapeutic target for angiogenesis, *Gland Surgery* 3 (2014) 203–206.
- [198] L.B. Nabors, E. Suswam, Y. Huang, X. Yang, M.J. Johnson, P.H. King, Tumor necrosis factor α induces angiogenic factor up-regulation in malignant glioma cells: a role for RNA stabilization and HuR1, *Cancer Res.* 63 (2003) 4181–4187.
- [199] Vladimir L. Gabai, Le Meng, Geunwon Kim, Teresa A. Mills, Ivor J. Benjamin, Michael Y. Sherman, Heat shock transcription factor Hsf1 is involved in tumor progression via regulation of hypoxia-inducible factor 1 and RNA-binding protein HuR, *Mol. Cell. Biol.* 32 (5) (2012) 929–940.
- [200] J. Zhang, Y. Modi, T. Yarovinsky, J. Yu, M. Collinge, T. Kyriakides, Y. Zhu, W.C. Sessa, R. Pardi, J.R. Bender, Macrophage 2 integrin-mediated, HuR-dependent stabilization of angiogenic factor-encoding mRNAs in inflammatory angiogenesis, *Am. J. Pathol.* 180 (2012) 1751–1760.
- [201] Harinder Gill, Raymond S.M. Wong, Yok-Lam Kwong, From chronic immune thrombocytopenia to severe aplastic anemia: recent insights into the evolution of eltrombopag, *Ther. Adv. Hematol.* 8 (5) (2017) 159–174.
- [202] Yuying Zhu, Liuqing Yang, Jiazhen Xu, Xiyang Yang, Pengwei Luan, Qianfei Cui, Pei Zhang, Feiyun Wang, Ruixiang Li, Xinyue Ding, Lixian Jiang, Guoqiang Lin, Jiange Zhang, Discovery of the anti-angiogenesis effect of eltrombopag in breast cancer through targeting of HuR protein, *Acta Pharm. Sin. B* 10 (8) (2020) 1414–1425.
- [203] R. Ahmed, R. Muralidharan, A. Srivastava, S.E. Johnston, Y.D. Zhao, S. Ekmekcioglu, A. Munshi, R. Ramesh, Molecular targeting of HuR oncoprotein suppresses MITF and induces apoptosis in melanoma cells, *Cancer* 13 (2021) 166.
- [204] Teena Dhir, Christopher W. Schultz, Aditi Jain, Samantha Z. Brown, Alex Haber, Austin Goetz, Chunhua Xi, Gloria H. Su, Liang Xu, James Posey, Wei Jiang, Charles J. Yeo, Talia Golan, Michael J. Pishvaian, Jonathan R. Brody, Abemaciclib is effective against pancreatic cancer cells and synergizes with HuR and YAP1 inhibition, *Mol. Cancer Res.* 17 (10) (2019) 2029–2041.
- [205] C von Roretz, X. Jin Lian, A.M. Macri, N. Punjani, E. Clair, O. Drouin, V. Dormoy-Raclet, J.F. Ma, I.-E. Gallouzi, Apoptotic-induced cleavage shifts HuR from being a promoter of survival to an activator of caspase-mediated apoptosis, *Cell Death Differ.* 20 (1) (2013) 154–168.
- [206] Danielle M. Pineda, David W. Rittenhouse, Christopher C. Valley, Joseph A. Cozzitorto, Richard A. Burkhart, Benjamin Leiby, Jordan M. Winter, Matthew C. Weber, Eric R. Londin, Isidore Rigoutsos, Charles J. Yeo, Myriam Gorospe, Agnieska K. Witkiewicz, Jonathan N. Sachs, Jonathan R. Brody, HuR's post-transcriptional regulation of death receptor 5 in pancreatic cancer cells, *Cancer Biol. Ther.* 13 (10) (2012) 946–955.
- [207] Sun-Lung Tsai, Fat-Moon Suk, Chun-I. Wang, Der-Zen Liu, Wen-Chi Hou, Pei-Jung Lin, Ling-Fang Hung, Yu-Chih Liang, Anti-tumor potential of 15,16-dihydrotanshinone I against breast adenocarcinoma through inducing G1 arrest and apoptosis, *Biochem. Pharmacol.* 74 (11) (2007) 1575–1586.
- [208] J. Fan, N.M. Heller, M. Gorospe, U. Atasoy, C. Stellato, The role of post-translational regulation in chemokine gene expression in inflammation and allergy, *Eur. Respir. J.* 26 (2005) 933–947.
- [209] Yuko Seko, Steven Cole, Wojciech Kasprzak, Bruce A. Shapiro, Jack A. Ragheb, The role of cytokine mRNA stability in the pathogenesis of autoimmune disease, *Autoimm. Rev.* 5 (5) (2006) 299–305.
- [210] D.S. Vinay, E.P. Ryan, G. Pawelec, W.H. Talib, J. Stagg, E. Elkord, T. Lichter, W.K. Decker, R.L. Whelan, H.M.C. Shantha Kumara, E. Signori, K. Honoki, A.G. Georgakilas, A. Amin, W.G. Helderich, C.S. Boosani, G. Guha, M.R. Ciriolo, S. Chen, S.I. Mohammed, A.S. Azmi, W.N. Keith, A. Bilsland, D. Bhakta, D. Halicka, H. Fujii, K. Aquilano, S. Salman Ashraf, S. Nowshheen, X. Yang, B.K. Choi, B.S. Kwon, Immune evasion in cancer: Mechanistic basis and therapeutic strategies, *Seminars Cancer Biol.* 35 (2015) 5185–5198.
- [211] Thilo F. Brauß, Sofia Winslow, Sebastian Lampe, Anica Scholz, Andreas Weigert, Nathalie Dehne, Kristoffer von Stedingk, Tobias Schmid, Bernhard Brüne, The RNA-binding protein HuR inhibits expression of CCL5 and limits recruitment of macrophages into tumors, *Mol. Carcinogen.* 56 (12) (2017) 2620–2629.
- [212] Feng-Yen Lin, Yung-Hsiang Chen, Yi-Wen Lin, Jen-Sung Tsai, Jaw-Wen Chen, Hsiao-Jung Wang, Yuh-Lien Chen, Chi-Yuan Li, Shing-Jong Lin, The role of human antigen R, an RNA-binding protein, in mediating the stabilization of

- Toll-like receptor 4 mRNA Induced by endotoxin, *Arterioscler. Tromb. Vasc. Biol.* 26 (12) (2006) 2622–2629.
- [213] D. Kontoyiannis, M. Pasparakis, T.T. Pizarro, F. Cominelli, G. Kollias, Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies, *Immunity* 10 (1999) 387–398.
- [214] Hoanh Tran, Fabienne Maurer, Yoshikuni Nagamine, Stabilization of urokinase and urokinase receptor mRNAs by HuR Is linked to its cytoplasmic accumulation induced by activated mitogen-activated protein kinase-activated protein kinase 2, *Mol. Cell. Biol.* 23 (20) (2003) 7177–7188.
- [215] E.A. Suswam, L.B. Nabors, Y. Huang, X. Yang, P.H. King, IL-1 induces stabilization of IL-8 mRNA in malignant breast cancer cells via the 3' untranslated region: Involvement of divergent RNA-binding factors HuR, KSRP and TIAR, *Int. J. Cancer* 113 (2005) 911–919.
- [216] Hye Jin Choi, Hyun Yang, Seong Hwan Park, Yuseok Moon, HuR/ELAVL1 RNA binding protein modulates interleukin-8 induction by muco-active ribotoxin deoxyvalenol, *Toxicol. Appl. Pharmacol.* 240 (1) (2009) 46–54.
- [217] José A. González-Feliciano, Marimar Hernández-Pérez, Luis A. Estrella, Daisy D. Colón-López, Armando López, Marina Martínez, Kirla R. Maurás-Rivera, Clarivel Lasalde, Daviana Martínez, Félix Araujo-Pérez, Carlos I. González, Carol J. Wilusz, The role of HuR in the post-transcriptional regulation of interleukin-3 in T cells, *PLoS ONE* 9 (3) (2014) e92457.
- [218] B. Herdy, T. Karonitsch, G.I. Vladimer, C.S.H. Tan, A. Stukalov, C. Trefzer, J.W. Bigenzahn, T. Theil, J. Holinka, H.P. Kiener, J. Colinge, K.L. Bennett, G. Superti-Furga, The RNA-binding protein HuR/ELAVL1 regulates IFN- β mRNA abundance and the type 1 IFN response, *Eur. J. Immunol.* 45 (2015) 1500–1511.
- [219] K. Ouhara, S. Munenaga, M. Kajiya, K. Takeda, S. Matsuda, Y. Sato, Y. Hamamoto, T. Iwata, S. Yamasaki, K. Akutagawa, N. Mizuno, T. Fujita, E. Sugiyama, H. Kurihara, The induced RNA-binding protein, HuR, targets 3'-UTR region of IL-6 mRNA and enhances its stabilization in periodontitis, *Clin. Exp. Immunol.* 192 (2018) 325–336.
- [220] T. Matsumiya, K. Ota, T. Imaizumi, H. Yoshida, H. Kimura, K. Satoh, Characterization of synergistic induction of CX3CL1/ fractalkine by TNF- α and IFN- γ in vascular endothelial cells: an essential role for TNF- α in post-transcriptional regulation of CX3CL1, *J. Immunol.* 184 (2010), 4205–4142.
- [221] Liang Lu, Lei Zheng, Ying Si, Wenyi Luo, Gwendal Dujardin, Thaddaeus Kwan, Nicholas R. Potochick, Sunnie R. Thompson, David A. Schneider, Peter H. King, Hu antigen R (HuR) is a positive regulator of the RNA-binding proteins TDP-43 and FUS/TLS, *J. Biol. Chem.* 289 (46) (2014) 31792–31804.
- [222] V. Sedlyarov, J. Fallmann, F. Ebner, J. Huemer, L. Sneezum, M. Ivin, K. Kreiner, A. Tanzer, C. Vogl, I. Hofacker, P. Kovarik, Tristetraprolin binding site atlas in the macrophage transcriptome reveals a switch for inflammation resolution, *Mol. Syst. Biol.* 12 (2016) 868.
- [223] C. Tiedje, M.D. Diaz-Munoz, P. Trulley, H. Ahlfors, K. Laaß, P.J. Blakeshear, M. Turner, M. Gaestel, The RNA-binding protein TTP is a global post-transcriptional regulator of feedback control in inflammation, *Nucleic Acids Res.* 44 (2016) 7418–7440.
- [224] Avijit Goswami, Kamalika Mukherjee, Anup Mazumder, Satarupa Ganguly, Ishita Mukherjee, Saikat Chakrabarti, Syamal Roy, Shyam Sundar, Krishnananda Chattopadhyay, Suvendra N Bhattacharyya, MicroRNA exporter HuR clears the internalized pathogens by promoting pro-inflammatory response in infected macrophages, *EMBO Mol. Med.* 12 (3) (2020), <https://doi.org/10.15252/emmm.201911011>.
- [225] E. Christodoulou-Vafeiadou, F. Ioakeimidis, M. Andreadou, G. Giagkas, G. Stamatakis, M. Reczko, M. Samiotaki, A.D. Papanastasiou, I. Karakasioti, D.L. Kontoyiannis, Divergent innate and epithelial functions of the RNA-binding protein HuR in intestinal inflammation, *Front. Immunol.* 9 (2018) 2732.
- [226] Marieke de Graaf, Janet de Beukelaar, Jelmer Bergsma, Jaco Kraan, Martin van den Bent, Markus Klimek, Yvette van Norden, Arinardi Kusuma, Peter Sillevius Smitt, Jan Willem Gratama, B and T cell imbalances in CSF of patients with Hu-antibody associated PNS, *J. Neuroimmunol.* 195 (1-2) (2008) 164–170.
- [227] Tomasz Herjan, Peng Yao, Wen Qian, Xiao Li, Caini Liu, Katarzyna Bulek, Dongxu Sun, Wen-Pin Yang, Jun Zhu, Aiqing He, Julie A. Carman, Serpil C. Erzurum, Howard D. Lipshitz, Paul L. Fox, Thomas A. Hamilton, Xiaoxia Li, HuR is required for IL-17-induced Act1-mediated CXCL1 and CXCL5 mRNA stabilization, *J. Immunol.* 191 (2) (2013) 640–649.
- [228] Jing Chen, Jennifer L. Martindale, Carole Cramer, Myriam Gorospe, Ullus Atasoy, Paul D. Drew, Shiguang Yu, The RNA-binding protein HuR contributes to neuroinflammation by promoting C-C chemokine receptor 6 (CCR6) expression on Th17 cells, *J. Biol. Chem.* 292 (35) (2017) 14532–14543.
- [229] Noémie Legrand, Dan A Dixon, Cyril Sobolewski, AU-rich element-binding proteins in colorectal cancer, *World, J. Gastrointest. Oncol.* 11 (2) (2019) 71–90.
- [230] S. Liu, Z. Huang, A. Tang, X. Wu, J. Aube, L. Xu, C. Xing, Y. Huang, Inhibition of RNA-binding protein HuR reduces glomerulosclerosis in experimental nephritis, *Clinical Sci.* 134 (2020) 1433–1448.

2. Aims and topic of this Thesis

This PhD thesis aims to investigate the key residues involved in the interaction between selected RNA binding proteins (HuR and human Musashi-1) containing single or multiple RRM domains and specific RNA strands, in order to design novel chimeric proteins for possible biotechnological and therapeutic applications. In particular, it focuses on the expression, purification, and characterization of the selected RRM-containing proteins and the study of the RNA-protein binding.

To carry out this research, biophysical and structural biology techniques like Nuclear Magnetic Resonance (NMR) and Size Exclusion Chromatography with Multi Angle Light Scattering (SEC-MALS) were used. Additionally, Surface Plasmon Resonance (SPR) was used to deepen on the kinetics of the interaction.

In detail, the specific objectives of this thesis related to each RRM-containing protein are described below:

- Human Musashi-1 (MSI-1) protein:
 - o Expression and characterization of both the isolated and tandem domains of MSI-1
 - o Identification of the key residues involved in the interaction with ssRNA targets
 - o Understanding how the binding between MSI-1 and RNA is affected by RNA secondary structures (hairpins)
 - o Modulation of the RNA-protein interaction to gain or lose specificity and affinity by proposing and performing specific single point mutations (both on the RNA and on the SI-1 side) and evaluate its effect on the binding.

- Bacterial HuR protein:
 - o Expression and characterization of the HuR isolated RRM domain coded in the *Acinetobacter Baumannii* genome
 - o Investigation of its RNA binding activities by solution NMR.

Furthermore, this PhD thesis also focuses on the development of a new application of the LigandTracer technology, a highly flexible platform that allows you to perform a broad range of interaction studies. This technique is generally used to monitor molecular interactions in real-time on different targets, usually cellular receptors expressed on the surface of living cells. We found that this technology has huge potential to be slightly modified and could be exploited for the study of protein-RNA interaction. Hence, we aimed to develop a new application to monitor protein-RNA interaction over time in bacterial cells to study the Musashi-1 protein in *E. coli* cells.

Structural biology is emerging as a key component to design new biotherapeutics with improved pharmacological activity and safety. In this regard, the role of NMR spectroscopy is particularly relevant

for the development of biologics because of its potential applications for the structural characterization of biomolecules, in protein-protein and protein-ligand interaction studies, and for the development of suitable formulations for protein drug delivery. The development of suitable formulations is obviously a key step for the real use of modified RBP and or RBP-RNA complexes in living systems. However, the characterization of formulations for protein drug delivery is still a challenging problem and a new field also for NMR spectroscopists. Therefore, during this research project efforts were devoted also to develop new strategies based on the use of NMR spectroscopy to characterize protein-based drugs and formulations with a view of using modified RBP. Two proteins Human transthyretin (TTR) protein, a physiological protein acting as a hormone carrier and the therapeutic protein L-Asparaginase-II (ANSII) from *E. coli*, approved against acute lymphoblastic leukaemia, already available in our laboratory have been selected to speed-up the development of these new NMR-based strategies. In detail, the activities have been focused on the NMR characterization of proteins embedded in hydrogels and protein-drug conjugates.

3. Methodological aspects

3.1 PROTEIN EXPRESSION AND PURIFICATION:

Protein production is a biotechnological process that leads to the expression and isolation of a recombinant protein by introducing a recombinant DNA plasmid encoding for the target protein into the host. After the insertion of the plasmid into the expression system, an over expression of the chosen protein is performed and followed by its isolation from the rest of the proteosome.

Commonly used hosts for protein expression are bacteria, yeast, insect, plants, and mammalian cells. All the proteins used in this PhD thesis have been expressed in *E. coli* cells.

In order to obtain a high yield of a recombinant protein, proper selection and manipulation of the expression conditions is essential. The main variables that can influence the expression are the host strain, growth medium and expression parameters such as the temperature, the expression promoter concentration (usually isopropyl β -D-thiogalactoside, IPTG) and induction time.

To select the best conditions to obtain the desired protein, a preliminary expression test in a small volume scale is performed, where different conditions are tested in order to evaluate them and choose the optimal one.

Expression results are checked on a SDS polyacrylamide gel (SDS-PAGE), an approach that allows us to compare the different expression conditions tested and evaluate which ones gives as the highest yield of soluble protein before scaling-up to a culture in a bigger volume.

In case the desired protein is expressed in the inclusion bodies as an insoluble protein, it is also possible to test multiple fusion tags, the kind of vector carrying the DNA and the expression promoters in order to try to find some conditions that allow the soluble expression of the protein. If the main fraction of the protein is still insoluble, another approach could be to try an *in vitro* re-folding screening.

As already mentioned, also different media can be used for cell growth, but this mainly depends by the assay we want to do on the protein. Usually, the cultures are performed in a so-called “rich medium” such as Luria Bertani (LB) medium not labelled, using this media is possible to get proteins, usually in good amount, that can be used with almost all the techniques. If there is the need of isotopically enriched proteins, chemically defined and isotopically enriched medium such as M9, a “minimal medium”, are used. Using minimal medium for the expression of isotopically enriched samples is easier since it's enough to supply the medium with enriched nitrogen and carbon sources in the form of $^{15}\text{NH}_4\text{Cl}$ and ^{13}C Glucose, respectively instead of the unlabelled one.

Once the most favourable conditions for the expression of the desired protein are obtained, and the protein has been expressed, the protein isolation from the whole host proteosome is required in order to work with the pure protein. A pure protein sample is essential for the characterization of the structure, function, and interactions of the protein of interest.

In theory, it is possible to use many purification strategies to get a pure sample, but due to the fact that each purification step, or in general each operation on a protein sample, usually results, in some protein loss, there is always the need to balance final yield and purity.

The purification process usually takes advantage of differences in size, physicochemical properties, binding affinity and biological activity between the desired protein and the rest of the proteosome, choosing an inadequate methodology, can affect the integrity and activity of the protein, or eventually also the final purity of the sample, which might result in its degradation or precipitation.

The preliminary step in protein purification is the disruption of the cells, or cell lysis. This can be achieved by several techniques, some of the most commonly used being sonication, osmotic lysis, homogenization by high pressure (French press), repeated freeze-thaw cycle lysis or detergent lysis. Afterwards, cell debris are removed by centrifugation.

The location of the expressed protein within the bacterial host will affect the approach towards its isolation and purification. In fact, bacterial host may express a cytosolic protein or store it as insoluble inclusion bodies. When working with an insoluble protein, extra steps for the extraction of the protein from the inclusion bodies and its re-folding might be needed.

The subsequent steps of purification involve several chromatographic techniques. In order to simplify the purification process and increase protein purity, affinity purification tags can be fused to the target protein. The most used chromatographic techniques are used in this work have been Ion Exchange Chromatography (IEX), size Exclusion Chromatography (SEC) and affinity Chromatography, such as Strep-Tag affinity Chromatography and Ni²⁺-affinity Chromatography.

IEX involved the separation of charged biological molecules based on their total charge. The separation is based on the reversible interaction between a charged protein and an oppositely charged chromatographic resin. Elution is performed by changing the ionic strength of the elution buffer, or by changing the pH.

SEC is a chromatographic technique that separated proteins based on their size, or better, their hydrodynamic radius. The column matrix is composed by a variety of beads with slightly different pore sizes. When the dissolved molecules flow into the column, the larger ones migrate quickly whereas the small ones fall into the pores and migrate slowly.

Affinity Chromatography is one of the most used techniques. It separates proteins based on the reversible interaction between the sidechains of specific amino acids and chromatographic matrix. Strep-Tag affinity chromatography can be used to separate recombinant proteins that contain a fused Strep-tag, an eight-residue minimal peptide sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) that exhibits intrinsic affinity towards Strep-Tactin, an engineered streptavidin attached to the resin of the column. Elution is performed using a buffer with desthiobiotin which competes with the Strep-tag for the Strep-Tactin.

Ni²⁺-affinity Chromatography can be used instead to separate recombinant proteins that contain a fused His-tag (a sequence of multiple His, usually between 4 and 8). In this case, proteins with a fused His-

Tag binds Ni²⁺ ions attached to the resin of the column. Elution is usually performed by using increasing concentrations of imidazole as a competitive agent.

3.1.1 Site directed mutagenesis:

Site-directed mutagenesis was used to obtain the construct of MSI-1 RRM-1 and all mutants of MSI-1 RRM₁₋₂ and MSI-1 RRM-2 genes. Mutations were performed using the QuickChange Mutagenesis Kit (Agilent Technologies). All mutations were confirmed by DNA sequencing. Amino acid substitutions were confirmed through ¹H¹⁵N heteronuclear single-quantum coherence (HSQC) NMR spectra.

3.1.2 UV-Visible spectroscopy (UV-Vis):

Proteins concentrations were estimated using UV-Visible spectroscopy (UV-Vis) by measuring the absorbance at 280 nm (A₂₈₀).

UV-Vis is an analytical technique used to measure the absorption of electromagnetic radiation by molecules in the ultraviolet-visible region of the electromagnetic spectrum and is commonly used to quantify the concentration of proteins and nucleic acids. Protein concentration can be in fact determined by its absorption around 280 nm using the Beer-Lambert equation. Theoretical molar extinction coefficients were computationally estimated using the ProtParam software from ExPASy (<http://ca.expasy.org/tools/protparam.html>).

UV-Vis can be also used to determine protein contamination by RNA/DNA by observing a shift of the main peak absorption around from 280nm to around 260nm or by the simultaneous presence of two peaks (one at 280nm and one at 260nm).

3.1.3 Poly (ethyleneimine) (PEI) treatment:

Is quite common to find nucleic acids produced in *E. coli* associated to an RNA binding protein (RBPs). This can be observed by UV-visible spectroscopy (UV-Vis) of the purified protein; if some nucleic acids are still present, the maximum of the UV absorption is in fact shifted from 280nm (typical of a free RNA/DNA bound protein) to 260 nm instead.

To avoid having intrinsic nucleic acids attached to the protein while performing further RNA-Protein binding experiments, an additional step during the purification of the protein is implemented.

Precipitation of nucleic acids in *E. coli* with Poly (ethyleneimine) has been investigated.⁶²

PEI (Sigma Aldrich) as a 5 % (w/v) solution, pH 7.9, is added slowly, under stirring to the cell extract and then the resulting suspension is centrifuged at 7500 rpm for 30 min. The following equation allow to calculate the proper amount in ml of PEI needed based on the cell extract volume:

$$(V_{cell\ extract} + x) \cdot 0.8 = PEI\% \cdot x$$

The pellet is discarded, while the supernatant is treated with solid ammonium sulphate, which is added slowly under stirring at 4 °C until reaching 70% (w/v) to eliminate the excess of PEI in solution. The precipitate, containing the protein, is collected by centrifugation and washed several times (around 4-5 times) with 30-60 ml of a 70% (w/v) solution of ammonium sulphate. After each washing step, the precipitate containing the protein is collected by centrifugation and re-suspended in 70% (w/v) ammonium sulphate solution. Finalized the washing, protein precipitate is re-suspended in the desired buffer.

Successful precipitation of nucleic acids and purification of the protein is confirmed by (UV-Vis).

3.1.4 Preparation of RNase free water:

To prepare Nuclease free water, 1 L of milliQ water is treated with 1 ml of Diethylpyrocarbonate (DEPC) (to a final concentration 0.1%) and mix thoroughly. The DEPC-mixed water is incubated for 12h at 37°C and further autoclaved, for 45 min or 3 x 15 min at 121°C, to remove the residual of DEPC.

3.1.5 AB-Elavl protein expression and purification protocol

Recombinant AB-Elavl (rAB-Elavl) protein encoded in plasmid pET-30a(+) was overexpressed in BL21(DE3) GOLD *E. coli* cells. Cells were grown in LB or M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ or $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose, at 37 °C until optical density (OD_{600}) reached 0.6–0.8. Subsequently, protein expression was induced with 0.2 mmol dm^{-3} of isopropyl β -D-thiogalactoside (IPTG), cells were incubated at 18 °C overnight and harvested by centrifugation at 4 °C, 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 , Proteases Inhibitor Cocktail), ruptured by sonication and separated by centrifugation at 30,000 rpm for 35 min at 4 °C. Soluble fraction was collected and a treatment with 5% PEI solution was performed to remove DNA/RNA attached to the protein. The protein was re-suspended in the lysis buffer. Soluble protein was filtered with a 0.22 μm membrane and purified by a Ni^{2+} -affinity chromatography using a His-Trap HP 5 cm^3 column previously equilibrated in 50 mmol dm^{-3} HEPES pH 6.8, 300 mmol dm^{-3} NaCl, 3 mmol dm^{-3} MgCl_2 , Proteases Inhibitor Cocktail. rAB-Elavl was eluted with increasing imidazole concentration (20–50–100–250 mmol dm^{-3}) in the buffer and subsequently dialyzed overnight against 4 dm^3 of 20 mmol dm^{-3} HEPES buffer at pH 6.8, containing 150 mmol dm^{-3}

NaCl and 3 mmol dm⁻³ MgCl₂. 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 MgCl₂ and Proteases Inhibitor Cocktail.

3.1.6 Crystallization of rAB-Elavl

rAB-Elavl was concentrated up to 6 mg/mL in 20 mmol dm⁻³ HEPES buffer pH 6.8, containing 150 mmol dm⁻³ NaCl, 3 mmol dm⁻³ MgCl₂ and Proteases Inhibitor Cocktail. Crystals diffracting at 1.6 Å were obtained by sitting drop vapor diffusion at 293 K, in which 5 µL of protein solution were mixed with 5 µL of reservoir solution and suspended over 600 µL of the same reservoir solution. 0.1 mol dm⁻³ sodium acetate trihydrate pH 4.5, 3 mol dm⁻³ NaCl).

3.1.7 Human Musashi-1 (MSI-1) (1-200) RRM₁₋₂ expression and purification protocol

Recombinant human MSI-1 RRM₁₋₂ (1-200) protein in plasmid pET29b was overexpressed in BL21(DE3) GOLD *E. coli* cells. Cells were grown in LB, or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose, at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Subsequently, expression was induced with 0.5 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 37 °C for 4 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer 50 mmol dm⁻³ Tris-HCl, pH 8.0, 500 mmol dm⁻³ NaCl, 10% glycerol, Proteases Inhibitor Cocktail, ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was collected and a treatment with 5% PEI solution was performed in order to remove DNA/RNA attached to the protein. Protein was re-suspended in 20 mmol dm⁻³ Tris-HCl, pH 9.0, Proteases Inhibitor Cocktail.

Soluble protein was filtered with a 0.22 µm membrane and purified by an ion exchange chromatography using an Anion exchange Q FF 16/10 column previously equilibrated in 20 mmol dm⁻³ Tris-HCl, pH 9.0, Proteases Inhibitor Cocktail. Elution was performed against the buffer 20 mmol dm⁻³ Tris-HCl, pH 9.0, 1 mol dm⁻³ NaCl, Proteases Inhibitor Cocktail. 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 50 mmol dm⁻³ Tris-HCl, pH 7.8, 150 mmol dm⁻³ NaCl, Proteases Inhibitor Cocktail. The purified fractions were collected, and buffer exchange chromatography was performed using a HiPrep 26/10 Desalting column previously equilibrated with the RNase free buffer for the NMR assignment experiments [20 mmol dm⁻³ MES, pH 6.0, 100 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] or the RNase free buffer for NMR titration experiments [50 mmol dm⁻³ Tris-HCl, pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail].

3.1.8 Human MSI-1 RRM-1 (1-103) domain expression and purification protocol

To produce the MSI-1 human RRM-1 construct (1-103), a pET29b plasmid containing a Strep-Tag at N-terminus followed by the first 261 residues of Musashi-1 human protein was modified by Site-Directed Mutagenesis to replace Met-104 with a stop codon.

Recombinant human MSI-1 RRM-1 (1-103) protein in plasmid pET29b was overexpressed in BL21(DE3) *E. coli* cells. Cells were grown in LB, or M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ or $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose, at 37 °C until optical density (OD_{600}) reached 0.6-0.8. Subsequently, expression was induced with 0.5 mmol dm⁻³ of isopropyl β -D-thiogalactoside (IPTG), cells were incubated at 37 °C for 3 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer 100 mmol dm⁻³ Tris-HCl pH 8.0, 150 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail, ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was filtered with a 0.22 μm membrane and purified using a Strep-Tag HP 5 ml column. Eluted fractions containing the protein were collected and purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column that was previously equilibrated in 50 mmol dm⁻³ Tris-HCl pH 7.5, 150 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail. The purified fractions were collected, 1 mmol dm⁻³ DTT was added to the solution and treatment with a 5% PEI solution was performed in order to remove DNA/RNA attached to the protein. Protein was re-suspended in RNase free final buffer [50 mmol dm⁻³ Tris-HCl, pH 7.5, 140 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] and filtered with a 0.22 μm membrane.

3.1.9 Human MSI-1 RRM-2 (104-200) domain expression and purification protocol

Recombinant human MSI-1 RRM-2 (104-200) protein in plasmid pET21a containing a Strep-Tag at N-terminus was overexpressed in BL21(DE3) *E. coli* cells. Cells were grown in LB, or M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ or $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose, at 37 °C until optical density (OD_{600}) reached 0.6-0.8. Subsequently, expression was induced with 0.5 mmol dm⁻³ of isopropyl β -D-thiogalactoside (IPTG), cells were incubated at 37 °C for 4 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer [100 mmol dm⁻³ Tris-HCl pH 8.0, 150 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail], ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was filtered with a 0.22 μm membrane and purified using a Strep-Tag HP 5 ml column. Elution fractions containing the protein were collected and purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column that was previously equilibrated in 50 mmol dm⁻³ Tris-HCl pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail. The purified fractions were collected and treatment

with a 5% PEI solution was performed in order to remove DNA/RNA attached to the protein. The protein was re-suspended in RNase free final buffer [50 mmol dm⁻³ Tris-HCl, pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] and filtered with a 0.22 µm membrane.

3.1.10 Human MSI-1 mutants' expression and purification protocol

Site-directed mutagenesis was performed on the genes coding for the MSI-1 (1-200) RRM₁₋₂ WT and the MSI-1 RRM-2 (104-200) to replace either the E180 by an asparagine, the K182 by a methionine, or both, to obtain the constructs MSI-1 RRM₁₋₂ E180N, MSI-1 RRM₁₋₂ K182M and MSI-1 RRM₁₋₂ K182M E180N, MSI-1 RRM-2 E180N, MSI-1 RRM-2 K182M and MSI-1 RRM-2 K182M E180N respectively. Expression and purification of all mutants have been performed as previously described in *section 3.1.7* for the mutants on the tandem domain protein and in *section 3.1.9* for the ones of the second recognition motif.

3.1.11 L-Asparaginase II (ANSII) expression and purification protocol.

Recombinant L-Asparaginase II (ANSII) protein in plasmid pET21a(+) was overexpressed in C41(DE2) *E. coli* cells. Cells were grown in LB, or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose. Different expression approaches were used for the expression of different isotopic labelled ANSII.

For the expression of natural abundance (NA) ANSII, cells were grown at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Expression was induced with 0.75 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 25 °C overnight and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm.

For the expression of [U-¹³C, ¹⁵N] ANSII, a Marley-like method was applied. Cells were grown at 37°C overnight in LB and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Pellet was re-suspended in 1L M9 minimal medium and incubated at 37 °C for 1h. Expression was induced with 0.75 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 25 °C for 5 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm.

Despite the differences in the expression protocol, all samples of ANSII were extracted and purified in the same way. Cell pellet was resuspended in lysis buffer [10 mmol dm⁻³ Tris-HCL, pH 8.0, 15 mmol dm⁻³ EDTA, 20% (m/v) sucrose buffer (60 mL per litre of culture)]. The suspension was stirred at 4 °C for 20 min, then centrifuged at 10000 rpm for 30 min. The supernatant was discarded, whereas the pellet was redissolved in H₂O milli-Q (30-60 ml per litre of culture) and incubated at 4 °C for 20 min under stirring. The mixture was again centrifuged at 10000 rpm for 30-60 min. The pellet was discarded, while the supernatant was treated with solid ammonium sulphate, which was added slowly under stirring at 4 °C until reaching 50% of saturation. The precipitate was collected by centrifugation and then discarded,

whereas the supernatant was similarly treated with solid ammonium sulphate until reaching 90% of saturation. This time, the supernatant was discarded, while the precipitate, containing the protein, was redissolved in a minimal amount of 20 mmol dm⁻³ Tris-HCl, pH 8.6 buffer. The solution was dialyzed against 4 L of 20 mmol dm⁻³ Tris-HCl, pH 8.6 buffer, which was refreshed once after a night to remove the excess of ammonium sulphate that might still be present.

ANSII was further purified by anionic exchange chromatography using an Anion exchange Q FF 16/10 column previously equilibrated in 20 mmol dm⁻³ Tris-HCl, pH 8.6 buffer. Elution was performed with a 0-300 mmol dm⁻³ NaCl gradient in 15 CV. An SDS-PAGE was performed to identify which fractions contained the protein. The protein was further purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column previously equilibrated in the final buffer [150 mmol dm⁻³ NaPi, pH 7.5]

3.1.12 Transthyretin (TTR) expression and purification protocol

Recombinant human transthyretin (TTR) protein in plasmid pET-28a was overexpressed in BL21(DE3) Codon Plus RIPL *E. coli* cells. Cells were grown in LB, or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose, at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Expression was induced with 1 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 37 °C overnight and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer [20 mmol dm⁻³ Tris-HCl, pH 8.6, 5 mmol dm⁻³ DTT, Proteases Inhibitor Cocktail], ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was collected and purified by an ion exchange chromatography using an Anion exchange Q FF 16/10 column previously equilibrated in the lysis buffer. Elution was performed with increasing gradients of NaCl (60 mL of 0-200 mmol dm⁻³, 300 mL of 200-500 mmol dm⁻³, 60 mL of 500-1000 mmol dm⁻³ and 60 mL of 1 mol dm⁻³) in the buffer.

An SDS-PAGE was performed to identify which fractions contained the protein. The protein was further purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column previously equilibrated in the final buffer [50 mmol dm⁻³ MES-NaOH, pH 6.5, 100 mmol dm⁻³ NaCl, 5 mmol dm⁻³ DTT, Proteases Inhibitor Cocktail].

3.2 PROTEIN CHARACTERIZATION:

3.2.1 Nuclear Magnetic Resonance (NMR) spectroscopy:

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the main techniques used for structural characterization in chemistry, biology, and material sciences. NMR is based upon the concept that many

atomic nuclei contain a spin, or “inherent angular momentum”, and all nuclei are electrically charged. In presence of an external magnetic field, different spin states with different energy exist, and hence an energy transfer is possible between the different energy levels. This energy difference between two spin states depends on the gyromagnetic moment (γ) that is unique for each nucleus. Hence, each NMR active nucleus gives rise to an individual signal in the spectrum. When working with molecules containing hundreds of atoms such as proteins, they require multi-dimensional NMR experiments to be resolved.⁶³

Developments in NMR methodology and hardware (cryoprobes, magnetic field...) together with developments in the sample preparation and labelling of recombinant proteins have radically improved the use of NMR for the characterization of biological macromolecules.

In this PhD thesis, ^1H 1D and ^1H - ^{15}N HSQC 2D NMR experiments were mainly used in order to characterize the proteins. ^1H - ^{15}N - ^{13}C 3D NMR experiments have been acquired to solve the assignment. The ^1H 1D NMR is a mono-dimensional spectra that it is based on the principle that protons in different chemical environments give signals at a different position in the spectra. It is performed as a very first step of any NMR-based study and although gives the first wealth of information on the sample, such as if your protein is folded or not, in case of biological macromolecules, that contain hundreds or thousands of resonances, ^1H spectra are not enough to describe the system and multidimensional NMR spectra can be used to provide complementary information.⁶⁴

Heteronuclear Single Quantum Coherence (^1H - ^{15}N HSQC) is a bi-dimensional NMR spectra used to correlate two different types of nuclei, generally ^1H with ^{15}N . Each amino acid of the protein (except for the proline and the last residue at the C-terminal) has an amide proton attached to a nitrogen within the peptide bond where the proton with the ^{15}N labelled nitrogen of the amide group can be correlated.⁶⁴

If the protein is folded, peaks are well dispersed in the proton/nitrogen dimensions and most of them can be distinguished. The number of peaks in the spectra should match the number of residues present in the protein with an addition of some extra peaks corresponding to several side chains with nitrogen-bound protons.

In order to assign each peak of the ^1H - ^{15}N HSQC to a specific residue, complementary three-dimensional spectra have to be acquired. That is to find out the chemical shift that corresponds to each atom. Three-dimensional spectra of ^{15}N - ^{13}C labelled proteins transfer magnetization over the peptide bond, connecting different spin systems through bonds. Combining different 3D spectra gives complementary information used to solve the assignment. The most commonly used ones are 3D HN(CO)CA, 3D HNCA, 3D HNCO, 3D HN(CA)CO, 3D CBCA(CO)NH and 3D HNCACB spectra.

3.2.2 X-ray Crystallography:

The crystallization of proteins depends on the creation of a solution that is supersaturated in the macromolecule but exhibits conditions that do not significantly perturb its natural state. Protein crystallization can be obtained using the hanging drop or sitting drop techniques, however, obtaining a crystal of a protein can be difficult to achieve due to many factors that can affect this process, such as purity of the sample, pH, temperature, protein concentration, protein dynamics, presence of substrates and co-factors, nature of the crystallizing agent and many other minor variables such as light, presence of oxygen or gravity acceleration.

Nevertheless, when a crystal of a protein is obtained, X-ray crystallography can be used to obtain a three-dimensional molecular structure from the crystal. Therefore, it can help to determine the structure of a protein.

3.3 PROTEIN – RNA INTERACTION:

3.3.1 NMR spectroscopy to study of protein -RNA interactions:

NMR is a powerful tool to probe and target molecular interactions and to investigate the behaviour of a number of biological systems. Interactions change the physical and chemical properties of molecules, and NMR can detect these subtle perturbations upon the complex formation. During the complex formation of a protein-RNA complex, significant structural modifications caused by the complex formation alter the chemical environment of the amino acids of the protein that are involved in the complex-formation domain. The residues that are mainly involved in complex formation will experience the most environmental changes, that will correspond to a bigger change in the NMR signals, hence is possible to record those changes and use them to map the interaction interface of the protein.⁶⁵

NMR titrations with a combination of ^1H 1D and ^1H - ^{15}N HSQC spectra or similar bi-dimensional spectra is one of the most commonly used approach to study these interactions. These experiments start with a mono and bi-dimensional spectra of the protein in its free-form. Then, further spectra are acquired after each addition of increasing amounts of ligand, in this case RNA until reaching a 1:1, 1:2 or even higher molar ratio of protein/RNA. The acquisition of these spectra allows to follow the formation of the complex and to track the changes of each peak in the bi-dimensional spectra.

Through the analysis of the Intensity decrease and Chemical shift perturbation (CSP) of the peaks, these experiments allow to detect the binding area and therefore, the residues of the protein affected by the interaction and complex formation.

3.3.2 Size Exclusion Chromatography with Multi Angle Light Scattering (SEC-MALS):

Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) can be used to obtain a deeper insight into the stoichiometry of the complexes.

Size exclusion chromatography (SEC) may be used together with standard concentration detectors, such as UV absorbance, differential refractive index (dRI) and/or light scattering detectors, in order to monitor the elution process. UV detector measures the protein or complex concentration at a wavelength of 260 or 280 nm. dRI detector determines the concentration based on the change in the refractive index in solution due to the presence of an analyte. Light scattering detectors such as multi-angle light scattering (MALS) can enhance the analytical information by determining accurately the molecular weight of soluble and insoluble protein aggregates.⁶⁶

SEC-MALS in structural biology is widely used to investigate the stoichiometry of isolated biomolecules and complexes. It can determine the MW of all types of complexes, non-globular or inherently disordered and heterocomplexes including protein-protein and protein-nucleic acid complexes.

For protein-nucleic acid complexes, a specific analysis is performed, using the Protein Conjugate function. This method is frequently used when both components in the complex differ in either the UV coefficient (ϵ) or the dRI. It combines MALS, UV and dRI signals to quantify both component systems. The addition of MALS data allows to determine the molar mass of each component of the complex (protein and modifier) as well as the entire complex.

3.3.3 Ligand Tracer for monitoring protein-RNA binding in vivo in real time:

Ligand Tracer (LT) technology has been widely used for real-time cell binding assays (RT-CBA) and typically used to measure the binding kinetics and calculate the affinity of proteins for receptors on living cells that are adhered or tethered to a surface of a Petri dish. We have exploited this technology to design a platform that offers the possibility monitoring protein-RNA binding throughout a reporter system. In our approach, we have a fluorescent readout, Green Fluorescence Protein (GFP) whose repression will be an indicator of the interactions between the protein and its RNA target sequence. Fluorescence was detected through the use of difference detectors. For our experiments we used the BlueGreen (488-535 nm) detector to detect GFP.

Two RBP systems were tested with our method; one based on bacteriophage MS2 coat protein and a second based on Musashi-1 protein. Each system is composed of two plasmids: p15A and pSC101, that are jointly transformed to *E. coli* cells. p15A contains a constitutive promoter with 3 different elements: an RNA binding motif close to sfGFP sequence, an sfGFP-encoding sequence, and an RFP-encoding

sequence. pSC101 contains an inducible plasmid, pLac (inducible by lactose or analogues such as Isopropyl β -D-1-thiogalactopyranoside, IPTG, in this case), encoding a Musashi-1 sequence comprising both RNA-recognition motifs (RRMs) (1-196 residues).

Untreated MultiDishes (MD) 2x2 (1-04-202, Ridgeview Instruments AB) were coated with LB agar. Agar was coated by applying 4 mL LB agar in each of the four sections and letting it solidify at RT for 1 h. Different techniques were used for seeding: uniform seeding of the whole section with 20 μ L of bacteria suspension in LB medium (OD₆₀₀= 0.5); and drop seeding of just the detection area followed by incubation at RT until the drop was completely absorbed by the agar.

3.3.4 Surface Plasmon Resonance (SPR)

Surface Plasmon Resonance is a powerful tool to study the dynamics of RNA-protein interactions and provides both equilibrium and kinetic information about intermolecular interactions.⁶⁷

This biophysical methodology⁶⁷ is based on an incident light that hits a metal (typically gold) surface and is reflected in a specific angle between two layers. At this particular angle, the light is absorbed and is converted to a wave that propagates into the medium on the nonilluminated side of the metal surface. The defined angle, at which resonance occurs, is dependent on the refractive index of the material near the metal surface, when working with a constant wavelength and metal thin surface.⁶⁸

When there is a small change in the reflective index of the sensing medium such as biomolecules interaction, it can be detected. Thus, this technique allows to monitor intermolecular interactions.^{67,68}

In our case, the ligand is coated onto the surface matrix, and the protein is injected in the flowing buffer. The experimental cycle consists of several steps.^{67,68} Firstly, the protein is injected in the buffer solution. Once the protein is close to the RNA surface, binding occurs and it causes a change on the refractive index near the metal surface, which results in a change in the angle at which the light is absorbed. This change is recorded and displayed in the form of a sensogram. This sensogram provides information on the association rate. Afterwards, buffer alone is injected which leads to the dissociation phase; the protein gradually dissociates from the RNA and the absorbed angle returns to its previous value. Surface can be regenerated, and several cycles can be performed increasing the protein concentration.

These real-time interaction analyses allow to study the kinetics of an interaction.

4. Results

4.1 ARTICLE 2:

4.1.1 Identification and Characterization of an RRM-Containing, RNA Binding Protein in Acinetobacter Baumannii

Caterina Ciani ¹, [Anna Pérez-Ràfols](#) ², Isabelle Bonomo ¹, Mariachiara Micaelli ¹, Alfonso Esposito ³, Chiara Zucal ¹, Romina Belli ⁴, Vito Giuseppe D'Agostino ¹, Irene Bianconi ¹, Vito Calderone ², Linda Cerofolini ², Orietta Massidda ¹, Michael Bernard Whalen ⁵, Marco Fragai ^{2,*} and Alessandro Provenzani ^{1,*}

¹ Department of Cellular, Computational and Integrative Biology, DeCIBIO, University of Trento, 38123 Trento, Italy;

² Magnetic Resonance Center (CERM)—Department of Chemistry “Ugo Schiff”, University of Florence, Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP) and Giotto Biotech, S.R.L, Sesto Fiorentino, 50019 Florence, Italy

³ International Centre for Genetic Engineering and Biotechnology, (ICGEB), Padriciano, 99, 34149 Trieste, Italy;

⁴ Proteomics and MS Core Facility, Department of Cellular, Computational and Integrative Biology, DeCIBIO, University of Trento, 38123 Trento, Italy

⁵ Institute of Biophysics (IBF), National Research Council (CNR), FBK Nord, 38123 Trento, Italy

* Correspondence: fragai@cerm.unifi.it (M.F.); alessandro.provenzani@unitn.it (A.P.)

Biomolecules, 2022, 12, 922

DOI: <https://doi.org/10.3390/biom12070922>

Acinetobacter baumannii due to its ability to be highly adaptable in antibiotics is considered as one of the most carbapenem-resistance bacteria. It is a Gam-negative opportunistic pathogen that can cause infections associated with high mortality.

The relevance of RNA binding proteins in this host is far from being elucidated, and so far, just a few examples have been reported. However, to our knowledge, the presence and role of an RRM-containing RBP in *Acinetobacter baumannii* has not yet been reported.

Human HuR is one of the most studied RRM-containing RNA binding protein due to its many implications in diseases (see *section 1.5.2*, and *Article 1 in section 1.7*). However, no information regarding these proteins has been found in *Acinetobacter baumannii*.

With all of this in mind, we aim to uncover the existence of an RRM-containing RNA binding protein from the ELAVL family in *Acinetobacter baumannii* that we call AB-Elavl. Our goal is to identify, express and characterize this RRM protein and demonstrate via biochemical and structural biology techniques its ability to recognize and bind RNA targets.

In order to accomplish this, a similarity search for homologous proteins to HuR protein in *Acinetobacter baumannii* has been used to detect a gene encoding for a RRM domain protein. To prove its expression and translation in *Acinetobacter baumannii* we have employed PCR, Mass Spectrometry and Western blot analysis. Subsequently, the expression in *E. coli* cells and characterization using liquid NMR and X-Ray Crystallography has been used to obtain confirm the presence of a typical RRM structure in our protein. Finally, RNA binding studies have been performed with RNA ElectroMobility Shift (REMSA) assays, liquid NMR, HTRF-FRET and AlphaScreen technology.

In particular, my contribution to this project has been the expression and purification of AB-Elavl RRM-containing protein in *E. coli* cells (unlabelled, ^1H - ^{15}N and ^1H - ^{15}N - ^{13}C), the characterization of the AB-Elavl protein in liquid NMR by solving the assignment and the experimental obtention of protein crystals from which the structure has been solved by other collaborators of the project. Regarding the RNA studies, I have also been involved in all RNA titrations performed with liquid NMR

Article

Identification and Characterization of an RRM-Containing, RNA Binding Protein in *Acinetobacter baumannii*

Caterina Ciani ¹, Anna Pérez-Ràfols ^{2,3}, Isabelle Bonomo ¹, Mariachiara Micaelli ¹, Alfonso Esposito ⁴, Chiara Zucal ¹, Romina Belli ¹, Vito Giuseppe D'Agostino ¹, Irene Bianconi ¹, Vito Calderone ^{2,5}, Linda Cerofolini ^{2,5}, Orietta Massidda ¹, Michael Bernard Whalen ⁶, Marco Fragai ^{2,5,*} and Alessandro Provenzani ^{1,*}

- ¹ Department of Cellular, Computational and Integrative Biology, DeCIBIO, Proteomics and MS Core Facility, University of Trento, 38123 Trento, Italy; caterina.ciani@unitn.it (C.C.); isabelle.bonomo@unitn.it (I.B.); mariachiara.micaelli@unitn.it (M.M.); chiara.zucal@gmail.com (C.Z.); romina.belli@unitn.it (R.B.); vito.dagostino@unitn.it (V.G.D.); irene.bianconi@unitn.it (I.B.); orietta.massidda@unitn.it (O.M.)
- ² Magnetic Resonance Center (CERM), Department of Chemistry “Ugo Schiff”, University of Florence, 50019 Florence, Italy; rafols@giottobiotech.com (A.P.-R.); vito.calderone@unifi.it (V.C.); cerofolini@cerm.unifi.it (L.C.)
- ³ Giotto Biotech s.r.l, Sesto Fiorentino, 50019 Florence, Italy
- ⁴ International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano, 99, 34149 Trieste, Italy; alfonso.esposito@icgeb.org
- ⁵ Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), 50019 Florence, Italy
- ⁶ Institute of Biophysics (IBF), National Research Council (CNR), FBK Nord, 38123 Trento, Italy; mwhalen@fbk.eu
- * Correspondence: fragai@cerm.unifi.it (M.F.); alessandro.provenzani@unitn.it (A.P.)



Citation: Ciani, C.; Pérez-Ràfols, A.; Bonomo, I.; Micaelli, M.; Esposito, A.; Zucal, C.; Belli, R.; D'Agostino, V.G.; Bianconi, I.; Calderone, V.; et al. Identification and Characterization of an RRM-Containing, RNA Binding Protein in *Acinetobacter baumannii*. *Biomolecules* **2022**, *12*, 922. <https://doi.org/10.3390/biom12070922>

Academic Editor: Francesc Rabanal Anglada

Received: 23 May 2022
Accepted: 28 June 2022
Published: 30 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Acinetobacter baumannii* is a Gram-negative pathogen, known to acquire resistance to antibiotics used in the clinic. The RNA-binding proteome of this bacterium is poorly characterized, in particular for what concerns the proteins containing RNA Recognition Motif (RRM). Here, we browsed the *A. baumannii* proteome for homologous proteins to the human HuR(ELAVL1), an RNA binding protein containing three RRM. We identified a unique locus that we called *AB-Elavl*, coding for a protein with a single RRM with an average of 34% identity to the first HuR RRM. We also widen the research to the genomes of all the bacteria, finding 227 entries in 12 bacterial phyla. Notably we observed a partial evolutionary divergence between the RNP1 and RNP2 conserved regions present in the prokaryotes in comparison to the metazoan consensus sequence. We checked the expression at the transcript and protein level, cloned the gene and expressed the recombinant protein. The X-ray and NMR structural characterization of the recombinant AB-Elavl revealed that the protein maintained the typical $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ and three-dimensional organization of eukaryotic RRM. The biochemical analyses showed that, although the RNP1 and RNP2 show differences, it can bind to AU-rich regions like the human HuR, but with less specificity and lower affinity. Therefore, we identified an RRM-containing RNA-binding protein actually expressed in *A. baumannii*.

Keywords: *Acinetobacter baumannii*; RNA recognition motif; ELAVL1

1. Introduction

The RNA Recognition Motif (RRM) is the most diffused RNA-binding domain present in eukaryotes and in proteins is often found in association with other domains (that could be all RRM-like or different). The single RRM domain is characterized by the presence of two consensus sequences: a highly conserved sequence (RNP1) of eight amino acids and a less conserved sequence of six (RNP2) [1–3]. The typical secondary structure consists of a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology and the RNP1 and RNP2 are localized in the internal $\beta_1\beta_3$ strands [4]. The role of the eukaryotic RRM-containing protein is associated with many functions in the cell, as pre-mRNA processing, mRNA stability, translation and

degradation [5]. In prokaryotes, instead, its role is still not totally elucidated, and the RRM domain-containing protein is mainly composed by a single domain of around 90 amino acids [1].

Acinetobacter baumannii (*A. baumannii*), is considered as one of the most concerning carbapenem-resistant bacteria, for its ability to be highly adaptable to antibiotics. *A. baumannii* is a Gram-negative opportunistic pathogen able to cause infections associated with high mortality, thanks to its exchange mobile genetic elements as transposons and plasmids [6,7].

The relevance of RBPs in *A. baumannii* is far from being elucidated, but some examples have been reported: the silencing of Csr(A), a global post-transcriptional regulator responsible for metabolisms of glucose, leads to the impairment of the growing abilities of the bacterium [8,9]; the RNA chaperone Hfq, has an important role as a virulence factor, since its knockout leads to reduced growth rate and stress tolerance [10,11]. Several RBP were described to be overexpressed in resistant strains such are enolase, RNase E and NusA, all involved in mRNA processing and gene expression modulation [12]. However, to the best of our knowledge, the presence and role of RRM-containing RNA-binding protein in *A. baumannii* has not been reported.

One of the most studied RRM-containing RNA-Binding Protein (RBP) families among the eukaryotes is the Elav-like family (Elavl). The Elavl proteins are widely spread across all metazoans and are characterized by a high degree of conservation between different species². In humans there are four paralogous genes (*ELAVL1–4*) that encode for four different proteins, HuR (or HuA), Hel-N1 (or HuB), HuC, and HuD, with different roles and cellular localization; they are constituted of three distinct RRM domains, with the most conserved sequences at the level of RNP1 and RNP2 [13]. ELAV-Like proteins bind specific sequences of RNA called AU-Rich Elements (ARE) characterized by the enriched presence of adenylates and uridylates. AREs can be located either in the intronic regions as well as in coding or non-coding parts of the mature mRNA, and contribute to mRNA splicing, maturation, stabilization, and translation [5,14,15].

Here, we identified an RRM-containing protein in *Acinetobacter baumannii* that we called AB-Elavl, starting from a protein similarity search with the human ELAV-like protein HuR. We cloned the corresponding gene, expressed the encoded protein, and characterized its biochemical function and protein structure.

2. Material and Methods

2.1. Similarity Search for Homologous Proteins to HuR Protein in *A. baumannii*

The human HuR protein sequence (NP_001410.2) was used as a query to search for the most similar protein in the *A. baumannii* genome, using tblastn on the NCBI web server (last accessed date 9 November 2019). 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, F3P16_RS16475) was searched in all *A. baumannii* genomes available using tblastn, and in all other bacterial genomes using the Ortholuge database [16]. The sequences of the bacterial Elav-like proteins were submitted to the MEME-suite tool MEME v5.3.3 [17], 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).

2.2. Cloning and Expression of the rAB-Elavl for Biochemical Characterization

The mRNA of the orthologous of the *A. baumannii* HuR, AB-Elavl, was retro-transcribed into cDNA and the sequence was amplified and inserted into the pET30a(+) vector (GenScript, Piscataway, NJ, USA) by using the forward (5'-CGGC CATATG AACTCAAATGTATA-3') and reverse (5'-ATAT CTCGAG CTCTTCAGCTGCCCTT-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(+)-AB-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 OD₆₀₀ 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 20 mM HEPES pH 7.5, 300 mM NaCl, 3 mM MgCl₂, and 0.5 mg/mL Proteases Inhibitor Cocktail (Leupeptin, Aprotinin and Pepstatin from Sigma Aldrich, St. Louis, MI, USA) and then centrifuged at 16,000 × g for 30 min at 4 °C. The supernatant was incubated with Ni-NTA Agarose beads (Ni-NTA Agarose, Qiagen GmbH, Hilden, Germany) for 2 h at 4 °C. After washing the beads with buffer A (20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM MgCl₂ and 20 mM imidazole), buffer B (as buffer A with 50 mM imidazole) and buffer C (as buffer A with 100 mM imidazole), protein was eluted with buffer D (as buffer A with 250 mM imidazole). The eluted protein was dialyzed against storage buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM MgCl₂, 5% glycerol) and stored at −80 °C [18]. 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.

2.3. Expression and Purification of rAB-Elavl for X-ray and NMR Analysis

Recombinant AB-Elavl (rAB-Elavl) protein encoded in plasmid pET-30a(+) was over-expressed in BL21(DE3) GOLD cells. Cells were grown in LB or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose at 37 °C until optical density (OD₆₀₀) reached 0.6–0.8. Subsequently, protein production was induced with 0.2 mM of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 18 °C overnight and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer (50 mM HEPES, pH 6.8, 300 mM NaCl, 3 mM MgCl₂, Proteases Inhibitor Cocktail), ruptured by sonication and separated by centrifugation at 30,000 rpm for 35 min at 4 °C. Soluble fraction was collected and treatment with 5% PEI solution was performed 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 Ni²⁺-affinity chromatography step using a His-Trap HP 5 cm³ column previously equilibrated in 50 mM HEPES pH 6.8, 300 mM NaCl, 3 mM MgCl₂, Proteases Inhibitor Cocktail. rAB-Elavl was eluted with increasing concentration of imidazole (20–50–100–250 mM) in the buffer and subsequently dialyzed overnight against 4 dm³ of 20 mM HEPES buffer at pH 6.8, containing 150 mM NaCl and 3 mM MgCl₂. 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 mM HEPES pH 6.8, 150 mM NaCl, 3 mM MgCl₂ and Proteases Inhibitor Cocktail.

2.4. Crystallization of rAB-Elavl

rAB-Elavl was concentrated to 6 mg/mL in 20 mM HEPES buffer pH 6.8, containing 150 mM NaCl, 3 mM MgCl₂ and Proteases Inhibitor Cocktail. Crystals diffracting at 1.6 Å were obtained by sitting drop vapor diffusion at 293 K, in which 5 μL of protein solution was mixed with 5 μL of reservoir solution and suspended over 600 μL of the same reservoir solution. The reservoir solution consisted of 0.1 M sodium acetate trihydrate pH 4.5, 3 M sodium chloride.

2.5. X-ray Data Collection and Refinement

The dataset was collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K (Bruker, Billerica, MA, USA); the crystals 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 [19], reduced and scaled using XSCALE [19] and amplitudes were calculated using XDSCONV [19]. The structure was solved using the molecular replacement technique; the model used was obtained through MODELLER [20] by using 1FXL as the template. The successful orientation and translation of the molecule within the crystallographic unit cell was determined with MOLREP [21]. The refinement was carried out using PHENIX [22], applying TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT [23]. The quality of the refined structures was assessed using the program MOLPROBITY [24]. Data collection and refinement statistics are summarized in Table 1. The relevant coordinates and structure factors have been deposited at the Protein Data Bank under the accession code 7QZP.

Table 1. X-ray structure parameter of rAB-ELAV, Pdb code: 7QZP. Statistics for the highest-resolution shell are shown in parentheses.

Parameter	Values	Parameter	Values
Wavelength (Å)	1.541	R-free	0.2280 (0.3590)
Resolution range	17.39–1.654 (1.713–1.654)	CC (work)	0.953 (0.523)
Space group	I 4 ₁	CC (free)	0.935 (0.489)
Unit cell (Å)	69.56 69.56 32.46	Number of non-hydrogen atoms	642
Total reflections	96620 (2480)	Protein	601
Unique reflections	8226 (502)	Solvent	41
Multiplicity	11.7 (4.9)	Protein residues	78
Completeness (%)	87.01 (53.41)	RMSD (bonds) (Å)	0.013
Mean I/sigma(I)	18.74 (1.52)	RMSD (angles) (°)	1.74
Wilson B-factor	24.38	Ramachandran favored (%)	96.05
R-merge	0.07861 (0.8945)	Ramachandran allowed (%)	1.32
R-meas	0.08199 (0.9341)	Ramachandran outliers (%)	2.63
R-pim	0.02281 (0.4356)	Rotamer outliers (%)	6.15
CC1/2	0.999 (0.364)	Clashscore	12.87
		Average B-factor (Å ²)	36.18
Reflections used in refinement	8214 (501)	protein (Å ²)	36.45
Reflections used for R-free	411 (25)	solvent (Å ²)	32.25
R-work	0.2049 (0.3354)		

2.6. NMR Measurements and Protein Assignment

Experiments for backbone assignment were performed on samples of the ¹³C, ¹⁵N isotopically enriched RRM domain of rAB-Elavl1 at protein concentration of 300 μM in buffer solution (20 mM HEPES, pH 6.8, 150 mM NaCl, 3 mM MgCl₂, Proteases Inhibitor Cocktail). NMR spectra were recorded at 298 K on a Bruker AVANCE NEO 900 spectrometer, equipped with a triple-resonance Cryo-Probe (Bruker, Billerica, MA, USA). 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,

time course. Association and dissociation rate constants were determined from nonlinear regression fits of the data according to the association kinetic model of multiple ligand concentration in GraphPad Prism[®], version 6.1. The resulting K_D values obtained by k_{off}/k_{on} ratio [29].

2.11. Western Blot from *A. baumannii* and HEK293 Cells Lysate

A. baumannii strain ATCC 19606 (American Type Culture Collection, Manassas, VA, USA), was grown in Luria-Bertani (LB) medium, in the incubator at 37 °C, 200 rpm shaking. Inoculum was grown overnight and the next day it was diluted to a final concentration of 0.05 OD₆₀₀. The bacteria were allowed to grow to a final OD₆₀₀ of 0.5 (they were measured at the spectrophotometer at a λ : 600 nm) and pelleted at 4000 rpm for 20 min at 4 °C. The pellet was incubated for 30 min in ice with lysis buffer (50 mM Tris HCl pH 7.5, 100 mM NaCl, 10% glycerol, 0.1% Triton, 1 mM DTT, 1 mM EDTA, Leupeptine, Aprotinin, Lysozyme, 2.5 U/ μ L) to a final volume equal to 1/20 of the initial culture, and then sonicated. HEK293 cells were rinsed with PBS and lysed in ice-cold RIPA buffer, while bacteria were lysed in a bacterial lysis buffer (20 mM Tris HCl pH 8, 150 mM KCl, 1 mM MgCl₂, 1 mM DTT, DNase, Proteinase inhibitors and RNase inhibitors). Proteins were boiled in SDS gel sample buffer, separated by SDS-PAGE and immunoblotted onto a polyvinylidene difluoride membrane. The primary antibody against AB-Elavl was developed by Davids biotechnologie in rabbit, while the antibody against HuR was purchased from Santa Cruz Biotechnologies (6A97) (Santa Cruz, Dallas, TX, USA). Bands were visualized with anti-rabbit or anti-mouse HRP-conjugated secondary antibodies and scanned on Biorad Chemidoc (Biorad, Hercules, CA, USA).

2.12. Time Resolved Fluorescence Resonance Energy Transfer (HTRF-FRET)

All assays were performed in 20 μ L in 96-well low-volume white plates, in triplicate. The EC₅₀ calculation was performed by adding 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 donor beads XL665-conjugated for biotin detection at a final concentration of 35 nM), potassium fluoride buffer and FRET reaction buffer 1x provided by the manufacturer. After brief spinning (1000 rpm, 1 min), the plate was incubated for 1 h at 4 °C. The signals of acceptors and donors were detected using Tecan Spark (Tecan, Zürich, Switzerland) and the results were calculated using the following equation:

$$\text{Acceptors/Donors} \times 10,000.$$

2.13. Immunoprecipitation (IP) Assay of AB-Elavl

For each IP, 2.5 mg of total protein lysate from *A. baumannii* was used. Bacteria were lysed in RIP lysis buffer (20 mM Tris HCl pH 8, 150 mM KCl, 1 mM MgCl₂, 1 mM DTT, DNase, Proteinase inhibitors and RNases inhibitors) [30]. The lysate was incubated with Pierce A/G beads (Thermo Scientific Pierce 88847–88848, Waltham, MA, USA) for pre-clearing steps 2 h at 4 °C; in parallel, 50% A and 50% G beads were incubated either with 10 μ g of anti-rAB-Elavl antibody or 10 μ g of IgG antibodies for antibody-coating step for 2 h at RT. At the end of the 2 h of incubation, the protein lysate was incubated with antibodies and beads overnight at 4 °C. Finally, samples were washed (5 times, 5 min each wash) with NT2 buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM MgCl₂, 0.05% NP40). The pellet was then analyzed by western blot or mass spectrometry assay.

2.14. Mass Spectrometry (MS) Analysis

To perform the MS analyses of rAB-Elavl protein and *A. baumannii* ATCC 19606 lysate, the samples were separately resolved on 10% polyacrylamide gel. After Coomassie stain, a protein band corresponding to rAB-Elavl and a region representing proteins with molecular masses of 6–14 kDa were cut from the gel. Excised gel bands were cut into small

pieces (~1 mm³) and subjected to reduction and alkylation with 10 mM DTT and 55 mM iodoacetamide, respectively. Gel pieces were then dehydrated with acetonitrile and dried in a speed-vac. Gel plugs were rehydrated with 50 mM NH₄HCO₃ solution containing 12.5 ng/mL trypsin (Promega, Madison, WI, USA) on ice for 30 min. The digestion was continued at 37 °C overnight. The supernatant was collected, and the peptides were sequentially extracted from the gels with 30% ACN/3% TFA and 100% ACN. All of the supernatants were combined and dried in a SpeedVac. The peptides were then acidified with 1% TFA, desalted on C18 stage-tips and resuspended in 20 µL of 0.1% formic acid buffer for LC-MS/MS analysis.

To perform the IP-MS analyses, the co-immunoprecipitated complexes were eluted with Laemmli buffer containing 5% β-mercaptoethanol at 80 °C for 10 min. The samples were loaded on a 10% SDS-PAGE and run for about 1 cm. Gels were then stained with Coomassie and the entire stained area was excised as one sample. The stained bands were then subjected to in gel digestion and peptide desalting process as described above. Samples Digested peptides were separated using an Easy-nLC 1200 system (Thermo Scientific, San Jose, CA, USA) on a reversed-phase column (25 cm column, inner diameter of 75 µm, packed in-house with ReproSil-Pur C18-AQ material: 3 µm particle size, Dr. Maisch, GmbH), heated at 40 °C, with a two-component mobile phase system of 0.1% formic acid in water (buffer A) and 0.1% formic acid in acetonitrile (B). The 85-min gradient was set as follows: from 5% to 25% over 52 min, from 25% to 40% over 8 min and from 40% to 98% over 10 min at a flow rate of 400 nL/min. Peptides were analyzed in a Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) in data-dependent mode and positive mode (2100 V). The full-scan in the Orbitrap was performed at 120,000 fwhm resolving power (at 200 m/z) and followed by a set of (higher-energy collision dissociation) MS/MS scans over 3 s cycle time. The full scans were performed with in a mass range of 350–1100 m/z, a target value of 1 × 10⁶ ions and a maximum injection time of 50 ms. A dynamic exclusion filter was set at 40 sec. The MS/MS scans were performed at a collision energy of 30%, 150 ms of maximum injection time (ion trap) and a target of 5 × 10³ ions. Peptides searches were performed in Proteome Discoverer software version 2.2 (Thermo Fisher Scientific) against the *A. baumannii* database (uniprot, downloaded March 2021), the rAB-Elavl amino acid sequence, and a database containing common contaminants. Proteins were identified using MASCOT search engine, with a mass tolerance of 10 ppm for precursor 0.6 Da for product. Trypsin/P was chosen as the enzyme with 5 missed cleavages. Static modification of carbamidomethyl (C) with variable modification of oxidation (M) and acetylation (protein N-term) were incorporated in the search. The false discovery rate was set to 1% at both peptide and protein level. The results were filters to exclude potential contaminants. For protein quantification in IP-MS experiment, peak intensities were transformed into log₂ space. Data were normalized by the average of its abundance within each sample to account for variation in sampling volumes [31]. Significant abundance differences between conditions were determined using a *t*-test.

3. Results

3.1. Identification of a Putative RRM Containing RBP, AB-Elavl

We performed a similarity search for the HuR protein in the RefSeq Genome Database, limited to the species *A. baumannii*, using tblastn and we found 25 hits on the same gene locus from different *A. baumannii* genomes (Table S1, Supplementary Materials). However, in 11 cases, the region of homology was limited to the RRM1 (aa 20–98 on the human protein), in the remaining 14 hits instead covered the RRM3 (aa 244–322, Figure 1A). The percentage of identity ranged 32.90–35.48% (on average 34.08 ± 1.00%) when the subject sequence was the human RRM1, and it ranged 31.94–46.00% (on average 35.70 ± 3.67%) when the subject was the human RRM3 (Figure 1B). The positive matches 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 S1, Supplementary Materials). The identified gene, F3P16_RS16475 (Table S2, Supplementary Materials), encodes for a putative RNA binding

protein that has been named here AB-Elavl. The similarity scores between the human HuR's RRM1-RRM2: 32% id and 57% positives, RRM1-RRM3: 36% id and 53% positives, RRM2-RRM3: 30% id and 48% positives). The gene locus *AB-Elavl* was present in nearly all the deposited *A. baumannii* genomes (4946 out of 4972 available), suggesting that it belongs to the core genome of this species. The *AB-Elavl* gene was comprised between a gene encoding an ASCH domain containing protein, 8 bp downstream, and a gene encoding ATP-dependent helicase, 74 bp upstream (Figure 1C). We searched for AB-Elavl homologous proteins in all bacterial genomes contained in OrtholugeDB¹⁷ and found 227 hits in genomes spanning 12 bacterial phyla. The sequence length ranged 78–241 aa with an average of 106.68 ± 30.06 ; multiple sequence alignment consisted of a 375 aa alignment. The bacterial proteins homologous to the human ELAV-like family shared an identity score ranging from 9.1% (*Dyadobacter fermentans* versus *Shewanella pealeana*)–99.3% (between two *Shewanella* spp.), having on average $31.1 \pm 10.8\%$ (Table S3, Supplementary Materials). The visual evaluation of the multiple-sequence alignment suggested that there were conserved regions within the bacterial homologous of HuR (Figure 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 (Figure 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-1 and RNP-2 in each of the RRMs in HuR protein², further supporting its potential RNA binding ability. Collectively, we identified a conserved gene in the *A. baumannii* species and in many other bacteria phyla, containing a RRM domain with a different consensus sequence with respect to the metazoan one.

3.2. *AB-Elavl* Gene Is Expressed and Translated in *A. baumannii*

We checked whether the RNA transcript corresponding to *A. baumannii* Elav-like (*AB-Elavl*) was expressed. We retro-transcribed the total RNA of the *A. baumannii* reference strain (ATCC 19606) and amplified by PCR the surroundings of the gene of interest (Figure 2A) using three different pairs of primers. The amplicons' sequences were confirmed by Sanger sequence analysis and matched the DNA deposited sequence. We observed that our gene of interest is expressed and contained into a longer, polycistronic, mRNA of at least 764 bp (Figure S1, Supplementary Materials). We cloned the *AB-Elavl* sequence into the expression plasmid pET30a(+) in frame with a 6XHis tag in the C terminal region and expressed the recombinant *A. baumannii* Elav-like (rAB-Elavl) protein in *E. coli* Rosetta BL21. We purified the recombinant protein (predicted MW 12.8 kDa) from cell lysate by affinity purification using Nickel NTA agarose beads. The purity of the rAB-Elavl protein was evaluated by Coomassie staining of protein polyacrylamide gel of each purification step performed by increasing imidazole concentration (Figure 2B). The yield of protein expression was 12.5 mg/L. The purified rAB-Elavl protein was subjected to mass spectrometry analysis after trypsin digestion. We obtained 80.2% coverage of the entire recombinant sequence, missing the first 17 amino acids, and the detected peptides perfectly matched the predicted amino acid sequence (Figure 2C up). To evaluate whether the polycistronic mRNA is translated into a protein containing the domain of interest, we performed proteome analyses by mass spectrometry on the protein lysate of *A. baumannii* ATCC 19606. Protein cell lysate was separated into a polyacrylamide gel, a band (6–14 kDa) comprising the MW of the predicted protein (predicted MW 10.8 kDa) was cut, trypsin digested and submitted to LC-MS/MS analysis. Among the detected fragments, we obtained 35% coverage of the recombinant protein with complete matching of the experimental amino acid sequences with the reference (Figure 2C down). Interestingly, the detected protein fragments contained the region of the highly conserved octapeptide KGFGFVEM that we found conserved in the protozoans and that corresponds to the ribonucleoprotein motif 1 (RNP-1) in metazoans (Figure 2D). This analysis confirmed the presence of several small peptides belonging to our hypothesized protein but did not provide any information on the real length of

the translated protein. To obtain more insight on the presence and on the MW of the hypothetical AB-Elavl protein translated in the bacterium, we developed an antibody against the recombinant protein (α AB-Elavl). Rabbits were immunized with the denatured rAB-Elavl protein, and after two cycles of immunization, the serum was collected, and the IgG titer quantified. Specificity of the IgG in recognizing the protein of interest was investigated by performing western blot against the rAB-Elavl protein, the recombinant HuR (rHuR), human cell lysate (MCF7 cell line) and the total proteome of *A. baumannii* (Figure 2E). We confirmed that the immunized serum recognizes the rAB-Elavl, but not HuR, and observed a band at a slightly heavier MW compared to the calculated one in the total protein lysate of the bacterium. This suggests that the native protein is longer than the predicted one as well as the recombinant one is digested in any part during the protein production into *E. coli*. To obtain more proofs about the presence of AB-Elavl protein, we performed a protein immunoprecipitation from the lysate of *A. baumannii* using the immunized serum with the α AB-Elavl. Protein precipitate was run on polyacrylamide gel, but no bands were detected (Figure 3A). Therefore, we performed an immunoprecipitation followed by mass spectrometry (IP-MS) on the same material, to investigate which proteins were enriched, with respect to rabbit IgG, used as control. About 5675 proteins (Figure 3B) 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 *A. baumannii* (Figure 3C and Figure S2, Supplementary Materials) that are extremely similar to our protein of interest (percentage of homology 54.5–81%, Figure 3C). 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. 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 the RNP-2 protozoa sequence. 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 two RNPs amino acid sequences responsible for nucleic acid binding.

3.3. AB-Elavl Protein Has a Typical Single RRM Domain Structure

The X-ray structure (Table 1) showed the domain of an RRM domain: β 1- α 1- β 2- β 3- α 2- β 4 [32] with an additional β 5 segment present right before the C-terminus of the protein. The two conserved amino acid sequences, ILVNRL and KGFGFVE, are localized at the level of the two internal strands of the β -sheet: β 2- β 3 (Figure 4A and Figure S5, Supplementary Materials). The structure was solved using the molecular replacement method; the model showing the highest sequence identity (about 40%) was 1FXL (HuD, HuR human paralog in complex with 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 4A shows the superposition between rAB-Elavl (red) and 1FLX (green). It appears quite clear that the fold of rAB-Elavl is very similar to that of the model used for structure solution and, in turn, similar to the typical RRM (Figure 4B). 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 very high mobility. The average B-factor along the sequence is, in fact, about 35 Å², confirming an overall rigidity of the structure, with the only exception being the above-mentioned region 50–58 where temperature factor values are extremely high. This mobility is not present in the case 1FXL because this region interacts with RNA. The Ramachandran plot is good for all residues except for those in the mobile regions, where the chain tracing is very approximate.

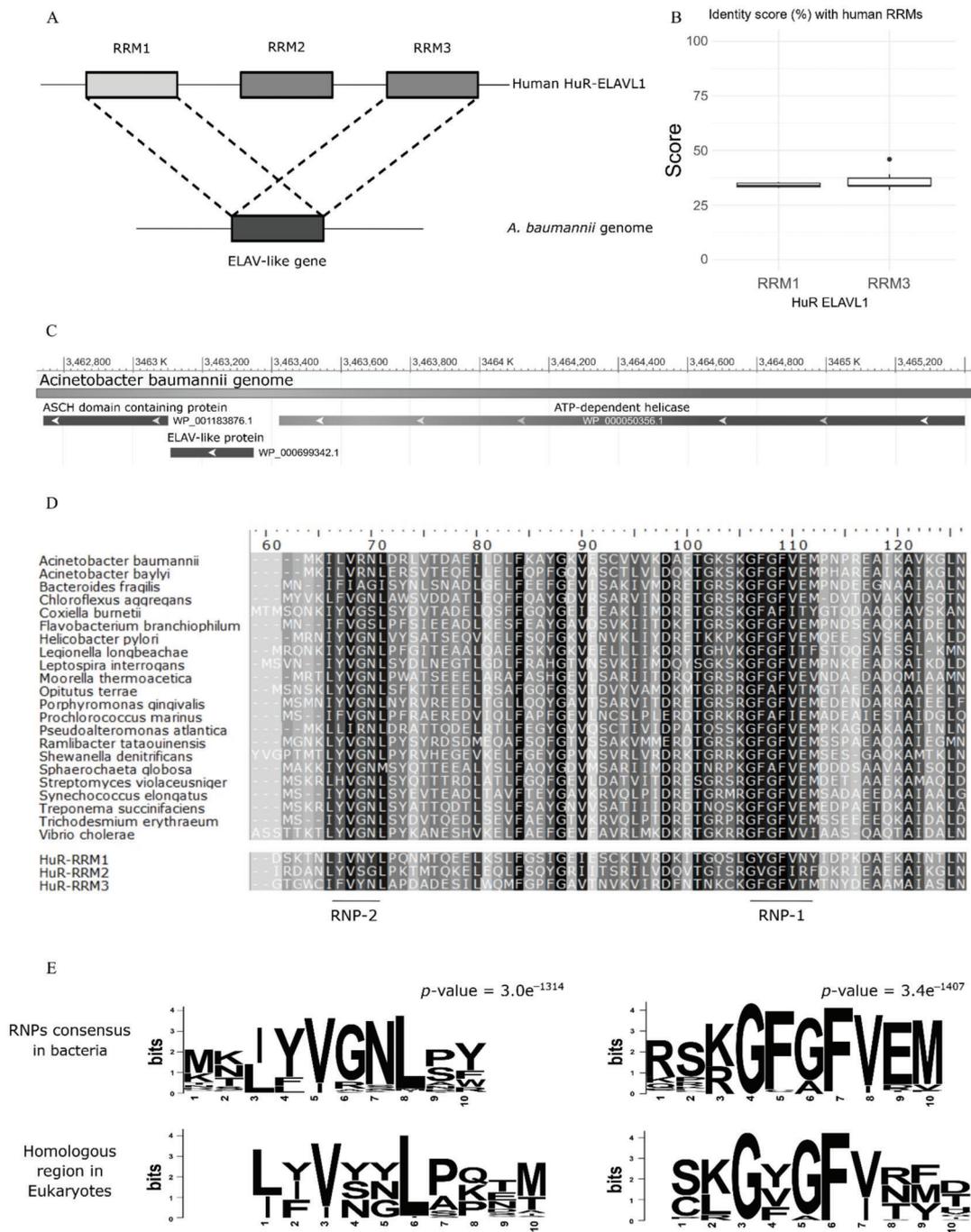


Figure 1. In silico analysis. (A) The tblastn search, using as query the human HuR and restricting the search to *A. baumannii* genomes, gave 25 hits corresponding to the same orthologous protein which share a high homology with both RRM1 and RRM3. (B) Boxplots showing the percentages of identity with the two RRM1 and RRM3, see Table S1 tblastn for the extended dataset. (C) Genomic context of the bacterial HuR, it is shown that the three genes, namely an ATP-dependent helicase, the AB-Elavl and the ASCH domain containing protein are arranged in proximity. (D) Alignment of the bacterial homologues of human ELAV in selected bacterial species with clinical or environmental relevance

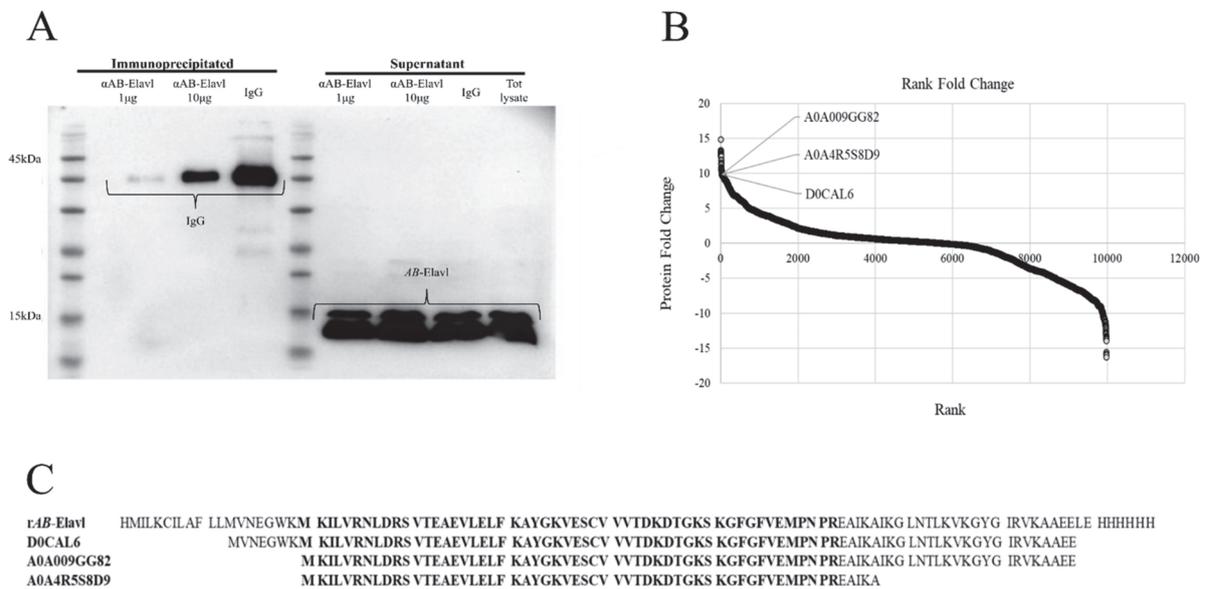


Figure 3. Molecular characterization of AB-Elavl. (A) Immunoprecipitation assay on the total protein lysate of *A. baumannii*. IgG was used as a control. No enrichment of the protein was visible by western blot analysis. (B) Protein ranking based on \log_2 fold-changes (IP/IgG) for all the proteins identified by MS showing an enrichment of three hypothetical and highly similar RNA-binding proteins (RBPs) in the top ten proteins. (C) Entry code and amino acid sequence of the three hypothetical RBPs based on the IP-MS analysis compared with the rAB-Elavl sequence. Bold retrieved peptides (sequence coverage: 62%, 67% and 91% for D0CAL6, A0A009GG82 and A0A4R5S8D9 proteins, respectively).

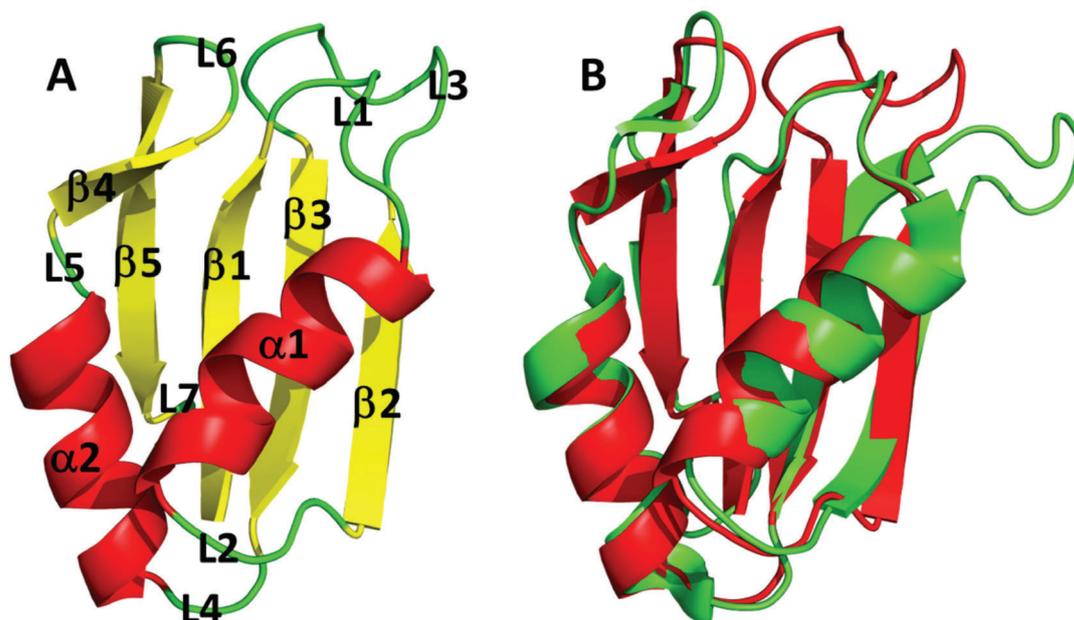


Figure 4. AB-Elavl protein structure. (A) Ribbon representation of the three-dimensional structure of the bacterial hypothetical HuR RRM domain. The secondary structural elements and loops have been annotated: helices ($\alpha 1$ – $\alpha 2$), strands ($\beta 1$ – $\beta 5$), loops (L1–L7). (B) Superposition of the crystal structure of the bacterial hypothetical HuR RRM domain (rAB-Elavl, red) and 1FLX (green).

The 2D ^1H - ^{15}N HSQC spectrum of the rAB-Elavl shows well-dispersed and resolved signals in agreement with a small, uniform and well-folded protein structure (Figure 5). 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. The NMR assignment of AB-Elavl (residues Lys22-Glu101) has been deposited in the BMRB database under the accession code: 51440. According to TALOS+ predictions, the RRM domain is constituted by two α -helices and four β -strands (Figure S3, Supplementary Materials), in agreement with the currently resolved crystal structure and with the reported structures of the RRM domains.

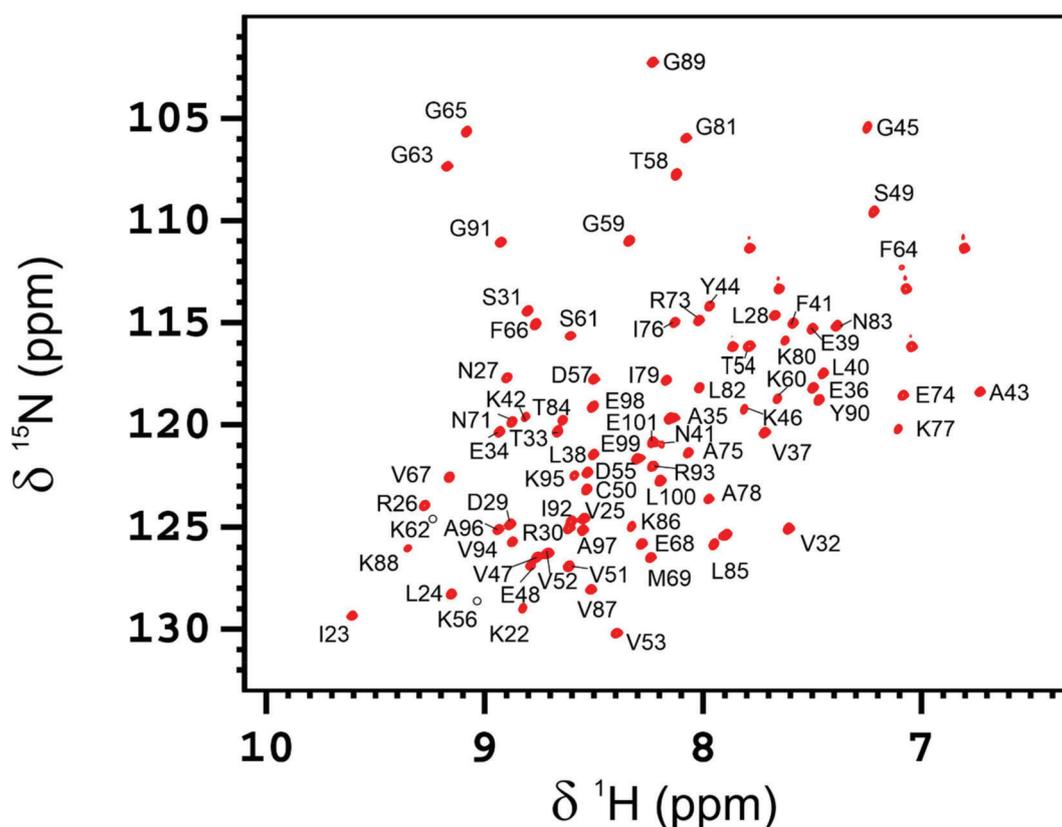


Figure 5. 2D ^1H - ^{15}N HSQC spectrum of AB-Elavl. The spectrum was recorded with a spectrometer operating at 900 MHz and 298 K. Assignment is reported on the signals.

3.4. AB-Elavl Binds AU Rich RNA Probes

Since AB-Elavl is an RRM domain containing protein characterized by the two conserved RNPs, we investigated whether the RRM domain of AB-Elavl protein could have RNA-binding abilities. We investigated whether in the proteome of *A. baumannii* there are proteins able to recognize and bind the ARE sequences taken by the 3'UTR of *TNF α* , a target of HuR, using non-denaturing and non-cross-linked RNA Electromobility Shift (REMSA) assay²⁸. The single strand (ss) RNA probe ARE probe was bound with an infra-red tag DY681 (IR-ARE pos). By mixing higher concentrations of the protein lysate with a fixed 2.5 nM concentration of the IR-ARE pos probe, we 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-ARE pos (Figure 6A). We then evaluated if the rAB-Elavl was able to bind to probes that contained the HuR consensus sequence (ARE pos) and probes without the HuR consensus sequence (ARE negative RNA probes, ARE neg), by

REMSA. We mixed increasing amounts (40 nM, 80 nM and 160 nM) of protein with 2.5 nM IR-ARE pos and 2.5 nM IR-ARE neg. As shown in the mobility shift assay, rAB-Elavl clearly caused the RNA probe electrophoretic retardation detectable as one prominent band, with both probes, showing a binding preference towards the ARE pos probe in this biochemical condition, since the shifted band is not clearly visible for the ARE neg (but the free RNA formation decreases with the incrementation of the protein) (Figure 6B). We evaluated the possible formation of a super-shift by adding the antibody against rAB-Elavl. We were expecting the formation of the heavy complex antibody–protein–RNA, but unfortunately, we could not observe any super shift (Figure 6C).

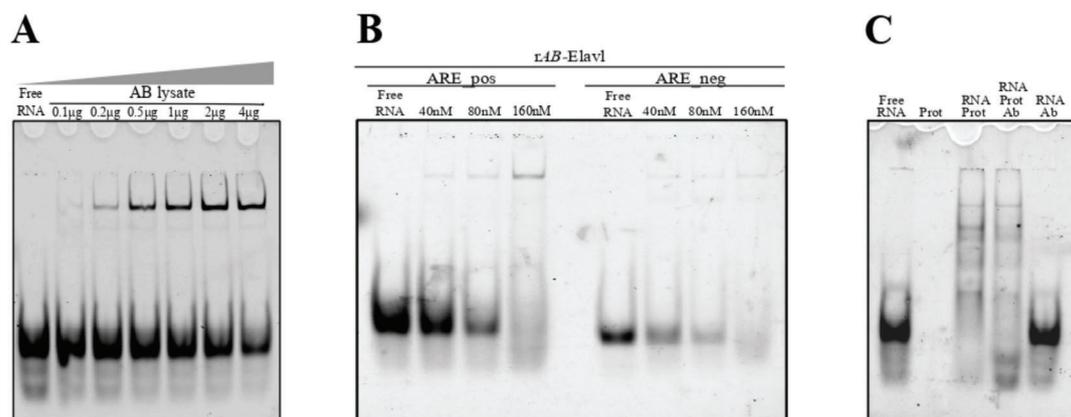


Figure 6. Analysis of the protein binding abilities. (A) REMSA assay on the total protein lysate of *A. baumannii* and a probe mimicking the AU rich sequence of TNF α (ARE pos) with an infrared tag. (B) REMSA assay on the recombinant protein incubated with different probes with an infrared tag: ARE pos and ARE neg. ARE pos is bound with a high affinity, while ARE neg shows a lower affinity. (C) REMSA assay for detection of the super-shift in presence of the antibody against AB-Elavl. The super-shift is not detectable.

The interactions of the AB-Elavl with ARE pos and ARE neg were also investigated through solution NMR. In the presence of ARE pos, at the protein/RNA molar ratio of 1:2 a global decrease of signal intensity was observed (Figures 7 and S4, Supplementary Materials). In particular, some residues located on the β -platform (Leu24, Val25, Asn27, Ser31, Val52, Thr58, Gly63, Gly65, Phe66, Lys80, Lys88, Gly89, Ile92) experience a larger decrease in signal intensity (Figure 7A and 7C top). Some of these residues (or the neighboring ones) are also affected by Chemical Shift Perturbation, CSP (Val25, Arg30, Ser31, Val52, Gly63, Gly89, Ile92, Glu99, Leu100, Glu101, Figure 7B and 7C top). It should be noted that Leu24, Val25 and Asn27 are located in RPN1, and Gly63, Gly65, Phe66 in RPN2. The superimposition of the X-ray structures (Figure S5, Supplementary Materials) shows that the same conserved protein regions are involved in the interaction of the HuR human paralog with the AU-rich element of the C-FOS RNA. In the presence of ARE neg, instead, the effect at the same protein/RNA molar ratio is much reduced. However, some residues of the β -platform still experience a decrease in signal intensity (Val25, Leu38, Phe41, Val51, Val53, Phe64, Gly65, Phe66, Tyr90, Glu101; Figure 7A and 7C bottom) and/or a CSP (Lys22, Val25, Arg30, Phe64, Gly65, Ile92, Glu98, Glu99, Leu100, Glu101 Figure 7B and 7C bottom).

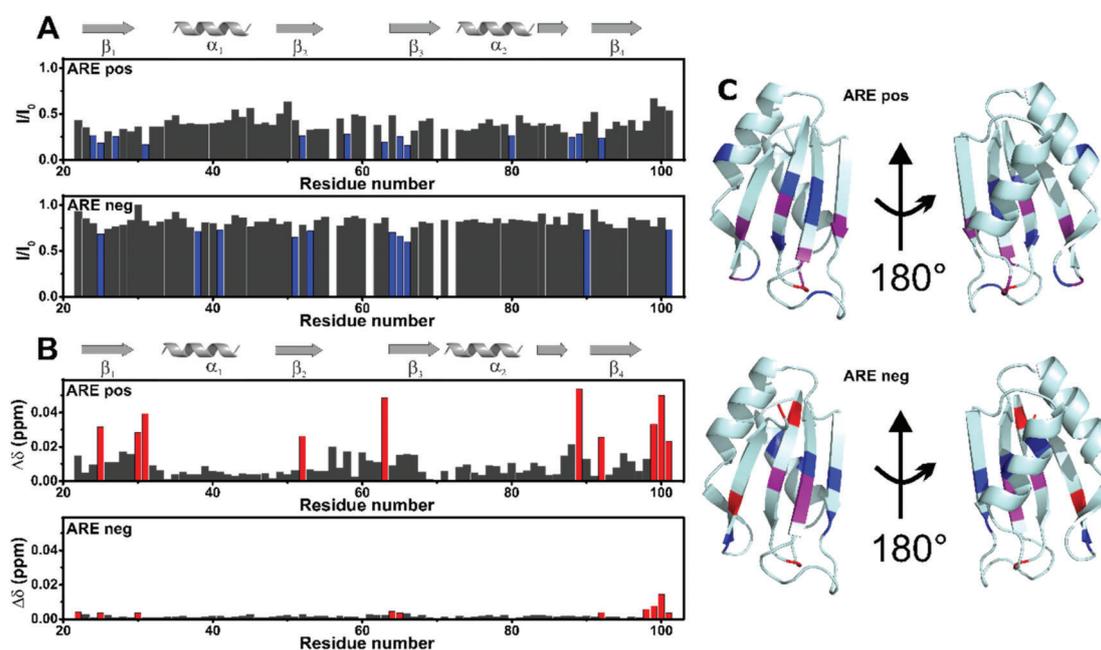


Figure 7. NMR analysis of the protein binding abilities. (A). Plots of decreases in signal intensity of rAB-Elavl RRM domain in the presence of 140 μM ARE pos (top), or 140 μM ARE neg (bottom) with respect to the free protein (70 μM). The residues experiencing the largest decreases have been highlighted in blue. (B). Chemical shift perturbation (CSP) of rAB-Elavl RRM domain (70 μM) with respect to the protein in the presence of 140 μM positive RNA (top), and 140 μM negative RNA (bottom). The CSP was evaluated with the formula: $\Delta\delta = \frac{1}{2} \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$. The residues experiencing the largest CSP have been highlighted in red. (C) Highlighted in blue are the residues experiencing the largest decreases in signal intensity, in red the residues experiencing the largest CSP, and in violet the residues experiencing the largest decreases in signal intensity and CSP, in the presence of 140 μM positive RNA (top), and 140 μM negative RNA (bottom).

3.5. AB-Elavl Binds AU Rich RNA Probes with Nanomolar Affinity

To quantitatively characterize the binding activity of the rAB-Elavl to different RNA probes, we applied AlphaScreen technology using 5'-biotinylated ssRNA probes as substrate. We used the 5'-biotinylated ARE pos (Bi-ARE pos) and the biotinylated ARE negative (Bi-ARE neg). We optimized the assay to identify the best molar ratio between the two interacting partners coupled with anti-His-Acceptor and Streptavidin-Donor beads; the optimal concentration, before the hooking effect, was observed at 250 μM and 50 nM for rAB-Elavl and Bi-probes, respectively (Figure S6, Supplementary Materials). We then evaluated the affinity of binding between the rAB-Elavl protein and probes with different sequences but the same length. The recombinant protein shows a high affinity for Bi-ARE Pos ($EC_{50} = 0.5$ nM) while it has low affinity for the Bi-ARE neg probe ($EC_{50} = 257.1$ nM) (Figure 8A,B). Coherently with REMSA, rAB-Elavl bound both probes, but we could quantify a stronger affinity for the ARE-positive probe. We then evaluated the minimal ARE sequence required for binding. We observed that rAB-Elavl, as its human orthologous¹⁴, has a higher affinity for longer ARE sequences than for shorter ones (EC_{50} ARE pos 29 nt = 35.62, ARE pos 19 nt = 64.76, ARE pos 11 nt = not converged) (Figure 8C). The affinity evaluation was confirmed using the HTRF-FRET assay. The probes used were the same of the AlphaScreen assay: Bi-ARE pos and Bi-ARE neg. This assay, as well, was optimized to identify the best molar ratio between the two interacting partners coupled with europium-labeled anti-6X His-Antibody and XL665-conjugated for biotin detection; the optimal protein concentration, before the hooking effect, was observed at around

200 nM for both the probes (Figure S7, Supplementary Materials). The EC_{50} (nM) are respectively: ARE pos $EC_{50} = 35.11$ nM and ARE neg- $EC_{50} = 945.5$, (Figure 8D). To further define the binding affinity between the rAB-Elavl protein and ARE sequence, we performed a time course experiment in which different concentrations of the Bi-ARE pos were mixed to the protein with different incubation time. The experiment shows that the binding of rAB-Elavl to Bi-AREpos probe was both time and dose dependent (Figure 8E). Data were globally fitted using the association kinetic model of multiple ligand concentration: derived association (k_{on} of $2.035 \text{ M}^{-1} \text{ min}^{-1}$) and dissociation (k_{off} of 0.02687 min^{-1}) rates indicated a very high affinity of the rAB-Elavl protein towards this RNA substrate and a low dissociation rate. According to the law of mass action, the equilibrium binding constant K_D calculated as k_{off}/k_{on} was obtained as K_D value of 13.2 nM. We performed the same type of assay for Bi-ARE neg, for which the binding resulted as ambiguous (Figure 8F).

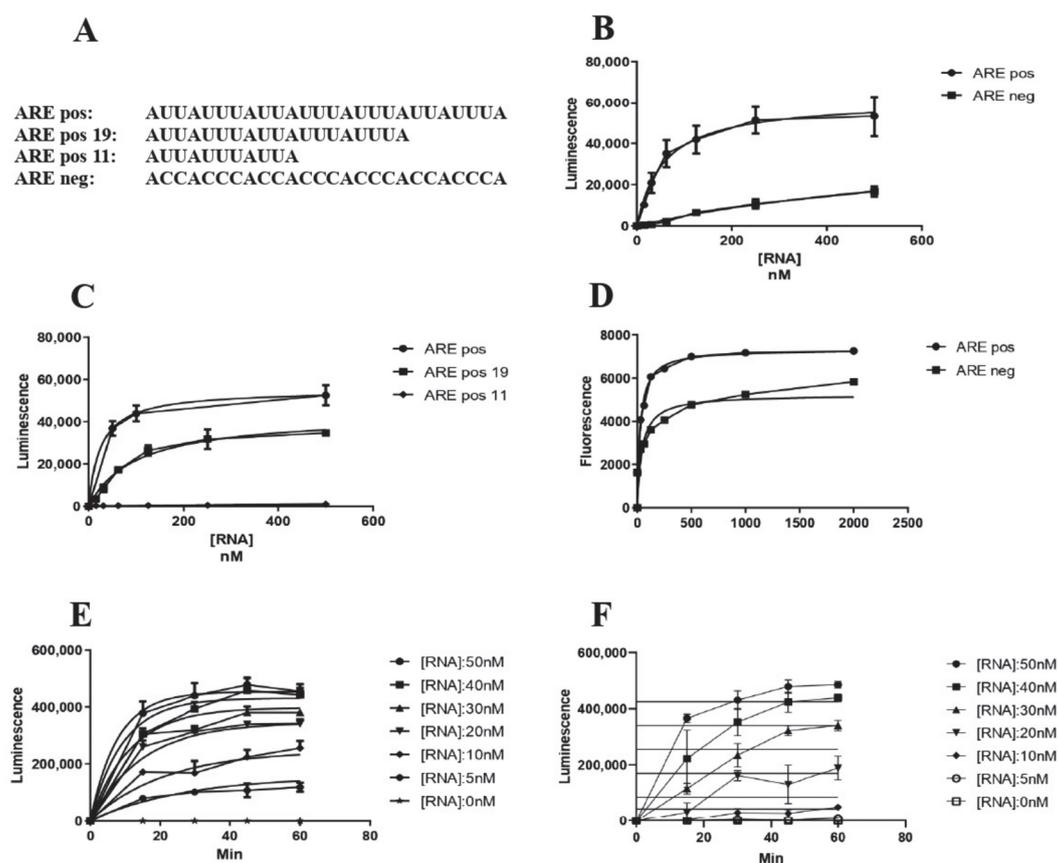


Figure 8. Biochemical characterization of the protein binding ability. (A) Sequences of the probes used in the different assays. (B) AlphaScreen saturation experiment between the recombinant protein and AREpos and AREneg. The EC_{50} was determined from non-linear regression fits of the data according to the dose–response model in GraphPad Prism®, version 6.1. (C) 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. The EC_{50} was determined from nonlinear regression fits of the data according to the dose–response model in GraphPad Prism®, version 6.1. (D) EC_{50} evaluation through saturation experiment by HTRF-FRET and AREpos and AREneg. AREpos is confirmed to have high affinity while AREneg is not well bound. The EC_{50} was

determined from nonlinear regression fits of the data according to the dose–response model in GraphPad Prism[®], version 6.1. (E,F) Kinetic experiment with rAB-Elavl and AREpos (E) is dose dependent, while for AREneg (F) the binding resulted ambiguous. Association (K_{on}) and dissociation (K_{off}) 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.

4. Discussion

We uncovered the existence of an RRM-containing, RNA-binding protein in *A. baumannii*, solved its structure and provided an initial characterization of its RNA-binding abilities. We started by performing a similarity search using the HuR protein sequence as seed and found a conserved gene locus encoding for a putative RBP conserved in most of the *A. baumannii* species. This supports the hypothesis that *AB-Elavl* found in the *A. baumannii* genome encodes for a protein that has a functional relevance in *A. baumannii* physiology. The length of the AB-Elavl roughly corresponds to a single RRM, while in Eukarya, the RRM domain is present in tandem with other heterologous or homologous domains. Indeed, evolution led to an increase in the number and specificity of eukaryotic RBPs; they are often characterized by a repetition of domains that collaborate for a better affinity to the target RNA [5,14,33]. On the contrary, bacteria tend to be more streamlined, with simpler RBPs composed by just one single domain but with wider functions, since they are less specific for their targets [4,8]. Bacterial RBPs, contrary to eukaryotic ones, can normally bind a wider number of sequences. For example, Hfq has a wide substrate selection, underlying the different roles covered by this protein: from RNA chaperone to ribosome biogenesis, DNA compaction, protein–protein interactions, and involvement in RNA degradation machinery [8,34,35]. We characterized the crystal structure of rAB-Elavl, which retraces the common structure of the classical RRM domain: $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\beta 3$ - $\alpha 2$ - $\beta 4$, in which RNP1 and RNP2 are in the internal strands of the beta-sheet ($\beta 1$ - $\beta 3$) [1,4]. We observed a partial divergency in the RNP1 and RNP2 sequences, the regions responsible of the binding with the RNA [1,2], between the HuR's RRM and the AB-Elavl RRM. Hence, the target consensus sequences between bacterial AB-Elavl and eukaryotic Elavl RNPs, are characterized by similar recognition motifs [2]; however, the two proteins do not show completely equivalent RNA binding properties. Indeed, we tested the binding affinities of RNA probes using the sequences known to be targeted by HuR (ARE pos) and the respective negative sequence (ARE neg). The presence in the proteome of *A. baumannii* of possible proteins binding the ARE sequences was confirmed by a REMSA assay, and the ability of rAB-Elavl to interact both with the ARE-positive and ARE-negative probes by AlphaScreen, FRET and NMR. All the assays showed a preference of the protein toward the AU-rich sequence in the nanomolar range, even though the AC-rich peptides are also bound but with a clear lower affinity. By a time course kinetic assay, we also calculated a K_D in the nanomolar range (13.2 nM), although an order of magnitude higher than the reported HuR K_D , in the same experimental condition (2.5 nM) [29]. The kinetic experiment using the ARE neg probe, instead, gave an ambiguous K_D calculation.

5. Conclusions

All these results suggest that *A. baumannii* express an RRM-containing RBP that shares RNA-binding properties and characteristics with the human HuR for the ability to bind RNA AU-rich region, although with lower affinity and specificity. However, the exact length of the protein produced has not been determined and it is likely that the RRM domain is contained into a longer protein. The structure of the bacterial RRM domain appears very similar to the eukaryotic one but for the presence of an additional short β strand and a more flexible central region. Notably the amino acids present in the RNP1 and RNP2 are different between protozoans and metazoans, but they are similarly involved in the RNA binding. In addition, functional studies are needed to understand the role of this protein in the bacteria.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12070922/s1>, Figure S1. Sanger sequence of the retro-transcribed amplicon of the transcriptome of *Acinetobacter baumannii*. We used the reverse primer R3 that was designed 235 nucleotides downstream the gene of interest. This data confirmed the insertion of the mRNA sequence into a bigger polycistronic mRNA and confirmed the possible expression of the protein AB-Elavl. Figure S2. Scatter plot comparing protein enrichment (log₂ FC IP-vs-IgG) compared with protein abundance within the IP proteome. Hypothetical and highly similar RNA binding proteins of *A. baumannii* are highlighted. Figure S3. Top, secondary structure prediction obtained by the program Talos+ using the experimental values of chemical shifts of HN, N, C', C α , and C β atoms as input data. The blue bars indicate the β -strand propensity while the red bars the α -helix propensity. Bottom, predicted order parameter (S₂) by the program Talos+ using the experimental values of chemical shifts of HN, N, C', C α , and C β atoms as input data. Figure S4. NMR analysis of the protein binding toward AREs. Superimposed 2D 1H-15N HSQC spectra of free AB-Elavl RRM domain (70 μ M, black) and in the presence of 140 μ M of ARE Pos (red). The spectra were acquired on a spectrometer operating at 950 MHz and 298 K. Figure S5. Superposition of the crystal structure of the bacterial hypothetical HuR RRM domain (rAB-Elavl, red) and 1FLX (green and blue) with the presence of RNA (magenta). Figure S6. AlphaScreen for detection of the hook point on the recombinant protein and three different concentrations of AREpos probe, to have the best signal/noise ratio. [protein]: 250 nM; [RNA]: 50 nM. Figure S7. Hook point established by HTRF-FRET between the recombinant protein and AREpos. Table S1: tblastn hits using as query the human HuR protein sequence on the *Acinetobacter baumannii* genomes, the search was performed restricted to the reference prokaryotic representative genomes database. Table S2: Retrieved target gene and protein name and related sequences. Table S3: Pairwise identity matrix of the 226 sequences found by searching the orthologous proteins of *A.baumannii* HuR gene on OrtholugeDB.

Author Contributions: Conceptualization, V.G.D., M.F. and A.P.; Data curation, C.C., R.B., V.C., L.C., A.P.-R. and C.Z.; Formal analysis, C.C., M.F., V.C., L.C. and A.E.; Funding acquisition, A.P. and M.F.; Investigation, C.C., A.P.-R., A.E., I.B. (Irene Bianconi), M.M., R.B., L.C., I.B. (Isabelle Bonomo) and V.C.; Project administration, A.P. and M.F.; Resources, C.C. and A.P.-R.; Supervision, A.P., O.M., M.B.W. and M.F.; Validation, C.C., I.B. (Irene Bianconi), I.B. (Isabelle Bonomo), M.M. and A.P.-R.; Visualization, C.C., A.E., M.F., L.C., V.C. and A.P.-R.; Roles/Writing—original draft, A.P., C.C., M.F. and A.E.; Writing—review & editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by Regione Toscana (CERM-TT and BioEnable), the “Progetto Dipartimenti di Eccellenza 2018–2022” to the Department of Chemistry “Ugo Schiff” of the University of Florence and the Recombinant Proteins JOYNLAB laboratory. The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy center. We acknowledge H2020-INFRAIA iNEXT-Discovery-Structural Biology Research Infrastructures for Translational Research and Discovery (contract No. 871037), EOSC-Life “Providing an open collaborative space for digital biology in Europe” (H2020, contract No. 824087) and “RNAct” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITN-2018 (contract No. 813239). AP wishes to thank AIRC IG 21548.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The coordinates and structure factors obtained by X-ray crystallography have been deposited at the Protein Data Bank under the accession code 7QZP. The NMR assignment of rAB-Elavl (residues Lys22-Glu101), generated during the current study, is available in the BMRB database under the accession code: 51440.

Acknowledgments: We thank D. Peroni of the MS and Proteomics Core Facility at Department CIBIO University of Trento.

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

References

1. Maris, C.; Dominguez, C.; Allain, F.H.T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **2005**, *272*, 2118–2131. [[CrossRef](#)] [[PubMed](#)]
2. Samson, M.L. Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins. *BMC Genom.* **2008**, *9*, 392. [[CrossRef](#)] [[PubMed](#)]
3. Corley, M.; Burns, M.C.; Yeo, G.W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell* **2020**, *78*, 9–29. [[CrossRef](#)] [[PubMed](#)]
4. Holmqvist, E.; Vogel, J. RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* **2018**, *16*, 601–615. [[CrossRef](#)]
5. Colombrita, C.; Silani, V.; Ratti, A. ELAV proteins along evolution: Back to the nucleus? *Mol. Cell Neurosci.* **2013**, *56*, 447–455. [[CrossRef](#)]
6. Gallagher, L.A.; Ramage, E.; Weiss, E.J.; Radey, M.; Hayden, H.S.; Held, K.G.; Huse, H.K.; Zurawski, D.V.; Brittnacher, M.J.; Manoil, C. Resources for Genetic and Genomic Analysis of Emerging Pathogen *Acinetobacter baumannii*. *J. Bacteriol.* **2015**, *197*, 2027–2035. [[CrossRef](#)]
7. Gonzalez-Villoria, A.M.; Valverde-Garduno, V. Antibiotic-Resistant *Acinetobacter baumannii* Increasing Success Remains a Challenge as a Nosocomial Pathogen. *J. Pathog.* **2016**, *2016*, 7318075. [[CrossRef](#)]
8. Farrow, J.M.; Wells, G.; Palethorpe, S.; Adams, M.D.; Pesci, E.C. CsrA Supports both Environmental Persistence and Host-Associated Growth of *Acinetobacter baumannii*. *Infect. Immun.* **2020**, *88*, 58–67. [[CrossRef](#)]
9. Hubloher, J.J.; Schabacker, K.; Müller, V.; Averhoff, B. CsrA Coordinates Compatible Solute Synthesis in *Acinetobacter baumannii* and Facilitates Growth in Human Urine. *Microbiol. Spectr.* **2021**, *9*, e01296-21. [[CrossRef](#)]
10. Moll, I.; Afonyushkin, T.; Vytvytska, O.; Kaberdin, V.R.; Bläsi, U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **2003**, *9*, 1308–1314. [[CrossRef](#)]
11. Sharma, A.; Dubey, V.; Sharma, R.; Devnath, K.; Gupta, V.K.; Akhter, J.; Bhandu, T.; Verma, A.; Ambatipudi, K.; Sarkar, M.; 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*, 13377–13388. [[CrossRef](#)]
12. Wang, P.; Li, R.Q.; Wang, L.; Yang, W.T.; Zou, Q.H.; Xiao, D. Proteomic Analyses of *Acinetobacter baumannii* Clinical Isolates to Identify Drug Resistant Mechanism. *Front. Cell Infect. Microbiol.* **2021**, *11*, 625430. [[CrossRef](#)]
13. Good, P.J. A conserved family of elav-like genes in vertebrates. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 4557–4561. [[CrossRef](#)]
14. Ma, W.J.; Cheng, S.; Campbell, C.; Wright, A.; Furneaux, H. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J. Biol. Chem.* **1996**, *271*, 8144–8151. [[CrossRef](#)]
15. Assoni, G.; La Pietra, V.; Digilio, R.; Ciani, C.; Licata, N.V.; Micaelli, M.; Facen, E.; Tomaszewska, W.; Cerofolini, L.; Pérez-Ráfols, A.; et al. HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models. *Adv. Drug Deliv. Rev.* **2021**, *181*, 114088. [[CrossRef](#)]
16. Whiteside, M.D.; Winsor, G.L.; Laird, M.R.; Brinkman, F.S.L. OrtholugeDB: A bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res.* **2013**, *41*, D366–D376. [[CrossRef](#)]
17. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, w202–w208. [[CrossRef](#)]
18. Kundu, P.; Fabian, M.R.; Sonenberg, N.; Bhattacharyya, S.N.; Filipowicz, W. HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. *Nucleic Acids Res.* **2012**, *40*, 5088–5100. [[CrossRef](#)]
19. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 2*, 125–132. [[CrossRef](#)]
20. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *54*, 5.6.1–5.6.37. [[CrossRef](#)]
21. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 1*, 22–25. [[CrossRef](#)]
22. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 2*, 213–221. [[CrossRef](#)]
23. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 4*, 486–501. [[CrossRef](#)]
24. Chen, V.B.; Arendall, W.B.; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 1*, 12–21. [[CrossRef](#)]
25. Kay, L.E.; Xu, G.Y.; Yamazaki, T. Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H₂O Saturation. *J. Magn. Reson. Ser. A* **1994**, *109*, 129–133. [[CrossRef](#)]
26. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **2009**, *44*, 213–223. [[CrossRef](#)]
27. Manzoni, L.; Zucal, C.; Di Maio, D.; D’Agostino, V.G.; Thongon, N.; Bonomo, I.; Lal, P.; Miceli, M.; Baj, V.; Brambilla, M.; et al. Interfering with HuR-RNA Interaction: Design, Synthesis and Biological Characterization of Tanshinone Mimics as Novel, Effective HuR Inhibitors. *J. Med. Chem.* **2018**, *61*, 1483–1498. [[CrossRef](#)]

28. D'Agostino, V.G.V.G.; Sighel, D.; Zucal, C.; Bonomo, I.; Micaelli, M.; Lolli, G.; Provenzani, A.; Quattrone, A.; Adami, V. Screening Approaches for Targeting Ribonucleoprotein Complexes: A New Dimension for Drug Discovery. *SLAS Discov.* **2019**, *24*, 314–331. [[CrossRef](#)]
29. D'Agostino, V.G.V.G.; Adami, V.; Provenzani, A. A Novel High Throughput Biochemical Assay to Evaluate the HuR Protein-RNA Complex Formation. *PLoS ONE* **2013**, *8*, e72426. [[CrossRef](#)]
30. Keene, J.D.; Lager, P.J. Post-transcriptional operons and regulons co-ordinating gene expression. *Chromosome Res.* **2005**, *13*, 327–337. [[CrossRef](#)]
31. Aguilan, J.T.; Kulej, K.; Sidoli, S. Guide for protein fold change and *p*-value calculation for non-experts in proteomics. *Mol. Omi.* **2020**, *16*, 573–582. [[CrossRef](#)] [[PubMed](#)]
32. Tang, Y.H.; Han, S.P.; Kassahn, K.S.; Skarshewski, A.; Rothnagel, J.A.; 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*, 214–228. [[CrossRef](#)] [[PubMed](#)]
33. 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*, 2029–2036. [[CrossRef](#)] [[PubMed](#)]
34. Asensio, N.C.; Rendón, J.M.; Burgas, M.T. Time-Resolved Transcriptional Profiling of Epithelial Cells Infected by Intracellular *Acinetobacter baumannii*. *Microorganisms* **2021**, *9*, 354. [[CrossRef](#)]
35. 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*, 104608. [[CrossRef](#)]

4.1.2 Supplementary Information

*Identification and Characterization of an RRM-Containing,
RNA Binding Protein in Acinetobacter Baumannii*

Supplementary Tables S1 and S3

See <https://www.mdpi.com/2218-273X/12/7/922>

Supplementary Table S2

Fasta format entry F3P16_RS16475 RNA-binding protein [*Acinetobacter baumannii*]

Gene ID: 66398575, updated on 15-Nov-2021

Gene symbol F3P16_RS16475

>NZ_CP043953.1:c3494438-3494199 *Acinetobacter baumannii* strain K09-14 chromosome, complete genome

```
ATGAAAATATTAGTTCGTAATTTAGATCGTTCAGTGACTGAAGCTGAAGTTTTAGAGCTGTTTAAAGCTT
ATGGTAAAGTTGAGTCTTGTGTCGTTGTAAGTAAAGATAACGGGTAAATCAAAGGGCTTTGGTTTTGT
CGAAATGCCGAATCCGCGTGAAGCCATTAAAGCAATCAAAGGTCTAAATACACTTAAAGTAAAAGGTTAC
GGTATTCGGGTTAAGGCAGCTGAAGAGTAA
```

Predicted protein sequence

> WP_000699342.1

1 mkilvrnldr svteaevlel fkaygkvesc vvvtkdtdgk skgfgfvemp npreaikaik

61 glntlkvkgky girvkaee

Supplementary figures of

Identification and characterization of an RRM-containing, RNA binding protein in *Acinetobacter Baumannii*

Caterina Ciani¹, Anna Pérez-Ràfols², Isabelle Bonomo¹, Mariachiara Micaelli¹, Alfonso Esposito³, Chiara Zucal¹, Romina Belli⁴, Vito Giuseppe D'Agostino¹, Irene Bianconi¹, Vito Calderone², Linda Cerofolini², Orietta Massidda¹, Michael Bernard Whalen⁵, Marco Fragai^{2*} and Alessandro Provenzani^{1*}

¹ Department of Cellular, Computational and Integrative Biology, DeCIBIO, University of Trento, Trento, Italy

² Magnetic Resonance Center (CERM) – Department of Chemistry “Ugo Schiff”, University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMMP) and Giotto Biotech, S.R.L, Florence, Italy

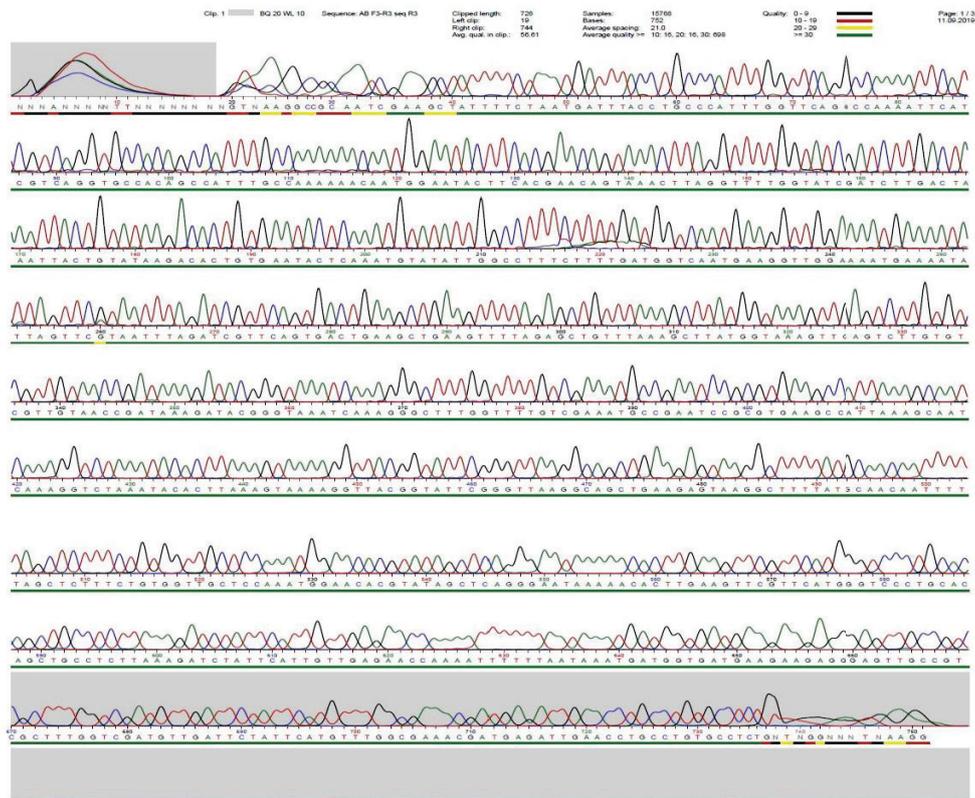
³ International Centre for Genetic Engineering and Biotechnology, (ICGEB), Padriciano, 99, Trieste, 34149 Italy

⁴ Proteomics and MS Core Facility, Department of Cellular, Computational and Integrative Biology, DeCIBIO, University of Trento, Trento, Italy

⁵ Institute of Biophysics (IBF), National Research Council (CNR), FBK Nord, Trento, Italy

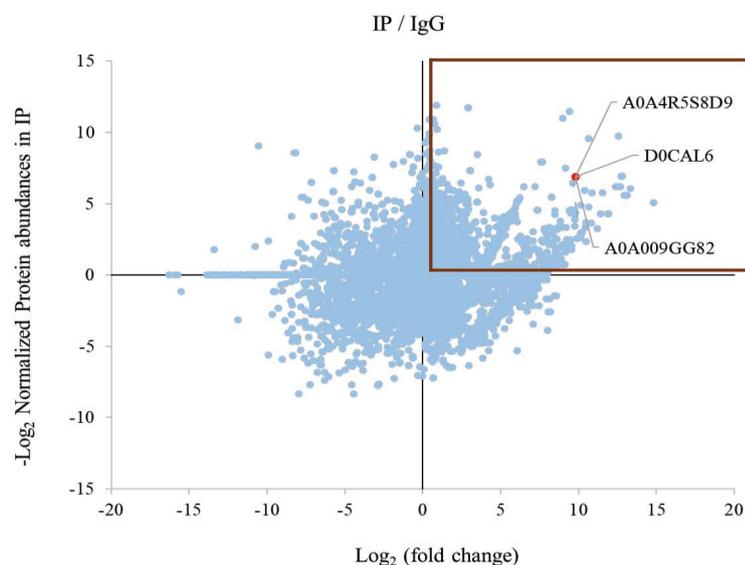
* Correspondence: alessandro.provenzani@unitn.it; fragai@cerm.unifi.it

Figure S1 1



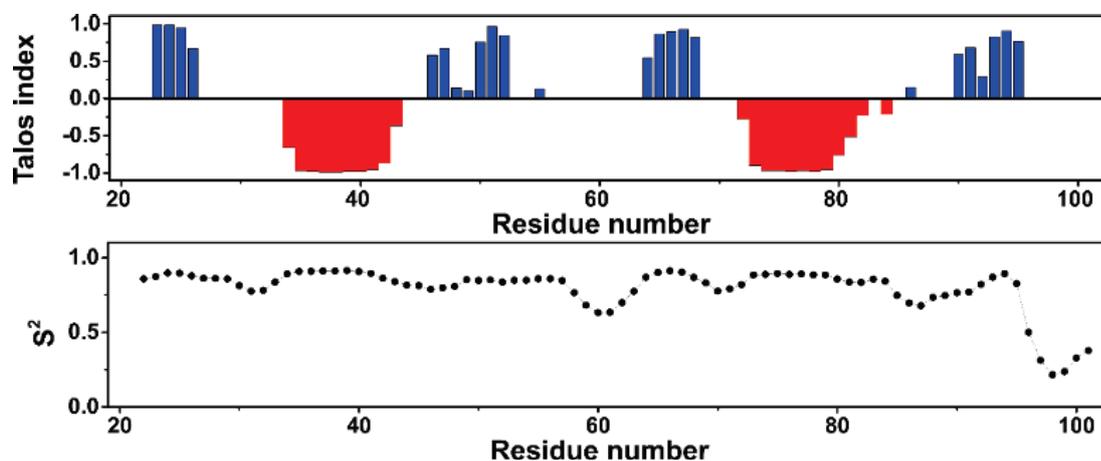
1) Sanger sequence of the retro-transcribed amplicon of the transcriptome of *Acinetobacter baumannii*. We used the reverse primer R3 that was designed 235 nucleotides downstream the gene of interest. This data confirmed the insertion of the mRNA sequence into a bigger polycistronic mRNA and confirmed the possible expression of the protein AB-Elavl.

2



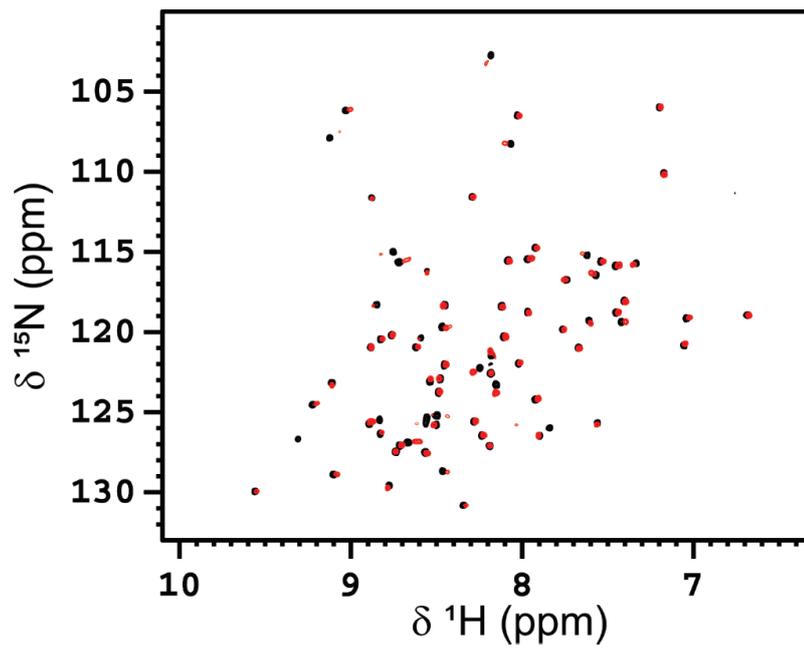
2) Scatter plot comparing protein enrichment (\log_2 FC IP-vs-IgG) compared with protein abundance within the IP proteome. Hypothetical and highly similar RNA binding proteins of *A. baumannii* are highlighted.

3



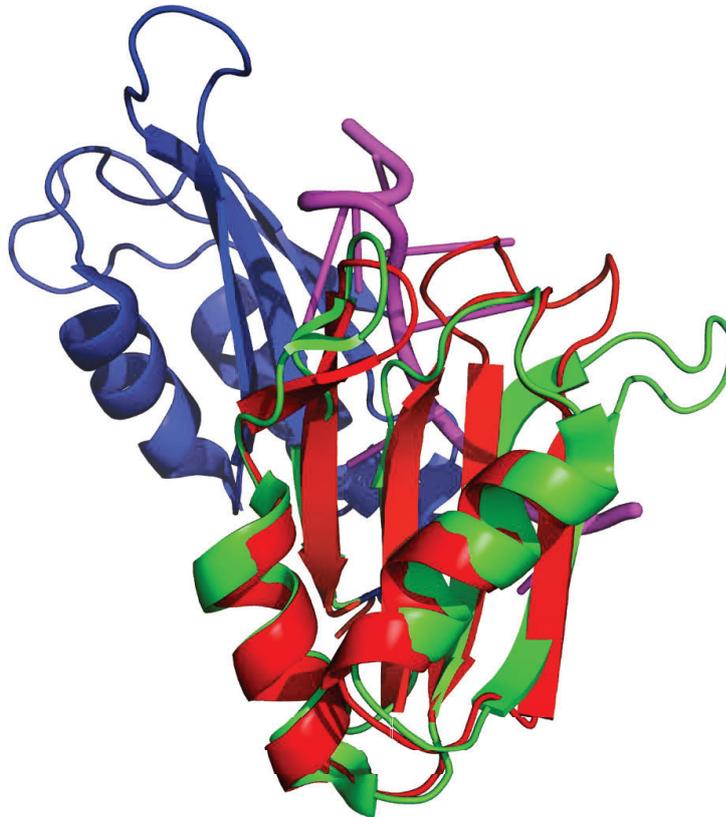
3) Top, secondary structure prediction obtained by the program Talos+ using the experimental values of chemical shifts of HN, N, C', C α , and C β atoms as input data. The blue bars indicate the β -strand propensity while the red bars the α -helix propensity. Bottom, predicted order parameter (S^2) by the program Talos+ using the experimental values of chemical shifts of HN, N, C', C α , and C β atoms as input data.

4



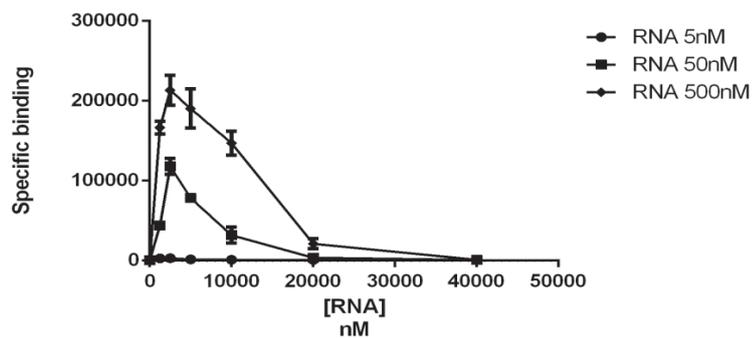
4) NMR analysis of the protein binding toward AREs. Superimposed 2D ¹H-¹⁵N HSQC spectra of free AB-Elavl RRM domain (70 μM , black) and in the presence of 140 μM of ARE Pos (red). The spectra were acquired on a spectrometer operating at 950 MHz and 298 K.

5



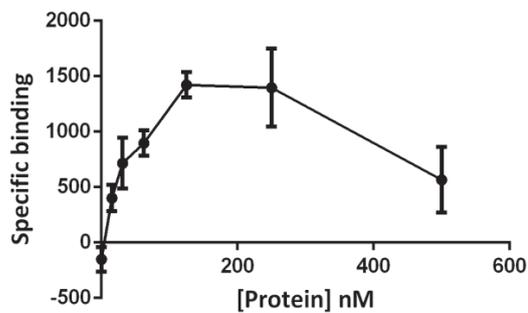
5) Superposition of the crystal structure of the bacterial hypothetical HuR RRM (rAB-Elavl, red) and 1FLX (green and blue) with the presence of RNA (magenta).

6



6) AlphaScreen for detection of the hook point on the recombinant protein and three different concentrations of AREpos probe, to have the best signal/noise ratio. [protein]: 250nM; [RNA]: 50nM.

7



7) Hook point established by HTRF-FRET between the recombinant protein and AREpos.

4.2 ARTICLE 3:

4.2.1 The interaction between Musashi-1 and RNA: a multiplayer game.

Anna Pérez-Ràfols, A.PR.,^{1,2} Guillermo Pérez-Ropero G.PR.,^{3,4} Rosa Anahí Higuera A.H.,⁵ Linda Cerofolini L.C.,^{1,6} Wolfgang Kaiser W.K.,⁵ Jos Buij J.B.,³ Claudio Luchinat, C.L.,^{1,2,6} Tommaso Martelli, T.M.² and Marco Fragai M.F.^{1,6*}

*Corresponding author: fragai@cerm.unifi.it

¹ Magnetic Resonance Center (CERM) and Department of Chemistry University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy

²Giotto Biotech S.R.L., Via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy.

³ Ridgeview Instruments AB, Vänge, Uppsala, Sweden

⁴ Sweden and Biochemistry Department, University of Uppsala, Sweden

⁵ Dynamic Biosensors GmbH, 82152 Martinsried, Germany.

⁶ Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy

In preparation

Abstract

RNA Recognition Motifs (RRMs) are key factors in the post-transcriptional regulation of gene expression in eukaryotic cells. Human Musashi-1 (MSI-1) is an RNA-binding protein that recognizes (G/A)U₁₋₃AGU and UAG sequences in diverse RNAs through two RRM domains and regulates the fate of target RNA. Understanding the interplay between the two domains in the mechanism of RNA binding is a prerequisite to develop ligands to hinder this interaction.

Here, we exploited an integrated approach to analyse the binding of the isolated domains and the tandem domain of human MSI-1 against two linear single stranded RNAs and two folded RNAs. A competition between the two RRM domains has been observed when recognizing the same RNA sequence, due to a similar affinity for the RNA motif. However, when a linear RNA strand with two binding sites is considered, an interaction of both RRM-1 and RRM-2 from a single MSI-1 has been observed. More intricately, instead, is the interaction with folded RNAs.

In particular, my contribution to this project has been the design of RNA sequences, the expression and purification of all constructs of human MSI-1 protein in *E. coli* cells (unlabelled, ¹H-¹⁵N and ¹H-¹⁵N-¹³C), I have been involved on the characterization of the MSI-1 protein in liquid NMR by solving the assignment of the tandem domain protein and on performing the protein-RNA interaction experiments, both with NMR and SEC-MALS. Regarding the kinetic studies, I have also been involved in the initial optimization phase of the SPR experiments.

Introduction

RNA binding proteins (RBPs) participate in each step of the RNA metabolism and play important roles in the regulation of gene expression.^{7,69} RBPs can have single or multi domains responsible for the recognition and binding of RNA. Within the most characterized RNA binding domain, we find the RNA Recognition Motif (RRM), also known as RBD,^{7,70} the most abundant RNA-binding domain in higher vertebrates and the most studied domain, both in terms of structure and biochemistry.⁴

RRMs typically comprise approximately 90 amino acids that consist of four-stranded antiparallel β -sheets packed against two alpha helices, adopting the canonical $\alpha\beta$ structure with a typical $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology^{4,7}. The two central β -strands once the domain is folded, β_3 and β_1 , contain two conserved sequences of eight and six amino-acids, called RNP1 and RNP2, respectively, that play important roles in nucleic acid recognition.¹ Because of their ability to interact with RNA, RNA Recognition Motifs (RRMs) are known to be key factors in the transcriptional and post-transcriptional gene expression in eukaryotic cells, and any dysregulation can result in a number of human disorders, including cancer and neurodegenerative diseases.⁷¹

The Musashi family is a widespread and highly conserved group of multi-domain RRM proteins with critical functions in stem cell maintenance and tumorigenesis. Overexpression of Musashi (MSI) protein has been found in several malignant tumors and, in some cases, a correlation between the levels of expression of the protein with the proliferative activity of cancer cells and a poor prognosis has been proposed.^{39,40} There are two members of the mammalian MSI family: Musashi-1 (MSI-1) and Musashi-2 (MSI-2) that share a 69% of sequence identity. MSI proteins contain N-terminal disordered region followed by two tandem RRM domains connected by a short inter-domain linker, about 10 amino acid long, and a large, disordered tail at C-terminal. Both RRM domains provide the surface for the interaction with RNA while the C-terminal region is known to bind poly(A)-binding proteins.³⁹

Musashi-1 is gradually down regulated during neural differentiation; it is involved in maintaining the undifferentiated state of neural stem cells through posttranscriptional control of downstream genes.^{38,43} It controls cell proliferation of cancer stem cells by modulating Notch and Wnt signalling pathways and promotes G0/G1 to S phase transition of cell proliferation by inhibiting translation of checkpoint regulators through the direct binding to p21, p27 and p53 mRNAs.⁴³ Therefore, any dysregulation of the expression or activation of the protein might lead to a disruption of these signalling pathways.

Information on the mRNA-bound structures of the mouse MSI-1 have been obtained from the complexes of the isolated domains with short single-stranded RNA fragments.^{32,34} In vitro SELEX experiments carried out by Okano and co-workers⁷² on mouse MSI-1 protein have provided information

on RNA binding specificity of each RRM. A consensus sequence (G/A)U₁₋₃AGU has been reported to have high affinity for the mouse MSI-1 RRM-1, whereas a preference for the UAG motif was observed for MSI-1 RRM-2.³⁴

However, nothing is reported for the human MSI-1 and, more important, the information on how the two RRMs interplay with RNAs of different length, composition and structure is completely missing. It is not trivial to unravel the mechanism of action of RRMs proteins, due to their versatile capability in their RNA recognition that often leads to complex systems.¹ Here, we have expressed in *E. coli* the two isolated domains (RRM-1 and RRM-2) and tandem domain human MSI-1 RRM₁₋₂ and characterized using solution NMR. To shed light on the binding mechanism and to dissect the different contributions, the interaction of the isolated RRMs and of the tandem domain with selected RNA strands was investigated with complementary biophysical methodologies. The integrated use of NMR spectroscopy, Surface Plasmon Resonance (SPR), SwitchSENSE[®] and Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) have provided information on the binding sites, stoichiometry and kinetic of the interactions.

Our results show a competition between the two RRMs present in the protein when recognizing the same RNA sequence, leading to an heterogeneous solution and thus, a more complex and dynamic interaction than expected.

Methods

RNA strands

Biotinylated RNA oligonucleotides (**Oligo-1**: 5'-CGG CGC CGC-3' and **Oligo-2**: 5'-UUGUUAGUUACCCCUU-3', **Oligo-4**: 5'-AAGCGUUAGUUAUUUAGUCGCUU-3' and **Oligo-5**: 5'-CACUCUGUAGUAUGUAGGGUUUAUUU-3') used for SPR kinetic experiments were purchased from Metabion, Planegg, Germany. Synthetic single stranded RNAs (**Oligo-2**: 5'-UUGUUAGUUACCCCUU-3', and **Oligo-3**: 5'-UUGUUAGUUAUUAGUU-3') and hairpin RNAs (**Oligo-4**: 5'-AAGCGUUAGUUAUUUAGUCGCUU-3' and **Oligo-5**: 5'-CACUCUGUAGUAUGUAGGGUUUAUUU-3'), for NMR experiments were purchased from Metabion international AG. Nucleic acid sequences (**Oligo-1**: 5'-CGG CGC CGC-3', **Oligo-2**: 5'-UUGUUAGUUACCCCUU -3', **Oligo-3**: 5'-UUGUUAGUUAUUAGUU-3', **Oligo-4**: 5'-AAGCGUUAGUUAUUUAGUCGCUU-3' and **Oligo-5**: 5'-CACUCUGUAGUAUGUAGGGUUUAUUU-3') used for switchSENSE[®] technology were purchased from Ella Biotech GmbH, Germany.

Expression and purification of recombinant Human Musashi-1 (MSI-1) RRM₁₋₂ domain

Recombinant human MSI-1 RRM₁₋₂ (1-200) protein in plasmid pET29b was overexpressed in BL21(DE3) GOLD *E. coli* cells. Cells were grown in LB or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Subsequently, expression was induced with 0.5 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 37 °C for 4 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer [50 mmol dm⁻³ Tris-HCl, pH 8.0, 500 mmol dm⁻³ NaCl, 10% glycerol, Proteases Inhibitor Cocktail (SIGMAFAST protease inhibitor cocktail tablets, EDTA-free)], ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was collected and a 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 buffer 20 mmol dm⁻³ Tris-HCl, pH 9.0, Proteases Inhibitor Cocktail.

Soluble protein was filtered with a 0.22 μm membrane and purified by an ion exchange chromatography using an Anion exchange Q FF 16/10 column previously equilibrated in 20 mmol dm⁻³ Tris-HCl, pH 9.0, Proteases Inhibitor Cocktail. Elution was performed against the buffer 20 mmol dm⁻³ Tris-HCl, pH 9.0, 1 mol dm⁻³ NaCl, Proteases Inhibitor Cocktail. The protein was collected on the flow-through. 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 50 mmol dm⁻³ Tris-HCl, pH 7.8, 150 mmol dm⁻³ NaCl, Proteases Inhibitor Cocktail. The purified fractions were collected, and the buffer was exchanged using a HiPrep 26/10 Desalting column, previously equilibrated with the RNase free buffer, for the NMR assignment experiments [20 mmol dm⁻³ MES, pH 6.0, 100 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] or the RNase free buffer for NMR titration experiments [50 mmol dm⁻³ Tris-HCl, pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail].

Expression and purification of recombinant the isolated domains RRM-1 and RRM-2 of Human Musashi-1 (MSI-1)

To produce the MSI-1 human RRM-1 construct (1-103), a pET29b plasmid containing a Strep-Tag at N-terminus followed by the first 261 residues of Musashi-1 human protein was modified by QuikChange Site-Directed Mutagenesis to replace Met-104 with a stop codon.

Recombinant human MSI-1 RRM-1 (1-103) protein in plasmid pET29b and MSI-1 RRM-2 (104-200) protein in plasmid pET21a were overexpressed in BL21(DE3) *E. coli* cells. Cells were grown in LB or M9 minimal media supplemented with ¹⁵NH₄Cl or ¹⁵NH₄Cl and ¹³C-glucose at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Subsequently, expression was induced with 0.5 mmol dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 37 °C for 3 h and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer [100 mmol dm⁻³ Tris-HCl pH 8.0,

150 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail], ruptured by sonication and separated by centrifugation at 30000 rpm for 35 min at 4 °C. Soluble fraction was filtered with a 0.22 µm membrane and purified using a Strep-Tag HP 5 mL column. Elution fraction was collected and purified to homogeneity by size exclusion chromatography using a Hi load 26/60 Superdex 75 pg column that was previously equilibrated in 50 mmol dm⁻³ Tris-HCl pH 7.5, 150 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail. The purified fractions were collected, and a treatment with 5% PEI solution was performed in order to remove DNA/RNA attached to the protein. Re-suspension of the RRM-1 protein was performed with RNase free final buffer [50 mmol dm⁻³ Tris-HCl, pH 7.5, 140 mmol dm⁻³ NaCl, 0.5 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] and filtered with a 0.22 µm membrane, whereas RRM-2 was resuspended in the RNase free final buffer [50 mmol dm⁻³ Tris-HCl, pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, Proteases Inhibitor Cocktail] due to stability conditions.

NMR measurements and protein assignment.

Experiments for the backbone resonance assignment (3D ¹H-¹⁵N-¹³C HNCA, HNCACB and HNCO) were performed on samples of the ¹³C, ¹⁵N isotopically enriched (MSI-1) RRM₁₋₂ domain at the protein concentration of 300 µmol dm⁻³ in buffer solution [20 mmol dm⁻³ MES pH 6.0, 100 mmol dm⁻³ NaCl, 1 mmol dm⁻³ DTT, 1 mmol dm⁻³ protease inhibitors]. NMR spectra were recorded at 298 K on a Bruker AvanceNEO NMR spectrometer operating at 1.2 GHz (¹H Larmor frequency) and equipped with a TCI 3 mm cryo-probe.

Spectra were processed with the Bruker TOPSPIN software packages and analysed with CARA (Computer Aided Resonance Assignment, ETH Zurich). The backbone resonance assignment of (MSI-1) RRM₁₋₂ was obtained comparing the assignments available in the literature for the individual domains from the mouse MSI-1 protein (BMRB code: 11450 and 36058)^{9,10} with the NMR spectra recorded on MSI-1 RRM₁₋₂ and analysing triple resonance spectra recorded on MSI-1 RRM₁₋₂.

R₁, R₂ and NOE measurements

Experiments for the determination of ¹⁵N longitudinal and transverse relaxation rates and ¹⁵N-¹H NOE were recorded on a Bruker AvanceNEO NMR spectrometer operating at 700 MHz (¹H Larmor frequency) using a ¹⁵N- enriched sample of the human MSI-1 RRM₁₋₂ tandem domain protein at 298 K [protein concentration of 140 µmol dm⁻³ in 20 mmol dm⁻³ MES pH 6.0, 100 mmol dm⁻³ NaCl, 1 mmol dm⁻³ DTT, 1 mmol dm⁻³ protease inhibitors]. ¹⁵N Longitudinal relaxation rates (R₁) were measured using a sequence modified to remove cross-correlation effects during the relaxation delay.⁷³ Inversion recovery times ranging between 20-2000 ms, with a recycle delay of 3.5 s, were used for the experiments. ¹⁵N transverse relaxation rates (R₂) were measured using a Carr-Purcell-Meiboom_Gill (CPMG) sequence^{73,74}, with delays ranging between 16.96 and 220.48 ms and a re-focusing delay of

450 μ s. Longitudinal and transverse relaxation rates were determined by fitting the cross-peak intensities as a function of the delay to a single-exponential decay using the Topspin Dynamic Center software. Heteronuclear NOE values were obtained from the ratio of the peak height for ^1H -saturated and unsaturated spectra.

Titration of Musashi with RNA strands.

The effect of two linear single stranded RNA constructs (**Oligo-2**: 5'-UUGUUAGUUACCCCUU-3') and (**Oligo-3**: 5'-UUGUUAGUUAUUAGUU-3') was evaluated on ^{15}N -isotopically enriched MSI-1 RRM₁₋₂, MSI-1 RRM-1 and MSI-1 RRM-2 proteins at the concentration of 100 $\mu\text{mol dm}^{-3}$ in the following experimental conditions: 50 mmol dm^{-3} Tris-HCl, 140 mmol dm^{-3} NaCl, 1 mmol dm^{-3} EDTA, 1 mmol dm^{-3} protease Inhibitors. The interaction with two other RNA strands **Oligo-4** (5'-AAGCGUUAGUUAUUUAGUCGCUU-3') and **Oligo-5** (5'-CACUCUGUAGUAUGUAGGGUUUAUUU-3') able to form a hairpin, was also investigated under the same experimental conditions. The pH was 7.2 in the case of MSI-1 RRM₁₋₂ and MSI-1 RRM-2, while 7.5 in the case of MSI-1 RRM-1, taking into account its lower isoelectric point. 2D ^1H ^{15}N HSQC and 2D ^1H ^{15}N TROSY NMR spectra were recorded at 298 K on the single domains and on the tandem domain, respectively, using a Bruker AvanceNEO NMR spectrometer, operating at 900 MHz (^1H Larmor frequency). During the NMR titration, increasing amounts of the RNA strands were added to the protein solution to reach the final concentrations of 6, 12, 25, 50, 120, 200 $\mu\text{mol dm}^{-3}$ of RNA. Spectra were processed with the Bruker TOPSPIN software packages and analysed with CARA (Computer Aided Resonance Assignment, ETH Zurich).

Size-Exclusion Chromatography with Multi-Angle Light scattering (SEC-MALS):

The isolated free MSI-1 RRM-1 and RRM-2 domains (140 $\mu\text{mol dm}^{-3}$) and in complex with oligo-2, oligo-3, oligo-4 and oligo-5 (in 1:0.5 and 1:1 protein/RNA molar ratio), were dissolved in 50 mmol dm^{-3} Tris-HCl, pH 7.5, 140 mmol dm^{-3} NaCl, 0.5 mmol dm^{-3} EDTA, 1 mmol dm^{-3} protease inhibitors, and used to perform SEC-MALS experiments. The effect of additional sub-stoichiometric concentrations of the RNA with respect to the protein was also evaluated in case of oligo-4 and oligo-5 (in 1:0.12, 1:0.25 protein/RNA molar ratio),

Samples of 100 μL were loaded at 0.6 mL/min on a Superdex 200 10/300 GL analytical size-exclusion column (GE Healthcare), and elution was monitored by the following in-line detectors: a light scattering diode array (DAWN EOS, Wyatt Technology UK Ltd.), a dynamic module (WYATT QELS, Wyatt Technology UK Ltd.), UV detector (Smartline UV Detector 2500, Knauer) and a differential refractive index detector (Optilab rEX, Wyatt Technology UK Ltd.).

Chromatograms were analysed using the ASTRA software (v7.3.2.19, Wyatt Technology UK Ltd.) and the interaction chromatograms were analysed using the Protein Conjugate template.

Parameters of the specific refractive index increment dn/dc (mL/g) and UV Extinction coefficient (mL/(mg·cm)) of each domain (RRM-1 and RRM-2) and of the modifiers (oligo-2, oligo-3, oligo-4, and oligo-5) are found in the Table S1.

Surface Plasmon Resonance (SPR)

SPR experiments were performed with a Biacore 3000 (Cytiva, Uppsala, Sweden). RNA oligonucleotides were biotinylated in the 3' end and immobilized to a streptavidin coated SPR chips (SAD200M, Xantec, Dusseldorf, Germany). Streptavidin excess was removed by five injections of 10 μ L of 50 mmol dm^{-3} NaOH, 1 mol dm^{-3} NaCl with a flow rate of 10 μ L/min. After that, RNA oligos diluted in running buffer (50 mmol dm^{-3} Tris pH 7.2, 140 mmol dm^{-3} NaCl, 0.5 mmol dm^{-3} EDTA, 0.05%(v/v) Tween) were immobilized on the streptavidin-coated chips by flowing them over the chip with a 2 μ L/min flow rate until a signal of 80-90 RU was attained. Oligos were immobilized in flow channels 1 to 3 and channel 4 was left as a reference. A 1 μ mol dm^{-3} stock solution in running buffer was prepared for each protein and further diluted 2-fold to create a concentration series down to 2 nmol dm^{-3} . Kinetic experiments were performed at 25 °C with a 100 μ L/min flow rate for human Musashi-1 RRM₁₋₂ and 60 μ L/min for RRM-1 and RRM-2, respectively. An assay comprised a 2 min association phase followed by 3 min dissociation for each concentration. After each assay, the chip was regenerated by injecting 10 μ L of 2 mol dm^{-3} $MgCl_2$, 2 mol dm^{-3} NaCl with a 10 μ L/min flow rate.

Data analysis was performed using TraceDrawer 1.9.2 (Ridgeview Instruments, Uppsala, Sweden) by using the background subtracted curves and applying different mathematical fittings. The simplest interaction model that describes the reversible interaction between a ligand (L) and a target (T) is the so-called 1:1 Langmuir binding model (Eq. 1). This model assumes that an interaction is described by a single association rate constant (k_a) representing the molecular recognition and a single dissociation rate constant (k_d) that reflects the stability of the formed complex. The strength of the interaction given by the affinity (K_D) and is equal to the ratio of the rate constants; k_d/k_a . When the ligand concentration is higher than the target concentration, the affinity represents the concentration at which 50% of the targets is bound by a ligand.



In real-time interaction analysis, the measured signal (B) is proportional to the number of complexes. Given that the total number of targets is constant, Eq. 1 can be rewritten as a function of time (t) and ligand concentration ([L]) to describe the change in the number of complexes.

$$\frac{\delta B}{\delta t} = k_a \cdot [L] \cdot (B_{max} - B) - k_d \cdot B \quad [2]$$

More complex interaction models are required to describe heterogenous interactions with more than one k_a and k_d . For example, the 1:2 model assumes that the ligand can bind two different types of

target populations, resulting in two independent 1:1 interactions with distinct kinetic parameters while the bivalent model assumes that, after being bound to one target, all ligands have the capacity to bind a second target of the same type and are thereby stabilized.

InteractionMap (IM) (Ridgeview Instruments, Uppsala, Sweden) is a data analysis method that estimates the presence and rate constants of multiple 1:1 interactions and their contribution to the experimental binding curves. In the resulting InteractionMap, all individual interactions are presented in a k_a (x-axis) - k_d (y-axis) plot with heat-map coloration to indicate the relative abundance of each interaction to the overall experimental data.

A 9 nucleotide RNA sequence (oligo-1) that didn't display the binding motif was also used as a negative control to test if the protein binding was truly specific.

switchSENSE®

All experiments were performed on adapter chips on the heliX® biosensor platform (Dynamic Biosensors GmbH, Germany) using switchSENSE® proximity sensing^{75,76} measurement mode (details of the apparatus are reported in the Supplementary Material).

Kinetic experiments were performed at 25°C with a 100 $\mu\text{L}/\text{min}$ flow rate for human Musashi-1 RRM₁₋₂ in buffer 50 mmol dm^{-3} Tris, 140 mmol dm^{-3} NaCl, 1 mmol dm^{-3} EDTA, 0.05% Tween 20 and 1 mmol dm^{-3} TCEP at pH 7.2. An assay comprised a 5 min association phase followed by 20 min dissociation for each concentration.

Data analysis was performed in heliOS® software (Dynamic Biosensors GmbH), where association and dissociation rate constants, as well as the derived equilibrium dissociation constants K_d , were calculated by performing a global fitting using monophasic or bi-phasic binding models. All data was normalized and referenced with blank injection. Data for MSI-1 RRM₁₋₂ was real-time referenced with the spot 2 carrying a negative control sequence oligo-1.

Results

NMR characterization of free MSI.

The 2D ^1H - ^{15}N -HSQC spectra of the three constructs (MSI RRM₁₋₂ tandem domain, RRM-1 and RRM-2 isolated domains), are superimposed in Figure 1. The spectra show well-dispersed signals in agreement with uniform and folded protein structures. The NMR spectrum of MSI RRM₁₋₂ tandem domain is largely superimposable to the spectra of the isolated domains, as the majority of the signals in the spectrum of MSI RRM₁₋₂ overlap with either a signal of RRM-1 or RRM-2. The absence of a large chemical shift perturbation when passing from the isolated domains to the tandem construct suggest a large interdomain flexibility with few interactions, occurring among the two domains and with

the linker, in the full-length protein. In the spectrum of MSI RRM₁₋₂, a number of additional signals, that can be attributed to the portion of the interdomain linker, is also visible.

The backbone assignment of the MSI RRM₁₋₂ tandem domain was obtained starting from the published assignment of the murine isolated RRMs domains (BMRB codes: 11450 and 36058)^{32,34} considering the high sequence homology between the mouse and human MSI-1 protein (99.44%). This was then confirmed and complemented by the analysis of triple resonance spectra. 92.5% of all residues including those forming the linker region, have been assigned. The present assignment is the most complete. The assignment for the single domains was as well obtained from the data available in the BMRB and by superimposition of the spectra with the spectra of the MSI RRM₁₋₂ tandem domain. The 95.14% and the 90.72% of the total sequences have been assigned for RRM-1 and RRM-2, respectively.

To characterize the protein dynamics, measurements of longitudinal (R_1) and transvers (R_2) relaxation rates of the backbone amide nitrogen were performed on a ¹⁵N – enriched sample of the MSI-1 RRM₁₋₂ construct. Theoretical estimates of R_1 and R_2 values and of the rotational correlation time were calculated by HydroNMR from the X-ray structures of the isolated RRM-1 (residues 20-96, PDB 2RS2) and RRM-2 (residues 110-186, PDB 5X3Z) domains and from a model of the RRM₁₋₂ tandem domain (residues 20-186) and are displayed together with the experimental data in Figure S1. The experimental rotational correlation time obtained from R_1 , R_2 and NOE experimental data (8.4 ± 0.4 ns) is in between the calculated values for the isolated (5.8 ns for both domains) and tandem domains (14.9 ns). This indicates that the RRM1–RRM2 construct does not behave as a rigid body but displays instead interdomain flexibility, simulating a protein of lower molecular weight. In details, the comparison of experimental R_1 and R_2 data with theoretical values calculated for the isolated RRM-1 and RRM-2 domains shows that experimental R_1 values are smaller and R_2 values are larger than their theoretical counterparts (Figure S1, panels A, B). At the same time, experimental R_1 values are higher and experimental R_2 values are lower than theoretical estimates calculated for the MSI-1 RRM₁₋₂ construct (Figure S1, panels C, D), indicating interdomain dynamics. Further indication of the presence of interdomain flexibility is provided by the ¹⁵N–¹H NOE for the linker residues, experiencing values lower than 0.5 (from Ala95 up to Lys109, see Figure S1, panel E). The small NOE values of the residues in the linker between the two domains are evidence of fast motions on ps-ns timescale (faster than the overall protein-tumbling rate). Very small NOE values are found also for the N- and C-terminal tails (from Thr3 up to His17 and from Gln185 up to Ser200), indicating fast motions also at this level.

Interaction of the isolated MSI-1 RRM1 (1-103) and MSI-1 RRM2 (104-200) domains with a linear single-binding site RNA (Oligo-2).

First the interaction of Musashi-1 with linear RNA strands was investigated. The effect on the isolated RRMs domains due to an oligonucleotide with a single consensus binding motif (G/AU₁₋₃AGU)⁷² has been initially analysed. ¹³During the NMR titration, upon the addition of increasing

concentrations of the oligo-2 to the solution of RRM-1, the intensity of the signals of the free protein in the 2D ^1H - ^{15}N HSQC decrease in intensity, while new cross-peaks, corresponding to RRM-1 in complex with oligo-2, appear and increase in intensity (Figure 2). This indicates that the interaction between RRM-1 and oligo-2 is in the slow exchange regime on the NMR timescale. As expected, the signals of the free protein experiencing the largest decreases in intensity after the addition of RNA at the concentration of 25 μM to the protein solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platform as well as in loop-1 (β_1 - α_1) and loop-3 (β_2 - β_3) of the isolated domain (Lys21, Ile24, Gly26, Ser28, Trp29, Gln30, Thr31, Arg61, Gly64, Phe65, Asp91, Ala95 and Phe96, Figure 3, panel A and C). In the presence of RNA at the concentration of 100 μM (1:1 protein/RNA molar ratio), the signals of the free protein have completely disappeared, while the signals of a new single homogeneous species are visible. The new signals have been partially reassigned and the analysis of the chemical shift perturbation (CSP) performed (Figure 3, panel B). The residues corresponding to the assigned signals experiencing the largest changes (Lys21, Met22, Gln30, Thr31, Arg61, Gly64, Lys93 and Val 94), are located in the β -platform as well as in loop-1 (β_1 - α_1) and loop-3 (β_2 - β_3) of the isolated domain and correspond to the already described binding area for the isolated domain (PDB: 2RS2,³²). Unfortunately, many signals could not be reassigned because of their too large CSP. These signals belong as well to residues located on the β -platform (Phe23, Ile24, Gly26, Trp29, Glu30, Thr31, Leu56, Thr89, Asp91, Ala95, Phe98 and Ala100).

The interaction of RRM-2 with oligo-2 was then performed. Unexpectedly, the same behaviour of RRM-1 was also observed for RRM-2, with the decrease in intensity of the signals of the free protein during the NMR titration and the appearance of new cross-peaks (Figure 2). Also in this case, the signals of the free protein experiencing the largest decreases in intensity after the addition of RNA at the concentration of 25 μM to the protein solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platform as well as in loop-1 (β_1 - α_1) and loop-3 (β_2 - β_3) of the isolated domain (Ile111, Val113, Gly115, Leu116, Asp124, Leu140, Gly153, His172, Glu180, Lys182, Ala184, Gln185, Glu188 and Val189, Figure 3, panel A and C). The analysis of the chemical shift perturbation (CSP) at the end of the titration (Figure 3, panel B), also showed that the reassigned residues experiencing the largest changes (Lys110, Ile111, Val113, Gly114, Leu140, Thr146, Asp161 and Met150) are located in the same region. Also, for RRM-2 many signals of the β -platform could not be reassigned in the spectrum of the complex because of their too large CSP (Thr108, Gly115, Leu116, Val118, Met139, Gly153, His172, Lys182, Ala184, Glu188 and Val189).

Binding was corroborated in both cases by SEC-MALS analysis, which also provided information about the binding stoichiometry. First, the isolated domains and oligo-2 were analysed separately to ascertain the monomeric form of each isolated macromolecule (Suppl. Figure 3). Then, the interaction of each RRM with oligo-2 at the protein/RNA molar ratios of 1:0.50 and 1:1 was evaluated.

Protein Conjugate Analysis shows a clear 1:1 interaction, confirming the ability of both domains to bind the same RNA sequence (Figure 4).

To decipher the kinetics of both RNA recognition motifs, RRM-1 and RRM-2, with oligo-2, we performed different Surface Plasmon Resonance (SPR) experiments. Both RRM-1 and RRM-2 showed a rapid and rather similar association and dissociation rate when binding to oligo-2 that are well described by a 1:1 model as shown in Figure 5A and 5B. The association rate constants (k_a) were $2.32 \text{ M}^{-1}\text{s}^{-1}$ (CV 31.14%) for RRM-1 and $2.86 \text{ M}^{-1}\text{s}^{-1}$ (CV 24.37%) for RRM-2, with the respective dissociation rate constants (k_d) being $4.7\text{E-}2 \text{ s}^{-1}$ (CV 10.31%) and $5.26\text{E-}2 \text{ s}^{-1}$ (CV 24.37%). Affinities (K_D), as calculated from the ratio between the dissociation and association rate constants, were also very similar and amounted to $22.3 \text{ nmol dm}^{-3}$ (CV 33.42%) for RRM-1 and $19.6 \text{ nmol dm}^{-3}$ (CV 40.98%) for RRM-2. Results from the InteractionMap (IM) calculations confirm that the RRM-1 and RRM-2 interactions are well described by a 1:1 model as a single binding event is shown for both RRMs (Figure 5) with affinities of $87.2 \text{ nmol dm}^{-3}\text{M}$ (CV 7.23%) and $62.8 \text{ nmol dm}^{-3}$ (CV 16.9%) respectively. No binding of the RRMs to the negative control was observed (data not shown).

Recognition of Oligo-2 by human MSI-1 RRM₁₋₂ tandem domain.

Next, we studied the interaction of the MSI-1 RRM₁₋₂ tandem domain with oligo-2 by solution NMR. With this construct, upon increasing the concentration of RNA, some cross-peaks of the free protein broaden and decreased in intensity (Figure 2) (at 1:0.25 protein/RNA molar ratio, mostly affected residues are: Ser15, Ile24-Gly26, Ser28, Leu50, Arg53, Gly64, Phe68, Lys76, Val94-Ala95, Thr106, Ile111, Gly115, Met139, Asp143, Lys183, Gln185 and Glu188-Val189, Figure 6, panel A and B) with some signals affected by only small CSP (Asp14, Met22, Gly25-Gly26, Trp29, Gly41, Lys58, Phe63, Phe68, Ile90, Val94-Ala95, Thr106-Arg107, Val113, Val118-Asn119, Arg150, Gly153, Glu160, Val166, Lys183-Ala184, Glu188, Thr193 and Arg197-Gly198, Figure 6, panel A and B). However, the cross-peaks of a new species could not be clearly detected; even with an excess of RNA with respect to the protein (protein/RNA molar ratio of 1:2), we could hardly see few new signals with very low intensity. Interestingly, the signals of residues belonging to the beta sheet surface of both domains are affected by changes after the addition of oligo-2. This behaviour suggests a competition between the two domains for the binding of oligo-2.

Kinetics of the tandem domain protein against oligo-2 were also performed with SPR. The 200 amino acid-long Musashi-1 protein, that includes both RRM 1 and 2, showed a complex binding trace to oligo-2 that didn't follow a 1:1 interaction, as it is shown in Figure 7, panel A. The dissociation phase is clearly biphasic with a rapid initial dissociation that slows down considerably after roughly one minute. Interestingly, at higher binding levels prior to starting dissociation, a larger fraction of the bound Musashi-1 dissociates rapidly. Although the isolated RRMs did not show a significant difference in their binding kinetics the two binding domains of Musashi-1 could independently bind the RNA sequence

that potentially explains the biphasic dissociation pattern. Therefore, the binding curves were fitted with a 1:2 model. This model fitted the data well as shown in Figure 7, panel B. Similar k_a -values were obtained for both interactions (k_{a1} $1.6E+6$ $M^{-1}s^{-1}$ (CV 23.43%) and k_{a2} $2.91E+6$ $M^{-1}s^{-1}$ (CV 20.87%)) with the major differences observed in the k_d values (k_{d1} $6.21E-02$ s^{-1} (CV 17.71%) and k_{d2} $1.17E-03$ s^{-1} (CV 36.1%)). The calculated affinities were $K_{D1} = 40.6$ $nmol\ dm^{-3}$ (CV 23.53%) and $K_{D2} = 1.07$ $nmol\ dm^{-3}$ (CV 18.21%). Based on the obtained data, both binding events contributed equally to the observed binding traces at higher concentrations while, at lower concentrations, the one displaying the higher affinity becomes dominating. This indicates that the recognition of both interactions is rather similar but that a fraction of the bound Musashi-1 is stabilized. This stabilized fraction is dependent on the concentration used during dissociation indicating that the interactions are not independent as assumed for the 1:2 model. The Interaction Map (IM) analysis shown in Figure 7, panel A confirms the results by showing two different binding events with similar recognition values with a clear peak (with an affinity of 21.5 $nmol\ dm^{-3}$) for a rapid dissociating fraction while the other peak (with an affinity of 1.53 $nmol\ dm^{-3}$), for the slower dissociating fraction, is smeared out over a range of dissociation rates.

A hypothesis that could explain the binding traces is that MSI-1, upon binding to one RNA, can bind another RNA strand with its second domain, thereby stabilizing the interaction. At high concentrations, MSI-1 binds rapidly thereby limiting the number of unbound RNA for the bivalent binding. Upon dissociation, however, free RNA strands become available for bivalent binding, thereby stabilizing the interaction and slowing down the dissociation rate. At low concentrations, Musashi-1 is able to bind in a bivalent manner during the association phase. In view of the similarities of both RRM domains to bind oligo-2, fitting the experimental data to a bivalent model (2:1) would be more appropriate. Unfortunately, this model is not capable of extracting all rate constants in a meaningful manner. No binding of the MSI-1 RRM₁₋₂ to the negative control was observed (data not shown).

Additional kinetic experiments with switchSENSE® were performed on the tandem domain protein. MSI-1 RRM₁₋₂ showed a kinetic profile with two different binding interactions, presented in the right panel of Suppl. Figure 4A. The interaction 1 has ≈ 4 -fold and ≈ 8 -fold faster on-rate and off-rate, respectively, compared to interaction 2 (see Suppl. Figure 4B). Moreover, the interaction 1 has the strongest contribution to the binding interaction (represented as a bigger data point on the rate plot map). The K_D of the interactions is 0.445 ± 12 $nmol\ dm^{-3}$ and 0.252 ± 10 $nmol\ dm^{-3}$, respectively.

Interaction of isolated and tandem RRM domains with a linear RNA with two-binding sites (Oligo-3).

After having detected this competitiveness effect, we next wondered how the RRM domains, both as isolated and in tandem, would behave with a single stranded RNA that contained two binding sites: one site with the consensus sequence known to bind the first RRM, G/AU₁₋₃AG, and one with the UAG

motif, known to bind the second RRM. In order to test if both RRMs would be able to bind both binding sites, we first investigated through solution NMR the interaction of the isolated RRM-1 and RRM-2 domains against oligo-3.

During the NMR titrations, upon the addition of increasing concentrations of oligo-3 to either RRM-1 or RRM-2 domain, a decrease in the intensity of the cross-peaks of the free protein is observed, while new cross-peaks, corresponding to the RRMs in complex with oligo-3, appear and increase in intensity (Figure 2). Therefore, also the interaction of RRM-1 and RRM-2 with oligo-3 is in the slow exchange regime on the NMR timescale. As observed for oligo-2, the signals of the free proteins experiencing the largest decreases in intensity also after the addition of oligo-3 at the concentration of 25 μM to the proteins solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platform (RRM-1: Lys21, Met22, Gly26, Leu50, Arg61, Gly64, Phe65, Val94, Ala95 and Phe96; and RRM-2: Thr108, Ile111, Gly115, Leu116, Met139, Leu140, Val155, Val166, Ile174, Glu180, Lys182, Gln185, Glu188 and Val189, Figure 8, panel A and C). Interestingly, after the addition of oligo-3 at the concentration of 50 μM (1:0.5 protein/RNA molar ratio), for both RRMs, the signals of the free protein have almost completely disappeared, while the new appeared signals are very broad. This phenomenon can be explained by the occurrence of multiple heterogenous species in solution. Indeed, RRM-1 and RRM-2 can bind both RNA binding sites, with the formation of complexes with different protein/RNA stoichiometric ratios. Instead, in the presence of RNA at the concentration of 100 μM (1:1 protein/RNA molar ratio) the linewidth of these new signals sharpens (Figure 2). The CSP between the resonances of the new partially reassigned signals of the complexes and the resonances of the free proteins is displayed in Figure 8, panel B and C. The residues corresponding to the assigned signals experiencing the largest changes (RRM-1: Lys21, Gln30, Thr31, Arg61, Gly64, Lys93, Val94 and Ala100; and RRM-2: Lys110, Ile111, Val113, Gly114, Leu140, Asp143, Thr146 and Asp161), are located in the β -platform. Unfortunately, also in this case, many signals could not be reassigned because of their too large CSP. These signals belong as well to residues located on the β -platform (RRM-1: Phe23, Ile24, Trp29, Thr89, Asp91, Ala95 and Phe96; and RRM-2: Arg107, Thr108, Gly115, Leu116, Met139, Met141, Gly152, Val155, His172, Glu180, Lys182, Ala184, Glu188, Val189). As expected, most of the affected residues are in common between oligo-2 and oligo-3.

SEC-MALS analysis was also carried out to shed light on the stoichiometry of the interaction between RRM-1 and oligo-3 and support the NMR data. The interaction of RRM-1 with oligo-3 was evaluated at the protein/RNA molar ratio of 1:0.50. The Protein Conjugate Analysis confirmed the presence of two species in solution, the most abundant one with a 2:1 protein/RNA stoichiometry and a minor one with a 1:1 protein/RNA stoichiometry Suppl. Figure 5. Moreover, no free protein was observed.

Next, we proceeded to see how the tandem domain RRM1-RRM2 protein would interact with the single stranded oligo-3 and if the competitiveness effect would be observed. A decrease in the intensity of the cross-peaks of the free protein is observed also during this NMR titration. New cross-peaks, corresponding to the MSI-1 RRM_{1,2} tandem domain in complex with oligo-3, appear and increase in intensity as well (Figure 2). Therefore, the interaction of the full-length construct with oligo-3 is in the slow exchange regime on the NMR timescale, contrary to what has been observed for oligo-2. The signals of the free protein experiencing the largest decreases in intensity after the addition of oligo-3 at the concentration of 25 μM to the protein solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platforms of both RRM domains (Lys21, Met22, Gly26, Leu50, Arg61, Gly64, Phe65, Val94, Ala95 and Phe96, Thr108, Ile111, Gly115, Leu116, Met139, Leu140, Val155, Val166, Ile174, Glu180, Lys182, Gln185, Glu188 and Val189, Figure 9). Interestingly, after the addition of oligo-3 at the concentration of 50 μM (1:0.5 protein/RNA molar ratio), the signals of the free protein have completely disappeared, while the new appeared signals are very broad (Figure 2). In the presence of RNA at the concentration of 100 μM (1:1 protein/RNA molar ratio), the new signals are still broad, thus the absence of multiple conformational states cannot be excluded. Assignment of the new shifted signals is unfeasible because of the higher ambiguity due to the higher crowding in the spectra of the MSI-1 RRM_{1,2} tandem domain.

Kinetic experiments for the MSI-1 RRM_{1,2} with oligo-3 were performed with switchSENSE® technology. The tandem protein showed a kinetic profile with two different binding interactions, presented in the right panel of Figure 10A. The interaction 2 has ≈ 4 -fold faster on-rate and a remarkable ≈ 600 -fold slower off-rate, respectively, compared to interaction 1 (see Figure 10B). Moreover, the interaction 2 has the strongest contribution to the binding (represented as a bigger data point on the rate plot map). This combined effects lead to a K_D of $1.55 \pm 0.08 \text{ nmol dm}^{-3}$ and $687 \pm 13 \text{ fmol dm}^{-3}$, accordingly. This seems to indicate a very strong cooperativity effect only present when both domains are fully functional in the tandem protein, combined with the accessibility of both binding sites on the RNA ligand.

Interaction of isolated and tandem RRM domains with folded RNAs bearing two-binding sites (Oligo-4 and Oligo-5).

The binding of Musashi-1 to folded RNA strands (i.e., hairpins) was then investigated through solution NMR. Oligo-4 and oligo-5 were selected because of the different location of the consensus motifs, known to bind RRM-1 and RRM-2, respectively, on these RNA sequences. Indeed, oligo-4 is an RNA fragment from the numb mRNA⁷² that contains both the binding motifs (the G/AU₁₋₃AGU motif for RRM-1⁷² and the UAG motif for RRM-2³⁴) within the loop region of a hairpin folding. Conversely, oligo-5 is an RNA fragment from the doublecortin (dcx) mRNA⁴⁴ and contains one consensus motif

(G/AU₁₋₃AGU) in the loop region, while the second one (UAG) in the double stranded region of the hairpin.

The isolated RRM domains display a similar behaviour to what observed with the linear RNA sequence with two binding sites (oligo-3). The interactions of isolated RRM-1 and RRM-2 with both oligo-4 and oligo-5 are in the slow exchange regime on the NMR timescale (Figure 2). The effects due to oligo-4 and oligo-5 are also located on the β -platforms of the isolated domains (Figure 11 and 12). Also in these cases, the formation of multiple heterogenous species cannot be excluded in the presence of RNA in sub-stoichiometric concentrations with respect to the protein.

SEC-MALS experiments were performed to better shed light on the stoichiometry of the complexes. First, the isolated RRM-1 domain and oligo-4 and oligo-5 were analysed separately to ascertain the monomeric form of each isolated macromolecule (Suppl. Figure 6). For both the oligo-4 and RRM-1, a monomeric form was detected, however, aggregates of RNA were observed with oligo-5. Then, the interaction of RRM-1 with oligo-4 and oligo-5 at different protein/RNA molar ratios was evaluated. The presence of species with different reciprocal protein/RNA stoichiometric ratio was observed in solution (Figure 13), confirming the heterogeneity detected through NMR experiments. We selected the chromatogram corresponding to an interaction of RRM-1 and oligo-4 with a molar ratio of 1:0.5 and attempted to analyse it with the Protein Conjugate Analysis. A mixture of complexes with the protein/RNA stoichiometric ratios of 2:1 and 1:1 was detected. Additionally, free RNA was present. However, the amount of each component cannot be determined with precision (Suppl. Figure 7), due to the overlapping of the peaks.

Kinetic experiments for the isolated domains with oligo-4 and oligo-5 were performed with SPR (Figure 14, panel A and B section i and ii). Binding of RRM-1 and RRM-2 to both hairpin structures fitted well with a 1:1 model. Also, InteractionMap analysis (Figure 14, panel A and B section i and ii) displays a single peak that indicates that the experimental data are dominating by an interaction with distinct dynamics and affinity. These results indicate that, even if multiple binding motifs exist on the hairpin structures, RRM-1 and RRM-2 are primarily binding to one of them or, alternatively, to both of them with similar kinetics. The binding affinities of RRM-1 and RRM-2 for both hairpin structures are in the 20-150 nM range. Regarding RRM-1, binding to RNA oligo-4 was two-times faster than RNA oligo-5, with both having a similar dissociation rate. As a result, the calculated affinity values (K_D) revealed a higher affinity of RRM-1 for RNA oligo-4 than oligo-5, as it is shown in Table 1. For both RNAs, RRM-2 had similar association and dissociation rates. Accordingly, the calculated affinities were similar, as shown in Table 1.

Recognition of Oligo-4 and Oligo-5 by human MSI-1 RRM₁₋₂ tandem domain.

We next investigated the interaction of oligo-4 and oligo-5 with the tandem domain MSI-1 RRM₁₋₂ protein through solution NMR. In the NMR titration of the MSI-1 RRM₁₋₂ with oligo-4 we observed only a decrease in signal intensity without the appearance of new cross-peaks (Figure 2). In the presence of RNA, at the protein/RNA molar ratio of 1:0.25, a general decrease of the protein signals is observed (Figure 15) with some residues experiencing a larger effect (Ser15, Ile24, Gly25, Gly26, Leu50, Met52, Gly64, Phe65, Val66, Val94, Ala95, Arg107, Thr108, Val113, Gly115, Leu140, Met141, Phe142, Thr146, Arg150, Lys183, Ala184, Gln185, Glu188 and Met190) and few residues experiencing minor CSP (Asp14, Lys21, Met22, Ile24, Cys49, Met52, Lys58, Gly64, Phe65, Phe68, Ala100, Lys109, Gly115, Lys134, Gly153, Phe154, Val155, Ile162, Glu164, Lys165, Val166, Glu180, Ala184, Gln185, Glu188 and Met190).

In the titration of MSI-1 RRM₁₋₂ tandem domain with oligo-5, instead, in the presence of the sub-stoichiometric concentrations of oligo-5 with respect to Musashi-1, the appearance of new cross-peaks corresponding to the protein in complex with the RNA has been observed (Figure 2). However, also after the addition of RNA in excess with respect to the protein (~ 200 μ M, protein/RNA ratio of 1:2) the signals of the new species do not increase in intensity, and the signals of the free protein are still present. Competition in the binding of the same RNA site between the two domains may be present, and multiple species formed in solution. Furthermore, the interaction landscape may be complicated by the possibility of an opening of the hairpin structure. Therefore, NMR data can give information only about the binding regions, but not about the strength of the interaction. In this respect, the residues experiencing the largest changes are located in the same region of those interacting with the other oligos (Ser15, Ile24, Gly26, Met52, Gly64, Val94, Ser117, Asp124, Val135, Met139, Leu140, Met141, Thr146, His149, Gly153, Val163, Ala184, Val189, see Figure 15).

Kinetic experiments with SPR and with switchSENSE® were performed on the tandem domain protein. MSI-1 RRM₁₋₂ interaction with both oligo-4 and oligo-5. SPR data of MSI-1 RRM₁₋₂ (Figure 14, panel A and B section iii) did not follow a 1:1 binding pattern and a clear biphasic dissociation was seen for both hairpin structures. Interestingly, this initial rapid dissociation was only observed after incubation with relatively high Musashi-1 concentrations. In the IM analysis (Figure 14, panel A and B section iii), interactions with both hairpin structures show two distinct peaks with a similar recognition but a difference in stability. Therefore, data were analysed with a 1:2 model assuming two independent interactions. Analysis of oligo-4 revealed two interactions with similar association rates, reflecting similar recognition, and a 20-fold difference in the dissociation rates while analysis of oligo-5 revealed two interactions with a little more than 7-fold difference in the dissociation rate constant. The association rate constants of Musashi-1 are somewhat slower than the association rates for the individual RRM-1

and RRM-2 for both hairpin structures but also the dissociation rates for the more rapid dissociation fraction is slower compared to the individual RRMs.

switchSENSE® data showed a kinetic profile with two different binding interactions, presented in the panel A and C of Suppl. Figure 8. When interacting with oligo-4, the interaction 1 has ≈ 3.5 -fold and ≈ 12 -fold faster on-rate and off-rate, respectively, compared to interaction 2 (see Suppl. Figure 8, panel B). The K_D of the interactions is $1.44 \pm 0.3 \text{ nmol dm}^{-3}$ and $425 \pm 20 \text{ pmol dm}^{-3}$, respectively. Instead, when interacting with oligo-5 the interaction 1 has ≈ 4.7 -fold and ≈ 39 -fold faster on-rate and off-rate, respectively, compared to interaction 2 (see Suppl. Figure 8, panel D). The K_D values for the interactions with oligo-5 are $53.4 \pm 18.1 \text{ nmol dm}^{-3}$ and $6.35 \pm 0.77 \text{ pmol dm}^{-3}$, respectively. In both cases, the interaction 1 has the strongest contribution to the binding interaction (represented as a bigger data point on the rate plot map).

Discussion

As previously described for the mouse Musashi-1, the minimal RNA sequences, recognized by RRM-1 and RRM-2 domains, are (G/AU₁₋₃AGU) and UAG, respectively. However, the presence of the UAG sequence in the (G/AU₁₋₃AGU) motif suggested an affinity of RRM-2 also for this site. Our experimental data supports this view also for the human MSI-1 with the two isolated domains showing a similar affinity in the nanomolar range for the (G/AU₁₋₃AGU) motif. In this regard, the NMR spectra recorded on the two isolated RRMs show an interaction in the slow exchange regime in the NMR time-scale, and very similar effects upon the addition of the oligo-2. The characterization of the interaction of the tandem domain with oligo-2 proves the presence of competition of the two domains for this site revealed also by a biphasic dissociation pattern in the SPR and switchSENSE® sensorgrams. The competition of the two RRMs for this site and presence in solution of different species is also supported by the analysis of the NMR data recorded on the tandem domain in the presence of the oligo-2. The intensity of the signals belonging to the protein binding sites drops down upon the addition of the RNA without the appearance of the cross-peaks of the complexes. The analysis carried out on the RNA strand containing the two recognition sites (oligo-3) has provided an additional piece of the puzzle. The two RRMs bind both sites with high affinity as shown by the evolution of the NMR spectra of the isolated RRMs upon the addition of the RNA strand. More important, two RRM-1 are able to bind the two sites of oligo-3 at the same time by forming a 2:1 RRM-1/RNA complex as proved by the SEC-MALS analysis. Interestingly, the data collected on RRM-1 in the presence of oligo-3 provide a possible way to decipher the NMR data collected on the tandem domain upon the addition of this RNA strand. In fact, at the MSI-1/oligo-3 molar ratio, new cross-peaks are visible in the spectra but they appear to be broad if not split. These findings seem to be in agreement with the presence of 1:1 MSI-1/Oligo-3 complexes with the two RRMs interacting with the two sites of the same RNA strand.

The experimental data collected on the isolated RRM-1 and RRM-2 in the presence of the oligo-4 provide interesting hints on the possible effects of the RNA structure on MSI-1 recognition process and binding. The NMR data shows that the two domains bind the hairpin with high affinity and as previously observed for the linear strands. More important, the SEC-MALS analysis shows that two isolated RRM-1 can bind oligo-4 at the same time provide a 2:1 complex. These results provide us with two interesting pieces of information: i) the presence of the binding sites in the loop of numb mRNA does not prevent the binding; ii) the position of the two binding sites within this mRNA strand allows to accommodate two RRM domains at the same time. Further interesting information were obtained from the analysis of the experimental data obtained by monitoring the interaction of the tandem domain with oligo-4. The general decrease in signal intensity observed for the cross-peaks of the free protein (stronger for the residues of the protein binding sites) during the NMR titration without the appearance of new signals indicates competition of the two domains for the two binding sites present in the hairpin. However, in contrast to what was previously observed with oligo-2, the general decrease in signal intensity indicates the presence of species involving two or more MSI-1 proteins. More important, these findings seem to rule out the capability of the two RRMs of a single MSI-1 to bind the two sites of a hairpin at the same time as observed with the oligo-3. Interesting it is also the behaviour of MSI-1 in the presence of the oligo-5. The high affinity of the isolated domains for this hairpin is proved by the NMR data. Also, the tandem domain binds the oligo-5 with high affinity but the evolution of the spectra upon the addition of the hairpin suggests a further different behaviour. The persistence of the signals of the free protein together with the signals of the new species, even in the presence of an excess of the hairpin, arises several questions on how MSI-1 can bind this oligo. A possible explanation is the binding at only one site of the hairpin (the G/AU₁₋₃AGU site on the loop, presumably) with this binding preventing the second domain present in the MSI-1 to interact with any other oligo. This behaviour could be related to the involvement of the UAG in the double strand structure which could make this site unavailable for the interaction with the two domains, and to the structure to the MSI-1-RNA complex which could make the second domain of the protein unavailable to bind a second oligo.

Collectively, these data show that MSI-1 can bind the consensus binding motif G/AU₁₋₃AGU and also the UAG sequence when placed in linear RNA sequences or in folded RNA structures. However, the involvement of a recognition motif in double RNA strands probably hamper the interaction. Finally, the protein binds all the four investigated oligos with high affinity, however only the oligo-3 seems to be able to interact with the RRM-1 and RRM-2 of a single MSI-1.

Acknowledgements

This work has been supported by Regione Toscana (CERM-TT, BioEnable and PANCREAS-AD bando salute 2018), the Italian Ministero dell'Istruzione, dell'Università e della Ricerca through PRIN 2017A2KEPL, the "Progetto Dipartimenti di Eccellenza 2018-2022" to the Department of

Chemistry “Ugo Schiff” of the University of Florence, the Recombinant Proteins JOYNLAB laboratory, and the project FISR2021_SYLCOV. The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy centre. We acknowledge This work has been supported by “RNAct” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITNAnalytical 2018 (contract n° 813239).

Figure Captions

Figure 1. 2D ^1H - ^{15}N HSQC and TROSY spectra of MSI-1 RRM-1, RRM-2 and RRM₁₋₂, spectra superimposition, respectively.

Figure 2. Portions of the 2D ^1H - ^{15}N HSQC and TROSY spectra, recorded on the isolated RRM-1 and RRM-2 domains and on the tandem domain, respectively. In black are the spectra of the free proteins, in blue the spectra of the proteins in the presence of sub-stoichiometric concentrations of oligo-2, oligo-3, oligo-4 or oligo-5 (protein/RNA ratio of about 1:0.5), and in red the spectra of the proteins in the presence of oligo-2, oligo-3, oligo-4 or oligo-5 in the protein/RNA ratio of 1:1.

Figure 3. Interaction of RRM-1 and RRM-2 with oligo-2. A) Plot of the decreases in intensity of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-2 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue. B) Plot of the Chemical Shift Perturbation (CSP) of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-2 at the protein: RNA molar ratio of 1:1 have been highlighted in red (residues with ambiguity in the assignment have been pointed at with a blue star). C) Mapping of the residues experiencing the largest effect in each RRM on the structure (PDB:2RS2 for RRM-1 and PDB: 5X3Z for RRM-2). Highlighted in blue the residues experiencing the largest Intensity decrease and in pink the one experiencing the largest CSP.

Figure 4. SEC–MALS Protein Conjugate analysis of the isolated RRM-1 and RRM-2 with oligo-2 at protein/RNA molar ratio of 1:0.50. Binding observed in both RRMs is in 1:1 stoichiometry.

Figure 5. A) Sensorgram and interaction map of MSI-1 RRM-1 binding to RNA oligo-2, concentrations used were 3.9, 7.8, 15.65, and 31.25 nmol dm⁻³. C) Sensorgram and interaction map of MSI-1 RRM-2 binding to RNA oligo-2, concentrations used were 7.8, 15.65, 31.25 and 62.5 nmol dm⁻³.

Figure 6. Interaction of MSI-1 RRM₁₋₂ with oligo-2. A) Plot of the decreases in intensity of the signals of MSI-1 RRM₁₋₂ protein. The residues experiencing the largest changes during the interaction with oligo-2 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue. B) Plot of the Chemical Shift Perturbation (CSP) of the signals of the MSI-1 RRM₁₋₂ protein. The residues experiencing the largest changes during the interaction with oligo-2 at the protein: RNA molar ratio of 1:0.25 have been highlighted in red. C) Mapping of the residues experiencing the largest effect on the RRM₁₋₂ structure (Alpha fold model) at molar ratio 1:0.25. Highlighted in blue the residues experiencing the largest

Intensity decrease, in pink the one experiencing the largest CSP, and in green the ones experiencing the largest effect in both Intensity decrease and CSP.

Figure 7. Sensorgram and interaction map of MSI-1 (1-200) binding to RNA oligo-2, concentrations used were 7.8, 15.65, 31.25 and 62.5 nmol dm⁻³.

Figure 8. Interaction of RRM-1 and RRM-2 with oligo-3. A) Plot of the decreases in intensity of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-3 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue. B) Plot of the Chemical Shift Perturbation (CSP) of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-2 at the protein: RNA molar ratio of 1:1 have been highlighted in red (residues with ambiguity in the assignment have been pointed at with a blue star). C) Mapping of the residues experiencing the largest effect in each RRM on the structure (PDB:2RS2 for RRM-1 and PDB: 5X3Z for RRM-2). Highlighted in blue the residues experiencing the largest Intensity decrease and in pink the one experiencing the largest CSP.

Figure 9. Interaction of MSI-1 RRM₁₋₂ with oligo-3. A) Plot of the decreases in intensity of the signals of MSI-1 RRM₁₋₂ protein. The residues experiencing the largest changes during the interaction with oligo-3 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue. B) Mapping of the residues experiencing a larger intensity decrease effect on the RRM₁₋₂ structure (Alpha fold model) at molar ratio 1:0.25. Highlighted in blue the residues experiencing a larger Intensity decrease.

Figure 10. Kinetic study of MSI-1 RRM₁₋₂ protein towards oligo-3 containing one binding motif for RRM-1 and one for RRM-2, on the 5' and 3', accordingly. A) Kinetic traces of MSI-1 RRM₁₋₂. The tandem protein displays a bi-phasic kinetic profile, with a step-like binding (interaction 1) and a slower binding (interaction 2) likely derived from RRM1 and RRM2 binding, respectively. B) Graphic displaying the on-rates, off-rates, and affinity values. The wild-type protein exhibit two distinct interactions. The biggest data point indicates the dominant interaction for each protein.

Figure 11. Interaction of RRM-1 and RRM-2 with oligo-4. A) Plot of the decreases in intensity of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-4 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue (RRM-1: Ile24, Gly26, Ser60, Arg61, Gly64, Phe65, Asp91, Lys93 and Ala95; and RRM-2: Trh106, Arg107, Ile111, Val113, Gly115, Leu116, Leu140, Thr146, Glu180, Lys182, Ala184, Gln185, Glu188 and Val189). B) Plot of the Chemical Shift Perturbation (CSP) of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-4 at

the protein: RNA molar ratio of 1:1 have been highlighted in red (RRM-1: Cys20, Lys21, Met22, Phe23, Trp29, Thr31, Glu48, Leu50, Leu56, Gly62, Phe65, Val66, Lys93, Val94 and Phe96; and RRM-2: Lys110, Ile111, Val113, Gly114, Leu140, Asp143, Thr146, Glu158, Asp161, Glu164, Val166 and Ile174; residues with ambiguity in the assignment have been pointed at with a blue star). Unfortunately, many signals could not be reassigned because of their too large CSP (RRM-1: Ile24, Gly26, Gln30, Ser60, Arg61, Gly64, Thr89 and Ala95; and RRM-2: Arg107, Thr108, Gly115, Leu116, Val118, Met139, Met141, Arg150, Gly153, Val155, His172, Lys182, Ala184, Glu188 and Val189). C) Mapping of the residues experiencing a larger effect in each RRM on the structure (PDB:2RS2 for RRM-1 and PDB: 5X3Z for RRM-2). Highlighted in blue the residues experiencing a larger Intensity decrease and in pink the one experiencing a larger CSP.

Figure 12. Interaction of RRM-1 and RRM-2 with oligo-5. A) Plot of the decreases in intensity of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-5 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue (RRM-1: Ala5, Met22, Ile24, Gly26, Ser28, Trp29, Gln30, Arg61, Gly64, Phe65, Lys93, Val94, Ala95 and Phe96; and RRM-2: Arg107, Ile111, Val113, Gly114, Gly115, Leu116, His149, Gly153, Glu180, Ala184, Glu188 and Val189). B) Plot of the Chemical Shift Perturbation (CSP) of the signals of the isolated RRM-1 and RRM-2 domains. The residues experiencing the largest changes during the interaction with oligo-5 at the protein: RNA molar ratio of 1:1 have been highlighted in red (RRM-1: Ala11, Cys20, Lys21, Met22, Phe23, Glu48, Asp54, Leu56, Gly62, Phe65, Val66, Lys76 and Lys93; and RRM-2: Lys110, Ile111, Val113, Gly114, Asp143, Thr145, Thr146, Asn147, Asp161, Ile174 and Met190; residues with ambiguity in the assignment have been pointed at with a blue star). Unfortunately, many signals could not be reassigned because of their too large CSP (RRM-1: Ile24, Gly26, Trp29, Gln30, Ser60, Arg61, Gly64, Val94, Ala95 and Phe96; and RRM-2: Arg107, Thr108, Gly115, Leu116, Met141, Gly153, Val155, His172, Glu180, Lys182, Ala184, Glu188 and Val189). C) Mapping of the residues experiencing the largest effect in each RRM on the structure (PDB:2RS2 for RRM-1 and PDB: 5X3Z for RRM-2). Highlighted in blue the residues experiencing the largest Intensity decrease and in pink the one experiencing a larger CSP.

Figure 13. SEC-MALS chromatograms of the interaction of RRM-1 with oligo-4 and oligo-5 at protein/RNA molar ratios of 1:0, 1:0.12, 1:0.25, 1:0.5, 1:1. Both show heterogeneous solutions with different stoichiometric species in solution. On the left, chromatograms of the interaction of RRM-1 with oligo-4. On the right, chromatograms of the interaction of RRM-1 with oligo-5.

Figure 14. Sensorgrams and InteractionMap corresponding to i) RRM-1, ii) RRM-2 and iii) RRM1-RRM2 binding to RNA A) oligo-4 and B) oligo-5.

Figure 15. Interaction of RRM1-2 with oligo-4 and oligo-5. A) Plot of the decreases in intensity of the signals of the MSI-1 RRM_{1,2} protein. The residues experiencing the largest changes during the interaction with oligo-4 and oligo-5 at a protein: RNA molar ratio of 1:0.25 have been highlighted in blue. B) Mapping of the residues experiencing the largest effect in the tandem domain protein on the structure (AF-O43347-F1) upon the binding. Highlighted in blue the residues experiencing the largest Intensity decrease.

Table 1. Kinetic and affinity values of RRM 1 and RRM 2 with oligo-4 and oligo-5

Table 2. Kinetic and affinity values of Msi-1 (1-200) with oligo-4 and oligo-5

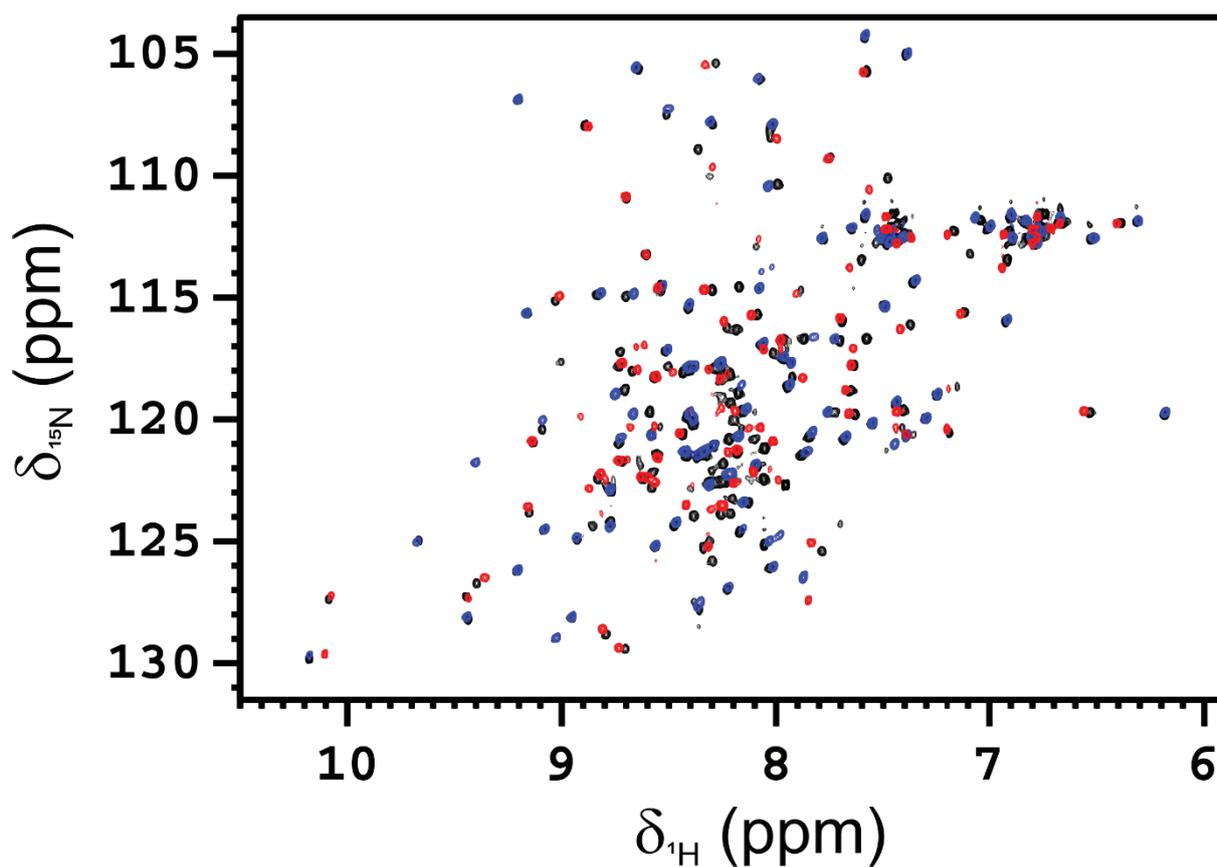


Figure 1.

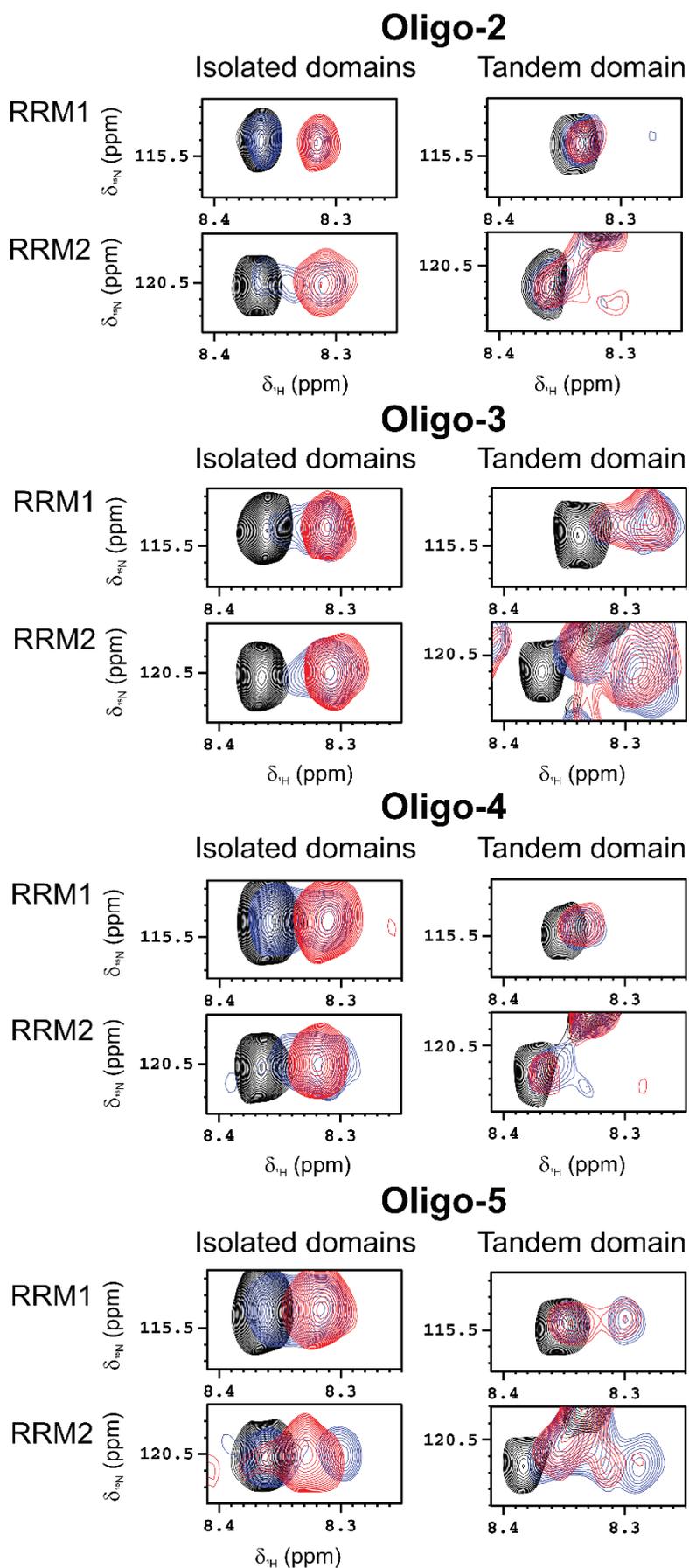


Figure 2.

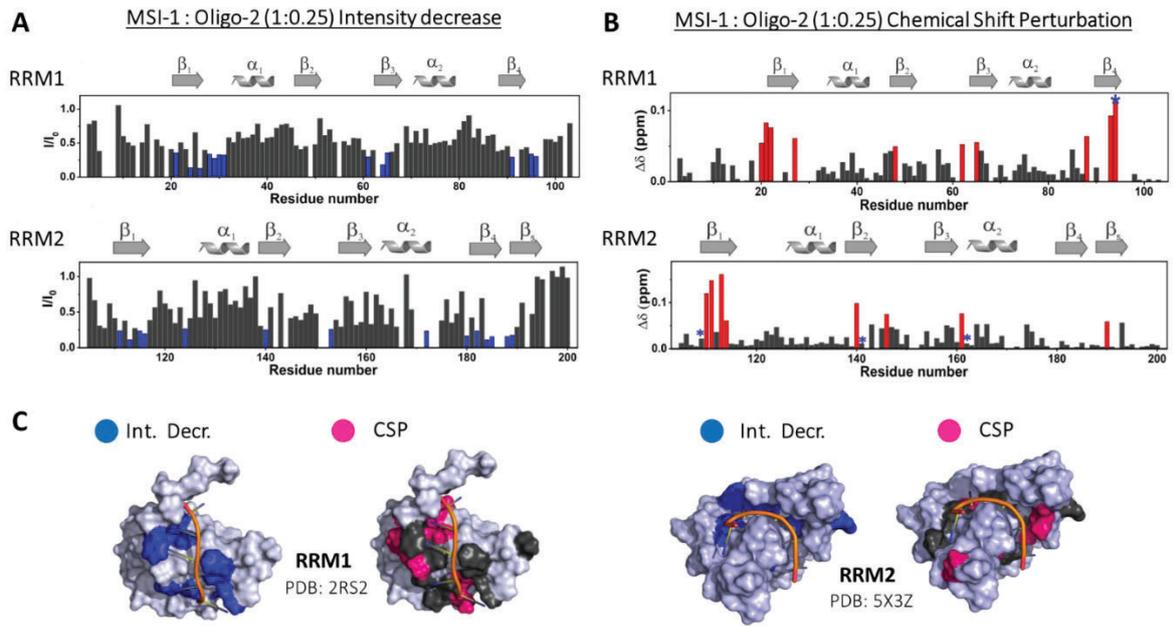


Figure 3.

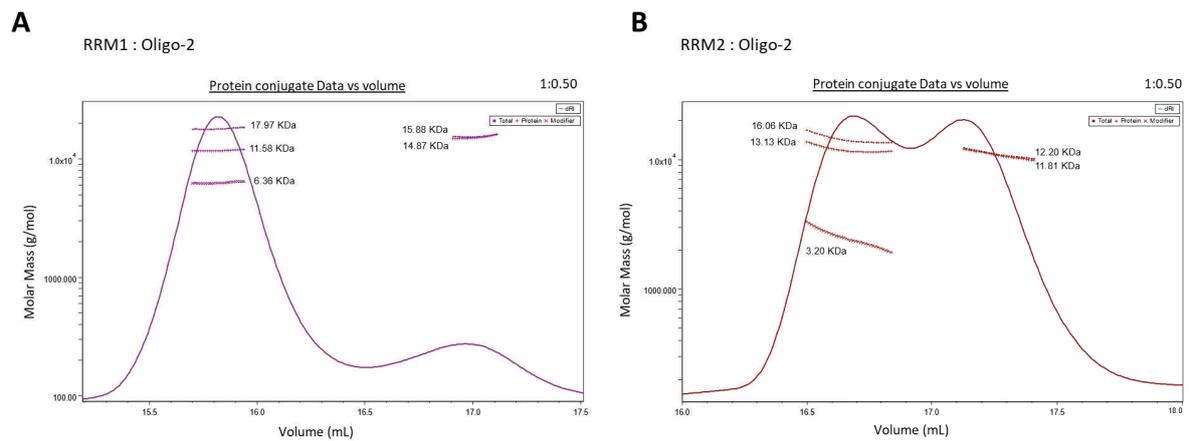


Figure 4.

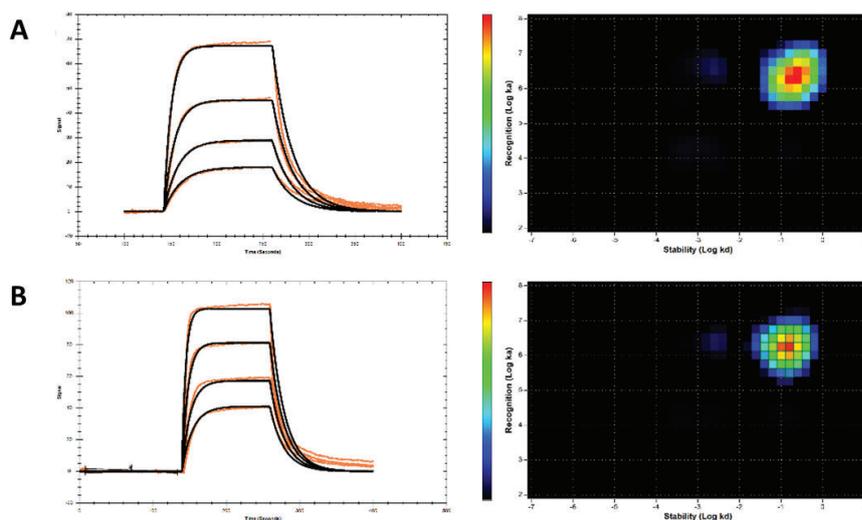


Figure 5.

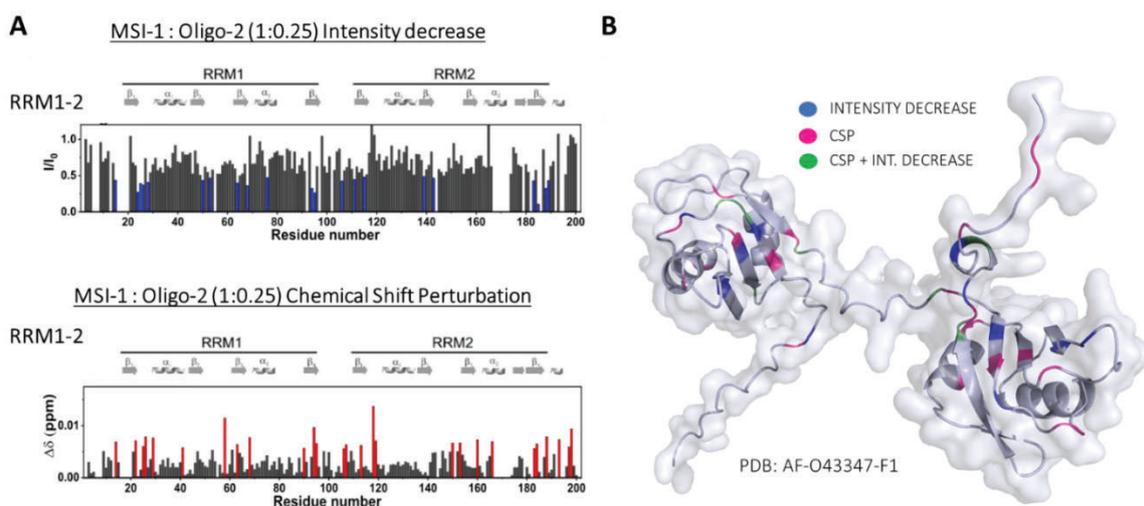


Figure 6.

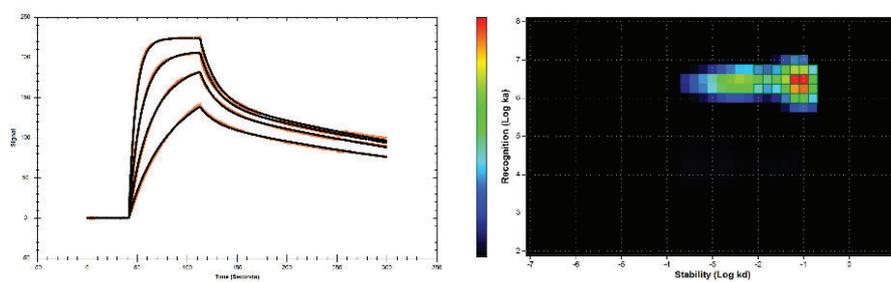


Figure 7.

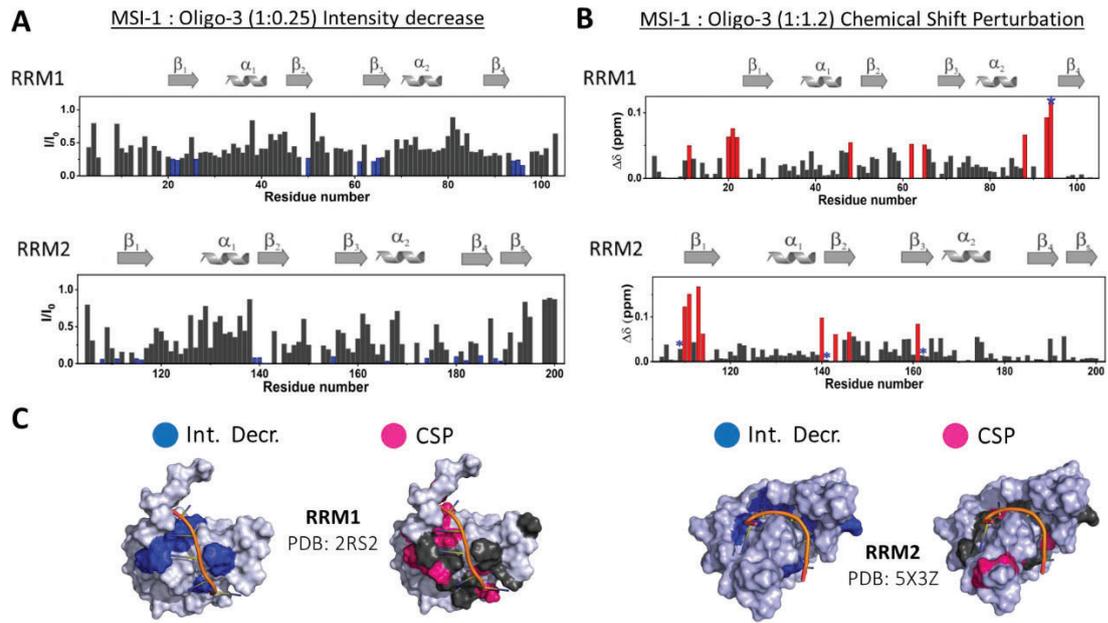


Figure 8.

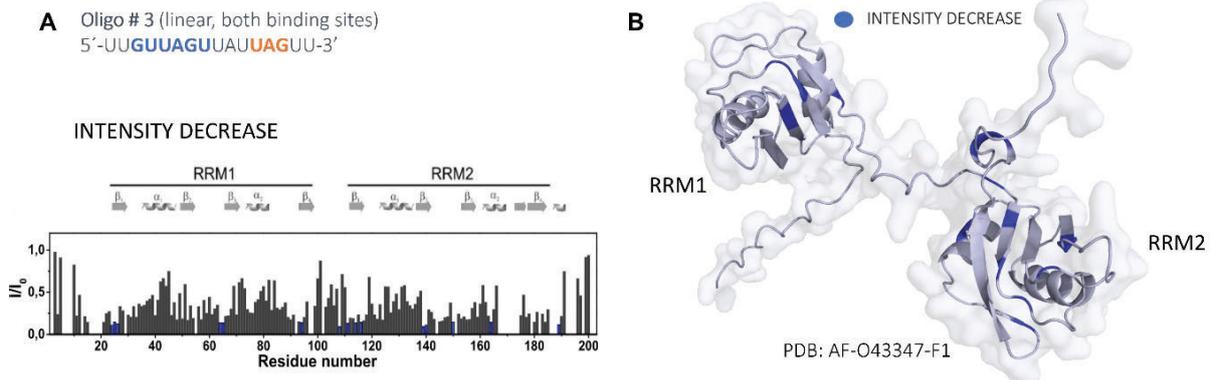


Figure 9.

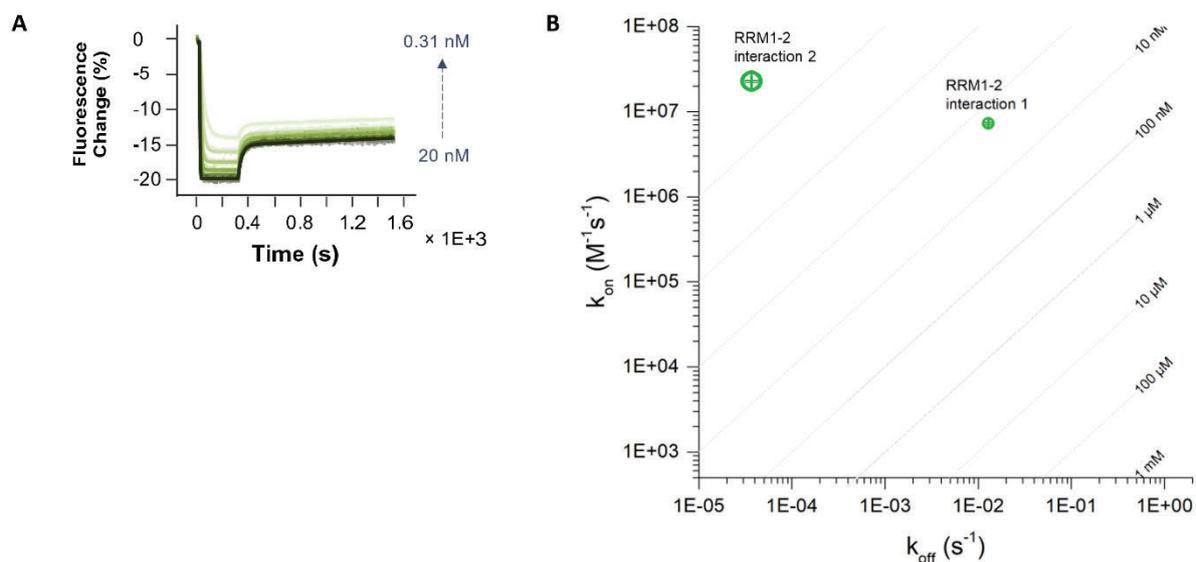


Figure 10.

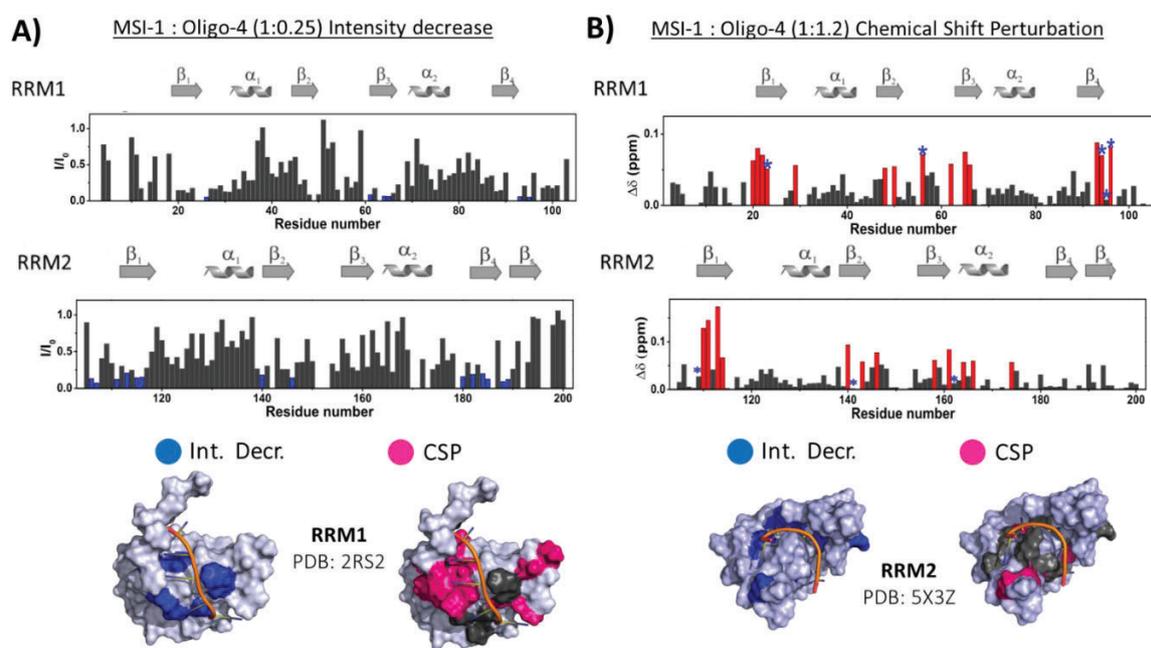


Figure 11.

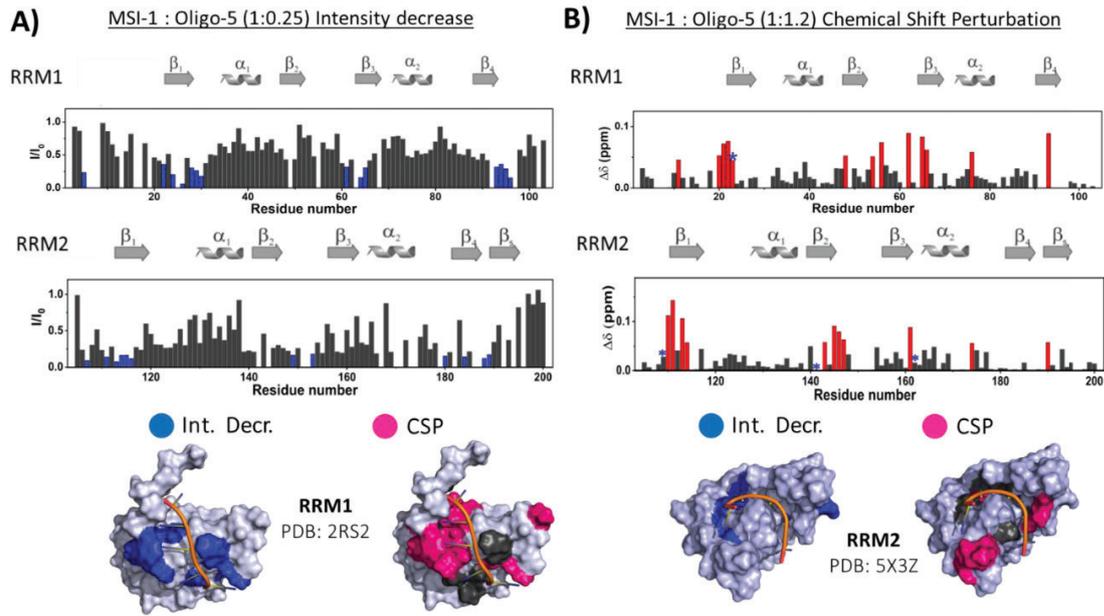


Figure 12.

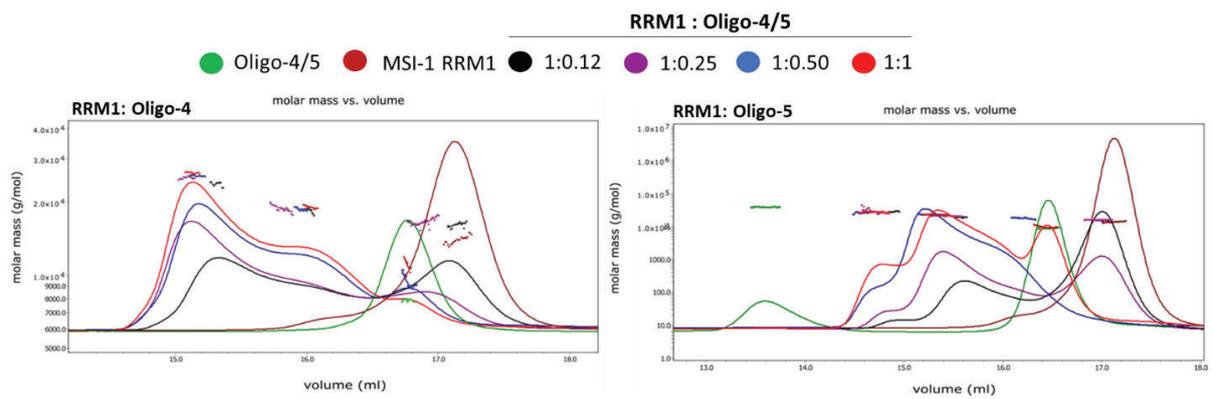


Figure 13.

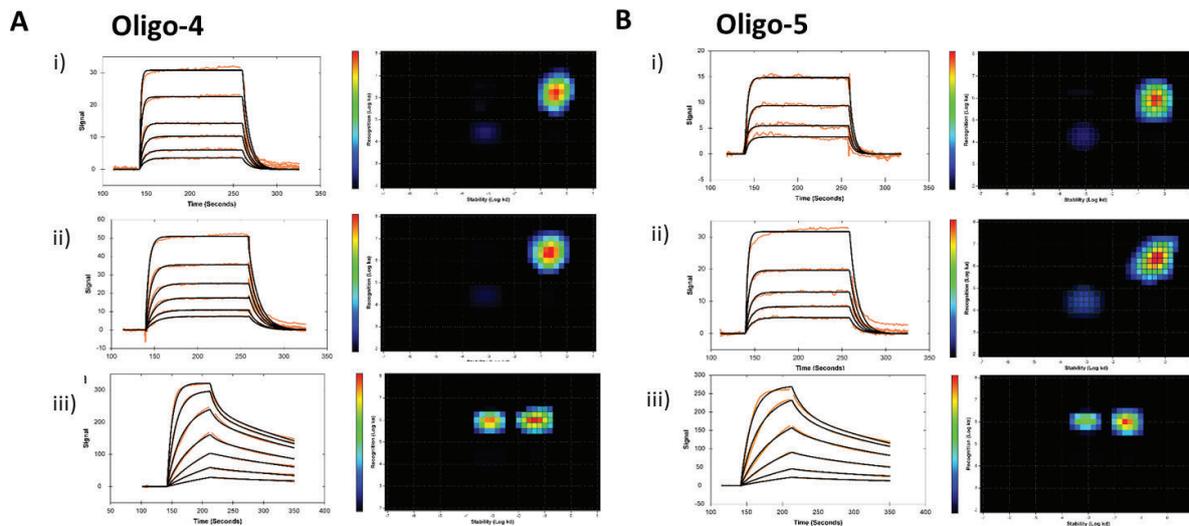


Figure 14.

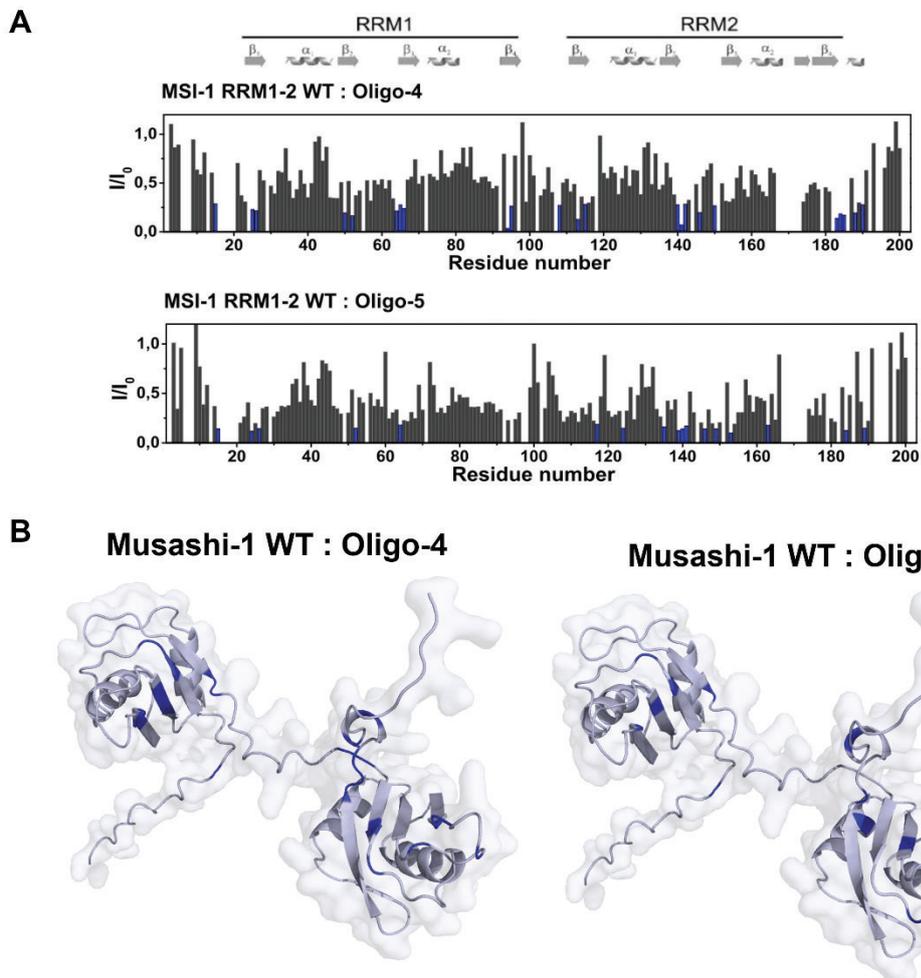


Figure 15.

Interactions	$k_{a1}(M^{-1}s^{-1})$ (CV)	$K_{d1}(s^{-1})$ (CV)	$K_{D1}(nM)$ (CV)
RRM 1– oligo-2	2.32 (31.14%)	4.7E-2 (10.31%)	22.3 (33.42%)
RRM 1– oligo-4	3.26E+06 (29.28%)	1.37E-01 (13.36%)	48 (46.04%)
RRM 1– oligo-5	1.80E+06 (26.01%)	2.39E-01 (31.61%)	145 (40.48%)
RRM 2– oligo-2	2.86 (24.37%)	5.26E-2 (24.37%)	19.6 (40.98%)
RRM 2– oligo-4	2.71E+06	8.42E-02	31.1
RRM 2– oligo-5	4.52E+06	1.13E-01	25

Table 1.

Interactions	$k_{a1}(M^{-1}s^{-1})$ (CV)	$K_{d1}(s^{-1})$ (CV)	$K_{D1}(nM)$ (CV)	$k_{a2}(M^{-1}s^{-1})$ (CV)	$K_{d1}(s^{-1})$ (CV)	$K_{D1}(nM)$ (CV)
Msi-1(1-200) – oligo-4	9.67E+05 (4.95%)	1.98E-03 (72.17%)	19.8 (72.48%)	4.68E+05 (34.26%)	4.25E-02 (35.74%)	111 (57.28%)
Msi-1(1-200) – oligo-5	1.07E+06 (5.91%)	3.90E-03 (33.87%)	3.64 (34.00%)	3.73E+05 (58.86%)	2.85E-02 (66.74%)	63.5 (40.48%)

Table 2.

References

1. Lorkovic, Z. From Structure to Function of RNA Binding Domains. *RNA Binding Proteins* 149–170 (2020) doi:10.1201/9781498713368-15.
2. Gebauer, F., Schwarzl, T., Valcárcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat Rev Genet* **22**, 185–198 (2021).
3. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* **19**, 327–341 (2018).
4. Cléry, A., Blatter, M. & Allain, F. H. T. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290–298 (2008).
5. Maris, C., Dominguez, C. & Allain, F. H. T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal* vol. 272 2118–2131 Preprint at <https://doi.org/10.1111/j.1742-4658.2005.04653.x> (2005).
6. Birney, E., Kumar, S. & Krainer, A. R. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* **21**, 5803–5816 (1993).
7. Oliveira, C., Faoro, H., Alves, L. R. & Goldenberg, S. RNA-binding proteins and their role in the regulation of gene expression in trypanosoma cruzi and saccharomyces cerevisiae. *Genet Mol Biol* **40**, 22–30 (2017).
8. Daubner, G. M., Cléry, A. & Allain, F. H. T. RRM-RNA recognition: NMR or crystallography...and new findings. *Curr Opin Struct Biol* **23**, 100–108 (2013).
9. Joel Roca-Martínez^{1, 2}, Hrishikesh Dhondge³, Michael Sattler^{4, 5}, Wim Vranken^{1, 2}. Deciphering the RRM-RNA recognition code: A computational analysis (Manuscript submitted). *PLoS Comput Biol*.
10. Assoni, G. *et al.* HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models. *Adv Drug Deliv Rev* **181**, 114088 (2022).
11. Lorkovic, Z. From Structure to Function of RNA Binding Domains. *RNA Binding Proteins* 149–170 (2020) doi:10.1201/9781498713368-15.
12. Corley, M. *et al.* mechanisms. **78**, 9–29 (2021).
13. Bono F, Ebert J, Unterholzner L, Guttler T, I. & E, E. & C. Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO report* **5**, 304–310 (2004).
14. Agrawal, S. *et al.* RNA recognition motifs of disease-linked RNA-binding proteins contribute to amyloid formation. *Sci Rep* **9**, 1–12 (2019).
15. Vanderweyde, T., Youmans, K., Liu-Yesucevitz, L. & Wolozin, B. Role of stress granules and RNA-binding proteins in neurodegeneration: A mini-review. *Gerontology* **59**, 524–533 (2013).
16. Lukong, K. E., Chang, K. wei, Khandjian, E. W. & Richard, S. RNA-binding proteins in human genetic disease. *Trends in Genetics* **24**, 416–425 (2008).

17. Hermann, T. Strategies for the design targeting RNA and RNA - Protein complexes. *Angewandte Chemie - International Edition* **39**, 1890–1904 (2000).
18. Mohibi, S., Chen, X. & Zhang, J. Cancer the‘RBP’eutics–RNA-binding proteins as therapeutic targets for cancer. *Pharmacol Ther* **203**, (2019).
19. Nahalka, J. The role of the protein–RNA recognition code in neurodegeneration. *Cellular and Molecular Life Sciences* **76**, 2043–2058 (2019).
20. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Hum Genet* **135**, 851–867 (2016).
21. Horisawa, K., Imai, T., Okano, H. & Yanagawa, H. The Musashi family RNA-binding proteins in stem cells. *Biomol Concepts* **1**, 59–66 (2010).
22. Tan, A. Y. & Manley, J. L. The TET family of proteins: Functions and roles in disease. *J Mol Cell Biol* **1**, 82–92 (2009).
23. Clingman, C. C. *et al.* Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *Elife* **2014**, (2014).
24. Lal, P. *et al.* Regulation of HuR structure and function by dihydrotanshinone-I. *Nucleic Acids Res* **45**, 9514–9527 (2017).
25. Ghidini, A., Cléry, A., Halloy, F., Allain, F. H. T. & Hall, J. RNA-PROTACs: Degraders of RNA-Binding Proteins. *Angewandte Chemie - International Edition* **60**, 3163–3169 (2021).
26. Ke, Y. *et al.* PARP1 promotes gene expression at the post-transcriptional level by modulating the RNA-binding protein HuR. *Nat Commun* **8**, (2017).
27. Rigo, F., Hua, Y., Krainer, A. R. & Frank Bennett, C. Antisense-based therapy for the treatment of spinal muscular atrophy. *Journal of Cell Biology* **199**, 21–25 (2012).
28. Schultz, C. W. *et al.* protein HuR. **11**, 1–31 (2021).
29. Mascini, M., Palchetti, I. & Tombelli, S. Nucleic acid and peptide aptamers: Fundamentals and bioanalytical aspects. *Angewandte Chemie - International Edition* **51**, 1316–1332 (2012).
30. Lucchesi, C. A., Zhang, J., Ma, B., Chen, M. & Chen, X. Disruption of the RBM38-eIF4E complex with a synthetic peptide PEP8 increases p53 expression. *Cancer Res* **79**, 807–818 (2019).
31. Abdelmohsen, K. *et al.* Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol* **14**, 361–369 (2017).
32. Ohyama, T. *et al.* Structure of Musashi1 in a complex with target RNA: The role of aromatic stacking interactions. *Nucleic Acids Res* **40**, 3218–3231 (2012).
33. Miyanoiri, Y. *et al.* Origin of Higher Affinity to RNA of the N-terminal RNA-binding Domain than That of the C-terminal One of a Mouse Neural Protein, Musashi1, as Revealed by Comparison of Their Structures, Modes of Interaction, Surface Electrostatic Potentials, and Backbone. *Journal of Biological Chemistry* **278**, 41309–41315 (2003).
34. Iwaoka, R. *et al.* Structural Insight into the Recognition of r(UAG) by Musashi-1 RBD2, and Construction of a Model of Musashi-1 RBD1-2 Bound to the Minimum Target RNA. *Molecules* **22**, 1–16 (2017).

35. Nagata, T. *et al.* Structure, backbone dynamics and interactions with RNA of the C-terminal RNA-binding domain of a mouse neural RNA-binding protein, Musashi1. *J Mol Biol* **287**, 315–330 (1999).
36. Sakakibara, S. I. *et al.* Mouse-Musashi-1, a neural RNA-Binding protein highly enriched in the mammalian CNS stem cell. *Dev Biol* **176**, 230–242 (1996).
37. Kurihara, Y. *et al.* Structural properties and RNA-binding activities of two RNA recognition motifs of a mouse neural RNA-binding protein, mouse-Musashi-1. *Gene* **186**, 21–27 (1997).
38. Kawahara, H. *et al.* Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. *Journal of Cell Biology* **181**, 639–653 (2008).
39. Ruth Zearfoss, N. *et al.* A conserved three-nucleotide core motif defines musashi RNA binding specificity. *Journal of Biological Chemistry* **289**, 35530–35541 (2014).
40. Abreu, R. D. S. *et al.* Genomic Analyses of Musashi1 Downstream Targets Show a Strong Association with Cancer-related Processes * □. **284**, 12125–12135 (2009).
41. Bley, N. *et al.* Musashi-1—a stemness RBP for cancer therapy? *Biology (Basel)* **10**, 1–16 (2021).
42. Okano, H. *et al.* Function of RNA-binding protein Musashi-1 in stem cells. *Exp Cell Res* **306**, 349–356 (2005).
43. Forouzanfar, M. *et al.* Intracellular functions of RNA-binding protein, Musashi1, in stem and cancer cells. *Stem Cell Res Ther* **11**, 1–10 (2020).
44. Horisawa, K., Imai, T., Okano, H. & Yanagawa, H. 3'-Untranslated region of doublecortin mRNA is a binding target of the Musashi1 RNA-binding protein. *FEBS Lett* **583**, 2429–2434 (2009).
45. Kudinov, A. E., Karanicolas, J., Golemis, E. A. & Bumber, Y. Musashi RNA-binding proteins as cancer drivers and novel therapeutic targets. *Clinical Cancer Research* **23**, 2143–2153 (2017).
46. Chen, H. Y. *et al.* Musashi-1 promotes stress-induced tumor progression through recruitment of AGO2. *Theranostics* **10**, 201–217 (2020).
47. Ito, T. *et al.* HHS Public Access. **466**, 765–768 (2011).
48. Jones, R. & Zweier. 基因的改变 NIH Public Access. *Bone* **23**, 1–7 (2014).
49. Samson, M. L. Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins. *BMC Genomics* **9**, 1–11 (2008).
50. Good, P. J. A conserved family of elav-like genes in vertebrates. *Proc Natl Acad Sci U S A* **92**, 4557–4561 (1995).
51. Fialcowitz-White, E. J. *et al.* Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. *Journal of Biological Chemistry* **282**, 20948–20959 (2007).
52. Kim, H. S. *et al.* Different modes of interaction by TIAR and HuR with target RNA and DNA. *Nucleic Acids Res* **39**, 1117–1130 (2011).
53. Wang, H. *et al.* The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr D Biol Crystallogr* **69**, 373–380 (2013).

54. Díaz-Quintana, A., García-Mauriño, S. M. & Díaz-Moreno, I. Dimerization model of the C-terminal RNA Recognition Motif of HuR. *FEBS Lett* **589**, 1059–1066 (2015).
55. Lixa, C. *et al.* Oligomeric transition and dynamics of RNA binding by the HuR RRM1 domain in solution. *J Biomol NMR* **72**, 179–192 (2018).
56. Pabis, M. *et al.* HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs. *Nucleic Acids Res* **47**, 1011–1029 (2019).
57. Scheiba, R. M. *et al.* The C-terminal RNA binding motif of HuR is a multi-functional domain leading to HuR oligomerization and binding to U-rich RNA targets. *RNA Biol* **11**, 1250–1261 (2014).
58. Ripin, N. *et al.* Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM. *Proc Natl Acad Sci U S A* **116**, 2935–2944 (2019).
59. Srikantan, S. & Gorospe, M. HuR function in disease. *Frontiers in Bioscience* **17**, 189–205 (2012).
60. Williams, M. A. *Protein-Ligand Interactions*.
61. Encarnação, J. C., Schulte, T., Achour, A., Björkelund, H. & Andersson, K. Detecting ligand interactions in real time on living bacterial cells. *Appl Microbiol Biotechnol* **102**, 4193–4201 (2018).
62. Cordes, R. M., Sims, W. B. & Glatz, C. E. Precipitation of Nucleic Acids with Poly (ethyleneimine). *Biotechnol Prog* **6**, 283–285 (1990).
63. Koutcher, J. A. & Burt, C. T. Principles of nuclear magnetic resonance. *Journal of Nuclear Medicine* **25**, 101–111 (1984).
64. Szakács, Z. & Sánta, Z. *NMR Methodological Overview. Anthropic Awareness: The Human Aspects of Scientific Thinking in NMR Spectroscopy and Mass Spectrometry* (2015). doi:10.1016/B978-0-12-419963-7.00007-9.
65. Fattori, J., Rodrigues, F. H. S., Pontes, J. G. M., Paula Espíndola, A. & Tasic, L. *Monitoring Intermolecular and Intramolecular Interactions by NMR Spectroscopy. Applications of NMR Spectroscopy* vol. 3 (2015).
66. Mahler, H. C., Friess, W., Grauschopf, U. & Kiese, S. Protein aggregation: Pathways, induction factors and analysis. *J Pharm Sci* **98**, 2909–2934 (2009).
67. Katsamba, P. S., Park, S. & Laird-Offringa, I. A. Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods* **26**, 95–104 (2002).
68. Nguyen, H. H., Park, J., Kang, S. & Kim, M. Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Switzerland)* **15**, 10481–10510 (2015).
69. Cléry, A. & Allain, F. H.-T. FROM STRUCTURE TO FUNCTION OF RNA BINDING DOMAINS. (2013).
70. Corley, M., Burns, M. C. & Yeo, G. W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol Cell* **78**, 9–29 (2020).
71. Gebauer, F., Schwarzl, T., Valcárcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat Rev Genet* **22**, 185–198 (2021).

-
72. Imai, T. *et al.* The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. *Mol Cell Biol* **21**, 3888–3900 (2001).
 73. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A. & Torchia, D. A. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. *Journal of Magnetic Resonance (1969)* **97**, 359–375 (1992).
 74. Ingram, D. J. E. Nuclear magnetic resonance (part ii). *Contemp Phys* **7**, 103–121 (1965).
 75. Ulrich Rant, Kenji Arinaga, Shozo Fujita, Naoki Yokoyama, Gerhard Abstreiter, and M. T. Structural Properties of Oligonucleotide Monolayers on Gold Surfaces Probed by Fluorescence Investigations. *Langmuir* **20**, 10086–10092 (2004).
 76. Kanga, H. S. *et al.* An autoinhibitory intramolecular interaction proofreads RNA recognition by the essential splicing factor U2AF2. *Proc Natl Acad Sci U S A* **117**, 7140–7149 (2020).

4.2.2 Supplementary Information

The interaction between Musashi-1 and RNA: a multiplayer game.

switchSENSE®

The adapter chip features a microfluidic channel with two gold electrodes, or measurement spots, which portray immobilized fluorophore-modified DNA nanolevers that serve as linkers between the sensor surface and the desired ligand (see Figure S1A). In the proximity sensing mode, the electrodes are under a constant negative potential that maintains the nanolevers in an upward orientation. The fluorescence of the probe is monitored over time, being affected when the analyte interacts with the ligand. The fluorescent signal changes represent the kinetic profile of the interaction, as displayed in Figure S1B.

Oligo-2, oligo-3, oligo-4, and oligo-5 were located in spot 1 and the RNA negative control without a binding motif oligo-1 on spot 2. All nucleic acid sequences were synthesized along with a standard DNA sequence called ligand strand on the 5' (5'-ATCAGTACTTGTCAACACGAGCAGCCCGTATATTCTCCTACAGCACTA-3') which is part of the DNA nanolever system serving as a linker between ligands and sensor surface. The upper half of the ligand strand was hybridized in solution with a DNA adapter strand which is decorated with the so called "Gb" green fluorescence probe on its 3' end. The bottom half of the adapter strand is complementary to a pre-immobilized DNA anchor strand on the chip surface. The functionalization of the surface with the ligand strand/adapter strand hybrid towards the anchor strand followed automated standard procedures within the heliX® platform.

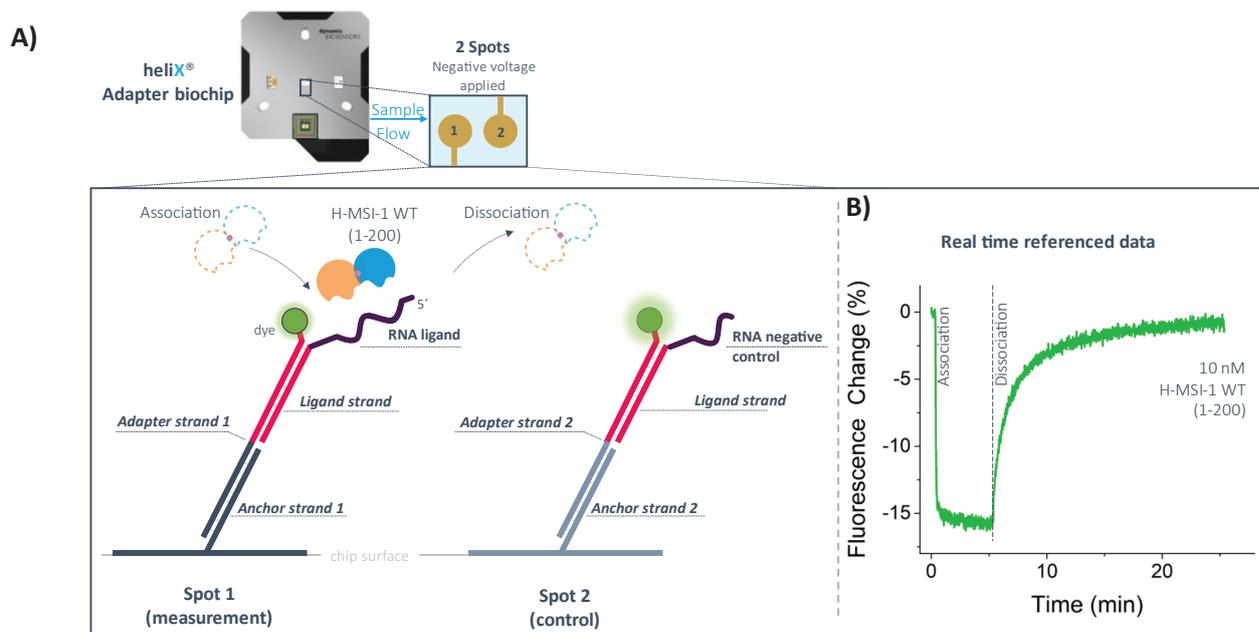


Figure S1. Experimental set up for measuring binding kinetics with the switchSENSE® technology. In A) the heliX® Adapter biochip portrays two gold electrodes, or measurement spots, where DNA nanolevers (adapter strands) with a fluorescent dye serve as linkers between the ligands of interest and the biosensor surface, while reporting on binding interactions. The DNA nanolevers consist of hybridized anchor, adapter and ligand strands. In this case, the ligands are RNA overhangs with the target sequence on spot 1 and a negative control sequence on spot 2. Kinetic experiments consist of an association phase where protein is automatically injected in the flow channel of the biochip, while in the dissociation phase the sensor is washed with buffer, all while recording the changes in fluorescence. In B) an example of the kinetic profile obtained from H-MSI-1 RRM₁₋₂ protein at a concentration of 10 nmol dm⁻³ binding to oligo-2 is shown. Data was blank and real time referenced with the spot 2. The association and dissociation phases are indicated. Image adapted with permission of Dynamic Biosensors GmbH.

	MSI-1 RRM1	MSI-1 RRM2	Oligo- 2	Oligo- 3	Oligo- 4	Oligo-5
dn/dc (ml/g)	0.185	0.185	0.180	0.180	0.180	0.180
UV extinction coefficient (ml/(mg·cm))	1.045	0.586	6.200	6.200	6.200	6.200
Theoretical MW (KDa)	13.491	12.133	4.975	5.023	7.287	8.245

Table S1. Size Exclusion Chromatography with Dynamic Light Scattering parameters of MSI-1 RRM-1 and RRM-2 and for oligo-2 and oligo-3.

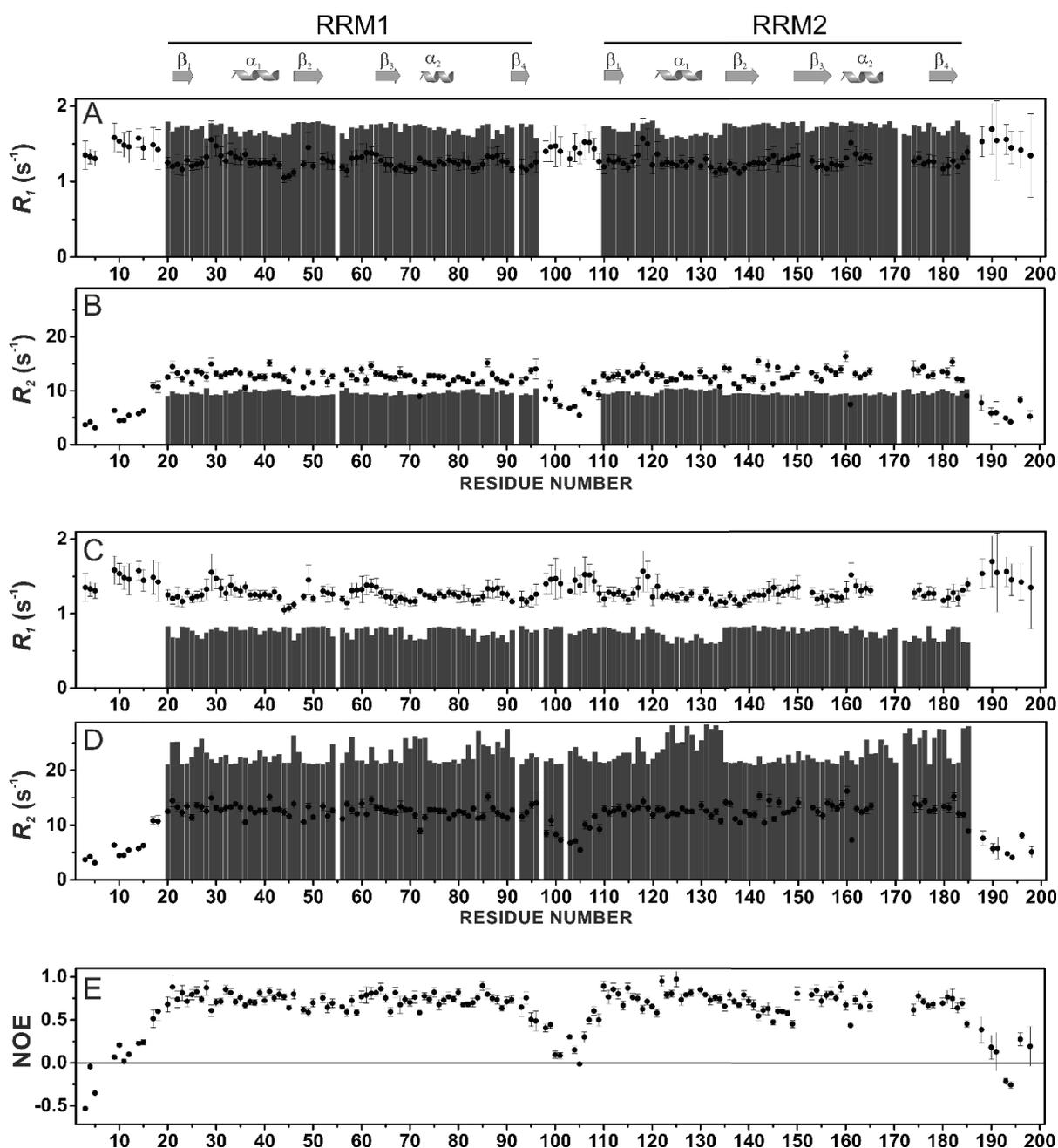


Figure S2. Comparison of experimental backbone ^{15}N R_1 values for MSI-1 RRM₁₋₂ construct (data collected at 298 K, black filled circles) with the calculated values (grey bars) for isolated RRM1 and RRM2 domains (A) and for the full RRM1-RRM2 construct (C). Comparison of experimental backbone ^{15}N R_2 values for MSI-1 RRM₁₋₂ construct (data collected at 298 K, black filled circles) with the calculated values (grey bars) for isolated RRM1 and RRM2 domains (B) and for the full RRM1-RRM2 construct (D). Experimental NOE values for MSI-1 RRM₁₋₂ construct (data collected at 298 K) (E).

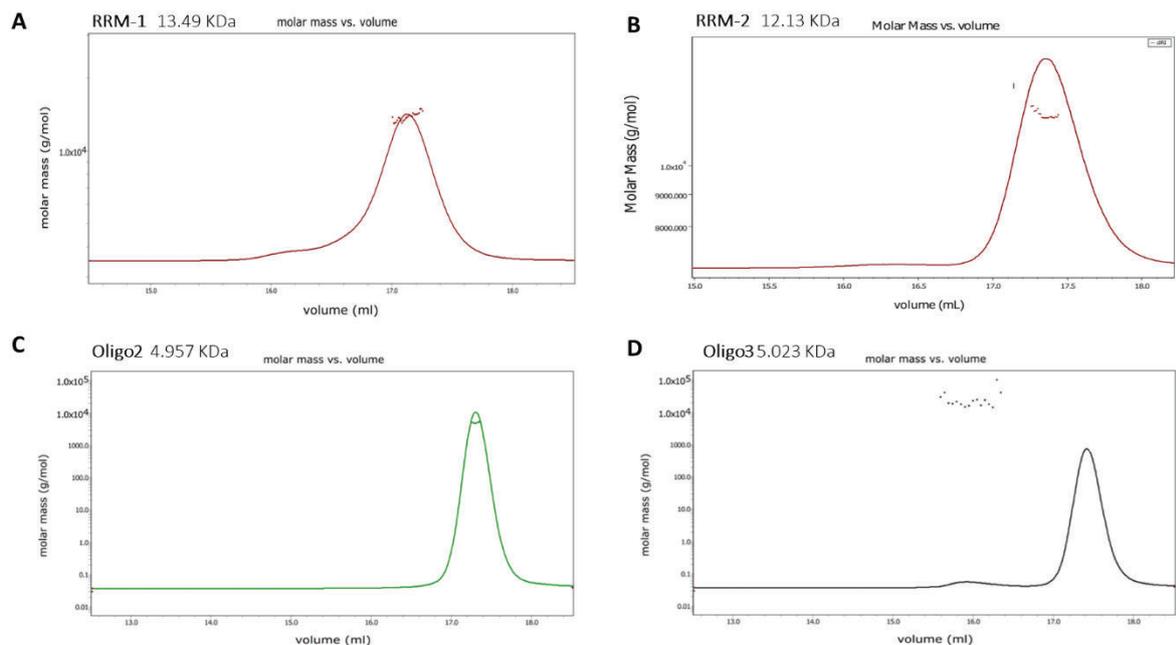


Figure S3. SEC-MALS chromatograms. A) Chromatogram of the MSI-1 RRM-1. Analysis of the peak correspond to a molecular weight (MW) value of 13.86 KDa close to the theoretical value of 13.49 KDa. The peak corresponds to monomeric form of the protein. B) Chromatogram of the MSI-1 RRM-2. Analysis of the peak correspond to a MW value of 12.16 KDa similar to its theoretical MW of 12.49 KDa, and therefore, corresponding to monomeric form of the protein. C) Chromatogram of oligo-2. Analysis shows one single peak with a MW of 5.14 KDa, close to the theoretical value of 4.95 KDa. Therefore, it shows the monomeric form of the RNA. D) Chromatogram of oligo-3. Analysis shows two peaks, one with really weak intensity of 16 KDa corresponding to some aggregate form of the RNA, and a second peak with much higher intensity with a MW of 5.12 KDa. The theoretical MW of this RNA is 5.023 KDa which indicates this second peak to be the monomeric form of Oligo-3.

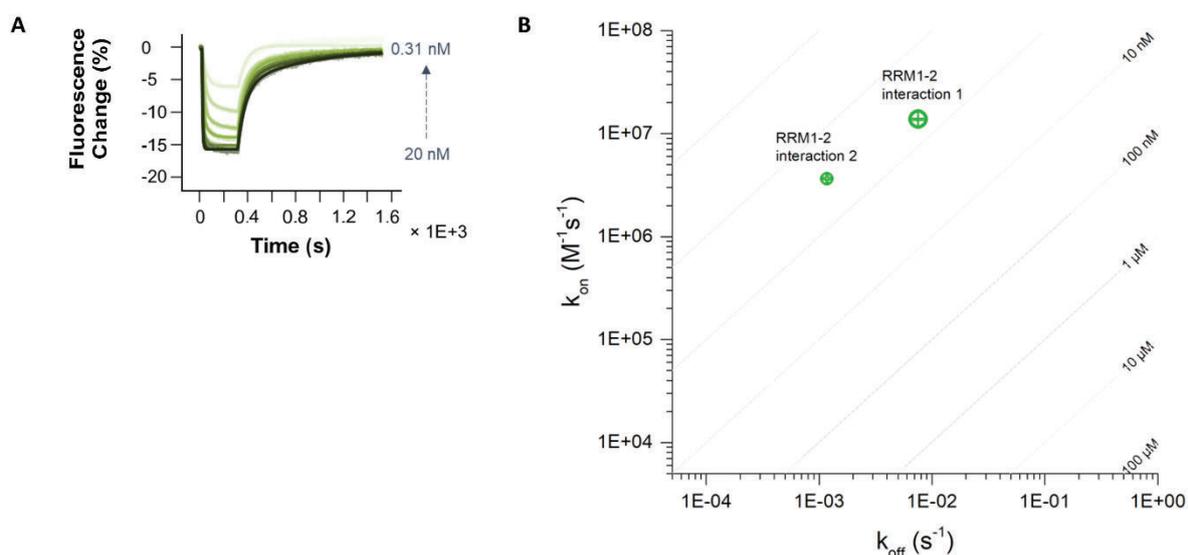


Figure S4. Kinetic study of MSI-1 RRM₁₋₂ protein towards oligo-2 containing one binding motif for RRM-1 and one for RRM-2, on the 5' and 3', accordingly. A) Kinetic traces of MSI-1 RRM₁₋₂. The tandem protein displays a bi-phasic kinetic profile, with a step-like binding (interaction 1) and a slower binding (interaction 2). B) Graphic displaying the on-rates, off-rates, and affinity values. The wild-type protein exhibit two distinct interactions. The biggest data point indicates the dominant interaction for each protein.

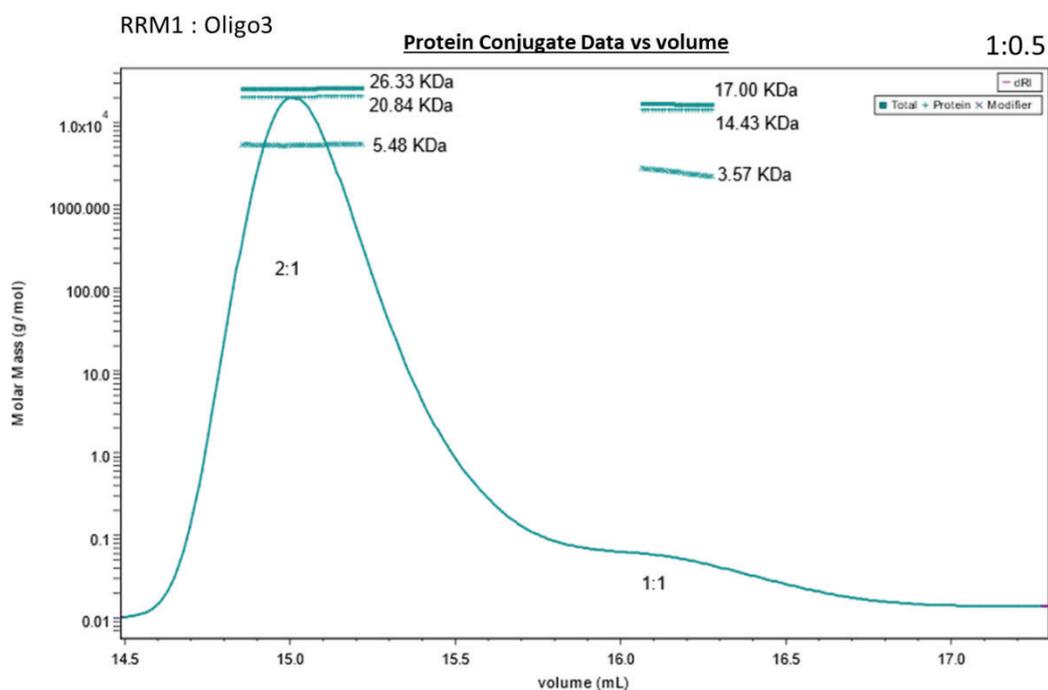


Figure S5. SEC –MALS Protein Conjugate analysis of the isolated RRM-1 with Oligo-2 at protein/RNA molar ratio of 1:0.50. Binding observed in both RRMs in 1:2 and 1:1 stoichiometry.

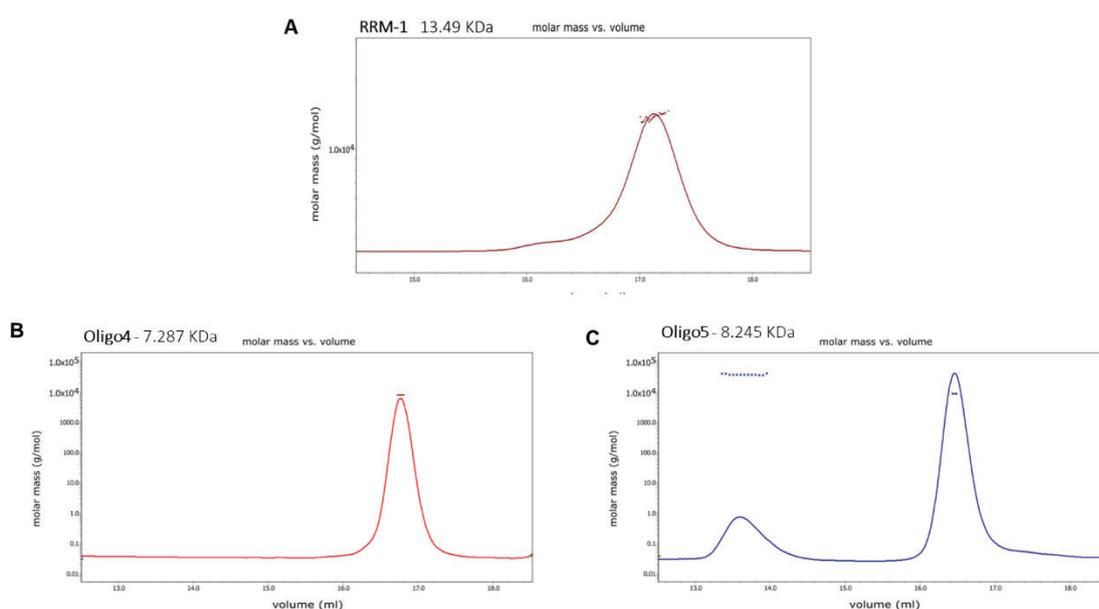


Figure S6. SEC-MALS chromatograms. A) Chromatogram of the MSI-1 RRM-1. Analysis of the peak correspond to a molecular weight (MW) value of 13.86 KDa close to the theoretical value of 13.49 KDa. The peak corresponds to monomeric form of the protein. B) Chromatogram of oligo-4. Analysis of the peak correspond to a MW value of 7.911 KDa similar to its theoretical MW of 7.287 KDa, and therefore, corresponding to monomeric form of the protein. C) Chromatogram of oligo-5. Analysis shows two clear peaks, one of 38.10 KDa corresponding to some aggregate form of the RNA, and a second peak with much higher intensity with a MW of 8.77 KDa. The theoretical MW of this RNA is 8.245 KDa which indicates this second peak to be the monomeric form of oligo-5.

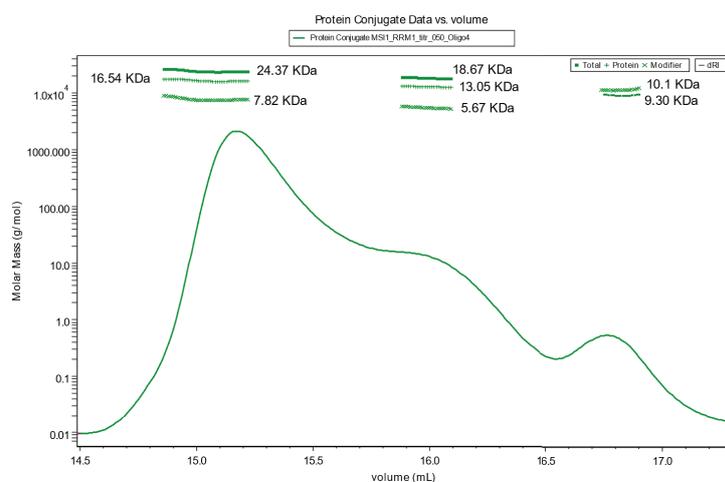


Figure S7. SEC –MALS Protein Conjugate analysis of the isolated RRM-1 with Oligo-4 at protein/RNA molar ratio of 1:0.50. Several species of different stoichiometric ratios are present in solution.

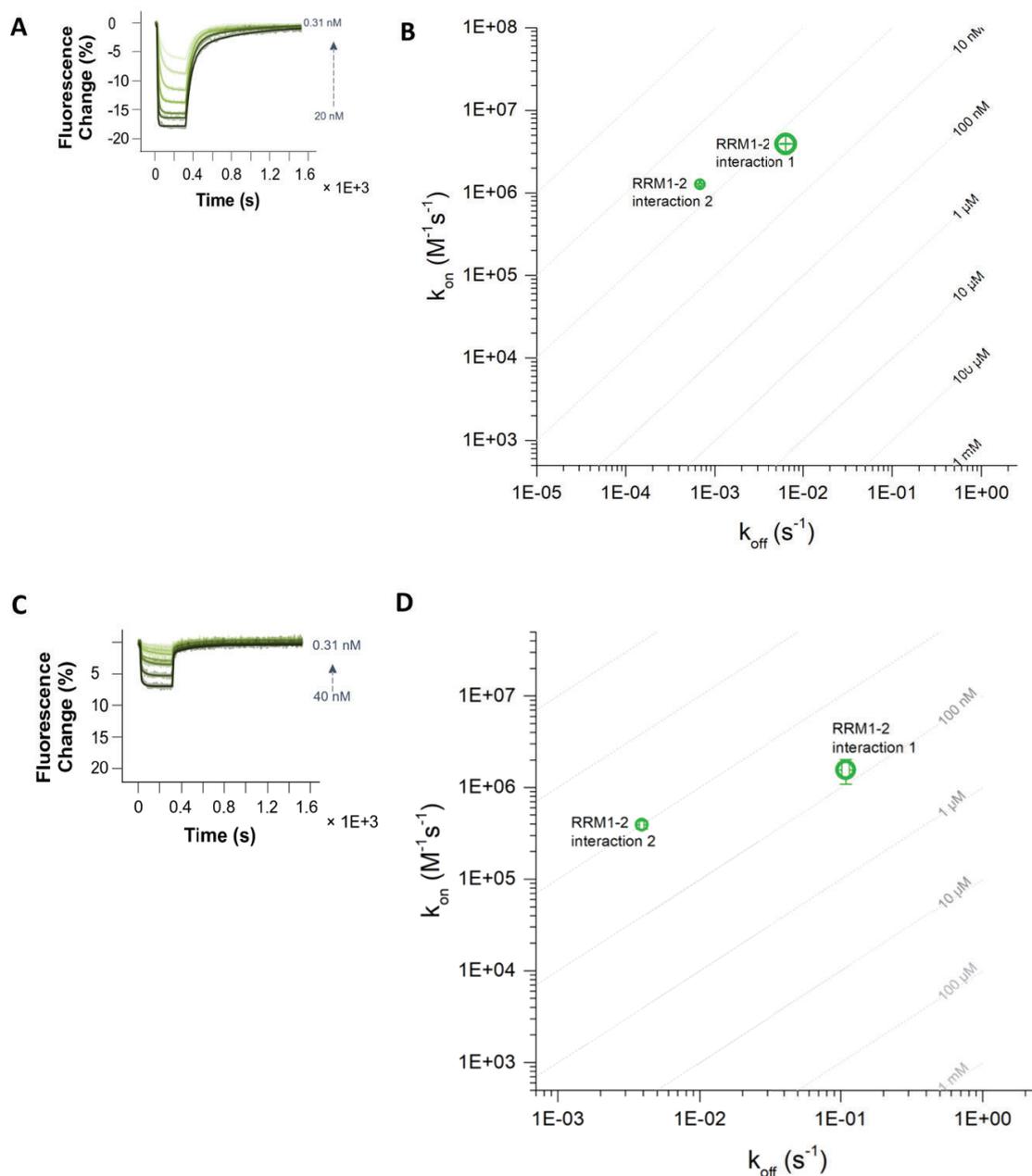


Figure S8. Contribution of each RRM in the binding of the tandem MSI1 protein towards oligo-4 and oligo-5 containing one binding motif for RRM-1 and one for RRM-2, on the 5' and 3', accordingly. In A) kinetic traces of the tandem protein with oligo-4. The tandem protein displays a kinetic profile that is a combination of both RRM domains, with a step-like binding (interaction 1) and a slower binding (interaction 2) likely derived from RRM-1 and RRM-2 binding, respectively. In B) comparison of the on-rates, off-rates and affinity values in the interaction with oligo-4. The tandem protein, exhibit two distinct interactions. In C) kinetic traces of the tandem protein with oligo-5. The tandem protein displays a kinetic profile that is a combination of both RRM domains, with a step-like binding (interaction 1) and a slower binding (interaction 2) likely derived from RRM-1 and RRM-2 binding, respectively. In D) comparison of the on-rates, off-rates and affinity values in the interaction with oligo-5. The tandem protein, exhibit two distinct interactions.

4.3 ARTICLE 4:

4.3.1 Deciphering the RNA recognition by Musashi-1 to develop new artificial biomolecules for in vitro and in vivo applications

Anna Pérez-Ràfols, A.PR.,^{1,2} Joel Roca J.R.,^{3,4} Guillermo Pérez-Ropero, G.PR.,^{5,6} Rosa Anahí Higuera A.H.,⁷ Linda Cerofolini L.C.,¹ Luca Sperotto L.S.,^{8,9} Jos Buij J.B.,⁵ Wolfgang Kaiser W.K.,⁷ Michael Sattler M.S.,^{8,9} Wim Vranken W.V.,^{3,4*} Claudio Luchinat, C.L.,^{1,2,10} Tommaso Martelli T.M.^{2*} and Marco Fragai M.F.^{1,*}

¹ Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy.

² Giotto Biotech S.R.L., Via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy.

³ Interuniversity Institute of Bioinformatics in Brussels, VUB/ULB, Brussels, 1050, Belgium.

⁴ Structural biology Brussels, Vrije Universiteit Brussel, Brussels, 1050, Belgium.

⁵ Ridgeview Instruments AB, Vänge, Uppsala, Sweden

⁶ Sweden and Biochemistry Department, University of Uppsala, Sweden

⁷ Dynamic Biosensors GmbH, 82152 Martinsried, Germany.

⁸ Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.

⁹ Bavarian NMR Center, Chemistry Department, Technical University of Munich, Garching, 85748 Munich, Germany

¹⁰ Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy

*Corresponding author

In Progress

Abstract

A new computational tool able to estimate and score the binding affinity of RNA-binding proteins for RNA has been used to design a set of mutants of the human Musashi-1 protein and new selected RNA strands to modulate binding selectivity and recognition. We have used this tool to investigate the competition between the two RRM domains of Musashi-1 (RRM-1 and RRM-2) for the same binding site within linear and an harpin RNA constructs. The *in-silico* predictions have been assessed by monitoring the interaction between *de-novo* designed proteins and RNA strands by solution NMR. The experimental results support the reliability of the predictions obtained *in silico* thus opening the way for the development of new biomolecules for *in vitro* and *in vivo* studies and applications.

In particular, my contribution to this project has been the design of RNA sequences and mutations, the obtention of experimental single point mutations and the expression and purification of all constructs of human MSI-1 protein in *E. coli* cells (unlabelled, ^1H - ^{15}N) and performing the protein-RNA interaction experiments with solution NMR. Computational scores have been calculated by our collaborators.

Introduction.

Musashi-1 (MSI-1) protein is a multi-domain RNA binding protein, highly expressed in neural precursors cells during the embryonic stage. It participates in brain development, maintenance of stem-cell state and cell differentiation.¹ MSI-1 is composed by two RNA Recognition Motifs (RRMs) located in tandem in the N-terminal region which contribute to the RNA binding. Once activated, it regulates hundreds of different 3'-UTR regions of mRNAs and therefore it is implicated in various signalling pathways, including Notch and Wnt. Due to its regulatory functions, any alteration of the level of expression of Musashi-1 often leads to a disruption of signalling pathways, leading to several diseases and cancer.⁵ This highlights the importance of MSI-1 and remarks its potential to be a marker and a promising therapeutic target in cancer disease.⁶⁻⁸ In this respect, the design of novel RRM could be relevant to investigate the regulation of gene expression, as well as in the discovery of *in vivo* RNA targets of RRM with a still unknown function or interaction.

In previous studies (*Article 3, section 4.2*), we investigated *in vitro* the interaction of the tandem domain of human Musashi-1 (MSI-1 RRM₁₋₂), and of its isolated domains (RRM-1 and RRM-2), with two single stranded linear RNAs and two hairpins, using solution NMR and kinetics experiments. We have demonstrated the capability of each RRM to recognize and bind with high affinity the (G/A)U₁₋₃AGU and UAG motifs placed in single linear RNA strands and in loops of harpins.

Very recently, a new computational tool able to estimate and score the binding affinity between any specific RRM and a given RNA sequence has been developed by J. Roca and co-workers (under review). Here, this tool (RRMscorer) has been used to address the competition between the two RRM domains of Musashi-1 (RRM-1 and RRM-2) for the UAG motif within a linear and an harpin RNA construct. The analysis allowed us to evaluate the contribution to the binding of the residues located in specific positions of the RRM-1 and RRM-2, as well as that of nucleotides within the two RNA strands. To prove the quality of the prediction and the reliability of RRMscorer, the data has been used to design mutations on the tandem domain of Musashi-1 protein and on the two RNA strands for a biophysical assessment. The RRM-2 mutant was designed to decrease its affinity for the UAG motif in order to reduce the competition with RRM-1. Then, the RRMscorer tool was used to design mutants of the two selected RNA strands where the UAG motif is replaced by a new set of three nucleotides with a high affinity for the mutated RRM-2 domain. The NMR analysis carried out on the mutated protein and RNA constructs proves the quality of the predictions provided by RRMscorer and suggest Musashi-1 as a promising candidate to develop mutants with high selectivity for specific RNA sequences for *in vitro* and *in vivo* studies and applications.

Methods

Computational design of protein and RNA mutants

For the design of mutations, scoring of RRM-RNA interaction in both Musashi and RNA targets were computed with RRMScorer method.¹¹ This method has been recently developed to estimate the binding between any specific RRM and a given RNA sequence (under revision data). Briefly, all the structures available for the RRM-RNA complexes were analysed to derive the scoring function. A set of entries, describing the canonical binding mode of RRM domains,¹² has been selected after a careful alignment of the RRM-RNA complexes. The contacts observed within this set of structures, and the positions of the residues involved in the binding on the RRM domains, have been integrated in a probabilistic framework. In this respect, meaningful propensities for residue-nucleotide contact preferences in specific positions can be extracted, and it becomes particularly suitable for the limited amount of the data and residue-level information currently available. The most relevant positions of the protein regarding RNA binding are short-listed based on RRMScorer database contacts analysis, and for each of them a preference matrix has been derived showing the binding preferences to a nucleotide for any residue in that specific position. To perform new predictions, the sequences for both the RRM and the RNA are required only.

RNA strands

Synthetic single stranded RNA (**Oligo-2:** 5'-UUGUUAGUUACCCCUU-3', **Oligo-3:** 5'-UUGUUAGUUAUUAGUU-3', **Oligo-3.2:** 5'-UUGAUAGUUAGCAGGU-3', **Oligo-4:** 5'-AAGCGUUAGUUAUUUAGUCGCUU-3', **Oligo-4.2:** 5'-AAGCGAUAGUUAUGCAGGCGCUU-3' and **Oligo-5:** 5'-CACUCUGUAGUAUGUAGGGUUUUAUUU-3'), for NMR experiments were purchased from Metabion international AG.

Expression and purification of recombinant isolated RRM-1 and RRM-2 domains of Human Musashi-1 (MSI-1)

To produce the MSI-1 human RRM-1 construct (1-103), a pET29b plasmid containing a Strep-Tag at N-terminus followed by the first 261 residues of Musashi-1 human protein was modified by QuikChange Site-Directed Mutagenesis to replace Met-104 with a stop codon. Recombinant human MSI-1 RRM1 (1-103) protein in plasmid pET29b and MSI-1 RRM-2 (104-200) protein in plasmid pET21a were overexpressed in BL21(DE3) *E. coli* cells. Expressions were performed as previously described in *Article 3, section 4.2*.

Expression and purification of recombinant wild-type Human Musashi-1 (MSI-1) RRM₁₋₂ tandem domain and of the E180N/K182M mutant

To produce the mutant, QuikChange Site-Directed Mutagenesis was performed on the gene coding for the MSI-1 (1-200) RRM₁₋₂ WT to replace the E180 by an asparagine, the K182 by a methionine, and both of them to obtain the double mutant construct MSI-1 RRM₁₋₂ E180N K182M (MSI-1 RRM₁₋₂ DM). Recombinant human MSI-1 RRM₁₋₂ (1-200) protein and MSI-1 RRM₁₋₂ DM in plasmid pET29b were overexpressed in BL21(DE3) GOLD *E. coli* cells. Expressions were performed as described in *Article 3, section 4.2* for the wild-type tandem domain protein.

Titration of Musashi with RNA strands.

The effects of the RNA strands (**Oligo-2:** 5'-UUGUUAGUUACCCCUU-3', **Oligo-3:** 5'-UUGUUAGUUAUUAGUU-3', **Oligo-3.2:** 5'-UUGAUAGUUAGCAGGU-3', **Oligo-4:** 5'-AAGCGUUAGUUAUUUAGUCGCUU-3', **Oligo-4.2:** 5'-AAGCGAUAGUUAUGCAGGCGCUU-3' and **Oligo-5:** 5'-CACUCUGUAGUAUGUAGGGUUUAUUU-3') were evaluated on the ¹⁵N-isotopically enriched double mutant (E180N K182M) of Musashi-1 (MSI-1 RRM₁₋₂ DM) at a concentration of 100 μmol dm⁻³ in the following experimental conditions: 50 mmol dm⁻³ Tris-HCl, pH 7.2, 140 mmol dm⁻³ NaCl, 1 mmol dm⁻³ EDTA, 1 mmol dm⁻³ protease Inhibitors. The effect of Oligo3.2 and Oligo4.2 was also evaluated on the wild-type protein under the same conditions. 2D ¹H ¹⁵N TROSY spectra were recorded at 298 K on a Bruker AvanceNEO NMR spectrometer, operating at 900 MHz (¹H Larmor frequency) in order to monitor the effect of increasing amounts (6, 12, 24, 50, 120 μmol dm⁻³) of the RNA added to the protein solution. Spectra were processed with the Bruker TOPSPIN software packages and analysed with CARA (Computer Aided Resonance Assignment, ETH Zurich).

Results and discussion

Mutant selection

First, the score of each RRM against all possible combinations of 3 nucleotide sequences was computed (Figure 1, panel A). A clear preference of both RRMs towards the UAG motif was observed and a better score for the affinity (0.1 difference in a logarithmic scale) was obtained for RRM-1. We next evaluated the best possible scores of each RRM towards the 5-nucleotide motif N_xUAGN_x, where N_x can be any nucleic base. The best scores are highlighted in Figure 1, panel A, whilst the positions with higher score and with higher difference between both RRMs are plotted in Supp. Figure 1. In this regard, we identified several RNA motifs that were predicted to have a reduced affinity (lower score) to bind RRM-1, while keeping a relatively high score towards RRM-2 (Figure 1, panel A). An RNA with two different short motifs to specifically bind the RRM-1 and RRM-2 is expected to show a stronger binding as it avoids the forementioned competitiveness. The CAG motif was identified as the most

promising candidates, being CCAGG, and GCAGG the top-scoring 5-mers, that were thus included in the experimentally tested RNAs (see Suppl. Figure 1).

Then, we wanted to identify protein mutations on the RRM-2, that helped to decrease the affinity of this domain for the UAG motif, whilst increasing the affinity to the CAG motif, previously identified as a good RRM-2 binder. For this purpose, we analyzed possible residues involved in the recognition of the first pyrimidine of the 3-mer motif, so we could switch the preference from the uracil of UAG to the cytosine of the CAG (Figure 1 panel C). Based on the published structure of the RRM-2 in complex with GUAGU (PDB ID: 5x3z), two of the key residues involved in the uracil recognition are E180 and K182, interacting with amine and carbonyl groups of the nucleotide base, respectively. Based on our scores (Figure 1, panel B and C), the most suitable mutations to switch the binding specificity towards a cytosine in that position are E180N and K182M. These mutations are supposed to have an impact on the charges in those positions, while keeping a similar overall bulkiness, to not disrupt the structure of the RRM domain. Moreover, these residues are already present in these positions for other RRMs and display a preference to bind a cytosine.

Experimental evaluation of Musashi mutant using NMR

We evaluated the effect of the two designed mutations for the Musashi-1 RRM₁₋₂ protein on the binding of different RNA strands, using solution NMR. First, to validate if the mutations decreased the affinity of RRM-2 for the (G/A)U₁₋₃AGU motif, we titrated the protein with a single stranded RNA only containing this consensus sequence (oligo-2). In previous studies (*Article 3, section 4.2*), in the presence of oligo-2, the residues of wild-type MSI-1 RRM₁₋₂, experiencing the largest effects, were located in both domains, since both RRMs were able to recognize this motif with similar affinities. However, only a decrease in the intensity of the protein signals was observed, without the appearance of new cross-peaks corresponding to the protein/RNA complex (Figure 2). Conversely, when titrating MSI-1 RRM₁₋₂ DM with oligo-2, the signals of the free protein in the 2D ¹H-¹⁵N TROSY spectrum decrease in intensity, while new cross-peaks, corresponding to MSI-1 RRM₁₋₂ DM in complex with oligo-2, appear and increase in intensity (Figure 2). More important, the signals of the free protein experiencing the largest decreases in intensity, after the addition of RNA at the concentration of 25 μM to the protein solution (1:0.25 protein/RNA molar ratio), correspond to residues located only (but two residues) on RRM-1 (Figure 3, position B and C) (Asp14, Ser15, Lys21, Phe23, Ile24, Gly25, Gly26, Leu27, Ser28, Thr31, Glu38, Val46, Glu48, Leu50, Met52, Asp54, Thr57, Arg61, Gly62, Gly64, Phe65, Val66, Lys76, Lys88, Thr89, Ile90, Asp91, Gln185 and Lys187). Although two residues at the C-terminal region of the RRM-2 domain still experience some effect, the double mutation on the protein seems to have largely shifted the binding preference of oligo-2 towards RRM-1. However, some heterogeneity seems to still

be present at a protein/RNA molar ratio of 1:1, as indicated by the broadening of the signals (Figure 2, and Suppl. Figure 2, position B)

Next, we studied by solution NMR the effect of the two mutations on the interaction of Musashi-1 with RNAs containing two binding sites: i) a linear single stranded RNA (oligo-3), and ii) a folded RNA bearing the two binding motifs within the loop region of a hairpin folding (oligo-4).

Upon the addition of increasing concentrations of the oligo-3 to the solution of MSI-1 RRM₁₋₂ DM, a decrease in intensity of the signals of the free protein and the appearance of new cross-peaks were observed, as previously obtained for the wild-type protein (Figure 2 and Suppl. Figure 3). The signals of the free protein experiencing the largest decreases in intensity after the addition of RNA at the concentration of 25 μ M to the protein solution (1:0.25 protein/RNA molar ratio) correspond to residues mainly located in the RRM-1 domain (Asp14, Ser15, Met22, Gly25, Gly26, Thr31, Leu50, Ser60, Arg61, Phe65, Val66, Phe68, Lys76, Leu85, Lys93, Ala95, Val155, Ala184, Gln185 and Lys187) (Figure 4, position A and B i and iii). More interestingly, in the presence of RNA in a 1:1 molar ratio with respect to the protein, the linewidth of the signals of the mutant is sharper than what observed for the wild-type protein (*Article 3, section 4.2*). These data suggest the formation of a single species in solution for the protein/RNA complex.

The interaction of MSI-1 RRM₁₋₂ DM with oligo-4 was then performed. In this case, we saw a different behavior with respect to what observed for the wild-type protein. Indeed, upon the addition of increasing concentrations of the oligo-4 to the solution of MSI-1 RRM₁₋₂ DM, the intensity of the signals of the free protein in the 2D ¹H-¹⁵N TROSY spectrum decrease in intensity, while new cross-peaks, corresponding to MSI-1 RRM₁₋₂ DM in complex with oligo-4, appear and increase in intensity (Figure 2, position and Suppl. Figure 4). As observed for oligo-3, the signals of the free protein experiencing the largest decreases in intensity after the addition of RNA at the concentration of 25 μ M to the protein solution (1:0.25 protein/RNA molar ratio) correspond to residues mainly located in the RRM-1 (Asp14, Ser15, Met22, Gly25, Gly26, Thr31, Leu50, Ser60, Arg61, Phe65, Val66, Phe68, Lys76, Leu85, Lys93, Ala95, Val155, Ala184, Gln185 and Lys187) (Figure 5, position A and B i and iii). Therefore, lower heterogeneity is observed for the double mutant protein in the interaction with oligo-4 with respect to the wild-type. However, more than one species seems to still be present in solution, as indicated by the broadening of the signals (Figure 2 and Suppl. Figure 4).

Experimental evaluation of the RNA mutants using NMR.

Next, the effect of the RNA mutants on the wild-type and DM tandem domain protein was evaluated using solution NMR. The interaction of oligo-3.2 with the wild-type protein was first investigated. During the NMR titration, upon the addition of increasing concentrations of oligo-3.2 to MSI-1 RRM₁₋₂ WT protein, a decrease in the intensity of the cross-peaks of the free protein is observed,

while new cross-peaks, corresponding to the wild-type tandem domain protein in complex with oligo-3.2, appear and increase in intensity (Figure 2 and Suppl. Figure 3 position B). Therefore, also the interaction of the wild-type tandem domain protein with oligo-3.2 is in the slow exchange regime on the NMR timescale, as observed with oligo-3 (*Article 3, section 4.2*). The signals of the free proteins experiencing the largest decreases in intensity after the addition of oligo-3.2 at the concentration of 25 μM to the protein's solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platform, as well as in loop-1 (β_1 - α_1) and loop-3 (β_2 - β_3) of both RRM domains (Figure 4, position A and B i and ii). (Asp4, Ile24, Gly25, Gly26, Gly64, Val94, Arg107, Val113, Gly114, Val118, Met139, Glu180 and Val189). However, in the presence of RNA in 1:1 molar ratio with respect to the protein, the signals of the protein in the 2D ^1H - ^{15}N TROSY spectrum appear split, suggesting the presence of multiple species in solution. This effect is more pronounced than what observed for the unmodified oligo-3 (*Article 3, section 4.2*) and the presence of a RRM domain, not involved in an interaction with RNA, cannot be excluded. As expected, the mutation on the RNA, weakens the interaction of the oligo with the wild-type MSI-1 RRM₁₋₂ protein. Interestingly, the residues experiencing the largest effect are spread on both domains and not only on RRM-1, as it happens instead in the interaction of the double mutant with oligo-3.

The effect of oligo-4.2 mutant on the wild-type MSI-1 RRM₁₋₂ protein was then investigated. Upon the addition of the RNA to the protein solution, the signals of the free protein in the 2D ^1H - ^{15}N TROSY decrease in intensity, while new cross-peaks, corresponding to MSI-1 RRM₁₋₂ WT in complex with oligo-4.2, appear and increase in intensity, unlike what was observed in the presence to oligo-4 (Figure 2, position and Suppl. Figure 4 position B). The signals of the free protein experiencing the largest decreases in intensity after the addition of oligo-4.2 at the concentration of 25 μM to the protein's solution (1:0.25 protein/RNA molar ratio) correspond to residues located in the β -platform as well as in in loop-1 (β_1 - α_1) and loop-3 (β_2 - β_3) of both RRM domains (Figure 5, position A and B i and ii) (Ser15, Met22, Ile24, Gly25, Gly26, Met52, Gly64, Ala95, Met104, Arg107, Thr108, Val113, Gly114, Gly115, Leu116, Ser117, Met139, Leu140, Phe142, Asp143, Thr146, Arg150, Gly153, Val155, Ser159, Glu164, Ala184, Gln185 and Val189). However, after the addition of RNA in a 1:1 ratio with respect to the protein (~100 μM , RNA) the signals of the new species do not increase further in intensity, and signals of the free protein seems to be still present. Therefore, the mutations on the RNA construct have weakened the interaction of the RRM domains for the second binding site and hampered the formation of possible species involving more than one MSI-1 protein. However, heterogeneity is still present in solution with different combinations of binding between the protein and the RNA: e.g. complexes with RRM-1 or RRM-2 bound to the G/AU₁₋₃AGU binding site.

The interaction of the double mutant tandem domain of Musashi with the mutant RNAs, oligo-3.2 and oligo-4.2, was evaluated with solution NMR. During both titration of MSI-1 RRM₁₋₂ DM with

oligo-3.2 and oligo-4.2, the cross-peaks of the free protein decrease in intensity while new cross-peaks, corresponding to the protein in complex with the RNA, appear and increase in intensity (Figure 2). The signals of the free proteins experiencing the largest decreases in intensity after the addition of oligo-3.2 and oligo-4.2 at the concentration of 25 μ M to the proteins solution (1:0.25 protein/RNA molar ratio) correspond to residues located on both RRM₁₋₂ DM with oligo-3.2: Asp18, Met22, Ile24, Gly26, Cys49, Val66, Val94, Ala95, Arg107, Ile111, Gly114, Gly115, Phe132, Met139, Asp143, Thr146 and Asn147, Figure 4 panel A and B iv; for MSI-1 RRM₁₋₂ DM with oligo-4.2: Ser15, Lys21, Met22, Ile24, Gly25, Gly26, Leu50, Met52, Arg61, Gly64, Phe68, Asp91, Lys93, Val94, Ala95, Lys103, Arg107, Val113, Met141, Val155, Ala184, Gln185, Lys187, Val189 and Met190, Figure 5 panel A and B iv). In the presence of oligo3.2 in 1:1 ratio with respect to the protein, a homogenous species is still visible in solution, as observed for the mutated protein with oligo-3. While, in the presence of oligo4.2 in 1:1 ratio with respect to the protein, the signals of the protein are split. In this case, the mutation on the RNA, also when combined with mutations on the protein, does not prevent the formation of different complexes were either one or the other RRM can interact with the (G/A)U₁₋₃AGU motif.

In conclusion, solution NMR has been here used to assess the quality of the predictions provided by the computational tool RRMscorer by testing the designed mutants of Musashi-1 protein and of two interacting RNA strands. In the presence of oligo-2, the RRM-2 domain within the mutated MSI-1 RRM₁₋₂ DM exhibits a weaker interaction with the (G/A)U₁₋₃AGU motif, thus, reducing the competition with RRM-1 for the recognition of this site. This finding is corroborated by the experiments carried out with the same mutant of MSI-1 in the presence of oligo-3 where the two binding sites ((G/A)U₁₋₃AGU and UAG motives) are both present in the construct. Indeed, the titration of MSI-1 RRM₁₋₂ DM with oligo-3 shows the presence of a single MSI-1-RNA complex with RRM-1 mostly involved in the interaction. These results suggest RRM-1 acting as a “guiding domain” of the RNA recognition since the mutated RRM-2 has a lower affinity for both the UAG motives, as predicted by the program. This behavior is different from what observed when the wild-type MSI-1 was titrated with oligo-3. In that case the competition between RRM-1 and RRM-2 for the two UAG sites leads to the formation two or more species. A similar behavior is also observed when the MSI-1 RRM₁₋₂ DM protein is titrated with the folded oligo-4. The decreased affinity of RRM-2 for the UAG motif results in a reduction of heterogeneity in solution, although two or more species are still observed. It should be noticed that, also in this case the NMR data show weak effects on RRM-2 domain. More important, the CAG mutation on the two RNA strands restores the affinity of the mutated RRM-2 domain for the two oligos leading to a higher selectivity in protein-RNA recognition with respect to what observed when the wild-type MSI-1 was titrated with non-mutated RNA strands. In particular a single species is obtained when a stoichiometric amount of oligo-3.2 is added to MSI-1 RRM₁₋₂ DM protein, Conversely, two predominant species seem to be present in solution when oligo-4.2 is added to MSI-1 RRM₁₋₂ DM protein. This behavior is sizably different from that of the wild-type MSI-1 in the presence of unmutated oligo-4

where a larger heterogeneity was observed. These results agree with a more selective interaction of the MSI-1 RRM₁₋₂ DM with oligo-4.2 where RRM-1 binds preferentially the UAG and RRM-2 the CAG motif. Summarizing, these results prove the reliability and the quality of the prediction provided by RRMscorer and, more important, that the MSI-1 and its RNA partners are promising candidates to develop new tools for biological studies, and possible biotherapeutics

The collection of kinetics experimental data is in progress to provide a better picture about these multi-player interactions.

Figure Captions:

Figure 1. Scoring values from RRMScorer calculations. A) Top-scoring RNAs for the isolated domains. B) Top scores and mutations proposed based on the protein. C) On the left, residues scoring values for the CAG motif in positions E180 and K182 of RRM-2. On the right, structural view of the second RNA Recognition Motif binding with GUAGU motif.

Figure 2. Portions of the 2D ^1H - ^{15}N TROSY spectra, recorded on the MSI-1 RRM₁₋₂ wild-type and double mutant (E180N, K182M). In black are the spectra of the free proteins, in blue the spectra of the proteins in the presence of sub-stoichiometric concentrations of RNA, and in red the spectra of the proteins in the presence the oligos in a protein/RNA ratio of 1:1.

Figure 3. Comparison of interaction between MSI-1 RRM₁₋₂ WT and MSI-1 RRM₁₋₂ DM proteins with oligo-2. A) Mapping of the residues experiencing a larger effect in the Wild-type tandem domain protein on the structure (AF-O43347-F1) upon the binding. Highlighted in blue the residues experiencing a larger Intensity decrease at a protein/RNA molar ratio of 1:0.25. B) Mapping of the residues experiencing a larger effect in the DM tandem domain protein on the structure (AF-O43347-F1) upon the binding. Highlighted in blue the residues experiencing a larger Intensity decrease at a protein/RNA molar ratio of 1:0.25. C) Graphic highlighting in blue on the WT and DM tandem domains the residues most affected by an intensity decrease effect during the interaction with oligo-2 at a molar ratio of protein/RNA of 1:0.25.

Figure 4. Comparison of interaction between MSI-1 RRM₁₋₂ WT and MSI-1 RRM₁₋₂ DM proteins with oligo-3 and 3.2. A) Mapping of the residues experiencing a larger effect in the Wild-type and DM tandem domain protein on the structure (AF-O43347-F1) upon the binding with oligo-3 and oligo-3.2. Highlighted in blue the residues experiencing a larger Intensity decrease at a protein/RNA molar ratio of 1:0.25. C) Graphic highlighting in blue on the WT and DM tandem domains the residues most affected by an intensity decrease effect during the interaction with oligo-3 and oligo-3.2 at a molar ratio of protein/RNA of 1:0.25.

Figure 5. Comparison of interaction between MSI-1 RRM₁₋₂ WT and MSI-1 RRM₁₋₂ DM proteins with oligo-4 and 4.2. A) Mapping of the residues experiencing a larger effect in the Wild-type and DM tandem domain protein on the structure (AF-O43347-F1) upon the binding with oligo-4 and oligo-4.2. Highlighted in blue the residues experiencing a larger Intensity decrease at a protein/RNA molar ratio of 1:0.25. C) Graphic highlighting in blue on the WT and DM tandem domains the residues most

affected by an intensity decrease effect during the interaction with oligo-4 and oligo-4.2 at a molar ratio of protein/RNA of 1:0.25.

	MSI-1 RRM1	MSI-1 RRM2	Oligo-4
dn/dc (ml/g)	0.185	0.185	0.180
UV extinction coefficient (ml/(mg·cm))	1.045	0.586	6.200
Theoretical MW	13.491	12.133	7.287

Table 1.

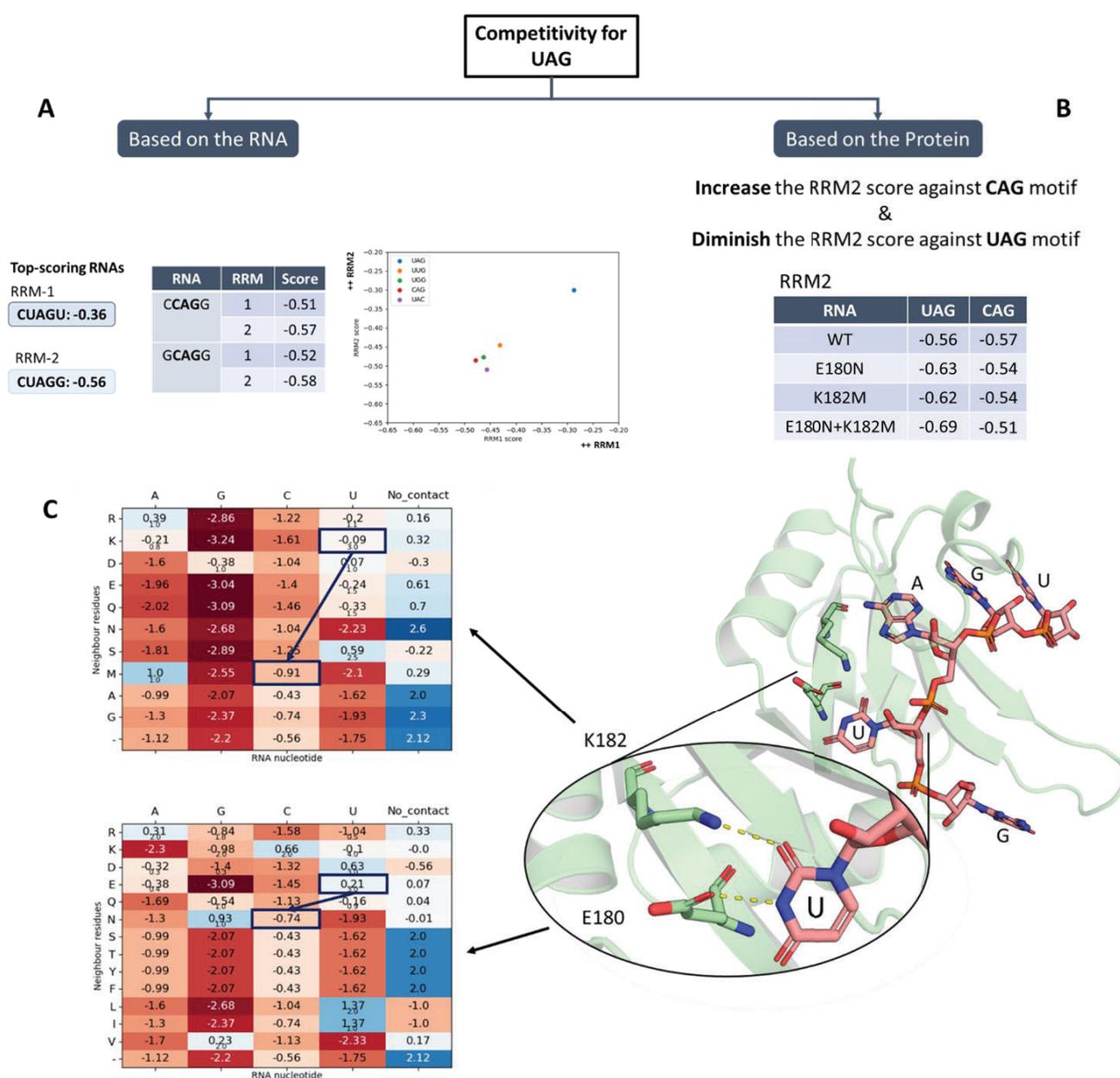


Figure 1.

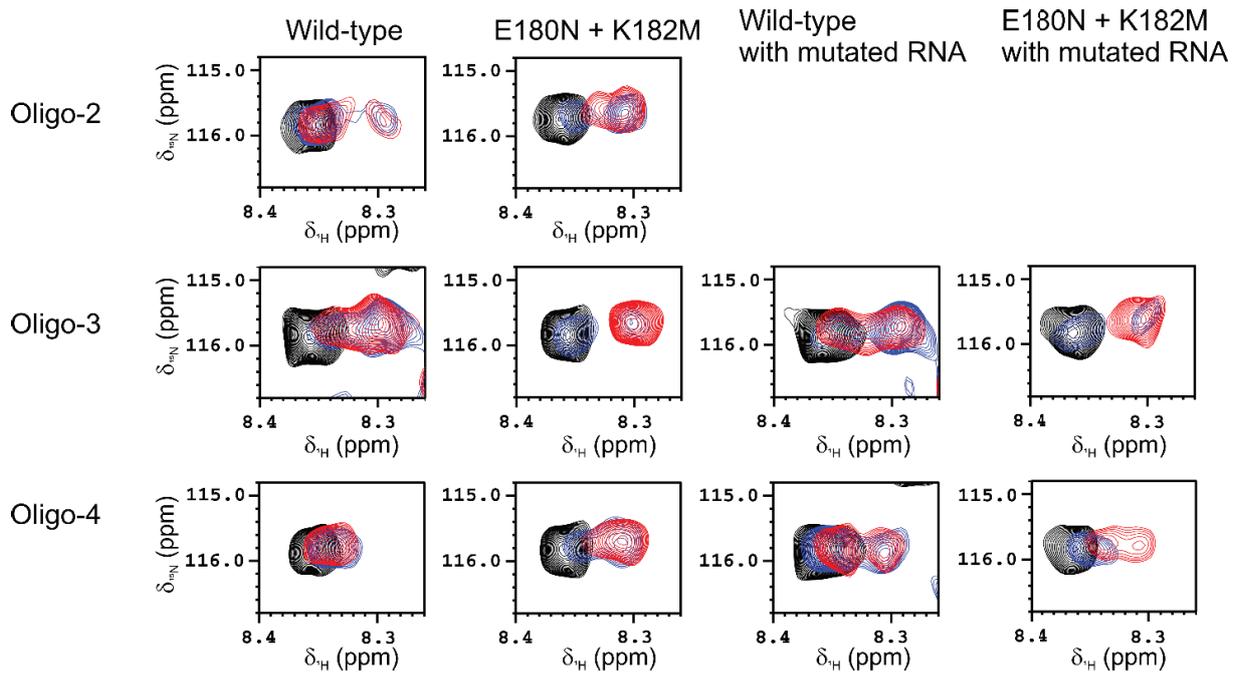


Figure 2.

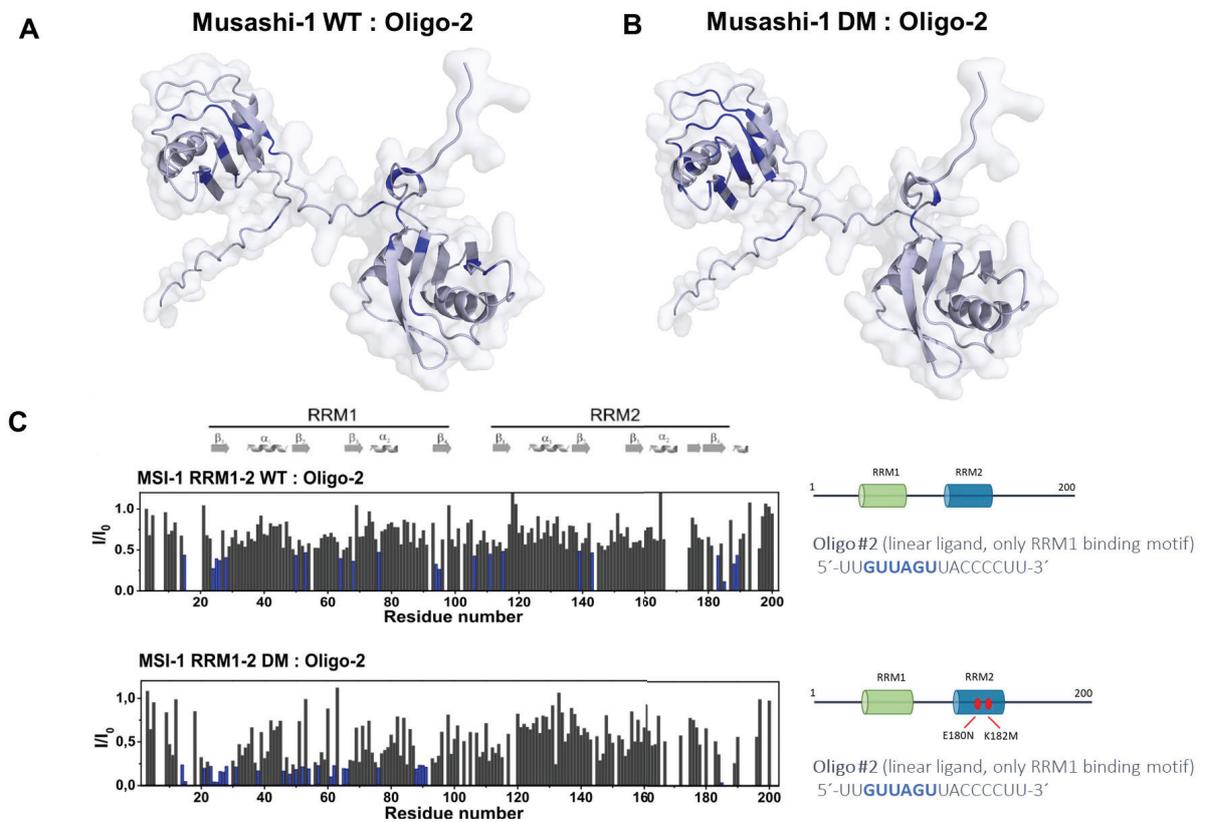


Figure 3.

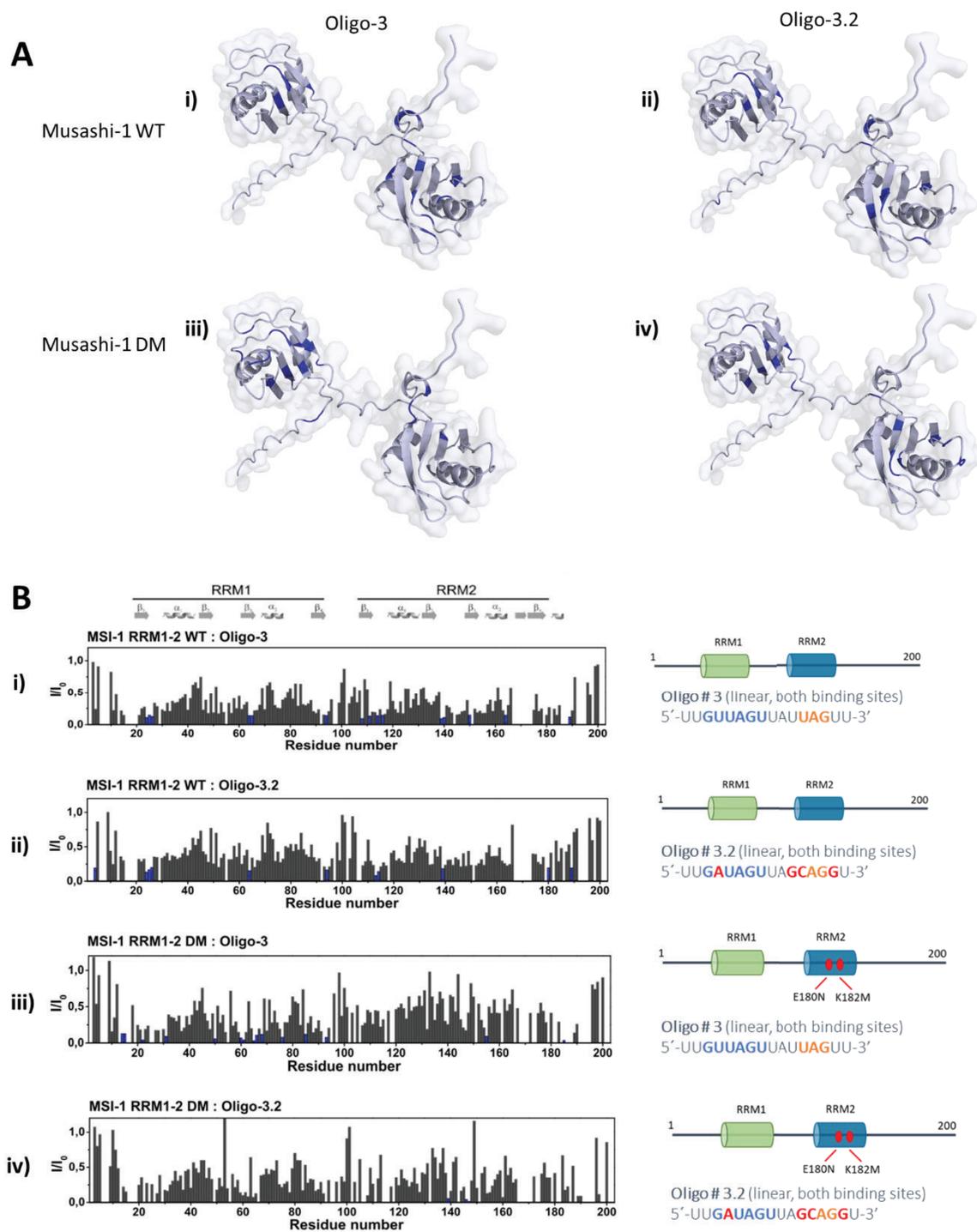


Figure 4.

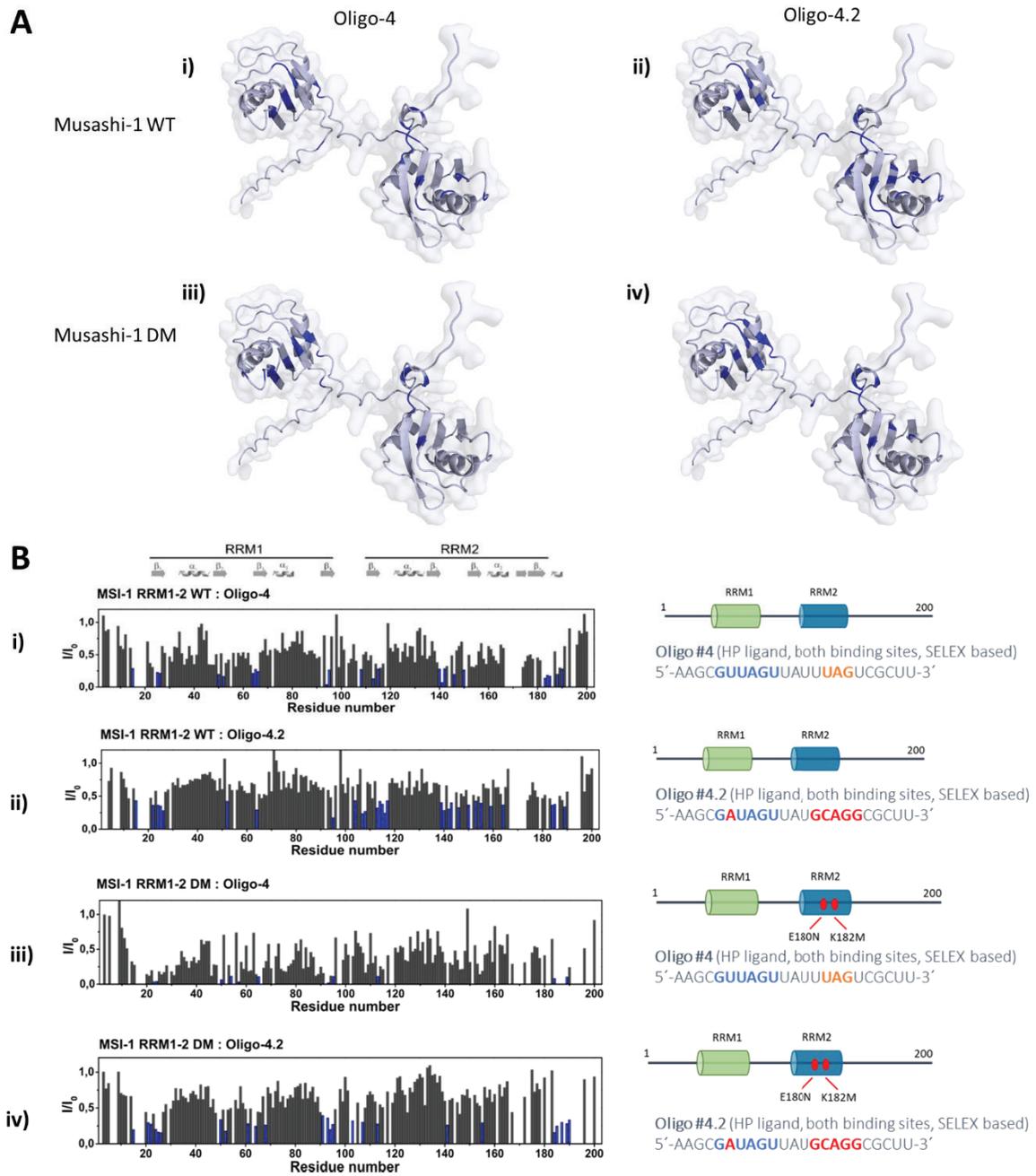


Figure 5.

4.3.2 Supplementary Information

Deciphering the RNA recognition by Musashi-1 to develop new artificial biomolecules for in vitro and in vivo applications

Figure S1. RRMScorer values of 5 nucleotide mutations scored against RRM-1 (x-axes) and RRM-2 (y-axes) of MSI-1 protein. Scores are computed for each residue-nucleotide interacting position individually and the final score is normalized from 0 to 1. Low values indicate that the RRM-RNA complex does not show very favourable contacts while higher values indicate that the contacts are very often observed in the training data set of the multiple sequence alignment.

Figure S2. One-dimensional NMR analysis comparison of MSI-1 RRM1-2 WT and MSI-1 RRM₁₋₂ DM proteins interaction with oligo-2. A) One-dimensional NMR titration steps of MSI-1 RRM1-2 WT interaction with oligo-2. B) One-dimensional NMR titration steps of MSI-1 RRM1-2 DM interaction with oligo-2.

Figure S3. One-dimensional NMR analysis comparison of MSI-1 RRM1-2 WT and MSI-1 RRM₁₋₂ DM proteins interaction with oligo-3 and oligo-3.2. A) One-dimensional NMR titration steps of MSI-1 RRM1-2 WT interaction with oligo-3. B) One-dimensional NMR titration steps of MSI-1 RRM1-2 WT interaction with oligo-3.2. C) One-dimensional NMR titration steps of MSI-1 RRM1-2 DM interaction with oligo-3. D) One-dimensional NMR titration steps of MSI-1 RRM1-2 DM interaction with oligo-3.2.

Figure S4. One-dimensional NMR analysis comparison of MSI-1 RRM1-2 WT and MSI-1 RRM₁₋₂ DM proteins interaction with oligo-4 and oligo-4.2. A) One-dimensional NMR titration steps of MSI-1 RRM1-2 WT interaction with oligo-4. B) One-dimensional NMR titration steps of MSI-1 RRM1-2 WT interaction with oligo-4.2. C) One-dimensional NMR titration steps of MSI-1 RRM1-2 DM interaction with oligo-4. D) One-dimensional NMR titration steps of MSI-1 RRM1-2 DM interaction with oligo-4.2.

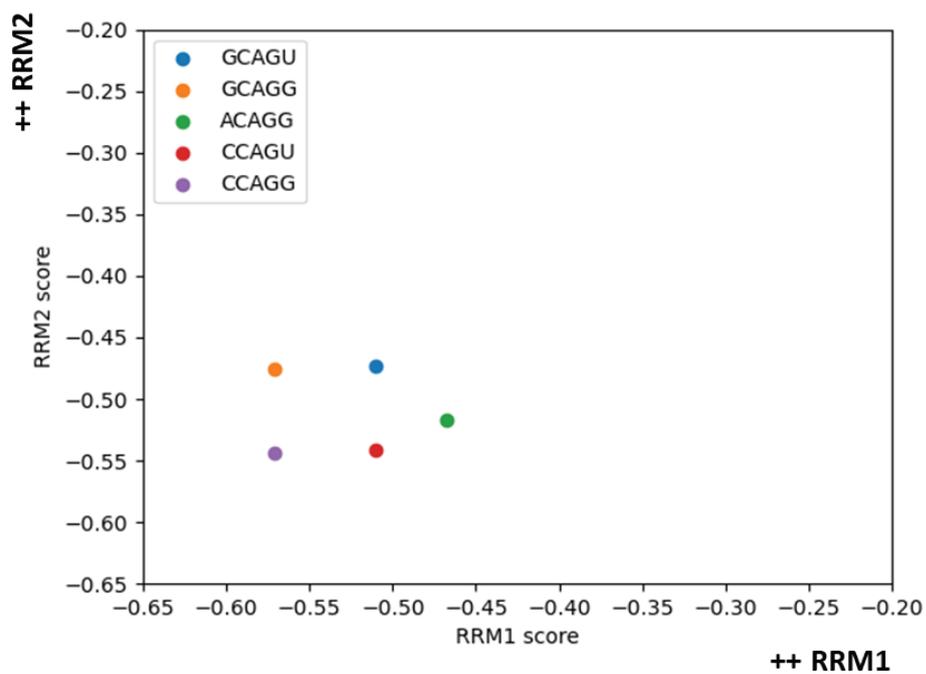
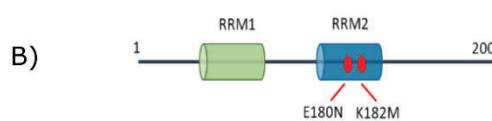
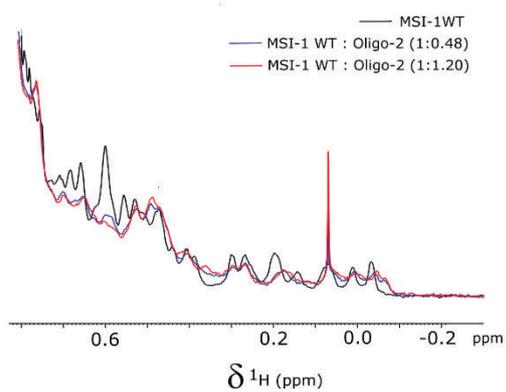


Figure S1.



Oligo #2 (linear ligand, only RRM1 binding motif)
5'-UUGUUAGUUACCCCUU-3'



Oligo #2 (linear ligand, only RRM1 binding motif)
5'-UUGUUAGUUACCCCUU-3'

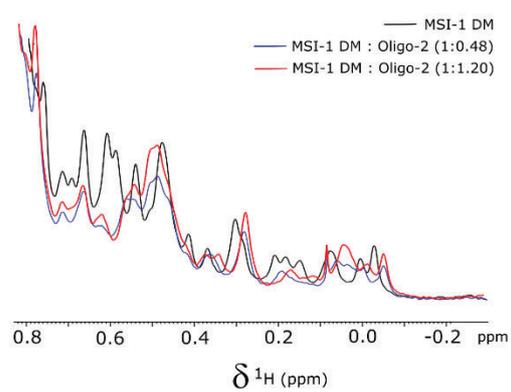


Figure S2.

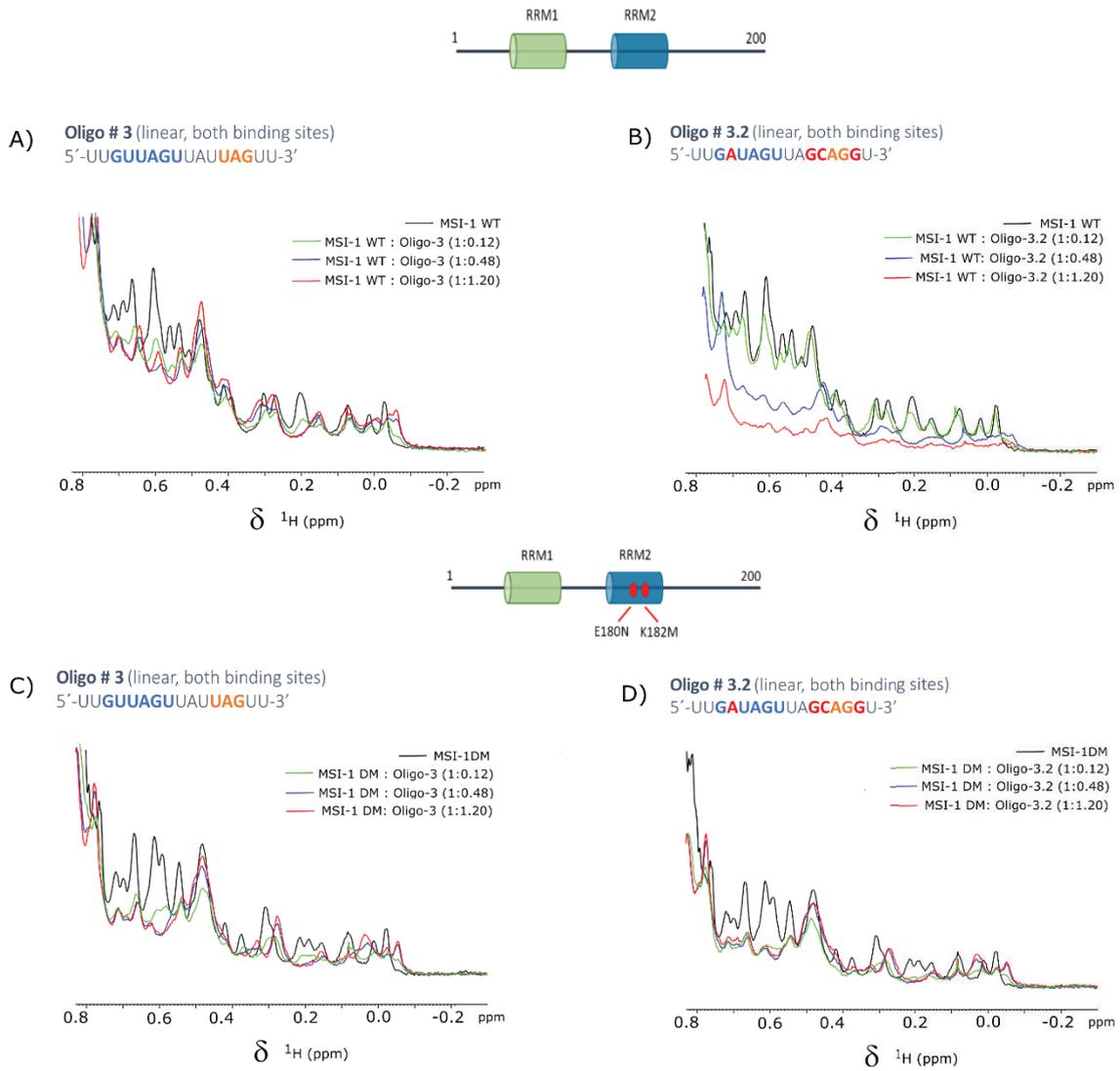


Figure S3.

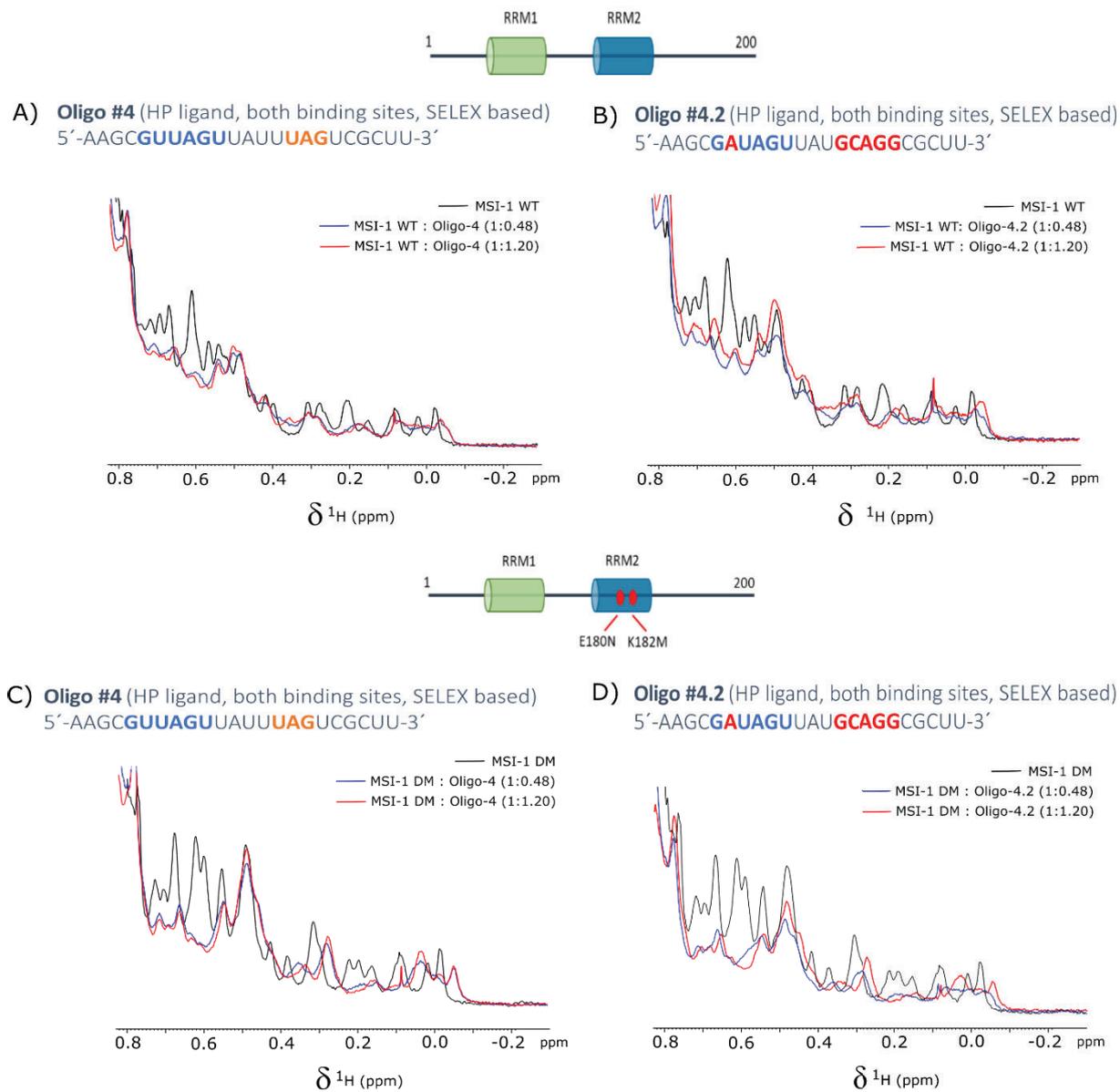


Figure S4.

4.4 ARTICLE 5:

4.4.1 Monitoring RNA-protein interactions in living bacteria in real-time

Guillermo Pérez-Ropero,^{1,2} Roswitha Dolcemascolo,³ Anna Pérez-Ràfols,^{4,5} Anahí Higuera,⁶ Jos Buij,^{2*}

¹ Sweden and Biochemistry Department, University of Uppsala, Sweden

² Ridgeview Instruments AB, Vänge, Uppsala, Sweden

³ Institute for Integrative Systems Biology (I2SysBio), CSIC-University of Valencia, València, Spain.

⁴ Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino (FI), Italy.

⁵ Giotto Biotech S.R.L., Via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy.

⁶ Dynamic Biosensors GmbH, 82152 Martinsried, Germany.

*Corresponding author

In Progress

Abstract

Musashi-1 is an RNA Binding Protein (RBP) which modulates key cellular functions by interacting with mRNA and affecting its translation. The interaction with RNA is complex as it has two RNA binding sites (RRM-1 and RRM-2) and its RNA binding, in turn, can be modulated by allosteric binders.

RNA Binding Proteins (RBP) act as modulators in key cellular functions by binding RNA and affecting its translation. Therefore, different techniques have been developed to study these interactions both *in vitro* and *in vivo*.

Here we present a new application to study RNA-protein interactions in living bacteria over time. Bacteria were transformed with a reporter system based on two plasmids; one encoding Musashi-1 and the other one an RNA sequence close to the promoter of a green fluorescent protein (GFP) sequence so that GFP expression is downregulated upon the binding. GFP expression is monitored for a bacteria cultured on a solid agar support using LigandTracer Green, an instrument designed for long time (> 24 h) monitoring of fluorescence in or on living cells in with high temporal resolution. We describe the method, its reliability, and results obtained from Musashi-1 binding to various RNA sequences and allosteric inhibitors.

In particular, my contribution to this project has been the experimental optimization of the system and performance of all replicates in collaboration with our colleagues in Sweden.

Introduction

RNA-binding proteins (RBP) mediate key interactions that modulate a variety of cellular processes such as cell differentiation¹, proliferation², and inflammation³ by modifying different aspects of the ribonucleic acid (RNA) lifecycle⁴. It is therefore no surprise that RBPs are involved in several pathologies such as neurological diseases⁵, cancer⁶, autoimmune⁷, and metabolic disorders⁸. Their biological relevance has fostered research to discover and characterize these interactions both *in vitro* and *in vivo*⁹ opening a broad horizon of applications, ranging from drug discovery¹⁰ to synthetic biology¹¹.

Several methods are used to study RNA-protein interactions¹². When expression of the protein is possible, *in vitro* techniques have been usually chosen to study binding kinetics, such as electrophoretic mobility shift assay (EMSA)¹³, fluorescence anisotropy¹⁴, and surface plasmon resonance (SPR)¹⁵. For a long time, these techniques have been the only way to extract kinetic information from RNA-protein interactions but a newly developed tool, kinetic crosslinking and immunoprecipitation (Kin-CLIP), nowadays offers the possibility of extracting association and dissociation rates and, thereby, affinity values for interactions in living cells¹⁶. Live cell interaction analysis has the advantage of studying interactions in a biological relevant environment but it also increases the complexity of the system, both for the experimental settings and the analysis. Before the appearance of this new method, the most widely used *in vivo* assays for RNA-protein interaction characterization were CLIP (crosslinking and immunoprecipitation) assays in which present RNA-protein complexes are crosslinked by UV light followed by immunoprecipitation and detection of the specific proteins using antibodies. Different modifications of the original CLIP protocol¹³, like eCLIP¹⁷, iCLIP¹⁸, HITCLIP¹⁹ have been developed. Another method described to characterize *in vivo* RNA-protein binding and extract affinity information is the use of reporter gene assays²⁰. Reporter gene assays are based on the detection of a fluorescence readout (or several readouts) associated with the expression of a certain gene or promoter²¹ and are widely used to study regulatory systems, especially in bacteria. Reporter gene assays are often conducted with bacteria in suspension but development towards reporter gene assays on solid supports has gained attention due to the strongly increased cell viability²² and lack of expression of major stress response genes^{23,24} thereby providing a superior platform for drug screening²⁵. Therefore, a method is required that monitors the expression of fluorescent proteins in sessile bacteria rather than planktonic cells²⁶.

LigandTracer (LT) technology has been widely used for real-time cell binding assays (RT-CBA) and is typically used to measure the binding kinetics and calculate the affinity of proteins for receptors on living cells that are adhered or tethered to a surface of a Petri dish. These measurements can be performed in environment-controlled cabinets like heat chambers or incubators and enable online monitoring of reporter gene expression with high temporal resolution. In contrast with previously mentioned *in vivo* techniques to study RNA-protein interactions, LT offers the possibility to monitor fluorescence in bacteria on solid support rather than in suspension²⁸. In this study, however, bacteria were captured on a Petri dish under conditions designed to minimize cell growth. This approach is not suitable for monitoring reporter gene assays as an environment that

allows bacteria to grow and express proteins is required. For this reason, an assay where bacteria are cultured on agar coated plates was developed so that the repression of Green Fluorescence Protein (GFP), as reported of the interaction between the protein and RNA target sequences, can be studied in biofilms.

The RBP studied with this new method was Musashi-1 binding to several RNA strands. These RNA strands were selected based on previous binding studies²⁹. Musashi-1 is an RBP, part of the RNA recognition motif (RRM) family, which is characterized because it harbours two RNA binding sites (RRM-1 and RRM-2)³⁰, which display a typical $\alpha\beta$ sandwich structure with a $\beta\alpha\beta\beta\alpha\beta$ distribution³¹. This protein also has a binding site for fatty acids (FA) (showing a higher affinity for mono-unsaturated FA)³². It plays a biological role as a regulator of stem cell differentiation in neural cells³³ and in pathological states has been associated with several diseases ranging from cancer³⁴⁻³⁶ to Alzheimer³⁷.

Material and Methods

Bacteria

Two strands of Gram-negative bacteria *Escherichia coli* were used: BL21 for the LigandTracer reporter assays and DH5 α for plasmid amplification. Bacteria came from a frozen stock and for both plasmid amplification and the assays they were previously inoculated in Luria-Broth (LB) medium (2 mL) and incubated overnight at 37°C, 220 rpm (InforsHT Minitron).

RNA strands

RNA oligonucleotides sequences in plasmid pREP6 with single-point mutations (highlighted in red):

WT: 5'-GGCAGCGTTAGTTATTTAGTTCGTATGCC-3';

Mutant-1: 5'-GGCAGCGTT**C**GTTATTTAGTTCGTATGCC-3';

Mutant-2: 5'-GGCAGCGTTA**C**TTATTTAGTTCGTATGCC-3';

Mutant-3: 5'-GGCAGCGTTAG**C**TATTTAGTTCGTATGCC-3';

Mutant-5: 5'-GGCAGCGTTAGTTAT**G**TTAGTTCGTATGCC-3' and

Mutant-6: 5'-GGCAGCGTTA**C**TTATTTTA**C**TTCGTATGCC-3') used for LigandTracer experiments were provided by CSIC, València.

Plasmids and plasmid amplification

Two RBP systems were tested with our method; one based on bacteriophage MS2 coat protein and a second based on Musashi-1 protein. Each system is composed of two plasmids: p15A and pSC101, that are jointly transformed to *E. coli* cells. p15A Contains a constitutive promotor with 3 different elements: an RNA binding motif close to sfGFP sequence, an sfGFP-encoding sequence, and an RFP-encoding sequence. pSC101 contains an inducible plasmid, pLac (inducible by lactose or analogues such as Isopropyl β -d-1-

thiogalactopyranoside, IPTG, in this case), encoding a Musashi-1 sequence comprising both RNA-recognition motifs (RRMs) (1-196 residues).

Preparation of chemo-competent cells

A single colony of *E. coli* BL21 strain was incubated in 5 mL LB-Miller medium at 37°C with shaking (220 rpm; InforsHT Minitron) for 24 h. Prewarmed LB-Miller medium (30 mL) was inoculated with 1% (v/v) of the overnight culture and grown for 3 hours to an $OD_{600} = 0.5$. Cells were harvested by centrifugation (4 000 x g, 4°C, 15 min; Heraeus Megafuge 8R). In the following, all solutions were precooled in ice. The pellet was resuspended in 9 mL Tfb1 buffer and 3.2 mL 1 mol dm⁻³ MgCl₂ was added. It was incubated on ice for 15 min. After centrifugation, the pellet was resuspended in 1.2 mL Tfb2 and incubated on ice for 15 min. After incubation, cells were divided in 50 µL aliquots that are snap-freezing in liquid nitrogen and then storage at -80°C.

Transformation

E. coli BL21 transformation was done using a heat-shock protocol. 50 µL of competent cells were incubated with 1 µL of each DNA plasmid (in Eppendorf tubes) for 30 min on ice. Heat-shock was performed by a 45 s incubation at 42 °C followed by 2 min incubation on ice followed by adding 1 mL of pre-warmed SOC medium and a 1 h incubation at 37 °C. Bacteria were seeded in an agar plate supplemented with appropriate antibiotics.

Dish coating

Untreated MultiDish 2x2 (1-04-202, Ridgeview Instruments AB) was coated with LB agar (see Figure 1). MultiDishes (MD) were chosen as a support because it has four independent sections that enables monitoring and comparing four different experimental conditions per assay. Agar was coated by applying 4 mL LB agar in each of the four sections and letting it solidify at RT for 1 h. Different techniques were used for seeding: uniform seeding of the whole section with 20 µL of bacteria suspension in LB medium ($OD_{600} = 0.5$); and drop seeding of just the detection area followed by incubation at RT until the drop was completely absorbed by the agar.

LigandTracer

After seeding the bacteria, the MultiDishes were placed on the rotating support in LigandTracer Green (1-04-002, Ridgeview Instruments AB). To detect fluorescence from GFP, the LigandTracer was equipped with a BlueGreen (488-535 nm) detector. The LigandTracer assay was set to detect fluorescence coming from the four different detection spots on each section of a MultiDish every 15 seconds.

Data evaluation

Gompertz model, a mathematical model used to fit sigmoidal curves, was used to fit the obtained curves. The used equation allowed us to obtain three different parameters: the reduction in the fluorescence signal (A); the growth-rate coefficient (k) and the time of inflection (T)³⁸.

$$\text{Signal}(t) = A e^{-e^{-k(t-T)}}$$

Results and Discussion

eGFP fluorescence signal measured in real-time in live bacteria.

E. coli BL21 cells were transformed with two different plasmids (see Figure 2). p15, inducible by IPTG, expresses Musashi-1 protein in presence of lactose analogues and a second one, pSC101, that contains an RNA binding motif close to the GFP promoter, constitutively expresses the RNA sequence. This reporter gene circuit was engineered to inhibit GFP expression upon MSI-1 binding by blocking RNA polymerase binding to the promoter. In absence of IPTG (Figure 2A), no expression of Musashi-1 is observed while high expression of GFP in cells is detected and its signal is monitored with LigandTracer. When IPTG is present in the medium (Figure 2B), Musashi-1 is expressed and binding to the RNA motif results in a reduced GFP signal compared to the situation where Musashi-1 expression is repressed.

Monitorization of the fluorescence signal over time up to 20-24 hours is recorded by LigandTracer and generates baseline-corrected fluorescence intensity curves as function of time for each of the four sections of the MultiDish. The typical signal output follows a sigmoidal pattern, which is characterized by a lag-phase of 4 to 5 hours followed by an exponential phase that ends in a plateau after roughly 20 h, as shown in Figure 3A.

Six different RNA strands were tested to monitor the effect of the MSI-1: RNA interaction on the fluorescence signal. The WT oligonucleotide sequence designed is an RNA fragment from the numb mRNA³⁹ that contains both the binding motifs (the (G/A)U₁₋₃AGU motif for RRM-1³⁹ and the UAG motif for RRM-2⁴⁰). Single point mutations at different positions were performed in the WT sequence (Mut-1, Mut-2, Mut-3, Mut-5 and Mut-6) to evaluate the implication that specific nucleic bases had on the binding.

MSI-1 binding to RNA has been proven to be fast, which allows us to assume that the binding process is mainly driven by the affinity and not by the binding kinetics. Due to this assumption, we concluded that the signal reduction is directly related to the fraction of MSI-1-RNA complexes, which at the same time is in function of the binding affinity and the MSI-1 and RNA concentrations.

The sigmoidal fluorescence curves were fitted using the Gompertz model (Figure 3B). The obtained results, that corresponds to MSI-1 RRM₁₋₂(1-200) binding to its RNA (Wild Type) sequence, showed an almost

80% reduction in fluorescence when the protein is expressed, in comparison with cells that were not treated with IPTG. This change was observed over all different replicates with a high consistency and allowed us to validate the LigandTracer platform as a device able to monitor gene reporter systems, and specifically, to qualitatively detect RNA-protein binding in living bacteria over time.

Regarding the mutants, the major effect has been detected in mutants 2,3 and 6 (Figure 4B). These constructs presented mutations on the second (Mut-2) and third (Mut-3) nucleotides of the three-nucleotides binding motif for RRM-1, while mutant 6 presented mutations both in the binding motif for RRM-1 and RRM-2.

Conclusion

RNA-protein characterization in living cells by using gene reporter assays is usually performed in suspension cells. In this study we were able to establish a methodology to measure fluorescence changes of bacteria in agar during long periods, showing its utility in the characterization of interactions in living cells.

Compared with usual protocols for suspension cells, the sample throughput, the impossibility to use liquid handling, and the limit of measuring several fluorescence readouts or absorbance at the same time are some of the most important limitations of our method. However, this method offers the possibility of tracking cell behaviour in the long run and in a relevant physiologic environment.

In conclusion, it is possible to use LigandTracer to track and characterize reporter gene assays over time, as it has been shown in the characterization of Musashi-1 binding to RNA.

Experiments with an allosteric inhibitor (oleic acid) of MSI-1 are currently being performed to evaluate the regulation of protein-RNA binding using this reporter system in LigandTracer technology.

Figure Captions:

Figure 1. Modification of the LigandTracer set-up. MultiDishes (MD) 2x2 (1-04-202, Ridgeview Instruments AB) coated with agar and used to test four different experimental conditions and Blue Green detector on a Ligand Tracer instrument were used for the fluorescence detection.

Figure 2. Reporter gene system of Musashi-1 protein. Representation of the system in absence A) or in presence B) of IPTG in the medium.

Figure 3. A) In black, fluorescence signal over time corresponding to bacteria where MSI-1 expression has not been induced. In red, fluorescence signal over time corresponding to bacteria where MSI-1 expression has been induced. B) In blue, Gompertz fitting corresponding to bacteria where MSI-1 expression has not been induced, the experimental data is in black. In green, Gompertz fitting corresponding to bacteria where MSI-1 expression has been induced, the experimental data is in red.

Figure 4. A) RNA sequences studied. B) Graphic representation of the fluorescent reduction signal observed for each of the studied RNA.



Figure 1.

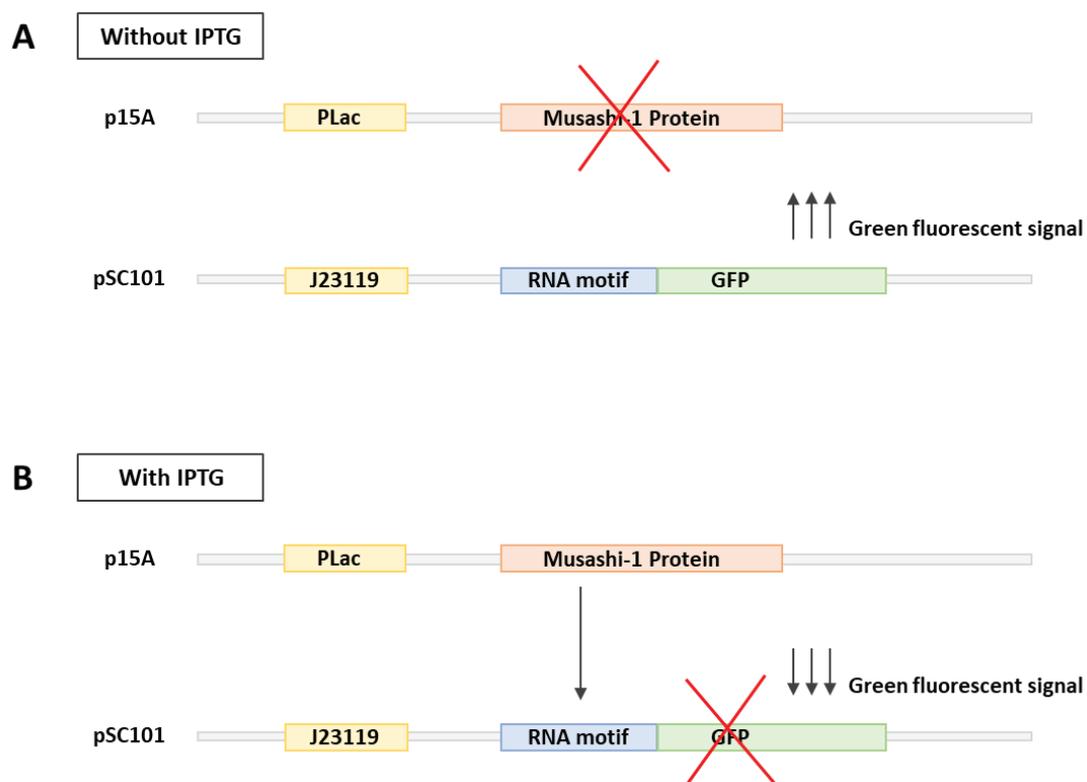


Figure 2.

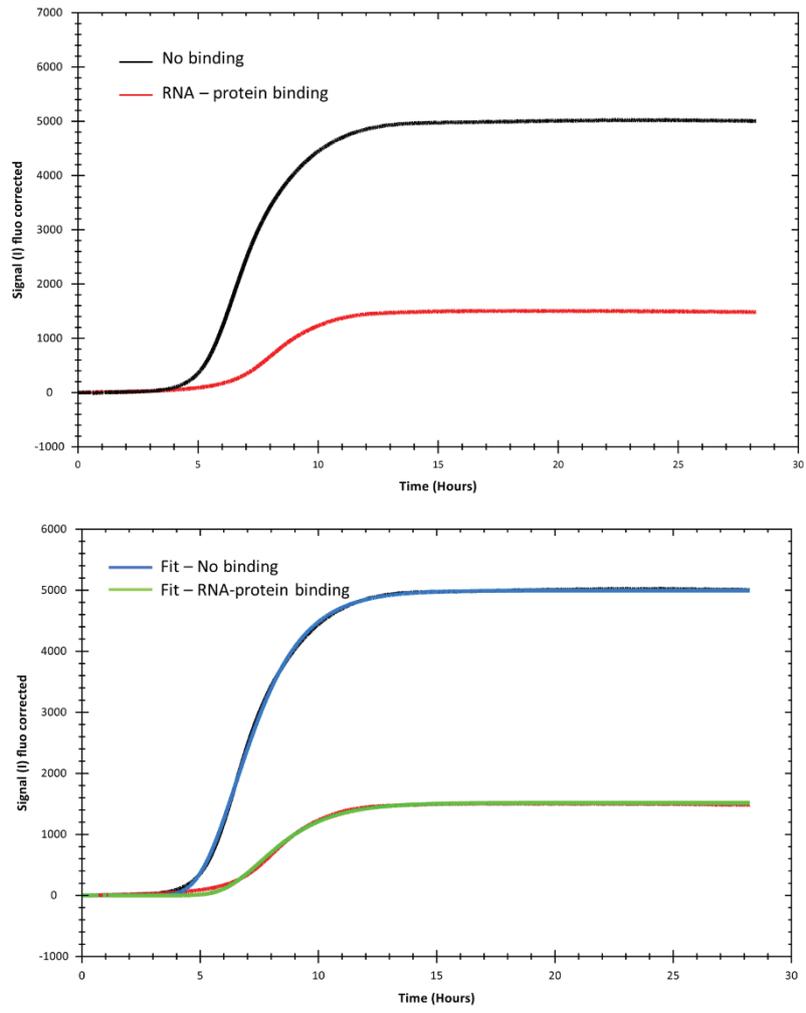


Figure 3.

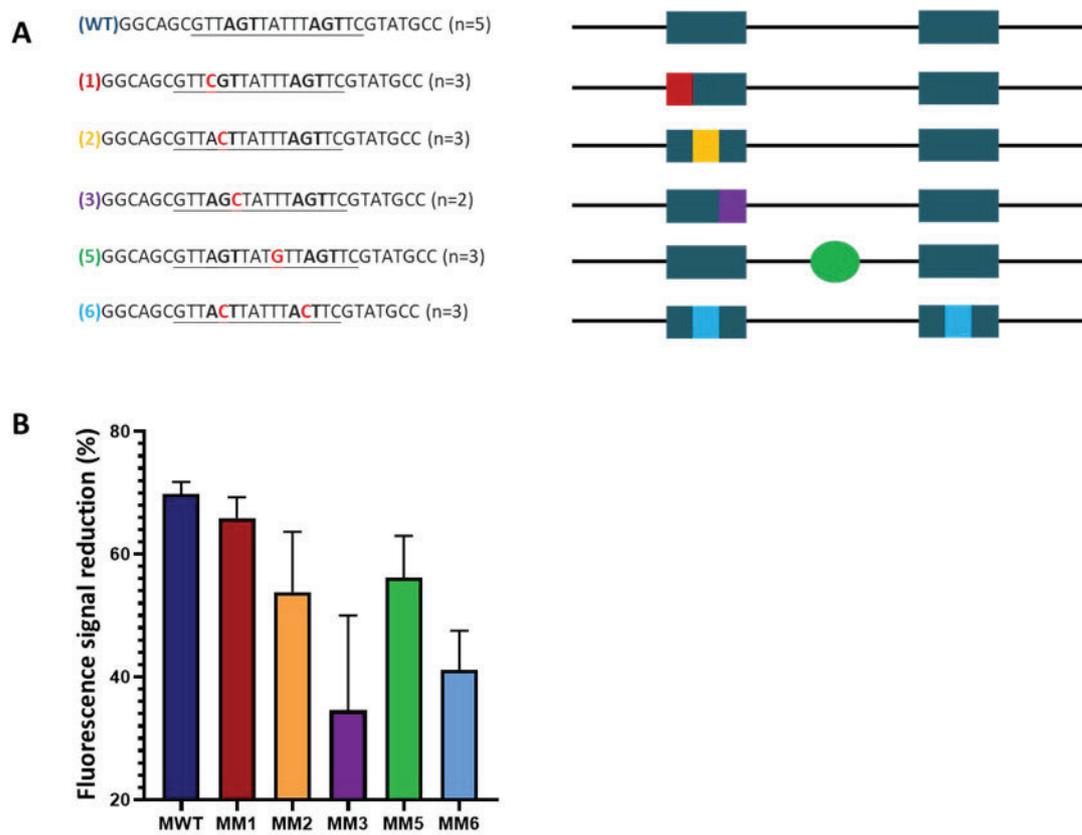


Figure 4.

References

1. Chen, J. et al. RNA-Binding Protein HuR Promotes Th17 Cell Differentiation and Can Be Targeted to Reduce Autoimmune Neuroinflammation. *J. Immunol.* 204, 2076–2087 (2020)
2. Cho, S. J., Jung, Y. S., Zhang, J. & Chen, X. The RNA-binding protein RNPC1 stabilizes the mRNA encoding the RNA-binding protein HuR and cooperates with HuR to suppress cell proliferation. *Journal of Biological Chemistry* 287, 14535–14544 (2012).
3. Nyati, K. K., Zaman, M. M. U., Sharma, P. & Kishimoto, T. Arid5a, an RNA-Binding Protein in Immune Regulation: RNA Stability, Inflammation, and Autoimmunity. *Trends in Immunology* 41, 255–268 (2020).
4. Liu, S. et al. Classification and function of RNA–protein interactions. *Wiley Interdisciplinary Reviews: RNA* 11, 1–27 (2020).
5. Meneses, A. et al. TDP-43 Pathology in Alzheimer 's Disease. *Molecular Neurodegeneration* 1–15 (2021).
6. Cava, C., Armaos, A., Lang, B., Tartaglia, G. G. & Castiglioni, I. Identification of long non-coding RNAs and RNA binding proteins in breast cancer subtypes. *Scientific Reports* 12, 1–13 (2022).
7. Lai, H.-C., Ho, U. Y., James, A., de Souza, P. & Roberts, T. L. RNA metabolism and links to inflammatory regulation and disease. *Cellular and Molecular Life Sciences* 79, (2022).
8. Good, A. L. & Stoffers, D. A. Stress-induced translational regulation mediated by RNA binding proteins: Key links to β -cell failure in diabetes. *Diabetes* 69, 499–507 (2020).
9. Marchese, D., de Groot, N. S., Lorenzo Gotor, N., Livi, C. M. & Tartaglia, G. G. Advances in the characterization of RNA-binding proteins. *Wiley Interdisciplinary Reviews: RNA* 7, 793–810 (2016).
10. Julio, A. R. & Backus, K. M. New approaches to target RNA binding proteins. *Current Opinion in Chemical Biology* 62, 13–23 (2021).
11. Nakanishi, H. Protein-based systems for translational regulation of synthetic mRNAs in mammalian cells. *Life* 11, (2021).
12. Ramanathan, M., Porter, D. F. & Khavari, P. A. Methods to study RNA–protein interactions. *Nature Methods* 16, 225–234 (2019).
13. Ryder, S. P., Recht, M. I. & Williamson, J. R. Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods in Molecular Biology* 488, 99–115 (2008).
14. Pagano, J. M., Clingman, C. C. & Ryder, S. P. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *Rna* 17, 14–20 (2011).
15. Katsamba, P. S., Park, S. & Laird-offringa, I. A. Kinetic studies of RNA – protein interactions using surface plasmon resonance. *Methods* 26, 95–104 (2002).
16. Sharma, D. et al. The kinetic landscape of an RNA-binding protein in cells. *Nature* 591, 152–156 (2021).
17. van Nostrand, E. L. et al. A large-scale binding and functional map of human RNA-binding proteins. *Nature* 583, 711–719 (2020).

18. Hauer, C. et al. Improved binding site assignment by high-resolution mapping of RNA-protein interactions using iCLIP. *Nature Communications* 6, 1–13 (2015).
19. Sugimoto, Y. et al. HiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* 519, 491–494 (2015).
21. Paraskeva, E., Atzberger, A. & Hentze, M. W. A translational repression assay procedure (TRAP) for RNA-protein interactions in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 95, 951–956 (1998).
22. Alvarez, G. S., Foglia, M. L., Copello, G. J., Desimone, M. F. & Diaz, L. E. Effect of various parameters on viability and growth of bacteria immobilized in sol-gel-derived silica matrices. *Applied Microbiology and Biotechnology* 82, 639–646 (2009).
23. Premkumar, J. R. et al. Fluorescent bacteria encapsulated in sol-gel derived silicate films. *Chemistry of Materials* 14, 2676–2686 (2002).
24. Powers, T. & Noller, H. F. Dominant lethal mutations in a conserved loop in 16S rRNA (site-directed mutagenesis/rRNA mutations/A PL promoter/ribosomal A site). *Biochemistry* 87, 1042–1046 (1990).
25. Eleftheriou, N. M. et al. Entrapment of living bacterial cells in low-concentration silica materials preserves cell division and promoter regulation. *Chemistry of Materials* 25, 4798–4805 (2013).
26. Fujikawa, H. & Morozumi, S. Modeling Surface Growth of *Escherichia coli* on Agar Plates. *Applied And Environmental Microbiology* 71, 7920–7926 (2005).
27. Ciofu, O., Moser, C., Jensen, P. Ø. & Høiby, N. Tolerance and resistance of microbial biofilms. *Nature Reviews Microbiology* 0123456789, (2022).
28. Encarnação, J. C., Schulte, T., Achour, A., Björkelund, H. & Andersson, K. Detecting ligand interactions in real time on living bacterial cells. *Applied Microbiology and Biotechnology* 102, 4193–4201 (2018).
29. Ruth Zearfoss, N. et al. A conserved three-nucleotide core motif defines musashi RNA binding specificity. *Journal of Biological Chemistry* 289, 35530–35541 (2014).
30. Kurihara, Y. et al. Structural properties and RNA-binding activities of two RNA recognition motifs of a mouse neural RNA-binding protein, mouse-Musashi-1. *Gene* 186, 21–27 (1997).
31. Maris, C., Dominguez, C. & Allain, F. H. T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal* 272, 2118–2131 (2005).
32. Clingman, C. C. et al. Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *eLife* 2014, 1–26 (2014).
33. Okano, H., Imai, T. & Okabe, M. Musashi: A translational regulator of cell fate. *Journal of Cell Science* 115, 1355–1359 (2002).
34. Lin, J. C., Tsai, J. T., Chao, T. Y., Ma, H. I. & Liu, W. H. Musashi-1 Enhances Glioblastoma Migration by Promoting ICAM1 Translation. *Neoplasia (United States)* 21, 459–468 (2019).
35. Troschel, F. M. et al. Knockdown of Musashi RNA binding proteins decreases radioresistance but enhances cell motility and invasion in triple-negative breast cancer. *International Journal of Molecular Sciences* 21, (2020).

-
36. Nikpour, P. et al. The RNA binding protein Musashi1 regulates apoptosis, gene expression and stress granule formation in urothelial carcinoma cells. *Journal of Cellular and Molecular Medicine* 15, 1210–1224 (2011).
 37. Montalbano, M. et al. RNA-binding proteins Musashi and tau soluble aggregates initiate nuclear dysfunction. *Nature Communications* 11, 1–16 (2020).
 38. Tjørve, K. M. C. & Tjørve, E. The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. *PLoS One* 12, 1–17 (2017).
 39. Imai, T. *et al.* The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA . *Mol. Cell. Biol.* **21**, 3888–3900 (2001).
 40. Iwaoka, R. *et al.* Structural Insight into the Recognition of r(UAG) by Musashi-1 RBD2, and Construction of a Model of Musashi-1 RBD1-2 Bound to the Minimum Target RNA. *Molecules* **22**, 1–16 (2017).

4.5 ARTICLE 6:

4.5.1 Evaluation of the Higher Order Structure of Biotherapeutics Embedded in Hydrogels for Bioprinting and Drug Release

^{1,2}Domenico Rizzo, ¹Linda Cerofolini, ^{2,3}Anna Pérez-Ràfols, ^{1,2}Stefano Giuntini, Fabio Baroni, ^{1,2}Enrico Ravera, ^{1,2}Claudio Luchinat, and ^{1,2}Marco Fragai*

¹ Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy;

² Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy

³ Giotto Biotech, S.R.L, Sesto Fiorentino, Florence 50019, Italy

⁴ Analytical Development Biotech Department, Merck Serono S.p.a, Merck KGaA, Guidonia, Rome 00012, Italy

Analytical Chemistry, 2021, 93, 11208–11214

DOI: <https://doi.org/10.1021/acs.analchem.1c01850>

New frontiers in medicine and new biotechnological opportunities are rising from the continuous development of new biocompatible materials. Biocompatible hydrogels like those formed by hyaluronic acid or mixtures of alginate and gelatine are used to load proteins as drugs or modulators of the biological activity. However, while loading and once the protein is introduced in the matrix, it needs to maintain its native structure, thus preserve its higher order structure (HOS) in order to be able to perform its therapeutic function.

Several biophysical techniques are used to characterize the protein in these matrixes. However, these techniques normally are not sensitive to local changes in the protein fold.

In line with this, we aim to exploit the use of solid-state NMR and demonstrate its applicability for studies of the conservation of protein's HOS when embedded in hyaluronic acid or alginate-gelatine hydrogels. To accomplish this goal, two well-known proteins used in therapeutics (TTR and ANSII) are chosen to investigate how the matrices used for 3D bioprinting interplay with embedded proteins have been used as models and demonstrate this potential application for solid-state NMR.

In this project, my contribution has been the expression and characterization of both U-¹³C-¹⁵N asparaginase II and U-¹³C-¹⁵N Transthyretin and the preparation of hydrogel matrices.

Evaluation of the Higher Order Structure of Biotherapeutics Embedded in Hydrogels for Bioprinting and Drug Release

Domenico Rizzo, Linda Cerofolini, Anna Pérez-Ràfols, Stefano Giuntini, Fabio Baroni, Enrico Ravera, Claudio Luchinat, and Marco Fragai*

Cite This: <https://doi.org/10.1021/acs.analchem.1c01850>

Read Online

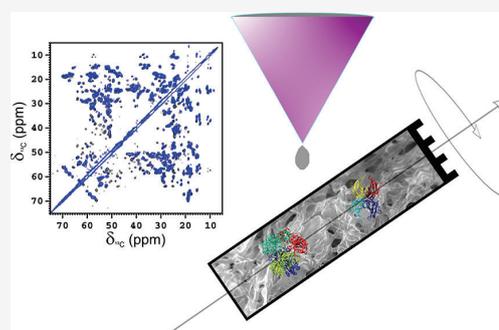
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Biocompatible hydrogels for tissue regeneration/replacement and drug release with specific architectures can be obtained by three-dimensional bioprinting techniques. The preservation of the higher order structure of the proteins embedded in the hydrogels as drugs or modulators is critical for their biological activity. Solution nuclear magnetic resonance (NMR) experiments are currently used to investigate the higher order structure of biotherapeutics in comparability, similarity, and stability studies. However, the size of pores in the gel, protein–matrix interactions, and the size of the embedded proteins often prevent the use of this methodology. The recent advancements of solid-state NMR allow for the comparison of the higher order structure of the matrix-embedded and free isotopically enriched proteins, allowing for the evaluation of the functionality of the material in several steps of hydrogel development. Moreover, the structural information at atomic detail on the matrix–protein interactions paves the way for a structure-based design of these biomaterials.



INTRODUCTION

The continuous development of new biocompatible materials is opening new frontiers in medicine and new biotechnological opportunities. Several biomaterials are currently used to replace/support non-functional tissues like those damaged or destroyed by injuries or diseases and in controlled drug release. Materials for tissue regeneration are designed to provide mechanical support to the surrounding tissue, to stimulate cell growth, and to modulate the immune response promoting an extensive cell colonization and matrix reabsorption.^{1,2} Composite scaffolds with a highly resolved architecture, incorporating proteins and seeding cells, can be obtained by three-dimensional (3D) bioprinting techniques starting from biocompatible hydrogels like those formed by hyaluronic acid^{3–7} or mixtures of alginate and gelatine.^{6,8–13} In this context, there is an increasing interest in loading proteins on hydrogels as drugs or modulators of the biological activity.^{14–25} The biological function of a protein is strictly related to its native folding, and the preservation of the higher order structure (HOS) in the composite biomaterial is crucial for its therapeutic function. Actually, the interaction of the protein with the matrix components can alter the protein structure leading to a loss of activity and immunological effects.

Several biophysical methodologies, such as attenuated total reflectance Fourier-transformed infrared and fluorescence spectroscopy, circular dichroism, and differential scanning calorimetry, are usually used to characterize the protein

component in heterogeneous materials.^{26–28} However, these analytical methods measure different aspects of the structure, either directly or indirectly, and are often not sensitive enough to small, local changes in the protein fold. Nuclear magnetic resonance (NMR) and mass spectrometry are well-established techniques to investigate the preservation of the HOS of biologics in solution.^{29–39} Solution NMR has been used previously on small proteins and peptides embedded in hydrogels to investigate the folding state in a confined environment⁴⁰ and for the structural characterization through residual dipolar couplings, since hydrogels behave as anisotropic external alignment media.^{41–43} However, when the size of the pores in the gels is too small or strong interactions between the gel matrix and the cargo protein take place, the rotational correlation time of the protein in solution increases and makes solution NMR ineffective in the analysis of the protein structure at the atomic level.

Recently, solid-state NMR has emerged as a tool to characterize the protein component and to reveal protein–matrix interactions in heterogeneous materials. In this respect,

Received: April 30, 2021

Accepted: July 20, 2021

the use of solid-state NMR has been described to characterize noncrystalline large protein assemblies,^{44–50} biomaterials,^{51,52} bioinspired silica matrix embedding enzymes,^{53–58} conjugated proteins,^{59–62} protein-grafted nanoparticles,⁶³ and vaccines.^{64–66} Here, we prove that solid-state NMR provides detailed information on the preservation of the HOS of proteins embedded into two popular matrices used for 3D bioprinting.

The therapeutic protein *E. coli* asparaginase-II (ANSII), clinically used against acute lymphoblastic leukemia, has recently shown its activity also against solid tumor when administered in long half-life formulations that reduce immunological adverse reactions.⁶⁷

Human transthyretin (TTR) is a physiological protein acting as a hormone carrier.^{68,69} Although some genetic variants of TTR lead to a systemic amyloidosis called familial amyloid polyneuropathy,⁷⁰ TTR is a potential drug carrier and has been recently proposed as a multivalency Fab platform for target clustering.⁷¹

Therefore, these two proteins are suitable models to investigate how the matrices used for 3D bioprinting interplay with embedded proteins and are used here to prove the potential of solid-state NMR (SSNMR) in the characterization of the protein components during the design of these composite hydrogels.

EXPERIMENTAL SECTION

Sample Preparation and NMR Measurements.

[U-¹³C-¹⁵N] ANSII was expressed and purified as previously described.^{59,61–64} The expression and purification protocol of [U-¹³C-¹⁵N] TTR is reported in the Supporting Information. All the hydrogels embedding the selected proteins (ANSII and TTR) were directly generated in Bruker 3.2 mm thin-walls zirconia rotors with bottom and top caps, starting from the dried materials prepared by using the different procedures described below.

The sample of [U-¹³C-¹⁵N] ANSII encapsulated in the alginate/gelatine hydrogel was prepared by incorporating the freeze-dried protein (4 mg) into a mixture of 1:1 alginate/gelatine powders (5 mg) and then by rehydrating the dried mixture within the rotor.⁷² A different procedure was used to prepare the sample of [U-¹³C-¹⁵N] TTR encapsulated in the alginate/gelatine hydrogel. The dry mixture containing TTR was prepared by lyophilizing a solution containing all the components (6 mg of protein and 5 mg of the 1:1 alginate/gelatine mixture). In both cases, the dried material was packed in the rotor and hydrated with MilliQ H₂O to reach a final concentration of ~5–7% w/w for alginate and gelatine. Finally, a concentrated solution of CaCl₂ (to reach a concentration of 100 mM in the rotor) was added to cross-link the hydrogel materials within the rotor.^{73,74}

A sample of [U-¹³C-¹⁵N] TTR protein encapsulated in the alginate/gelatine hydrogel was also analyzed by solution NMR. The gel was prepared by dissolving a mixture of alginate and gelatine (~7% w/w) in 600 μL of a solution of TTR (100 μM in 50 mM MES, pH 6.5, 100 mM NaCl, 5 mM DTT). Then, the material was transferred in a 5 mm tube and cross-linked by adding a concentrated solution of CaCl₂ (to reach a concentration of 100 mM) in the NMR tube. The 2D ¹H-¹⁵N TROSY-HSQC spectrum recorded on the encapsulated protein was superimposed with that of TTR collected in solution (see Figure S1).

The hyaluronic acid hydrogels encapsulating the selected proteins ([U-¹³C-¹⁵N] ANSII or TTR) were prepared by packing the rotor with consecutive layers of the freeze-dried protein (~4–6 mg) and freeze-dried hyaluronic acid (Jonexa, 7–9 mg), which had been previously dialyzed against MilliQ H₂O to remove the excess of salts. The material was finally rehydrated with MilliQ H₂O (from 10 to 20 μL). Sample homogeneity was obtained after rotor spinning and supported by the quality of the spectra that suggests the presence of a protein experiencing a single environment.

Samples of freeze-dried proteins were prepared as reference. The free proteins (~20 and 25 mg of ANSII and TTR, respectively) were freeze-dried in the presence of PEG1000 (4 and 2.5 mg for ANSII and TTR, respectively); the materials were packed into a Bruker 3.2 mm zirconia rotor and rehydrated with MilliQ H₂O (~9 and 16 μL for ANSII and TTR, respectively). CaCl₂ was not present in the samples of rehydrated freeze-dried proteins.

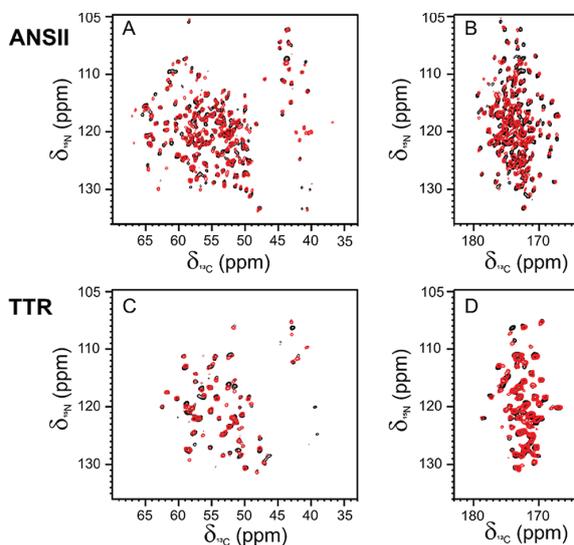


Figure 1. (A, C) 2D ¹⁵N ¹³C NCA and (B, D) NCO spectra of ANSII-HA (red, top) and TTR-HA (red, bottom) superimposed with NCA and NCO of the rehydrated freeze-dried reference proteins (black). The spectra were acquired at ~290 K, MAS 14 kHz and 800 MHz.

Silicon plugs (courtesy of Bruker Biospin) placed below the turbine cap were used to close the rotor and preserve hydration.

SSNMR experiments were recorded on a Bruker Avance III spectrometer operating at 800 MHz (18.8 T, 201.2 MHz ¹³C Larmor frequency) equipped with a Bruker 3.2 mm Efree NCH probe-head. The spectra were recorded at 14 kHz MAS frequency, and the sample temperature was kept at ~290 K. The sample of the alginate/gelatine hydrogel encapsulating TTR was also investigated at a higher spinning frequency (16 and 20 kHz).

Standard ¹³C-detected SSNMR spectra (2D ¹⁵N-¹³C NCA, ¹⁵N-¹³C NCO, and ¹³C-¹³C DARR, mixing time 50 ms) were acquired on all the samples (except for TTR encapsulated in the alginate/gelatine hydrogel) using the pulse sequences reported in the literature.⁷⁵ 2D ¹³C-¹³C CORDxy⁴⁶ was instead recorded for the sample of the alginate/gelatine

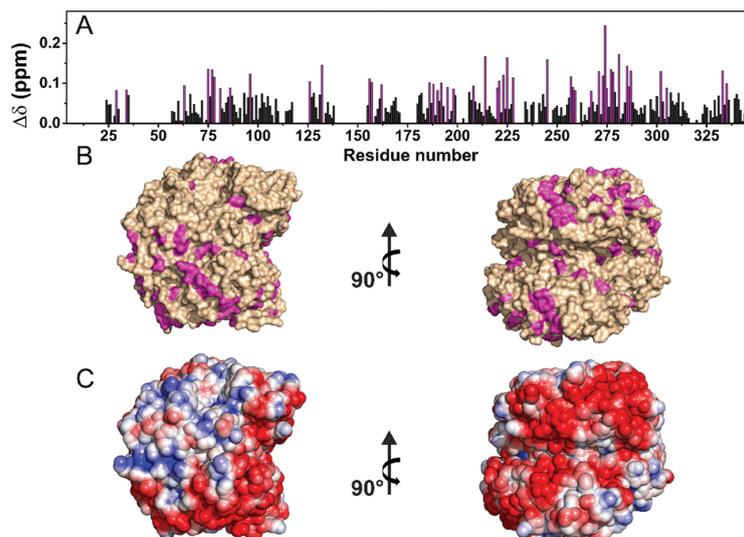


Figure 2. (A) Chemical shift perturbations (CSPs) of ANSII-HA with respect to rehydrated freeze-dried ANSII, evaluated according to the formula $\Delta\delta = \frac{1}{2}\sqrt{(\Delta\delta_{Ca}/2)^2 + (\Delta\delta_N/5)^2}$. The residues experiencing the largest variations have been highlighted in magenta. (B) CSP mapping on the protein surface (PDB code: 3ECA) with the region with the largest perturbation in magenta. (C) Electrostatic potential generated by APBS plugin in PyMOL on 3ECA with blue and red representing the regions of positive and negative electrostatic potential, respectively.

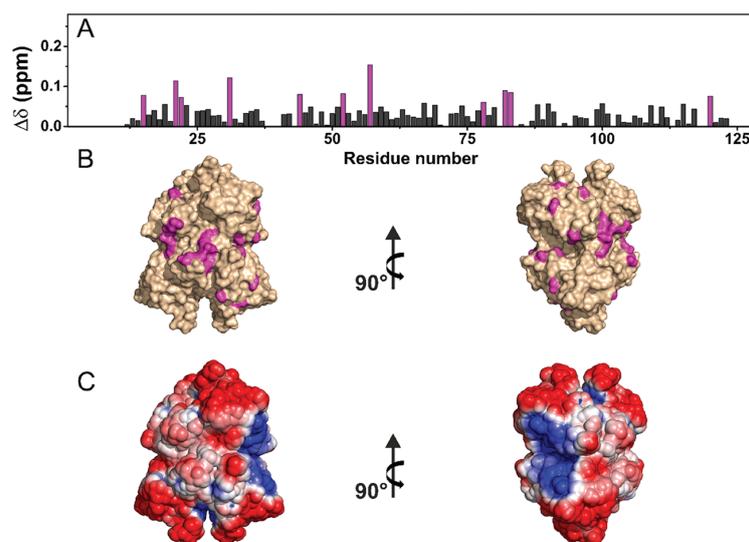


Figure 3. (A) Chemical shift perturbations (CSPs) of TTR-HA with respect to rehydrated freeze-dried TTR, evaluated according to the formula $\Delta\delta = \frac{1}{2}\sqrt{(\Delta\delta_{Ca}/2)^2 + (\Delta\delta_N/5)^2}$. The residues experiencing the largest variations have been highlighted in magenta. (B) CSP mapping on the protein surface (PDB code: 1BMZ) with the region with the largest perturbation in magenta. (C) Electrostatic potential generated by APBS plugin in PyMOL on 1BMZ with blue and red representing the regions of positive and negative electrostatic potential, respectively.

hydrogel encapsulating TTR at a higher frequency speed (20 kHz), to favor the protein sedimentation.

All the spectra were processed with the Bruker TopSpin 3.2 software package and analyzed with the program CARRA.⁷⁷

RESULTS AND DISCUSSION

Analysis of the Preservation of the HOS of the Proteins Encapsulated in the Hyaluronic Acid Hydrogel by SSNMR. The selected proteins (ANSII and TTR)

encapsulated in the hyaluronic acid hydrogel (ANSII-HA and TTR-HA, respectively) were first analyzed by SSNMR. The 1D $\{^1\text{H}\}$ - ^{13}C cross polarization spectra of ANSII-HA and TTR-HA show well-resolved and sharp signals with quality comparable with that of the spectra of the rehydrated freeze-dried materials (Figure S2).

Despite the limited concentration of the embedded proteins in the hydrogel, the 2D amide-carbon alpha (2D ^{15}N ^{13}C NCA) and amide-carbonyl (2D ^{15}N ^{13}C NCO) correlation spectra of ANSII-HA (Figure 1A,B) and TTR-HA (Figure

C

<https://doi.org/10.1021/acs.analchem.1c01850>
Anal. Chem. 2021, 93, 11208–11214

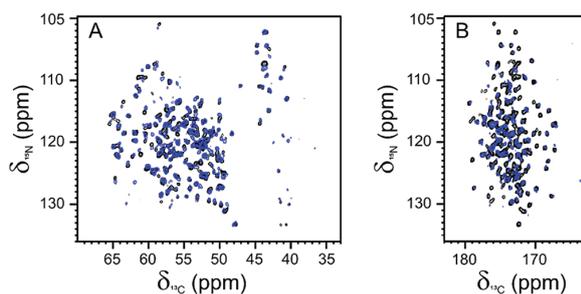


Figure 4. (A) 2D ^{15}N ^{13}C NCA and (B) NCO spectra of ANSII-AG (blue) superimposed with the NCA and NCO of the rehydrated freeze-dried reference protein (black). The spectra were acquired at ~ 290 K, MAS 14 kHz and 800 MHz.

IC,D) are of high quality and comparable, for the number of cross-peaks detected, with those of rehydrated freeze-dried proteins. For both proteins embedded in the hyaluronic acid matrix, the matching of the resonances of the 2D-NMR spectral fingerprints with those of their own reference allows us to assess the preservation of the HOS after encapsulation in the matrix.

The assignment of the 2D ^{15}N ^{13}C NCA and NCO spectra of ANSII-HA and TTR-HA was easily obtained by comparison with the 2D ^{15}N ^{13}C NCA and NCO collected for the rehydrated freeze-dried proteins and also using the information from the 2D ^{13}C - ^{13}C correlation spectrum (dipolar assisted rotational resonance, DARR) acquired for ANSII-HA and TTR-HA. The analyses of the chemical shift perturbation (CSP) of the NCA spectra of the proteins embedded in the hyaluronic acid hydrogels, with respect to the NCA of the corresponding rehydrated freeze-dried references, are reported in Figures 2 and 3. Most CSP values were less than 0.1 ppm for ANSII-HA and even lower for TTR-HA. The analysis of the

CSPs shows that for ANSII-HA, hydrophobic (Ala, Val, Ile, Tyr, and Phe) and neutral polar (Thr, Ser, Asn, and Gln) residues experience the largest effects (Figure 2). Minimal CSPs were observed in TTR-HA protein where the largest effects again involve hydrophobic residues and neutral polar surface patches (Figure 3).

Analysis of the Preservation of the HOS of the Proteins Encapsulated in the Alginate/Gelatin Hydrogel by SSNMR. The same analysis was also performed on the alginate/gelatin hydrogels encapsulating ANSII and TTR, respectively (ANSII-AG and TTR-AG). The 1D $\{^1\text{H}\}$ - ^{13}C cross polarization spectra of ANSII-AG and TTR-AG show the same spreading of the resonances of the corresponding rehydrated freeze-dried analogue. However, in particular for TTR-AG, the signals feature broader lines than in the rehydrated freeze-dried protein (Figure S2).

The NCA and NCO correlation spectra collected for ANSII-AG (Figure 4) are still of high quality and comparable, for the number of cross-peaks detected, with those collected on rehydrated freeze-dried ANSII. On the contrary, for TTR-AG, the fast decay of the NMR signal does not allow us to collect high quality and well-resolved 2D spectra. However, by increasing the spinning rate up to 16 and 20 kHz, the signals become sharper and increase in intensity (Figure S3), indicating a more efficient protein immobilization. Therefore, it was possible to acquire a 2D ^{13}C - ^{13}C correlation spectrum at 20 kHz, which allowed us to assess the folding state of the protein in the hydrogel and, after comparison with that acquired for the rehydrated freeze-dried reference (Figure S4), confirm the preservation of the HOS after encapsulation. The structural analysis of TTR encapsulated in the alginate/gelatin matrix was also attempted using solution NMR. However, all the signals, but the N- and C-termini (Thr3-Ser8; Lys126-Glu127), are broadened beyond detection (Figure S1).

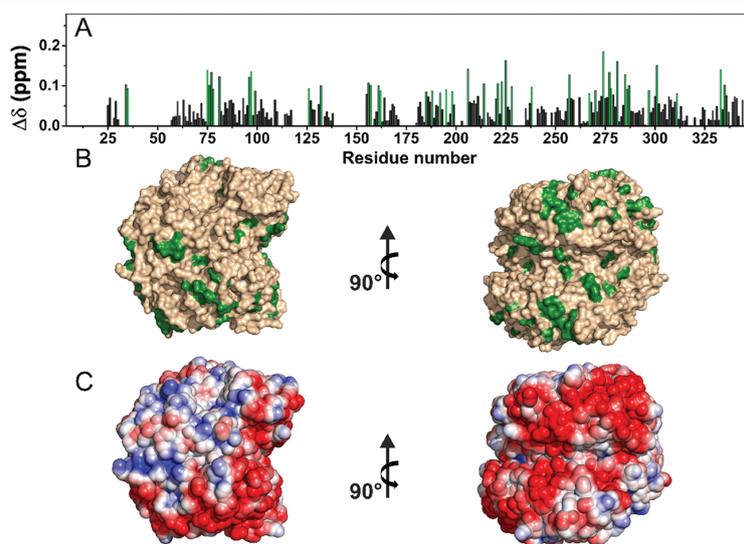


Figure 5. (A) Chemical shift perturbations (CSPs) of ANSII-AG with respect to rehydrated freeze-dried ANSII, evaluated according to the formula $\Delta\delta = \frac{1}{2}\sqrt{(\Delta\delta_{\text{Ca}}/2)^2 + (\Delta\delta_{\text{N}}/5)^2}$. The residues experiencing the largest variations have been highlighted in green. (B) CSP mapping on the protein surface (PDB code: 3ECA) with the region with the largest perturbation in green. (C) Electrostatic potential generated by APBS plugin in PyMOL on 3ECA with blue and red representing the regions of positive and negative electrostatic potential, respectively.

D

<https://doi.org/10.1021/acs.analchem.1c01850>
Anal. Chem. 2021, 93, 11208–11214

The assignment of the ANSII-AG spectra could be easily obtained by comparison with the spectra collected for the rehydrated freeze-dried protein and complemented with the information from the 2D ^{13}C - ^{13}C correlation spectrum acquired for ANSII-AG. The analysis of the CSP of the NCA spectrum of ANSII-AG with respect to the NCA of the rehydrated freeze-dried reference is reported in Figure 5.

The analysis of the CSPs shows that also for ANSII-AG, hydrophobic (Ala, Val, Ile, Tyr, and Phe) and neutral polar (Thr, Ser, Ans, and Gln) residues experience the largest effects. In particular, many threonine residues are affected by significant CSP, thus suggesting a possible interaction of these surface residues with the hydroxyl groups of alginate in the hydrogel.

Collectively, the good superimposition of the spectra and the small CSPs observed for the two proteins prove the preservation of their native HOS, thus providing the first fundamental information on the investigated biomaterial. Additional information on protein–matrix interactions is obtained from the line broadening of the signals in the spectra. For TTR, the large line broadening, its dependence from the spinning rate, and the small CSPs suggest a weaker protein–matrix interaction with respect to ANSII protein, although the different molecular weights may also play a role. The different behavior is probably related to the different sizes of the proteins and to the physical–chemical properties of the surface due to the different amino acid compositions. In this respect, the observation that hydrophobic and polar neutral amino acids on the protein surface experience the largest effects provides a way to design possible chemical modifications of the matrix in order to tune the protein–matrix interactions and the properties of the resulting biomaterial.^{78–82}

CONCLUSIONS

In summary, we demonstrate that 2D-SSNMR spectra can be exploited to assess the preservation of HOS of proteins when embedded in matrices used for 3D bioprinting and drug release. This analytical method can be integrated in the pipeline for the development of new composite hydrogels bearing biotherapeutics. In particular, when the assignment is available, the analysis of the residues experiencing chemical shift variations can provide information for a quality by design approach of these innovative biomaterials.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.1c01850>.

Protocols for the expression and purification of isotopically enriched *E. coli* asparaginase-II (ANSII) and human transthyretin (TTR) and figures showing additional spectra of *E. coli* asparaginase-II (ANSII) and human transthyretin (TTR) (PDF)

AUTHOR INFORMATION

Corresponding Author

Marco Fragai – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy; Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy;*

✉ orcid.org/0000-0002-8440-1690; Phone: +39 055 4574261; Email: fragai@cerm.unifi.it

Authors

Domenico Rizzo – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy; Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy*

Linda Cerofolini – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy*

Anna Pérez-Ràfols – *Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy; Giotto Biotech, S.R.L, Sesto Fiorentino, Florence 50019, Italy*

Stefano Giuntini – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy; Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy*

Fabio Baroni – *Analytical Development Biotech Department, Merck Serono S.p.a, Merck KGaA, Guidonia, Rome 00012, Italy*

Enrico Ravera – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy; Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy;* orcid.org/0000-0001-7708-9208

Claudio Luchinat – *Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Sesto Fiorentino 50019, Italy; Department of Chemistry “Ugo Schiff”, University of Florence, Sesto Fiorentino 50019, Italy*

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.1c01850>

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been supported by Regione Toscana (CERM-TT and BioEnable), Fondazione Cassa di Risparmio di Firenze, the Italian Ministero dell’Istruzione, dell’Università e della Ricerca through PRIN 2017A2KEPL, the “Progetto Dipartimenti di Eccellenza 2018-2022” to the Department of Chemistry “Ugo Schiff” of the University of Florence, and the Recombinant Proteins JOYNLAB laboratory. The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy centre. We acknowledge H2020 -INFRAIA iNEXT-Discovery - Structural Biology Research Infrastructures for Translational Research and Discovery (contract n° 871037), EOSC-Life “Providing an open collaborative space for digital biology in Europe” (H2020, contract n° 824087), and “RNAct” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITN-

2018 (contract n° 813239). The authors also acknowledge Mestrelab Research for providing Mnova software and Bruker BioSpin for AssureNMR software.

REFERENCES

- (1) Gaharwar, A. K.; Singh, I.; Khademhosseini, A. *Nat. Rev. Mater.* **2020**, *5*, 686–705.
- (2) Gu, L.; Mooney, D. J. *Nat. Rev. Cancer* **2016**, *16*, 56–66.
- (3) Suri, S.; Han, L.-H.; Zhang, W.; Singh, A.; Chen, S.; Schmidt, C. E. *Biomed. Microdevices* **2011**, *13*, 983–993.
- (4) Noh, I.; Kim, N.; Tran, H. N.; Lee, J.; Lee, C. *Biomater. Res.* **2019**, *23*, 3.
- (5) Wei, Y.-T.; He, Y.; Xu, C.-L.; Wang, Y.; Liu, B.-F.; Wang, X.-M.; Sun, X.-D.; Cui, F.-Z.; Xu, Q.-Y. *J. Biomed. Mater. Res. B Appl. Biomater.* **2010**, *95B*, 110–117.
- (6) Antich, C.; de Vicente, J.; Jiménez, G.; Chocarro, C.; Carrillo, E.; Montañez, E.; Gálvez-Martín, P.; Marchal, J. A. *Acta Biomater.* **2020**, *106*, 114–123.
- (7) Skardal, A.; Zhang, J.; Prestwich, G. D. *Biomaterials* **2010**, *31*, 6173–6181.
- (8) Chung, J. H. Y.; Naficy, S.; Yue, Z.; Kapsa, R.; Quigley, A.; Moulton, S. E.; Wallace, G. G. *Biomater. Sci.* **2013**, *1*, 763–773.
- (9) Wu, Z.; Su, X.; Xu, Y.; Kong, B.; Sun, W.; Mi, S. *Sci. Rep.* **2016**, *6*, 24474.
- (10) Poldervaart, M. T.; Wang, H.; van der Stok, J.; Weinans, H.; Leeuwenburgh, S. C. G.; Öner, F. C.; Dhert, W. J. A.; Alblas, J. *PLoS One* **2013**, *8*, No. e72610.
- (11) Duan, B.; Hockaday, L. A.; Kang, K. H.; Butcher, J. T. *J. Biomed. Mater. Res. A* **2013**, *101A*, 1255–1264.
- (12) Giuseppe, M. D.; Law, N.; Webb, B.; Macrae, R. A.; Liew, L. J.; Sercombe, T. B.; Dilley, R. J.; Doyle, B. J. *J. Mech. Behav. Biomed. Mater.* **2018**, *79*, 150–157.
- (13) Jia, W.; Gungor-Ozkerim, P. S.; Zhang, Y. S.; Yue, K.; Zhu, K.; Liu, W.; Pi, Q.; Byambaa, B.; Dokmeci, M. R.; Shin, S. R.; Khademhosseini, A. *Biomaterials* **2016**, *106*, 58–68.
- (14) Buwalda, S. J.; Vermonden, T.; Hennink, W. E. *Biomacromolecules* **2017**, *18*, 316–330.
- (15) Chen, J.; Ouyang, J.; Chen, Q.; Deng, C.; Meng, F.; Zhang, J.; Cheng, R.; Lan, Q.; Zhong, Z. *ACS Appl. Mater. Interfaces* **2017**, *9*, 24140–24147.
- (16) Zhu, Q.; Chen, X.; Xu, X.; Zhang, Y.; Zhang, C.; Mo, R. *Adv. Funct. Mater.* **2018**, *28*, No. 1707371.
- (17) Wawrzyńska, E.; Kubies, D. *Physiol. Res.* **2018**, S319–S334.
- (18) Iqbal, S.; Blenner, M.; Alexander-Bryant, A.; Larsen, J. *Biomacromolecules* **2020**, *21*, 1327–1350.
- (19) Shigemitsu, H.; Kubota, R.; Nakamura, K.; Matsuzaki, T.; Minami, S.; Aoyama, T.; Urayama, K.; Hamachi, I. *Nat. Commun.* **2020**, *11*, 3859.
- (20) Lau, C. M. L.; Jahanmir, G.; Yu, Y.; Chau, Y. *J. Controlled Release* **2021**, *335*, 75–85.
- (21) Tae, H.; Lee, S.; Ki, C. S. *J. Ind. Eng. Chem.* **2019**, *75*, 69–76.
- (22) Chang, D.; Park, K.; Famili, A. *Drug Discovery Today* **2019**, *24*, 1470–1482.
- (23) Ziegler, C. E.; Graf, M.; Beck, S.; Goepferich, A. M. *Eur. Polym. J.* **2021**, *147*, No. 110286.
- (24) Gombotz, W. R.; Wee, S. F. *Adv. Drug Delivery Rev.* **2012**, *64*, 194–205.
- (25) Vermonden, T.; Censi, R.; Hennink, W. E. *Chem. Rev.* **2012**, *112*, 2853–2888.
- (26) Kumar, S.; Koh, J. *Int. J. Mol. Sci.* **2012**, *13*, 6102–6116.
- (27) Dong, A.; Jones, L. S.; Kerwin, B. A.; Krishnan, S.; Carpenter, J. F. *Anal. Biochem.* **2006**, *351*, 282–289.
- (28) Kirkitadze, M.; Sinha, A.; Hu, J.; Williams, W.; Cates, G. *Procedia Vaccinol.* **2009**, *1*, 135–139.
- (29) Arbogast, L. W.; Brinson, R. G.; Marino, J. P. *Anal. Chem.* **2015**, *87*, 3556–3561.
- (30) Brinson, R. G.; Marino, J. P.; Delaglio, F.; Arbogast, L. W.; Evans, R. M.; Kearsley, A.; Gingras, G.; Ghasriani, H.; Aubin, Y.; Pierens, G. K.; Jia, X.; Mobli, M.; Grant, H. G.; Keizer, D. W.; Schweimer, K.; Stähle, J.; Widmalm, G.; Zartler, E. R.; Lawrence, C. W.; Reardon, P. N.; Cort, J. R.; Xu, P.; Ni, F.; Yanaka, S.; Kato, K.; Parnham, S. R.; Tsao, D.; Blomgren, A.; Rundlöf, T.; Trieloff, N.; Schmieder, P.; Ross, A.; Skidmore, K.; Chen, K.; Keire, D.; Freedberg, D. I.; Suter-Stahel, T.; Wider, G.; Ilc, G.; Plavec, J.; Bradley, S. A.; Baldissieri, D. M.; Sforça, M. L.; de Zeri, A. C. M.; Wei, J. Y.; Szabo, C. M.; Amezcuca, C. A.; Jordan, J. B.; Wikström, M. *mAbs* **2019**, *11*, 94–105.
- (31) Jones, L. M.; Zhang, H.; Cui, W.; Kumar, S.; Sperry, J. B.; Carroll, J. A.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 835–845.
- (32) Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F. *Anal. Chem.* **2014**, *86*, 2657–2664.
- (33) Ehkirch, A.; Hernandez-Alba, O.; Colas, O.; Beck, A.; Guilleme, D.; Cianféroni, S. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2018**, *1086*, 176–183.
- (34) Brinson, R. G.; Ghasriani, H.; Hodgson, D. J.; Adams, K. M.; McEwen, I.; Freedberg, D. I.; Chen, K.; Keire, D. A.; Aubin, Y.; Marino, J. P. *J. Pharm. Biomed. Anal.* **2017**, *141*, 229–233.
- (35) Ghasriani, H.; Hodgson, D. J.; Brinson, R. G.; McEwen, I.; Buhse, L. F.; Kozłowski, S.; Marino, J. P.; Aubin, Y.; Keire, D. A. *Nat. Biotechnol.* **2016**, *34*, 139–141.
- (36) Arbogast, L. W.; Brinson, R. G.; Formolo, T.; Hoopes, J. T.; Marino, J. P. *Pharm. Res.* **2016**, *33*, 462–475.
- (37) Arbogast, L. W.; Delaglio, F.; Tolman, J. R.; Marino, J. P. *J. Biomol. NMR* **2018**, *72*, 149–161.
- (38) Arbogast, L. W.; Delaglio, F.; Brinson, R. G.; Marino, J. P. *Curr. Protoc. Protein Sci.* **2020**, *100*, No. e105.
- (39) Wu, K.; Luo, J.; Zeng, Q.; Dong, X.; Chen, J.; Zhan, C.; Chen, Z.; Lin, Y. *Anal. Chem.* **2021**, *93*, 1377–1382.
- (40) Pastore, A.; Salvadori, S.; Temussi, P. A. *J. Pept. Sci.* **2007**, *13*, 342–347.
- (41) Sass, H.-J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. *J. Biomol. NMR* **2000**, *18*, 303–309.
- (42) Barrientos, L. G.; Dolan, C.; Gronenborn, A. M. *J. Biomol. NMR* **2000**, *16*, 329–337.
- (43) Tycko, R.; Blanco, F. J.; Ishii, Y. *J. Am. Chem. Soc.* **2000**, *122*, 9340–9341.
- (44) Lecoq, L.; Fogeron, M.-L.; Meier, B. H.; Nassal, M.; Böckmann, A. Solid-State NMR for Studying the Structure and Dynamics of Viral Assemblies. *Viruses* **2020**, *12* (), DOI: 10.3390/v12101069.
- (45) Wiegand, T.; Lacabanne, D.; Torosyan, A.; Boudet, J.; Cadalbert, R.; Allain, F. H.-T.; Meier, B. H.; Böckmann, A. *Front. Mol. Biosci.* **2020**, *7*, 17.
- (46) Hassan, A.; Quinn, C. M.; Struppe, J.; Sergeev, I. V.; Zhang, C.; Guo, C.; Runge, B.; Theint, T.; Dao, H. H.; Jaroniec, C. P.; Berbon, M.; Lends, A.; Habenstein, B.; Loquet, A.; Kuemmerle, R.; Perrone, B.; Gronenborn, A. M.; Polenova, T. *J. Magn. Reson.* **2020**, *311*, No. 106680.
- (47) Lu, M.; Russell, R. W.; Bryer, A. J.; Quinn, C. M.; Hou, G.; Zhang, H.; Schwieters, C. D.; Perilla, J. R.; Gronenborn, A. M.; Polenova, T. *Nat. Struct. Mol. Biol.* **2020**, *27*, 863–869.
- (48) Eddy, M. T.; Yu, T.-Y.; Wagner, G.; Griffin, R. G. *J. Biomol. NMR* **2019**, *73*, 451–460.
- (49) Gupta, R.; Zhang, H.; Lu, M.; Hou, G.; Caporini, M.; Rosay, M.; Maas, W.; Struppe, J.; Ahn, J.; Byeon, I.-J. L.; Oschkinat, H.; Jaudzems, K.; Barbet-Massin, E.; Emsley, L.; Pintacuda, G.; Lesage, A.; Gronenborn, A. M.; Polenova, T. *J. Phys. Chem. B* **2019**, *123*, 5048–5058.
- (50) le Paige, U. B.; Xiang, S.; Hendrix, M. M. R. M.; Zhang, Y.; Folkers, G. E.; Weingarth, M.; Bonvin, A. M. J. J.; Kutateladze, T. G.; Voets, I. K.; Baldus, M.; van Ingen, H. *Magn. Reson.* **2021**, *2*, 187–202.
- (51) Mroue, K. H.; MacKinnon, N.; Xu, J.; Zhu, P.; McNerny, E.; Kohn, D. H.; Morris, M. D.; Ramamoorthy, A. *J. Phys. Chem. B* **2012**, *116*, 11656–11661.

- (52) Azais, T.; Von Euw, S.; Ajili, W.; Auzoux-Bordenave, S.; Bertani, P.; Gajan, D.; Emsley, L.; Nassif, N.; Lesage, A. *Solid State Nucl. Magn. Reson.* **2019**, *102*, 2–11.
- (53) Cerofolini, L.; Giuntini, S.; Louka, A.; Ravera, E.; Fragai, M.; Luchinat, C. *J. Phys. Chem. B* **2017**, *121*, 8094–8101.
- (54) Louka, A.; Matlahov, I.; Giuntini, S.; Cerofolini, L.; Cavallo, A.; Pillozzi, S.; Ravera, E.; Fragai, M.; Arcangeli, A.; Ramamoorthy, A.; Goobes, G.; Luchinat, C. *Phys. Chem. Chem. Phys.* **2018**, *20*, 12719–12726.
- (55) Ravera, E.; Cerofolini, L.; Martelli, T.; Louka, A.; Fragai, M.; Luchinat, C. *Sci. Rep.* **2016**, *6*, 27851.
- (56) Martelli, T.; Ravera, E.; Louka, A.; Cerofolini, L.; Hafner, M.; Fragai, M.; Becker, C. F. W.; Luchinat, C. *Chemistry* **2016**, *22*, 425–432.
- (57) Fragai, M.; Luchinat, C.; Martelli, T.; Ravera, E.; Sagi, I.; Solomonov, I.; Udi, Y. *Chem. Commun.* **2014**, *50*, 421–423.
- (58) Ravera, E.; Schubeis, T.; Martelli, T.; Fragai, M.; Parigi, G.; Luchinat, C. *J. Magn. Reson.* **2015**, *253*, 60–70.
- (59) Cerofolini, L.; Giuntini, S.; Carlon, A.; Ravera, E.; Calderone, V.; Fragai, M.; Parigi, G.; Luchinat, C. *Chem. – Eur. J.* **2019**, *25*, 1984–1991.
- (60) Cerofolini, L.; Fragai, M.; Ravera, E.; Diebolder, C. A.; Renault, L.; Calderone, V. *Biomolecules* **2019**, *9*, 370.
- (61) Giuntini, S.; Balducci, E.; Cerofolini, L.; Ravera, E.; Fragai, M.; Berti, F.; Luchinat, C. *Angew. Chem., Int. Ed.* **2017**, *56*, 14997–15001.
- (62) Ravera, E.; Ciambellotti, S.; Cerofolini, L.; Martelli, T.; Kozyreva, T.; Bernacchioni, C.; Giuntini, S.; Fragai, M.; Turano, P.; Luchinat, C. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 2446–2449.
- (63) Giuntini, S.; Cerofolini, L.; Ravera, E.; Fragai, M.; Luchinat, C. *Sci. Rep.* **2017**, *7*, 17934.
- (64) Cerofolini, L.; Giuntini, S.; Ravera, E.; Luchinat, C.; Berti, F.; Fragai, M. *npj Vaccines* **2019**, *4*, 20.
- (65) Viger-Gravel, J.; Paruzzo, F. M.; Cazaux, C.; Jabbour, R.; Leleu, A.; Canini, F.; Florian, P.; Ronzon, F.; Gajan, D.; Lesage, A. *Chemistry* **2020**, *26*, 8976–8982.
- (66) Jaudzems, K.; Kirsteina, A.; Schubeis, T.; Casano, G.; Ouari, O.; Bogans, J.; Kazaks, A.; Tars, K.; Lesage, A.; Pintacuda, G. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 12847.
- (67) Wang, H.; Wang, L.; Li, C.; Wuxiao, Z.; Chen, G.; Luo, W.; Lu, Y. *Oncologist* **2020**, *25*, e1725–e1731.
- (68) Herbert, J.; Wilcox, J. N.; Pham, K.-T. C.; Fremeau, R. T.; Zeviani, M.; Dwork, A.; Soprano, D. R.; Makover, A.; Goodman, D. S.; Zimmerman, E. A.; Roberts, J. L.; Schon, E. A. *Neurology* **1986**, *36*, 900–900.
- (69) Hamilton, J. A.; Benson, M. D. *Cell. Mol. Life Sci.* **2001**, *58*, 1491–1521.
- (70) Connors, L. H.; Lim, A.; Prokaeva, T.; Roskens, V. A.; Costello, C. E. *Amyloid* **2003**, *10*, 160–184.
- (71) Walker, K. W.; Foltz, I. N.; Wang, T.; Salimi-Moosavi, H.; Bailis, J. M.; Lee, F.; An, P.; Smith, S.; Bruno, R.; Wang, Z. *J. Biol. Chem.* **2020**, *295*, 10446–10455.
- (72) Fragai, M.; Luchinat, C.; Parigi, G.; Ravera, E. *J. Biomol. NMR* **2013**, *57*, 155–166.
- (73) Sarker, B.; Papageorgiou, D. G.; Silva, R.; Zehnder, T.; Gul-E-Noor, F.; Bertmer, M.; Kaschta, J.; Chrissafis, K.; Detsch, R.; Boccaccini, A. R. *J. Mater. Chem. B* **2014**, *2*, 1470–1482.
- (74) Wang, Q.-Q.; Liu, Y.; Zhang, C.-J.; Zhang, C.; Zhu, P. *Mater. Sci. Eng. C Mater. Biol. Appl.* **2019**, *99*, 1469–1476.
- (75) Schuetz, A.; Wasmer, C.; Habenstein, B.; Verel, R.; Greenwald, J.; Riek, R.; Böckmann, A.; Meier, B. H. *ChemBioChem* **2010**, *11*, 1543–1551.
- (76) Lu, X.; Guo, C.; Hou, G.; Polenova, T. *J. Biomol. NMR* **2015**, *61*, 7–20.
- (77) Keller, R. The Computer Aided Resonance Assignment Tutorial (CARA); *The CARA/Lua Programmers Manual*. DATONAL AG.; CANTINA Verlag: Goldau, Switzerland, 2004.
- (78) Leach, J. B.; Schmidt, C. E. *Biomaterials* **2005**, *26*, 125–135.
- (79) Jia, J.; Richards, D. J.; Pollard, S.; Tan, Y.; Rodriguez, J.; Visconti, R. P.; Trusk, T. C.; Yost, M. J.; Yao, H.; Markwald, R. R.; Mei, Y. *Acta Biomater.* **2014**, *10*, 4323–4331.
- (80) Ouyang, L.; Yao, R.; Zhao, Y.; Sun, W. *Biofabrication* **2016**, *8*, No. 035020.
- (81) Gao, T.; Gillispie, G. J.; Copus, J. S.; Pr, A. K.; Seol, Y.-J.; Atala, A.; Yoo, J. J.; Lee, S. J. *Biofabrication* **2018**, *10*, No. 034106.
- (82) Liao, Y.-H.; Jones, S. A.; Forbes, B.; Martin, G. P.; Brown, M. B. *Drug Delivery* **2005**, *12*, 327–342.

4.5.2 Supplementary Information

*Evaluation of the Higher Order Structure of Biotherapeutics
Embedded in Hydrogels for Bioprinting and Drug Release*

Supporting Information

Evaluation of the Higher Order Structure of biotherapeutics embedded in hydrogels for bioprinting and drug release

Domenico Rizzo^{1,2}, Linda Cerofolini¹, Anna Pérez-Ràfols^{2,3}, Stefano Giuntini^{1,2}, Fabio Baroni⁴, Enrico Ravera^{1,2}, Claudio Luchinat^{1,2}, Marco Fragai^{*1,2}

1. Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP) Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy
2. Department of Chemistry “Ugo Schiff”, University of Florence, Via della Lastruccia 3, 50019, Sesto Fiorentino, Italy
3. Giotto Biotech, S.R.L, Via Madonna del piano 6, 50019, Sesto Fiorentino (FI), Italy
4. Analytical Development Biotech Department, Merck Serono S.p.a, Guidonia, RM, Italy; an affiliate of Merck KGaA

*Corresponding author; Phone: +39 055 4574261; E-mail: fragai@cerm.unifi.it

Table of Contents

Expression and purification of uniformly isotopically enriched ANSII [U- ¹³ C- ¹⁵ N]	S2
Expression and purification of uniformly isotopically enriched TTR [U- ¹³ C- ¹⁵ N]	S2
Figure S1	S4
Figure S2	S5
Figure S3	S6
Figure S4	S7

Expression and purification of uniformly isotopically enriched ANSII [U - ^{13}C - ^{15}N]. Escherichia coli C41(DE3) cells were transformed with pET-21a(+) plasmid encoding ANSII gene. The cells were cultured in ^{13}C -, ^{15}N -enriched minimal medium (M9) containing 0.1 mg/mL of ampicillin, and grown at 310 K until OD_{600} reached 0.6–0.8. Then, the cells were induced with 1 mM isopropyl β -D-1-thiogalactopyranoside and further grown at 310 K overnight. Finally, they were harvested by centrifugation at 6500 rpm (JA-10 Beckman Coulter) for 15 min at 277 K. The pellet obtained from 1 liter of culture was suspended in 60 mL of 10 mM Tris-HCl buffer, at pH 8.0, with 15 mM EDTA, 20% sucrose and incubated at 277 K for 20 min, under magnetic stirring. The suspension was centrifuged at 10,000 rpm (F15-6x100y Thermo Scientific) for 30 min, and the supernatant discarded. The recovered pellet was re-suspended in H_2O milli-Q and newly incubated with the Tris-HCl buffer solution at 277 K for 20 min under magnetic stirring. Again, the suspension was centrifuged at 10,000 rpm (F15-6x100y Thermo Scientific) for 30 min. The pellet was discarded, whereas the supernatant was treated with ammonium sulfate. Still under magnetic stirring, aliquots of solid ammonium sulfate were added up to 50% saturation. Then, the precipitate was removed by centrifugation, and ammonium sulfate added again up to 90% saturation to trigger the precipitation of ANSII, which was recovered by centrifugation. The precipitated ANSII was re-dissolved in a minimal amount of 20 mM Tris-HCl buffer at pH 8.6 and dialyzed extensively against the same buffer. ANSII was purified by anionic-exchange chromatography using a HiPrep Q FF 16/10 column (GE Healthcare Life Science). The protein was eluted in 20 mM Tris-HCl buffer at pH 8.6 with a linear 0–1 M NaCl gradient. Fractions containing pure ANSII were identified by Coomassie staining SDS-PAGE gels, then joined and dialyzed extensively against 0.5 mM Tris-HCl buffer at pH 7.5. Aliquots of 0.5 mL, each containing 1 mg of protein, were freeze-dried to be used for SSNMR analysis.

Expression and purification of uniformly isotopically enriched TTR [U - ^{13}C - ^{15}N]. Escherichia coli BL21(DE3) RIPL pLysS cells were transformed with pET-28a(+) plasmid encoding TTR gene. The cells were cultured in ^{13}C -, ^{15}N -enriched minimal medium (M9) containing 0.1 mg/mL of kanamycin, grown at 310 K, until OD_{600} reached 0.6–0.8 and then induced with 1 mM isopropyl β -D-1-thiogalactopyranoside. The cells were further grown at 310 K overnight and then harvested by centrifugation at 6500 rpm (JA-10 Beckman Coulter) for 15 min at 277 K. The pellet was suspended in 20 mM Tris-HCl at pH 8.5 (60 mL per liter of culture) and sonicated at 277 K for 40 min. The suspension was centrifuged at 40,000 rpm (F15-6x100y Thermo Scientific) for 40 min and the pellet discarded. The protein was purified by anionic-exchange chromatography using a HiPrep Q FF 16/10 column (GE Healthcare Life Science). The protein was eluted in 20 mM Tris-HCl buffer at pH 8.6 with a linear 0–1 M NaCl gradient. Fractions containing pure TTR were identified by Coomassie

staining SDS-PAGE gels, then joined and purified by Size Exclusion Chromatography using HiLoad Superdex 26/60 75pg in 50 mM phosphate buffer at pH 7.5; then dialyzed extensively against 2 mM Tris-HCl buffer at pH 7.5. Aliquots of 1 mL, each containing 6 mg of protein, were freeze-dried to be used for SSNMR analysis.

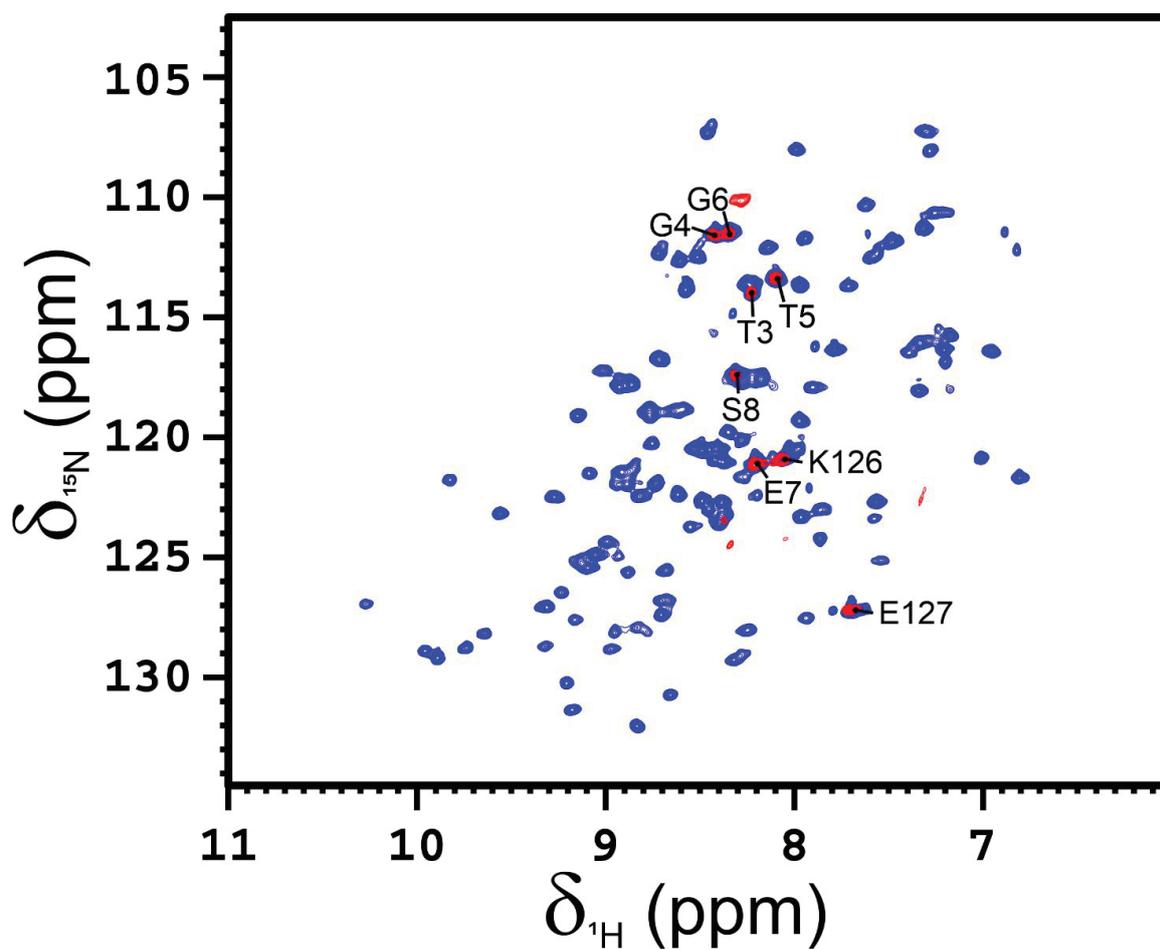


Figure S1. 2D ^1H - ^{15}N TROSY-HSQC spectrum of TTR (blue) collected on a solution of the protein at the concentration of 100 μM superimposed with the same spectrum collected on TTR (100 μM) encapsulated in alginate/gelatine hydrogel (red). The spectra were recorded on a 950 MHz spectrometer at 310 K.

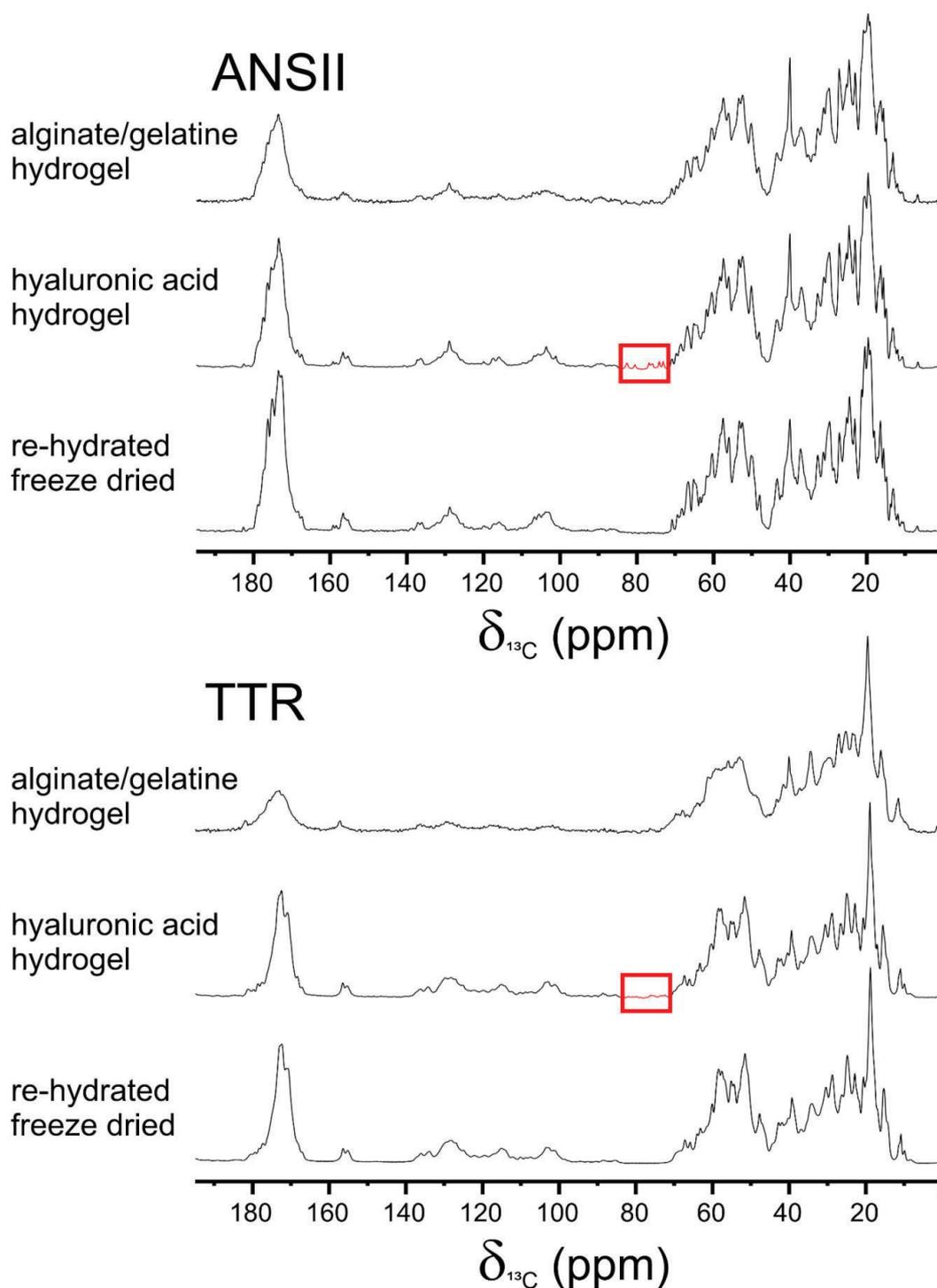


Figure S2. 1D $\{^1\text{H}\}^{13}\text{C}$ cross-polarization spectra acquired on the re-hydrated freeze-dried proteins, on the proteins embedded in the hyaluronic acid hydrogels, and in the alginate/gelatine hydrogels (ANSII, top, and TTR, bottom). The signals of the hyaluronic acid are highlighted by a red box. The spectra were acquired at ~ 290 K, MAS 14 kHz and 800 MHz (number of scans: 512, recycle delay: 3 sec).

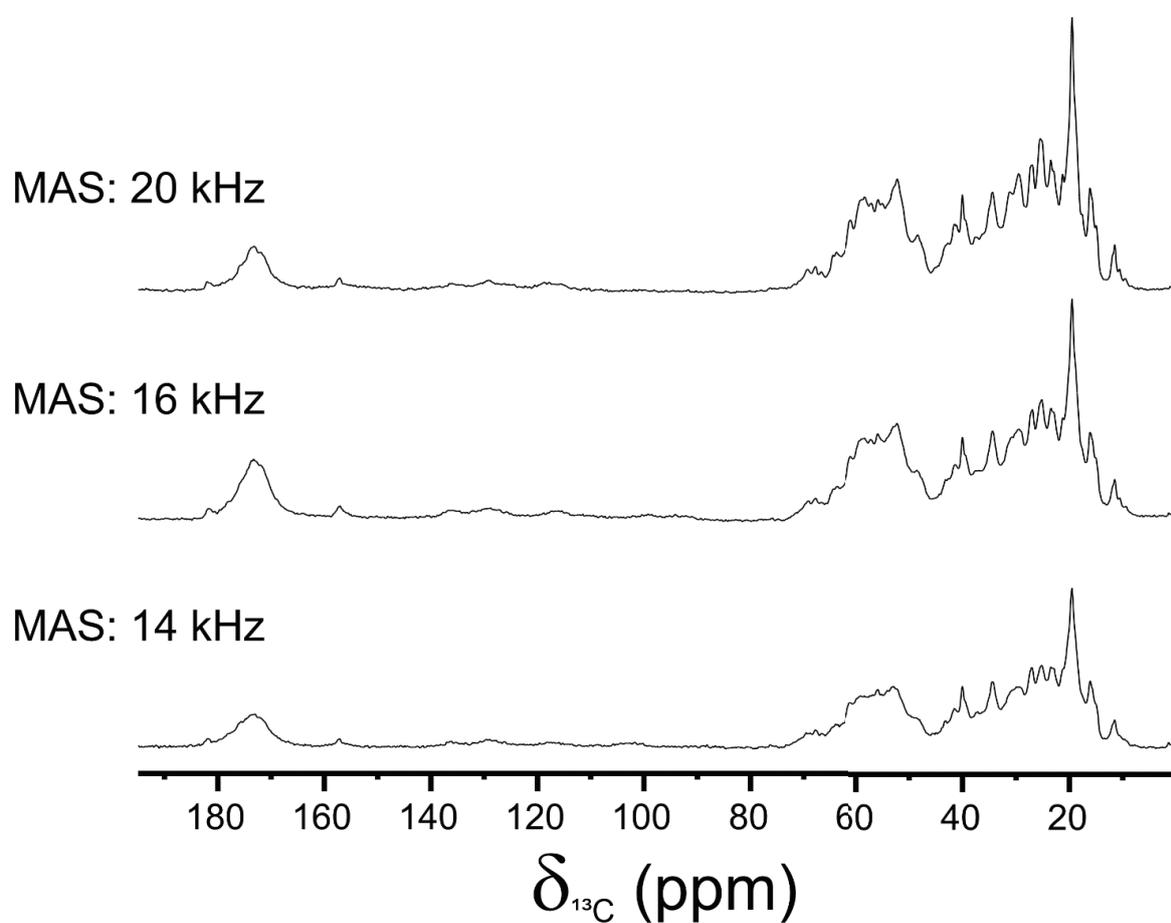


Figure S3. 1D $\{^1\text{H}\}^{13}\text{C}$ cross-polarization spectra acquired on the TTR-AG at the different MAS speeds indicated in the figure. The spectra were acquired at ~ 290 K and 800 MHz (number of scans: 512, recycle delay: 3 sec).

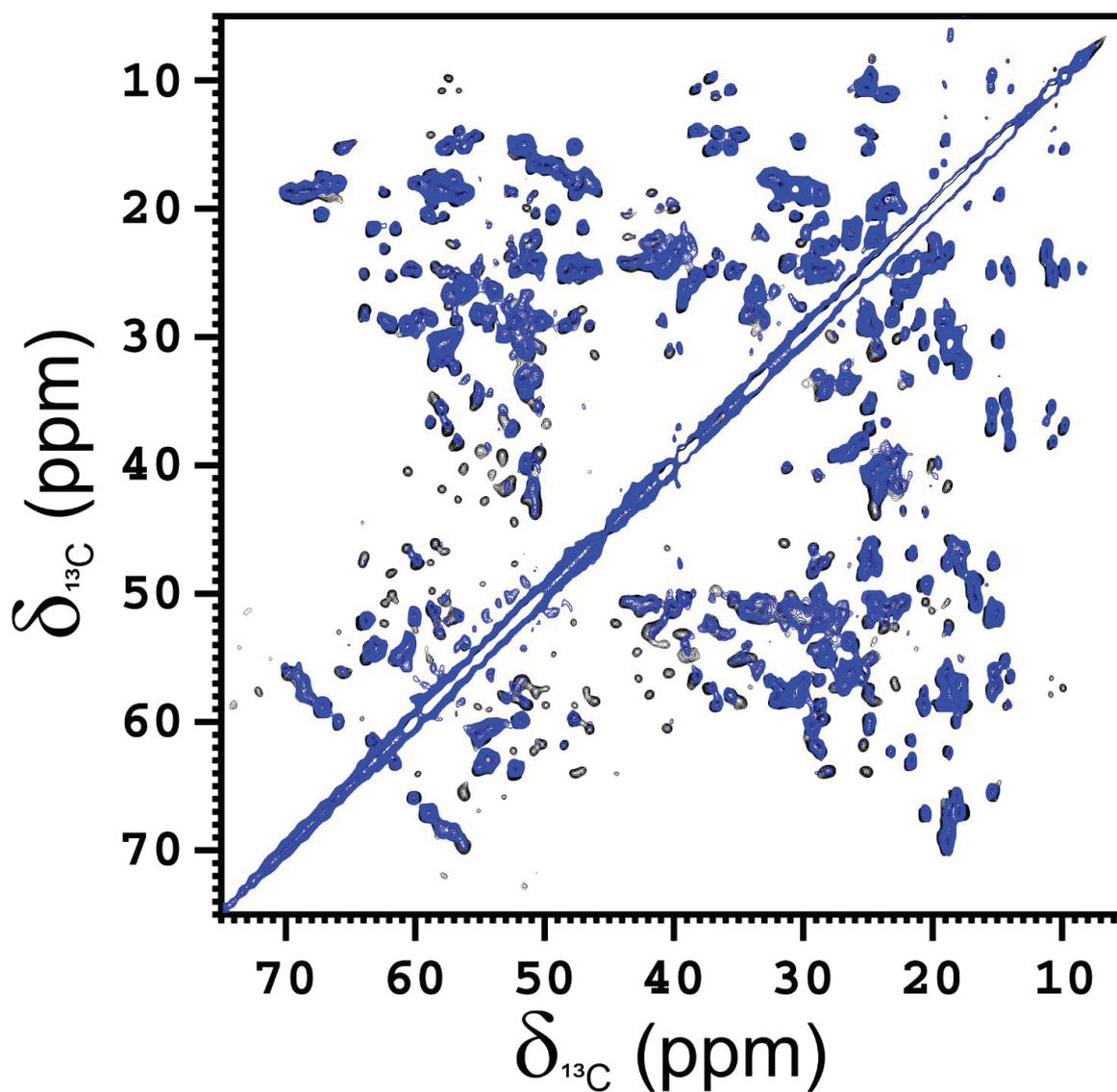


Figure S4. Aliphatic region of the 2D ^{13}C - ^{13}C CORDxy4 recorded on TTR-AG hydrogel (blue) superimposed with the CORDxy4 recorded on the re-hydrated freeze-dried TTR (black). The spectra were acquired at ~ 290 K, MAS 20 kHz and 800 MHz.

4.6 ARTICLE 7:

4.6.1 Transthyretin shows new possibilities for the structure-based molecular design of protein-drug conjugates

Linda Cerofolini, Kristian Vasa, Maria Salobehaj, Giulia Cappelli, Alice Bonciani, Giulia Licciardi, **Anna Pérez-Ràfols**, Luis David Padilla Cortes, Sabrina Antonacci, Domenico Rizzo, Enrico Ravera, Vito Calderone, Giacomo Parigi, Claudio Luchinat, Stefano Menichetti, Marco Fragai

In preparation

Abstract

Protein-drug conjugates represent one of the new frontiers for drug design and promising tools for cancer therapy. Several of such pharmaceutical products are already in clinical use and rely on cytotoxic organic molecules covalently bound to the carrier proteins or interacting with them by van der Waals force and hydrogen-bond mediated interactions. A physiological protein, human transthyretin has been already considered as carrier protein to deliver cytotoxic agents to cancer cells. Here transthyretin is used to show a new structure-based strategy for the design a non-covalent high affinity cytotoxic molecule able to generate a stable protein-drug conjugate. The approach is based on the integration of different structural biophysical methodologies and shows the pivotal contribution of NMR for the development of this new class of challenging biological drugs.

In this project, my contribution has been the expression and characterization of the U-¹³C-¹⁵N Transthyretin and I have been involved with the experimental part of the titration experiments with TTR/Taf-Ptx complexes in solution NMR and lyophilization of the sample complexes.

Introduction.

The development of a suitable drug delivery system is a crucial step in drug design, and an extended half-life together with an efficient targeting capability are often required to achieve a high therapeutic efficacy.^[1,2] The conjugation of drugs to natural or synthetic large polymers is often used to increase the half-life by reducing the renal excretion.^[3] When the polymer is conjugated to a protein, or is a large protein itself that recognizes a specific receptor, an efficient targeting can be also achieved. Human serum albumin is currently used as drug carrier and component of nanoparticles to deliver cytotoxic molecules to cancer cells.^[4,5] Albumin has also been used in genetic fusion proteins as a carrier for therapeutic peptides to prevent the fast clearance, the proteolytic degradation, and to improve the solubility.^[6-8]

Human transthyretin is a physiological protein present in blood plasma and cerebrospinal fluid where it carries the holo-retinol binding protein and the thyroxine T₄ hormone. The functional protein is an assembly of four identical subunits, forming a tetramer (dimer of dimers) of 55 kDa with D₂ symmetry. Mutations in the gene coding for the transthyretin protein decrease the stability of the assembly leading to the dissociation of the tetramer into monomers that partially unfold and polymerize to form amyloid fibrils.^[9] Extracellular accumulation of transthyretin amyloid fibrils in different tissues and organs, including eyes, kidney, heart and peripheral nerves, leads to severe disorders, and ultimately to fatal multiorgan failure. A recent therapeutic approach to treat the familial amyloid polyneuropathy relies on small organic molecules that fit the two channels where the thyroxine T₄ hormone binds the tetramer.^[10,11] One of these molecules, 2-(3,5-dichlorophenyl)benzo[d]oxazole-6-carboxylic acid (Tafamidis hereafter), binds transthyretin in a negative cooperative manner but still with low nanomolar affinity, thus preventing its monomerization.^[12] The conjugation of hydrophobic cytotoxic drugs to high affinity ligands of transthyretin has been already extensively developed by Alhamadsheh and co-workers to generate drug-protein conjugates with an improved selectivity against the cancer cells.^[13-15]

There is an increasing interest in rational design of protein-drug conjugates to optimize efficacy, pharmacokinetics and stability *in vivo*, and to limit their structural complexity. Structural biology already plays a pivotal role in rational design of organic molecules and promises to be a game-changer in the development of new biologics. X-ray crystallography, NMR spectroscopy and cryo-electron microscopy are currently used to characterize the structure of macromolecules and macromolecular complexes.^[16-25] However, proteins covalently bound or strongly interacting with large drugs by long linkers can be hard to crystalize and, often, are too large for NMR spectroscopy in solution. Conversely, the use of cryo-electron microscopy is often prevented by the limited molecular weight and by the intrinsic flexibility of these systems.^[26] Solid-state NMR is currently used to investigate non-crystalline protein samples biologics and biomaterials^[27-46] and nowadays enhancement in sensitivity has been obtained by the recent achievements in NMR probes and by biomolecular Dynamic Nuclear Polarization (DNP).^[47-51]

Here, we show that solution and solid-state NMR provides complementary structural information which can be integrated with X-ray crystallography to design new protein-drug conjugates (PDC). The structural data have been used to design and synthesize a paclitaxel derivative conjugated to the Tafamidis molecule by a long linker containing an easily hydrolysable ester bond. The development of this molecule, that preserves a high affinity for transthyretin, shows the way for the integration of the different techniques to develop protein drug conjugates and its advantages over the single structural methodologies.

Methods.

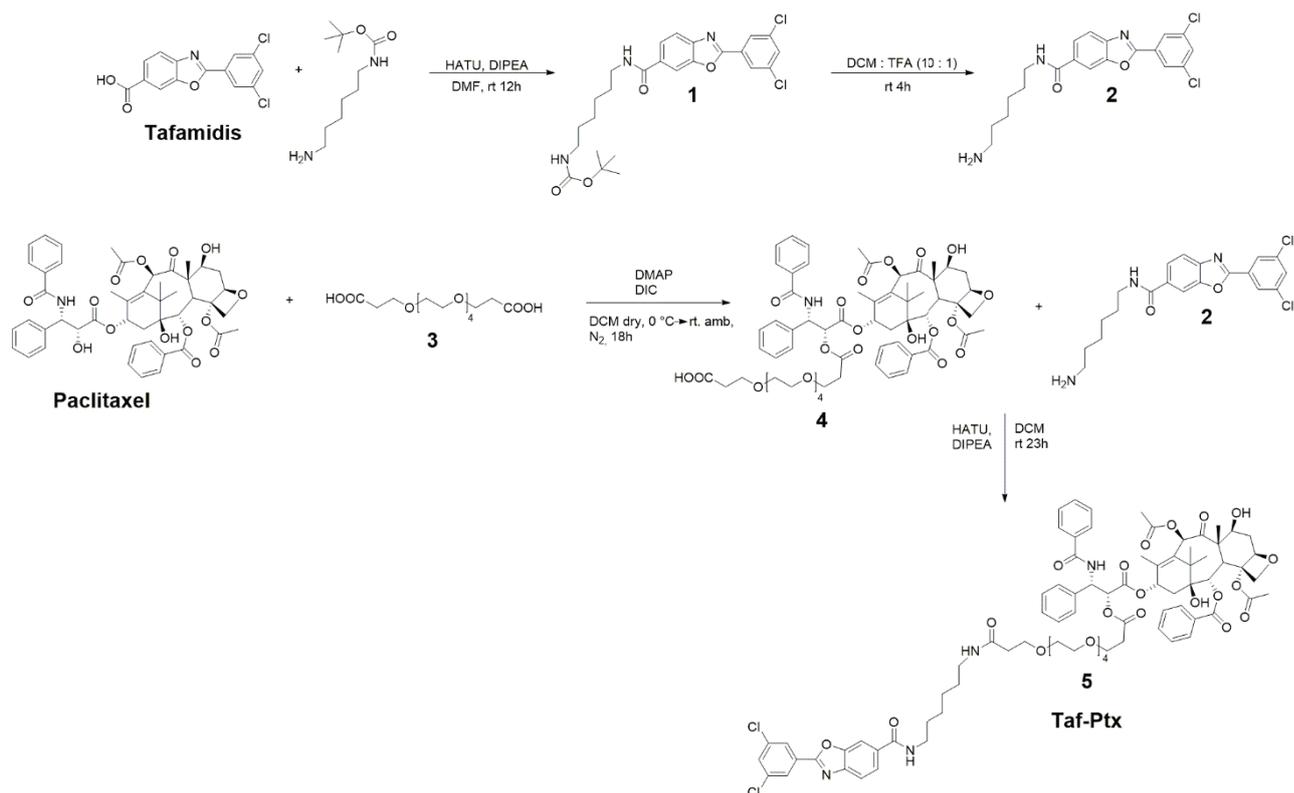
Expression and purification of [U-²H, ¹³C, ¹⁵N] and [U-¹³C, ¹⁵N] TTR.

Escherichia coli BL21(DE3) RIPL PLYS cells were transformed with pET-28a(+) plasmid encoding human TTR gene. For the expression of [U-¹³C, ¹⁵N] TTR, cells were grown in M9 minimal media supplemented with ¹⁵NH₄Cl and ¹³C-glucose. For the expression of [U-²H, ¹³C, ¹⁵N] TTR, cells were inoculated into 10 cm³ of Silantes OD2 medium and scaled up to 1 dm³ of the same enriched medium. The same experimental protocol was used to purify the samples of [U-¹³C, ¹⁵N] TTR and [U-²H, ¹³C, ¹⁵N] TTR. Cells were grown at 37 °C until optical density (OD₆₀₀) reached 0.6-0.8. Expression was induced with 1 mmol·dm⁻³ of isopropyl β-D-thiogalactoside (IPTG), cells were incubated at 37 °C overnight and harvested by centrifugation at 4 °C, for 15 min at 7500 rpm. Cell pellet was resuspended in lysis buffer [20 mmol·dm⁻³ Tris-HCl, pH 8.6, 5 mmol·dm⁻³ DTT, 1 mmol·dm⁻³ protease inhibitors], physical disrupted by sonication and centrifugated at 30000 rpm for 35 min at 4 °C. Soluble fraction was collected and purified by using an Anion exchange Q FF 16/10 column previously equilibrated in the lysis buffer. Elution was performed with increasing gradients of NaCl (60 cm³ of 0-200 mmol·dm⁻³, 300 cm³ of 0.2-0.5 mol·dm⁻³, 60 cm³ of 0.5-1.0 mol·dm⁻³ and 60 cm³ of 1 mol·dm⁻³) in the buffer. An SDS-PAGE was performed to identify which fractions contained the protein. The protein was further purified by size exclusion chromatography using a Hi Load 26/60 Superdex 75 pg column previously equilibrated in the final buffer [50 mmol·dm⁻³ MES-NaOH, pH 6.5, 0.1 mol·dm⁻³ NaCl, 5 mmol·dm⁻³ DTT, 1 mmol·dm⁻³ protease inhibitors].

Synthesis of Taf-Ptx.

Taf-Ptx was synthesized as reported in Scheme 1. In a reaction flask, HATU (190 mg, 0.50 mmol) was solubilized in 6 cm³ of DMF, then Tafamidis (170 mg, 0.55 mmol), DIPEA (142 mg, 1.1 mmol) and a solution of tert-butyl (6-aminohexyl) carbamate (108 mg, 0.50 mmol) in DMF were added. The solution was stirred at room temperature for 12 h. Control TLC was performed (EP : AcOEt = 2 : 1). Subsequently the mixture was diluted with H₂O, and NH₄Cl saturated solution was added until pH = 7. Extractions were performed with AcOEt (10x10 cm³) and then washes with H₂O (3x20 cm³) were

carried out. The organic layer was dried under anhydrous Na_2SO_4 , following filtration and evaporation of the solvent under reduced pressure to give the crude product **1** in 90% yield. Then, in a vial product **1** (50 mg, 0.1 mmol) was solubilized in 3.2 cm^3 of DCM and the solution was cooled to $0 \text{ }^\circ\text{C}$. Subsequently 0.3 cm^3 of TFA was added dropwise. The mixture was stirred at room temperature for 4 h. TLC control was performed to verify reaction completion (EP : AcOEt = 2 : 1). The mixture was diluted with 20 cm^3 of DCM and pH adjusted to 7-8 with NaHCO_3 saturated solution. Extractions were performed with DCM ($3 \times 20 \text{ cm}^3$) and the organic layer was dried over anhydrous Na_2SO_4 , following filtration and evaporation of the solvent under reduced pressure to give product **2** in 66% yield. To obtain product **3**, DMAP (1.35 mg, 0.012 mmol) was added to a solution of paclitaxel (100 mg, 0.12 mmol) in 16 cm^3 of anhydrous DCM under nitrogen flow. The solution was cooled to $0 \text{ }^\circ\text{C}$ and 4,7,10,13,16-pentaoxonadecanedioic acid (81.20 mg, 0.24 mmol) and then DIC (19 mm^3 , 0.12 mmol) were added. The reaction mixture was stirred at room temperature for 24 h. Control TLC were performed to monitor the reaction using as eluent (DCM : MeOH : AcOEt = 8 : 1 : 2). The mixture was diluted with 20 cm^3 of DCM, washed with a saturated solution of NH_4Cl ($3 \times 40 \text{ cm}^3$) and H_2O ($6 \times 20 \text{ cm}^3$). The organic layer was dried over anhydrous Na_2SO_4 , following filtration and evaporation of the solvent under reduced pressure to give the crude product. After purification by silica column chromatography (DCM : MeOH : AcOEt = 8 : 1 : 4) product **3** was obtained in 65% yield. To obtain the final product **4** (Taf-PtxIn) a reaction flask, HATU (8 mg, 0.021 mmol) was solubilized in 0.25 cm^3 of DCM then **7** (27 mg, 0.023 mmol), DIPEA (6 mg, 0.046 mmol) and a solution of **2** (10 mg, 0.025 mmol) in 0.25 cm^3 of DCM were added. The reaction mixture was stirred at room temperature for 24 h. Control TLC was performed to monitor the reaction (DCM : MeOH : AcOEt = 10 : 1 : 2). The mixture was diluted with 10 cm^3 of DCM and washed with a saturated solution of NH_4Cl ($3 \times 20 \text{ cm}^3$), H_2O ($3 \times 20 \text{ cm}^3$). Solvent was dried under Na_2SO_4 , following filtration and evaporation of the solvent under reduced pressure to give the crude product. After purification by silica column chromatography (DCM : MeOH : AcOEt = 6 : 1 : 2) Taf-Ptx was obtained in 22% yield.



Scheme 1. Synthesis of Taf-Ptx.

Crystallization, Data Collection and Structure Solution.

Crystals of apo TTR were obtained in sitting drop by adding an aliquot of 10 mm³ of protein solution in the final buffer [50 mmol·dm⁻³ MES-NaOH, pH 6.5, 0.1 mol·dm⁻³ NaCl, 5 mmol·dm⁻³ DTT, 1 mmol·dm⁻³ protease inhibitors] to 10 mm³ of crystallization buffer [100 mmol·dm⁻³ HEPES – NaOH, pH 7.5, 400 mmol·dm⁻³ CaCl₂, 34% PEG400].^[52] The reservoirs were filled with 800 mm³ of crystallization buffer and the plates were incubated at 20 °C. The protein concentration in the sample was 10 mg·cm⁻³. The native crystals of apo TTR were afterwards soaked in a solution containing the compound **2** with a 1 mmol·dm⁻³ concentration with respect to the protein for about some days. 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 10% ethylene glycol in the mother liquor. The crystal diffracted up to 1.5 Å resolution but the structure has been refined at 1.6 Å: it belongs to space group P2₁2₁2 with two molecules in the asymmetric unit, a solvent content of about 50%, and a mosaicity of 0.3°. The data were processed using the program XDS,^[53] reduced and scaled using XSCALE^[53] and amplitudes were calculated using XDSCONV.^[53] The structure has been solved using the molecular replacement technique; the model used was 3TCT.^[12] The successful orientation hand translation of the molecule within the crystallographic unit cell was determined with MOLREP.^[54] The refinement and water molecule fitting have been carried out using PHENIX.^[55] In

between the refinement cycles, the model was subjected to manual rebuilding using COOT.^[56] The quality of the refined structure was assessed using the program MOLPROBITY.^[57] Data processing and refinement statistics are shown in Table S1. Coordinates and structure factors have been deposited at the PDB under the accession code 8AWW.

NMR measurements

NMR relaxometry measurements. ¹H nuclear magnetic relaxation dispersion (NMRD) profiles were obtained for water solution samples of the free-TTR and of the protein in complex with Tafamidis, by measuring the water proton relaxation rates, R_1 , as a function of the applied magnetic field. The profiles were recorded with a SPINMASTER2000 fast field cycling relaxometer (Stelar, Mede (PV), Italy) operating in the 0.01–40 MHz, ¹H Larmor frequency range. The measurements are affected by an error of about $\pm 1\%$, as obtained in the field cycling experiment from the fit to a mono-exponential decay/recovery of the magnetization.

The profiles were fitted as the sum of multiple relaxation contributions arising from different correlation times τ_i , according to the model-free approach:^[58,59]

$$R_1 = \alpha + \beta \sum_i^N c_i \left(\frac{\tau_i}{1 + \omega^2 \tau_i^2} + \frac{4\tau_i}{1 + 4\omega^2 \tau_i^2} \right)$$

where c_i are weight coefficients summing to 1. The parameter β depends on the squared proton-proton dipole–dipole interaction energy and on the protein concentration, and the coefficients c_i report on the contributions from protons with the associated correlation times. The parameter α takes into account the contribution to water relaxation from protons with correlation times smaller than few ns (*i.e.*, with a dispersion occurring beyond the highest magnetic field). In order to reduce the covariance among the many unknown parameters, ¹H NMRD profiles at the same temperature were fitted simultaneously with common values of the correlation times. This permitted to better monitor differences in the values of c_i . The parameter β was also constrained to be the same because of the same protein conditions and concentration.

Solution NMR experiments on free-TTR. Solution NMR experiments for backbone resonance assignment with a TROSY scheme^[60] [3D tr-HNCA, tr-HNCACB, tr-HNCO, tr-HN(CA)CO]^[61–66] were performed on perdeuterated [²H, ¹³C, ¹⁵N] samples of native TTR (at the concentration of 500 $\mu\text{mol dm}^{-3}$ with respect to the monomer) in water buffer solution [50 $\text{mmol} \cdot \text{dm}^{-3}$ MES, pH 6.5, 100 $\text{mmol} \cdot \text{dm}^{-3}$ NaCl, 5 $\text{mmol} \cdot \text{dm}^{-3}$ DTT, 0.1% NaN₃, protease inhibitors (Roche)]. For 3D tr-HNCACB and tr-HN(CA)CO nonuniform random sampling at 57% and 25%, respectively, and compressed-sensing reconstruction were used.^[67] A 3D ¹H-¹⁵N NOESY-TROSY spectrum (mixing time 100 ms) was also acquired to help and confirm the sequential assignment. All the spectra were recorded at 310 K on Bruker AVANCE MHD and AVANCE NEO NMR spectrometers, operating at 950 and 900 MHz (¹H

Larmor frequency), respectively, and equipped with triple resonance cryo-probes. Two-dimensional carbon-detected solution ^{13}C - ^{15}N CON NMR spectrum was acquired on a Bruker AVANCE NEO 700 MHz spectrometer, equipped with a triple-resonance cryo-probe optimized for ^{13}C -direct detection.^[68,69]

Solution NMR experiments on TTR/Tafamidis and TTR/Taf-Ptx complexes. The titration with Tafamidis was performed using uniformly ^{13}C , ^{15}N -isotopically enriched [^{13}C , ^{15}N] TTR at the concentration of $800\ \mu\text{mol}\cdot\text{dm}^{-3}$ (with respect to the monomer) in $50\ \text{mmol}\cdot\text{dm}^{-3}$ MES buffer at pH 6.5, with $100\ \text{mmol}\cdot\text{dm}^{-3}$ NaCl, $5\ \text{mmol}\cdot\text{dm}^{-3}$ DTT, 0.1% NaN_3 and protease inhibitors (Roche). Increasing aliquots of the ligand (solubilized in DMSO-d_6), to reach the final concentrations in solution of 25, 50, 100, 200, 400 and $800\ \mu\text{mol}\cdot\text{dm}^{-3}$, were added to TTR solution and 2D ^1H - ^{15}N TROSY-HSQC acquired at 950 MHz after each addition. The binding of Taf-Ptx to TTR was also evaluated by an NMR titration in solution. Increasing aliquots of the ligand (solubilized in DMSO-d_6) were added to the solution of [^{13}C , ^{15}N] TTR [$\sim 400\ \mu\text{mol}\cdot\text{dm}^{-3}$, with respect to the monomer, in $50\ \text{mmol}\cdot\text{dm}^{-3}$ MES buffer at pH 6.5, with $100\ \text{mmol}\cdot\text{dm}^{-3}$ NaCl, $5\ \text{mmol}\cdot\text{dm}^{-3}$ DTT, 0.1% NaN_3 and protease inhibitors (Roche)] and 2D ^1H - ^{15}N TROSY-HSQC acquired at 950 MHz after each addition. After both titrations, the excess of the ligands was removed using PD10 column and the buffer exchanged to $10\ \text{mmol}\cdot\text{dm}^{-3}$ MES, pH 6.5 and $20\ \text{mmol}\cdot\text{dm}^{-3}$ NaCl. PEG1000 (in 1:10 weight ratio with respect to the protein) was added to protect the protein during the lyophilization process. The solutions (containing ~ 6 - 8 mg of protein/ligand, 0.6-0.8 mg of PEG1000, 1 mg MES, 0.7 mg NaCl) were freeze-dried and the materials used to pack 3.2 mm zirconia thick wall rotors. The materials were then rehydrated by multiple additions of MilliQ H_2O until the resolution of the 1D $\{^1\text{H}\}^{13}\text{C}$ CP solid-state NMR spectra stopped changing.^[70] Silicon plugs (courtesy of Bruker Biospin) placed below the turbine cap were used to close the rotor and preserve hydration.

A sample of [^{13}C , ^{15}N] free TTR lyophilized in the presence of PEG1000 (in the same weight ratio, 25 mg TTR: 2.5 mg PEG1000) was also analyzed in 3.2 mm zirconia rotor after rehydration, as reference for solid-state NMR.

The solid-state NMR spectra of free TTR and TTR in the presence of Tafamidis or Taf-Ptx were collected on a Bruker Avance III spectrometer operating at 800 MHz (18.8 T, 201.2 MHz ^{13}C Larmor frequency) equipped with a Bruker 3.2 mm Efree NCH probe-head and on a Bruker Avance III 850 MHz wide-bore spectrometer (20 T, 213.6 MHz ^{13}C Larmor frequency), equipped with 3.2 mm DVT MAS probe head in triple-resonance mode. The spectra were recorded at 14 kHz MAS frequency and the sample temperature was kept at ~ 290 K.

Standard ^{13}C -detected solid-state NMR spectra [2D ^{15}N - ^{13}C NCA, ^{15}N - ^{13}C NCO and ^{13}C - ^{13}C DARR; 3D NCACX and NCOCX] were acquired on the samples in 3.2 mm rotors, using the pulse sequences reported in the literature.^[71-77] 3D CANCO experiment was also acquired on the sample of free TTR.

All the spectra were processed with the Bruker TopSpin 3.2 software and analyzed with the program CARRA.^[78]

Results and Discussion.

Tafamidis was selected to anchor Paclitaxel onto TTR because of its high affinity for the two binding sites of the tetrameric assembly and for the availability of the X-ray structure of the Tafamidis-TTR complex (PDB code: 3TCT).^[12] The X-ray structure shows the binding mode of Tafamidis and suggests a suitable position on the molecule to host the linker needed to connect a Paclitaxel moiety. The visual inspection of the X-ray structure (see Figure 1) suggests that the functionalization of Tafamidis on carboxylic acid in position 6 does not significantly interfere with the interaction between the ligand and the tetrameric assembly. Therefore, a derivative of Tafamidis bearing a six carbon atoms linker at position 6 (compound **2**) was synthesized and its complex with TTR investigated by X-ray crystallography. The molecule was small enough to be used in soaking experiments with TTR. The X-ray analysis of the soaked crystals reveals that the polypeptide structure is totally superimposable with several others already deposited on the PDB and, of course, to the one used as the molecular replacement model that bears a very similar ligand (3TCT). Figure S1 shows the superposition between the structure presented in this work and 3TCT (just one monomer for both structures is presented for clarity), with a RMSD of backbone atoms as low as 0.82 Å: it is clearly visible that the only deviations are in external loops whereas the other regions show a negligible RMSD. As in the case of 3TCT, there is a peculiar crystallographic feature concerning the ligand: it sits with the longitudinal axis crossing the aromatic rings almost coincident with the crystallographic two-fold axis. This situation implies that the refinement has to be carried out placing the ligand in the density at half occupancy; the application of the crystallographic symmetry itself generates a symmetry mate for the ligand, which is slightly tilted about the above mentioned two-fold axis (Figure 1). Concerning the ligand, it still has to be pointed out that the electron density is well defined for the aromatic rings only, whereas the long hydrocarbon chain that is attached to the rings shows very faint patches here and there. This is certainly due to the mobility of this chain that is totally projected into the solvent. The interactions between the protein and the ligand are displayed in Figure S2. Summarizing, the X-ray structure proves that the synthesized molecule preserves the binding pose of Tafamidis with the linker sticking-out from the central channel toward the solvent (see Figure 1) thus supporting the correct design of this first ligand. Then, the Paclitaxel molecule was conjugated to compound **2** through an additional nineteen-atoms PEG spacer to minimize any possible steric clash with the protein and to increase the solubility of the final molecules. Paclitaxel is connected to the PEG spacer through a labile ester linkage, which ensures its release into the cell allowing the inhibition of tubulin polymerization.^[79,80] The molecule (Taf-Ptx) is too bulky to be used in soaking experiments and does not allow for the co-crystallization with TTR. However, the molecular weight of the tetrameric assembly of TTR is still within the range affordable by solution NMR although

deuterated samples and TROSY-based pulse sequences are required for resonance assignment by 3D experiments.

To investigate the binding mode of Taf-Ptx to TTR and the preservation of the stabilizing effects on the tetramer, a comparative and extensive NMR analysis was performed on the free TTR and on the protein in the presence of Tafamidis and Taf-Ptx. The analysis was carried out in solution and in the solid-state to compare the quality and the relevance of the structural information of the two methodologies and to evaluate the efficacy of the use of solid-state NMR alone in the development of large protein-drug conjugates, when the application of the other methodologies is hampered.

NMR relaxometry

Firstly, a relaxometric analysis was carried out to investigate the stability of the tetramer and the overall steadiness of protein dynamics upon ligand binding. The ^1H NMRD profiles of 1.2 mM wild-type TTR in water solutions, with and without Tafamidis, are shown in Figure 2. The profiles were acquired at 15, 25 and 37 °C and, at each temperature, they appear close to one another, indicating that no sizable changes in the overall dynamics of the protein occur upon addition of the ligand.

The ^1H NMRD profiles report the field dependent relaxation rates of water protons, which interact with the protein protons.^[31,58,59,81] Their field dependence is determined by the correlation times modulating the proton-proton dipole–dipole interactions. Multiple correlation times must be considered to account for the many motional processes of the different water protons interacting with the protein. These correlation times are the fastest between the proton exchange time and the reorientation times, comprising both the overall protein tumbling and faster protein local dynamics. Three correlation times were needed for the fit of the profiles (see Table S2), the shortest of the order of nanoseconds, thus reporting on the internal protein mobility and/or the lifetime of proton exchange processes, and the longest of the order of hundreds or thousands of nanoseconds. The weight coefficients of the latter are very low (< 1%), and thus correspond to contributions from aggregated forms of the protein. The intermediate correlation time is of few tens of nanoseconds, hence corresponding to the overall reorientation time of the protein. This correlation time in fact corresponds to 30 ns at 25 °C, and is in nice agreement with the overall reorientation time expected for tetrameric TTR, as it can also be calculated with HydroNMR^[82] (29 ns).

The weight coefficients c_i indicate a slightly higher propensity of the free protein to form aggregates than in the presence of Tafamidis. Although within the experimental error, the data also suggest that the presence of Tafamidis determines a slightly larger contribution from the intermediate correlation time (c_2 of 0.27 instead of 0.24), which may correspond to a more rigid tumbling of the tetrameric protein assembly.

The ^1H NMRD profiles acquired at the other two temperatures could be nicely fitted in agreement with the values obtained for the samples at 25 °C, with an intermediate correlation time of about 40 and 20 ns at 15 and 37 °C, respectively, in close agreement with the HydroNMR values of 38 ns and 22 ns calculated for these temperatures.

NMR analysis and assignment of free native TTR in solution

The assignment of free tetrameric TTR in solution was obtained comparing the assignments available in the literature for the monomeric and tetrameric states of the protein^[83–86] and analyzing triple resonance spectra recorded on the perdeuterated sample of TTR. All residues (but the N-terminus, Gly-1 and Asn-98) have been assigned in the spectra (Figure S3). The present assignment is the most complete and has been deposited in the bmrB under the accession code YYYYY.

The number of cross-peaks present in the 2D ^1H - ^{15}N TROSY-HSQC spectrum, and the absence of signal splitting, are both consistent with the preservation of the D_2 symmetry of TTR in the tetrameric assembly also supported by the relaxometric analysis. Interestingly, the signals within the same 2D spectra are characterized by different line broadening. In particular, sharp and intense signals were observed for the residues forming loops and on the external surface of the tetramer, while broad signals were observed for the residues at the interfaces between the monomers (i.e., Cys10-Lys15; His91-Phe95; Tyr105-Val121). Also in the 2D ^1H - ^{15}N TROSY-HSQC spectrum recorded on the deuterated sample, these signals are broad, suggesting the occurrence of a conformational heterogeneity/exchange for the protein. These features also affect the quality of the 3D ^1H - ^{15}N NOESY spectrum where only few NOE correlations are visible. As expected for a folded protein of 50 kDa, in the 2D ^{13}C - ^{15}N CON spectrum only signals of the flexible regions, which are not well-defined in most X-ray structures, can be observed (Thr3-Cy10; Ala37-Thr40; Glu51-Ser52; Asn124-Glu127).

NMR analysis and assignment of free native TTR in the solid-state

The solid-state NMR spectra of re-hydrated freeze-dried tetrameric TTR are of good quality and characterized by a good signal resolution (Figure S4). Nevertheless, around 20% of the expected resonances are missing and some signals are characterized by large line-broadening. Assignment of the free tetrameric protein has been also obtained in the solid-state (Figure 3). The available assignment of the free protein in solution has been used as starting point and complemented by the analysis of new carbon-detected spectra acquired in the solid-state. The residues whose signals are missing in the spectra, are mainly located at the N-terminus (up to Lys15) and in flexible regions: Asp38-Thr40, Gly57-Leu58, Phe64, Ser117-Thr119, Thr123-Asn124, Lys126. However, 80% of the spin systems of the protein have been reassigned in the solid-state NMR spectra.

NMR analysis and assignment of TTR-Tafamidis and TTR-Taf-Ptx in solution

The binding-mode of Tafamidis to [U- ^{13}C , ^{15}N] TTR was first analyzed by solution NMR. During the NMR titration, in the 2D ^1H - ^{15}N TROSY-HSQC spectra, the cross-peaks of the free protein decrease in intensity upon the addition of increasing concentrations of the ligand, while new cross-peaks, corresponding to the complex between TTR and Tafamidis, appear and increase in intensity. This behavior indicates that the ligand is in slow exchange regime on the NMR timescale, and confirms its expected high affinity towards the protein ($K_{\text{dS}} \sim 2 \text{ nM}$ and $\sim 200 \text{ nM}$).^[12,87] In the presence of Tafamidis at $100 \mu\text{M}$ concentration (tetramer:ligand ratio equal to 1:0.5) the cross-peaks corresponding to the free protein and to the protein bound to the ligand have similar intensities (Figure 4A), as visible for the signal of Ser112 located at the interface between the dimers of the tetrameric assembly (PDB code: 3TCT).^[12] The analysis of the chemical shift perturbation (CSP) at the end of the titration (Figure S5A) confirmed that the residues experiencing the largest changes (Lys15, Leu17, Ala19, Val20, Arg21, Gly22, Ser23, Ile26, Gly53, His88, Val94, Tyr105, Thr106, Ile107, Leu111, Ser112, Ala120, Val122) are located in the expected Tafamidis binding site (Figure 5), that is shared by the natural ligand L-thyroxine (T_4). Some of the signals experiencing large perturbations (Ala19, Arg21, Gly22, Leu111) have been tentatively reassigned with some uncertainty (Figure 5A). Interestingly the signals corresponding to residue Ser117 and Thr118 which are almost missing in the spectra of the free protein appear with increased intensity in the spectrum of TTR in the presence of Tafamidis.

NMR titration in solution was also performed with the new designed ligand, Taf-Ptx. The evolution of the spectra upon the addition of increasing amounts of Taf-Ptx was superimposable to that previously observed for Tafamidis, with several protein resonances experiencing a slow exchange regime on the NMR timescale. After the addition of Taf-Ptx, in the presence of a tetramer:ligand ratio equal to 1:0.5, the signals of the free protein and those of the protein bound to the ligand have similar intensities (Figure 4B). The analysis of CSP between the protein signals in the complex and the free protein (Figure S5B) confirmed that the residues experiencing the largest changes (Lys15, Leu17, Val20, Arg21, Gly22, Ala25, Ile26, Asp99, Leu111, Ser112, Ala120, Val121, Val122) are in the same protein region affected by the presence of Tafamidis (Figure 5). It is important to point out that the signals of some residues mostly affected by CSP in the spectrum of TTR/Tafamidis (Ala19, His88, Val94, Tyr105, T106) are broaden beyond detection in the spectrum of TTR/Taf-Ptx. This is probably due to the sizably lower concentration of the sample of TTR/Taf-Ptx with respect to that of TTR/Tafamidis and to their large line-broadening.

NMR analysis and assignment of TTR-Tafamidis and TTR-Taf-Ptx in the solid-state

The complexes between TTR and Tafamidis or Taf-Ptx were analyzed also by solid-state NMR. The 2D solid-state NMR spectra of these complexes exhibit a higher number of cross-peaks with respect

to those of the free protein. In particular, in the 2D ^{13}C - ^{13}C DARR spectrum of TTR in the presence of either Tafamidis or Taf-Ptx, some new signals appear or increase in intensity (Gly57, Arg103, Ile107, Ala108, Ser117, Thr118, Thr119, Ala120, Thr123; see Figure 6, panel A and B). This also occurs in 2D ^{15}N - ^{13}C NCA spectra (Figure 6, panel C and D). These signals belong to residues located at the tetramer interface, where Tafamidis binds (Figure 7C/D). The increase in signal intensity can be explained by a higher rigidity of this region after the binding of Tafamidis, which is known to freeze the tetrameric form of the protein. As shown by the high similarity of the spectra of TTR bound to one ligand or to the other, the high affinity of the Tafamidis unit, and its stabilization effect on the tetramer, are still present also when the Tafamidis unit is conjugated to the paclitaxel unit. However, a lower intensity of the signals corresponding to residues S117 and T118 has to be pointed out when TTR binds to Taf-Ptx compared to when it binds to Tafamidis.

The analysis of the CSP of TTR bound to Tafamidis or Taf-Ptx with respect to the free protein indicates that the signals influenced by the ligands are largely the same (Ala19, Phe44, Gly47, Lys48, Val65, Glu66, Gly83, Arg103, Leu110, Ser115, Val121, for Tafamidis and Ala19, Phe44, Gly47, Lys48, Val65, Glu66, Gly67, Arg103, Ser115, Val121 for Taf-Ptx). Most of the signals experiencing the largest perturbations correspond to residues at the dimer/tetramer interface (Figure 7).

Impact of solid-state NMR on drug design

CSP in solution are extensively used to analyze and compare the binding mode of ligands interacting with proteins or nucleic acids. For TTR the spectral changes resulting from the binding of Tafamidis or Taf-Ptx are both in the slow exchange regime on the NMR time scale, in agreement to the high affinity of Tafamidis for the protein previously observed and well characterized. Most of the residues experiencing the largest chemical shift variations on the NH resonances are placed at the interface of the two dimers forming the central channel which usually hosts the thyroxine T4 hormone. In particular, Lys15, Leu17, Val20, Arg21, Gly22, Ile26, Leu111, Ser112, Ala120, Val122 show chemical shift perturbation in the presence of both ligands, Tafamidis and Taf-Ptx. More important, the visual inspection of the X-ray structure 3TCT shows that among the residues experiencing the largest CSP, only Lys15, Leu17 and Thr106 seem to interact directly with the two ligands within the tetramer. Therefore, the structural rearrangement associated to the tetramer stabilization induced by the two investigated ligands seems to play the major role in determining the largest CSPs. These systems thus represent interesting examples where the typical approach based on chemical shift mapping to design and improve interacting molecules is extremely challenging.

The stabilizing effect of Tafamidis was unambiguously proved *in vitro* and *in vivo*. However, the chemical shift mapping obtained by NMR in solution provides only indirect evidence of this important structural effect, which can be inferred from the slow exchange regime on the NMR timescale

of the signals corresponding to the residues where the binding occurs. At this regard, more informative is the observation of the slightly better quality of the spectra recorded on TTR in the presence of Tafamidis and Taf-Ptx. Another important finding of the comparative NMR analysis carried out in solution is the experimental evidence of preservation of the binding mode of Tafamidis also for Taf-Ptx, that proves the success of the functionalization with the Paclitaxel moiety.

It is important to point out that solid state NMR can be used to investigate very large biomolecular systems which are not accessible to NMR in solution. More informative has been the analysis of the data recorded on the solid-state samples of TTR in the presence and in the absence of the two ligands. The comparative analysis on Tafamidis and Taf-Ptx has been carried out by using ^{15}N - ^{13}C isotopically enriched samples of TTR and ^{13}C -detected experiments. The analysis of the NCA spectra shows that TTR in complex with Tafamidis and Taf-Ptx exhibits the largest chemical shift variations on the same residues with very few differences. This is a further experimental evidence about the very similar binding mode of Taf-Ptx on TTR with respect to Tafamidis. The chemical shift variations affect several residues placed at the interface of the two dimers around the central channel as previously observed by NMR in solution, with three residues (Leu110, Ser115 and Val121) near the ligands in the X-ray structure 3TCT. The mismatch between solution and solid-state data about the residues experiencing the largest effects is not real because they are close to each other, and in both cases localized at the interface between the two dimers forming the assembly. This can be clearly inferred from Figure 8 where the residues experiencing the largest CSPs in solution and in the solid-state are shown together.

NMRD measurements indicate that the tetrameric assembly of TTR is maintained, and possibly reinforced, in the presence of Tafamidis. However, relaxometry is not sensitive to the presence of multiple conformational states with similar reorientation time. Important information on this respect has been obtained from the qualitative analysis of the signal intensity on DARR and NCA spectra. Several residues, including some experiencing also large chemical shift variation, increase in intensity or appear in the spectra in the presence of Tafamidis and Taf-Ptx. The increase in signal intensity or the appearance of a signal previously undetectable in the solid-state spectra is conclusive evidence of an equilibrium of the atom/residue shifted toward a unique conformation that in this case is associated with the structural stabilization of the tetrameric assembly resulting from the interaction of two high affinity ligands with TTR. Therefore, the analysis provides a map of the residues adopting a unique conformation thus providing a different and more informative parameter to monitor the binding mode and to evaluate the effects of ligands in large proteins and in multimeric systems.

Summarizing, this study shows that structure-based molecular design strategies can be fruitfully used to design protein-drug conjugates. Our results show that the integration of different structural biophysical methodologies is particularly suitable for the development of such systems where the classic

strategies based on single techniques are not applicable and resolutive. More important, we realize that solid-state NMR can provide new detailed information not only on the binding pose of the ligands but also on the structural effects of the interaction opening new ways for the development of new protein-drug conjugates.

Acknowledgements

This work has been supported by Regione Toscana (CERM-TT, BioEnable and PANCREAS-AD bando salute 2018), the Italian Ministero dell'Istruzione, dell'Università e della Ricerca through PRIN 2017A2KEPL, the “Progetto Dipartimenti di Eccellenza 2018-2022” to the Department of Chemistry “Ugo Schiff” of the University of Florence, the Recombinant Proteins JOYNLAB laboratory, and the project FISR2021_SYLCOV. The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy centre. We acknowledge H2020 -INFRAIA iNEXT-Discovery - Structural Biology Research Infrastructures for Translational Research and Discovery (contract n° 871037), H2020 EOSC-Life "Providing an open collaborative space for digital biology in Europe" (contract n° 824087), H2020 FET-Open project HIRES-MULTIDYN “Multiscale Dynamics with Ultrafast High-Resolution Relaxometry” (contract n° 899683), RNAct” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITNAnalytical 2018 (contract n° 813239) and “Glytunes” Marie Skłodowska-Curie Action (MSCA) Innovative Training Networks (ITN) H2020-MSCA-ITN-2020 (contract n° 956758). The authors acknowledge also Mestrelab Research for providing Mnova software and Bruker BioSpin for AssureNMR software.

Figure Captions.

Figure 1. Detail of the protein and ligand electron density map for the complex between TTR and Compound **2** (the two possible orientations of compound **2** within the binding site are shown) (A). Superposition between the structure presented (XXX) in this work and 3TCT (B). Surface representation of TTR interacting with one of the two possible conformations of compound **2** displayed as stick (C). In all the three panels, the conformation of the aliphatic chain of compound **2** is arbitrarily defined.

Figure 2: ^1H NMRD profiles of wild-type TTR, in the absence and in the presence of Tafamidis at 15, 25 and 37 °C.

Figure 3. 2D ^{15}N ^{13}C NCA spectrum of rehydrated freeze-dried [$\text{U-}^{13}\text{C}$, ^{15}N] TTR with the assignment reported on the signals. The spectrum was acquired on a spectrometer operating at 800 MHz (^1H Larmor frequency) with MAS of 14 kHz and temperature of ~ 290 K.

Figure 4. Region of the 2D ^1H - ^{15}N HSQC spectra of free TTR (200 μM with respect to the tetramer, blue) and TTR in the presence of ~ 100 μM Tafamidis (A, red) or ~ 100 μM Taf-Ptx (B, red). The signals of the free and bound species display almost equal intensities as indicated by the assigned peak of Ser112.

Figure 5. (A, B) Chemical shift perturbation (CSP) of TTR in the presence of Tafamidis (A) and Taf-Ptx (B) (in 1:1 ratio) with respect to free TTR, evaluated according to the formula $\Delta\delta = 1/2 \sqrt{\Delta\delta_H^2 + (\Delta\delta_N/5)^2}$. The residues experiencing the largest variations have been highlighted in red and blue, respectively. The residues that were tentatively assigned have been indicated by a star. (C, D) CSP mapping on the X-ray structure of TTR in complex with Tafamidis (PDB code: 3TCT)^[12] with the residues experiencing the largest perturbations in the presence of Tafamidis or Taf-Ptx colored in red (C) and blue (D), respectively. The residues missing in the 2D ^1H - ^{15}N HSQC spectra are colored in grey. The monomers are in different colors (wheat, green, violet, pink) and Tafamidis as yellow sticks.

Figure 6. Comparison of a region of 2D ^{13}C - ^{13}C DARR spectra (mixing time 50 ms, A, B) and ^{15}N - ^{13}C NCA spectra (C, D) of free rehydrated freeze-dried TTR (blue) and TTR in the presence of the ligand (red), Tafamidis (A, C) or Taf-Ptx (B, D). The assignment of some signals, with increased intensity in the spectra of TTR in complex with the ligands with respect to the free protein, has been reported.

Figure 7. (A) Chemical shift perturbation (CSP) of rehydrated freeze-dried TTR in the presence of Tafamidis (A) and Taf-Ptx (B) with respect to free rehydrated freeze-dried TTR, evaluated according to the formula $\Delta\delta = 1/2 \sqrt{(\Delta\delta_{C\alpha}/2)^2 + (\Delta\delta_N/5)^2}$.^[88] The residues experiencing the largest variations have been highlighted in violet (Ala19, Phe44, Gly47, Lys48, Val65, Glu66, Gly83, Arg103, Leu110,

Ser115, Val121, for Tafamidis) and orange (Ala19, Phe44, Gly47, Lys48, Val65, Glu66, Gly67, Arg103, Ser115, Val121, for Taf-Ptx), respectively. (C, D) CSP mapping on the X-ray structure of TTR in complex with Tafamidis (PDB code: 3TCT)^[12] with the residues experiencing the largest CSP in the presence of Tafamidis or Taf-Ptx colored in red (C) and blue (D), respectively. The residues experiencing an increase in signal intensity after Tafamidis or Taf-Ptx binding have been colored in magenta. The monomers are in different colors (wheat, green) and Tafamidis is shown as yellow sticks.

Figure 8. Comparative analysis of the NMR data collected in solution and in the solid-state of the effects of Tafamidis (A) and Taf-Ptx (B) on TTR. The residues experiencing in solution the largest CSP in the presence of Tafamidis or Taf-Ptx are colored in red (A) and blue (B), respectively. The residues experiencing in the solid-state the largest CSP in the presence of Tafamidis or Taf-Ptx are colored in violet (A) and orange (B), respectively. Residues experiencing in the solid-state an increase in signal intensity after the binding of Tafamidis or Taf-Ptx are colored in magenta. Residues missing in solution and solid-state spectra are colored in grey. Residues experiencing the largest CSP both in solution and in the solid state are reported in black. Monomers are in different colors (wheat, green, violet, pink) and Tafamidis as yellow sticks.

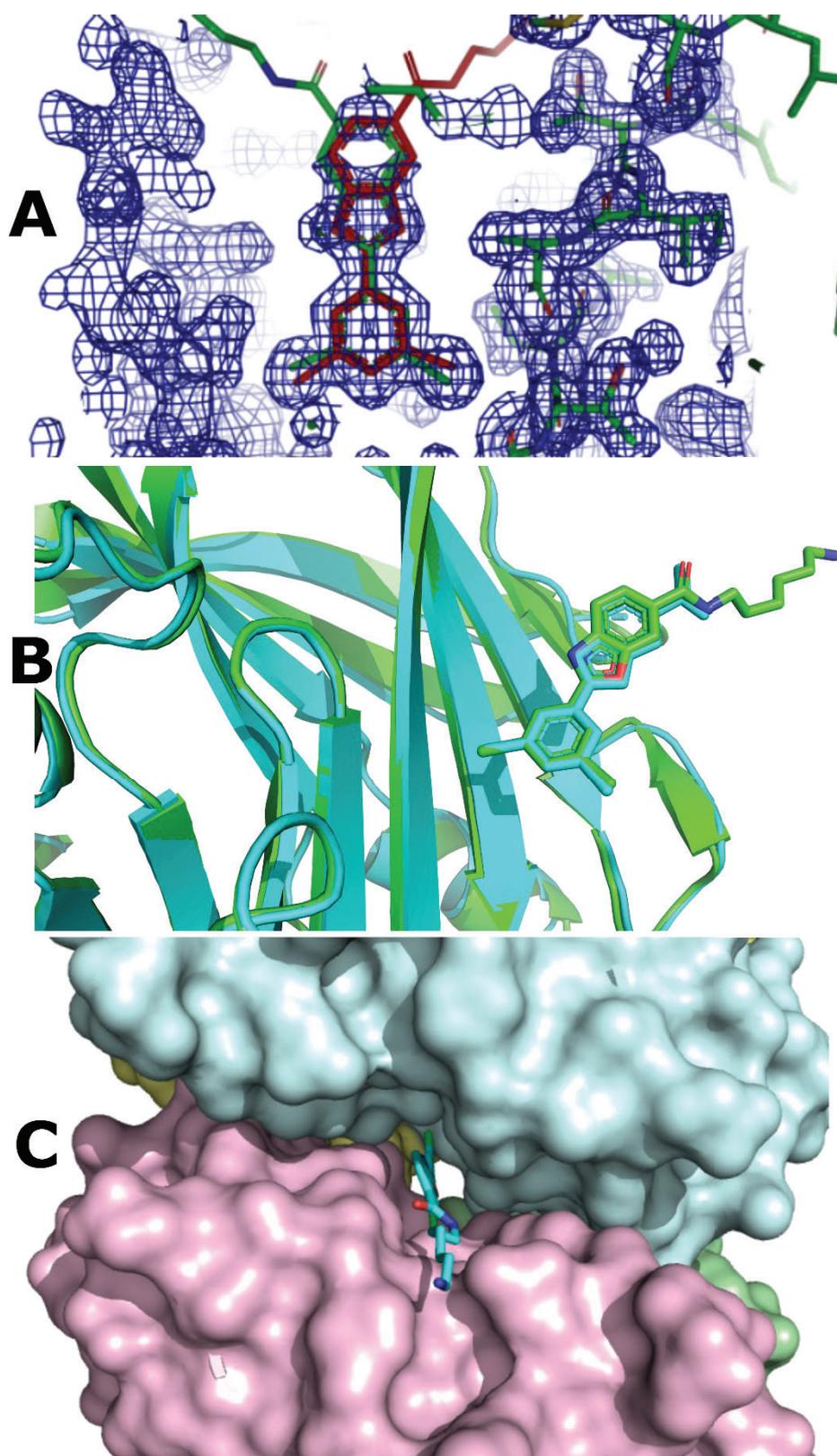


Figure 1.

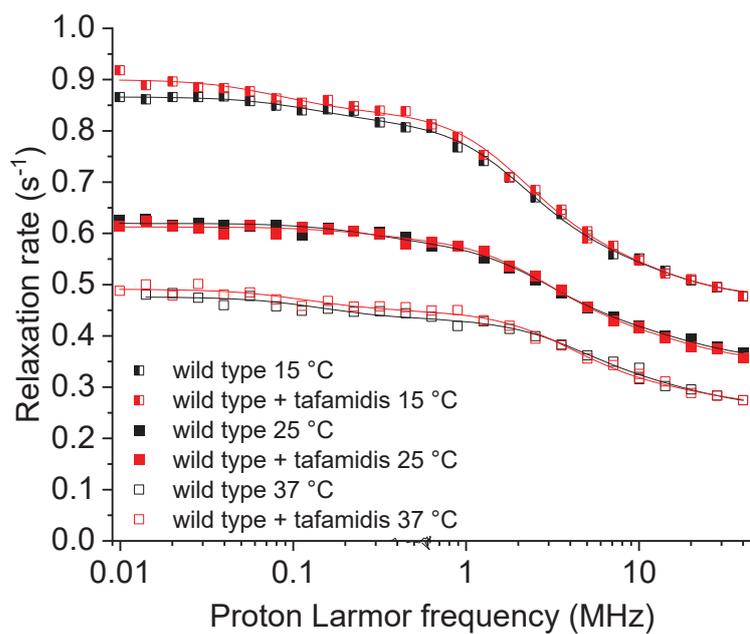


Figure 2.

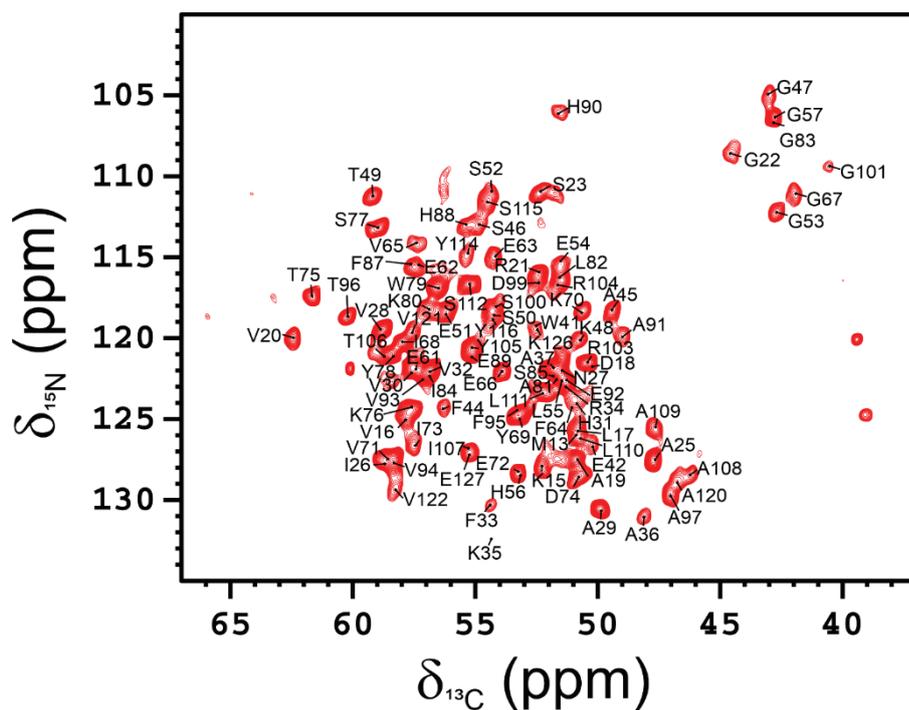


Figure 3.

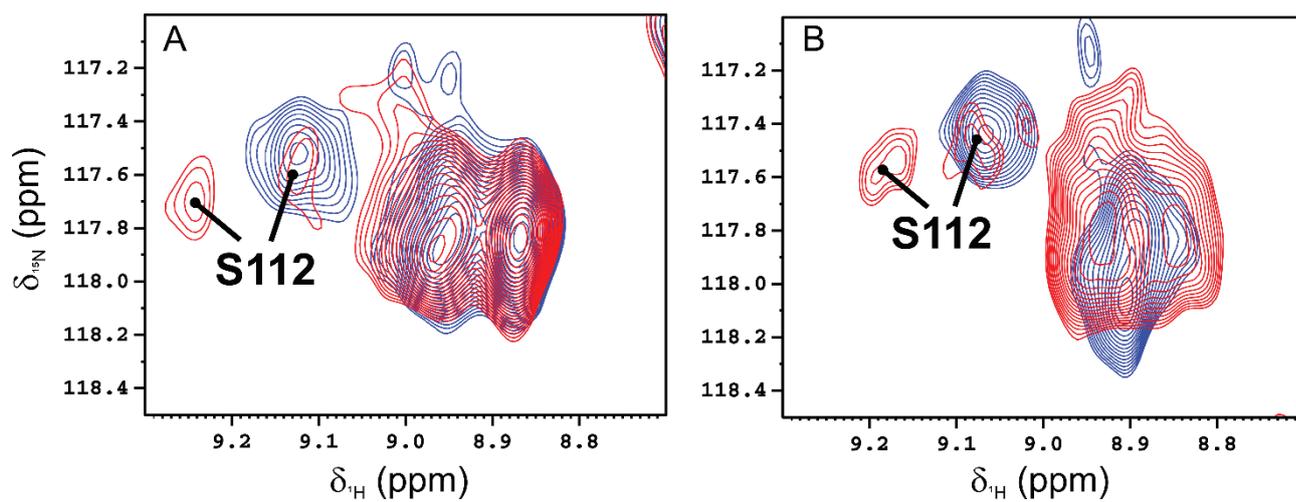


Figure 4.

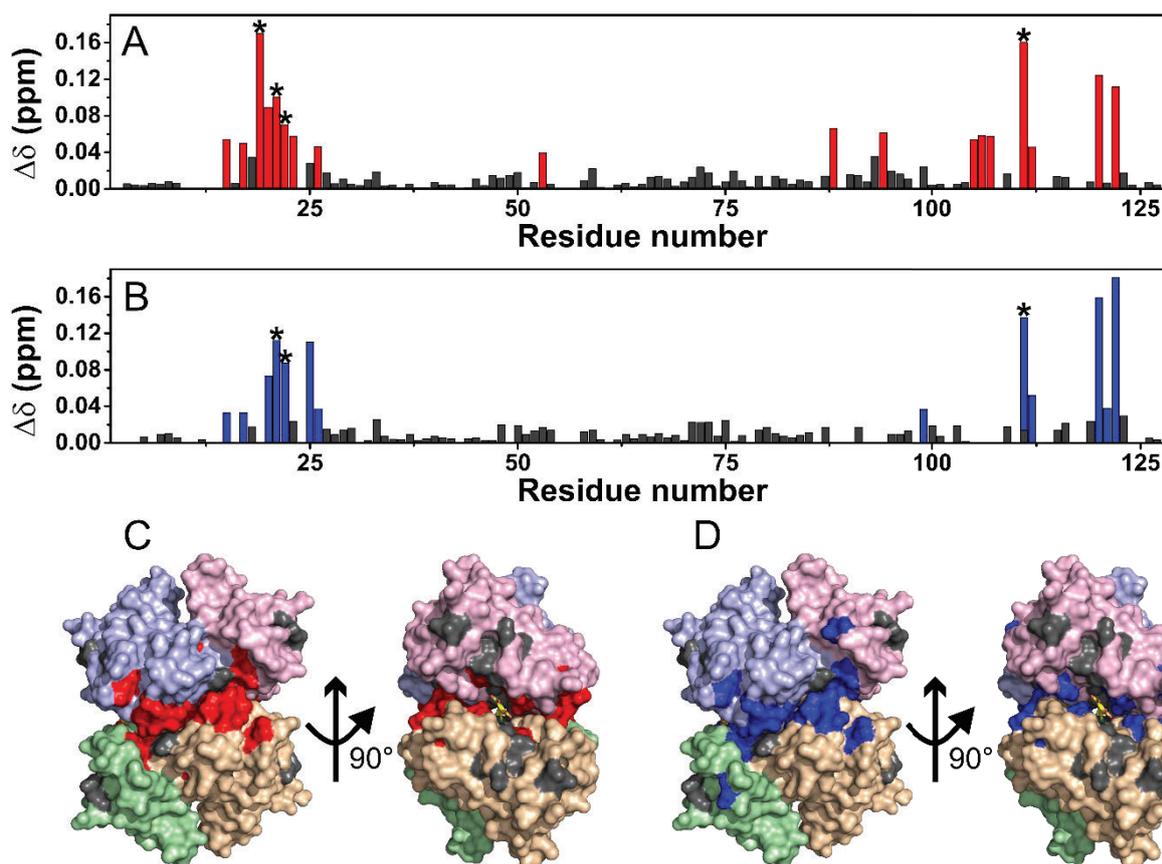


Figure 5.

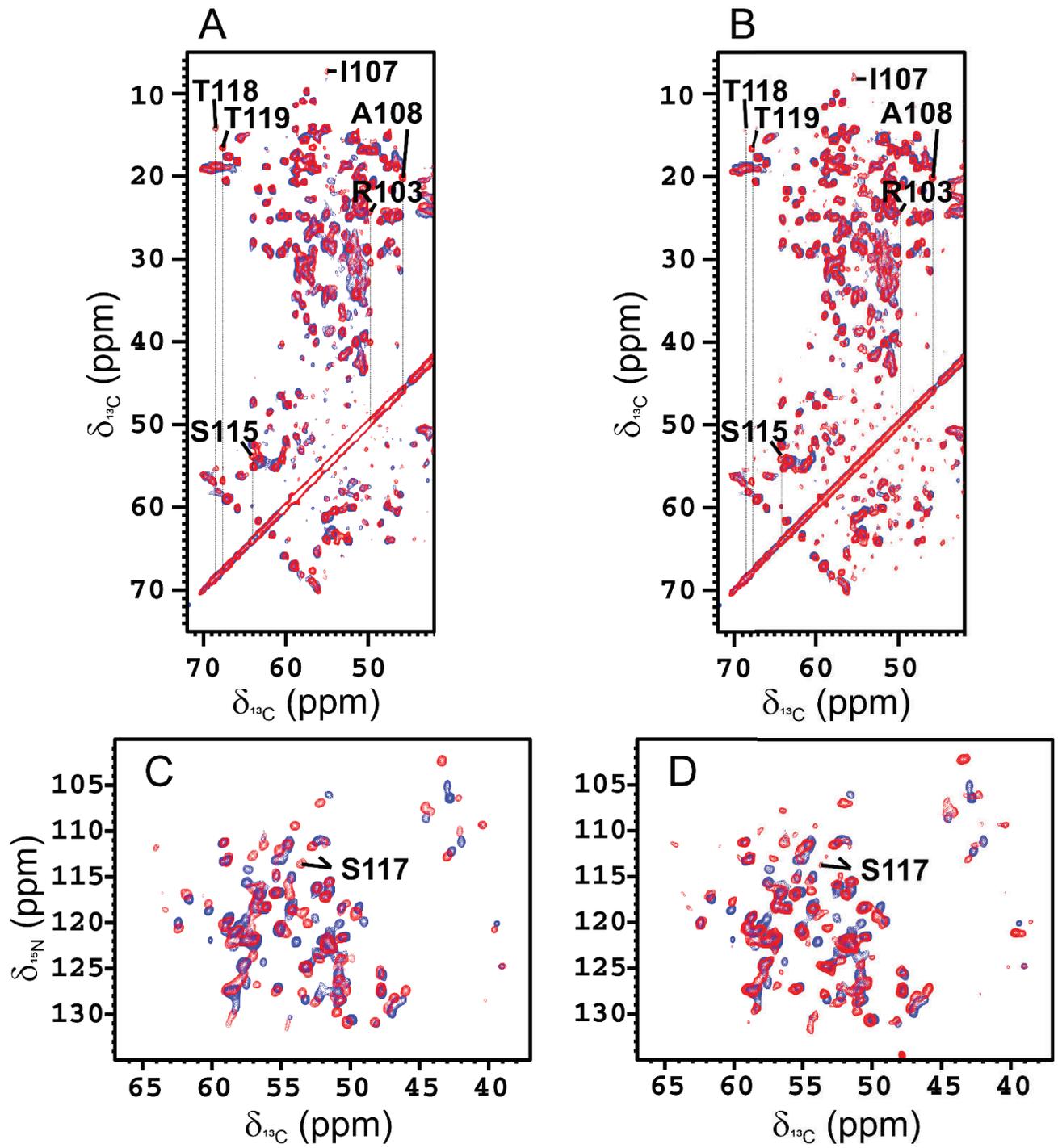


Figure 6.

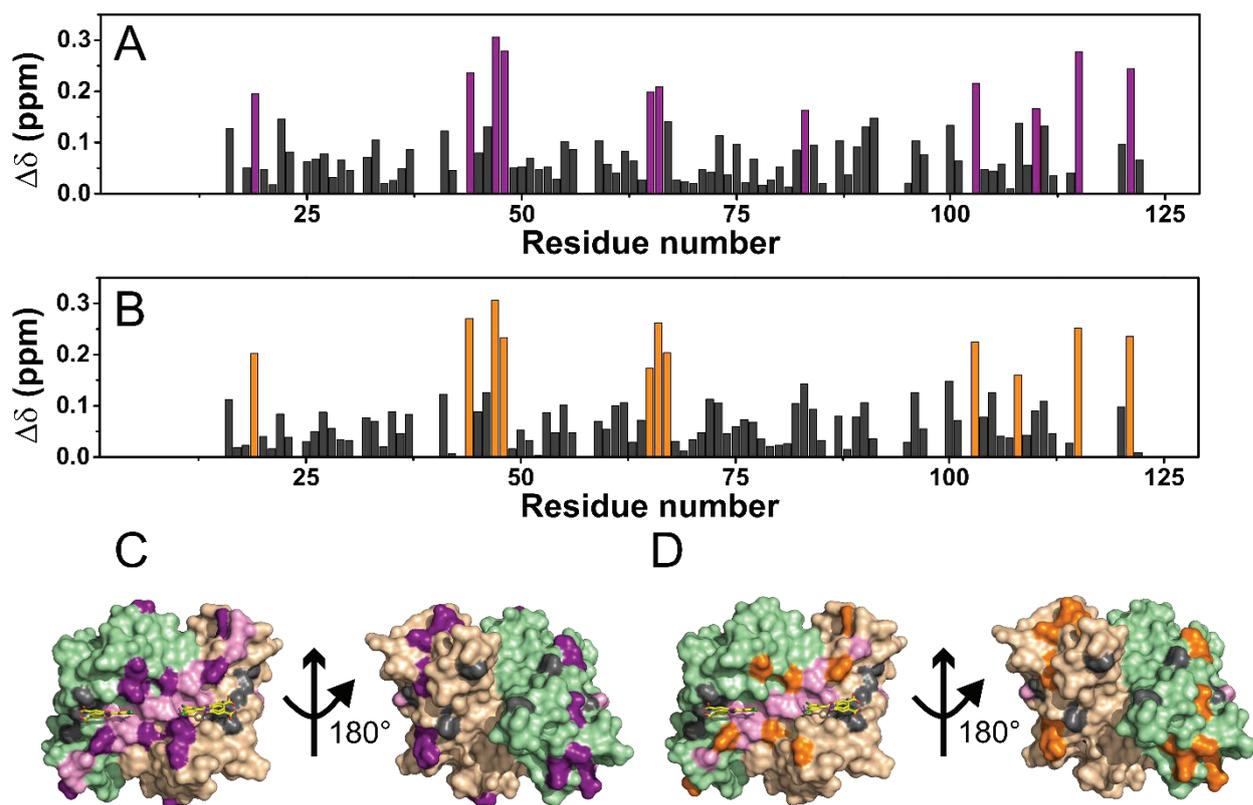


Figure 7.

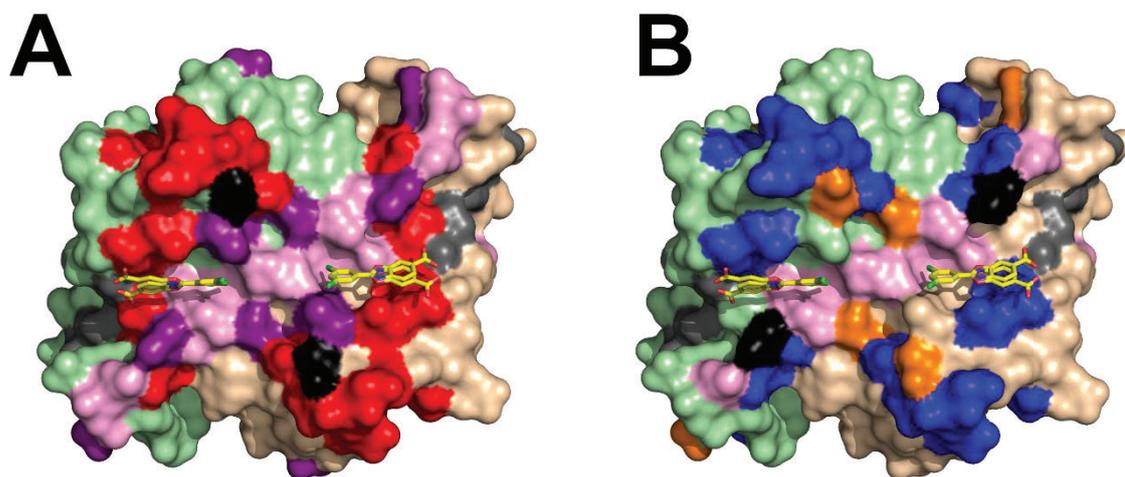


Figure 8.

References:

- [1] P. Akkapeddi, S.-A. Azizi, A. M. Freedy, P. M. S. D. Cal, P. M. P. Gois, G. J. L. Bernardes, *Chem. Sci.* **2016**, *7*, 2954–2963.
- [2] R. V. J. Chari, M. L. Miller, W. C. Widdison, *Angew Chem Int Ed Engl* **2014**, *53*, 3796–3827.
- [3] J. M. Harris, R. B. Chess, *Nat Rev Drug Discov* **2003**, *2*, 214–221.
- [4] J. Huang, B. Wu, Z. Zhou, S. Hu, H. Xu, Y. Piao, H. Zheng, J. Tang, X. Liu, Y. Shen, *Nanomedicine: Nanotechnology, Biology and Medicine* **2019**, *21*, 102058.
- [5] O. A. Mandrup, S. C. Ong, S. Lykkemark, A. Dinesen, I. Rudnik-Jansen, N. F. Dagnæs-Hansen, J. T. Andersen, L. Alvarez-Vallina, K. A. Howard, *Commun Biol* **2021**, *4*, 1–11.
- [6] A. Zorzi, S. Linciano, A. Angelini, *Medchemcomm* **2019**, *10*, 1068–1081.
- [7] A. Spada, J. Emami, J. A. Tuszynski, A. Lavasanifar, *Mol. Pharmaceutics* **2021**, *18*, 1862–1894.
- [8] A. D. AlQahtani, D. O'Connor, A. Domling, S. K. Goda, *Biomedicine & Pharmacotherapy* **2019**, *113*, 108750.
- [9] A. W. Yee, M. Aldeghi, M. P. Blakeley, A. Ostermann, P. J. Mas, M. Moulin, D. de Sanctis, M. W. Bowler, C. Mueller-Dieckmann, E. P. Mitchell, M. Haertlein, B. L. de Groot, E. Boeri Erba, V. T. Forsyth, *Nat Commun* **2019**, *10*, DOI 10.1038/s41467-019-08609-z.
- [10] M. M. Alhamadsheh, S. Connelly, A. Cho, N. Reixach, E. T. Powers, D. W. Pan, I. A. Wilson, J. W. Kelly, I. A. Graef, *Science Translational Medicine* **2011**, *3*, 97ra81-97ra81.
- [11] M. Miller, A. Pal, W. Albusairi, H. Joo, B. Pappas, M. T. Haque Tuhin, D. Liang, R. Jampala, F. Liu, J. Khan, M. Faaij, M. Park, W. Chan, I. Graef, R. Zamboni, N. Kumar, J. Fox, U. Sinha, M. Alhamadsheh, *J. Med. Chem.* **2018**, *61*, 7862–7876.
- [12] C. E. Bulawa, S. Connelly, M. DeVit, L. Wang, C. Weigel, J. A. Fleming, J. Packman, E. T. Powers, R. L. Wiseman, T. R. Foss, I. A. Wilson, J. W. Kelly, R. Labaudinière, *Proc Natl Acad Sci USA* **2012**, *109*, 9629–9634.
- [13] A. Pal, W. Albusairi, F. Liu, M. T. H. Tuhin, M. Miller, D. Liang, H. Joo, T. U. Amin, E. A. Wilson, J. S. Faridi, M. Park, M. M. Alhamadsheh, *Mol. Pharmaceutics* **2019**, *16*, 3237–3252.
- [14] S. C. Penchala, M. R. Miller, A. Pal, J. Dong, N. R. Madadi, J. Xie, H. Joo, J. Tsai, P. Batoon, V. Samoshin, A. Franz, T. Cox, J. Miles, W. K. Chan, M. S. Park, M. M. Alhamadsheh, *Nat Chem Biol* **2015**, *11*, 793–798.
- [15] F. Liu, T. Ul Amin, D. Liang, M. S. Park, M. M. Alhamadsheh, *J. Med. Chem.* **2021**, *64*, 14876–14886.
- [16] J.-P. Renaud, A. Chari, C. Ciferri, W. Liu, H.-W. Rémygy, H. Stark, C. Wiesmann, *Nat Rev Drug Discov* **2018**, *17*, 471–492.

- [17] J. M. Lamley, D. Iuga, C. Öster, H.-J. Sass, M. Rogowski, A. Oss, J. Past, A. Reinhold, S. Grzesiek, A. Samoson, J. R. Lewandowski, *J. Am. Chem. Soc.* **2014**, *136*, 16800–16806.
- [18] D. Rizzo, L. Cerofolini, S. Giuntini, L. Iozzino, C. Pergola, F. Sacco, A. Palmese, E. Ravera, C. Luchinat, F. Baroni, M. Fragai, *J. Am. Chem. Soc.* **2022**, DOI 10.1021/jacs.2c03232.
- [19] A. Mainz, T. L. Religa, R. Sprangers, R. Linser, L. E. Kay, B. Reif, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 8746–8751.
- [20] A. Mainz, S. Jehle, B. J. van Rossum, H. Oschkinat, B. Reif, *J. Am. Chem. Soc.* **2009**, *131*, 15968–15969.
- [21] U. B. le Paige, S. Xiang, M. M. R. M. Hendrix, Y. Zhang, G. E. Folkers, M. Weingarth, A. M. J. J. Bonvin, T. G. Kutateladze, I. K. Voets, M. Baldus, H. van Ingen, *Magnetic Resonance* **2021**, *2*, 187–202.
- [22] B.-M. E, H. Ct, D. V, H. St, R. B, *Angewandte Chemie (International ed. in English)* **2015**, *54*, DOI 10.1002/anie.201409393.
- [23] L. Eshun-Wilson, R. Zhang, D. Portran, M. V. Nachury, D. B. Toso, T. Löhr, M. Vendruscolo, M. Bonomi, J. S. Fraser, E. Nogales, *Proc Natl Acad Sci U S A* **2019**, *116*, 10366–10371.
- [24] H. Kato, H. van Ingen, B.-R. Zhou, H. Feng, M. Bustin, L. E. Kay, Y. Bai, *Proceedings of the National Academy of Sciences* **2011**, *108*, 12283–12288.
- [25] A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, G. M. Clore, *Science* **1991**, *253*, 657–661.
- [26] Q. Wang, H. Yang, X. Liu, L. Dai, T. Ma, J. Qi, G. Wong, R. Peng, S. Liu, J. Li, S. Li, J. Song, J. Liu, J. He, H. Yuan, Y. Xiong, Y. Liao, J. Li, J. Yang, Z. Tong, B. D. Griffin, Y. Bi, M. Liang, X. Xu, C. Qin, G. Cheng, X. Zhang, P. Wang, X. Qiu, G. Kobinger, Y. Shi, J. Yan, G. F. Gao, *Sci Transl Med* **2016**, *8*, 369ra179.
- [27] I. Bertini, C. Luchinat, G. Parigi, E. Ravera, B. Reif, P. Turano, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10396–10399.
- [28] S. Giuntini, E. Balducci, L. Cerofolini, E. Ravera, M. Fragai, F. Berti, C. Luchinat, *Angew. Chem. Int. Ed.* **2017**, *56*, 14997–15001.
- [29] S. Giuntini, L. Cerofolini, E. Ravera, M. Fragai, C. Luchinat, *Sci Rep* **2017**, *7*, 17934.
- [30] L. Cerofolini, S. Giuntini, E. Ravera, C. Luchinat, F. Berti, M. Fragai, *npj Vaccines* **2019**, *4*, 20.
- [31] L. Cerofolini, S. Giuntini, A. Carlon, E. Ravera, V. Calderone, M. Fragai, G. Parigi, C. Luchinat, *Chemistry – A European Journal* **2019**, *25*, 1984–1991.
- [32] E. Ravera, S. Ciambellotti, L. Cerofolini, T. Martelli, T. Kozyreva, C. Bernacchioni, S. Giuntini, M. Fragai, P. Turano, C. Luchinat, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 2446–2449.
- [33] L. Cerofolini, M. Fragai, E. Ravera, C. A. Diebold, L. Renault, V. Calderone, *Biomolecules* **2019**, *9*, E370.
- [34] L. Lecoq, M.-L. Fogeron, B. H. Meier, M. Nassal, A. Böckmann, *Viruses* **2020**, *12*, E1069.

- [35] T. Wiegand, D. Lacabanne, A. Torosyan, J. Boudet, R. Cadalbert, F. H.-T. Allain, B. H. Meier, A. Böckmann, *Front Mol Biosci* **2020**, *7*, 17.
- [36] M. Lu, R. W. Russell, A. J. Bryer, C. M. Quinn, G. Hou, H. Zhang, C. D. Schwieters, J. R. Perilla, A. M. Gronenborn, T. Polenova, *Nat Struct Mol Biol* **2020**, *27*, 863–869.
- [37] M. T. Eddy, T.-Y. Yu, G. Wagner, R. G. Griffin, *J Biomol NMR* **2019**, *73*, 451–460.
- [38] L. Cerofolini, S. Giuntini, A. Louka, E. Ravera, M. Fragai, C. Luchinat, *J. Phys. Chem. B* **2017**, *121*, 8094–8101.
- [39] A. Louka, I. Matlahov, S. Giuntini, L. Cerofolini, A. Cavallo, S. Pillozzi, E. Ravera, M. Fragai, A. Arcangeli, A. Ramamoorthy, G. Goobes, C. Luchinat, *Phys Chem Chem Phys* **2018**, *20*, 12719–12726.
- [40] E. Ravera, L. Cerofolini, T. Martelli, A. Louka, M. Fragai, C. Luchinat, *Sci Rep* **2016**, *6*, 27851.
- [41] T. Martelli, E. Ravera, A. Louka, L. Cerofolini, M. Hafner, M. Fragai, C. F. W. Becker, C. Luchinat, *Chemistry – A European Journal* **2016**, *22*, 425–432.
- [42] M. Fragai, C. Luchinat, T. Martelli, E. Ravera, I. Sagi, I. Solomonov, Y. Udi, *Chem. Commun. (Camb.)* **2014**, *50*, 421–423.
- [43] E. Ravera, T. Schubeis, T. Martelli, M. Fragai, G. Parigi, C. Luchinat, *J. Magn. Reson.* **2015**, *253*, 60–70.
- [44] K. Jaudzems, A. Kirsteina, T. Schubeis, G. Casano, O. Ouari, J. Bogans, A. Kazaks, K. Tars, A. Lesage, G. Pintacuda, *Angew Chem Int Ed Engl* **2021**, *60*, 12847–12851.
- [45] D. Rizzo, L. Cerofolini, A. Pérez-Ràfols, S. Giuntini, F. Baroni, E. Ravera, C. Luchinat, M. Fragai, *Anal Chem* **2021**, *93*, 11208–11214.
- [46] S. Gupta, R. Tycko, *J Biomol NMR* **2018**, *70*, 103–114.
- [47] R. Gupta, H. Zhang, M. Lu, G. Hou, M. Caporini, M. Rosay, W. Maas, J. Struppe, J. Ahn, I.-J. L. Byeon, H. Oschkinat, K. Jaudzems, E. Barbet-Massin, L. Emsley, G. Pintacuda, A. Lesage, A. M. Gronenborn, T. Polenova, *J Phys Chem B* **2019**, *123*, 5048–5058.
- [48] T. Azaïs, S. Von Euw, W. Ajili, S. Auzoux-Bordenave, P. Bertani, D. Gajan, L. Emsley, N. Nassif, A. Lesage, *Solid State Nucl Magn Reson* **2019**, *102*, 2–11.
- [49] J. Viger-Gravel, F. M. Paruzzo, C. Cazaux, R. Jabbour, A. Leleu, F. Canini, P. Florian, F. Ronzon, D. Gajan, A. Lesage, *Chemistry – A European Journal* **2020**, *26*, 8976–8982.
- [50] A. Hassan, C. M. Quinn, J. Struppe, I. V. Sergeyev, C. Zhang, C. Guo, B. Runge, T. Theint, H. H. Dao, C. P. Jaroniec, M. Berbon, A. Lends, B. Habenstein, A. Loquet, R. Kuemmerle, B. Perrone, A. M. Gronenborn, T. Polenova, *J Magn Reson* **2020**, *311*, 106680.
- [51] R. Zhang, Y. Hong, T. Ravula, Y. Nishiyama, A. Ramamoorthy, *Journal of Magnetic Resonance* **2020**, *313*, 106717.
- [52] T. Yokoyama, Y. Kosaka, M. Mizuguchi, *J Med Chem* **2014**, *57*, 1090–1096.
- [53] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 133–144.

- [54] A. Vagin, A. Teplyakov, *Acta Crystallogr D Biol Crystallogr* **2000**, *56*, 1622–1624.
- [55] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213–221.
- [56] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486–501.
- [57] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallographica Section D Biological Crystallography* **2009**, *66*, 12–21.
- [58] I. Bertini, M. Fragai, C. Luchinat, G. Parigi, *Magn. Reson. Chem.* **2000**, *38*, 543–550.
- [59] G. Parigi, E. Ravera, M. Fragai, C. Luchinat, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2021**, *124–125*, 85–98.
- [60] Y. Xu, S. Matthews, *Top Curr Chem* **2013**, *335*, 97–119.
- [61] S. Grzesiek, A. Bax, *Journal of Magnetic Resonance (1969)* **1992**, *96*, 432–440.
- [62] J. Schleucher, M. Sattler, C. Griesinger, *Angewandte Chemie International Edition in English* **1993**, *32*, 1489–1491.
- [63] L. E. Kay, G. Y. Xu, T. Yamazaki, *Journal of Magnetic Resonance, Series A* **1994**, *109*, 129–133.
- [64] M. Wittekind, L. Mueller, *Journal of Magnetic Resonance, Series B* **1993**, *101*, 201–205.
- [65] D. R. Muhandiram, L. E. Kay, *Journal of Magnetic Resonance, Series B* **1994**, *103*, 203–216.
- [66] S. Grzesiek, A. Bax, *J Biomol NMR* **1993**, *3*, 185–204.
- [67] M. Bostock, D. Nietlispach, *Concepts in Magnetic Resonance Part A* **2017**, *46A*, e21438.
- [68] W. Bermel, I. Bertini, I. C. Felli, Y.-M. Lee, C. Luchinat, R. Pierattelli, *J. Am. Chem. Soc.* **2006**, *128*, 3918–3919.
- [69] W. Bermel, I. Bertini, L. Duma, I. C. Felli, L. Emsley, R. Pierattelli, P. R. Vasos, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 3089–3092.
- [70] M. Fragai, C. Luchinat, G. Parigi, E. Ravera, *J Biomol NMR* **2013**, *57*, 155–166.
- [71] L. B. Andreas, T. Le Marchand, K. Jaudzems, G. Pintacuda, *Journal of Magnetic Resonance* **2015**, *253*, 36–49.
- [72] E. Barbet-Massin, A. J. Pell, K. Jaudzems, W. T. Franks, J. S. Retel, S. Kotelovica, I. Akopjana, K. Tars, L. Emsley, H. Oschkinat, A. Lesage, G. Pintacuda, *J. Biomol. NMR* **2013**, *56*, 379–386.
- [73] E. Barbet-Massin, A. J. Pell, J. S. Retel, L. B. Andreas, K. Jaudzems, W. T. Franks, A. J. Nieuwkoop, M. Hiller, V. Higman, P. Guerry, A. Bertarello, M. J. Knight, M. Felletti, T. Le Marchand, S. Kotelovica, I. Akopjana, K. Tars, M. Stoppini, V. Bellotti, M. Bolognesi, S. Ricagno, J. J. Chou, R. G. Griffin, H. Oschkinat, A. Lesage, L. Emsley, T. Herrmann, G. Pintacuda, *J. Am. Chem. Soc.* **2014**, *136*, 12489–12497.

- [74] A. Schuetz, C. Wasmer, B. Habenstein, R. Verel, J. Greenwald, R. Riek, A. Böckmann, B. H. Meier, *ChemBioChem* **2010**, *11*, 1543–1551.
- [75] M. BALDUS, A. T. PETKOVA, J. HERZFELD, R. G. GRIFFIN, *Molecular Physics* **1998**, *95*, 1197–1207.
- [76] M. Hong, R. G. Griffin, *J. Am. Chem. Soc.* **1998**, *120*, 7113–7114.
- [77] K. Takegoshi, S. Nakamura, T. Terao, *Chemical Physics Letters* **2001**, *344*, 631–637.
- [78] R. Keller, *The Computer Aided Resonance Assignment Tutorial (CARA)*, CANTINA Verlag, Goldau, Switzerland, **2004**.
- [79] J. Brunetti, S. Pillozzi, C. Falciani, L. Depau, E. Tenori, S. Scali, L. Lozzi, A. Pini, A. Arcangeli, S. Menichetti, L. Bracci, *Scientific Reports* **2015**, *5*, 17736.
- [80] J. Brunetti, S. Piantini, M. Fragai, S. Scali, G. Cipriani, L. Depau, A. Pini, C. Falciani, S. Menichetti, L. Bracci, *Molecules* **2020**, *25*, E1088.
- [81] E. Ravera, G. Parigi, A. Mainz, T. L. Religa, B. Reif, C. Luchinat, *J Phys Chem B* **2013**, *117*, 3548–3553.
- [82] J. García de la Torre, M. L. Huertas, B. Carrasco, *Journal of Magnetic Resonance* **2000**, *147*, 138–146.
- [83] K. Liu, J. W. Kelly, D. E. Wemmer, *J. Mol. Biol.* **2002**, *320*, 821–832.
- [84] J. Oroz, J. H. Kim, B. J. Chang, M. Zweckstetter, *Nat. Struct. Mol. Biol.* **2017**, *24*, 407–413.
- [85] B. I. Leach, X. Zhang, J. W. Kelly, H. J. Dyson, P. E. Wright, *Biochemistry* **2018**, *57*, 4421–4430.
- [86] Y.-T. Liu, Y.-J. Yen, F. Ricardo, Y. Chang, P.-H. Wu, S.-J. Huang, K.-P. Lin, T.-Y. Yu, *Ann Clin Transl Neurol* **2019**, *6*, 1961–1970.
- [87] A. Corazza, G. Verona, C. A. Waudby, P. P. Mangione, R. Bingham, I. Uings, D. Canetti, P. Nocerino, G. W. Taylor, M. B. Pepys, J. Christodoulou, V. Bellotti, *J. Med. Chem.* **2019**, *62*, 8274–8283.
- [88] S. Grzesiek, A. Bax, G. M. Clore, A. M. Gronenborn, J. S. Hu, J. Kaufman, I. Palmer, S. J. Stahl, P. T. Wingfield, *Nat. Struct. Biol.* **1996**, *3*, 340–345.

4.6.2 Supplementary Information

Transthyretin shows new possibilities for the structure-based molecular design of protein-drug conjugates

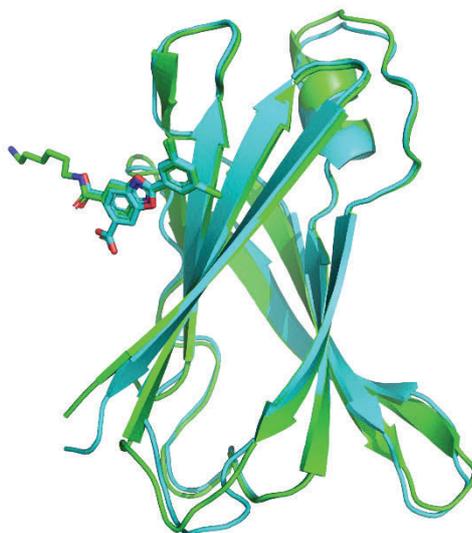


Figure S1. Superposition between the structure presented in this work and 3TCT. One monomer for both structure is presented for clarity.

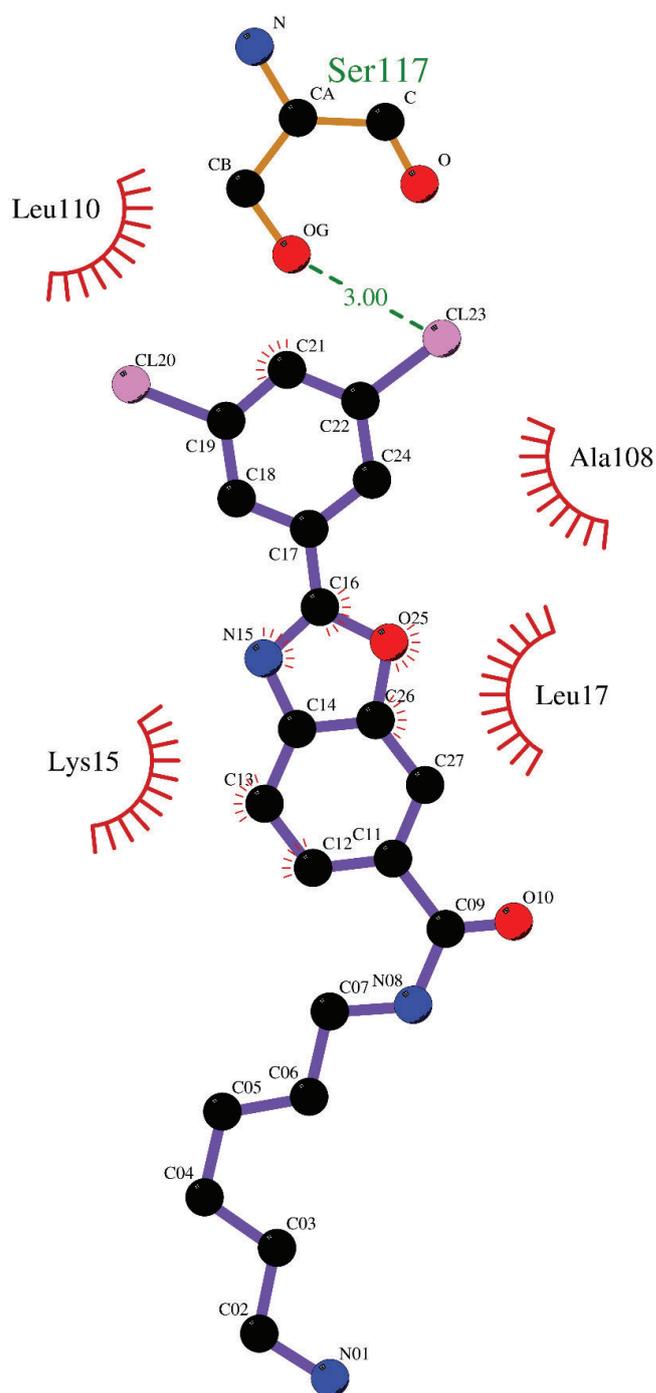


Figure S2. Interactions between the protein and compound 2 observed in the X-ray structure solved in the present paper.

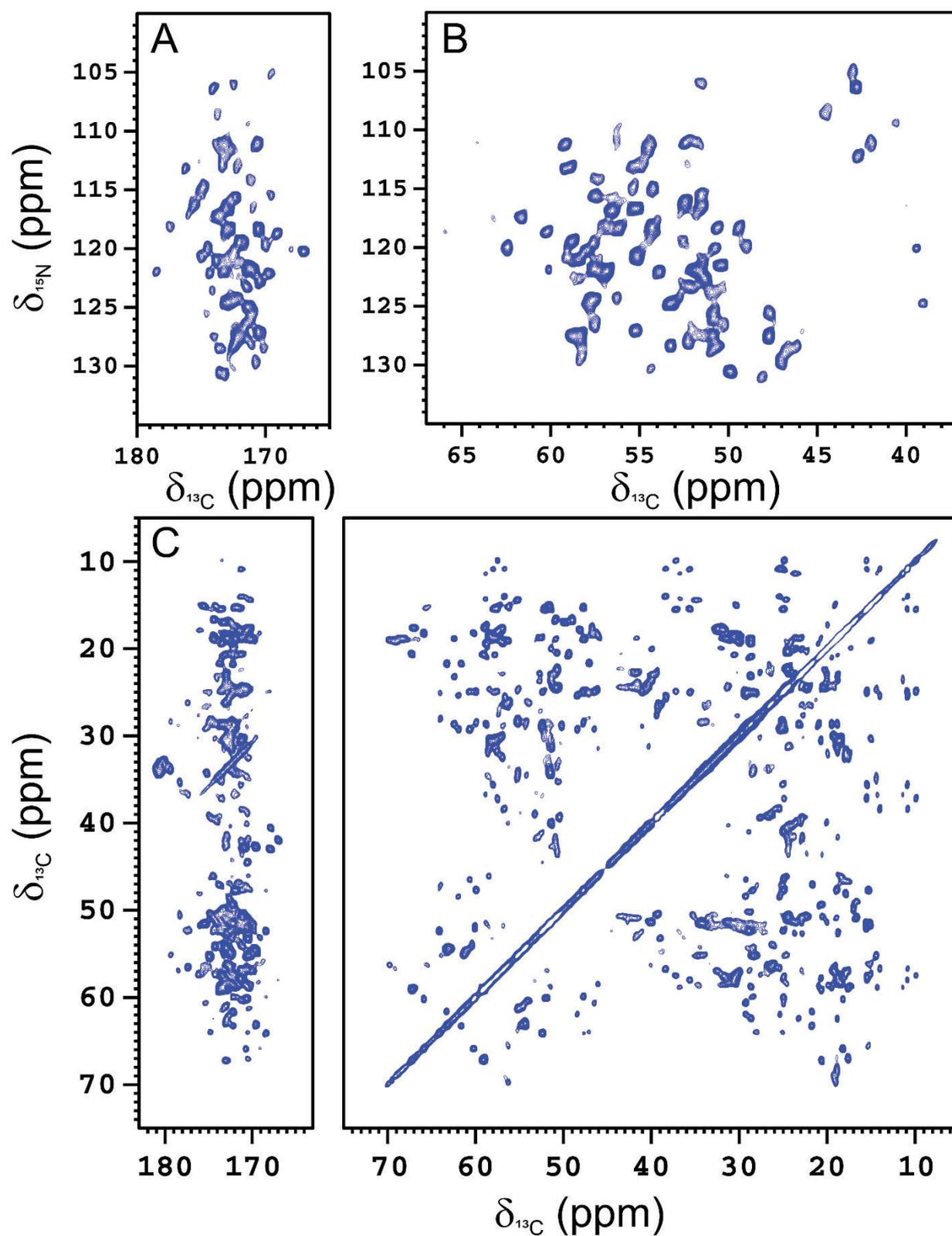


Figure S4. 2D ^{15}N ^{13}C NCO (A), ^{15}N ^{13}C NCA (B) and ^{13}C - ^{13}C DARR (C) spectra acquired after rehydration on a sample of freeze-dried $[\text{U-}^{13}\text{C}, ^{15}\text{N}]$ TTR. Spectra were acquired on a spectrometer operating at 800 MHz (^1H Larmor frequency) with MAS of 14 kHz and temperature of ~ 290 K.

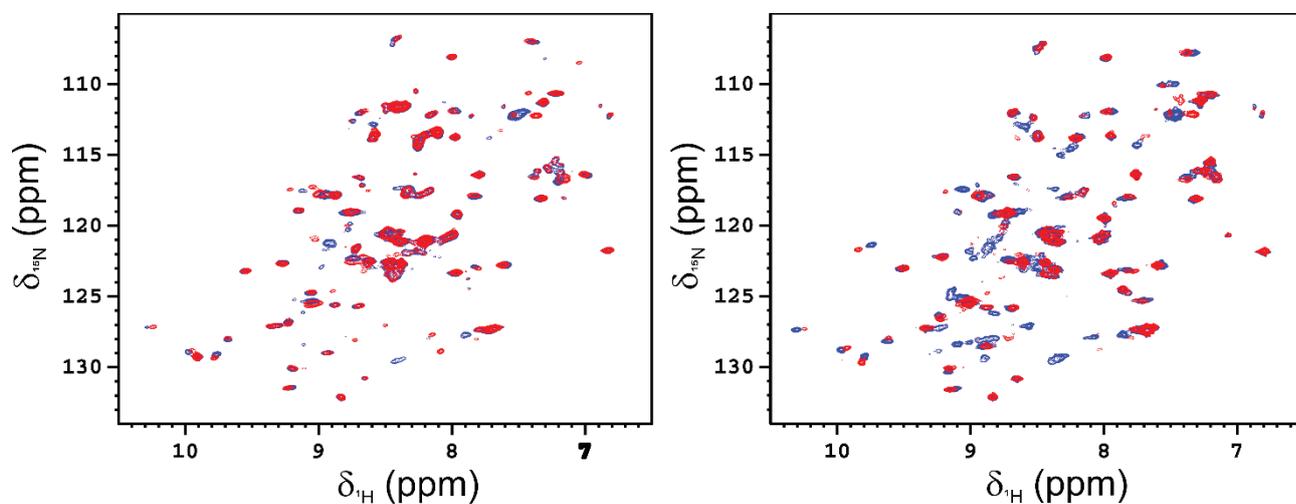


Figure S5. 2D ^1H ^{15}N TROSY-HSQC of free TTR (blue) and TTR in the presence of tafamidis (red, left) and in the presence of Taf-Ptx (red, right). The spectra were acquired on an NMR spectrometer operating at 950 MHz (^1H Larmor frequency) and 310 K.

Table S1.

Table S1. Data processing and refinement statistics for the X-ray structure solved in the present paper.

	TTR
Wavelength	1.541
Resolution range	35.47 - 1.6 (1.657 - 1.6)
Space group	P 21 21 2
Unit cell	43.48 85.34 63.79 90 90 90
Total reflections	384543 (15355)
Unique reflections	31504 (2783)
Multiplicity	12.2 (5.5)
Completeness (%)	98.21 (88.60)
Mean I/sigma(I)	22.29 (2.08)
Wilson B-factor	16.87
R-merge	0.08045 (0.9225)
R-meas	0.08391 (1.016)
R-pim	0.02336 (0.4066)
CC1/2	1 (0.668)
CC*	1 (0.895)
Reflections used in refinement	31492 (2783)
Reflections used for R-free	1574 (139)
R-work	0.1783 (0.2452)
R-free	0.2035 (0.2458)
CC(work)	0.968 (0.839)
CC(free)	0.977 (0.802)
Number of non-hydrogen atoms	2068
macromolecules	1785
ligands	54
solvent	229
Protein residues	231
RMS(bonds)	0.014

RMS(angles)	1.28
Ramachandran favored (%)	98.68
Ramachandran allowed (%)	1.32
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	3.92
Average B-factor	23.91
macromolecules	22.45
ligands	31.26
solvent	33.63

Table S2. Best fit values of the ^1H NMRD profiles of wild-type TTR with and without tafamidis. Unless differently specified, the values refer to the profiles collected at 25 °C.

	wild type TTR	wild type TTR + Tafamidis	
α	0.34 ± 0.01	0.34 ± 0.01	s^{-1}
β	1.9 ± 0.7		10^7 s^{-2}
c_1	0.005 ± 0.002	0.003 ± 0.002	
τ_1	350 ± 100		10^{-9} s
c_2	0.24 ± 0.07	0.27 ± 0.08	
τ_2	30 ± 3		10^{-9} s
c_3	0.75 ± 0.07	0.73 ± 0.08	
τ_3	4 ± 2		10^{-9} s
<hr/>			
$\tau_2(15\text{ °C})$	40 ± 1	40 ± 1	10^{-9} s
$\tau_2(37\text{ °C})$	19 ± 1	22 ± 1	10^{-9} s

5. Conclusions and perspectives

The relationship between structure and binding capability toward specific RNA oligonucleotides is of fundamental importance to reveal the biological role of the RNA binding proteins (RBP), to design potential drug candidates and mutants for application in synthetic biology and bio-analytics. The research activity of my PhD project here described, was focused on the expression of bacterial and human RBPs and on the biophysical characterization of their binding to selected physiological and engineered RNA sequences. Concerning the RBPs, as a part of a collaborative project, we have discovered, expressed and characterized an AB-Elavl RRM-containing protein from the ELAVL family in *Acinetobacter baumannii*. After the development of an optimized experimental protocol, I obtained crystals of the protein suitable for x-ray diffraction. The experimental structure confirmed the presence of a typical RRM fold in the protein, and the NMR data recorded in solution proved its ability to bind RNA AU-rich oligonucleotides. These results open the way to reveal the biological function of ELAVL proteins in *Acinetobacter baumannii* and possible new drug targets to face this pathogen.

More closely related to the RNAct project is the research activity carried out on the human Musashi-1 protein. This RBP play an important role in stem cell maintenance, tumorigenesis and is a potential drug target. The results here reported demonstrate the capability of each isolated domain of MSI-1 (RRM-1 and RRM-2) to bind the RNA sequences containing the consensus binding motif G/AU₁₋₃AGU and also the simple UAG motif placed in linear RNA sequences or in folded RNA structures with high affinity. We have also demonstrated that the two domains within MSI-1 compete for the UAG motif decreasing the selectivity and leading to the presence of multiple species in solution. However, when the AUG recognition motif is involved in the formation of a double RNA strands the affinity for the two domain is probably lost. These results have provided the necessary insights for the analysis in silico by the computational tool RRMScorer, of the affinity of RRM-2 for the AUG motif and to design mutants of the MSI-1 an improved selectivity. The decreased affinity of RRM-2 for the UAG motif induced by the mutations in this domain within the MSI-1, results in a reduction of heterogeneity in the solutions containing the mutated protein and the selected RNA oligonucleotides. The same computational tool has been used to identify RNA motif with high affinity for the RRM-2 domain. In this respect, is interesting the motif CAG. The replacement of the AUG motif with CAG sequence on the two previously investigated RNA strands restores the affinity of the mutated RRM-2 domain for the two mutated RNA oligonucleotides leading to a higher selectivity of the mutated MSI-1 in the recognition of the mutated RNA strands, and to a lower heterogeneity in solution. Collectively, the NMR characterization proves the reliability and the quality of the prediction provided by RRMScorer, suggesting MSI-1 and its RNA partners are promising candidates to develop new mutants for application in synthetic biology and bio-analytics.

Always within the RNAct project and related to RBPs, is the research focused on the development of a new application for the Ligand technology for the monitorization of RNA-protein interactions in living bacteria cells in real time. An adaptation of the dish coating combined with the design of a reporter system with RBPs, this work has allowed to track and characterize gene reporter assays over time in a relevant physiologic environment. At the same time, the development of this new methodology opens up tons of possibilities to exploit LigandTracer to regulate gene expression throughout RBPs and small molecules.

Moreover, during this PhD project new strategies for the characterization of protein-based drugs and formulations by NMR spectroscopy have been developed and tested. To speed-up the development of these new NMR-based strategies, we have used the therapeutic protein ANSII and the potential drug carrier TTR, already available in the lab. At this regard, solid-state NMR (SSNMR) spectroscopy has been used to investigate the preservation of the higher order structure (HOS) of the proteins in heterogenous materials suitable for drug release and bioprinting applications. In particular, the two ANSII and TTR have been embedded in hydrogel matrices by using a simple sample preparation protocol. As reported in the manuscript the recorded NMR data prove that SSNMR can be employed to evaluate the preservation of the protein folding providing new perspectives for the development and characterization of new biotherapeutics and pharmaceutical formulations.

Finally, SSNMR has been used within an integrated approach to develop a new strategy to design protein-drug conjugates (PDC). In this research Transthyretin was selected as potential carrier for a paclitaxel derivative conjugated to Tafamidis. Our results highlight the benefits and show the potential of using an integrative approach using different complementary structural techniques like SSNMR and X-ray to develop PDC.

6. References

1. Lorkovic, Z. From Structure to Function of RNA Binding Domains. *RNA Binding Proteins* 149–170 (2020) doi:10.1201/9781498713368-15.
2. Gebauer, F., Schwarzl, T., Valcárcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat Rev Genet* **22**, 185–198 (2021).
3. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* **19**, 327–341 (2018).
4. Cléry, A., Blatter, M. & Allain, F. H. T. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290–298 (2008).
5. Maris, C., Dominguez, C. & Allain, F. H. T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal* vol. 272 2118–2131 Preprint at <https://doi.org/10.1111/j.1742-4658.2005.04653.x> (2005).
6. Birney, E., Kumar, S. & Krainer, A. R. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* **21**, 5803–5816 (1993).
7. Oliveira, C., Faoro, H., Alves, L. R. & Goldenberg, S. RNA-binding proteins and their role in the regulation of gene expression in trypanosoma cruzi and saccharomyces cerevisiae. *Genet Mol Biol* **40**, 22–30 (2017).
8. Daubner, G. M., Cléry, A. & Allain, F. H. T. RRM-RNA recognition: NMR or crystallography...and new findings. *Curr Opin Struct Biol* **23**, 100–108 (2013).
9. Joel Roca-Martínez^{1, 2}, Hrishikesh Dhondge³, Michael Sattler^{4, 5}, Wim Vranken^{1, 2}. Deciphering the RRM-RNA recognition code: A computational analysis (Manuscript submitted). *PLoS Comput Biol*.
10. Assoni, G. *et al.* HuR-targeted agents: An insight into medicinal chemistry, biophysical, computational studies and pharmacological effects on cancer models. *Adv Drug Deliv Rev* **181**, 114088 (2022).
11. Lorkovic, Z. From Structure to Function of RNA Binding Domains. *RNA Binding Proteins* 149–170 (2020) doi:10.1201/9781498713368-15.
12. Corley, M. *et al.* mechanisms. **78**, 9–29 (2021).
13. Bono F, Ebert J, Unterholzner L, Guttler T, I. & E, E. & C. Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO report* **5**, 304–310 (2004).
14. Agrawal, S. *et al.* RNA recognition motifs of disease-linked RNA-binding proteins contribute to amyloid formation. *Sci Rep* **9**, 1–12 (2019).
15. Vanderweyde, T., Youmans, K., Liu-Yesucevitz, L. & Wolozin, B. Role of stress granules and RNA-binding proteins in neurodegeneration: A mini-review. *Gerontology* **59**, 524–533 (2013).
16. Lukong, K. E., Chang, K. wei, Khandjian, E. W. & Richard, S. RNA-binding proteins in human genetic disease. *Trends in Genetics* **24**, 416–425 (2008).
17. Hermann, T. Strategies for the design targeting RNA and RNA - Protein complexes. *Angewandte Chemie - International Edition* **39**, 1890–1904 (2000).

18. Mohibi, S., Chen, X. & Zhang, J. Cancer the ‘RBP’ eutics–RNA-binding proteins as therapeutic targets for cancer. *Pharmacol Ther* **203**, (2019).
19. Nahalka, J. The role of the protein–RNA recognition code in neurodegeneration. *Cellular and Molecular Life Sciences* **76**, 2043–2058 (2019).
20. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Hum Genet* **135**, 851–867 (2016).
21. Horisawa, K., Imai, T., Okano, H. & Yanagawa, H. The Musashi family RNA-binding proteins in stem cells. *Biomol Concepts* **1**, 59–66 (2010).
22. Tan, A. Y. & Manley, J. L. The TET family of proteins: Functions and roles in disease. *J Mol Cell Biol* **1**, 82–92 (2009).
23. Clingman, C. C. *et al.* Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *Elife* **2014**, (2014).
24. Lal, P. *et al.* Regulation of HuR structure and function by dihydrotanshinone-I. *Nucleic Acids Res* **45**, 9514–9527 (2017).
25. Ghidini, A., Cléry, A., Halloy, F., Allain, F. H. T. & Hall, J. RNA-PROTACs: Degraders of RNA-Binding Proteins. *Angewandte Chemie - International Edition* **60**, 3163–3169 (2021).
26. Ke, Y. *et al.* PARP1 promotes gene expression at the post-transcriptional level by modulating the RNA-binding protein HuR. *Nat Commun* **8**, (2017).
27. Rigo, F., Hua, Y., Krainer, A. R. & Frank Bennett, C. Antisense-based therapy for the treatment of spinal muscular atrophy. *Journal of Cell Biology* **199**, 21–25 (2012).
28. Schultz, C. W. *et al.* protein HuR. **11**, 1–31 (2021).
29. Mascini, M., Palchetti, I. & Tombelli, S. Nucleic acid and peptide aptamers: Fundamentals and bioanalytical aspects. *Angewandte Chemie - International Edition* **51**, 1316–1332 (2012).
30. Lucchesi, C. A., Zhang, J., Ma, B., Chen, M. & Chen, X. Disruption of the RBM38-eIF4E complex with a synthetic peptide PEP8 increases p53 expression. *Cancer Res* **79**, 807–818 (2019).
31. Abdelmohsen, K. *et al.* Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol* **14**, 361–369 (2017).
32. Ohyama, T. *et al.* Structure of Musashi1 in a complex with target RNA: The role of aromatic stacking interactions. *Nucleic Acids Res* **40**, 3218–3231 (2012).
33. Miyanoiri, Y. *et al.* Origin of Higher Affinity to RNA of the N-terminal RNA-binding Domain than That of the C-terminal One of a Mouse Neural Protein, Musashi1, as Revealed by Comparison of Their Structures, Modes of Interaction, Surface Electrostatic Potentials, and Backbone. *Journal of Biological Chemistry* **278**, 41309–41315 (2003).
34. Iwaoka, R. *et al.* Structural Insight into the Recognition of r(UAG) by Musashi-1 RBD2, and Construction of a Model of Musashi-1 RBD1-2 Bound to the Minimum Target RNA. *Molecules* **22**, 1–16 (2017).
35. Nagata, T. *et al.* Structure, backbone dynamics and interactions with RNA of the C-terminal RNA-binding domain of a mouse neural RNA-binding protein, Musashi1. *J Mol Biol* **287**, 315–330 (1999).

36. Sakakibara, S. I. *et al.* Mouse-Musashi-1, a neural RNA-Binding protein highly enriched in the mammalian CNS stem cell. *Dev Biol* **176**, 230–242 (1996).
37. Kurihara, Y. *et al.* Structural properties and RNA-binding activities of two RNA recognition motifs of a mouse neural RNA-binding protein, mouse-Musashi-1. *Gene* **186**, 21–27 (1997).
38. Kawahara, H. *et al.* Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. *Journal of Cell Biology* **181**, 639–653 (2008).
39. Ruth Zearfoss, N. *et al.* A conserved three-nucleotide core motif defines musashi RNA binding specificity. *Journal of Biological Chemistry* **289**, 35530–35541 (2014).
40. Abreu, R. D. S. *et al.* Genomic Analyses of Musashi1 Downstream Targets Show a Strong Association with Cancer-related Processes * □. **284**, 12125–12135 (2009).
41. Bley, N. *et al.* Musashi-1—a stemness RBP for cancer therapy? *Biology (Basel)* **10**, 1–16 (2021).
42. Okano, H. *et al.* Function of RNA-binding protein Musashi-1 in stem cells. *Exp Cell Res* **306**, 349–356 (2005).
43. Forouzanfar, M. *et al.* Intracellular functions of RNA-binding protein, Musashi1, in stem and cancer cells. *Stem Cell Res Ther* **11**, 1–10 (2020).
44. Horisawa, K., Imai, T., Okano, H. & Yanagawa, H. 3'-Untranslated region of doublecortin mRNA is a binding target of the Musashi1 RNA-binding protein. *FEBS Lett* **583**, 2429–2434 (2009).
45. Kudinov, A. E., Karanicolas, J., Golemis, E. A. & Bumber, Y. Musashi RNA-binding proteins as cancer drivers and novel therapeutic targets. *Clinical Cancer Research* **23**, 2143–2153 (2017).
46. Chen, H. Y. *et al.* Musashi-1 promotes stress-induced tumor progression through recruitment of AGO2. *Theranostics* **10**, 201–217 (2020).
47. Ito, T. *et al.* HHS Public Access. **466**, 765–768 (2011).
48. Jones, R. & Zweier. 基因的改变 NIH Public Access. *Bone* **23**, 1–7 (2014).
49. Samson, M. L. Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins. *BMC Genomics* **9**, 1–11 (2008).
50. Good, P. J. A conserved family of elav-like genes in vertebrates. *Proc Natl Acad Sci U S A* **92**, 4557–4561 (1995).
51. Fialcowitz-White, E. J. *et al.* Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. *Journal of Biological Chemistry* **282**, 20948–20959 (2007).
52. Kim, H. S. *et al.* Different modes of interaction by TIAR and HuR with target RNA and DNA. *Nucleic Acids Res* **39**, 1117–1130 (2011).
53. Wang, H. *et al.* The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr D Biol Crystallogr* **69**, 373–380 (2013).
54. Díaz-Quintana, A., García-Mauriño, S. M. & Díaz-Moreno, I. Dimerization model of the C-terminal RNA Recognition Motif of HuR. *FEBS Lett* **589**, 1059–1066 (2015).

55. Lixa, C. *et al.* Oligomeric transition and dynamics of RNA binding by the HuR RRM1 domain in solution. *J Biomol NMR* **72**, 179–192 (2018).
56. Pabis, M. *et al.* HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs. *Nucleic Acids Res* **47**, 1011–1029 (2019).
57. Scheiba, R. M. *et al.* The C-terminal RNA binding motif of HuR is a multi-functional domain leading to HuR oligomerization and binding to U-rich RNA targets. *RNA Biol* **11**, 1250–1261 (2014).
58. Ripin, N. *et al.* Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM. *Proc Natl Acad Sci U S A* **116**, 2935–2944 (2019).
59. Srikantan, S. & Gorospe, M. HuR function in disease. *Frontiers in Bioscience* **17**, 189–205 (2012).
60. Williams, M. A. *Protein-Ligand Interactions*.
61. Encarnaç o, J. C., Schulte, T., Achour, A., Bj rkelund, H. & Andersson, K. Detecting ligand interactions in real time on living bacterial cells. *Appl Microbiol Biotechnol* **102**, 4193–4201 (2018).
62. Cordes, R. M., Sims, W. B. & Glatz, C. E. Precipitation of Nucleic Acids with Poly (ethyleneimine). *Biotechnol Prog* **6**, 283–285 (1990).
63. Koutcher, J. A. & Burt, C. T. Principles of nuclear magnetic resonance. *Journal of Nuclear Medicine* **25**, 101–111 (1984).
64. Szak acs, Z. & S anta, Z. *NMR Methodological Overview. Anthropic Awareness: The Human Aspects of Scientific Thinking in NMR Spectroscopy and Mass Spectrometry* (2015). doi:10.1016/B978-0-12-419963-7.00007-9.
65. Fattori, J., Rodrigues, F. H. S., Pontes, J. G. M., Paula Esp ndola, A. & Tasic, L. *Monitoring Intermolecular and Intramolecular Interactions by NMR Spectroscopy. Applications of NMR Spectroscopy* vol. 3 (2015).
66. Mahler, H. C., Friess, W., Grauschopf, U. & Kiese, S. Protein aggregation: Pathways, induction factors and analysis. *J Pharm Sci* **98**, 2909–2934 (2009).
67. Katsamba, P. S., Park, S. & Laird-Offringa, I. A. Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods* **26**, 95–104 (2002).
68. Nguyen, H. H., Park, J., Kang, S. & Kim, M. Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Switzerland)* **15**, 10481–10510 (2015).
69. Cl ery, A. & Allain, F. H.-T. FROM STRUCTURE TO FUNCTION OF RNA BINDING DOMAINS. (2013).
70. Corley, M., Burns, M. C. & Yeo, G. W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol Cell* **78**, 9–29 (2020).
71. Gebauer, F., Schwarzl, T., Valc arcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat Rev Genet* **22**, 185–198 (2021).
72. Imai, T. *et al.* The Neural RNA-Binding Protein Musashi 1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. *Mol Cell Biol* **21**, 3888–3900 (2001).

-
73. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A. & Torchia, D. A. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. *Journal of Magnetic Resonance (1969)* **97**, 359–375 (1992).
 74. Ingram, D. J. E. Nuclear magnetic resonance (part ii). *Contemp Phys* **7**, 103–121 (1965).
 75. Ulrich Rant, Kenji Arinaga, Shozo Fujita, Naoki Yokoyama, Gerhard Abstreiter, and M. T. Structural Properties of Oligonucleotide Monolayers on Gold Surfaces Probed by Fluorescence Investigations. *Langmuir* **20**, 10086–10092 (2004).
 76. Kanga, H. S. *et al.* An autoinhibitory intramolecular interaction proofreads RNA recognition by the essential splicing factor U2AF2. *Proc Natl Acad Sci U S A* **117**, 7140–7149 (2020).

7. Acknowledgements

First of all, I would like to thank my supervisors, Dr. Tommaso Martelli and Prof. Marco Fragai for letting me have the opportunity to work in their company and research lab and for all the scientific knowledges acquired under their supervision.

Di GiottoBiotech, vorrei ringraziare Stefano, Laura e Tommaso per essere stati sempre molto disponibili per tutto ciò di cui avevo bisogno, dalla burocrazia alla supervisione del lavoro. Fin dal primo momento mi avete fatto sentire la benvenuta e parte del gruppo.

Inoltre vorrei ringraziare Tania e Mercia. Grazie per l'infinita pazienza, per l'aiuto e gli insegnamenti che mi avete dato in lab. Mi avete supportata e aiutata nella realizzazione di questo complesso progetto. Grazie anche per avermi fatto ridere ogni giorno e aver creato un bellissimo ambiente in lab.

Del CERM, sono innumerevoli le persone che mi hanno accompagnato in questo percorso. Stefano, vorrei ringraziarti per l'aiuto iniziale che mi hai dato in lab. Mi hai insegnato tutte le basi per lavorare bene e con efficienza. Vorrei fare anche un speciale ringraziamento a Linda, per l'enorme pazienza e aiuto dato fine alla fine.

A Maxime per avermi integrato nel gruppo e presentato quelli che oggi sono amici che porto nel cuore. A Dafne, Sara, Lucia, Alessio, Lorenzo, Luis, Francesca, Vanessa, e tutti gli altri per tutte le gioie condivise dentro e fuori il lab, e per tutte le "pause caffè e sigaretta".

A le mie amiche del "trio disagate del CERM": Letizia e Vale. Per tutte le discussioni scientifiche con le nostre proteine problematiche, per il sostegno mutuo e pioggia d'idee che abbiamo avuto.

A Giovanni, sei stato una persona fondamentale per me e la tua amicizia è una gioia. Mi hai aiutata quando sono arrivata in questo nuovo paese, mi hai ascoltato e aiutato quando le cose in lab non andavano bene e abbiamo condiviso mille esperienze e sorrisi. Per tutto questo e mille altre cose, rimarrai sempre nel mio cuore.

Uno speciale ringraziamento alle mie amiche del cuore, la stanza 2 (Deborah, Lucrezia e Silvia) e a Bora. Non so cosa avrei fatto senza di voi. Voi mi avete sopportato, ascoltato nelle innumerevoli occasioni per qualsiasi cosa in qualsiasi momento, mi avete accompagnato in tutte le gioie e dolori, e siete quelle che credete più in me. Abbiamo condiviso mille esperienze, belle pause al lavoro e innumerevoli pranzi, cene, padel e gite fuori porta. Mi avete accompagnata durante questi tre anni e siete diventati una famiglia per me. Voi siete "ESTUPENDE".

Grazie a tutti voi, per avermi fatto considerare Firenze la mia casa.

To my RNAct family, the best consortium one could have ended up to. Thank you or all your friendship and support throughout these three years. And for all the experiences we have shared.

A special thanks to other two members of the trio LaLaLa Roswi and Hrishi. You have been true friends to me, my personal cheerleaders and the ones who have make me laugh the most, even in stressful days.

To Isaure and Arha, thanks for everything you did during my short secondment in Nancy, France.

A special thanks to Jos, not only you have welcomed me in your company for three months and have taught me a lot, but you have also been an amazing person to work with. Without you and your commitment and dedication I would not have been able to write this thesis as it is. For all of it, I am truly grateful.

Y sobretodo, gracias a ti Guille. Creo que es justo decir que sin ti mi tesis no hubiera sido posible. Tu amistad, dedicación, compromiso, y tu profesionalidad han hecho de esta tesis lo que es hoy. Has sido un gran apoyo durante los 3 años pero por encima de todo, un gran amigo, y por esto, te estoy muy agradecida.

To all the people I met during my 3-month stay in Uppsala, Sweden that welcomed me with open arms and help me grow as a person and as a scientist. Specially to all the people at Ridgeview and at the Chemistry department. And a special thanks to Nadine, Livia, Max, Emily, Javi and Tiscar.

E a le mie coinquiline Laura e Kri per ascoltarmi e per i vostri consigli. Grazie anche per il raccozzamento durante i numerosi fine settimana e notti che ho dovuto lavorare. Ma soprattutto, per la vostra compagnia e amicizia.

I per últim, gràcies als de casa. Per l'ajuda muntant la tesi i per escoltar-me totes les vegades que parlava de la feina i de la meva vida a Itàlia. I gràcies àvia per ser sempre la que creu més amb mi. I sobretot, gràcies a les tres que m'heu aguantat més, a la Marta, a la Etna i a la Iman. Sense vosaltres aquesta experiència i la meva vida no serien igual. Vosaltres sou la meva roca, el meu lloc de confort on anar i sentir-me sempre bé. Vosaltres m'heu aguantat, m'heu consolat, m'heu fet riure, heu confiat en mi i m'heu estimat. No podria haver demanat millors amigues que vosaltres tres. Vosaltres cada dia em feu créixer professional i personalment. Ho sou tot per mi.