

**Versatile use of advanced NMR techniques for the identification of compounds able to  
interfere with ELAV protein–mRNA complexes**

Dissertation  
zur Erlangung des Grades  
des Doktors der Naturwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät  
der Universität des Saarlandes

von  
Serena Della Volpe

Saarbrücken  
2020

**Tag des Kolloquiums:** 20. Februar 2020

**Dekan:** Prof. Dr. Guido Kickelbick

**Berichterstatter:** Prof. Dr. Anna K. H. Hirsch

Dr. Gregor Fuhrmann

Prof. Dr. Simona Collina

**Weitere Mitglieder:** Prof. Dr. Roccaldo Sardella

Dott. Silvia Franchini

**Vorsitz:** Prof. Dr. Patrizia Santi

**Versatile use of advanced NMR techniques for the identification of compounds able to interfere with ELAV protein–mRNA complexes**

PhD in Chemical and Pharmaceutical Sciences

University of Pavia

XXXII cycle

Academic year 2016/2017–2018/2019



**Tutor**

**Prof. Simona Collina**

**Co-tutors**

**Prof. Anna K. H. Hirsch**

**Dr. Francesca Vasile**

**PhD Student**

**Serena Della Volpe**



## ***Index***

Short Abstracts	1
English – Short Abstract	1
Deutsch – Kurzzusammenfassung	2
Abstracts	3
English - Abstract	3
Italiano - Riassunto	5
Deutsch - Zusammenfassung	7
SECTION 1: Versatile use of advanced NMR techniques for the identification of compounds able to interfere with ELAV protein–mRNA complexes	9
1. Introduction	9
1.1 RNA-binding proteins	9
1.2 ELAV/Hu protein family	12
1.3 Compounds able to interfere with ELAV–mRNA complexes	13
1.4 ELAV–RNA interface	16
1.4.1 ELAV–RNA co-crystals	17
1.4.2 Studies on ELAV peptides (RRM1/2 RNP1/2)–RNA	18
1.5 Postulated binding site for small molecules: preliminary docking of literature compounds on subtype HuR	20
2. Aim	22
3. Methodologies and Strategies	23
3.1 NMR techniques to study small molecule–macromolecule interactions	23
3.1.1 Ligand-observed NMR: Saturation Transfer Difference (STD) NMR	24
3.1.2 Ligand-observed NMR: Differential epitope mapping (DEEP) by STD NMR	26
3.2 Other biophysical approaches	29
3.2.1 Surface Plasmon Resonance	32
3.2.2 Thermal Shift Assay	33
3.3 In silico methodologies: Docking and Molecular Dynamics	35
3.4 Drug Discovery Strategies	35
3.4.1 Structure-based ligand design	36
3.4.2 Fragment-based ligand design	36
3.4.3 Multi-component Reactions	37
3.4.4 Dynamic Combinatorial Chemistry	38
3.4.5 Microwave-assisted procedures	41
4. Exploration of HuR–ligand binding modes	43
4.1 Set up and application of differential epitope mapping (DEEP) by STD NMR on selected natural products: 5, 8 and 9 (rutin, aesculin and novobiocin)	46
5. Structure-based approach to obtain new HuR ligands.	52

5.1	Resolution of ( $\pm$ ) 1 enantiomers and differential interaction study	56
6.	Design of new derivatives of synthetic hit 4 via fragment-based approach	63
6.1	Fragment screening on the native protein	63
6.2	Molecular hybridization of HuR ligand 4 and hit fragments	65
7.	Fragment linking by protein templated dynamic combinatorial chemistry to identify new HuR ligands	70
7.1	Initial studies on the two new protein constructs	70
7.2	Fragment screening on protein constructs by SPR	71
7.3	Dynamic Combinatorial Chemistry studies	73
7.4	Preliminary interaction studies	74
8.	Discussion	76
	Conclusion and research perspectives	85
	Experimental Section 1	87
	Chapter 4: Exploration of HuR–ligand binding modes	87
	Chapter 5: Structure-based approach to obtain new HuR ligands	99
	Chapter 6: Design of new derivatives of synthetic hit 4 via fragment-based approach	124
	Chapter 7: Fragment linking by protein templated dynamic combinatorial chemistry to identify new HuR ligands	126
	SECTION 2: Application of Dynamic Combinatorial Chemistry approach to obtain Bacterial Metalloprotease LasB ligands potentially acting as inhibitors	139
1.	Introduction and Aim	139
2.	Protein-templated DCC to obtain new LasB ligands	140
3.	Conclusion	142
	Experimental section 2	143
	Appendix	i
I.	References	i
II.	Abbreviations	xvii
III.	Publications related to the thesis	xix
	Article 1: Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD-NMR and Molecular Modelling approach	xix
	Article 2: Novel compounds targeting the RNA-binding protein HuR. Structure-based design, synthesis and interaction studies	xxx
	Book Chapter: 18. The use of microwaves in drug discovery	xxxvi
	Article 3: (R)-(-)-Aloesaponol III Methyl Ether from <i>Eremurus persicus</i> : A Novel Compound against Leishmaniosis	xxxvii
	Article 4: The use of a microwave assisted solvent extraction coupled with HPLC UV/PAD to assess the quality of <i>Marrubium Vulgare</i> L. (white horehound) herbal raw material	lii
V.	Contributions to the research	lx

## ***Short Abstracts***

### ***English – Short Abstract***

The ultimate goal of the project in which this thesis took part is to develop a new series of small, drug-like organic molecules able to modulate the stability of protein–RNA complexes involved in several diseases, thus regulating gene expression with an unprecedented mode of action and opening the way to a new class of therapeutic agents.

I focused my research on ELAV protein HuR due to its involvement in several pathologies (*i.e.*, cancer, retinopathy, *etc.*) in which the protein is overexpressed/hyper-activated with consequences on target RNAs and their activity (*i.e.*, transcription). In this context, the main aim would be to develop small molecules able to bind HuR and interfere with the formation and/or the stability of its complexes with target RNAs.

Given the early stage of this research topic, the work was articulated in three main blocks:

1. Implementation of a systematic biophysical and *in silico* approach exploiting complementary and informative methods;
2. Design and synthesis of new ligands according to both a structure-based (*in silico*) and fragment-based (based on experimental fragment screening) approach;
3. Biophysical assessment of the interaction between compounds and target protein (HuR).

### ***Deutsch – Kurzzusammenfassung***

Das Ziel des Projekts, zu welchem diese Dissertation gehörte, ist die Entwicklung einer neuen Serie kleiner, wirkstoffartiger organischer Moleküle, die die Stabilität von Protein–RNA Komplexen, welche in etlichen Krankheiten eine Rolle spielen, modulieren können. Dadurch würde die Gen-Expression auf eine bisher unbekannte Weise reguliert und die Entwicklung einer neuen Klasse therapeutischer Wirkstoffe ermöglicht werden.

Ich habe meine Forschung auf das ELAV Protein HuR fokussiert, da es in verschiedenen Krankheitsbildern (z.B. Krebs, Retinopathie) involviert ist. In diesen Fällen ist das Protein überexprimiert/hyperaktiviert, wodurch die Ziel-RNA und ihre Aktivität (z.B. Transkription) beeinflusst werden. In diesem Zusammenhang wäre das Hauptziel die Entwicklung kleiner Moleküle, die HuR binden und die Bildung und/oder die Stabilität der Komplexe mit Ziel-RNA beeinträchtigen können. Da sich das Forschungsthema noch in einem frühen Stadium befindet, wurde die Arbeit in drei Blöcken formuliert:

1. Implementierung eines systematischen biophysikalischen und *in silico* Ansatzes, der komplementäre und informative Methoden nutzt;
2. Design und Synthese neuer Liganden basierend sowohl auf einem struktur- (*in silico*), als auch auf einem fragmentbasierten Ansatz (experimentelles Fragmentscreening)
3. Biophysikalische Bewertung der Wechselwirkungen zwischen Molekülen und Ziel-Protein (HuR).

## ***Abstracts***

### ***English - Abstract***

The research for the present PhD thesis was carried out in co-tutelle between the University of Pavia (Italy) and Saarland University (Germany), on a collaborative project concerning the identification of new compounds as ligands of RNA-binding proteins. The work was conducted at the Medicinal Chemistry Laboratory of the Department of Drug Sciences (University of Pavia – I), the Department of Drug Design and Optimization – DDOP (Helmholtz Institute for Pharmaceutical Research Saarland – HIPS, Saarbrücken – D), and the NMR Unit of the Department of Chemistry (University of Milano – I).

The ultimate goal of the project in which this thesis took part is to develop a new series of small, drug-like organic molecules able to modulate the stability of protein–RNA complexes involved in several diseases, thus regulating gene expression with an unprecedented mode of action and opening the way to a new class of therapeutic agents.

RNA-binding proteins (RBPs) have been widely recognized for their pivotal role in the regulation of post-transcriptional processes. Particularly, their complexes with RNA are involved in numerous dysfunctions (*i.e.*, cancer, inflammation and neurodegeneration) and thus pose the interesting question of whether they could be used as therapeutic targets with clinical relevance. In this context, particular attention deserves the recently discovered Embryonic Lethal Abnormal Vision (ELAV) or Hu protein class, characterized by the ability to stabilize mRNA by binding specific Adenine Uridine Rich elements (ARE); compounds able to modulate the complex stability could represent an innovative pharmacological strategy for both treatment and prevention of numerous diseases.

Over the past few years, these considerations have led to a small number of studies aimed at investigating these proteins in their binding to mRNA to develop novel strategies against different diseases. The few compounds so far identified as complex modulators largely derive from different HTS approaches and the comparison between their results is often limited by the lack of homogeneity of the assays used. Along with HTS approaches, very few other investigations have been conducted.

With the final aim to obtain compounds able to act on the stability of ELAV–RNA complexes in mind, I focused my research on ELAV protein HuR due to its involvement in several pathologies (*i.e.*, cancer, retinopathy, *etc.*) in which the protein is shown as overexpressed/hyper-activated with consequences on target RNAs and their activity (*i.e.*, transcription). In this context, the main aim would be to develop small molecules able to

bind HuR and interfere with the formation and/or the stability of its complexes with target RNAs.

Given the early stage of this research topic, the work was articulated in three main blocks:

1. Implementation of a systematic biophysical and *in silico* approach exploiting complementary and informative methods;
2. Design and synthesis of new ligands according to both a structure-based (*in silico*) and fragment-based (based on experimental fragment screening) approach;
3. Biophysical assessment of the interaction between compounds and target protein (HuR).

The methodologies implemented in the first block, and exploited throughout the thesis, include saturation transfer difference (STD) NMR, molecular dynamics (MD) and docking studies.

The second block exploits free-web based platform AnchorQuery (NucleoQuery) focused on multi-component reactions for the structure-based strategy which led to the first series of synthetic HuR ligands. Another series of compounds were, instead, developed according to a fragment-based approach which consisted in the following steps: (i) biophysical screening of a small halogen-enriched fragment library (HEFL) and selection of hit fragments (exploiting Surface Plasmon Resonance – SPR, STD NMR, and Thermal Shift Assay – TSA); (ii) fragment-based design of new HuR candidate ligands: (ii-a) molecular hybridization of hit fragments and one selected HuR binder; (ii-b) establishment and performance of a Dynamic Combinatorial Chemistry (DCC) approach based on the chemistry of acylhydrazones, using hit fragments to guide building block selection.

The hit compounds identified and synthesized throughout the PhD project were submitted to biological and biophysical characterization against the protein and protein–RNA complex target. The significance of the obtained small molecules in the modulation of gene expression, for treating clinically relevant diseases in which HuR–RNA complexes are involved, is currently under evaluation.

Given the challenging nature of the main target HuR, and the many efforts spent in optimizing the experimental conditions on several protein models, the first DCC experiments were performed in the context of a DDOP – HIPS infection target, bacterial metalloprotease LasB. Since the purpose of the related project goes beyond the aim of this thesis, the associated section will be mainly focused on the DCC approach carried out.

## ***Italiano - Riassunto***

La presente tesi di dottorato è il risultato di una cotutela tra l'Università degli Studi di Pavia e l'Università del Saarland (Germania), nel contesto di una collaborazione focalizzata sull'identificazione di nuovi ligandi per proteine che legano RNA. Il lavoro è stato svolto presso il laboratorio di chimica farmaceutica del dipartimento di scienze del farmaco (Università degli Studi di Pavia, Italia), il dipartimento di drug design e optimization – DDOP (Helmholtz Institute for Pharmaceutical Research Saarland – HIPS, Saarbrücken, Germania), e l'unità NMR del dipartimento di chimica (Università degli Studi di Milano, Italia).

Obiettivo ultimo del progetto in cui questa tesi si inserisce è di costruire una nuova serie di piccole molecole capaci di modulare la stabilità di quei complessi proteina–RNA che sono implicati in diverse patologie, regolando l'espressione genica con un meccanismo d'azione peculiare ed innovativo, e dando il via allo sviluppo di una nuova classe di agenti terapeutici.

Le proteine che legano RNA (RBP) hanno un ruolo fondamentale nella regolazione dei processi post-trascrizionali. In particolare, i loro complessi con RNA sono coinvolti in numerose disfunzioni (ad esempio, cancro, processi infiammatori e neurodegenerativi); recentemente è quindi sorta la domanda se RBP e i loro complessi con RNA possano essere utilizzati come bersagli terapeutici con rilevanza clinica. In questo contesto, ha richiamato particolare attenzione la classe di proteine Embryonic Lethal Abnormal Vision (ELAV) o Hu, in grado di stabilizzare RNA messaggeri (mRNA) interagendo con particolari sequenze ricche di adenina e uracile (ARE); composti capaci di modulare la stabilità dei complessi potrebbero rappresentare una strategia farmacologica innovativa per il trattamento e la prevenzione di diverse malattie.

Negli ultimi anni, queste considerazioni hanno stimolato un piccolo numero di studi volti ad investigare le proteine ELAV e il loro legame con mRNA per sviluppare nuove strategie contro varie patologie. I pochi composti finora identificati come modulatori dei complessi derivano principalmente da screening ad alta portata (HTS) e un confronto critico dei loro risultati è spesso limitato dalla scarsa omogeneità dei saggi utilizzati. Oltre agli screening HTS, solo pochi altri studi sono stati effettuati.

Al fine di ottenere composti in grado di agire sulla stabilità dei complessi ELAV–RNA, ho concentrato la mia ricerca sul sottotipo HuR dato il suo coinvolgimento in diverse malattie (ad esempio, cancro e retinopatia), in cui la proteina è sovra espressa e iperattivata con conseguenze sui suoi RNA target e sulla loro attività. In questo contesto, l'obiettivo principale è quello di sviluppare piccole molecole capaci di legarsi a HuR ed interferire con la formazione e/o la stabilità dei suoi complessi con RNA bersaglio.

Dato lo stadio precoce della ricerca in questo campo, il lavoro è stato articolato in tre macro-blocchi:

1. Definizione, set up ed esecuzione di un sistema combinato basato su tecniche biofisiche e approcci *in silico*, sfruttando approcci complementari e informativi;
2. Disegno e sintesi di nuovi ligandi sulla base di due approcci, structure-based (*in silico*) e fragment-based (basato sullo screening sperimentale di frammenti);
3. Valutazione biofisica dell'interazione tra composti e proteina target (HuR).

Le metodologie implementate nel primo blocco, e sfruttate durante tutto il lavoro di tesi, includono saturation transfer difference (STD) NMR, dinamica molecolare (MD) e studi di docking. Il secondo blocco comprende l'utilizzo della piattaforma web gratuita AnchorQuery (NucleoQuery) che ha portato al disegno e sintesi della prima serie di ligandi sintetici di HuR. Le serie seguenti sono state invece sviluppate secondo un approccio fragment-based formato dai seguenti step: (i) screening biofisico di una piccola libreria di frammenti alogenati (HEFL) e selezione di frammenti hit (sfruttando Surface Plasmon Resonance – SPR, STD NMR, and Thermal Shift Assay – TSA); (ii) design di ligandi sulla base dei frammenti selezionati: (ii-a) ibridazione molecolare di frammenti e un ligando di HuR selezionato; (ii-b) set up ed esecuzione di un approccio dynamic combinatorial chemistry (DCC) basato sulla chimica degli acilidrazoni, utilizzando frammenti hit per la selezione dei building block.

I composti identificati e sintetizzati nel corso del progetto di dottorato sono stati sottomessi a caratterizzazione biofisica e biologica dell'interazione con la proteina e l'attività sui complessi proteina–RNA. È attualmente in corso la valutazione della rilevanza delle piccole molecole ottenute nella modulazione dell'espressione genica, per il trattamento di patologie legate alla sregolazione dei complessi HuR–RNA.

Data la natura complessa del target HuR, e considerati i diversi sforzi spesi per ottimizzare le condizioni sperimentali su diversi modelli proteici, i primi esperimenti DCC sono stati eseguiti nel contesto di un target per l'infezione del DDOP – HIPS, la metallo proteasi batterica LasB. Poiché l'obiettivo del progetto correlato esula dallo scopo di questa tesi, la sezione relativa sarà prevalentemente dedicata all'approccio DCC sviluppato.

## ***Deutsch - Zusammenfassung***

Die Forschung zur vorliegenden Dissertation wurde im Cotutelle-Verfahren zwischen der Universität Pavia (Italien) und der Universität des Saarlandes (Deutschland) durchgeführt. Das Kollaborationsprojekt befasste sich mit der Identifizierung neuer Verbindungen als Liganden für RNA-bindende Proteine. Die Arbeit erfolgte im Labor für Medizinalchemie der Abteilung „Drug Sciences“ (Universität Pavia – I), der Abteilung „Drug Design and Optimization“ – DDOP (Helmholtz-Institut für pharmazeutische Forschung Saarland – HIPS, Saarbrücken – D) und der NMR-Abteilung der chemischen Fakultät in Mailand (Universität Mailand – I).

Das Ziel des Projekts, zu welchem diese Dissertation gehörte, ist die Entwicklung einer neuen Serie kleiner, wirkstoffartiger organischer Moleküle, die die Stabilität von Protein–RNA Komplexen, welche in etlichen Krankheiten eine Rolle spielen, modulieren können. Dadurch würde die Gen-Expression auf eine bisher unbekannte Weise reguliert und die Entwicklung einer neuen Klasse therapeutischer Wirkstoffe ermöglicht werden.

RNA-bindende Proteine (RBPs) nehmen eine zentrale Rolle in der Regulierung post-transkriptionaler Prozesse ein. Besonders ihre Komplexe mit RNA sind an vielfältigen Fehlfunktionen (z.B. Krebs, Entzündungen und Neurodegeneration) beteiligt, weshalb sich die interessante Frage stellt, ob sie als therapeutische Ansatzpunkte mit klinischer Relevanz genutzt werden können. In diesem Zusammenhang verdient das erst kürzlich entdeckte „Embryonic Lethal Abnormal Vision“ (ELAV) oder auch Hu-Proteinklasse besondere Aufmerksamkeit. Es zeichnet sich durch die Fähigkeit aus, mRNA zu stabilisieren, indem es spezifische Adenin-Uridin reiche Elemente (ARE) bindet. Verbindungen, die diese Komplexstabilität modulieren können, könnten eine innovative pharmakologische Strategie für die Behandlung und Vorbeugung zahlreicher Krankheiten darstellen.

In den vergangenen Jahren haben diese Überlegungen zu einer kleinen Anzahl an Studien geführt, die die Bindung dieser Proteine an mRNA untersuchten, mit dem Ziel neue Strategien für verschiedene Krankheiten zu entwickeln. Die wenigen Komplex-modulierenden Verbindungen, die bisher identifiziert wurden, stammen aus verschiedenen Hochdurchsatz-Screenings und der Vergleich der Ergebnisse ist häufig limitiert wegen geringer Vergleichbarkeit der Assays. Bis jetzt wurden wenig andere Ansätze verfolgt.

Mit dem Ziel vor Augen, Verbindungen zu erhalten, die die Stabilität der ELAV–RNA Komplexe beeinflussen, habe ich meine Forschung auf das ELAV Protein HuR fokussiert, da es in verschiedenen Krankheitsbildern (z.B. Krebs, Retinopathie) involviert ist. In diesen Fällen ist das Protein überexprimiert/hyperaktiviert, wodurch die Ziel-RNA und ihre Aktivität (z.B. Transkription) beeinflusst werden. In diesem Zusammenhang wäre das Hauptziel die

Entwicklung kleiner Moleküle, die HuR binden und die Bildung und/oder die Stabilität der Komplexe mit Ziel-RNA beeinträchtigen können. Da sich das Forschungsthema noch in einem frühen Stadium befindet, wurde die Arbeit in drei Blöcken formuliert:

4. Implementierung eines systematischen biophysikalischen und *in silico* Ansatzes, der komplementäre und informative Methoden nutzt;
5. Design und Synthese neuer Liganden basierend sowohl auf einem struktur- (*in silico*), als auch auf einem fragmentbasierten Ansatz (experimentelles Fragmentscreening)
6. Biophysikalische Bewertung der Wechselwirkungen zwischen Molekülen und Ziel-Protein (HuR).

Die Methoden des ersten Blocks, die in dieser Dissertation genutzt wurden, sind „saturation-transfer difference“ (STD) NMR, Molekulardynamik (MD) und Docking Studien.

Der zweite Block nutzt die frei zugängliche Plattform AnchorQuery (NucleoQuery), welche für die strukturbasierte Strategie auf Multikomponenten Reaktionen fokussiert ist. Die folgende Molekülserie wurden hingegen nach ein fragmentbasiertem Ansatz entwickelt, der folgende Schritte enthielt: (i) biophysikalisches Screening einer kleinen Halogen-angereicherten Fragment Bibliothek (HEFL) und Auswahl der Fragment-Hits (Nutzung von Oberflächenplasmonenresonanz-Spektroskopie – SPR, STD NMR und thermische Verschiebung Assay – TSA); (ii) Fragment-basiertes Design neuartiger potentieller HuR Liganden. (ii-a) molekulare Hybridisierung der Fragment-Hits und eines ausgewählten HuR-Binders; (ii-b) Etablierung und Durchführung der „Dynamic Combinatorial Chemistry“ (DCC) Methode, die auf der Chemie der Acylhydrazone basiert, mit den Fragment-Hits als Vorlage für die Bausteinauswahl. Das Verhalten, der während der Dissertation identifizierten und synthetisierten Hitverbindungen, zu dem Protein und dem Protein–RNA Komplex wurde biologisch und biophysikalisch charakterisiert. Die Bedeutung der erhaltenen kleinen Moleküle in der Modulierung der Gen-Expression für die Behandlung von klinisch relevanten Krankheiten, in denen HuR–RNA Komplexe involviert sind, wird derzeit evaluiert.

Es stellte sich als sehr schwierig heraus sowohl für das HuR-Protein als auch für diverse Proteinmodelle die experimentellen Bedingungen zu etablieren. Daher wurden die ersten DCC Experimente an einem Infektionstarget der DDOP-Gruppe, der bakteriellen Metalloprotease LasB, durchgeführt. Da das Ziel dieses verwandten Projekts über das Ziel dieser Arbeit hinausgeht, liegt der Schwerpunkt des entsprechenden Abschnitts auf dem DCC-Ansatz im Allgemeinen.

## ***SECTION 1: Versatile use of advanced NMR techniques for the identification of compounds able to interfere with ELAV protein–mRNA complexes***

### ***1. Introduction***

Starting from the development of modern biology and biochemistry, continuous advances in proteomics have been unveiling new proteins with pivotal functions in the physiological and pathological pathways of our bodies. From a medicinal and pharmacological standpoint, these proteins represent potential or factual targets of clinically relevant therapies that need to be continuously studied and developed to keep up with the overwhelming understanding that a longer life-span is linked with a series of dysfunctions we do not know yet how to face. This is not only producing a rise in the number of protein targets to consider, but also a paradigm shift, from traditional targets, such as receptors and enzymes, to novel targets,<sup>1</sup> *i.e.* protein–protein interactions,<sup>2</sup> intrinsically-disordered proteins<sup>3</sup> and, finally, non-protein targets such as nucleic acids;<sup>4–6</sup> this rather drastic shift is forcing us to redefine and redesign research concepts and techniques of classical pharmacology and medicinal chemistry to better respond to this growing demand.

One of these unusual targets, and main topic of this doctoral dissertation, is represented by RNA-binding proteins (RBPs). The involvement of this protein class in the post-transcriptional modifications related to many physio-pathological mechanism makes it an interesting target for uphill gene regulation.

#### ***1.1 RNA-binding proteins***

RNAs in cells are bound to RBPs to form ribonucleoprotein (RNP) complexes, which influence their structure and interactions in their journey from the nucleus (*i.e.*, alternative splicing, polyadenylation, nuclear export) to the ribosome (*i.e.*, cytoplasmic localization, stability, translation rate). Eukaryotic cells encode a large number of RBPs (thousands in vertebrates), each with a unique RNA-binding activity, and type of RNA-binding domain (RBD) and protein–protein interaction characteristics.<sup>7–13</sup>

Focusing solely on the role of RBPs in the stabilization of target mRNAs, we can recognize the possible implications of their dysregulation<sup>9,10,12,14–17</sup> in different tissues and organs of the human body, as a change in mRNA function and stability can cause inhibition/augmentation in the production and function of their codified proteins. This can often relate directly with the evolution of pathologies such as neurodegenerative diseases, inflammation and cancer.<sup>18–25</sup> This realization has recently raised the question of whether RBPs and their complexes with RNAs could constitute therapeutically relevant targets to modulate gene expression.<sup>26,27</sup>

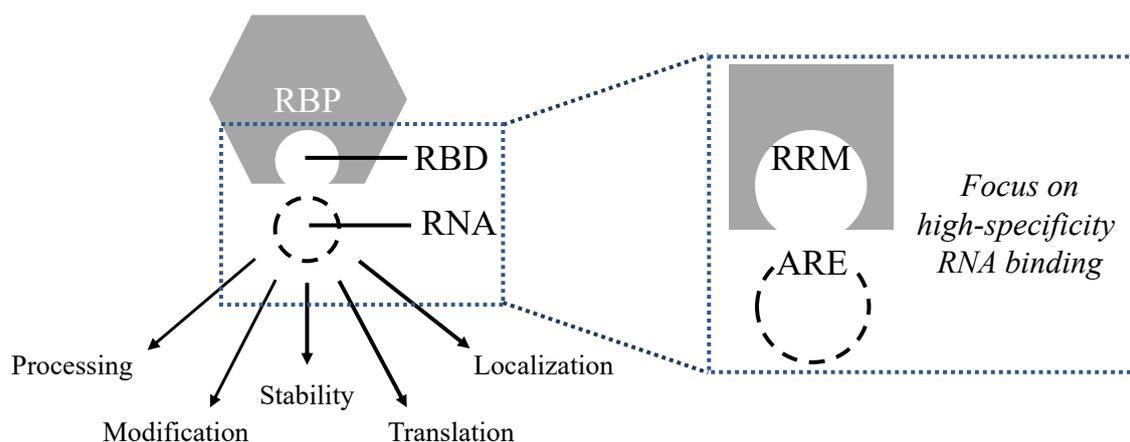


Figure 1. An RNA-binding protein (RBP) can interact with RNA through defined RNA-binding domains (RBDs) to regulate RNA metabolism and function. For certain ribonucleoprotein complexes (RNP), the interaction occurs with high specificity and is mediated, on the protein side, by a RNA-recognition motif type domain (RRM) and, on the RNA side, by adenine-uridine rich sequences (ARE). (Adapted from Hentze *et al.*, 2018<sup>11</sup>)

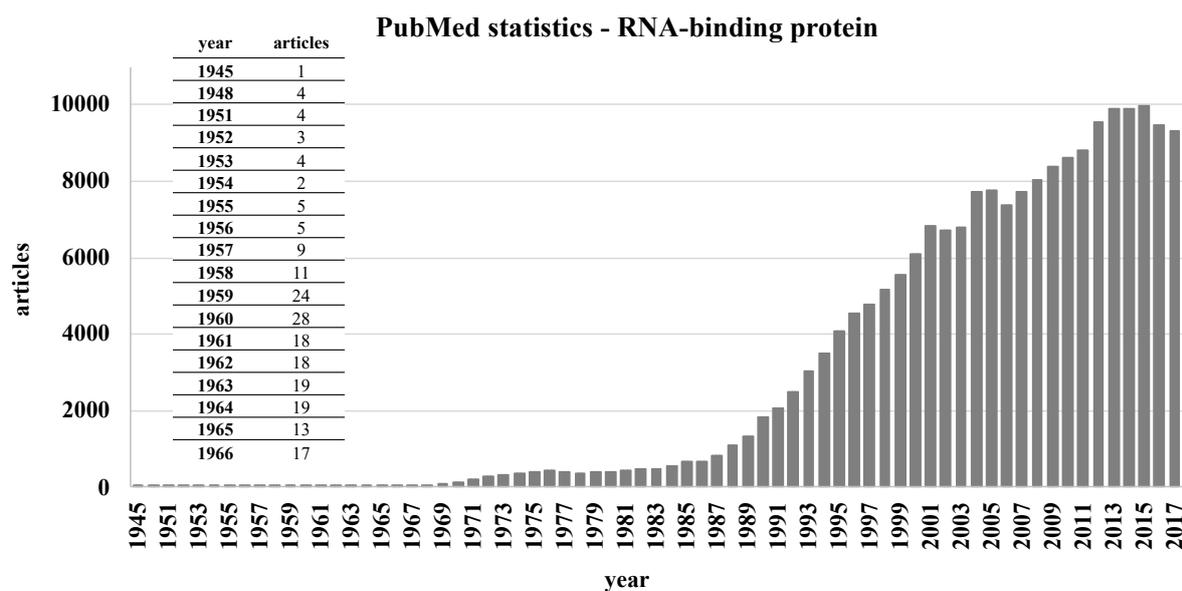
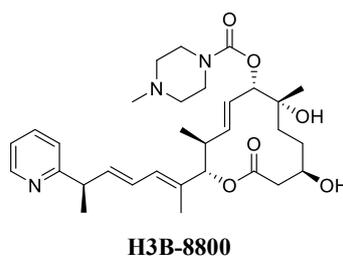


Figure 2. Statistics reported on PubMed for the number of publications including RBPs along the years since their discovery (complete data reported until the end of 2018; platform accessed in September 2019 - <https://www.ncbi.nlm.nih.gov/pubmed/>).

Few approaches have been envisioned to target RBPs and modulate RNP complexes and, so far, most of them have included RNA interference-based oligonucleotides (*e.g.* siRNA, microRNAs and RNA decoys) targeting different types of RBDs.<sup>28-32</sup> A more recent, strictly MedChem, approach involves the identification of small molecules able to modulate the stability and/or formation of RNP complexes or directly of the target RNA itself. This approach is intrinsically challenging, due to the fact that this type of proteins often do not have a recognized or defined small molecule binding site, thus, to selectively target RNA sequences

or RBP by developing small molecule ligands in the classical sense, is nowadays still quite a difficult task. Nonetheless, few pioneering examples of rationally designed compounds targeting RBPs can be found in literature; perhaps, the most promising is represented by H3B-8800<sup>33-35</sup> (see Figure 3), the first RBP-targeting small molecule to reach, in 2018, the first stage of clinical phase for acute myelogenous leukemia and chronic myelomonocytic leukemia. This compound was developed through an iterative medicinal chemistry effort starting from a scaffold based on pladienolide, a naturally occurring antitumor macrolide targeting splicing factor 3B1 (SF3B1) protein.<sup>36</sup> H3B-8800 preferentially targets cells with spliceosome complexes containing mutant SF3B1, modulating intron splicing and leading to increased death in cancer cells while having little effect on the viability cells with wild-type SF3B1; it is thought to bind to a site similar to pladienolide on the SF3B complex within the spliceosome-1. Once bound, it induces increased retention of short (< 300 nucleotides) GC-rich introns through modulation of pre-mRNA processing. These intron-retained mRNA sequences are then thought to be destroyed through the nonsense-mediated decay pathway. 41 of 404 genes encoding spliceosome proteins contain GC-rich sequences whose retention was induced by H3B-8800, and this is presumed to be key to the specificity of H3B-8800 lethality. Since cancer cells, as in myelodysplasia, experience SF3B1 mutations much more frequently than host cells, this allows H3B-8800 to be used to preferentially target these cells by inducing intron-retention in critical spliceosome component pre-mRNA, leading to destruction of nonsense mature RNA and ultimately cell death due to the lack of critical proteins.<sup>33-35</sup>



*Figure 3.* Structure of H3B-8800, the first RBP-targeting small molecule to reach, in 2018, the first stage of clinical phase for acute myelogenous leukemia and chronic myelomonocytic leukemia.

The success of H3B-8800 is an encouragement to keep pursuing RBPs as pharmacological targets with clinical relevance. In this study field, since 2007, the focus of our research group revolves around the Embryonic Lethal Abnormal Vision (ELAV) or Hu protein family, a class with a pivotal role in neuron development<sup>37,38</sup> and maintenance as well as in cancer,<sup>20,39,40</sup> and inflammation,<sup>40</sup> recently proposed as an interesting pharmacological target.<sup>12,26</sup>

## 1.2 ELAV/Hu protein family

The ELAV protein class is thoroughly studied and described from a biochemical and functional point of view. In particular, four ELAV proteins are known in mammals: ubiquitous HuR (ELAVL1), and prevalently neuronal subtypes HuB (ELAVL2), HuC (ELAVL3), and HuD (ELAVL4).<sup>39</sup> These proteins are characterized by a high degree of sequence homology (70–85%) and are approximately 40 kDa in size. They contain three RNA recognition motif-type (RRM) domains, each approximately 90 amino acids long:<sup>41</sup> the first two consecutive domains (RRM1 and 2) are near the N-terminus and connect to the third domain (RRM3) by an unconserved hinge region, mainly responsible for the nuclear/cytoplasmic shuttling occurring after protein activation.<sup>37,42,43</sup> The RRM1 and RRM2 domains directly interact with the ARE sequences of target transcripts through the highly conserved RNP1 and RNP2 sequences (8 and 6 amino acids, respectively),<sup>44</sup> while the RRM3 domain is involved in binding to the poly(A) tail of the target mRNAs and has a crucial role in the homo- and heteromultimerization processes of ELAV proteins.<sup>45–47</sup>

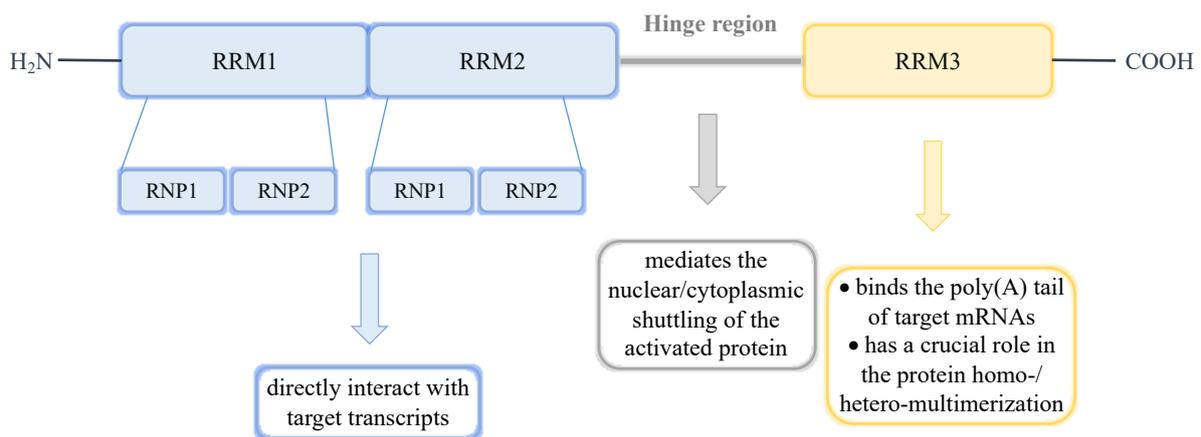


Figure 4. ELAV domains and their involvement in protein function (adapted from Vasile *et al.*, 2018, see appendix III, article 1).

While in basal conditions, neuronal ELAVs are mainly found in the cytoplasm and ubiquitous HuR mostly in the nucleus, there is no strict compartmentalization for both localization and function, since ELAV proteins can shuttle between the cytoplasm and the nucleus under specific *stimuli*, and thus affect any step of the post-transcriptional gene regulation. In fact, neuronal ELAVs in the nucleus of neurons regulate splicing and polyadenylation, whereas, in the cytoplasm, they control mRNA target stability, translation, and transport.<sup>48,49</sup> All mammalian ELAVs can interact with many mRNAs, giving rise to various ELAV protein–mRNA complexes with different physiological roles<sup>38,50–52</sup> and it has been suggested that both structure-specificity and subcellular localization could affect protein–target RNA specificity.<sup>53</sup> In fact, mRNAs encoding proteins involved in neuronal plasticity, outgrowth, and neuronal

metabolism maintenance are mainly targeted by neuronal ELAV proteins, while HuR has been described to possess a greater binding affinity for transcripts implicated in oxidative stress defense, inflammation, cell growth, and cell cycle regulation.<sup>48,49</sup> Notably, the importance of HuR in these processes is so relevant that HuR has been proposed as a prognostic factor for different types of cancer,<sup>21,54,55</sup> while neuronal ELAV impairments have been suggested to contribute to the development of neurodegenerative disorders, such as Alzheimer's disease<sup>56</sup> and paraneoplastic neurological syndrome.<sup>15</sup> Overall, this evidence suggests that ELAV proteins have potential as pharmacological targets in several pathologies. Each pathology presents a different impairment in the function, activation and/or expression of one or more ELAV subtypes, thus the search for modulators of ELAV–RNA complexes can be in principle aimed at either increasing the stabilization of the complex (or of the target RNA), *i.e.* in some neurodegenerative diseases, or at preventing the formation/destabilizing the complex, *i.e.* in some cancers, inflammatory processes and diabetic retinopathy. In the context of this thesis, the focus will revolve around building technical and pharmacological tools to develop a therapeutic approach which can interfere with or disrupt selected ELAV–RNA complexes.

### ***1.3 Compounds able to interfere with ELAV–mRNA complexes***

As underlined in the previous paragraph, the ELAV protein class is thoroughly characterized from a biochemical and functional point of view, and several studies have gained us knowledge on the structural features of the protein domains interacting with some target RNAs. Nonetheless, regardless of the potential of ELAV proteins as pharmacological targets in various diseases, related medicinal chemistry research is still at an early stage: only few compounds able to interfere with ELAV–RNA complexes have been identified and very little is known about the features of the ELAV–small molecule interaction, even for compounds which are proven interferers of the ELAV–RNA binding. When approaching a new target, the elucidation of the structural hot-spots is fundamental for the rational discovery of new ligands but, at the time this thesis was started, all literature-described ligands for these targets had been identified as a result of high-throughput screenings (HTS); a thorough description of the state-of-the-art on this topic at the time this thesis was started can be found in a perspective review of our research group (see Figure 5).<sup>26</sup> The main findings are briefly reported hereinafter.

Meisner's team was the first to approach the ELAV–RNA system from a medicinal chemistry perspective in 2007, focusing on subtype HuR.<sup>57</sup> A screening and purification of 50,000 extracts

from natural sources, against both the full-length protein and a truncated variant, based on fluorescence intensity distribution analysis (FIDA),<sup>58,\*</sup> allowed isolation and characterization

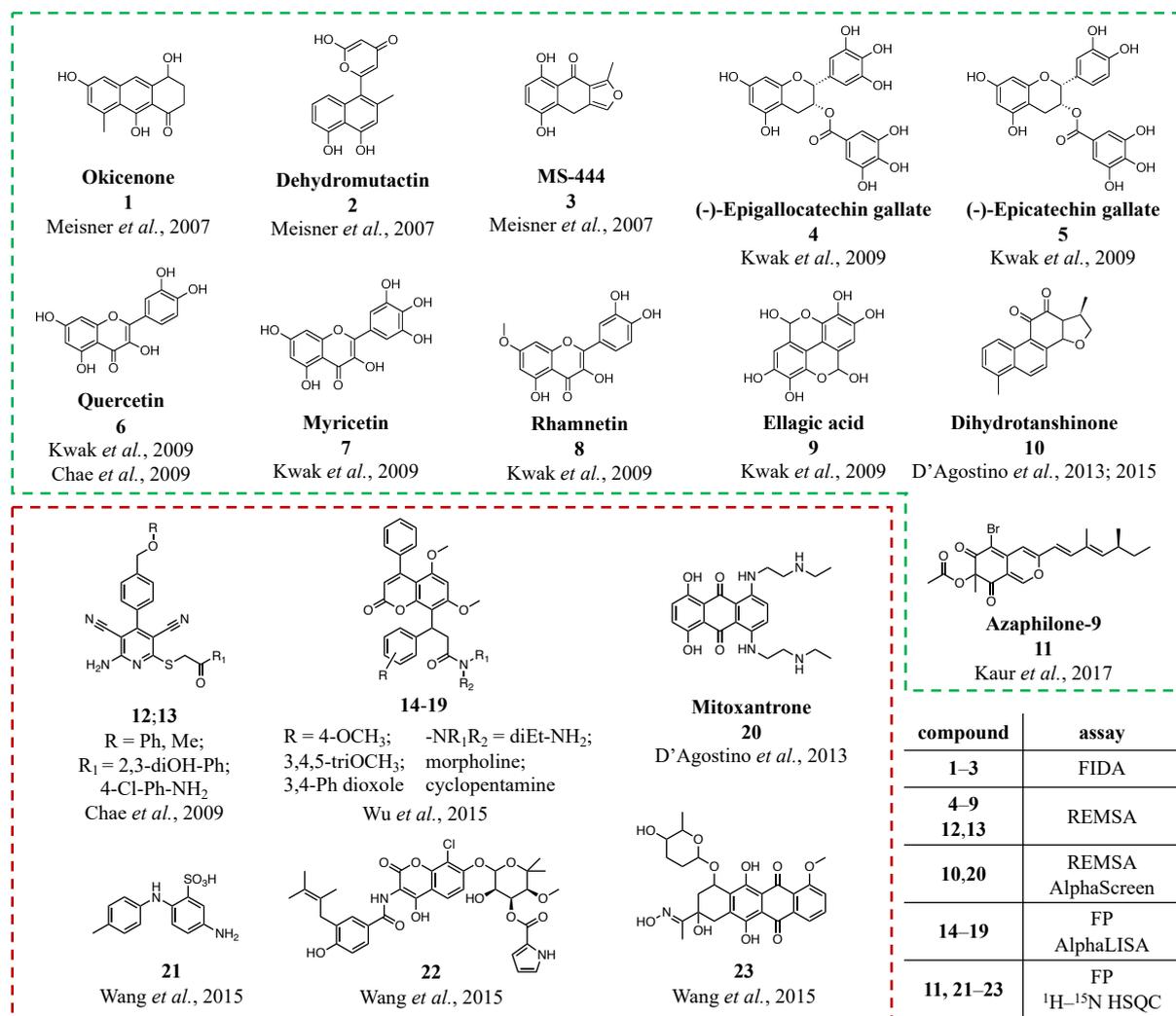


Figure 5. Compounds able to interfere with ELAV–RNA complexes from high-throughput screenings (HTS);<sup>26,59</sup> natural products are enclosed in a green rectangle, synthetic compounds in a red rectangle.

of 1–3; these exerted concentration-dependent *in vitro* inhibitory activity, attributed to their interaction with RRM1 and 2 domains. The postulated mechanism of interaction was further investigated on the HuR RRM1/2–RNA complex combining experimental and *in silico* results, allowing to advance the following points: a) HuR exists as a homodimer; b) it binds RNA in a 2:1 or 1:1 stoichiometric ratio; c) RRM domains may mediate homodimerization; d) certain compounds (1 and 2) can interfere with homodimerization. These findings partially clarified the interaction mechanism of exogenous ligands and ELAV proteins, stressing the importance

\* FIDA is a fluorescence-based technique that detects the variation of fluorescence of a compound bound to the target by considering the statistics of the intensities of the generated photon bursts. A typical biochemical application in an HTS is a ligand binding assay, which measures the binding of a ligand to the target and its displacement by tested compounds.

of RRM sequences, posing a milestone in this field and opening the concept of chemical druggability of the ELAV protein family.

In 2009, a set of flavonoids able to inhibit HuC–RNA complexes were identified by screening 52 natural compounds with anti-inflammatory and anti-tumorigenic activities through a non-radioactive RNA electrophoretic mobility gel shift assay (REMSA),<sup>†</sup> a technique affirmed for characterizing protein–RNA interactions.<sup>60,61</sup> IC<sub>50</sub> values were assessed for the best compounds, **4–9**, ranging from 0.2 to 1.8 μM.

The same technique was applied on HuR–RNA complexes involved in inflammatory diseases,<sup>62</sup> affording nine candidates based on their inhibitory effect; the determination of IC<sub>50</sub> values by a filter-binding assay allowed the further selection of compounds **6,12** and **13**.

In 2013, Provenzani's team proposed amplified luminescent proximity homogeneous assay screen (AlphaScreen)<sup>‡</sup> to identify new compounds able to interfere with HuR–RNA complexes.<sup>63</sup> The screening of a library of 2,000 molecules, afforded 16 compounds further evaluated by REMSA to exclude false-positives. Interestingly, in agreement with results previously reported by Meisner, mitoxantrone (**20**) was found to disrupt HuR–RNA complexes. Upon approach validation, a new compound library was screened against the HuR–ARE (RNA<sup>TNF $\alpha$</sup> ) complex using AlphaScreen to select hits and REMSA for confirmation.<sup>64</sup> Among eight hits identified, **10** emerged as the most potent in preventing complex formation (IC<sub>50</sub> = 0.15 μM).

In 2015, a screening was performed on 6,000 compounds by using a fluorescence polarization (FP)<sup>§,65</sup> binding assay on HuR and a particular ARE sequence related to malignancy and proliferation of gliomas and melanomas.<sup>66</sup> This study led to the discovery of six coumarins (**14–19**) able to disrupt the complex. All compounds showed lower IC<sub>50</sub> and *K<sub>i</sub>* values for the full-length protein than for the RRM1 and 2 domains, although related data did not provide information about the binding region involved (direct binding to HuR was confirmed by Surface Plasmon Resonance, SPR).<sup>\*\*</sup>

---

<sup>†</sup> REMSA is an affinity electrophoresis technique used to study protein–RNA interactions. It can determine if a protein or mixture of proteins is capable of binding a given RNA sequence, and can sometimes indicate the stoichiometry of the binding complex. A set up where compounds to test are added to the complexes can highlight displacement. Detection can be performed through radioactivity and fluorescence (using the appropriate probe).

<sup>‡</sup> AlphaScreen is a bead-based proximity assay developed from diagnostic assay LOCI (Luminescent Oxygen Channeling Assay). It is a channeling immunoassay in which an immune reaction directs two bead components into close proximity on a defined surface. One produces a product which stimulates a second component, thereby yielding a detectable signal. Its extensive use results from both the simplicity of its protocols and its high sensitivity. These features, as well as the absence of wash-up and separation steps, allow for the automation of a range of related assays.

<sup>§</sup> FP is a solution-based technique that allows rapid and quantitative analysis of diverse molecular interactions and enzyme activities. The principle of FP derives from the fact that the intrinsic fluorescence intensity of a fluorophore may change upon binding.

<sup>\*\*</sup> A more thorough description of SPR can be found in chapter 3, paragraph 3.2.1.

In the same year, 1,597 compounds were screened by FP based on their ability to inhibit HuR–RNA<sup>c-fos</sup> complex formation; twelve gave IC<sub>50</sub> values ranging from 2.7 to 97.4 μM, (among them, **21–23**).<sup>67</sup> A successive NMR-based screening, <sup>1</sup>H–<sup>15</sup>N heteronuclear single quantum correlation (HSQC),<sup>††</sup> provided evidence of direct interaction with the target protein and with RRM1 and 2 domains for **21**.

The most recent studies in 2017 and 2018, brought to the discovery of natural product azaphilone-9 (**11**)<sup>59</sup> and suramin, an antitrypanosomal drug;<sup>68</sup> due to its high molecular weight (> 1000 Da) suramin does not comply to the small molecule definition and is thus excluded from this recollection. As for **11**, similarly to **21–23**, it was discovered through a FP assay against HuR–ARE–RNA<sup>c-fos</sup> with a IC<sub>50</sub> of 1.2 μM; direct binding was confirmed by SPR (against HuR) and <sup>1</sup>H–<sup>15</sup>N HSQC (against HuR RRM1/2).

Despite their undeniable significance, the obtained outcomes could not afford enough data to lead to a concrete discovery for the identification of new small molecules, directed at ELAV–RNA complexes, with therapeutic value. In fact, these studies were carried out separately not following a consistent method or taking into account respective results, and as a consequence, aside for few overlapping data, their results are often hard to compare and exploit. Namely, because of the structural variability of the derived compounds and the different techniques and experimental conditions employed for the screenings, a direct definition of a structure–activity relationship (SAR) cannot be accomplished.<sup>26</sup>

#### 1.4 *ELAV–RNA interface*

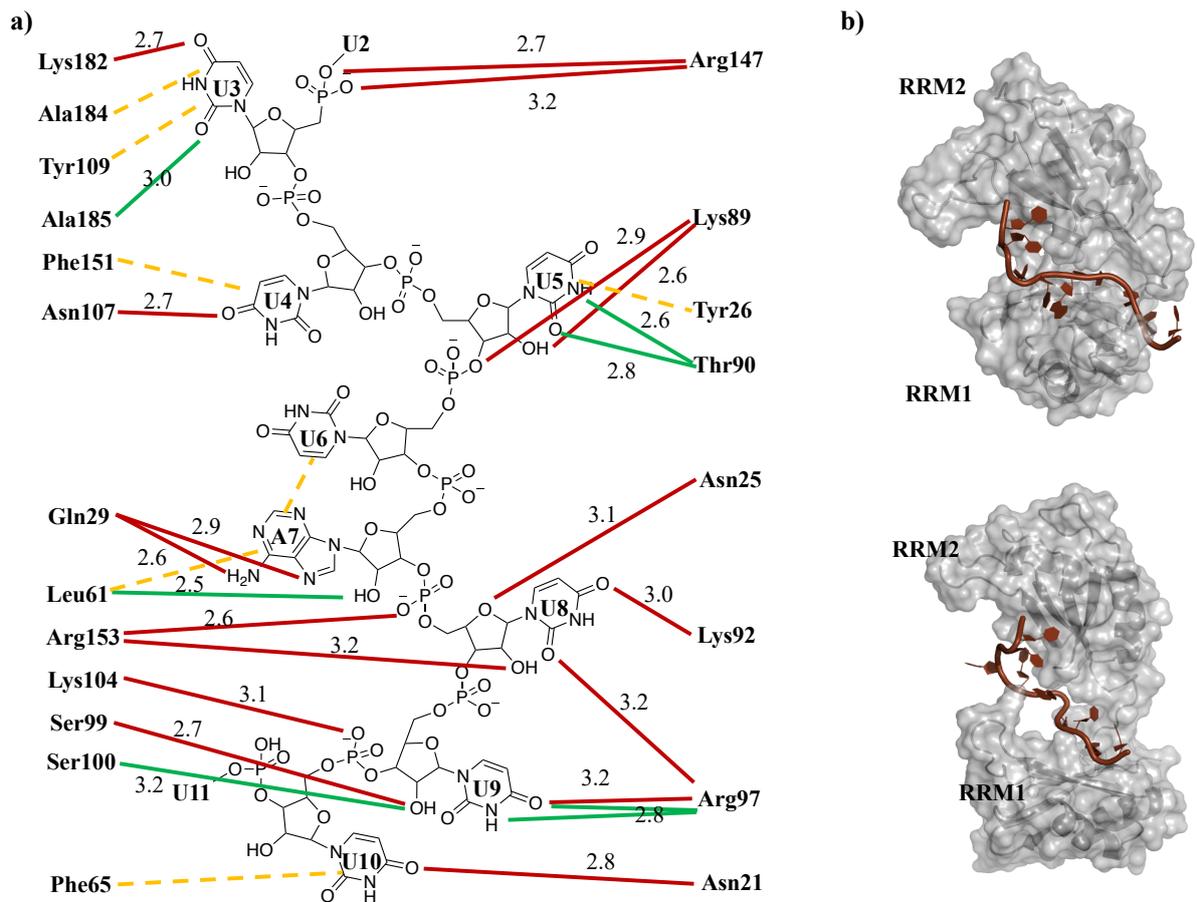
The importance of elucidating in detail the relationship between structural and functional hot-spots of a new target, for the rational discovery of new ligands, has been stressed in the previous paragraph, along with the lack thereof for this interesting family of proteins and small interacting molecules. Nonetheless, as reported in paragraph 1.2, several studies have allowed to obtain this information for the interaction between protein and target RNAs.<sup>44,69–75</sup> Given the large size of the protein–RNA interface and the high flexibility of this protein class, it is hard to directly pinpoint the breaking point of the complex interaction, but, at the same time, we can, among superficial and deeper interactions, recognize some useful hotspots. The main findings, from the co-crystal structures to a functional and structural study of selected ELAV peptides, will be illustrated in the next paragraphs.

---

<sup>††</sup> <sup>1</sup>H–<sup>15</sup>N HSQC is a bidimensional protein-observed NMR technique to study protein-ligand interactions and to detect, thanks to selective signal shifts, which residues of the protein are in contact with the ligand. It allows to detect direct binding and pinpoint the binding site as only the resonances of interacting residues will shift. It requires <sup>15</sup>N-labeled protein.

### 1.4.1 *ELAV–RNA co-crystals*

As reported in paragraph 1.2 and Figure 2, the structure of ELAV proteins includes three RRM domains. The two N-terminal tandem RRM domains can selectively bind AU-rich elements, while the third RRM domain (RRM3) contributes to interactions with the poly-A tail of target mRNAs. Regardless of the importance of RRM3,<sup>46,47,74</sup> focus of the attention for specific RNA binding is on the first two RRMs and the highly conserved RNP1 and 2 sequences of each RRM1/2, of 8 and 6 amino acids, respectively.<sup>44,75</sup> Studies reported on the co-crystal structure of HuD (PDB 1FXL)<sup>44</sup> and HuR (PDB 4ED5)<sup>75</sup> RRM1/2 and AREs of target transcript *c-fos*, the comparison of the latter with the RRMs free form (PDB 4EGL), and additional FP assays afforded interesting and useful information: a) both the co-crystal structure and the FP show that RRM1 is the main ARE binding domain in HuR and that the conformational changes cause subsequent contacts of the RNA substrate with the inter-domain linker and RRM2, greatly improving the complex binding affinity. In the HuR RRM1/2–RNA complex, RRM1 recognizes five nucleotides U5–U8 and U10, the inter-domain linker interacts with U9, and RRM2 with U3–U4. The capability to recognize four or five consecutive uracil-recognition groups gives RRM1 a high RNA-binding capacity in solution, while the ability of the inter-domain linker and RRM2 to recognize one or two uracil residues is not sufficient to subtract the RNA segment from solution (see Figure 6 a).



**Figure 6.** (a) HuR–RNA contacts (PDB code 4ED5). Red lines show side-chain contacts; green lines show main-chain contacts; dashed orange lines show stacking interactions. Distances are expressed in Å (adapted from Wang *et al.*, 2013<sup>75</sup>); (b) Comparison of closed and open form of HuR RRM1/2 (the open form shown was reproduced with Molecular Dynamics simulations from PDB 4ED5 to show the quick loss in contacts with RNA).<sup>18</sup>

b) Due to the high sequence homology of ELAV proteins, NMR studies on HuC and SPR experiments on HuD revealed the same RNA-binding properties, which are consistent with the above results.<sup>70</sup> In the HuD RRM1/2–RNA complex the two individual RRM domains also recognize the RNA substrate in the same conserved manner as in HuR.

c) The two crystal structures of free HuR RRM1/2 and its RNA complex present two distinct conformations: an open state (free form) and a closed state (bound form). Though the protein folds of the single RRM domains in the bound form are basically the same as the free form, the tertiary structure undergoes dramatic conformational changes upon RNA binding (see Figure 6 b) to form the RNA-binding cleft.<sup>75</sup>

#### 1.4.2 Studies on ELAV peptides (RRM1/2 RNP1/2)–RNA

In parallel with the study of the previous paragraph, the first involvement of our research group with the ELAV puzzle focused on the investigation of the significance of the RNP sequences in binding and stabilizing two target RNAs, RNA<sup>NOVA1</sup> and RNA<sup>VEGF</sup>. Briefly, four peptides (**P1–P4**) corresponding to the RNP1/2 sequences in the RRM1/2 domains of HuD (see Table 1

and Figure 7) were synthesized and their behavior towards the RNA transcripts was evaluated through: a) PCR (stabilization of RNA);<sup>76</sup> b) *in silico* methods (prediction of binding modes);<sup>77</sup> c) STD NMR and diffusion-ordered spectroscopy (DOSY) (binding and structural considerations).<sup>78</sup>

PEPTIDE	DOMAIN	AMINO ACID SEQUENCE
P1	RRM1_RNP1	Leu-Gly-Tyr-Gly-Phe-Val-Asn-Tyr
P2	RRM1_RNP2	Leu-Ile-Val-Asn-Tyr-Leu
P3	RRM2_RNP1	Arg-Gly-Val-Gly-Phe-Ile-Arg-Phe
P4	RRM2_RNP2	Leu-Tyr-Val-Ser-Gly-Leu

Table 1. Amino acid sequence of the four ELAV-derived peptides synthesized for the study.

The four peptides were able to stabilize the target RNAs when used as a equimolar mixture of all four or two by two (**P1–P2** and **P3–P4**). In particular, in accordance with the study reported in paragraph 1.4.1, **P1** and **P2** play greater roles in stabilizing RNA. This evidence was further confirmed by *in silico* predictions and NMR binding studies.

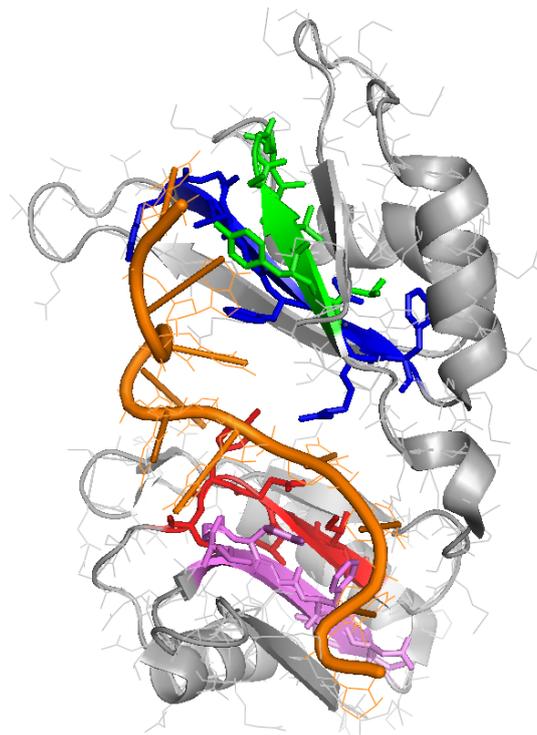


Figure 7. Co-crystal structure of HuD RRM1/2–RNA<sup>TNF $\alpha$</sup>  (PDB 1G2E). The protein is shown as gray and RNA as orange cartoons. RRM1\_RNP1 is shown in pink, RRM1\_RNP2 in red, RRM2\_RNP1 in blue, and RRM2\_RNP2 in green. (Image generated with PyMOL v1.7.4.5 Edu, Schrödinger LLC)

It is very important at this point to underline once again that RNP sequences are highly conserved in the ELAV protein class. In fact, by comparing the four sequences for subtype HuD and HuR we notice only two differences in residue substitution, namely Ala for Gly in RRM2\_RNP1 and Ile for Val in RRM2\_RNP2. Taking this into account, together with the

conclusions of the study illustrated in paragraph 1.4.1, the considerations made for the HuD peptides could be extended to the whole family.

### ***1.5 Postulated binding site for small molecules: preliminary docking of literature compounds on subtype HuR***

Based on these considerations and the literature recollection of compounds interfering with ELAV–RNA complexes shown in paragraph 1.3 (Figure 5), a preliminary molecular modeling study was performed by our research group on subtype HuR RRM1/2, chosen as model (PDB 4ED5).<sup>26</sup> Aim of the investigation was to gather all the literature compounds and, despite their structural diversity and different experimental data, submit them to the same docking process to hypothesize a common binding site. This region can reasonably be identified between the RRM domains, especially in the site hosting the RNA recognition sequence (AUUUA). Exactly at this point, the two asymmetric units of the protein are in proximity, providing a deeper and narrower pocket, which is partially basic due to the presence of arginine and asparagine residues.

Starting point was the co-crystal structure of HuR RRM1/2–RNA<sup>c-fos</sup> (PDB 4ED5). The compounds were subjected to the investigation using the FlexX docking module in the LeadIT suite (BioSolveIT GmbH) and the following evaluations exploiting the scoring function HYDE in SeeSAR (BioSolveIT GmbH). The binding site in the protein was restricted to 6.5 Å around the co-crystallized RNA, the 30 top-scored (FlexX) solutions were retained and subsequently post-scored with SeeSAR, and the best-scored poses were then selected. Results of these explorations revealed that, though they occupy different sites of the interaction site, most compounds bind to the same pocket on HuR, benefiting from interactions with the RNP1/2 sequences of the RRM1/2 domains of the protein in the deep part of the pocket. In accordance with experimental data, **23** did not show interaction with the RRM1/2 domains. Aside for **23**, each compound can be superimposed with at least one nucleotide of the RNA<sup>c-fos</sup> sequence, with U4/U8 as the most represented.<sup>26</sup>

Accordingly, these scouting docking studies identified the possible interactions of compounds already recognized as inhibitors of ELAV–RNA and hinted at possible complex breaking points; moreover, they showed the potential of using computer-aided drug design (CADD) in future drug-discovery programs.

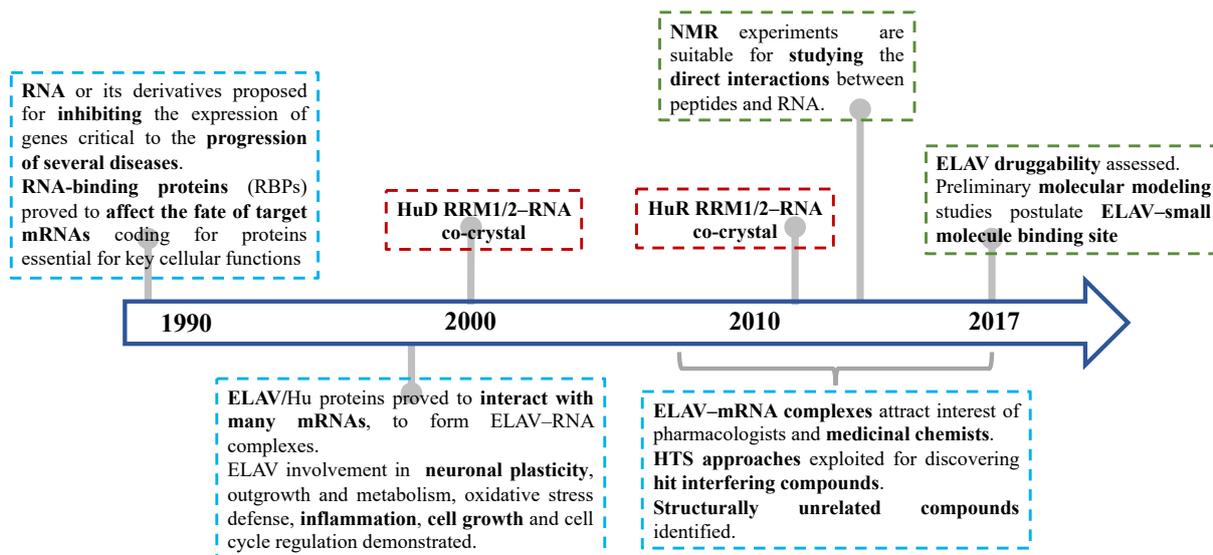


Figure 8. Summary timeline to recap the milestones in ELAV research at the time this thesis was started.

## 2. *Aim*

The ultimate goal of the project in which this thesis took part is to develop a new series of small, drug-like organic molecules able to modulate the stability of protein–RNA complexes involved in several diseases, thus regulating gene expression with an unprecedented mode of action and opening the way to a new class of therapeutic agents.

Within this context, the present thesis focuses on the ELAV RBP family, and in particular on subtype HuR and its complexes with RNA due to their involvement in numerous dysfunctions (*i.e.* cancer, retinopathy, *etc.*).

Given the early stage of this research topic, the research gaps in the state-of-the-art at the moment this thesis was started (October 2016), and the highly challenging character of the target, the aim of the present PhD project was split into three “sub” aims to reach ideally in series and consequentially but with the potential to be carried out in parallel and separately:

1. Establishment of a systematic biophysical and *in silico* approach exploiting complementary and informative methods to apply to all obtained compounds;
2. Design and synthesis of new ligands according to (a) a fragment-based (fragment screening and linking) approach followed by the application of DCC, or, in case of failure, (b) a structure-based (*in silico*) approach.

Finally, to perform the biophysical assessment of the interaction between compounds and target protein (HuR), to select compounds for biological evaluation.

### ***3. Methodologies and Strategies***

As stated in the previous chapter, aim of the present thesis is to establish a systematic and informative approach to evaluate the interaction between potential ligands/inhibitors and the chosen target HuR, and to exploit this approach and the deriving data to design and synthesize new compounds and iteratively evolve them to develop new molecules as HuR ligands able to modulate its complexes with RNA. From the recap provided in introduction, it is clear that our knowledge of the HuR–RNA–small molecule system is still very preliminary. For this reason, to tackle this challenging target, different and complementary techniques and approaches were employed in the research work related to this thesis.

Here below, the main methodologies and strategies applied will be briefly introduced and described for clarity purposes. The chapter articulates in different paragraphs, starting from the main technique chosen for the project, NMR, additional biophysical approaches exploited, surface plasmon resonance and thermal shift assay, and a brief paragraph on supporting *in silico* methodologies used. The following paragraphs concern the two approaches exploited for HuR ligand discovery, structure-based and fragment-based, and the chapter ends highlighting three useful strategies, multi-component reactions, dynamic combinatorial chemistry and microwave heating which could be applied to accelerate ligand discovery by affording diverse and rich compound libraries.

#### ***3.1 NMR techniques to study small molecule–macromolecule interactions***

Biological functions generally depend on the interaction between ligands and macromolecular targets (proteins, nucleic acids, *etc*) and to understand the biological function we need to study the underlying recognition events at an atomic level. Several biophysical methods have been developed to supply to this very need, as understanding how a certain macromolecule interacts with natural or known ligands, allows us to develop other more potent substances that may then serve as pharmacological tools or even drug candidates. Among these, NMR techniques are recognized as powerful techniques to understand the binding process at a molecular or atomic level as they occur in solution. They can be exploited for several purposes in medicinal chemistry and drug discovery: a) screening compound libraries, b) obtaining detailed identification of the groups involved in the interaction, thus gaining essential structural information, c) characterizing thermodynamic and kinetic parameters of the binding events, and finally d) identifying competitive binding as well as allosteric effects.<sup>79</sup>

While, in principle, all NMR parameters could be exploited for monitoring the binding activity of a ligand to a macromolecule, in practice, only those that can be obtained easily and with high sensitivity are of significance; among these, are changes in chemical-shift, relaxation times,

diffusion constants, nuclear Overhauser effect (NOE), or transfers of saturation. Different experimental formats have been thus developed, which can be divided into two main groups, those based on observing the binding from the small molecule perspective (ligand-based or ligand-observed) and those focused on the macromolecule signals (target-based or target-observed). The first approach usually utilizes NOE effects between macromolecule and ligand, while the second focuses on chemical-shift changes of the target upon binding of the ligand; the two approaches can thus be considered complementary to build a complete map of the interaction and have different advantages and limitations.

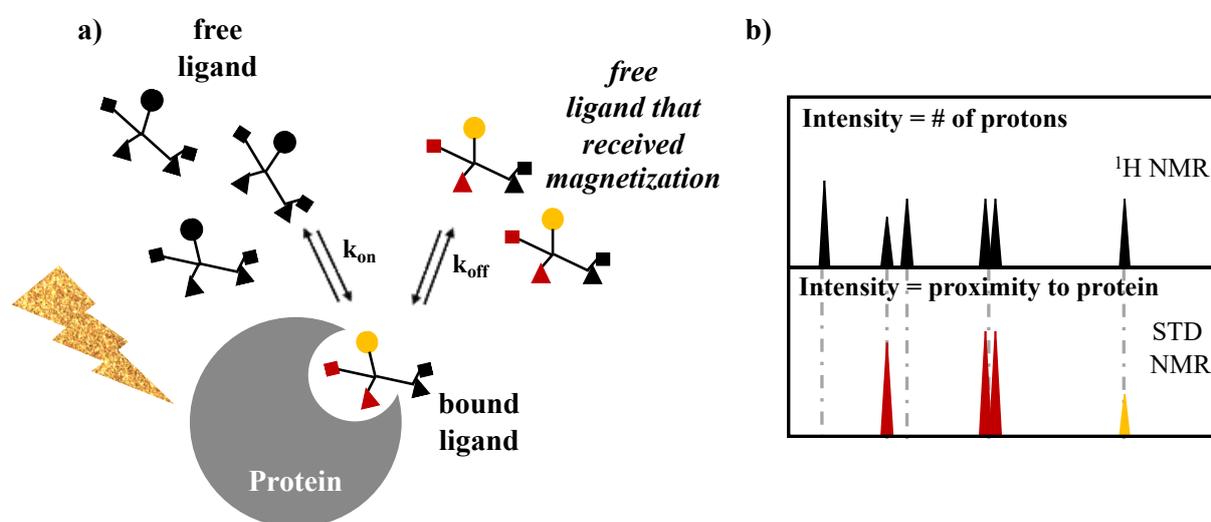
A complete and more detailed panoramic on this subject is reported in several articles and reviews (Meyer *et al.*, 2003;<sup>79</sup> Ciulli, 2013;<sup>80</sup> Mittermaier *et al.*, 2013;<sup>81</sup> Li *et al.*, 2017;<sup>82</sup> Sugiki *et al.*, 2018;<sup>83</sup> Barile *et al.*, 2019;<sup>84</sup> Maity *et al.*, 2019;<sup>85</sup>). In the following sections, the focus will revolve around ligand-observed technique saturation transfer difference (STD) NMR and its very recent derivation differential epitope mapping (DEEP) by STD NMR, exploited to study the interaction between small molecules and target protein HuR.

### **3.1.1 Ligand-observed NMR: Saturation Transfer Difference (STD) NMR**

In the context of ligand-based NMR, saturation transfer difference (STD) NMR has become a leading technique to characterize ligand–macromolecule interactions as they occur in solution.<sup>79,86</sup> For this reason, it is considered a powerful tool for screening compounds, identifying lead structures or ligand moieties important for binding to target.<sup>87</sup> The STD experiment is based on the NOE effect and exploits the transfer of the magnetization from the macromolecule, selectively irradiated, to the ligand. Rather than gaining a only yes/no answer on the binding, STD can also afford information on the structural features of the binding event, as those regions of the ligand which are in closer contact with the macromolecule receive the magnetization more efficiently than those which are farther from it. Practically, this means that protons of the ligand which receive a higher magnetization produce more intense signals on a 1D <sup>1</sup>H-STD spectrum (see Figure 9). In the STD NMR literature, these protons as a whole are often referred to as binding epitope; the epitope mapping can be performed on single ligands, as well as to rank and determine best binding ligands in a mixture (with the condition that each component of the mixture has been assigned).<sup>88</sup> A sample containing the macromolecule at low concentration and a large molar excess of the ligand (1:50 up to 1:1000) is generally employed to preclude the perturbations of absolute STD intensities due to rebinding effects (*i.e.*, a ligand already saturated associates with the target again, without prior full relaxation), which would compromise the correct group epitope mapping. Additionally, STD NMR is ideal for studying

relatively weak binding events (dissociation constant in a low  $\mu\text{M}$  to  $\text{mM}$  range) in which there is exchange between the bound and the free ligand state.

To obtain the STD spectrum two spectra are acquired: one in which the macromolecule has been selectively saturated (on-resonance spectrum obtained by irradiating at a region of the spectrum that contains only resonances of the macromolecule with signal intensities  $I_{\text{SAT}}$ ), and one recorded without macromolecule saturation (off-resonance spectrum, with signal intensities  $I_0$ ). Though the resonances of macromolecules can be quite shifted from those of small molecules, the chemical shift ranges considered for a small molecule  $^1\text{H}$  NMR still apply to a certain degree. If we consider a protein as macromolecular target, formed by different amino acids, we can irradiate in both aromatic and aliphatic regions of the spectrum where the variety of amino acid side chains can receive the magnetization. Most often, the selected frequency for STD experiments lies in the range between  $-2$  ppm and  $0.5$  ppm where most of the aliphatic resonances are found within proteins, with a relatively low chance of irradiating the small molecule contained in the sample, thus preventing artifacts. Nonetheless, blank experiments (spectra acquired on samples containing the sole small molecules) should always be performed to optimized the irradiation frequency.



*Figure 9.* The STD NMR experiment. (a) The macromolecule (in the example shown, a protein) is selectively irradiated and transfers the magnetization via spin diffusion, through the NOE effect, to the bound ligands. The ligand which has received the magnetization goes back into solution; (b) Comparison between 1D  $^1\text{H}$  NMR and  $^1\text{H}$ -STD NMR spectra: though visually similar, peak intensities in the  $^1\text{H}$  NMR are related to the number of protons resonating at that chemical shift while, in the  $^1\text{H}$ -STD NMR spectrum, peak intensities are related to the vicinity of single protons to the protein. Protons closer to the protein give higher signals while protons which have lower interaction with the protein give progressively lower signals (down to no signal with non-interacting protons).

The STD spectrum is thus the difference spectrum between off-resonance and on-resonance spectra ( $I_{\text{STD}} = I_0 - I_{\text{SAT}}$ ) (see Figure 9 b). Further processing for building epitope maps within

single ligands involves the calculation of two related values, absolute and relative STD intensities; both values are normally given as percentages. To obtain absolute STD %, the value of each peak area in the STD spectrum needs to be compared with the same peak area in the off-resonance spectrum: for each peak, absolute STD % =  $(\text{area}_{\text{STD}} / \text{area}_0) * 100$ ; in case of multiple protons under the same chemical shift this number needs to be normalized. To facilitate comparison of protons within a single molecule, and to compare epitope maps within a series of structurally-related compounds, relative STD % can be subsequently calculated. Within a molecule, the proton with the highest absolute STD % is given the arbitrary value of 100%; the values of the other protons are then calculated relative to this proton. By dividing different ranges of relative STD % and specific codes (*e.g.* color or size), it is possible to give visual representations of these maps for quick examination.

While within the same spectrum, higher absolute STD values are symptomatic of a higher affinity for the macromolecular target and this is often exploited for the screening of compound/fragment mixtures, it is not commonly accepted to compare absolute STD values of different spectra of different compounds to rank their affinity to the macromolecule. Nevertheless, if the solubility and stability of the tested compounds are favorable, STD NMR can be used, like many other biophysical techniques, for estimation of the dissociation constant ( $K_D$ ) which can be utilized for ranking.<sup>89</sup> Compared to other techniques, NMR generally needs relatively high concentrations of compound and macromolecule in the aqueous environment utilized for an experiment, thus limited solubility and macromolecule instability/aggregation can represent limitations for some systems. Additionally to  $K_D$  determination, STD NMR can also be used for competition experiments to evaluate displacement of known ligands/inhibitors or to confirm binding sites.<sup>90</sup>

Regardless of the many advantages, in most cases STD NMR gives only partial information (only from the ligand perspective) and benefits from the coupling with either target-based NMR (when available) and, more commonly, with *in silico* approaches such as molecular dynamics (when X-ray crystallography data on the binding site are available). In the first case, complementary information about small molecule–macromolecule contacts can be compared to build a map of the interaction on both counterparts. In the second case, STD NMR can help to validate predicted binding poses.

### **3.1.2 Ligand-observed NMR: Differential epitope mapping (DEEP) by STD NMR**

Very recently, a new STD-based methodology was developed by Angulo and collaborators called differential epitope mapping (DEEP) by STD NMR.<sup>91</sup> While it remains a STD ligand-based experiment with its own limitations, it can be exploited to gain additional data to the

simple STD experiment thanks to the application of differential conditions. Namely, it can identify the type of protein residues contacting the ligand through the generation of differential epitope maps, it can readily afford pharmacophore information on the protein, and, if the 3D structure of the protein is known, it helps in orienting the ligand in the binding pocket, with a higher precision compared to the single STD experiment. In the first article reporting this new method, two proofs of principle were performed on two protein–ligand systems with available co-crystals; the results of the two methods (DEEP STD and X-ray diffraction) were in agreement, showing the great potential of this new technique.<sup>91</sup> Since then, few more papers exploiting this technique were published.<sup>92,93</sup>

Two types of differential experiments can be performed and they are complementary to each other: a) exploitation of differential irradiation frequencies; b) utilization of differential solvent composition.

a) Differential irradiation frequency: As mentioned in the previous paragraph, in an STD experiment, the on-resonance spectrum requires selective irradiation of the macromolecular target. Within a binding site, different amino acids are usually present. By performing different experiments (two are sufficient for this protocol) varying the irradiation frequency, we can obtain different STD spectra which, properly processed and plotted, can afford differential epitope maps. Briefly, while a proton close to the protein will in both cases have an enhanced STD signal, if it interacts with higher affinity with a specific type of amino acid side chain, we should be able to identify a difference in the two spectra showing us the different relevance of the irradiated residue-type on the binding of that proton to the protein.

Clearly, while it is possible to pick the general type of amino acid to irradiate, two points must be raised to take into account: i) the irradiation cannot selectively be directed at the binding site, so all those amino acids will be irradiated. Moreover, to a certain degree, the magnetization can move along the protein. Nevertheless, though this could cause some aspecific components to emerge, a “good” ligand, orthosteric to the binding site considered, generally shows higher affinity for that binding site. For this reason, the population of ligands that have received the magnetization from one specific binding site, where the encounters have higher probability to occur, is generally the main component of the signal. ii) To selectively irradiate the macromolecule, the correct frequencies have to be picked where only the macromolecule is irradiated directly while the ligand is not. This is of particular importance to avoid artifacts, in fact, set up experiments to pick the irradiation should always include blanks – only compound samples. As mentioned above, for proteins, both aromatic and aliphatic regions can be irradiated; at the same time, these are structural features very common in small organic

molecules for drug discovery. The experiments can still be performed thanks to subtle shifts in macromolecule resonances but, as already stated, special care has to be put into irradiation frequency selection.

b) Differential solvent composition: macromolecular systems, as many proteins, can generally only be studied within a small range of solvents which do not alter their native conformation motion range and in a prevalently aqueous environment very little variations are thus allowed. Nonetheless, it is still possible to perform a differential experiment through utilization of a completely deuterated water buffer and of a prevalently non-deuterated water buffer. In D<sub>2</sub>O, the polar side chains in the binding pocket have their exchangeable protons replaced by deuterium which is inefficient for transferring the saturation. On the other hand, in a prevalently H<sub>2</sub>O environment, these protons can contribute to an additional transfer of saturation. This process depends on their exchange rate with bulk water, with slowly exchanging polar protons producing the largest variations. For this reason, in H<sub>2</sub>O, ligand protons contacting polar residues show a relative increase in STD values compared to the binding epitopes in D<sub>2</sub>O. Once again, acquiring the spectra in both conditions can afford us differential epitope maps, very useful to correctly place the ligands in the binding site confirming *in silico* poses. This approach can thus give complementary results to the previous by highlight the influence of polar residues.

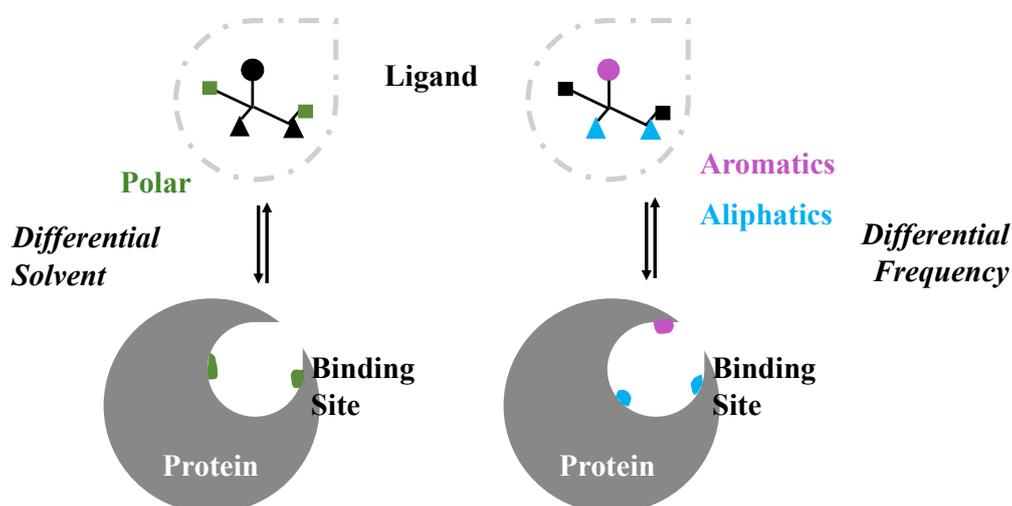


Figure 10. Graphic representation of the two protocols of DEEP-STD NMR. Differential Solvent composition (D<sub>2</sub>O vs H<sub>2</sub>O) can highlight ligand interactions with polar amino acid residues in the binding site. Differential irradiation frequencies can show interactions of the small molecule with amino acids with different structural features (*e.g.* aromatic vs aliphatic) (Adapted from Monaco *et al.*, 2017)<sup>91</sup>

Correct processing of the spectra is key to obtaining valuable information about the macromolecule–small molecule interaction. Briefly, after acquisition of the two spectra at the two different conditions to be compared, absolute STD percentages should be calculated for each proton; from these values,  $\Delta\text{STD}_i$  can be calculated for each proton:

$$\Delta\text{STD}_i = \frac{\text{STD}_{exp1,i}}{\text{STD}_{exp2,i}} - \frac{1}{n} \sum_i^n \left( \frac{\text{STD}_{exp1,i}}{\text{STD}_{exp2,i}} \right)$$

$\Delta\text{STD}_i$  = differential STD value  
 $exp1$  and  $2$  = two different experimental conditions  
 $i$  = each proton

To obtain a consistent scale of  $\Delta\text{STD}_i$  factors,  $exp1$  should always be the experiment which shows larger total ligand saturation. The highlighted section of the equation represents the intrinsic differences in saturation levels under different conditions. For this reason, contrary to STD intensities which are always positive numbers or 0, in DEEP STD processing, values can be positive and negative, as well as numbers close to 0. Positive values indicate higher influence of the  $exp1$  conditions, while negative values imply higher impact of the  $exp2$  conditions on a determined proton. Finally, the protons showing low variation due to exp conditions will give values in a range around zero; during processing, it is quite important to define a cut-off of statistical significance for these values. The DEEP epitope maps can be built similarly to STD epitope maps with color and size codes identifying the different type of amino acid residue interacting with each interacting proton and thus improving the predicted placement of the ligand inside the binding site.

### 3.2 Other biophysical approaches

Biophysical techniques are of great importance to study small-molecule ligand interactions, from the screening of libraries to select hits, to their validation from both structural (binding mode) and energetic (binding affinity) stand-points. Though this PhD thesis mainly relies on ligand-observed NMR techniques for studying ligand–protein interactions, due to their highly informative nature, these also present some limitations. Applying a combination of different methods is often considered to produce the best chance for leading a robust and effective identification of high-quality starting points for medicinal chemistry campaigns for both drug discovery and development. For this reason, other methods were considered in different steps of this thesis. The following paragraphs will present surface plasmon resonance (SPR) and differential scanning fluorimetry or thermal shift assay (DSF, TSA). A brief comparison of their main features with NMR techniques (discussed in the previous paragraph), and other additional methods, isothermal titration calorimetry and protein X-ray crystallography, is also reported in tables 2 and 3 to provide a quick yet detailed panoramic view over the characteristics, advantages and disadvantages of some of the most informative techniques available.

<b>Technique</b>	<b>Screening throughput</b>	<b>Material consumption</b>	<b>Detectable <math>K_D</math> range</b>	<b>Binding site information</b>
<b>DSF (TSA)</b>	High	Medium	Up to 5 mM	–
<b>NMR</b>	Medium	Medium	Low nM–10 mM	Good
<b>ITC</b>	Low	High	Low nM–5 mM	(Limited to competition)
<b>SPR</b>	Medium	Low	pM–2 mM	(Limited to competition)
<b>X-ray crystallography</b>	Low	Medium	Up to ligand solubility	Excellent

*Table 2.* Summary of parameters of different biophysical techniques utilized for small molecule screening and interaction studies with macromolecular targets (adapted from Ciulli, 2013).<sup>80</sup>

Technique	Advantages	Disadvantages
<b>DSF (TSA)</b>	<ul style="list-style-type: none"> <li>High throughput</li> <li>Applicable to most target proteins</li> <li>Direct binding assay</li> </ul>	<ul style="list-style-type: none"> <li>Prone to false positives and negatives</li> <li>High material consumption</li> <li>Across-plate variability</li> </ul>
<b>ligand-observed NMR</b>	<ul style="list-style-type: none"> <li>Medium throughput</li> <li>Applicable to most target proteins (&gt; 10 kDa)</li> <li>Provides quality control</li> </ul>	<ul style="list-style-type: none"> <li>Prone to false positives (compound aggregation or aspecific effects)</li> </ul>
<b>protein-observed NMR</b>	<ul style="list-style-type: none"> <li>Intermediate throughput</li> <li>Can identify binding site (need peak assignment)</li> <li>Can measure <math>K_D</math> from ligand titrations</li> </ul>	<ul style="list-style-type: none"> <li>Limited to small (&lt; 30 kDa) and soluble target proteins</li> <li>Requires <math>^{15}\text{N}</math>-labelling of target proteins</li> </ul>
<b>ITC</b>	<ul style="list-style-type: none"> <li>Direct/competition binding assays</li> <li>Applicable to most target proteins</li> <li>High information content (<math>K_D</math>, <math>\Delta H</math>, <math>\Delta S</math> and <math>n</math>)</li> </ul>	<ul style="list-style-type: none"> <li>High material consumption</li> <li>Low throughput</li> <li>High material consumption</li> <li>Requires large heat changes upon binding for reliable measurements</li> </ul>
<b>SPR</b>	<ul style="list-style-type: none"> <li>Label-free detection and ease of automation</li> <li>Applicable to most target proteins</li> <li>Direct/competition binding assays</li> <li>Low material consumption</li> <li>High information content (<math>K_D</math>, <math>k_{on}</math>, <math>k_{off}</math> and <math>n</math>)</li> </ul>	<ul style="list-style-type: none"> <li>Requires immobilization of macromolecular target</li> <li>Possible artifacts due to compound aggregation or aspecific effects</li> <li>Time-consuming optimization</li> </ul>
<b>X-ray crystallography</b>	<ul style="list-style-type: none"> <li>Intermediate throughput</li> <li>Can rapidly identify binding site and binding modes</li> <li>Can directly identify conformational changes</li> </ul>	<ul style="list-style-type: none"> <li>Limited to soluble target proteins that can be crystallized</li> <li>Requires expensive X-ray sources (in-house, access to synchrotrons)</li> <li>Crystal packing may occlude binding site</li> <li>High occupancy of the ligand binding site required</li> </ul>

Table 3. Summary of relative advantages and disadvantages of different biophysical techniques with regard to small molecule screening, validation, and characterization of the binding to macromolecular targets (adapted from Ciulli, 2013).<sup>80</sup>

### 3.2.1 Surface Plasmon Resonance

Surface plasmon resonance (SPR) is an optical technique based on the transfer of energy to electrons in a thin layer of metal in contact with a solution. In the standard SPR set up, a beam of polarized monochromatic light is shone through a prism at a thin-layer of gold coating one surface of the prism. The light is then reflected at the gold-coated surface, but, since it is not reflected precisely at the prism-gold junction, its electromagnetic field penetrates into and beyond the gold for a certain distance (formation of the evanescent wave). At a specific angle of incidence, the electrons in the gold absorb some of the light exciting charged density waves or “surface plasmons” which propagate along the metal surface. At a resonance condition where transfer of momentum matches that of the plasmons *i.e.*, at a specific incident angle, the intensity of the reflected light is significantly decreased. Since the evanescent wave extends only ~100–200 nm into the solution (after which it decays quickly) if the gold layer is sufficiently thin, the resonance condition/angle is dependent on the refractive index of the adjacent medium. For this reason, SPR is highly sensitive to changes in the environment at the interface between gold and aqueous flowing medium; a change in the refraction index at the surface of the sensor (due to interactions occurring near the surface) can be monitored as a shift in the resonance angle.<sup>80,94,95</sup>

A typical SPR setup involves a “chip” composed by a glass surface coated by a thin layer of gold; the solution side of the gold film is functionalized with a dextran matrix (covalently attached through a linker layer) to allow immobilization of target molecules (usually called ligands in SPR experiments) such as proteins (see Figure 11 a). Immobilization of the protein to the sensor surface is required and is of primary importance to the design of a successful assay; several immobilizations can be considered depending on the studied target to obtain a highly stable association and prevent signal drift.<sup>96,97</sup>

As mentioned above, when an analyte (the ligand studied) in solution binds to the protein, the refractive index near the surface changes and an SPR shift is detected, which can be monitored in real time producing a “sensogram”. A schematic representation of sensogram is shown in Figure 11 b where signal increases until the protein binding sites are saturated (equilibrium), and successively only buffer is flowed over the chip progressively removing the interacting analyte. Due to the dependence of the refractive index, and thus the SPR signal to the mass bound at the surface, it is possible to measure the thermodynamic and kinetic parameters of the interaction.

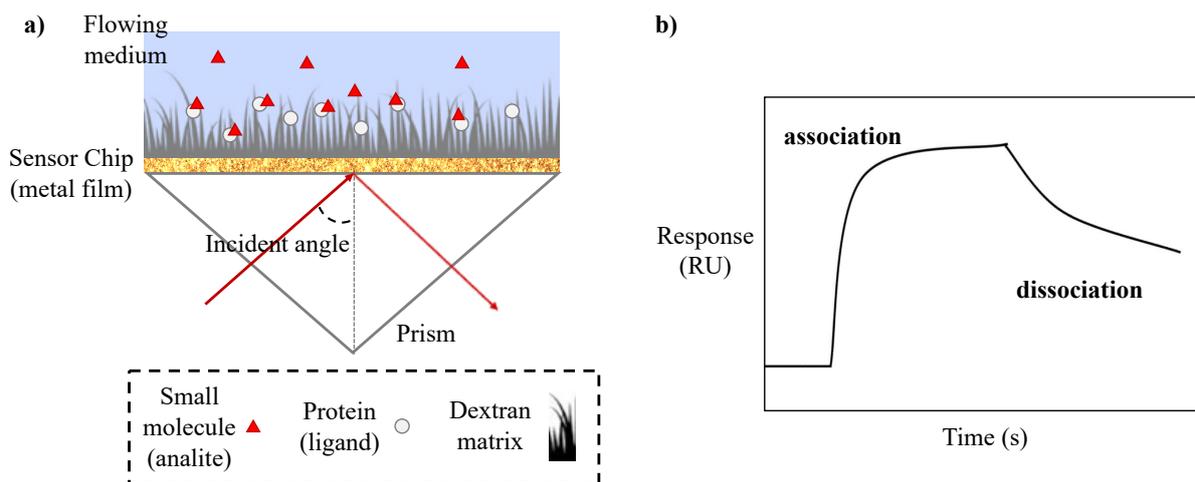


Figure 11. (a) A typical SPR biosensor setup. The glass of the sensor chip is coated with a thin layer of gold, a dextran matrix is attached via a linker layer to the gold and receptors are cross-linked to the dextran. An incident light is directed onto a SPR sensor chip via a prism, and the reflected beam is detected via a photodetector or imager. At an appropriate angle (resonance angle), the incident light excites the surface plasmons in the sensor chip (metal film) and the intensity of the reflected light drops to a minimum. The electromagnetic field produced penetrates the medium and probes molecular binding processes taking place on the surface: changes in the refractive index of the flowing medium are observed that are proportional to the amount of material bound. A plot of resonance signal vs time, the sensogram (b), can be monitored in real time (adapted from <http://biosensingusa.com>, 2019).

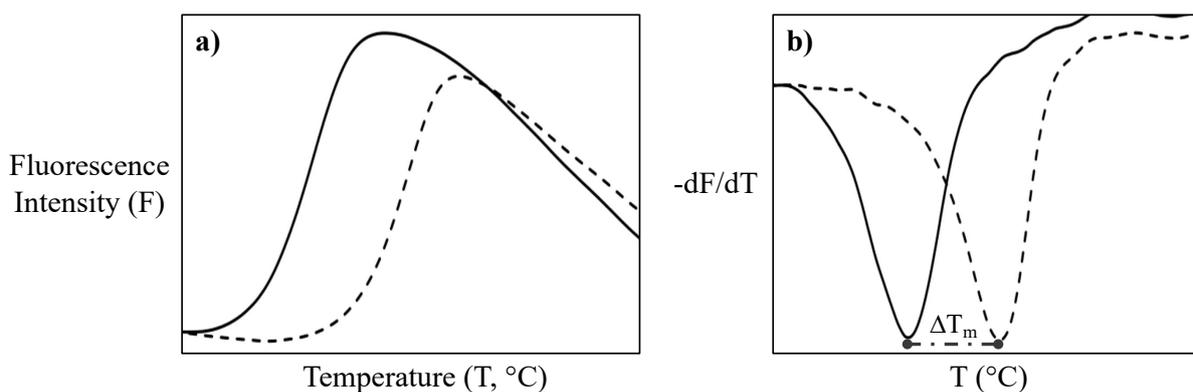
Both the on-rate and off-rate constants of the binding process can be determined and their ratio gives an accurate estimate of affinity. Generally, the latter can also be evaluated through injection of a range of analyte concentrations producing different sensorgrams, and affinity can be determined from the ligand concentration dependence of the response at equilibrium of the ligand-complex formation. SPR is now an established method for primary fragment<sup>98–100</sup> and small molecule library screening in addition to its exploitation for studying several binding events (*i.e.*, protein–protein interactions)<sup>94,101,102</sup> thanks to its experimental versatility, customization of chip matrixes and immobilization strategies available.

### 3.2.2 Thermal Shift Assay

Since its first description in 2001,<sup>103</sup> differential scanning fluorimetry or thermal shift assay (DSF, TSA, also found as thermofluor) has been used increasingly for drug discovery. In this methodology, proteins are incubated with a fluorescent dye which binds specifically to hydrophobic sites; when this binding occurs, a progressive change in the dye fluorescence can be recorded. When a protein undergoes a thermal cycle (heating), it gradually exposes its hydrophobic parts due to temperature-dependent unfolding. The dye will consequently bind to a growing number of hydrophobic regions of the protein; monitoring the change in fluorescence against the rising temperature gives characteristic traces (see Figure 12) allowing to determine the temperature of 50% protein unfolding (commonly called melting point,  $T_m$ ). In solution,

proteins exist in thermodynamic equilibrium between several different conformational states; upon interaction with a ligand the populations of these states can be altered. If we consider a two-state system with a folded (native) and an unfolded (denatured) state, the population of the unfolded state will grow with increasing temperature. Binding of a small-molecule to a structurally defined site of the native state of a protein can stabilize it, increasing the population of the protein native state during a thermal cycle, producing a shift in the protein  $T_m$  to higher values. By comparing  $T_m$  measurements, in the presence and absence of a small molecule, it is possible to determine protein–ligand binding. Similarly, TSA can be applied to determine optimal experimental conditions for protein stability (*i.e.*, type of buffer, pH, ionic strength, DMSO percentage, and so on).<sup>104–106</sup>

This experiment can be carried out in small volumes in any commercial quantitative PCR instrument, using different formats (48, 96 or 384-well plates). Given the length of the experiments (1–2 hours), this method can afford high throughput analyses.<sup>107,108</sup> A plot of fluorescence signal against temperature should give a sigmoidal plot (see Figure 12 a). The melting temperature is determined by the point of inflection of this curve. This can most easily be assessed by plotting the derivative of the fluorescent signal against temperature (see Figure 12 b). Though this method is mainly utilized for primary ligand screenings, it also has the potential for  $K_D$  determination, by preparing wells with ranges of ligand concentration, and, thanks to advances in related software, these measurements can now be performed with increased ease and reliability.<sup>107</sup>



*Figure 12.* Differential melting behavior of a protein and a protein–ligand complex by fluorescence. The solid line identifies the protein melting curve and dashed line that of the protein–ligand complex. a) TSA trace of a interaction study. b) Derivative of the fluorescent signal plotted against temperature. The minimum of the derivative plot (corresponding to the inflection point of the F vs T plot) allows a more convenient identification of the melting temperatures  $T_m$ , necessary for the measurement of the thermal shift  $\Delta T_m$  (adapted from Ciulli *et al.*, 2008).<sup>109</sup>

### **3.3 *In silico methodologies: Docking and Molecular Dynamics***

As stressed in previous chapters, the binding of small molecule ligands to macromolecular targets is central to numerous biological processes and the accurate prediction of the binding modes between the ligand and target, is of great importance for the study of these mechanisms, for the rationalization of biophysical studies and for related drug design. The number of algorithms available for assessing and rationalizing ligand–target interactions is large and ever increasing. Though several do share common methodologies, their diversity in both complexity and computational speed has afforded a range of techniques to solve modern structural puzzles.<sup>110</sup> Assuming the structure of the target macromolecule is available, a primary challenge is to predict both ligand orientation and binding affinity, generally referred to as molecular docking. Molecular docking aims at computationally simulating the molecular recognition process; it has been commonly described as a “lock-and-key” model, as we want to find the correct relative orientation of the “key” (ligand) which will open up the “lock” (target macromolecule) finding out where the key hole is, how to turn the key inside the hole, and so on. Nonetheless, since both ligand and macromolecular target are flexible, a “hand-in-glove” analogy might be more fitting than “lock-and-key”.<sup>111</sup>

During the interaction, both ligand and macromolecule adjust their conformation; in fact, the atoms in a biomolecule are in constant motion and both molecular function and intermolecular interactions are correlated to the dynamics of the molecules involved. Rather than studying just a static snapshot, the ideal observation should consider these biomolecules in action, to perturb them at the atomic level, and to see how they respond. Since this would be rather difficult and time-consuming, an attractive alternative is to work with an atomic-level computer simulation of the target macromolecules under study.<sup>112</sup> In this context, molecular dynamics (MD) simulations can come in play, as they can predict how every atom in a protein or other macromolecular system will move based on a general model of the physics involved in interatomic interactions.<sup>113</sup> Exploiting these simulations we can capture a large variety of important biomolecular processes, including conformational change, ligand binding, and protein folding as a function of time.<sup>114</sup> MD simulations are often used in combination with a wide variety of experimental structural biology techniques, including X-ray crystallography, cryogenic electron microscopy (cryo-EM), and NMR.<sup>112</sup>

### **3.4 *Drug Discovery Strategies***

One of the main steps in the search for small-molecule drugs is to identify one or more new chemical leads, molecules whose good potency in biological assays reflects the targeted mechanism. Leads are also characterized by other chemical, physical and biological properties

important for drugs and drug-like molecules (*i.e.*, favorable absorption, distribution and metabolism profile, ADMET).

A number of medicinal chemistry approaches can be used to identify hits, generate leads, as well as to accelerate the development of high quality drug candidates,<sup>115</sup> such as modification of an existing lead or drug, a natural product or an endogenous ligand, and structure-based and fragment-based approaches.

In the following paragraphs, the approaches utilized for the generation of HuR ligands will be described.

### **3.4.1 Structure-based ligand design**

By understanding the structural and chemical binding features of drug targets pivotal in biologically relevant pathways, we can design small molecules able to modulate specific target functions closely connected to human diseases and dysfunction.<sup>116–118</sup> Identifying these promising hits to develop and optimize is a major challenge faced for researchers and although the trial and error nature is intrinsic in drug research, rationalization and *in silico* methodologies are widely exploited in this field. In this context, Structure-Based Drug Design (SBDD) is a well-established, positively successful strategy based on three-dimensional (3D) protein structure information to develop new active molecules.<sup>119–121</sup> Understanding of the spatial and energetic aspects of protein ligand interactions is crucial for structure-based studies; assuming the shape and chemical nature of the binding site of the target and the possible interactions between ligands and the protein have been identified, this information can be directly employed to identify new ligands or optimize lead compounds. The aim is to boost the search for lead molecules at the same time limiting the number of compounds to be evaluated experimentally. Hits can be identified by docking small molecule ligands (*i.e.*, from databases) into protein binding sites or by using target-based pharmacophore models. Moreover, potential drug candidates can be designed *de novo* by improving the complementary properties of lead compounds and the respective target. The target-based approach can, in principle, be carried out iteratively, through multiple computational and experimental steps until an optimized lead compound with high affinity, selectivity, and optimized pharmacokinetic properties is developed.

### **3.4.2 Fragment-based ligand design**

Fragment-based drug discovery (FBDD) is a relatively recent approach, emerged as an alternative to HTS,<sup>122,123</sup> increasingly used in medicinal research, for reducing attrition and providing leads for challenging biological targets. FBDD identifies low-molecular-weight ligands (~150 Da) which can bind to biologically relevant macromolecules. Compared with

HTS, the fragment-based approach requires fewer compounds to be screened, and, despite the lower initial potency of the screening hits, it can offer more efficient optimization campaigns.<sup>115</sup> In FBDD, fragment libraries are typically screened using a series of progressively more restrictive biophysical techniques. The fragments are between 150–250 Da MW and a general unfocused fragment library will be chosen to contain a diverse set of chemical scaffolds and functionalities that are well preceded in drugs.<sup>122</sup> The experimental binding mode of fragment hits is usually established in the protein target of interest, the 3D experimental binding mode of these fragments can be determined using X-ray crystallography or NMR spectroscopy, and used to facilitate their optimization into potent molecules with druglike properties. Fragments are grown to form new interactions using structure-based drug design. Fragment linking can also be used if different fragments are observed to bind to proximal sites in the protein. Although initial fragment hits have low potency due to their small size, they form high-quality interactions and can be readily optimized into potent lead molecules.<sup>115</sup> Regardless of the initial skepticism, more than ten FBDD leads targeting different protein families involved in diverse diseases have made it into clinical trials (see references<sup>123,124</sup> for examples). The fragment approach has been enthusiastically taken up in different areas of drug discovery research (academia, large pharma and biotech companies) and a number of rule-of-three-compliant<sup>125</sup> fragment libraries are now commercially available.

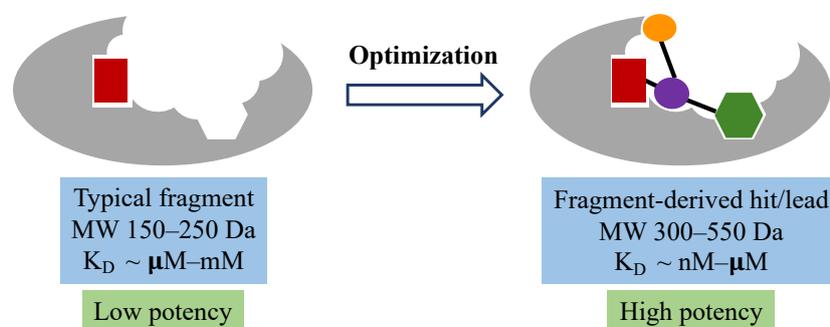


Figure 13. Schematic representation of the fragment-based ligand design approach (adapted from Murray *et al.*, 2009).<sup>115</sup>

### 3.4.3 Multi-component Reactions

Multicomponent reactions (MCRs) are chemical transformations in which three or more reactants form a product whose structure derives all of the components added.<sup>126</sup> Although ‘cascade’, ‘domino’ and ‘one-pot’ processes are generally included in MCRs, a more specific definition can help to understand the main differences and advantages. Performing successive reactions in a vessel can be an efficient way to access molecular diversity, but a carrying out a series of ‘two-component’ reactions following a specific order of addition is radically different than many MCRs in which the components can be added in any order.<sup>127–129</sup>

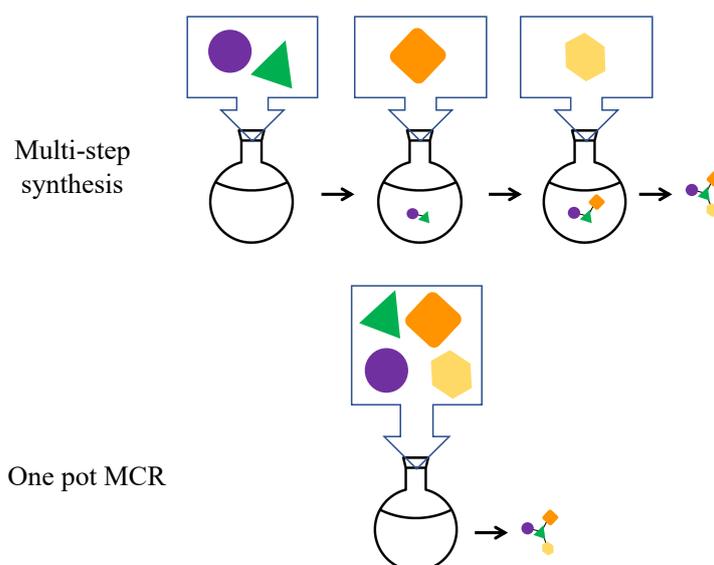


Figure 14. Schematic comparison of multi-step versus multi-component synthesis.

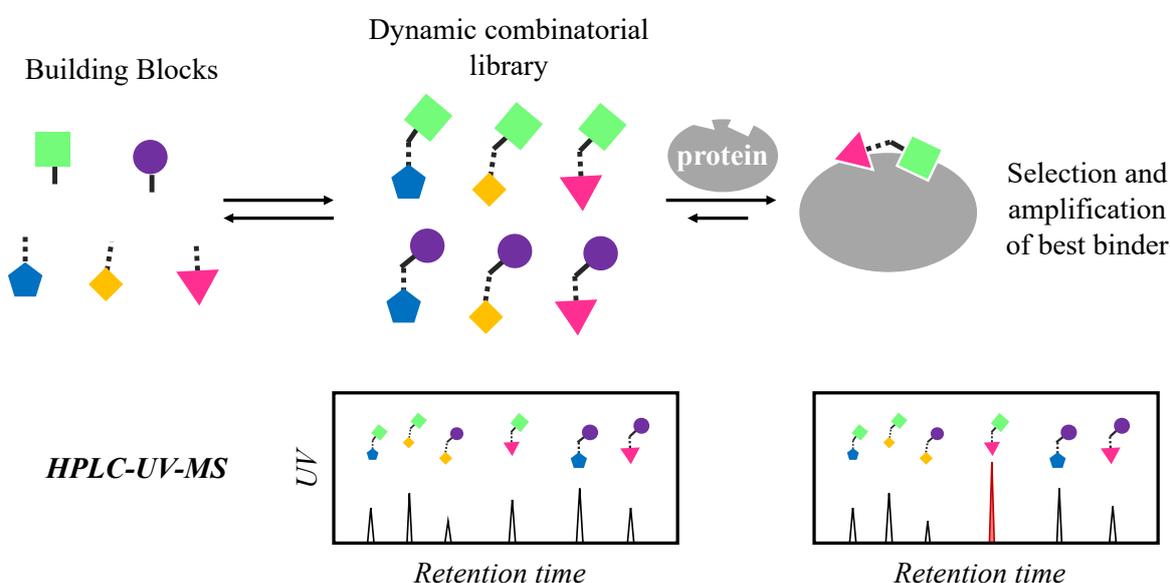
MCRs have been known for over a century, and although it may be hard to pinpoint the first example of an MCR, the Hantzsch dihydropyridine synthesis was reported in 1882,<sup>130</sup> followed by the Biginelli 3CR in 1893.<sup>131</sup> The first isocyanide-based MCRs were described by Passerini (3CR)<sup>132</sup> and Ugi (4CR)<sup>133</sup> in 1921 and 1959. Subsequently, several variations of the Passerini and Ugi reactions based on the unique reactivity of isonitriles have been described;<sup>129</sup> the interest in MCRs grew in the early 1990s contextually with the development of combinatorial chemistry and the establishment of academic HTS.<sup>134,135</sup> MCRs fill a unique niche in the synthesis of libraries and the complex follow-up associated with drug development or optimization of selective pharmacological tools, with several advantages over conventional synthesis.<sup>128</sup> First of all, MCRs can afford a high number of compounds with small synthetic effort; for example, a 3CR could provide 1000 products when 10 variants of each component are combined. Secondly, MCRs can provide an intrinsic SAR information within a single library by providing sets of compounds with related core structures. Finally, MCR library hits could represent a starting point for follow-up allowing for rapid preparation of ‘focused’ libraries. On the other hand, the use of MCRs carries the potential issue of having over-represented core-structures within a library; the diversity of a MCR library can be limited by the structure of the components. Nonetheless, this liability can be addressed by applying new variants of traditional MCRs that result in fundamentally different structures, as well as by using MCRs as a starting point for subsequent reactions, thus achieving efficiency and diversity.<sup>128</sup>

#### 3.4.4 Dynamic Combinatorial Chemistry

Dynamic combinatorial chemistry (DCC) is a powerful tool for the identification of bioactive compounds. Briefly, through the reversible combination of building blocks, DCC allows to

obtain dynamic combinatorial libraries (DCLs) composed by interchanging products.<sup>136–139</sup> Since the reaction between building blocks is reversible, the product distribution is determined by the thermodynamic stability of the compounds formed; a DCL is thus responsive to external *stimuli* such as the addition of a target. DCL members binding the target with the strongest affinity are subtracted from the equilibrium, causing a re-equilibration of the library (see Figure 15).<sup>137</sup> Ultimately, this leads to amplification of the best binders, which can be directly identified from the mixture, avoiding the synthesis, purification and characterization of each individual library member.

From an operational point of view, several important factors should be considered, including biocompatibility, pH dependence, temperature, solubility and stability of the components.<sup>140</sup> This is especially important to avoid artifacts, deriving, for example, from events such as precipitation of the protein or DCL members leading to false negatives or positives due to undesired shifts in the equilibrium. Additionally, to obtain meaningful results from DCC experiments, the protein sample should represent its native state.<sup>141–143</sup> For these reasons, experimental conditions, including type of buffer, pH, ionic strength, protein concentration and DMSO content should be determined prior to performing DCC experiments.



*Figure 15.* The principle of protein-templated DCC. A member of the DCL that interacts with the protein, will be amplified compared to the other members. The amplification can be monitored by an appropriate biophysical technique such as HPLC-UV. The product formed can be identified univocally with the coupling of the HPLC-UV system with a MS detector.

The effect of the buffer components on a protein can be measured in a straightforward way, by determining the melting point of the target protein, via a thermal-shift assay/differential scanning fluorimetry (TSA/DSF), and following its behavior over a range of time compatible with the experiment.<sup>144</sup> Other methods, like differential scanning calorimetry (DSC), isothermal

titration calorimetry (ITC) and circular dichroism spectroscopy (CD) could also be used to gain information on the buffer-related interactions and possible stabilization of the protein, but they require a high amount of protein and/or long measurement time; SDS-PAGE (Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis) can only be used to check for cleavage of the protein backbone. For enzymatic targets, a functional assay can be used instead of TSA and PAGE measurements for the assessment of long-term protein stability; given its univocal evaluation of the enzyme catalytic activity, if available, this should be the method of choice.<sup>140</sup> To date, only a limited number of reversible reactions have been used.<sup>140</sup> One of the most frequently used is the formation of acylhydrazones combining carbonyl (ketone or aldehyde) building blocks with acylhydrazides, as it can take place in water, and is thus biocompatible;<sup>145</sup> additionally, the preparation of the related building blocks is generally straightforward and many are commercially available. Acylhydrazone formation and exchange also depend on the buffer, more specifically on the pH, and at physiological conditions, neutral pH and room temperature, they are relatively slow, while at acidic pH, the equilibrium is reached quickly. This does not pose a conflict with the conditions required by the protein, as, if necessary, the pH-dependence can be influenced by adding a nucleophilic catalyst (*i.e.*, aniline and derivatives).<sup>146,147</sup>

To allow a comparative analysis of DCLs, the protein-templated DCC and a blank reaction (building blocks of the DCL without the protein), should be run in parallel. The reaction mixtures can then be “frozen”, to avoid possible shifts in the library equilibrium at the time of the analysis; focusing on acylhydrazone chemistry, this can be achieved by increasing the pH. Denaturation of the protein (by heat, addition of a solvent or ultracentrifugation) releases all binders protein before analysis.<sup>140</sup> Different techniques could be used to evaluate the samples to identify hits: a commonly applied method to analyze DCC experiments is HPLC-UV-MS. A proper chromatographic method can in fact separate the formed products from each other (as well as possible residual unreacted building blocks), while a MS-detector can help to univocally identify the formed products based on their exact mass. Finally, integration of UV peaks allows to accurately compare peak areas for the same products in blank and protein-templated reaction to detect amplification and identify hit compounds (see figure 17).

Protein-templated DCC thus offers an efficient and powerful approach for the acceleration of the identification and optimization of novel ligands for biological targets, and therefore holds a great potential for drug discovery combining synthesis of ligands and screening for affinity for the drug target in a single operation where the target selects its own ligands.<sup>139,148–151</sup>

### 3.4.5 *Microwave-assisted procedures*

To accelerate reactions, in the conventional one-molecule-at-a-time synthesis, as well as in combinatorial chemistry and multi-component reactions, heating is undoubtedly the most common practice. Yet, since some organic compounds are unstable at high temperatures, heat can lead to a number of issues; excessive heating might cause the formation of side products and need for further purification along with affected yields. In this context, Microwave-Assisted Organic Synthesis (MAOS),<sup>152</sup> together with other applications of microwaves (*e.g.* Microwave-Assisted Extraction, MAE),<sup>153</sup> has gained growing interest in the process of Drug Discovery, for both lead generation and optimization, due to its many interesting features, including faster chemistry (ability to reduce reaction times from days or hours to minutes), greater versatility and formation of cleaner products, compared to conventional synthetic methods, eventually potentially speeding the whole drug discovery process. In addition, the implementation of techniques combining microwave-assisted strategies with other high-performing approaches such as combinatorial chemistry, together with the optimization of associated instrumentation and procedures, has opened the way to the use of microwaves in even more effective applications. In the field of Drug Discovery, several microwave-assisted techniques are currently used with manifold applications not limited to organic synthesis related areas. Microwave-enhanced technologies are also used in target discovery, pharmacogenosy, screening, pharmacokinetics and even in the clinic.<sup>152–156</sup> Microwave heating differs from conventional heating in several ways. First of all, the energy is only transferred directly to those components that are susceptible to microwave polarization. This improves energy efficiency, decreasing the need to heat vessels or the heating apparatus itself; it allows the energy to be directed into specific parts of the reactive system such as metal particles, or susceptible solid supports, reducing in some cases the need for solvents, with obvious economic and environmental benefits. Secondly, through direct heating it is possible to raise the sample temperature much faster compared to conventional methods, frequently leading to quite different kinetics.<sup>157,158</sup>

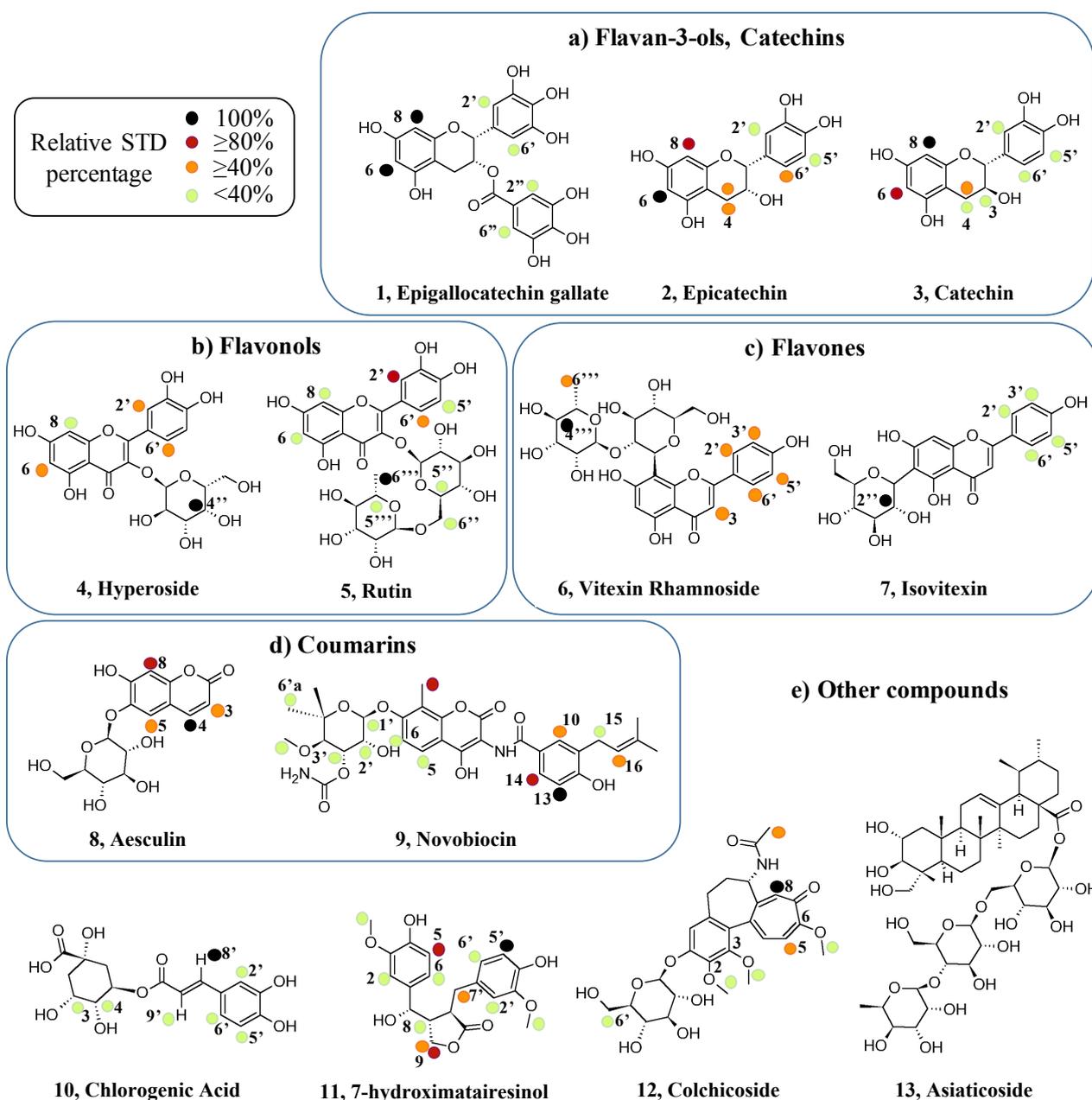
Within this thesis, the exploitation of microwave heating in the process of discovering new HuR ligands was considered from the beginning for both MAOS and MAE applications. Nonetheless, given the preliminary stage of the topic, though microwaves were indeed exploited for some synthesis and extraction procedures, a thorough and systematic study of the possible advantages of microwave-enhanced procedures was finally not carried out within the main topic of the thesis. The research done in preparation was exploited to write the book chapter “The use of microwaves in drug discovery” (see appendix IV, book chapter). Two MAE

applications, performed during a research period in the context of drug discovery from plants, are described in the two articles “(R)-(-)-Aloesaponol III Methyl Ether from *Eremurus persicus*: A Novel Compound against Leishmaniosis” (see appendix IV, article 3) and “The use of a microwave assisted solvent extraction coupled with HPLC-UV/PAD to assess the quality of *Marrubium Vulgare L.* (white horehound) herbal raw material” (see appendix IV, article 4).

Nevertheless, it is quite interesting to consider the possibility to apply microwaves to some of the reactions performed in this thesis, especially considering that microwave-assisted combinatorial synthesis (MICROCOS) and microwave-assisted high speed parallel synthesis can often find successful applications in the field of multi-component reactions (described in paragraph 3.4.3).<sup>152</sup> In fact, by combining multiple starting materials with different moieties, these can lead to a high efficient and “greener” formation of multifunctional/complex target compounds;<sup>159</sup> thus coupling MCRs with microwave techniques further increases the advantages of the multi-component method, creating molecular diversity with quicker one-step procedures that oppose the commonly costly and time-consuming traditional drug discovery process.<sup>152,159</sup> The most interesting applications found in literature, so far involve Hantzsch, Biginelli and Ugi reactions,<sup>160–162</sup> but the continuous evolution of research can grant us a growing number of applications.

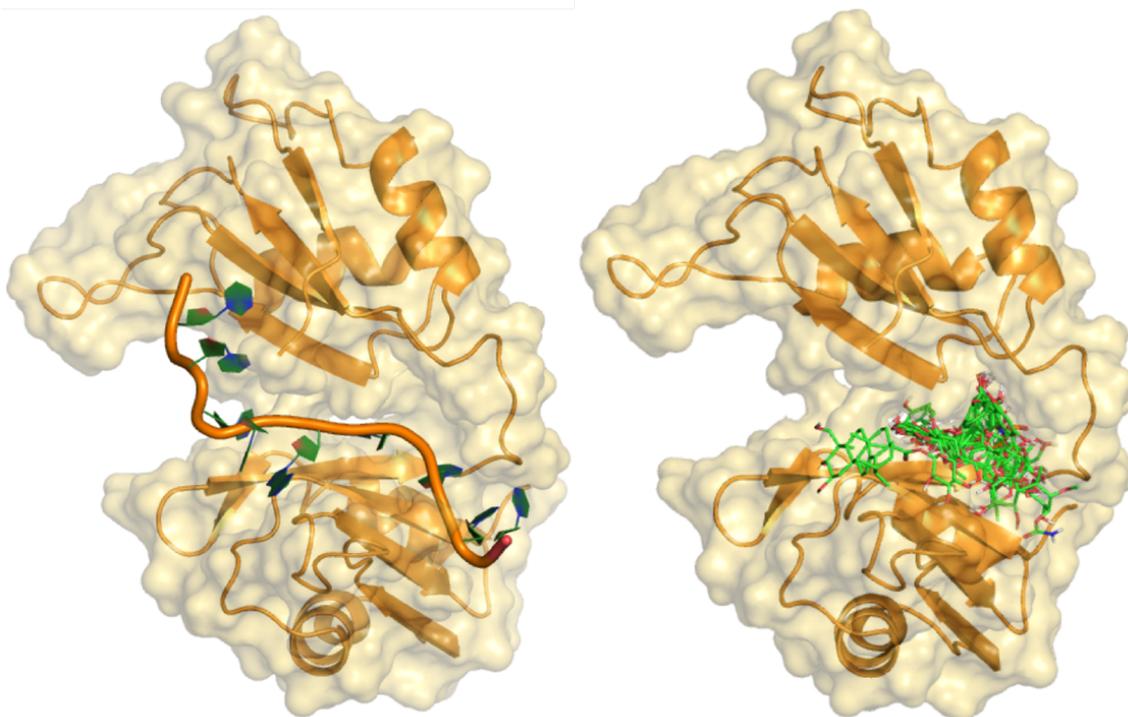
#### 4. *Exploration of HuR–ligand binding modes*

The present chapter represents the application of the first systematic medicinal chemistry approach towards a rational identification of compounds able to interfere with HuR–RNA complexes. In particular, we set out to study the interaction between HuR and a small series of compounds, all natural products including known HuR/ELAV–RNA interferers and derivatives, namely flavonoids, coumarins and structurally-related (see figure 16); the compounds subjected to the study were selected according to their structural features, solubility and stability in water-based buffers, with the aim to explore the chemical space around the structures of literature-reported ELAV–RNA interferers (see figure 5, paragraph 1.3). To achieve our goal, we applied ligand-based NMR technique STD NMR, which can be employed to investigate weak and transient protein–ligand complexes that are difficult to study by other structural methods (see paragraph 3.1.1). The epitope maps identified for each ligand, were compared with molecular dynamics and docking simulations, thus affording an improved understanding of ligand–HuR interactions. In particular, molecular dynamics simulations were necessary to reproduce the protein high mobility and flexibility, allowing us to study the possible conformational rearrangements occurring upon binding to HuR, according to results previously reported by Wang *et al.*, 2013<sup>75</sup> (see paragraph 1.4.1). All compounds which experimentally showed direct binding to HuR (**1–12**) were predicted to bind in the same pocket-like region within the RRM1/2 region of HuR, in its closed conformation, identifying repeated residues in the interaction site and showing, in some cases, similar or overlapping binding modes for some moieties, when superimposed. The same was not found for compound **13**; in fact, in accordance with the experimental data, *in silico* simulations could only place the structure outside of the cavity, where only superficial contacts are available (see figure 17). Similar results are found for open or closed conformation of HuR, though a slightly higher number of contacts could be established between **13** and open HuR; this could rationalize the absence of direct binding observed with the target protein, as it is hypothesized that active compounds should be able to stabilize a closed conformation of the protein.<sup>163</sup> Additionally, the results for all interacting compounds were analyzed within structural class (flavan-3-ols, flavonols, flavones, coumarins and structurally-unrelated, see figure 16), allowing us to elucidate how different substitutions on the same or a similar flavonoid-like core produce not only differences in affinities but differences in binding modes.



*Figure 16.* Structures, common names and epitope maps of all compounds submitted to the study with HuR. Relative STD percentages are conveyed by color code: black dots indicate the most intense signal (100% relative STD), dark red dots over 80%, orange dots over 40%, and lime green dots under 40% relative to the most intense STD signal. No epitope map is reported for Asiaticoside (**13**), for which we could not observe any direct binding with HuR. (Adapted from Vasile *et al.*, 2018, see appendix III, article 1).

Thus, the combination of NMR and *in silico* studies of the interaction process proved to be a reliable tool for understanding HuR–ligand binding modes and their exploitation laid the foundation for the design of *ad hoc* molecules acting as HuR ligands, potentially able to interfere with HuR–RNA complexes.



*Figure 17.* left side: 3D representation of HuR RRM1 and RRM2 domain-RNA<sup>c-fos</sup> complex in its “closed” conformation; right side: superimposition of all studied compounds in the “closed” conformation of the HuR–RNA interaction site. HuR RRM1/2 and RNA<sup>c-fos</sup> are represented as light-orange and bright-orange cartoon, respectively; all compounds are depicted as green sticks. All interacting compounds are predicted to bind within the same pocket-like region, while the structure of compound **13** is clearly laying on the outer side of the protein, establishing only superficial contacts. (Adapted from Vasile et al., 2018, see appendix III, article 1).

The full work is reported in the form of a published article: “Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD NMR and Molecular Modelling approach”, *Scientific Reports*, 2018 (see appendix III, article 1).

A recent follow up to this research, reported in paragraph 4.1, consists in the setup of newly reported DEEP-STD NMR methodology (see paragraph 3.1.2) on selected natural products (**5**, **8** and **9**, rutin, aesculin and novobiocin). While traditional STD NMR approaches can gain only partial information, this new methodology<sup>91</sup> can provide a more detailed picture of the small molecule–HuR interaction thanks to the development of differential epitope maps which add information on the nature of the amino acids involved in the binding event.

#### 4.1 *Set up and application of differential epitope mapping (DEEP) by STD NMR on selected natural products: 5, 8 and 9 (rutin, aesculin and novobiocin)*

The still current lack of experimental data concerning a small molecule binding site on HuR finds its causes in the absence of both small molecule–HuR co-crystals and the issues in applying protein-observed NMR encountered so far by other groups *i.e.* low solubility of small molecule inhibitors.<sup>163,164</sup> Though it can be argued that traditional STD NMR approaches can grant only partial information, new STD-based methodology DEEP-STD NMR<sup>91</sup> can help to fill some gaps by generating differential epitope maps which can aid the validation of predicted binding poses by adding information on the nature of the amino acids involved in the binding event (additional details on the technique are described in paragraph 3.1.2). This rather new methodology, is to be included in the systematic approach exploited for the study of small molecule–HuR interactions, especially for the study of novel synthetic compounds with still completely unreported profiles. The setup of the methodology was performed on the native protein, exploited for the studies so far described in the present thesis, and three selected natural products of the previous paragraph, **5**, **8** and **9**. These three compounds were chosen based on their structures including quite diverse moieties and substituents (aromatic, aliphatic, polar) with the idea that such widespread substitution could afford interactions with different types of amino acids, detected by DEEP-STD. Given the high aromaticity of the compounds studied, and the presence of additives in the native protein samples (necessary to insure protein stability and integrity), a differential frequency approach was avoided in order to reduce the possibility of artifacts (coming in this case from aspecific irradiation); in fact, only a differential solvent approach could be so far applied in this system. Nonetheless, given the recently introduced possibility to perform experiments on different protein constructs (see Chapter 7), characterized by different stability and solubility, more complete studies might be performed in the future pending optimization and readjustment of the experimental conditions.

All experiments reported in this section were thus performed applying an irradiation frequency of  $-30$  Hz ( $-0.05$  ppm); differential solvent conditions were obtained by measuring the samples in 100% D<sub>2</sub>O and H<sub>2</sub>O:D<sub>2</sub>O 90:10. The next paragraphs describe and discuss the results divided by compound.

**Rutin (5)** DEEP-STD data were analyzed according to the equation shown in paragraph 3.1.2: since in D<sub>2</sub>O the global saturation was larger than in H<sub>2</sub>O, we set it as *expl*; thus, positive  $\Delta$ STD factors would identify ligand protons with a relative increase in STD intensities in D<sub>2</sub>O, suggesting their proximity to apolar residues, while, negative  $\Delta$ STD factors, ligand protons with

a relative decrease in STD intensities in D<sub>2</sub>O to indicate that they are adjacent to slow exchanging protons and, consequently, polar residues. By calculating the ratio and the average ratio of STD intensities over all protons as well as their DEEP-STD factor (data reported in table 4) we observed several  $\Delta$ STD factors, confirming the ligand binding epitope map changes in the two different solvents (see figure 18 for graphic representation). Positive  $\Delta$ STD factors were found for proton 2', 6' and 6, while negative  $\Delta$ STD factors for proton 1''' and 6'''. Different protons of **5** thus seem to occupy distinct areas of HuR binding pocket lined by either apolar or polar protons (differential epitope map shown in figure 18). The DEEP-STD results are in agreement with the docking pose of the complex between rutin and HuR (see figure 18), where the ligand sits between apolar and polar patches in the protein binding pocket. In particular, the ligand protons 2', 6' and 6 point towards apolar residues (Ile23 and Ile133), while 1''' and 6''' are positioned near polar residues (Arg97, Ser99 and Lys104). To interpret the outcomes from the DEEP-STD NMR experiment, only the strongest  $\Delta$ STD were considered; we experimentally determined that  $\Delta$ STD greater than 0.4 in magnitude are significant (the threshold always depends on the STD values of the protein–ligand system under study and can be determined from the standard deviation).

Proton	STD% D <sub>2</sub> O	STD% H <sub>2</sub> O	Ratio STD D <sub>2</sub> O/H <sub>2</sub> O	$\Delta$ STD
2'	0.78	0.21	3.71	1.36
6'	0.77	0.24	3.21	0.86
5'	0.48	0.23	2.09	-0.26
8	0.60	0.25	2.4	0.05
6	1.03	0.30	3.43	1.08
1'''	0.27	0.21	1.29	-1.06
6''	0.25	-	-	-
6'''	0.68	2.12	0.32	-2.03
	<b>Sum</b>	<b>Sum</b>	<b>STD average</b>	
	4.61	3.56	2.35	

Table 4. Raw and processed data of the differential solvent (D<sub>2</sub>O/H<sub>2</sub>O) DEEP-STD NMR experiment of the **5**–HuR complex.

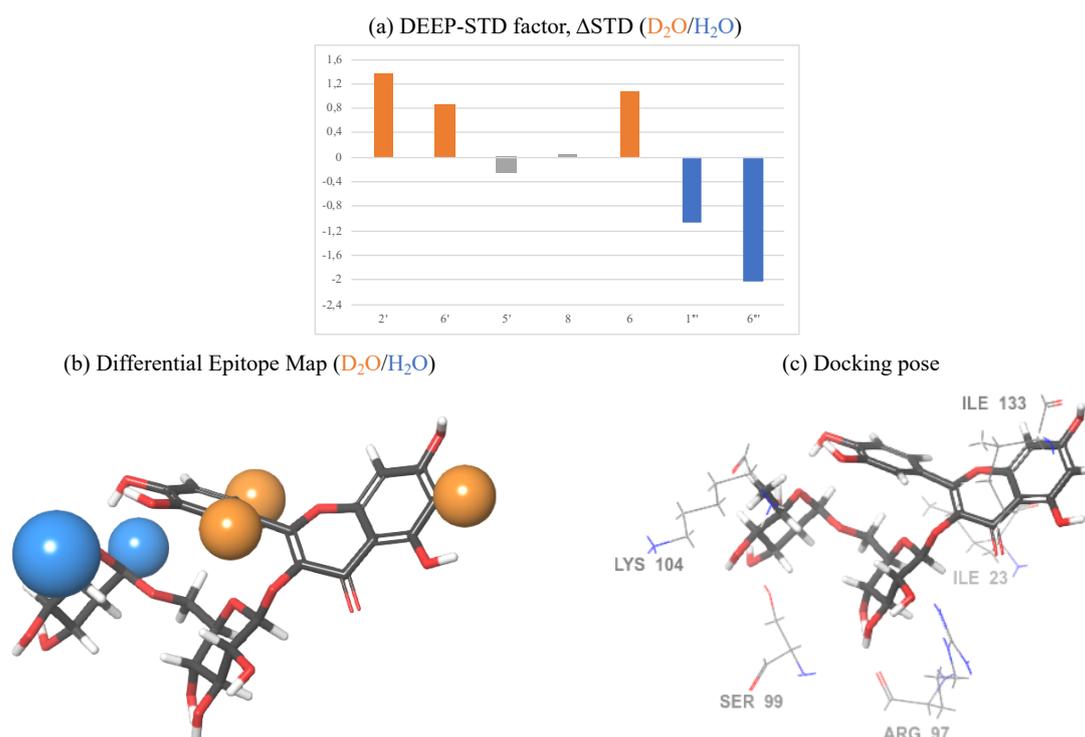
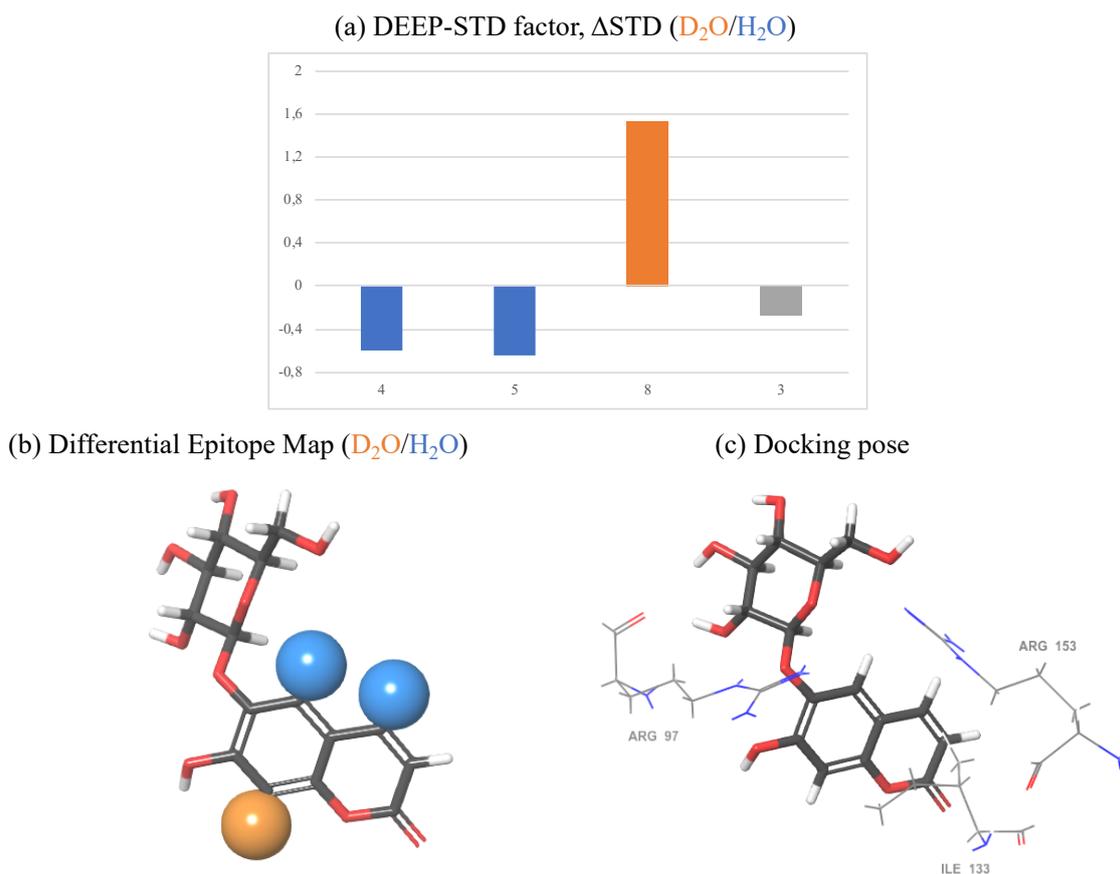


Figure 18. Graphic representation of the DEEP experiment of the 5–HuR complex. (a)  $\Delta\text{STD}$  histogram: significant negative  $\Delta\text{STD}$  factor ( $< -0.4$ ) are shown in blue, while significant positive  $\Delta\text{STD}$  factors ( $> 0.4$ ) in orange; (b) DEEP mapping of the ligand. Blue surfaces indicate ligand contacts with protein side chains carrying slowly exchanging protons; orange surfaces indicate ligand contacts with apolar residues. (c) Docking pose of the complex (only residues relevant to the DEEP mapping are shown).

**Aesculin (8)** For compound **8**,  $\Delta\text{STD}$ s calculated as described for **5** (data and graphic representation shown in table 5 and figure 19). Protons 4 and 5 show negative  $\Delta\text{STD}$  values, indicating that they are near to polar residues, while proton 8 display positive  $\Delta\text{STD}$  value, denoting that it looks towards apolar residues. This is in agreement with the docking pose of the complex, in which protons 4 and 5 sit between Arg97 and Arg153, and proton 8 is in proximity of Ile133 (see figure 19).

Proton	STD% $\text{D}_2\text{O}$	STD% $\text{H}_2\text{O}$	Ratio STD $\text{D}_2\text{O} / \text{H}_2\text{O}$	$\Delta\text{STD}$
4	0.50	0.30	1.67	-0.60
5	0.26	0.16	1.62	-0.65
8	0.99	0.26	3.80	1.53
3	0.68	0.34	2.00	-0.27
	<b>Sum</b>	<b>Sum</b>	<b>STD average</b>	
	2.43	1.06	2.27	

Table 5. Raw and processed data of the differential solvent ( $\text{D}_2\text{O}/\text{H}_2\text{O}$ ) DEEP-STD NMR experiment of the **8**–HuR complex.



*Figure 19.* Graphic representation of the DEEP experiment of the **8**-HuR complex. (a)  $\Delta$ STD histogram: significant negative  $\Delta$ STD factor ( $< -0.4$ ) are shown in blue, while significant positive  $\Delta$ STD factors ( $> 0.4$ ) in orange; (b) DEEP mapping of the ligand. Blue surfaces indicate ligand contacts with protein side chains carrying slowly exchanging protons; orange surfaces indicate ligand contacts with apolar residues. (c) Docking pose of the complex (only residues relevant to the DEEP mapping are shown).

**Novobiocin (9)** The experiment performed for compound **9** produced negligible  $\Delta$ STD factors for several protons indicating that the change of the solvent did not significantly affect them. Conversely, the methyl group in position 6 of the noviose moiety shows a great interaction with exchangeable protons, probably due to the presence of Lys55, Ser99, Ser100 and Lys104 as evidenced from docking studies (see figure 20). On the other hand, the methyl groups in position 17 and 18 show interactions with hydrophobic residues that, from docking studies, could be Ile78, Val 93, Ile132, Ile133.

Proton	STD% D <sub>2</sub> O	STD% H <sub>2</sub> O	Ratio STD D <sub>2</sub> O / H <sub>2</sub> O	ΔSTD
6	1.10	-	-	-
14	1.29	-	-	-
1'	0.81	0.64	1.27	-0.08
4'	1.60	-	-	-
2'	1.03	-	-	-
OCH <sub>3</sub>	0.66	0.44	1.5	0.15
15	0.70	0.45	1.56	0.21
9	1.13	0.69	1.64	0.29
17 - 18	0.96	0.52	1.85	0.5
6'a	1.31	1.02	1.28	-0.07
6'b	0.90	2.53	0.36	-0.99
	<b>Sum</b>	<b>Sum</b>	<b>STD average</b>	
	6.47	6.29	1.35	

Table 6. Raw and processed data of the differential solvent (D<sub>2</sub>O/H<sub>2</sub>O) DEEP-STD NMR experiment of the **9**-HuR complex.

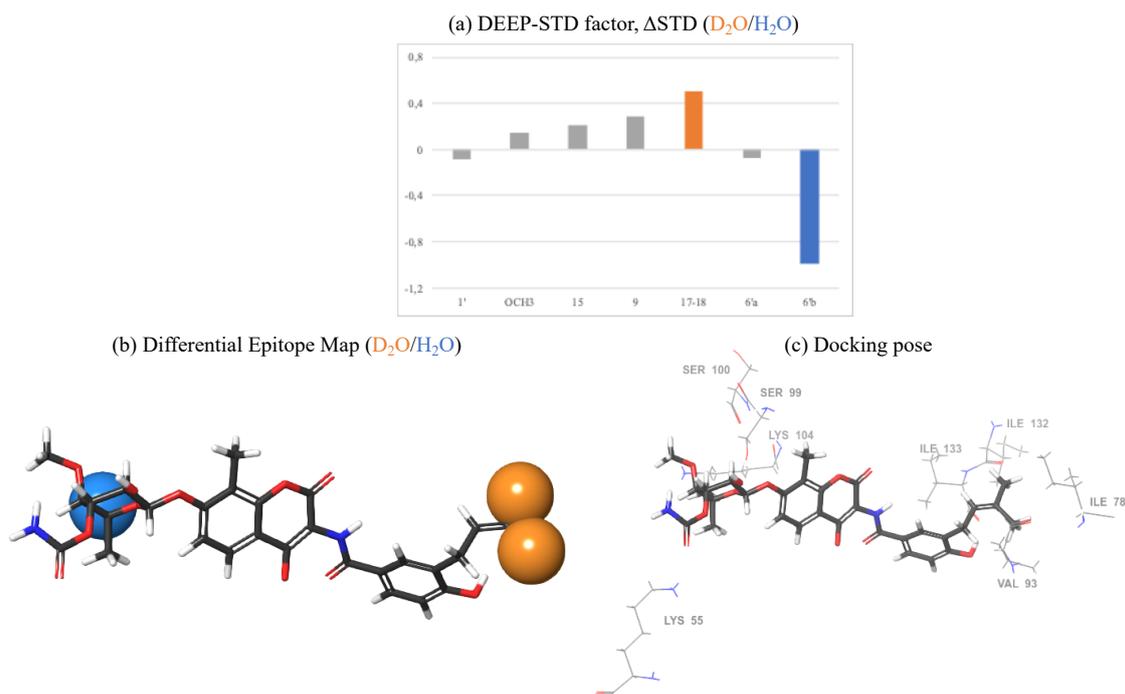


Figure 20. Graphic representation of the DEEP experiment of the **9**-HuR complex. (a) ΔSTD histogram: significant negative ΔSTD factor ( $< -0.4$ ) are shown in blue, while significant positive ΔSTD factors ( $> 0.4$ ) in orange; (b) DEEP mapping of the ligand. Blue surfaces indicate ligand contacts with protein side chains carrying slowly exchanging protons; orange surfaces indicate ligand contacts with apolar residues. (c) Docking pose of the complex (only residues relevant to the DEEP mapping are shown).

Based on the data so far available to us, following initial setup, the DEEP-STD experiment with differential solvent conditions seems to be a valid method to study the HuR–small molecule system coupled with *in silico* methodologies (molecular dynamics and docking). The additional information gathered on the interaction between certain protons and particular amino acid residues can help us understand the position of these small compounds in the binding site. Given the rather favorable solubility and stability in aqueous buffers of the three compounds chosen, these will be considered for setting up a protein-observed NMR methodology with the support of a newly launched partnership.<sup>‡‡</sup> The confirmation of the predicted poses and the building of a complete map of the HuR–small molecule interaction for the natural product series would produce the first experimentally validated binding site model for small molecules.

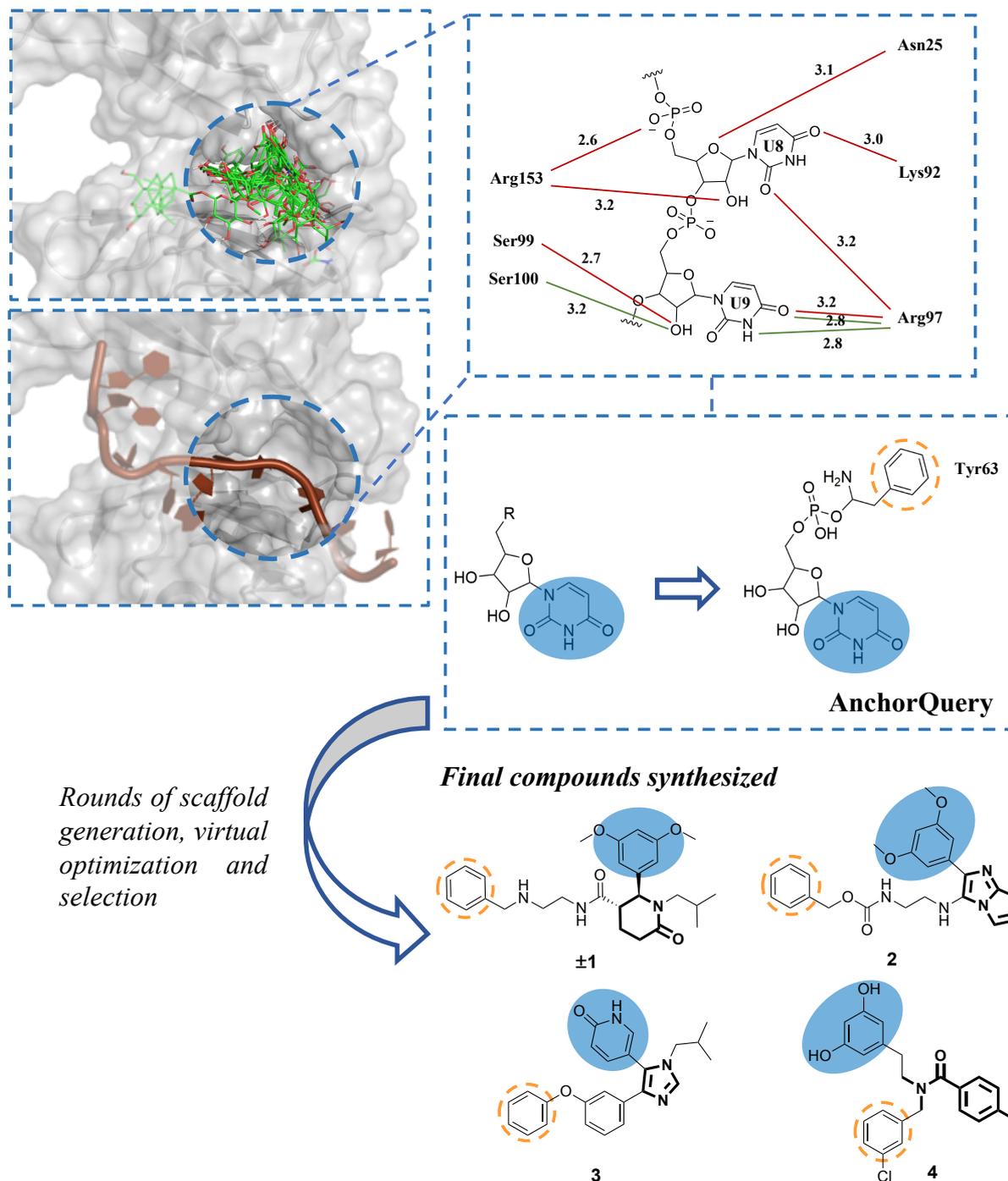
---

<sup>‡‡</sup> Prof. Dr. Michael Sattler Group, Institute of Structural Biology, Helmholtz Zentrum Munich - Germany

## 5. Structure-based approach to obtain new HuR ligands.

This chapter reports on the structure-based design and the synthesis of new HuR ligands based on four different scaffolds corresponding to four different MCRs. These compounds represent the first HuR MedChem campaign of our research group. The investigation of their interactions with HuR was performed using the combination of STD NMR and *in silico* studies set up and presented in the previous chapter.

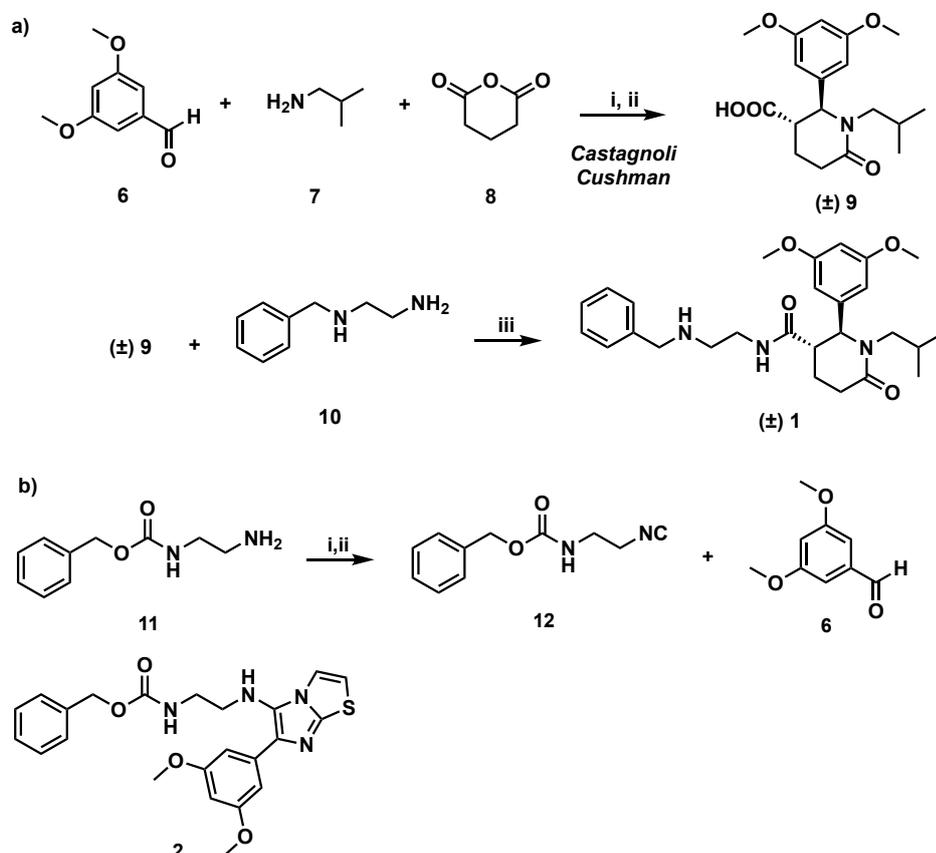
Briefly, based on the results of chapter 4, for the design of new compounds, we focused on a pocket-like region hosting the small HuR ligands, formed by the two asymmetric units of the protein, belonging to RNP1 and RNP2 of RRM1. Inspection of the cocrystal structure of the HuR RRM1/2 domains in complex with the ARE sequence of RNA<sup>c-fos</sup> (PDB 4ED5) confirms that this region corresponds to binding site of RNA uridine residues 8 and 9 (U8–U9) (see figure 21). We started our search for new scaffolds for HuR ligands from the key interactions shown in figure 21 and utilized the free web-based virtual screening platform AnchorQuery;<sup>165,166</sup> we selected U8 and U9 as anchors and investigated whether one of the two nucleotides may be a pharmacophore in terms of occupied position and interactions. The anchors were modified and grown to bear an additional phenyl group to afford a  $\pi$ – $\pi$ -stacking interaction with Tyr63 (see figure 21). Among several structures generated, the candidates for synthesis were selected based on visual inspection and molecular recognition studies, according to the following criteria: (a) the anchor should show good overlap with the corresponding uracil ring, preserving the same hydrogen bonds with Arg97 and Lys92, while (b) the additional aromatic rings should gain the desired interaction with Tyr63. We selected the most promising scaffolds and replaced the uracil moiety with an easy-to-handle starting material, thus designing compounds **1–4** (see figure 21), characterized by different scaffolds. The structures designed with AnchorQuery can be prepared through MCR or equally efficient processes. Thus, compound **1** was obtained by a Castagnoli–Cushman reaction (CC-3CR)<sup>167</sup> followed by amidation; **2** by a Groebke–Blackburn–Bienaymé (GBB-2CR);<sup>168</sup> **3** was synthesized by a Van Leusen reaction (vL-3CR),<sup>169</sup> and **4** by a reductive amination<sup>170</sup> followed by acylation. Herein, the synthetic strategies adopted to obtain compounds **1–4** are briefly described.



**Figure 21.** Structure-based design strategy for novel ligands of target protein HuR exploiting a nucleotide extension of the platform AnchorQuery. The docked poses of the compounds from chapter 4 (placement guided by STD NMR results) identify repeated sites on the protein. A comparison with RRM1/2–RNA<sup>c-fos</sup> co-crystal (PDB 4ED5) highlights residues U8–U9 in the same positions. This site is chosen for the structure-based design strategy. The four compounds selected for synthesis and interaction studies are shown on the bottom. Anchors, original and modified, are shown highlighted in blue, additional rings for Tyr 63 interaction are encircled in dashed orange, and the different core structures obtained by MCR are highlighted in bold.

Amidation of the acid intermediate *trans* ( $\pm$ ) **9** with amine **10** afforded compound ( $\pm$ ) **1** as racemate. The key intermediate *trans*-6-oxopiperidine-3-carboxylic acid **9** was obtained in good yield and diastereoselectivity, applying a microwave-aided one-pot CC-3CR on benzaldehyde **6**, isobutylamine

7, and glutaric anhydride **8** (see scheme 1a).<sup>167</sup> A microwave-assisted GBB-2CR,<sup>168</sup> involving isocyanide **12**, aldehyde **6**, and amino-thiazole **13**, produced compound **2**. Isocyanide **12** was obtained starting from amine **11**, which was converted into the corresponding formamide and subsequently dehydrated (see scheme 1b).

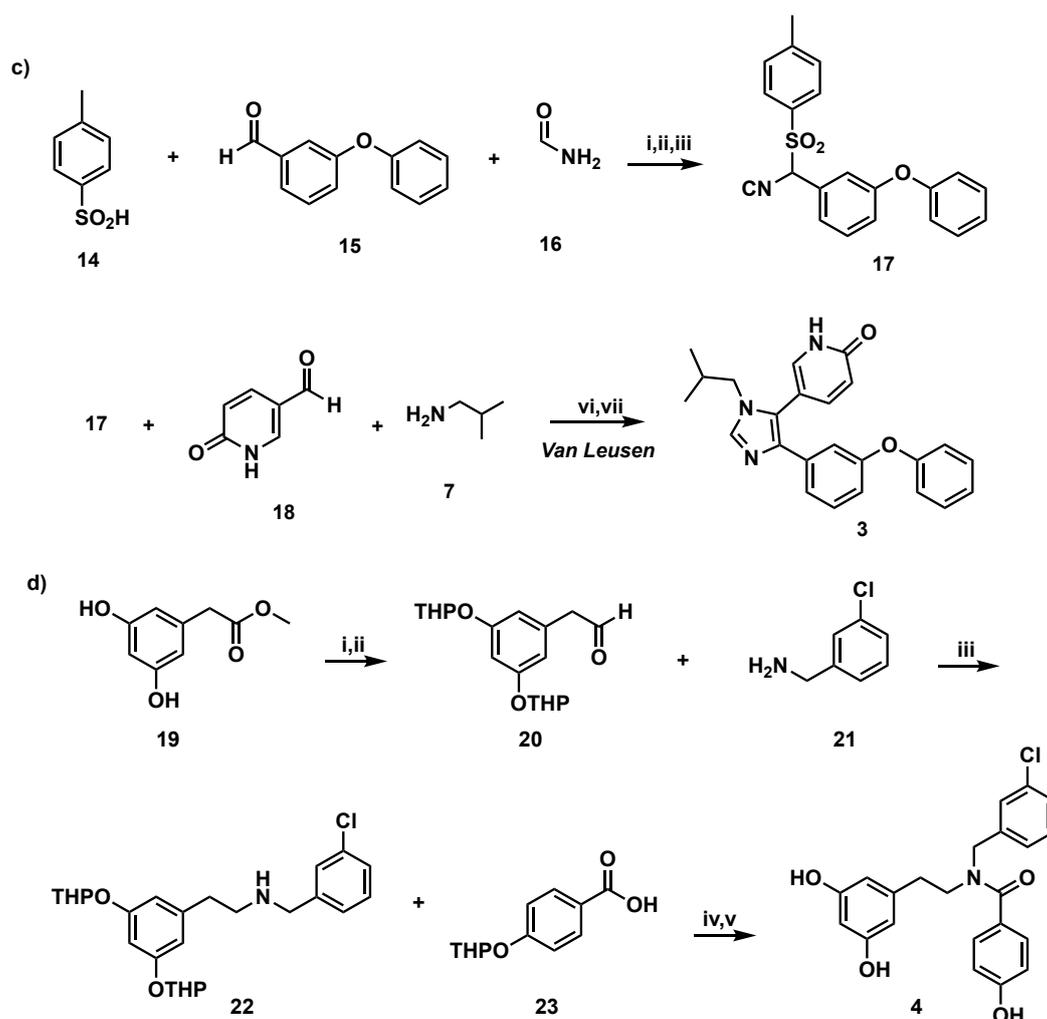


*Scheme 1.* Synthetic pathways of compounds **1** and **2**. Reagents and conditions:

a) Compound ( $\pm$ ) **1**: i) **6**, **7**, HCOOH, ACN, mw 120°C, 50 W, 30 min; ii) **8**, p-xylene, reflux, 10h, yield 62%; iii) TBTU, DIPEA, THF, rt, 12h, yield 71%.

b) Compound **2**: i) Ethyl formiate, 10h; ii) TEA, POCl<sub>3</sub>, DCM, rt, 4h, yield 62%; iii) **6**, **13**, HCOOH, ACN, mw 120°C, 30 min; iv) **12**, ZrCl<sub>4</sub>, 80 °C, 10h, yield 21% (adapted from Della Volpe *et al.*, 2019, see appendix III, article 2).

A vL-3CR27 on commercially available aldehyde **18**, isobutylamine **7**, and substituted TosMIC derivative **17** led to compound **3**. This MCR allows access to functionally rich imidazoles in a single pot via cycloaddition of TosMIC reagents on imines, generated in situ from an aldehyde and an amine under mildly basic conditions. We synthesized **17** starting from *p*-toluenesulfinic acid **14**, which afforded the intermediate formamide derivative upon reaction with commercially available 3-phenoxybenzaldehyde **15**, formamide **16**, and chlorotrimethylsilane; subsequent dehydration furnished isocyanide **17** (Scheme 2c). Compound **4** was obtained by reductive amination of aldehyde **20** followed by an amide coupling with acid **23** followed by removal of the THP protecting groups (Scheme 2d).<sup>170</sup>



*Scheme 2.* Synthetic pathways of compounds **3** and **4**. Reagents and conditions:

c) Compound **3**: i) H<sub>2</sub>O, 30 min; ii) HCl, MTBE, 30 min; iii) **15**, **16** Me<sub>3</sub>SiCl, ACN, Toluene, N<sub>2</sub>, 50°C, 5 h; iv) **14**, 16h, yield 62%; v) POCl<sub>3</sub>, triethylamine, THF, 10°C, 45 min, 85%; vi) **18**, **7** dry DMF, rt, 4h; vii) **17**, K<sub>2</sub>CO<sub>3</sub>, rt, 72h, yield 15%.

d) Compound **4**: i) PPTS, 3,4-DHP, dry DCM, rt, 1.5h, yield 94%; ii) DIBAL-H, dry DCM, N<sub>2</sub>, -78°C, 30 min, yield 49%; iii) NaBH(OAc)<sub>3</sub>, THF, N<sub>2</sub>, rt, 24h, yield 85%; iv) HOBt, EDC·HCl, DIPEA, dry DMF, N<sub>2</sub>, rt, 24h, yield 60%; v) PPTS, EtOH, 55°C, 16h, 75% (adapted from Della Volpe *et al.*, 2019, see appendix III, article 2).

Of four compounds synthesized, three (**1**, **3** and **4**) were subjected to the combined STD NMR and *in silico* investigation of their interactions with HuR (according to the format set up and presented in the previous chapter), thanks to their favorable solubility in water-based buffers. Finally, **1** and **4** were selected as promising compounds to develop in the next steps of our MedChem campaign. In particular, compound **4** shows interesting binding modes which appear so far to be in accordance with the initial design: while all rings interact with the target, the anchor, as hypothesized, appears to be the most interacting moiety.

On the contrary, compound **1**, though directly binding to the target, shows a different interaction mode which requires further investigation prior to the design of new derivatives. Several factors could

play into the differences of compound **1**: first of all, this was initially synthesized and tested as a racemic mixture of the two *trans* enantiomers,<sup>171</sup> so it is likely that the two enantiomers may have different interaction with the target (as it often happens with a chiral molecule within a chiral environment), thus they should be separated and investigated individually. Moreover, the modified anchor of compound **1**, though similar to **4**, has both hydroxyls substituted; though the additional methyl group is rather small, it changes the binding capacity of the anchor (*i.e.* the anchor now lacks hydrogen-bond donors) and could thus be responsible for the different binding mode. Finally, the compound could have higher affinity for a different protein site, which we cannot detect with STD NMR but need a protein observed NMR method, rather than X-ray diffraction.

The structure-based design, synthesis and preliminary results of the STD NMR and *in silico* interaction studies are described in the published article “Novel Compounds Targeting the RNA-Binding Protein HuR. Structure-Based Design, Synthesis, and Interaction Studies”. ACS Medicinal Chemistry Letters, 2019 (see appendix III, article 2).

Additional studies on the two hits **1** and **4** have thus been performed and some are still ongoing. Concerning compound **1**, a second part of the study is reported in paragraph 5.1, while a follow-up investigation to design derivatives of compound **4** is reported in chapter 6, paragraph 6.4.

### **5.1 Resolution of ( $\pm$ ) **1** enantiomers and differential interaction study**

Compound **1** was initially synthesized and studied as a racemic mixture of the two *trans* enantiomers but, given its interesting behavior, the obvious next step involved the resolution of the racemate in the two separate enantiomers by enantioselective HPLC. Both were studied by STD NMR and the differences in their binding modes elucidated by applying an enantio-application of DEEP-STD NMR. Additionally, differential solvent composition was used for DEEP-STD NMR experiments on each enantiomer to understand the importance of polar residues in their respective binding to HuR. Pending the assignment of their absolute configuration via crystallization and X-ray diffraction, the collected spectroscopic information will guide the synthesis of several derivatives with modified anchors (to understand the importance of anchor substitution), for improved affinity and RNA displacement ability.

#### **Chiral resolution via enantioselective HPLC**

To develop a separation method able to afford the two enantiomers of **1** on a semi-preparative scale, two immobilized polysaccharide-based chiral stationary phases were initially tested at the analytical scale, amylose and cellulose (REGISPACK IA and CHIRALPAK IC columns, respectively) with the screening protocol as reported in the experimental section (see experimental section). The best separations were obtained on Chiralpak IC by eluting with both *n*-Hexane-IPA 75:25 with

0.1 % DEA, and *n*-Hex-IPA 80:20 with 0.1 % DEA (details reported in the experimental section ). Looking at the chromatographic profiles it is quite evident that *n*-Hexane:IPA:DEA 80:20:0.1 (v/v/v) (see figure 22, top) affords the best resolution though with longer times, thus it is not suitable for scale-up; on the contrary, the elution with *n*-Hex:IPA:DEA 75:25:0.1 affords a good resolution in more appropriate timeframes (see figure 22, bottom).

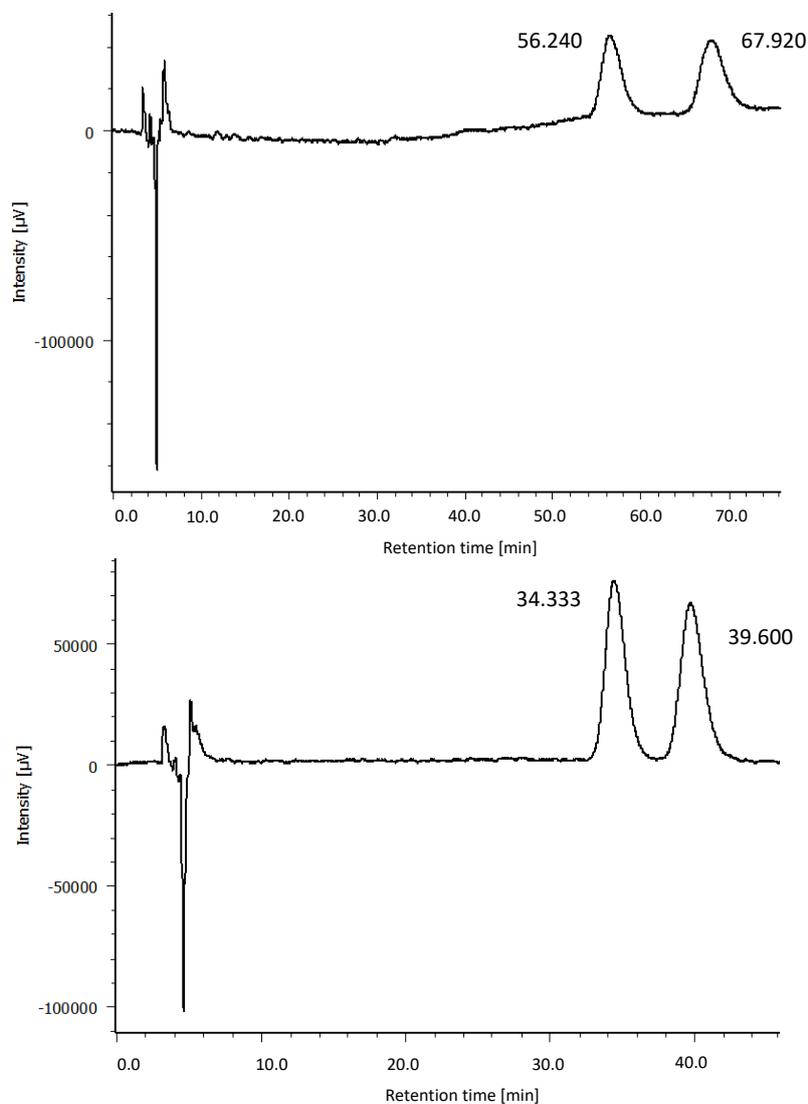


Figure 22. Top: Analysis performed using Chiralpak IC (4.6 mm I.D. x 250 mm), *n*-Hex:IPA:DEA 80:20:0.1 (v/v/v), flow: 1 mL/min, 25°C,  $\lambda = 220$  nm. Injection volume 10  $\mu$ L (1 mg/mL in IPA).  $R_{t1} = 56.240$  min;  $R_{t2} = 67.920$  min;  $\alpha = 1.231$ ;  $R_s = 2.530$ . Bottom: Analysis performed using Chiralpak IC (4.6 mm I.D. x 250 mm), *n*-Hex:IPA:DEA 75:25:0.1 (v/v/v), flow: 1 mL/min, 25°C,  $\lambda = 220$  nm. Injection volume 10  $\mu$ L (1 mg/mL in IPA).  $R_{t1} = 34.333$  min;  $R_{t2} = 39.600$  min;  $\alpha = 1.184$ ;  $R_s = 1.896$ .

To obtain each enantiomer in sufficient quantity for interaction studies, a semi-preparative column Chiralpak IC (1 cm I.D. x 25cm, s-5 $\mu$ m) was used. Four 1 mL samples of 10 mg/mL (*n*-Hex:IPA 75:25) were injected. Three fractions were collected in each run, and those with comparable enantiomeric excess (*ee*) of the same enantiomer were pooled and evaporated affording 25 mg of the first enantiomer eluted (*ee* = 99%) and 12.5 mg of the second enantiomer eluted (*ee* = 95.6%) (see figure 23). The specific rotation of each was determined in MeOH:  $[\alpha]_D^{20} = +74.0^\circ$  for the first enantiomer and  $[\alpha]_D^{20} = -74.6^\circ$  for the second enantiomer (see figure 23).

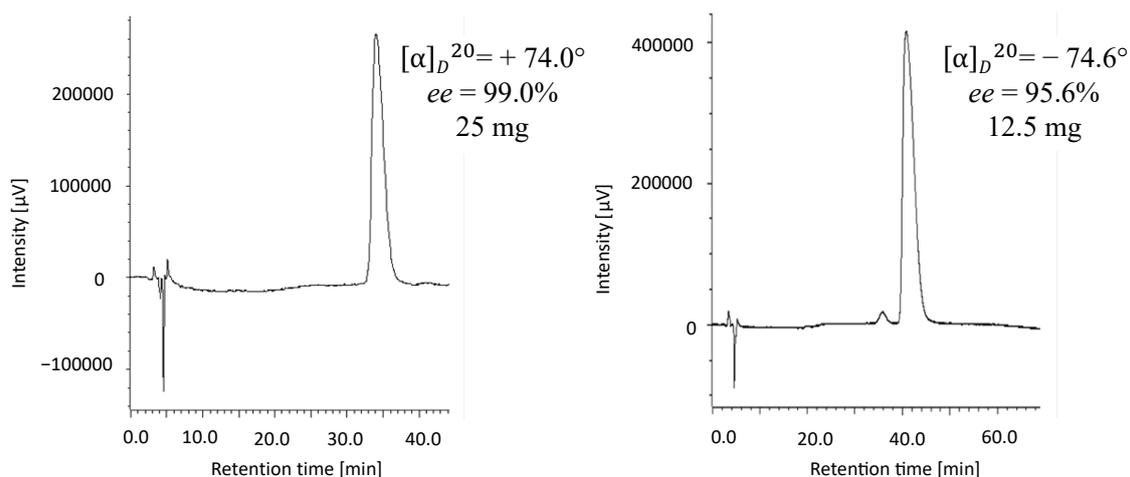


Figure 23. Chromatographic profile of the two enantiomers, first (*ee* 99%) and second elute (*ee* 95.66%) after separation. Analyses performed using Chiralpak IC (4.6 mm I.D. x 250 mm), *n*-Hex:IPA:DEA 80:20:0.1 (v/v/v), flow: 1 mL/min, 25°C,  $\lambda = 220$  nm. Injection volume 10  $\mu$ L (1 mg/mL in *n*-Hex:IPA 75:25).

### ***Interaction studies with HuR of (+) 1 and (-) 1 by enantio- and solvent- DEEP-STD***

The two enantiomers thus obtained were exploited for interaction studies with the target HuR according to the DEEP-STD protocol reported in the previous chapter for both comparison of the two enantiomers and to build a differential solvent epitope map. The results are hereby reported and discussed.

STD NMR spectra were recorded for both enantiomers of **1** in the same conditions as reported for all previous compounds (see experimental section). Absolute and relative STD % for each enantiomer are stated in tables 7 (for compound (+) **1**) and 8 (for compound (-) **1**), while a schematic representation of each enantiomer epitope mapping is shown in figure 24; the color code of the dots associated to each proton is once again conveying a range value of relative STD percentage (see figure 24).

atom	Absolute STD %	Relative STD %
Ring A (2, 3, 4)	0.62	100
Ring A (1, 5)	0.44	71
Ring B (3)	0.38	61
Ring B (1, 5)	0.34	55
16	0.15	24
OCH <sub>3</sub>	0.26	42
5	0.40	65
3	0.43	69
9/10	0.17	27

Table 7. Absolute and relative STD % of compound (+) **1**.

atom	Absolute STD %	Relative STD%
Ring A (2, 3, 4)	0.79	100
Ring A (1, 5)	0.32	41
Ring B (3)	0.21	27
Ring B (1, 5)	0.22	28
16	0.28	35
OCH <sub>3</sub>	0.15	19
5	0.56	71
3	0.34	43
9/10	0.27	34

Table 8. Absolute and relative STD % of compound (-) **1**.

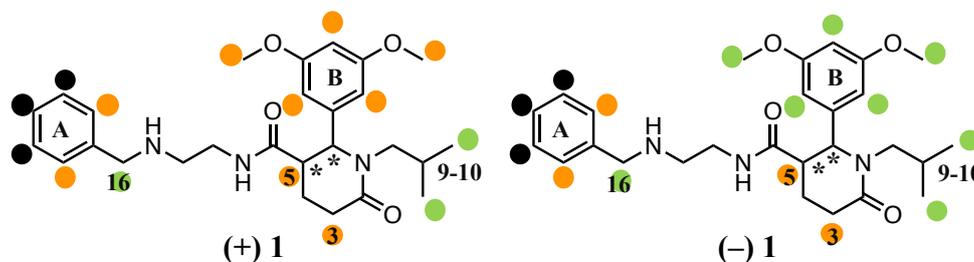


Figure 24. Representation of the group epitope map for (+) **1** and (-) **1** binding to HuR. Relative STD % ranges are conveyed by color code: black dots for 100%, orange dots for values between 40 and 80%, and green dots for values below 40%.

Nonetheless, as anticipated in paragraph 3.1, absolute STD values coming from different STD spectra, cannot be directly compared, due to possible differences in sample preparation or instrument setting. For this reason, in order to obtain a comparison between the interaction of the two separate enantiomers with HuR, we exploited the DEEP equation reported in paragraph 3.1.2. The experiment of (+) **1** was chosen as *expl* (since the global saturation is larger than that of (-) **1**), and the ratio and average ratio of STD intensities over all protons was calculated affording “enantio DEEP-STD”

values for each proton. The raw and processed data and  $\Delta$ STD factors are shown in table 9 and represented schematically in figure 25. We recorded positive  $\Delta$ STD factors for the protons of ring B and its methoxy group, while negative  $\Delta$ STD factors were observed for protons 5, 9/10 and 16. These differences suggest that the two enantiomers interact with the same protons but different orientation into the binding site. Positive  $\Delta$ STD factors suggest which moieties are most involved in the interaction of (+) **1** with HuR, while negative  $\Delta$ STD factors for (-) **1** and HuR.

Proton	STD%	STD%	Ratio STD	$\Delta$ STD
	(+) <b>1</b>	(-) <b>1</b>	(+) <b>1</b> / (-) <b>1</b>	
Ring A (2, 3, 4)	0.62	0.79	0.78	-0.37
Ring A (1, 5)	0.44	0.32	1.38	0.23
Ring B (3)	0.38	0.21	1.81	0.66
Ring B (1, 5)	0.34	0.22	1.55	0.4
16	0.15	0.28	0.54	-0.61
OCH <sub>3</sub>	0.26	0.15	1.73	0.58
5	0.40	0.56	0.71	-0.44
3	0.43	0.34	1.26	0.11
9/10	0.17	0.27	0.63	-0.52
<b>Sum</b>	<b>3.19</b>	<b>3.14</b>	<b>STD average</b>	
			1.15	

Table 10. Raw and processed data of the enantio DEEP-STD NMR experiment of the (+) **1**-HuR and (-) **1**-HuR complexes in D<sub>2</sub>O.

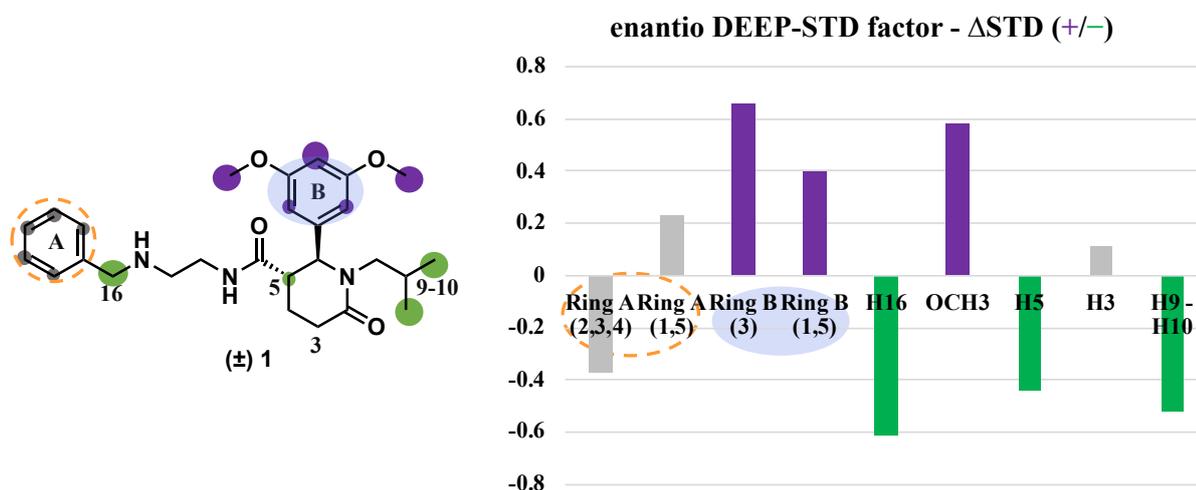


Figure 25. Enantio application of the DEEP-STD experiment. The STD spectra of (+) **1** and (-) **1** in D<sub>2</sub>O are compared thanks to the equation exploited in the differential approach. Stronger binding components are reported in purple for the first enantiomer eluted (+) **1** and green for the second (-) **1**. Grey dots indicate no significant difference in moiety contribution for the two enantiomers tested. The size of the colored circles reported on the structure is related to the  $\Delta$ STD modulus.

Since (+) **1** shows a stronger contribution of the anchor moiety (ring B) compared to (-) **1**, we decided to start differential solvent experiments on this enantiomer, following the procedure described in paragraph 4.1 for selected natural products. The experiment highlighted that that ring A points towards apolar residues, while protons 16 and 9/10 are projected towards polar residues. (All data are reported in table 10 and represented in figure 26).

Proton	STD% D <sub>2</sub> O	STD% H <sub>2</sub> O	Ratio STD D <sub>2</sub> O / H <sub>2</sub> O	ΔSTD
Ring A (2, 3, 4)	0.63	0.38	1.66	0.14
Ring A (1, 5)	0.44	0.12	3.67	2.15
Ring B (3)	0.38	0.29	1.31	-0.21
Ring B (1, 5)	0.34	0.21	1.62	0.1
16	0.15	0.31	0.48	-1.04
OCH <sub>3</sub>	0.27	0.20	1.35	-0.17
5	0.40	-	-	-
3	0.43	-	-	-
9/10	0.17	0.32	0.53	-0.99
	<b>Sum</b>	<b>Sum</b>	<b>STD average</b>	
	2.38	1.83	1.52	

Table 10. Raw and processed data of the differential solvent (D<sub>2</sub>O/H<sub>2</sub>O) DEEP-STD NMR experiment of the (+) **1**-HuR complex.

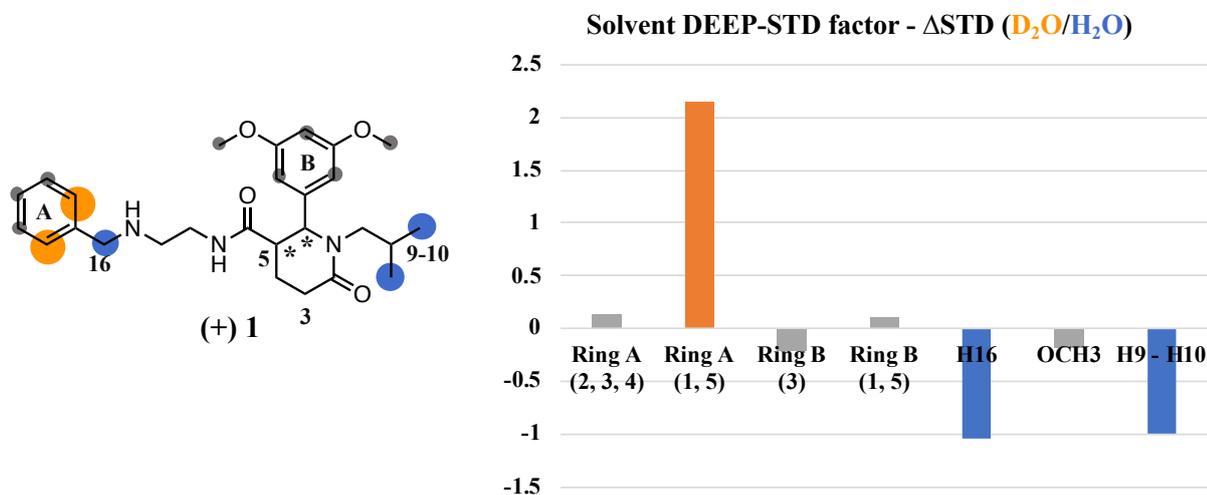


Figure 26. Graphic representation of the solvent DEEP-STD experiment on the (+) **1**-HuR complex. Blue surfaces indicate ligand contacts with protein side chains carrying slowly exchanging protons; orange surfaces indicate ligand contacts with apolar residues.

Differential solvent experiments are currently ongoing for the second enantiomer, (-) **1**. Pending assignment absolute configuration the binding modes of the two enantiomers will be rationalized with *in silico* techniques. The data thus obtained will be exploited for the development of new derivatives

with modified anchors (to understand the importance of anchor substitution) and specific stereochemistry, for improved affinity and RNA displacement ability.<sup>§§</sup>

---

<sup>§§</sup> The completed study, including assignment of the absolute configuration and complete DEEP-STD NMR evaluation, is described in the just accepted article “BOPC1 enantiomers preparation and HuR interaction study. From molecular modeling to a curious DEEP-STD NMR application”. ACS Medicinal Chemistry Letters. This content could not be reproduced yet and is thus not reported in the appendix.

## 6. Design of new derivatives of synthetic hit 4 via fragment-based approach

The present chapter describes the first fragment-based ligand design approach to obtain new HuR ligands based on the scaffold of synthetic hit 4 (described in chapter 5).

Briefly, a small halogen-enriched fragment library (HEFL, ~ 150 fragments<sup>\*\*\*</sup>), built to exploit halogen bonding in small substituted homo- and heterocycles, was selected to perform a biophysical fragment screening on native HuR. The primary screening technique chosen was Surface Plasmon Resonance (SPR) and STD NMR was used for confirmation of hit fragments. Three hits and few readily available (synthetically or commercially) derivatives were selected to build a library of analogs of synthetic hit compound 4 where the carboxylic acid moiety (ring C) is varied according to incorporate chosen fragments. The derivatives so designed are currently undergoing *in silico* selection (via the molecular dynamics and docking approach described in chapters 4 and 5). The structures with the best predicted affinities will be selected for synthesis and biophysical and biological evaluation. Structurally-related derivatives where ring B is replaced with benzylamine have also been designed and are under consideration for synthesis.

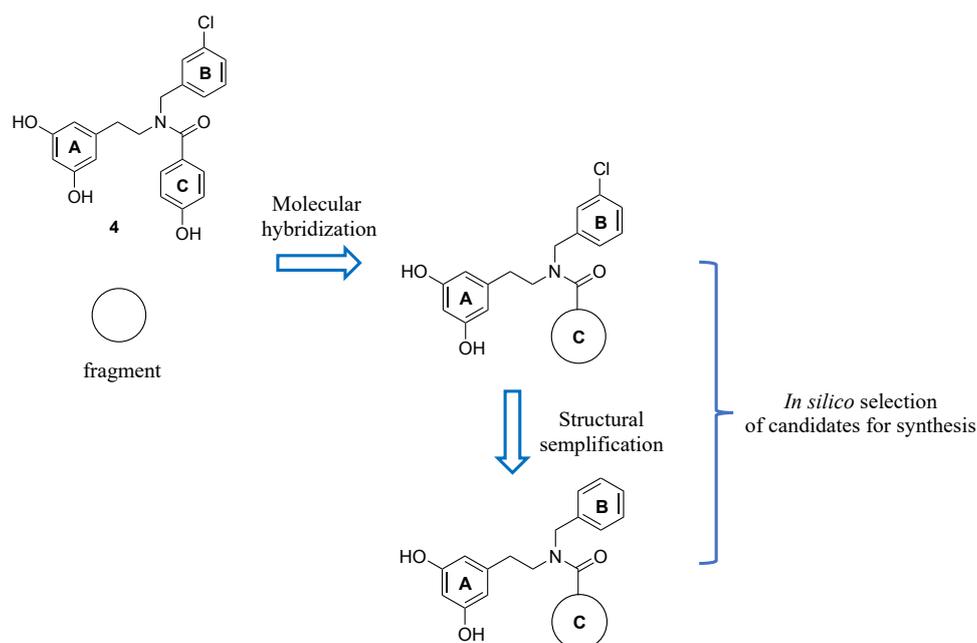


Figure 27. Design strategy of analogs of synthetic hit 4 by molecular hybridization with selected fragments.

### 6.1 Fragment screening on the native protein

Though three techniques were initially considered for fragment screening, Surface Plasmon Resonance (SPR), Thermal Shift Assay (TSA) and STD NMR, the choice of SPR as primary

<sup>\*\*\*</sup> The HEFL library was designed, synthesized and donated by Prof. Franck Boeckler and collaborators at the Pharmazeutische Chemie, Pharmazeutisches Institut Eberhard Karls, Universität Tübingen (Germany) based on the principles of halogen bonding.<sup>190-192</sup> The library comprises various homo/heterocyclic compounds diversely decorated with different functions and substituted with -Cl, -Br, -I.

screening technique was straightforward. In fact, despite its very informative nature and the fact that the technique had already been established, STD NMR requires comparably high amounts of protein; TSA suffers from similar drawbacks, thus both were only considered as secondary screening approach. In both cases in fact, the protein needs to be added into each tube (NMR) or well (TSA) and cannot be recovered. While protein amounts can be partly reduced with STD NMR by grouping the compounds to be screened in mixtures, spectra recording and full assignment of each mixture component is required prior to the experiment, as well as thorough mixture planning to avoid signal overlap. Additionally, utilizing mixtures is quite discouraged with TSA due to the high risk of artifacts.<sup>80</sup> On the contrary, SPR requires a much smaller amount of protein, due to the fact that one SPR chip with immobilized protein can be utilized for a whole screening.

Since no study was reported on the application of TSA on HuR, preliminary analyses were carried out to assess the optimal experimental conditions. Unfortunately, native HuR tested at different concentrations was unresponsive to the thermal cycles applied.

For these reasons, SPR was chosen as primary screening technique, followed by STD NMR to confirm selected hits.

#### ***Screening of the halogen-enriched fragment library by SPR***

The screening protocol was established based on a work by Wu *et al.* 2015<sup>66</sup> (see paragraph 1.3, compounds 14–19), which exploited amine coupling for protein immobilization of two different constructs, different than the native protein, at unreported concentration. The conditions thus needed to be optimized for the protein available to us; the immobilization was finally obtained through application of HuR protein to the activated chip at a concentration of 50 µg/ml in 10 mM acetate buffer pH 4.0 at a flow rate of 5 µl/min. Epigallocatechin gallate (see chapter 4, compound 1) was injected as reference compound to confirm integrity of the binding site prior to the screening. A three-concentration-point screening was performed in duplicates (250–500–1000 µM) on the (146 library members) allowing to select 10 fragments for  $K_D$  determination; the following curves were built with 10 concentration points (15 µM to 2 mM). Three fragments with estimated  $K_D \leq 1$  mM were selected as hits (see figure 28).

#### ***Validation of hit fragments by STD NMR***

Following the STD NMR protocol reported in the previous chapters, the interaction of the three hit fragments with HuR was confirmed (see figure 28). Given the small size of the fragments it was chosen not to determine the epitope mapping.

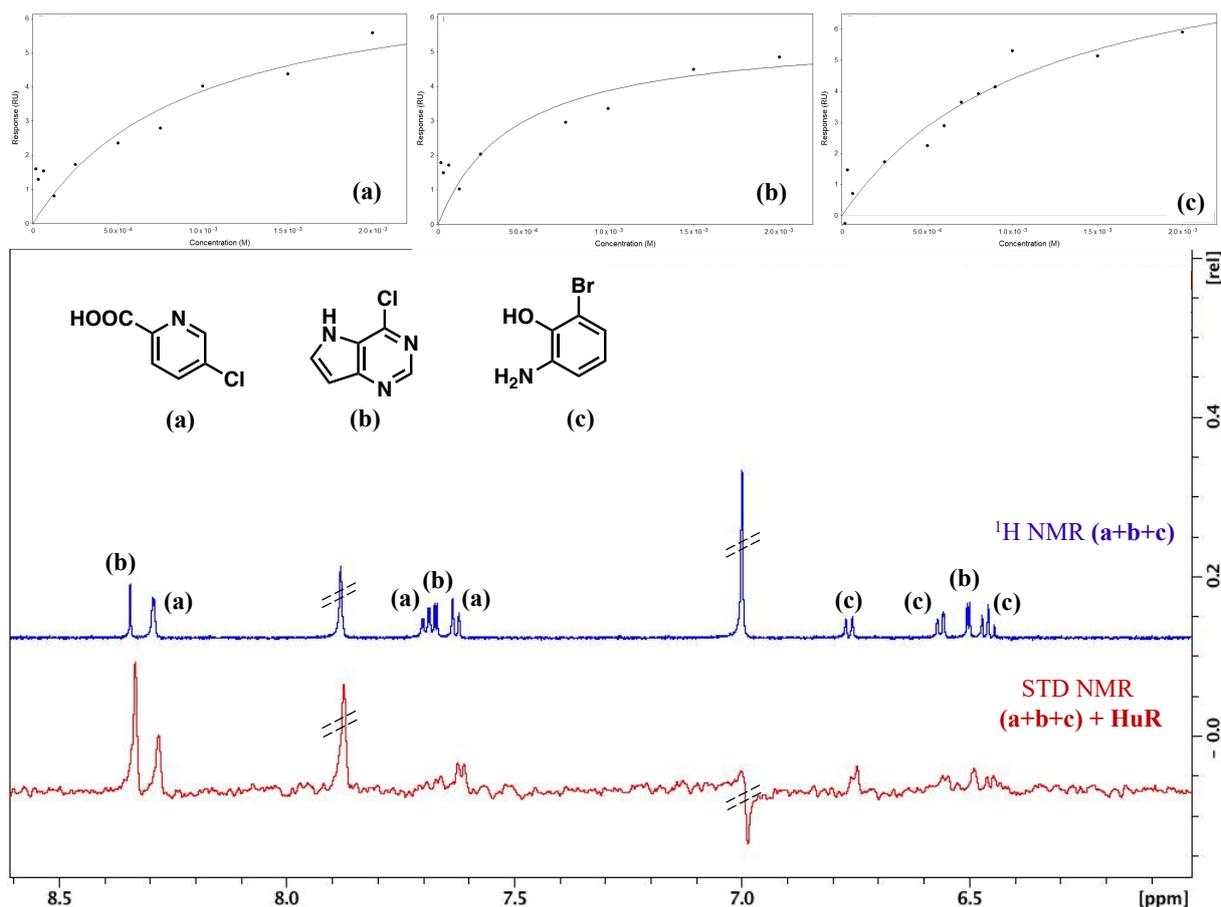


Figure 28. Structures,  $K_D$  curves and STD NMR spectra of the three hit fragments selected against native HuR, identified with the letters **a**, **b**, **c**.

## 6.2 Molecular hybridization of HuR ligand **4** and hit fragments

### Design of new derivatives of **4** for *in silico* screening

The linking step was devised starting from the original design of **4** maintaining a fixed anchor (3,5-dihydroxyphenyl – ring A) and side chain (3-chloro benzylamine – ring B), and varying the carboxylic acid moiety (ring C) according to the hit fragments identified and few derivatives. The commercial and synthetic availability of carboxylic acid derivatives of the 3 hit fragments, assessed through database search (SciFinder, PubChem), afforded the design of 24 derivatives of **4** where ring C is replaced (see figure 29). The analogs of **4** so designed are currently undergoing *in silico* selection through the molecular dynamics and docking approach described in the previous chapters. The most promising compounds will be selected for synthesis and evaluation according to their predicted affinity to the target HuR.

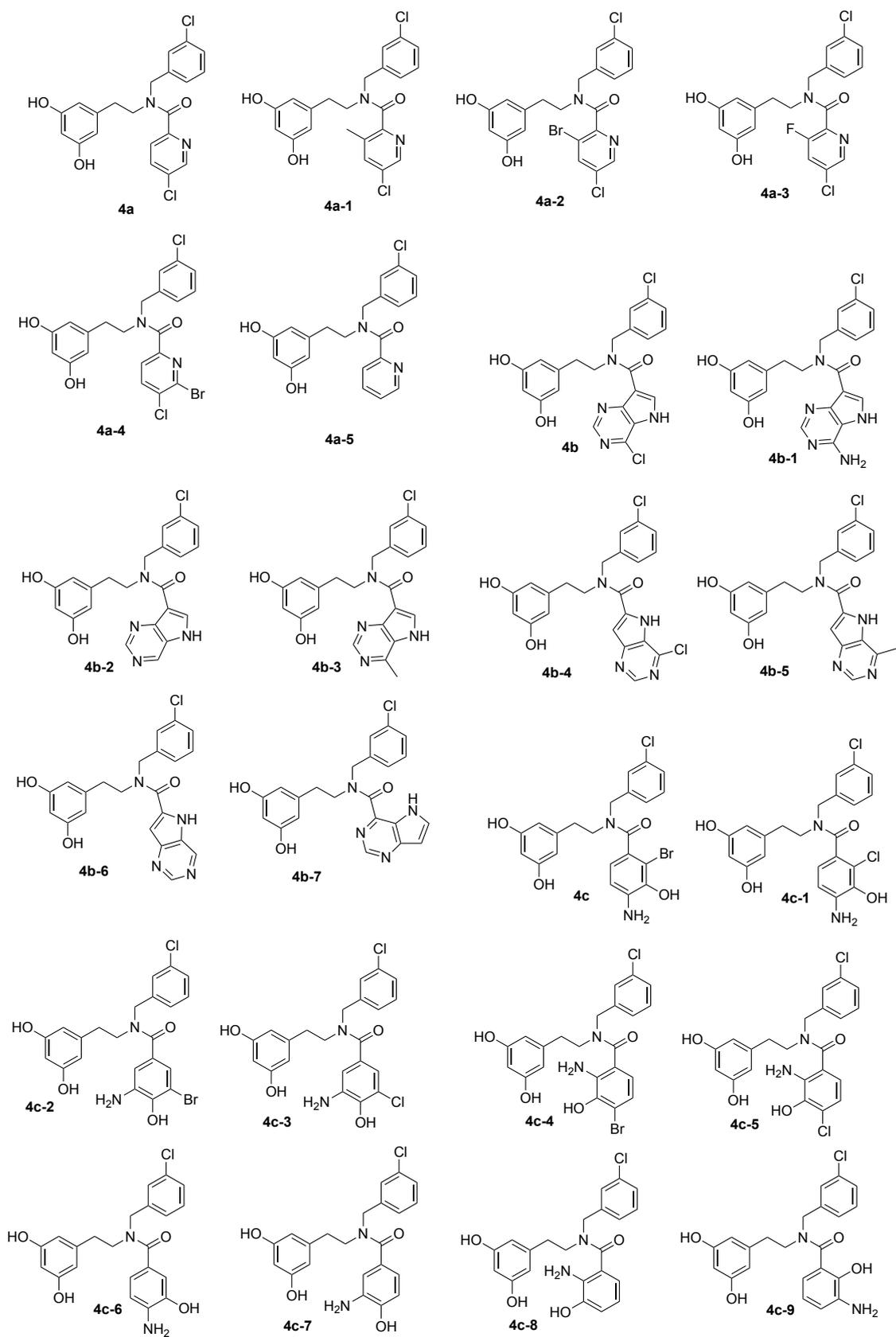
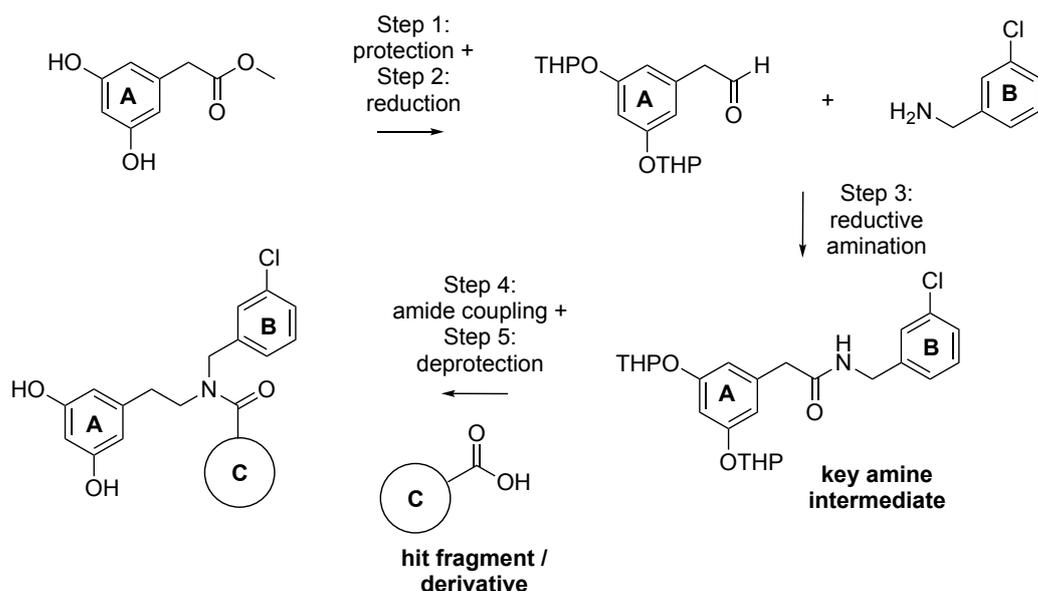


Figure 28. Structures of the designed analogs of compound 4 where ring C is replaced with hit fragments and their derivatives.

### Optimization of the synthetic pathway

In view of having to synthesize a library of selected analogues of **4**, the original synthetic pathway exploited, should be revised, to improve the general yield as well as reduce the number of steps. The main issue concerns the sensitivity of the O-THP groups, present on each intermediate through each step of the reaction, to any mildly acidic additive or purification media necessary to obtain the final compound, which calls for additional long purification steps and in the end can lead to poorly reproducible yields.

#### Original Synthetic Strategy



#### Alternative Synthetic Strategy

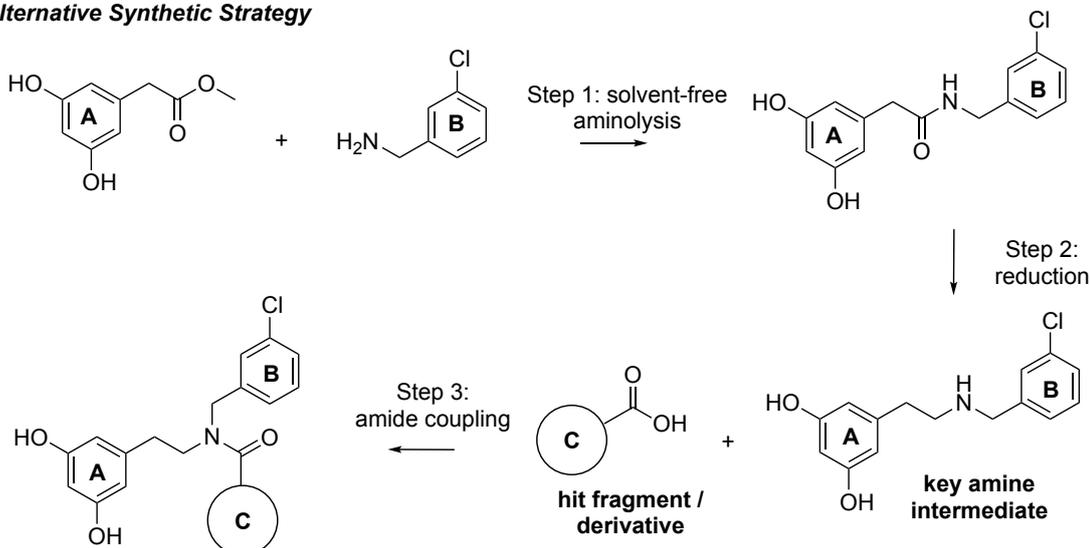


Figure 30. Comparison of original and alternative synthetic pathway devised for the synthesis of derivatives of **4**.

To reduce its impact, a new strategy has been devised to reach the key amine intermediate necessary for the synthesis of the analogs designed via amide coupling to the different carboxylic acids chosen.

The new synthesis involves a solvent-free aminolysis (step 1) of the ester to obtain the amide product; this is then reduced (step 2) to afford the amine ready for the amide coupling (step 3). So far, the first two steps were performed on a small scale (without column purification) and followed by TLC and  $^1\text{H}$  NMR to prove feasibility.

The synthesis is currently under scale up to obtain the key amine intermediate in quantity and purity suitable to synthesize the analogs of **4** currently under selection. A similar synthetic pathway could be applied for the synthesis of derivatives with a nude benzyl B ring. In case hydroxyl protection should once again be considered, to avoid side reactions occurring during the final step, this could be performed right before avoiding the inconvenience of having an unstable moiety throughout the whole pathway. In case of side protection on the secondary amine, a OH-selective protective group could be considered to avoid this interference.

### ***Experimental limits of the native HuR protein***

The fragment-based approach described in the present chapter follows a rather unusual pathway, mainly due to several issues encountered with the native protein. In fact, native HuR is a quite unstable protein, very sensitive to changes in pH (*i.e.*, pH necessary for the immobilization reaction), temperature (degradation seems to occur prior to TSA thermal cycle), prone to aggregation (cannot be concentrated in solution, the highest concentrations reported are  $\leq 10\ \mu\text{M}$ ) and with reduced stability in frozen storage. This instability is exacerbated by removal of glycerol from the preservation conditions, though this is necessary for most assays used. All of these characteristics made it an impossible candidate for DCC applications as well as additional assays. Moreover, while intra-batch reproducibility for SPR immobilization was quite favorable, the same was not found for inter-batch reproducibility. Such issues were not encountered within STD NMR experiments, probably due to the minimal manipulation and favorable experimental conditions necessary.

The stability of the protein could only be investigated by SDS-PAGE and Bradford assay, since TSA was not an option. SDS-PAGE can highlight changes in size of the protein (cleavage, fragmentation and aggregation), while the Bradford assay can quantify the total protein present in a sample concentration through the binding of protein molecules to Coomassie dye under acidic conditions, resulting in a shift in color from brown to blue which can be detected with a spectrophotometer. Both methodologies, were chosen due to their selectivity of protein detection, compared for example with UV readings normally performed at 280 nm for protein detection where several interferers also absorb (*i.e.*, residual imidazole from purification and “pollutant” non-protein macromolecules). Nonetheless, they could only afford partial information,<sup>†††</sup> confirming the limits of the native protein and clearly

---

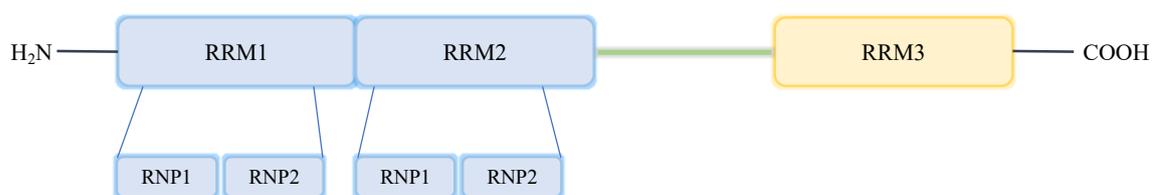
<sup>†††</sup> Experimental data not reported in the present thesis.

highlighting the quite urgent need for a solution for both current and future studies. The solution found and its applications are reported in chapter 7.

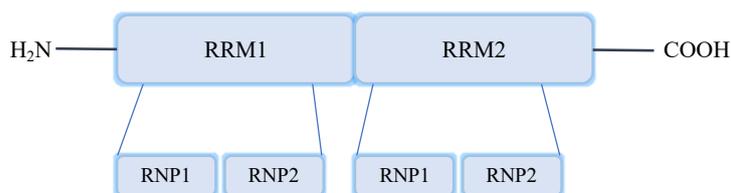
## 7. *Fragment linking by protein templated dynamic combinatorial chemistry to identify new HuR ligands*

The present chapter aims at solving the cliffhanger at the end of chapter 6, searching for a solution to the experimental limits of the native protein. Thanks to a newly started collaboration with the Sattler group (Institute of Structural Biology, Helmholtz Centre Munich), two new protein constructs were introduced to support the study, one modified full length HuR (RRM1+2+3\_GGS) and one truncated (RRM1+2, see figure 31).

**Construct 1: Full length protein, modified linker (Glycine–Serine); ~37 kD**



**Construct 2: Truncated protein; ~20 kD**



*Figure 31.* Schematic representation of the two protein constructs (full length, RRM1+2+3\_GGS and truncated, RRM1+2) expressed, purified and shared from the Sattler Group, utilized for the next steps of the study.

After initial study of the two protein constructs, both were immobilized for the fragment screening but, finally, only RRM1+2 gained reliable results and could be used with favorable conditions for the establishment and execution of the protein-templated DCC (pt-DCC) approach finally exploited for fragment linking. The pt-DCC experiments performed on two different dynamic combinatorial libraries (DCL) provided three hits and five additional derivatives to subject to interaction studies by STD NMR with the target protein. The optimization of the STD NMR experimental conditions for the new constructs are currently ongoing.

### 7.1 *Initial studies on the two new protein constructs*

The two constructs, expressed and purified by the Sattler group,<sup>74</sup> represent a modified full length HuR, RRM1+2+3\_GGS (1–326, 37 kDa), where the linker between RRM domains is substituted by a series of Glycine and Serine residues, and a truncated HuR, RRM1+2 (1–186, ~20 kDa), where only the first two RRM domains are present. Both constructs show higher stability compared to native protein and are less prone to aggregation and precipitation; for this reason they can be concentrated

in solution up to 27  $\mu\text{M}$  (1 mg/mL) and 125  $\mu\text{M}$  (2.5 mg/mL) respectively, and have so far shown better behavior in all assay performed (NMR-based assays carried out in Munich). To assess their potential for our objectives, both were tested in a thermal shift assay (TSA) in their native buffer (20 mM Phosphate buffer pH 7.0) at six protein concentrations (0.25–10  $\mu\text{M}$ ). Both constructs responded ( $T_m$  RRM1+2+3  $\sim 58^\circ\text{C}$ ,  $T_m$  RRM1+2  $\sim 56^\circ\text{C}$ ) at 2.5–5  $\mu\text{M}$  and 5–10  $\mu\text{M}$  respectively. Additionally, immobilization on a SPR chip was tried for both constructs following the same procedure exploited for the native protein via amine coupling. The immobilization was successful at the first trial and showed consistent signal in time ( $\sim 3000$  and  $2000$  RU respectively). For this reason, it was chosen to carry out the investigation in parallel on both constructs; the full length protein was planned to be used for the screening, and the truncated to select possible aspecific binders.

## 7.2 *Fragment screening on protein constructs by SPR*

The halogen enriched fragment library (HEFL) of chapter 6 was thus rescreened against the two constructs via SPR, following a similar protocol to the previous screening. Some of the hits selected show structural resemblance to the hits found for the native protein but, given the possibility to exploit at least one of the protein constructs for pt-DCC experiments, the hit selection was kept to a larger size (8 hit fragments, see figure 34) to guide choice of building blocks to design the DCLs.

### *Full protein construct: RRM1+2+3\_GGS linker*

After immobilization of the construct, epigallocatechin gallate (EGCG, see chapter 4, compound **1**<sup>60</sup>) was injected as reference compound prior to the screening to verify the integrity of the binding site. Fourteen concentration points in duplicate (25 nM–500  $\mu\text{M}$ ) could not reach saturation of the protein binding site (see figure 32).

A second reference compound (**mfm 49**<sup>164</sup>) was also tested (25 nM–100  $\mu\text{M}$ , due to limited solubility) obtaining similar results (reported estimated  $K_D$  for both compounds is  $\sim 0.1$   $\mu\text{M}$ ). An analogous behavior was later on observed for a selected fragment, we chose not to pursue the screening on the construct RRM1+2+3 construct. Moreover, obtained RU values are well over the expected response for 1:1 ligand–protein construct interaction.<sup>†††</sup>

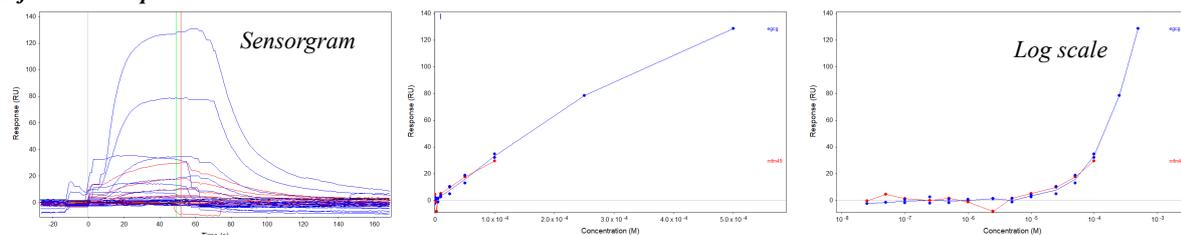
---

<sup>†††</sup> According to the equation:

$$\text{analyte binding capacity (RU)} = \frac{\text{analyte MW}}{\text{ligand MW}} * \text{immobilized ligand level (RU)}$$

(Biacore Sensor Surface Handbook BR-1005-71 Edition AB).

### Reference compounds



### Fragments

Figure 32. Sensorgrams and concentration curves for two reference compounds (EGCG, see chapter 4, compound **1** in blue, and synthetic compound **mfm 49** in red) and one fragment (in orange). No reference compound reaches saturation on this construct, the response just keeps growing in intensity. Similar behavior was observed with one selected fragment.

### Truncated protein construct: RRM1+2

The second construct was then immobilized for testing. Once again, epigallocatechin gallate (EGCG, see chapter 4, compound **1**) and **mfm 49** were injected as reference compounds prior to the screening to verify the integrity of the binding site. Saturation was reached for both compounds and estimated  $K_D$  reproduced reported values within an acceptable range (curves reported in figure 33).

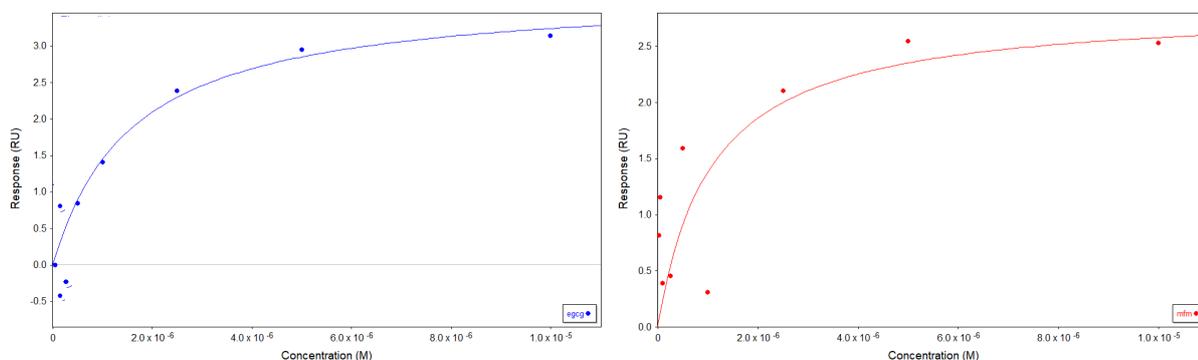


Figure 33.  $K_D$  curves for reference compounds (EGCG, see chapter 4, compound **1** in blue, and synthetic compound **mfm 49**,<sup>164</sup> in red). Estimated  $K_D$  reproduced reported values in an acceptable range.

A three-concentration-point screening was performed in duplicates (250–500–1000  $\mu\text{M}$ ) on the (146 library members) allowing to select 20 fragments for  $K_D$  estimation; the following curves were built with 5 concentration points (125  $\mu\text{M}$  to 2 mM). Eight fragments with estimated  $K_D$  below 1 mM were selected to guide the construction of DCLs.<sup>§§§</sup>

<sup>§§§</sup> Additional structurally-related derivatives tested within the HEFL could be considered in future studies for the development of additional DCL.

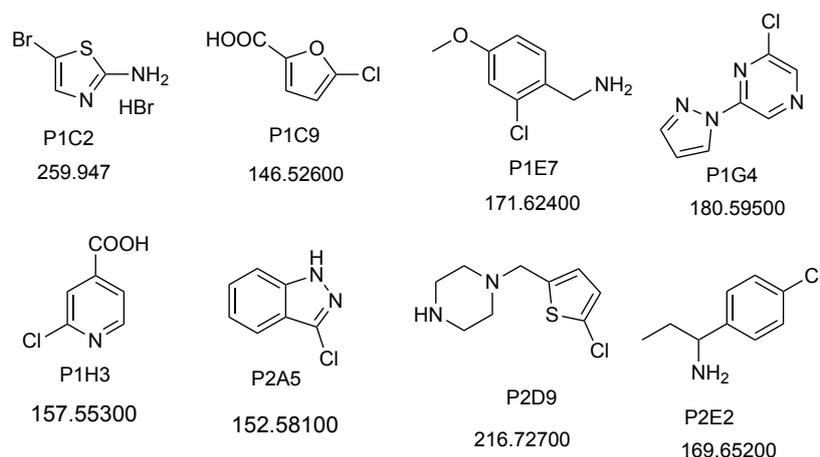


Figure 34. Structures and molecular weights of the fragments selected via SPR screening on protein construct RRM1+2.

### 7.3 Dynamic Combinatorial Chemistry studies

According to the chosen chemistry of acylhydrazones, two DCL were so far built, choosing as building blocks commercially available aldehyde and acylhydrazide derivatives inspired to the eight hit fragments coming from the SPR screening. pt-DCC experiments were conducted in two buffers, selected according to the protein construct stability, monitored through HPLC-UV-MS against the same blank reaction to allow for the discovery of hits (amplified compounds). Three DCC hits were chosen for synthesis and evaluations along with DCC negatives and additional derivatives.

#### Establishment of experimental conditions

TSA was used to monitor protein stability over time in different buffers (10–50 mM; pH range 5–7.5; 150 mM NaCl) with the aim of establishing the optimal experimental conditions for the pt-DCC experiment. The assay was performed at a compromise protein concentration between optimal TSA and DCC ranges. One 96-well plate was thus prepared where a 10  $\mu$ M solution of protein in sixteen buffers and H<sub>2</sub>O containing 5% DMSO, was stored at room temperature over a time range fully covering DCC timeframes, and sampled every day (day 0 to day 3) for TSA experiments. Within the pH range tested, the protein showed little  $T_m$  variation (see experimental section); nonetheless, two buffers, phosphate pH 7 and acetate pH 6 were chosen to perform the pt-DCC experiments.

#### Fragment-inspired dynamic combinatorial libraries

So far, two fragment-inspired DCL libraries were built and exploited for pt-DCC on RRM1+2, utilizing commercially available aldehydes and hydrazides as building blocks.

Based on past experience (both personal and of the research group), it was chosen to use small building blocks libraries to construct DCLs, based on the chemistry of acylhydrazones. Each building block library comprises two aldehydes (each 100  $\mu$ M) and eight hydrazides (each 200  $\mu$ M). The protein construct RRM1+2 was used at the concentration of 40  $\mu$ M and 1 mM aniline was added to

speed up the rate at which equilibrium is reached at the pH chosen for the experiment (see paragraph 3.4.4). The experiments were conducted in parallel for the same DCL in the two different buffers, and monitored via HPLC-UV-MS upon protein precipitation (by adding ACN) and pH-induced “freezing” of the reaction. As reported in paragraph 3.4.4, it is essential to separate, identify and quantify the amount of each DCL member formed univocally and without incurring in any artifacts to determine both equilibrium and pt-amplification.

Within the two libraries (see figure 35), three hits were selected (with amplification > 80%); additional derivatives were chosen for synthesis and testing, for a total of eight final compounds, chosen among DCC negatives and hit derivatives.

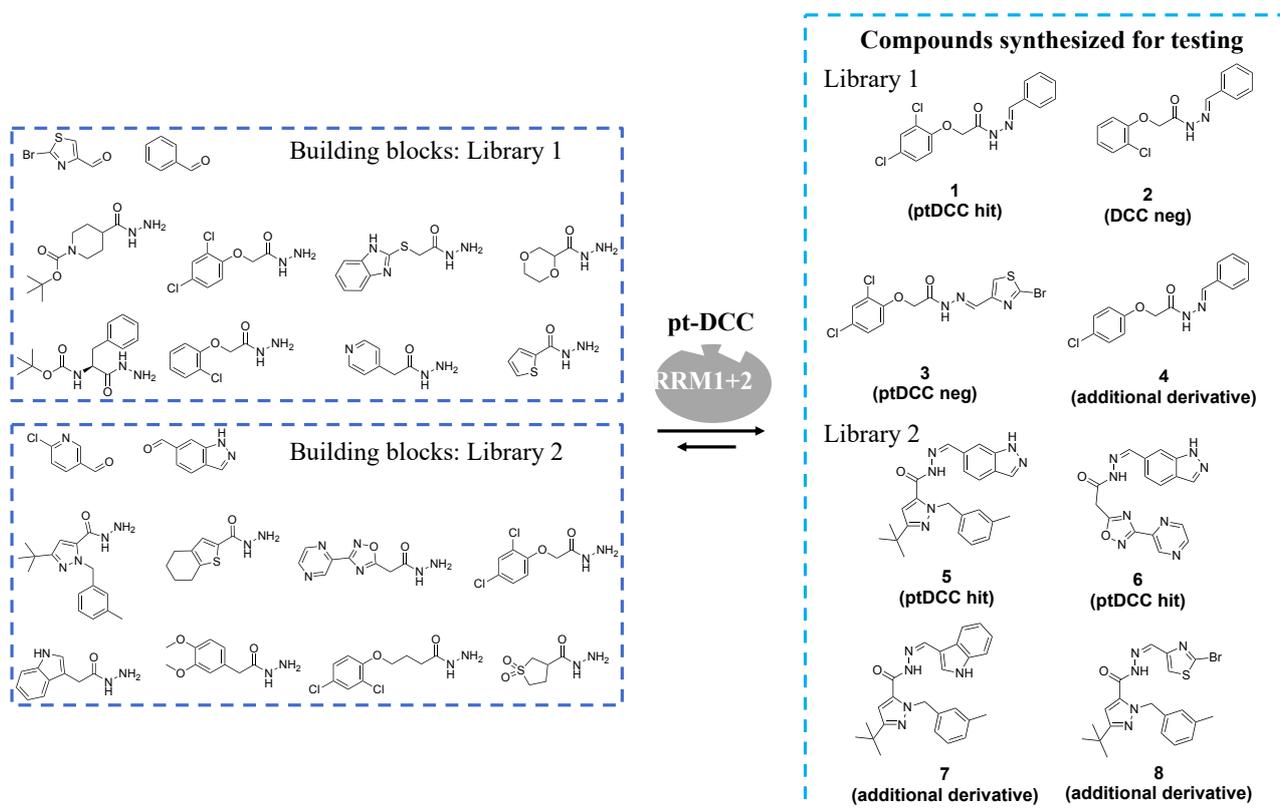


Figure 35. Fragment-inspired DCC libraries used for the pt-DCC experiment, and related hits and additional compounds synthesized for testing.

#### 7.4 Preliminary interaction studies

Though the pt-DCC experiment allows in principle to combine synthesis of ligands and screening for affinity in one step, it is still fundamental to evaluate the affinity, study the binding modes with the target and evaluate the biological activity. Since the optimization of the STD NMR experiment on the new constructs is ongoing, a preliminary study was performed through SPR, following the protocol utilized for fragment screening and  $K_D$  estimation. Nine concentrations were injected for each final compound (50 nM–500  $\mu$ M). Only one measurement could be performed so far, so the

results need further validation. Nonetheless, this preliminary study showed a promising improvement in binding affinity for the DCC hits compared to the single fragments tested.

A second affinity measurement and STD NMR study are in progress; their results will assess the validity of this approach for the target HuR thus, adding to the pool of strategies which can be exploited for the design and synthesis of new HuR ligands. In case of positive results, the compounds could be analyzed by DEEP-STD and protein observed NMR to build a complete map of their interaction with HuR. Additional assessment of their ability to displace RNA could then be performed (by STD NMR and Fluorescence Polarization\*\*\*\*) to select compounds able to interfere with HuR–RNA binding.

---

\*\*\*\* Assay under development for the new constructs.

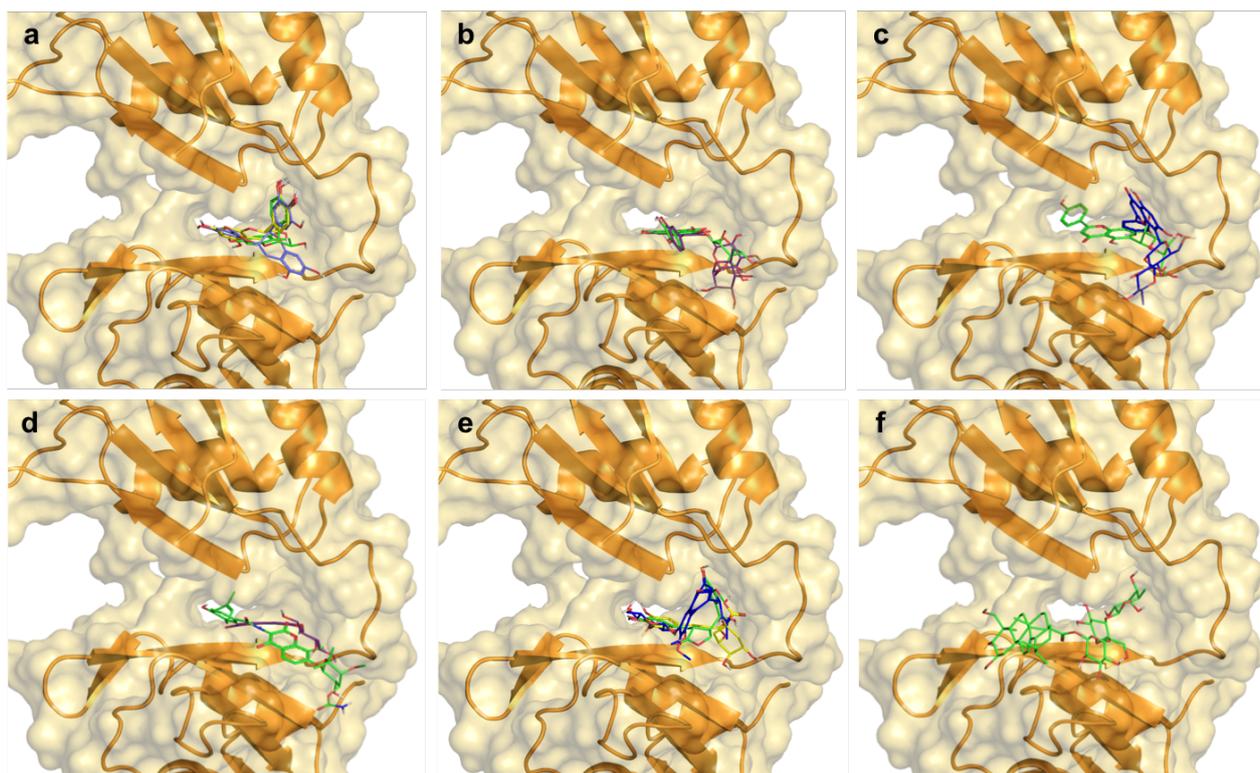
## 8. Discussion

The work reported in the present thesis focused mainly on adding some puzzle pieces to a research project aimed at developing a new series of small, drug-like organic molecules able to modulate the stability of protein–RNA complexes involved in several diseases, to regulate gene expression with an unprecedented mode of action. The main target of this project is the ELAV protein family, a biochemically very well characterized group of highly homologous proteins. This class is particularly interesting for medicinal chemists and pharmacologists due to the involvement of related protein and protein–RNA dysregulations in several diseases (including cancer, neurodegeneration and inflammatory diseases). Regardless of the high potential of the ELAV proteins, not much is known about small-molecule modulators of either the proteins themselves or their complexes with RNA; most literature results come from independent HTSs based on different techniques and libraries, and, at the time this thesis was started, neither rational design nor HTS-hit evolution had been performed. This can be mainly attributed to the fact that, while the protein–RNA interface has been studied, this interaction is rather far from the more common receptor or enzyme binding sites we are used to studying. In particular, RNA and protein interact through a large surface, most of the times through a high number of superficial contacts; since we do not have any experimental evidence of the presence of effective and targetable hotspots, the development of a small molecule able to displace such an interaction requires several iterative steps.

For this thesis, we decided to apply a strategy that focused mainly on developing compounds with affinity for the ELAV subtype HuR, chosen as a model protein and with high potential for therapeutic applications (*i.e.*, cancer and retinopathy), by targeting a site on the large protein–RNA interface. While RNA is clearly also a component of the interaction, and the RBD sequences on ELAV proteins (as well as several RBPs) are highly homologous, we believe that targeting the protein along the RBD could represent a safer approach rather than targeting RNA transcripts. First of all, the diversity in protein-binding domains of target RNAs is even lower, as high affinity RBDs are reported to mostly bind AU-rich sequences (see paragraph 1.1), thus posing a greater challenge for target RNA selectivity; moreover, targeting RNAs could also be considered potentially less safe, due to the similarity of nucleic acids in general and the risky impact on cell function (as seen, for example, with DNA intercalating drugs<sup>172,173</sup>).

Nevertheless, since a “proper” binding site for small molecule–HuR, as well as the related structural features, had not been elucidated at the time this thesis was started, we had to go back to the very beginning. First of all, exploiting the information coming from literature HTSs, we grouped a number

of flavonoid and flavonoid-like compounds all derived from natural sources as well as some additional natural-product derivatives. Some of these compounds were directly picked out from the HTS hits, as compounds already known to interfere with ELAV–RNA complexes, while others were chosen to explore the chemical space around the flavonoid scaffold and other related and unrelated structures, for a total of 13 compounds finally considered. The collection so generated was utilized in the first study performed for this thesis (described in chapter 4) which exploited a combination of NMR-based ligand-observed technique, STD NMR, coupled with *in silico* methodologies, molecular dynamics and docking, to gain a comprehensive study of HuR–small molecule interactions. As previously described, STD NMR allows us to identify (at the atomic level) the moieties of the small molecule involved in the interaction with a target macromolecule as it occurs in solution; these data can be utilized to validate the binding modes predicted by *in silico* studies thus providing a two-way map of the interaction. Out of 13 compounds studied, 12 showed direct binding with HuR in STD NMR studies and one, asiaticoside (compound **13** of chapter 4) did not show any interaction; moreover the NMR data elaborated in epitope maps allowed us to analyze the interaction with HuR for each compound, but also within each structural class, showing how different substitutions affect binding affinity as well binding modes. NMR data were compared with *in silico* results; in particular, given the high protein mobility and flexibility which oversees the conformational changes occurring upon ligand binding, the HuR–RNA complex (PDB 4ED5) was studied with molecular dynamics to reproduce the conformational rearrangements of the protein–interaction interface. Though longer and computationally more demanding than rigid docking methods, this strategy gave us the chance to simulate the interaction in a more easily relatable way to STD NMR studies. In the following docking studies, all interacting compounds were predicted to bind to repeated amino acid residues of the protein interaction site, within a deeper pocket-like region of the HuR–RNA interface; the NMR data were in best accordance with binding predictions of the compounds to a closed conformation of the protein (see figure 36). On the other hand, in accordance with experimental data, asiaticoside was predicted to only bind superficially to the protein (see figure 36); the interaction was investigated over different conformations with little difference. According to the results obtained, the compounds investigated seem to comply to the theory that interfering compounds could stabilize a closed conformation of HuR (and other ELAV proteins) thus impeding the RNA initial contacts to the protein which trigger the protein conformational rearrangements that bring to the formation of the HuR–RNA complex.<sup>75,163</sup> This study also allowed us to propose a postulated binding site which was then exploited in the next steps of the thesis work for SBLD/SBVS, and has the potential for the application of SBVS on numerous libraries for future applications.



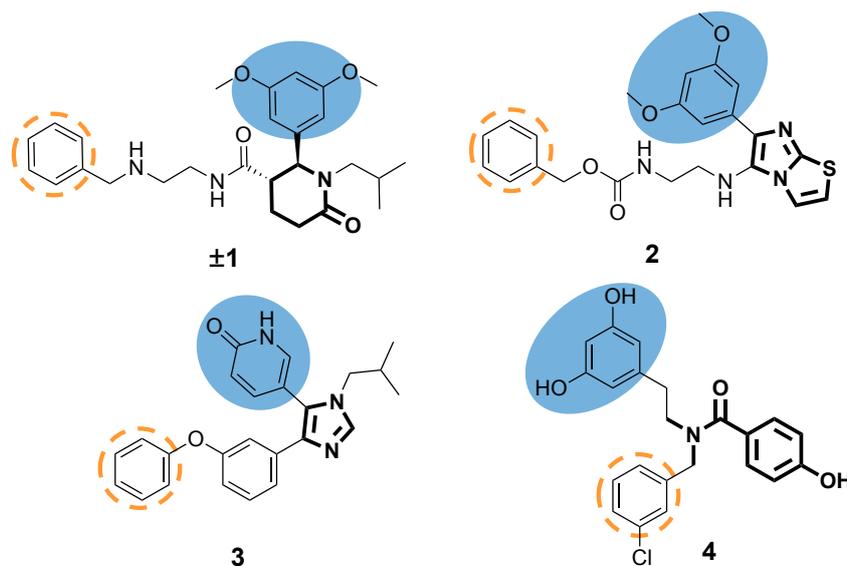
*Figure 36.* 3D representation of compounds divided by compound class in the HuR–RNA interaction site: (a) compounds **1**, **2**, and **3** (Epigallocatechin gallate, epicatechin and catechin, blue, green and yellow sticks, respectively); (b) compounds **4** and **5** (hyperoside and rutin, green and violet sticks respectively); (c) compounds **6** and **7** (vitexin rhamnoside and isovitexin, blue and green, respectively); (d) compounds **8** and **9** (aesculin and novobiocin, violet and green sticks); (e) compounds **10**, **11** and **12** (chlorogenic acid, 7-hydroxymatairesinol and colchicoside, yellow, green, blue sticks, respectively); (f) compound **13** (asiaticoside) shown as green sticks; HuR protein shown as orange cartoon (adapted from Vasile *et al.*, 2018, see appendix III, article 1).

The key features of the small molecule–HuR interaction so identified (for both ligand and protein counterparts) can be summarized as follows:

- a) unsaturated conjugated regions as well as aromatic rings on the ligands are of high importance in the interaction with HuR;
- b) these moieties appear to fit in repeated regions, and thus show hydrophobic interactions with repeated residues, namely Ile23, Asn25, Ser94, Ile133 and Arg153;
- c) for the sugar moieties, we observed that the rings directly attached to the aromatic ring of compounds **4**, **5**, **6**, **7** and **8** and the quinic acid moiety of **10**, overlap in the same region of the HuR binding site, establishing contacts with repeated residues (Arg97, Pro98, Ser99 and Ile 103);
- d) additional sugar moieties generally protrude outside the pocket-like region considered establishing more superficial contacts.

These observations were exploited in the next step of the thesis (described in chapter 5) for our first medicinal chemistry campaign. The visual inspection and comparison of the small molecule–HuR

hypothesized binding site and the RNA–HuR experimental binding site allowed us to identify two ribonucleotides involved with interactions in the same pocket-like site of the studied compounds, uridines 8 and 9. We focused on these two residues for the design of new compounds, searching for new scaffolds to build HuR-ligands utilizing the free web-based virtual screening platform AnchorQuery;<sup>165,166</sup> U8 and U9 were selected as anchors to examine whether one of the two nucleotides could represent a pharmacophore concerning occupied position and interactions. Modifications were applied to grow and develop the anchors, in particular to gain an additional a  $\pi$ – $\pi$ -stacking interaction with Tyr63, and to enhance synthetic feasibility. 800 candidate structures were generated and screened for synthesis through visual inspection and molecular recognition studies, to select compounds complying to the following criteria: (a) good overlap of the anchor with the corresponding uracil ring, preserving hydrogen bonds with Arg97 and Lys92; (b) the additional aromatic rings should gain the desired interaction with Tyr63. The most promising scaffolds thus selected brought to the synthesis of final compounds **1–4** based on four different MCRs (see figure 37).



*Figure 37.* Structures of the four compounds selected for the experimental study. The modified anchors are highlighted with light blue full circles, while additional aromatic rings designed to interact with Tyr63 in a dashed orange circle. The different scaffolds (piperidinones, heterocycles, N,N-disubstituted amides) are highlighted in bold (adapted from Della Volpe *et al.*, 2019, see appendix III, article 2).

Of four compounds synthesized, three (**1**, **3** and **4**) could be subjected to the combined STD NMR and *in silico* investigation of their interactions with HuR, owing to their favorable solubility in buffer. Finally, **1** and **4** were selected as hit compounds to develop in the next steps of our MedChem campaign; nonetheless, these two compounds show quite different binding modes. In particular, compound **4** shows interesting interaction features which seem so far to be in accordance with the

initial design: while all rings interact with the target, the anchor, as hypothesized, appears to be the most interacting moiety.

On the contrary, compound **1**, while directly binding to the target, shows a different interaction mode which required further investigation prior to the design of new derivatives. Several factors could be involved into the differences of compound **1**: first of all, we initially synthesized and tested **1** as a racemic mixture of the two trans enantiomers;<sup>171</sup> as it often happens with a chiral molecule within a chiral environment, it is likely that the two enantiomers could have different interaction with the target, thus they should be separated and investigated individually. Moreover, regardless of its similarities with **4**, the modified anchor of compound **1**, has both hydroxyls substituted; though the additional methyl group is rather small, it can change the binding capacity of the anchor (*i.e.* the anchor now lacks hydrogen-bond donors) and could thus produce a different binding mode. Finally, the compound could have higher affinity for a different protein site, which we cannot detect with STD NMR but need a protein-observed NMR method, rather than X-ray diffraction.

The first step for the elucidation of compound **1** binding modes involved the resolution of the racemate in the two separate enantiomers by enantioselective HPLC; a separation method was first established at the analytical scale and then transferred to the semi-preparative scale, exploiting a cellulose-based chiral stationary phase. We afforded both enantiomers in good enantiomeric excess and measured their specific rotation. (+) **1** and (-) **1** were thus subjected to STD NMR studies along with an enantio-application of DEEP-STD NMR to compare their STD spectra and elucidate the differences in their binding modes. Pending the assignment of their absolute configuration via crystallization and X-ray diffraction or NMR (based on the coupling with two enantiomers of a chiral derivatizing agent), these data will be compared with their *in silico* poses. This way, the collected spectroscopic information will guide the synthesis of several derivatives with modified anchors (to understand the importance of anchor substitution), for improved affinity and RNA-displacement ability. Nonetheless, we can already draw some conclusions, from the spectroscopic data obtained. In fact, the spectral differences processed through an enantio-DEEP-STD application suggest that the two enantiomers interact through the same protons but have different orientation into the binding site. In particular, (+) **1** shows a stronger contribution of the anchor moiety compared to (-) **1**, and is thus in better accordance with the initial anchor design. This consideration on stereochemistry will be taken into account for the development of derivatives.

The second approach applied in the present thesis for the design of new HuR-ligands exploited FBLD; according to this method, we screened and identified low-molecular-weight ligands (~150 Da) which can bind to the target protein. As anticipated in the dedicated paragraph, compared with HTS, the

fragment-based approach requires fewer compounds to be screened, and, despite the lower initial potency of the screening hits, it can offer more efficient optimization campaigns;<sup>115</sup> fragment libraries are typically screened using a series of progressively more restrictive biophysical techniques to refine the search for hit fragments and discriminate false positives and aspecific binders. Based on the expertise of the Hirsch group, the initial plan was to exploit pt-DCC for fragment linking to design and select new HuR ligands. This method allows to directly evaluate the different products which are part of the DCL and select the best binders through the pt-DCC experiment. Thus we could directly synthesize DCC hits and few derivatives, avoiding the synthesis of each DCL member for the next evaluation step. Based on our knowledge, literature-reported features and preliminary results we decided to apply SPR as primary fragment screening technique followed by STD NMR for hit fragment validation, and to utilize TSA only for studying the protein stability. In line with the previous studies reported in the thesis, the first experiments were performed on the native HuR protein. Though the fragment screening was successful and allowed for the selection of three hits (as described in chapter 6), we started to encounter several issues within experiments which could be reconducted to the stability of native HuR. In fact, native HuR is a quite unstable protein, very sensitive to changes in pH (*i.e.*, during the immobilization step in SPR), temperature (degradation seems to occur prior to TSA thermal cycle), prone to aggregation (cannot be concentrated in solution, the highest concentrations reported are  $\leq 10 \mu\text{M}$ ) and with reduced stability in frozen storage. This instability was exacerbated by removal of glycerol from the preservation conditions, though this action is necessary for most of the assays used. Moreover, we observed good intra-batch reproducibility for SPR immobilization but the inter-batch reproducibility was very low. These issues had not been identified in previous experiments exploiting STD NMR, probably due to the minimal manipulation and favorable experimental conditions utilized. All of these characteristics made it an impossible candidate for DCC applications as well as additional assays. For this reason, we decided to look for alternative constructs, and found two (one full-length modified construct, RRM1+2+3\_GGS, and one truncated construct composed by RRM1/2 domains, RRM1+2).

We decided to keep the results coming from the screening on native HuR and to design a series of derivatives of hit compound **4** (obtained through the structure-based study of chapter 5), where ring C is replaced with the three hit fragments and related derivatives (see figure 28). 24 analogs were designed based on the availability/synthetic feasibility of the related building blocks, and are currently under structure-based evaluation (VS) to select the predicted most interesting candidates for synthesis. In the meantime, a new synthetic pathway was thus devised and tried in small scale to

allow for the synthesis of a key intermediate to then generate in one step the derivatives selected (see figure 29).

In parallel, the two new constructs of HuR (provided by the Sattler group, see appendix V contributions to the research) were studied to assess their potential for both biophysical assay- and pt-DCC-applications. Both constructs proved to respond to TSA at reasonable concentrations for further studies (RRM1+2+3\_GGS at 2.5–5  $\mu$ M, RRM1+2 at 5–10  $\mu$ M) and showed immediate and stable over-time immobilization on the SPR chip. For this reason, they were both considered for repeating the screening step. At this stage, RRM1+2+3\_GGS showed to have some limits as well, as titration curves with two reference compounds could not reach saturation and did not reproduce reported  $K_D$  values (one hypothesis could take into account the possibility of aspecific interaction of the compounds with the modified linker). The same behavior was not encountered with RRM1+2, which proved to gain reliable results, thus this construct was used to repeat the whole screening, affording the selection of 8 hit fragments with  $K_D < 1$  mM. Its general characteristics (reported in chapter 7), including the possibility to concentrate it in solution and the experimentally demonstrated stability within a useful buffer range over time (studied through TSA, see chapter 7), made it an ideal candidate for pt-DCC experiments. So far, two fragment-inspired DCLs were designed and subjected to the pt-DCC experiment producing three positive hits; additional derivatives and negatives were synthesized for testing, for a total of eight compounds synthesized. Preliminary SPR data reported an enhancement in binding affinity compared to the single fragments and pending the setup of the STD NMR procedure on the newly selected construct (ongoing), the compounds will be subjected to the NMR–*in silico* coupled study, so far systematically applied throughout the thesis to elucidate the structural features of their interaction with HuR and guide the design of new HuR-ligands as well as building new DCLs.

As mentioned at the beginning of the present discussion, for this thesis, we decided to apply a strategy focused on developing compounds with affinity for the ELAV subtype HuR by targeting a site on the large protein–RNA interface and progressively assess and upgrade the NMR–*in silico*-built binding site model. Nonetheless, we are very well aware that the relevance of the compounds designed and synthesized in the present project in the interference/displacement of HuR–RNA complexes should be evaluated. Both REMSA (at single dose) and AlphaScreen (in a dose-response curve) could be employed to monitor the efficacy of the hits in displacing the protein–RNA complex using a protocol that has been already set up and reported.<sup>63</sup> Nonetheless, our aim is to simplify this procedure by obtaining the same results through one single assay based on the displacement of a fluorescent RNA probe followed by fluorescent polarization. The assay has already been described and exploited in a

high-throughput format for selected HuR constructs,<sup>66</sup> thus the in-house set up and execution should be straightforward. Additional structural studies of the displacement could be then performed only on active interferers by a STD NMR-based displacement/competition assay on HuR–small molecule–RNA. Our preliminary results on this approach have underlined the sensitivity of RNA to RNAses which can be found on any surface and buffer solution if not properly treated. Within the instrumental apparatus of NMR experiments, it can be quite hard to provide and maintain this environment over the experiment time-range, often leading to compromised results and representing a liability. To overcome this issue, we have started to develop an approach based on the exploitation of PNAs (Peptide Nucleic Acid);<sup>174,175</sup> while DNA and RNA have a deoxyribose and ribose sugar backbone, respectively, the PNA backbone is composed of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds. The various purine and pyrimidine bases are linked to the backbone by a methylene bridge and a carbonyl group, thus they are not labile to RNase action. Our preliminary study, performed with small derivatives and base units,<sup>††††</sup> have already shown promising results for the development of a STD NMR-based competition assay.<sup>‡‡‡‡</sup>

In the next steps, newly identified positive interferers of the HuR–RNA interaction will be subjected to biological studies to evaluate the modulation of HuR-binding by Ribonucleoprotein Immunoprecipitation (RIP) experiments, as well as the mRNA stability and translational efficiency of specific genes. The effects of the compounds on the stability of selected HuR target mRNAs will be evaluated by RT-qPCR in MCF7 cells in the presence of the transcriptional inhibitor Actinomycin-D, as already published.

To complete the construction of the experimental map of the HuR–small molecule interactions, several options have been evaluated and a pathway has been devised. Starting from the team expertise and the newly published DEEP-STD experiment, we have already set up and performed differential-solvent experiments against native HuR on three selected natural products (see chapter 4) as well as compound (+) **1**. By building differential maps we were able to highlight the nature of the amino acids present in the HuR binding site which are interacting with the compounds. Thanks to the features of the technique this information is gained for each interacting proton and not just for the general molecule. Moreover, since new constructs with different stability and different buffer requirements are now available to us, new experiments could be performed, including the possibility to try and exploit differential frequency DEEP-STD experiments. Finally, the very same collaboration which provided us with the new HuR constructs involves a structural biology team which performs

---

†††† Materials provided by Dr. Cauteruccio and collaborators of the University of Milano.

‡‡‡‡ Data not reported.

protein-observed NMR studies (see appendix V, contributions to the research). The procedures are currently under setup on selected compounds. The successful combinations of these approaches would overcome the still current lack of a defined small-molecule binding site on HuR and, compared to X-ray diffraction, it would elucidate the interaction modes as they occur in solution.

## Conclusion and research perspectives

The work presented in this thesis took part in a larger collaborative project, which sees several competences, techniques and “mindsets” merge, aimed at solving one of the current pharmacological and medicinal puzzles: is it possible to drug an unusual target, in this case RNA-binding proteins and their complexes with RNA, and how can we do it?

The PhD project, in particular, was carried out by focusing on the ELAV RBP family, exploiting subtype HuR as model protein. At the beginning of this work, much was unknown about HuR ligands and their interaction with HuR and its complexes with RNA, and a lot still is. Though this could appear as a limit, it also gave me the chance to apply, for the first time on this target, different ligand design approaches, structure-based and fragment-based, and to exploit different strategies for ligand discovery, such as multi-component reactions and protein-templated dynamic combinatorial chemistry.

In spite of the realization that several differences occur in research between experimental plan (expectations) and experiments (reality), all initially planned activities could be performed, leading to 12 new compounds developed to act as HuR ligands.

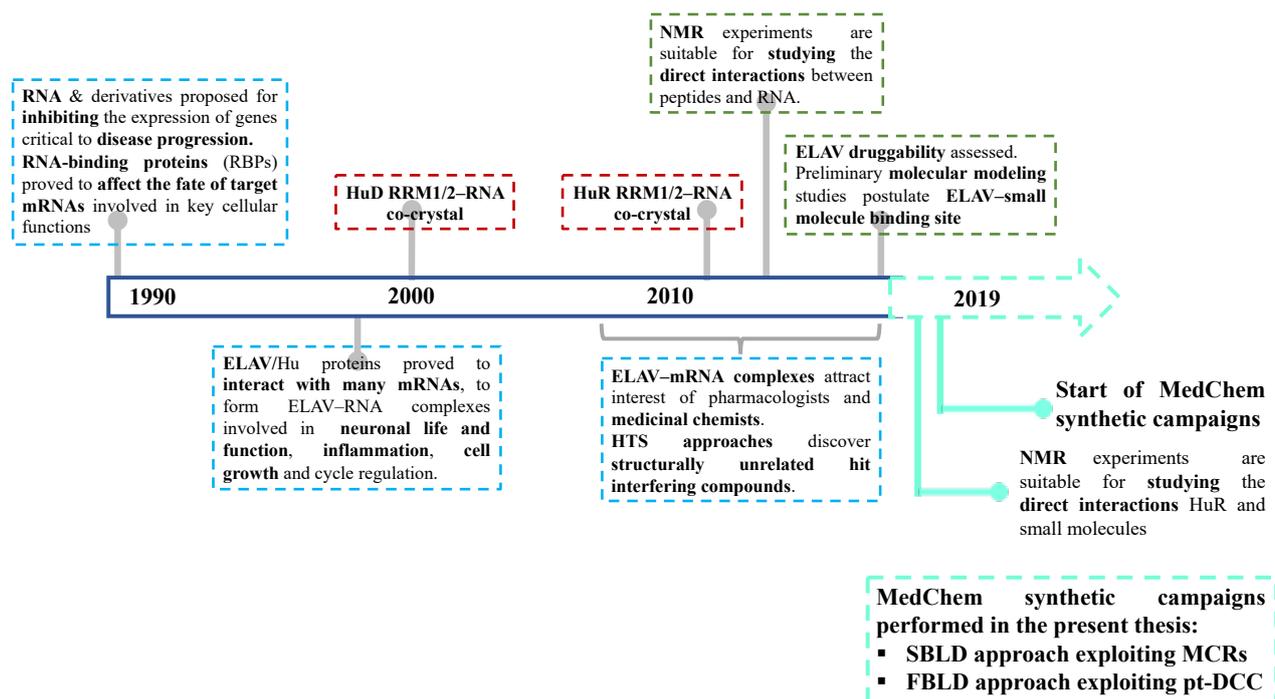
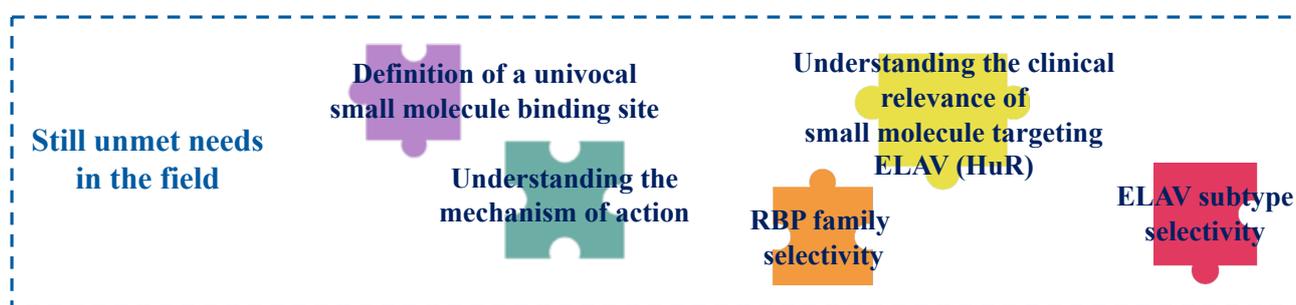


Figure 38. Updated summary timeline to recap the milestones in ELAV research at the time this thesis was finished.

The hit compounds identified and synthesized throughout the PhD project have been submitted for further biological and biophysical characterization against the protein and protein-RNA complex

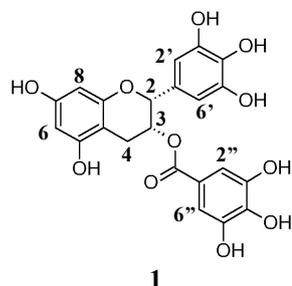
target, and the significance of the obtained small molecules in the modulation of gene expression, for treating clinically relevant diseases in which HuR–RNA complexes are involved, will be evaluated. To our knowledge, only one additional MedChem synthetic campaign was reported for HuR, following a different approach and gaining different results (see Manzoni *et al.*, 2018<sup>164</sup>). Nonetheless, several questions remain unanswered in this study field. In particular, the definition of a small molecule binding site as well and the understanding of the relevance of small molecules inhibitors of the ELAV–RNA interaction, are the most pressing. Though this is the conclusion of the present thesis, the data gathered (and those still being processed) will be exploited within our group and its research network for following up with this study, aimed at adding more pieces to solve the ELAV puzzle.



## Experimental Section 1

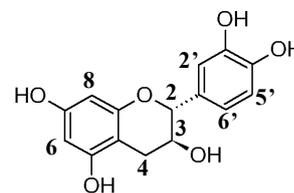
### Chapter 4: Exploration of HuR–ligand binding modes

#### NMR spectra assignment in D<sub>2</sub>O for all compounds (283K)



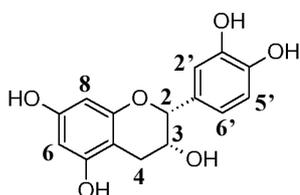
1

<sup>1</sup> H	C	
6.73	109.70	2''-6''
6.34	106.27	2'-6'
5.89	95.44	6-8
5.36	68.50	3
4.80	76.69	2
2.80		
2.68	24.40	4



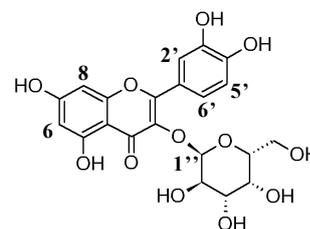
3

<sup>1</sup> H	C	
6.67	114.66	2'
	116.66	5'
6.60	119.69	6'
5.84	95.43	8
5.75	94.54	6
4.55	80.50	2
3.96	66.17	3
2.62		
2.27	25.56	4



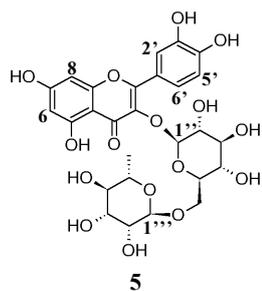
2

<sup>1</sup> H	C	
6.78	114.51	6'
	119.05	2'
6.68	116.21	5'
5.85	95.53	8
5.83	95.84	6
4.74	78.70	2
4.06	66.20	3
2.67		
2.52	27.60	4

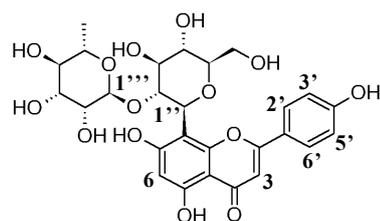


4

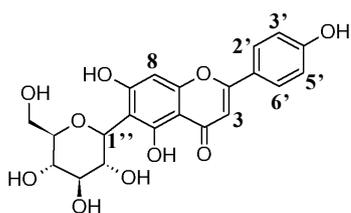
<sup>1</sup> H	C	
7.39	116.51	6'
7.28	122.47	2'
6.30	115.85	5'
6.03	95.84	8
5.85	101.40	6
4.63	102.77	1''
3.60	67.76	4''
3.51	71.00	2''
3.33	72.18	3''
3.32	59.80	6''
3.23	75.13	5''



<sup>1</sup> H	C	
7.33	116.5	2'
7.24	122.76	6'
6.65	115.38	5'
6.02	95.53	8
5.85	101.60	6
4.68	102.09	1''
4.28	100.62	1'''
3.51	67.57	6''
3.41	69.92	2'''
3.29	73.53	2''
3.25	69.92	3'''
3.20	75.60	3''
3.14	68.89	5''
3.12	74.73	4'''
3.06	69.74	4''
3.02	71.81	5'''
0.82	16.52	6'''

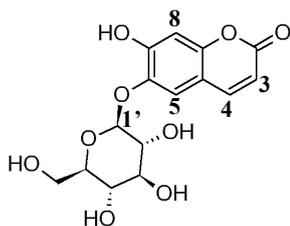


<sup>1</sup> H	C	
7.69	128.69	2'-6'
6.75	115.70	3'-5'
6.32	102.22	3
5.89	101.00	6
4.81	72.11	1''
4.75	100.22	1'''
4.00	76.20	2''
3.66	69.57	2'''
3.51	71.20	5''
3.48	79.23	3''
3.39	61.81	6''
3.29	79.89	4''
3.24	69.09	3'''
2.86	71.00	4'''
2.02	68.44	5'''
0.40	16.53	6'''



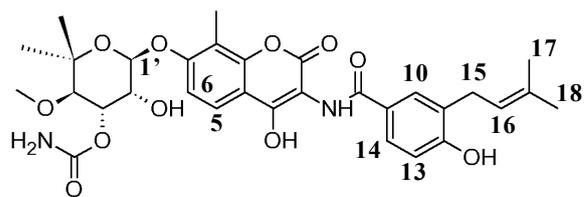
7

<sup>1</sup> H	C	
7.54	128.47	2'-6'
6.64	116.43	3'-5'
6.23	102.20	3
6.10	96.90	6
4.61	73.74	1''
4.16	70.16	2''
3.64		
3.54	60.83	6''
3.32	69.76	3''
3.30	78.10	4''
3.28	80.48	5''



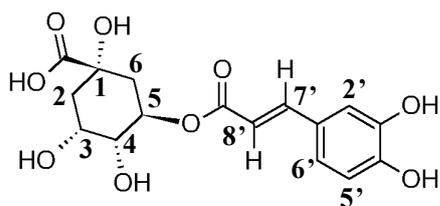
8

<sup>1</sup> H	C	
7.61	146.50	4
6.98	113.22	5
6.42	104.38	8
5.88	107.49	3
4.77	101.4	1'
3.70		
3.52	60.44	6'
3.40	75.71	2'
3.38	73.20	4'
3.35	72.88	5'
3.25	69.19	3'



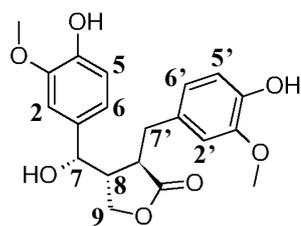
9

<sup>1</sup> H	C	
7.50	121.0	10
7.47	nd	5
7.44	nd	13
6.89	108.8	6
6.70	114.6	14
5.43	96.1	1
5.12	nd	16
5.09	70.5	4'
4.07	68.0	2'
3.35	80.7	3'
3.31	60.3	OCH <sub>3</sub>
3.08	27.4	15
2.03	nd	CH <sub>3</sub>
1.47	24.5	17-18
1.17	26.6	6'a
0.95	22.2	6'b



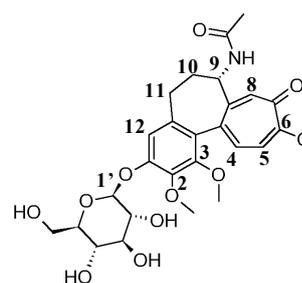
**10**

<sup>1</sup> H	C	
7.40	146.10	7'
6.93	114.81	2'
6.85	122.49	6'
6.66	116.28	5'
6.13	114.82	8'
5.06	70.90	5
3.97	70.32	3
3.61	72.51	2
1.93	37.70	6
1.80		
1.85	37.19	4
1.75		



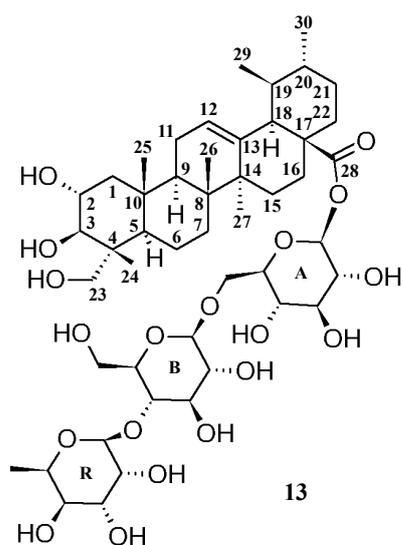
**11**

<sup>1</sup> H	C	
6.47	115.01	5
6.41	115.30	5'
6.37	117.74	6
6.27	109.33	2
6.23	122.03	2'
6.20	112.23	6'
4.54	73.72	
4.23	71.3	9
4.14		
3.41	55.15	OCH <sub>3</sub> - OCH <sub>3</sub> '
2.64	34.94	7'
2.33		
2.50	42.35	8
2.40	44.94	8'



**12**

<sup>1</sup> H	C	
7.26	137.50	5
7.20	130.00	8
7.05	115.94	4
6.65	111.60	12
4.95	100.07	1'
4.17	53.11	
3.71	56.71	6-OCH <sub>3</sub>
3.70	60.64	6'
3.49	61.8	2-OCH <sub>3</sub>
3.68	61.8	2-OCH <sub>3</sub>
3.36-3.43	72.99-75.95	Overlap, 2',3',4'
3.34	61.62	3-OCH <sub>3</sub>
3.26	69.51	5'
2.32	28.72	10
1.68		
2.00	35.29	11
1.93		
1.77	21.90	9- NHCOC <sub>3</sub> H <sub>7</sub>



**-CH:**

<sup>1</sup> H	C	
5.32	94.04	A1
5.22	125.52	12
4.79	101.00	R1
4.47	101.97	B1
3.95	69.13	R5
3.89	70.41	R2
3.74	68.62	2
3.66	70.24	R3
3.63	76.48	A3
3.48	77.28	<i>n.d.</i>
3.47	75.69	<i>n.d.</i>
3.48	74.73	<i>n.d.</i>
3.43	68.95	A
3.40	71.87	R4
3.34	75.17	<i>n.d.</i>
3.30	76.66	<i>n.d.</i>
3.21	73.44	B2
2.09	52.64	18
1.57	47.07	9
1.36	38.56	20
1.14	46.70	5
0.90	38.60	19

**-CH<sub>2</sub>:**

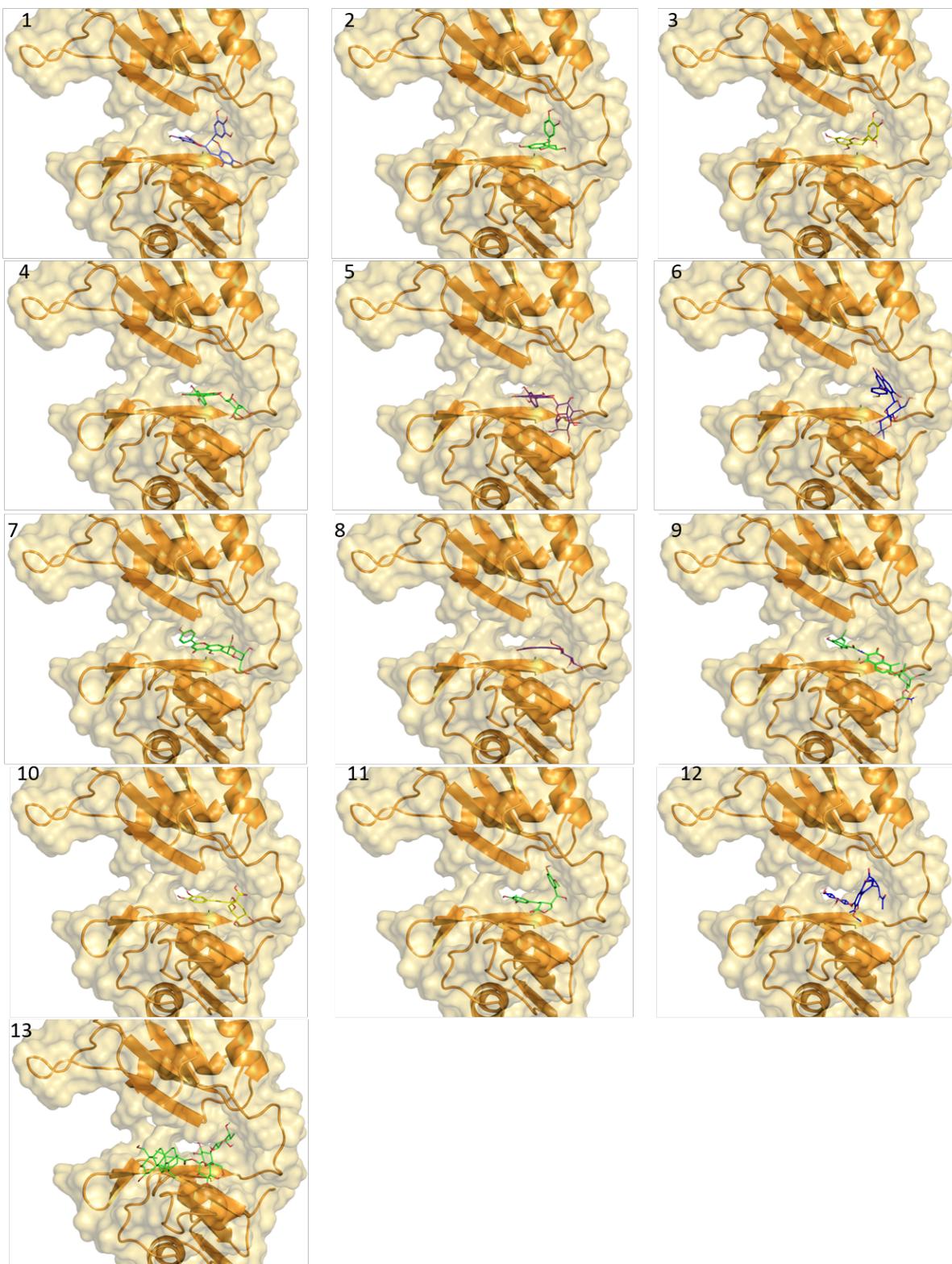
<sup>1</sup> H	C	
4.02		
3.85	67.57	A6
3.80		
3.64	60.00	B6
3.40		
3.30	64.16	23
2.09		
1.64	23.90	16
1.93		
0.89	46.10	1
1.91	23.00	11
1.68		
1.47	36.50	22
1.63		
1.07	27.80	15
1.48		
1.20	32.10	7
1.47		
1.31	30.00	21
1.33	17.70	6

**-CH<sub>3</sub>:**

<sup>1</sup> H	C	
1.20	16.59	R6
1.06	22.89	27
0.98	16.63	25
0.90	20.46	30
0.83	16.59	29
0.70	16.87	26
0.62	12.84	24

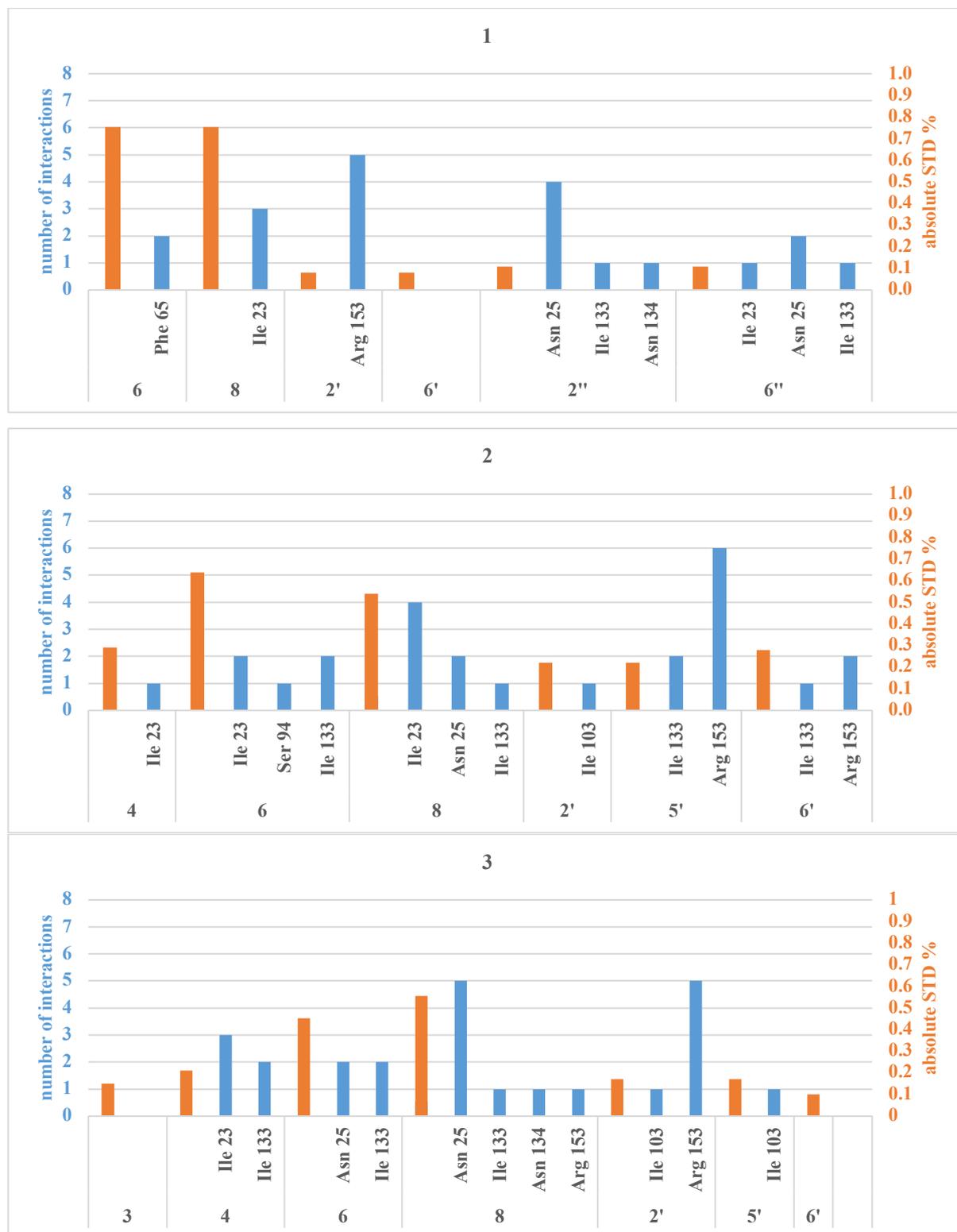
### ***Predicted binding modes***

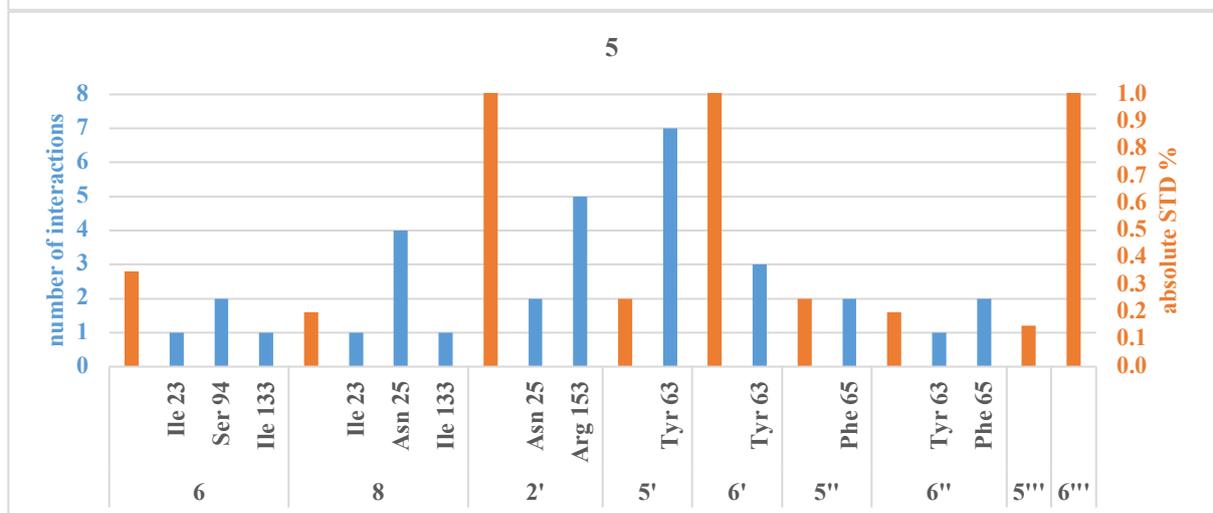
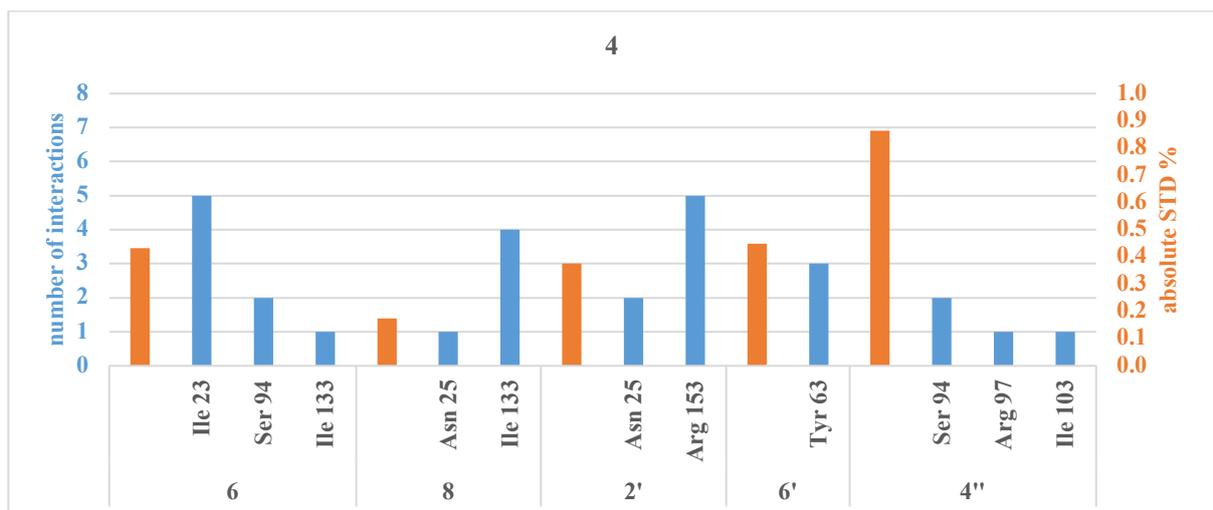
3D representation of each compound in the HuR interaction site. HuR is shown as orange cartoon.



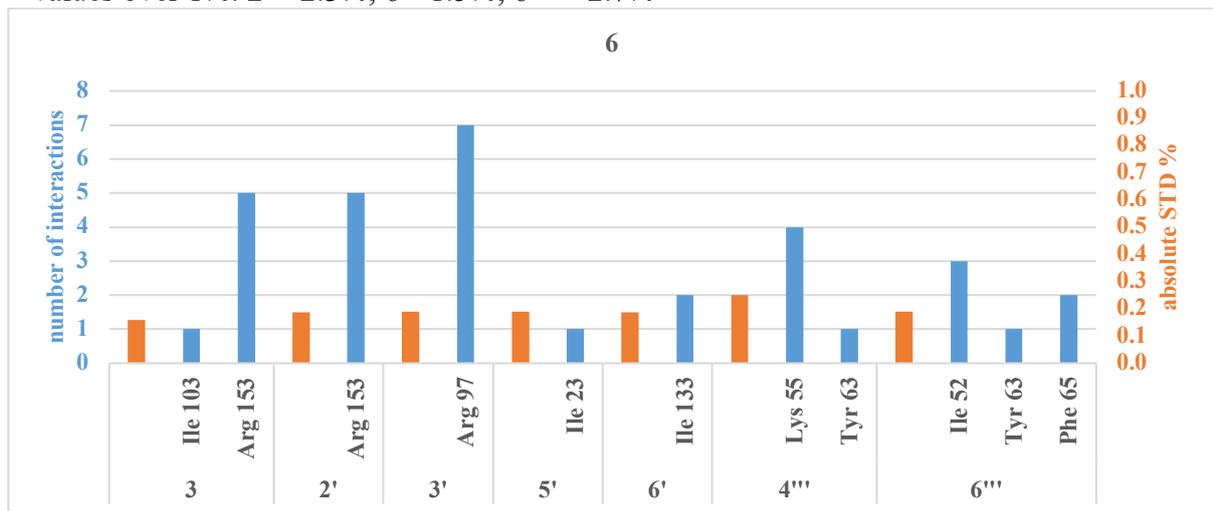
### Combined STD NMR-Docking histograms for all interacting compounds

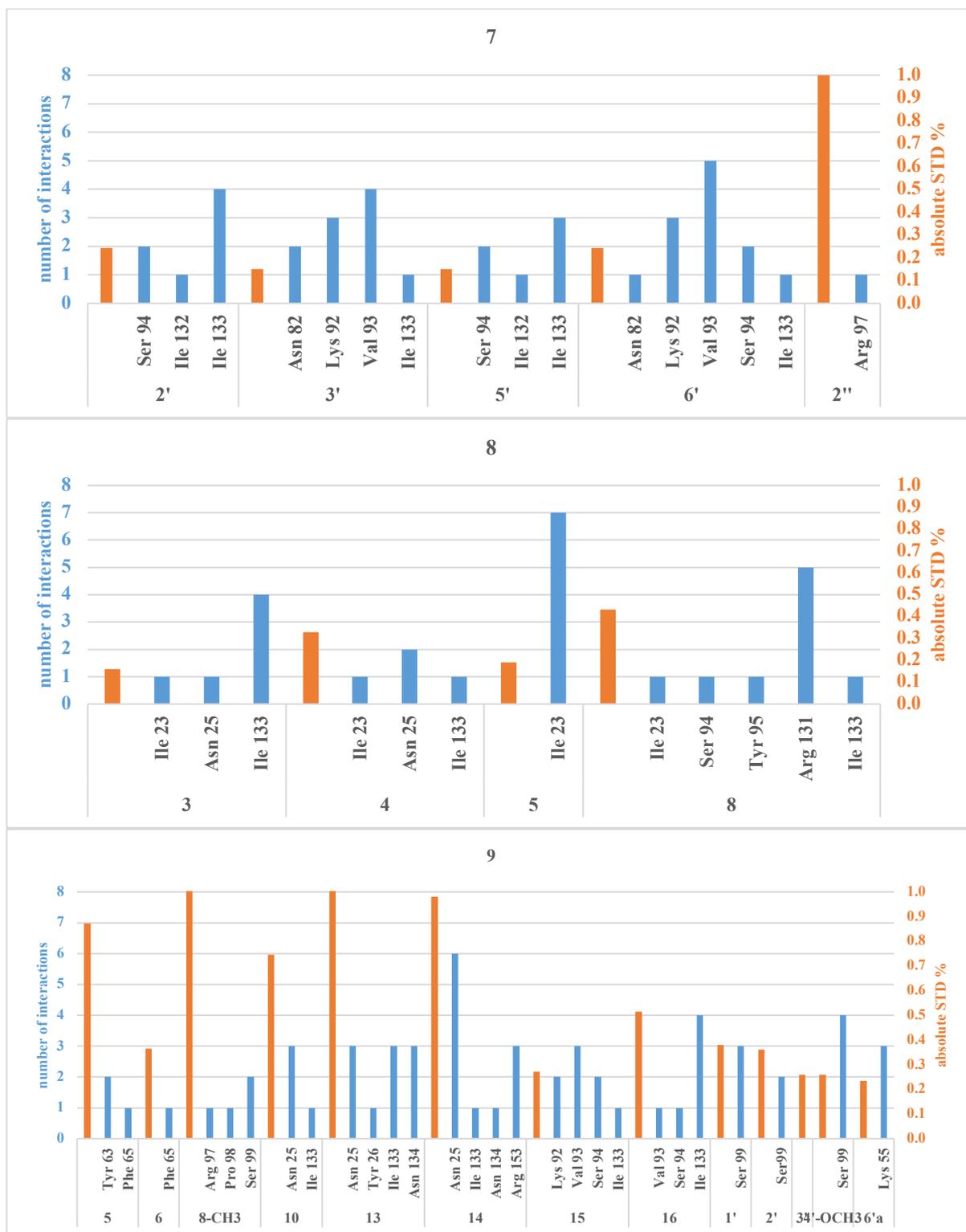
The following histograms plot the number of interactions and related HuR residue (as seen through molecular modeling) and absolute STD% for each interacting proton (as determined by STD NMR).



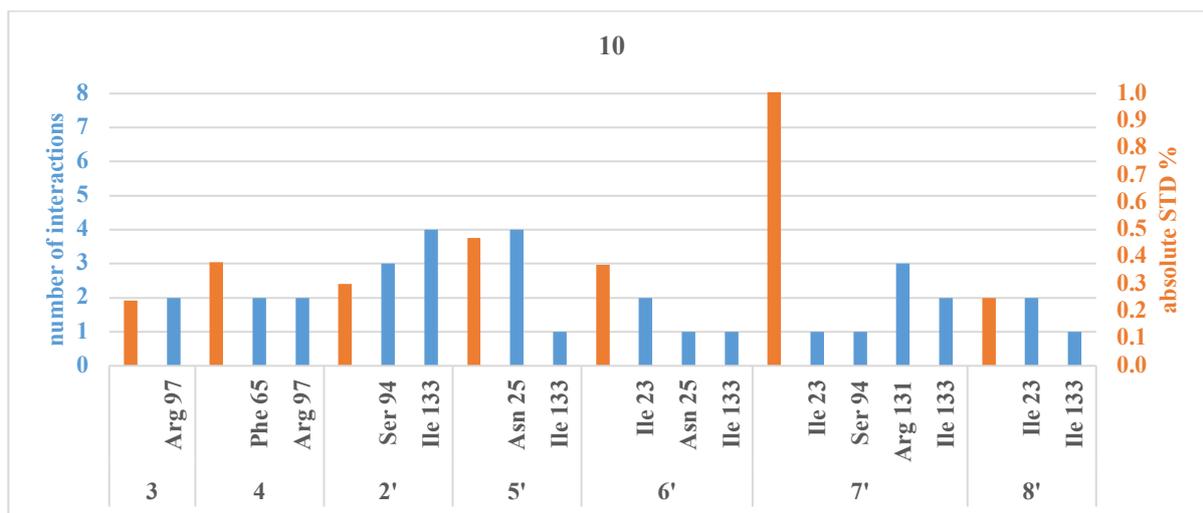


\* values over 1%: 2'= 2.3%; 6'=1.5%; 6'''= 2.7%

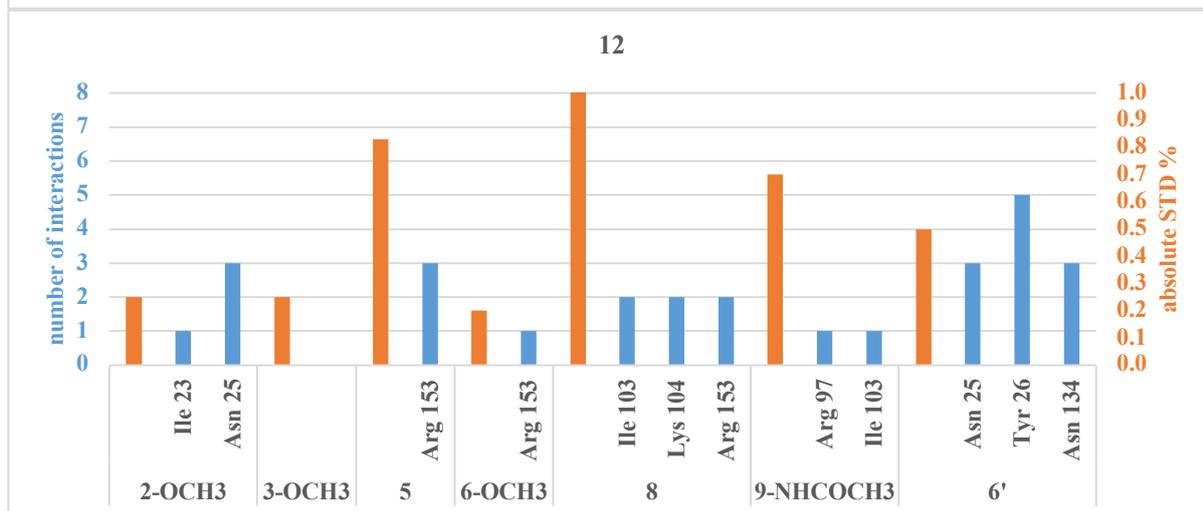
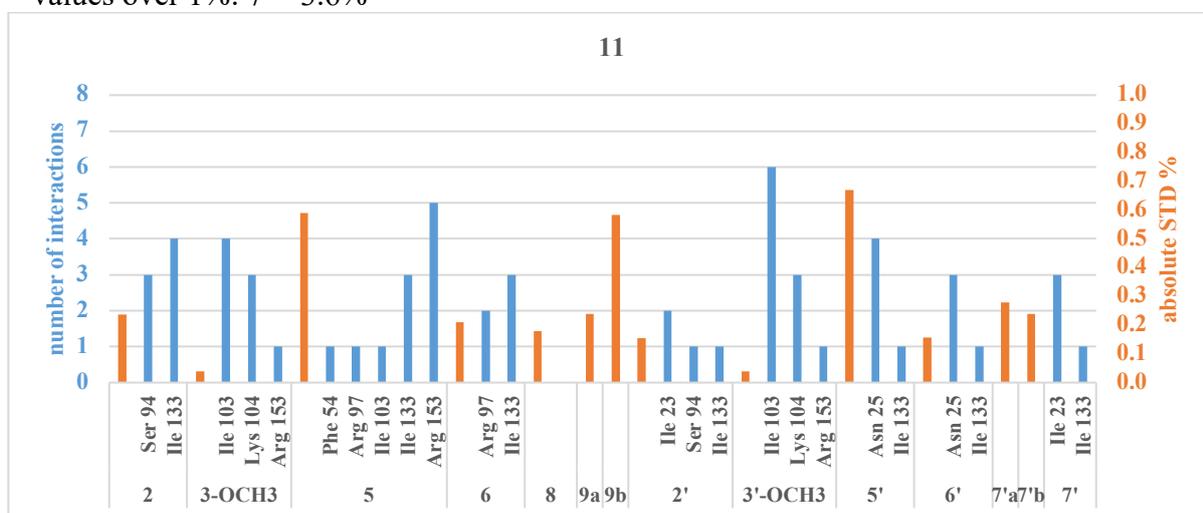




\* values over 1%: 8-CH<sub>3</sub>= 1.1%; 13=1.2%



\* values over 1%: 7'= 3.6%

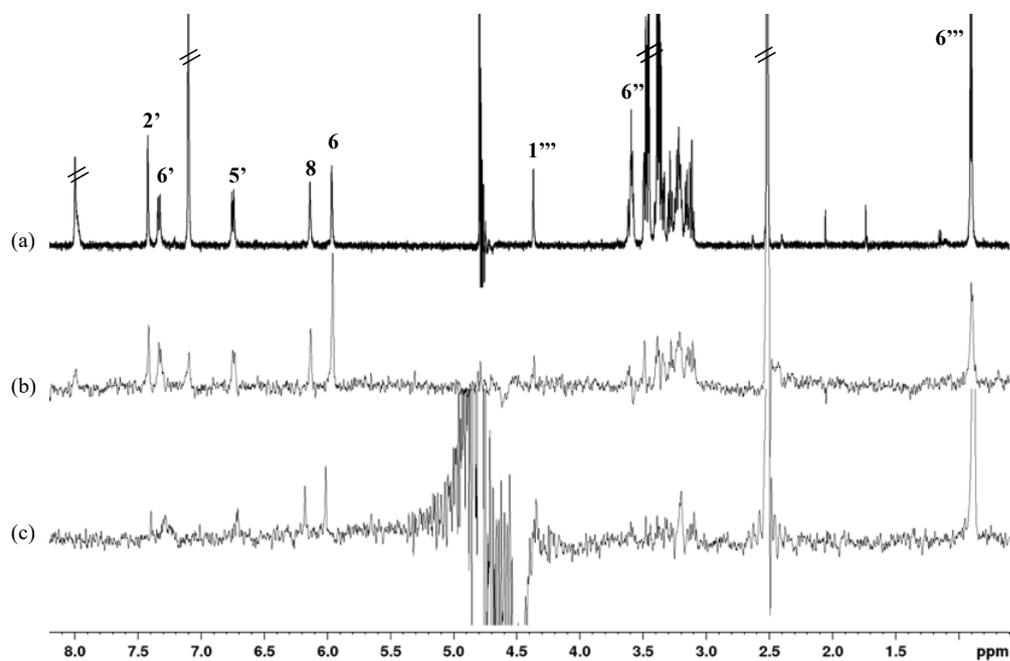


\* values over 1%: 8= 1.4%

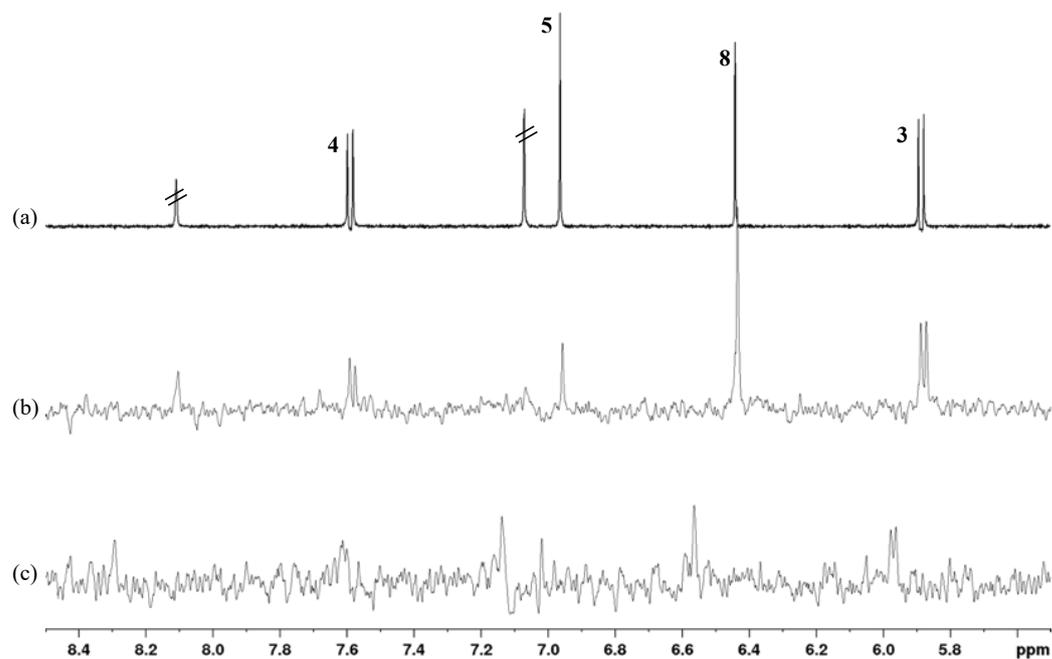
***STD NMR spectra of selected natural compounds acquired with differential solvent conditions (DEEP-STD experiment)***

Reference spectra; (b) STD spectrum in D<sub>2</sub>O; (c) STD spectrum in H<sub>2</sub>O:D<sub>2</sub>O 90:10.

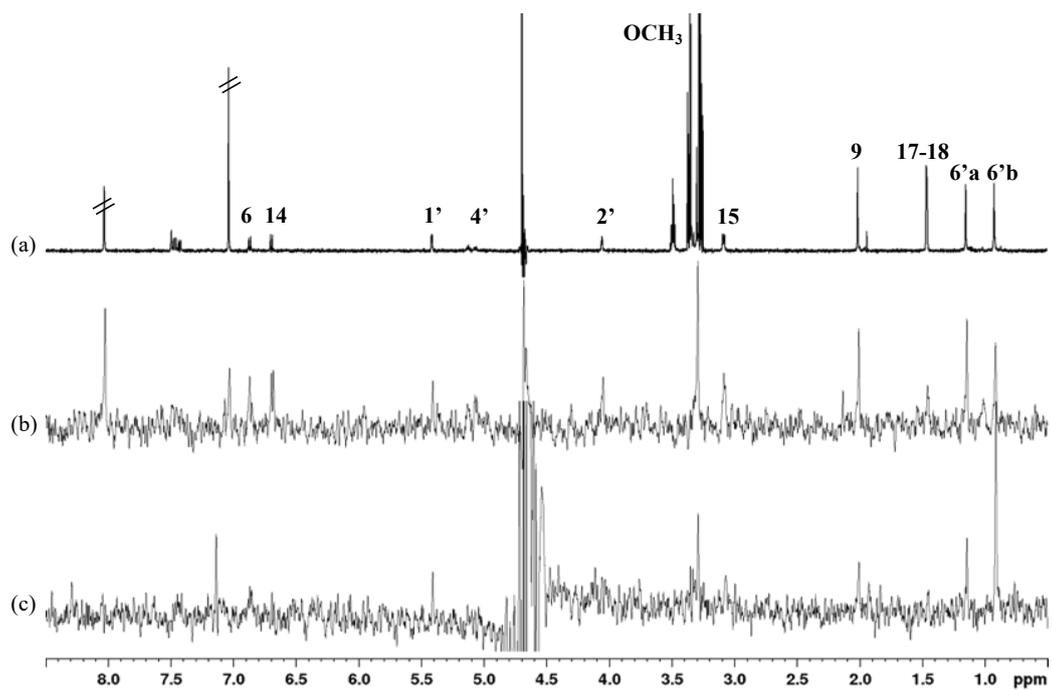
- DEEP-STD spectra of **5**



- DEEP-STD spectra of **8**



▪ DEEP-STD spectra of **9**



## ***Chapter 5: Structure-based approach to obtain new HuR ligands***

### ***Design of the new scaffolds, docking and optimization: methods, structures of the originally designed compounds, and docking poses of selected examples***

#### **Methods**

Protein HuR (PDB code: 4ED5) and uracil residues 8 or 9 (anchors U8-U9, PDB code: 4ED5) were individually uploaded into AnchorQuery. The anchor was modified by PyMol so as to bear an additional aromatic group to allow a  $\pi$ - $\pi$  interaction with Tyr63, not naturally present between RNA and HuR. The pharmacophore features were edited maintaining the aromatic character of the uracil ring as well as the hydrogen-bond acceptor features of the carboxylic moiety.

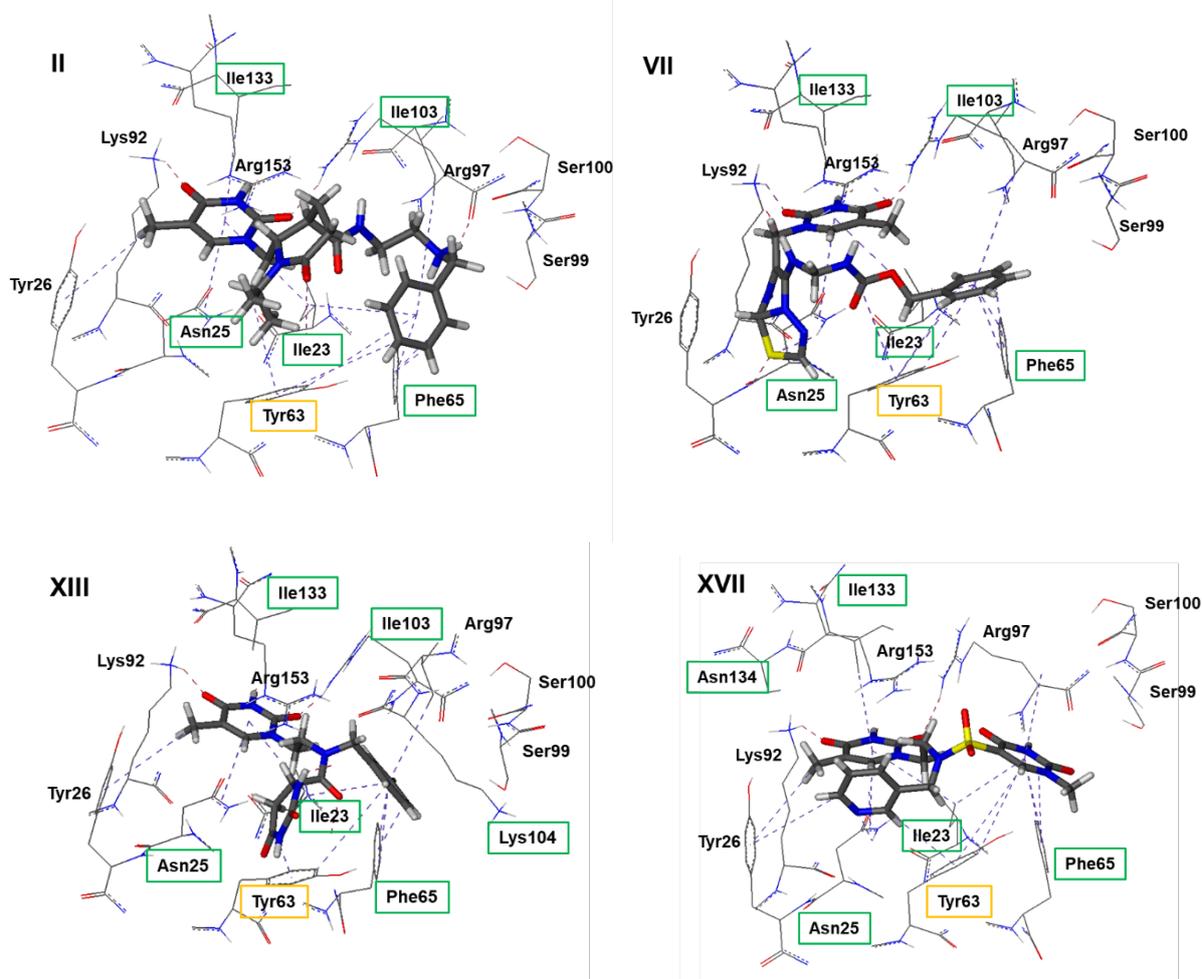
The query characteristics were selected under the filters tab. The maximum total hits value was set to 200 and ranking was done according to molecular weight (lower to higher). In our case, a total of four queries were submitted.

The resulting hit compounds were downloaded and their energy optimized by MOLOC. Further visual inspection allowed us to evaluate intermolecular clashes with the protein. The final selection of 17 molecules was carried out through a docking study performed by LeadIT and SeeSAR by generating and analyzing 30 docking poses for each molecule.



## Docking poses of selected exemplary compounds

Selected docking pose of one compound per scaffold class (II, VII, XII, XVII). Compounds are shown as gray sticks, HuR residues as gray lines. Green rectangles show additional interactions compared to those established by U8–U9, and the yellow rectangle highlights Tyr63, which new compounds are designed to target.



## Synthesis: general methods and experimental procedures

### General methods

Reagents and solvents for synthesis were purchased from Sigma-Aldrich (Italy, The Netherlands), Acros, Alpha Aesar or Apollo Scientific. Unless otherwise specified, commercially available reagents were used as received from the supplier. Solvents were purified according to the guidelines in Purification of Laboratory Chemicals. Microwave dielectric heating was performed in a Discover® LabMate instrument (CEM Corporation) specifically designed for organic synthesis and following an appropriate microwave program. Melting points were measured on SMP3 Stuart Scientific apparatus. Analytical thin-layer-chromatography (TLC) was carried out on silica gel precoated glass-backed plates (Fluka

Kieselgel 60 F254, Merck) and visualized by ultra-violet (UV) radiation, acidic ammonium molybdate (IV), or potassium permanganate. Flash chromatography (FC) was performed on Silica Gel 60 (particle size 230–400 mesh, purchased from Nova Chimica). IR spectra were recorded on a Jasco FT/IR-4100 spectrophotometer with ATR module; only noteworthy absorptions are reported. Unless otherwise specified,  $^1\text{H}$  and  $^{13}\text{C}$  nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance 400 MHz spectrometer and Bruker Avance 500 MHz spectrometer at 25°C.

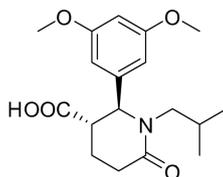
Proton chemical shifts ( $\delta$ ) are reported in ppm with the solvent reference relative to tetramethylsilane (TMS) employed as internal standard ( $\text{CDCl}_3$ ,  $\delta = 7.26$  ppm;  $\text{CD}_2\text{Cl}_2$ ,  $\delta = 5.32$  ppm; acetone- $d_6$ ,  $\delta = 2.05$  ppm). The coupling constant values are reported in Hz.  $^{13}\text{C}$ -NMR spectra were recorded with complete proton decoupling. Carbon chemical shifts ( $\delta$ ) are reported in ppm relative to TMS with the respective solvent resonance as internal standard ( $\text{CDCl}_3$ ,  $\delta = 77.23$  ppm;  $\text{CD}_2\text{Cl}_2$ ,  $\delta = 54.00$  ppm; acetone- $d_6$ ,  $\delta = 29.84$  ppm).

The compounds purity was determined by Ultra High Pressure Liquid Chromatography (UHPLC) coupled with UV-ESI/MS. The analyses were carried out on an Acquity UPLC Waters LCQ FLEET system using an ESI source operating in positive ion mode, controlled by ACQUITY PDA and 4 MICRO (Waters). Analyses were run on an Acquity BEH Shield RP18 (50 x 2,1 mm; 1,7  $\mu\text{m}$ ) column, at room temperature, with gradient elution (solvent A: water containing 0.1% of formic acid; solvent B: methanol containing 0.1% of formic acid; gradient: 10% B in A to 100% B in 3 minutes, followed by isocratic elution 100% B for 1.5 minutes, return to the initial conditions in 0.2 minutes) at a flow rate of 0.5 mL/min. All final compounds had a purity suitable for our purposes.

Exact mass analyses (HRMS) were performed using a Thermo Scientific LTQ Orbitrap-XL mass spectrometer.

## Experimental procedures

### 2-(3,5-Dimethoxyphenyl)-1-isobutyl-6-oxopiperidine-3-carboxylic acid (9)

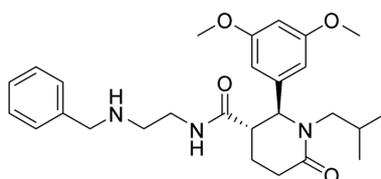


To a solution of aldehyde **6** (100 mg, 0.60 mmol) in acetonitrile (2 mL) at room temperature (rt), isobutylamine **7** (60  $\mu\text{L}$ , 0.60 mmol) and two drops of formic acid were sequentially added. The reaction mixture was irradiated with a microwave power of 140 W at 120 °C for 30 minutes. After that, the yellow solution was immediately evaporated to dryness, dissolved in *p*-xylene

(2 mL) and then glutaric anhydride **8** was added (54 mg, 0.60 mmol). The suspension was refluxed overnight under stirring. The solvent was decanted and the remaining brown solid was washed with ether affording a beige solid. The crude was further purified by flash chromatography eluting with DCM/MeOH (9:1, v/v), giving (±)-**9** as a white solid (125 mg, 62%).

Mp=176–179 °C. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.76 (s, 1H), 6.31 (t, *J* = 2.2 Hz, 1H), 6.24 (d, *J* = 2.2 Hz, 2H), 5.02 (d, *J* = 2.9 Hz, 1H), 3.88 (dd, *J* = 13.6, 8.8 Hz, 1H), 3.71 (s, 6H), 2.80 (dd, *J* = 7.5, 4.2 Hz, 1H), 2.64 (ddd, *J* = 17.7, 10.5, 6.9 Hz, 1H), 2.48 (ddd, *J* = 18.1, 6.0, 3.4 Hz, 1H), 2.19 (dd, *J* = 13.5, 6.1 Hz, 1H), 1.96 (t, *J* = 22.9 Hz, 1H), 1.85 (d, *J* = 35.1 Hz, 1H), 0.80 (d, *J* = 6.7 Hz, 3H), 0.76 (d, *J* = 6.6 Hz, 3H).

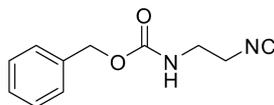
*N*-(2-(Benzylamino)ethyl)-2-(3,5-dimethoxyphenyl)-1-isobutyl-6-oxopiperidine-3-carboxamide (1)



To a solution of (±)-**9** (100 mg, 0.30 mmol) dissolved in THF (3 mL), TBTU (115.3 mg, 0.36 mmol) and DIPEA (0.109 mL, 0.62 mmol) were subsequently added. The reaction mixture was stirred at room temperature for 30 minutes. *N*-benzylethyldiamine (**10**) (0.045 mL, 0.30 mmol) was added and the solution was stirred overnight at room temperature. The reaction mixture was concentrated *in vacuo*, dissolved in DCM (10 mL) and washed with water (20 mL). The organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and the solvent evaporated *in vacuo*. The crude was purified by FC eluting with DCM/MeOH (8:2, v/v), giving **1** as yellow oil (99.6 mg, 71%)

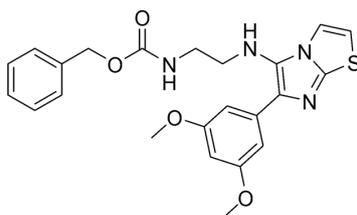
<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 7.25 (ddd, *J* = 7.4, 4.5, 1.5 Hz, 2H), 7.22-7.15 (m, 3H), 6.30-6.27 (m, 1H), 6.25 (d, *J* = 2.2 Hz, 2H), 5.96 (s, 1H), 4.75 (d, *J* = 7.5 Hz, 1H), 3.87-3.80 (m, 1H), 3.67 (d, *J* = 3.0 Hz, 6H), 3.62 (d, *J* = 4.5 Hz, 2H), 3.23 (ddd, *J* = 16.9, 11.9, 5.1 Hz, 1H), 3.18-3.08 (m, 1H), 2.61 (dq, *J* = 8.7, 4.9 Hz, 1H), 2.57-2.49 (m, 2H), 2.46–2.34 (m, 2H), 2.15 (dd, *J* = 13.6, 6.2 Hz, 1H), 2.02-1.83 (m, 4H), 0.78 (d, *J* = 5.4 Hz, 3H), 0.76 (s, 3H). <sup>13</sup>C-NMR (101 MHz, CDCl<sub>3</sub>): δ 175.26, 171.13, 161.21, 142.42, 104.53, 99.23, 61.68, 55.31, 52.91, 45.75, 28.78, 26.41, 20.18, 19.97, 18.50. IR (ν<sub>max</sub>/cm<sup>-1</sup>): 836.9, 910.4, 991, 1061.62, 1156.12, 1203.36, 1607.38, 1746.23, 3002.62, 3027.69, 3381.57. UHPLC-ESI-MS: ABS t<sub>R</sub> = 1.55 min, 96% pure (λ = 210 nm), *m/z* = 468.26 [*M* + H]<sup>+</sup>.

Benzyl-(isocyanoethyl)-carbamate (12)



The synthesis was carried out according to the reported procedure. Benzyl-2-(aminomethyl) carbamate (**11**) (500 mg, 2.57 mmol) was dissolved in ethyl formiate and the suspension was refluxed overnight. The reaction mixture was concentrated *in vacuo* giving the corresponding *N*-formylcarbamate as a white solid without further purification. Subsequently, to a solution of the carbamate thus prepared (530 mg, 2.38 mmol) and TEA (2 mL, 14.35 mmol) in DCM (10 mL) cooled to 0°C under stirring, POCl<sub>3</sub> (0.25 mL, 2.68 mmol) was dropwise added keeping the internal reaction temperature below 5°C. After the POCl<sub>3</sub> addition was complete, the reaction was left at rt for 4 hours. The change in color for the reaction mixture, from yellow to dark brown, was indicative of the end of the reaction (further confirmed by TLC). The reaction was thus purged in a 5% NaHCO<sub>3</sub> aq solution and icy-water, then extracted with DCM (40 mL) and the organic layer washed with a 5% NaHCO<sub>3</sub> aq solution. The collected organic layers were dried on dry MgSO<sub>4</sub>, filtrated and the solvent evaporated *in vacuo*. The crude was re-dissolved in DCM and filtered on silica pad. The filtrate was evaporated once more *in vacuo* yielding isocyanide **12** as a brown oil (271 mg, 52 %). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 7.43-7.28 (m, 5H), 5.35 (s, 1H), 5.14 (s, 2H), 3.52 (d, *J* = 10.6 Hz, 2H), 3.47-3.36 (m, 2H).

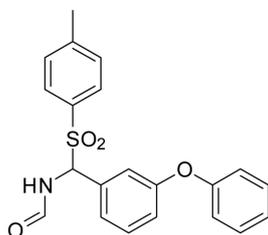
Benzyl-2-((6-(3,5-dimethoxyphenyl)-imidazo[2,1-*b*]thiazol-5-yl)amino)ethyl) carbamate (2)



Aldehyde **6** (100 mg, 0.60 mmol) and amine **13** (49 mg, 0.60 mmol) were dissolved in ACN (2 mL) in a microwave vessel and added two drops of formic acid. The reaction mixture was irradiated with a microwave power of 200 W, at 120 °C for 30 minutes. Then, zirconium chloride (10 mg, 0.07 mmol) and isocyanide **12** (120 mg, 0.60 mmol) were sequentially added. The reaction mixture was then quickly warmed in an oil bath at 70-80°C and stirred overnight. Successively, the reaction was concentrated *in vacuo*, dissolved in DCM (25 mL) and washed with a HCl aq solution (1 M). The aqueous layer was basified with a NaOH aq solution (1 M) and then extracted with EtOAc (40 mL). The organic layers were collected, dried over Na<sub>2</sub>SO<sub>4</sub>

and then filtered; the solvent was evaporated *in vacuo*. The crude was purified by FC eluting with DCM/MeOH (97:3, v/v) affording **2** as a yellow oil (57.5 mg, 21 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.42 (d, *J* = 2.3 Hz, 1H), 7.27 (s, 5H), 6.97 (d, *J* = 2.2 Hz, 2H), 6.67 (d, *J* = 4.4 Hz, 1H), 6.31 (t, *J* = 2.1 Hz, 1H), 5.01 (s, 2H), 3.73 (s, 6H), 3.47 (d, *J* = 5.2 Hz, 1H), 3.21 (dd, *J* = 11.1, 5.6 Hz, 2H), 3.14-3.07 (m, 2H). <sup>13</sup>C-NMR (101 MHz, CDCl<sub>3</sub>): δ 161.03, 156.70, 144.90, 136.43, 136.19, 135.70, 128.55, 128.18, 128.09, 116.83, 112.66, 108.58, 104.30, 99.56, 66.99, 55.43, 48.60, 41.23. IR (ν<sub>max</sub>/cm<sup>-1</sup>): 926.62, 992.19, 1204.33, 1295.93, 1318.11, 1510.95, 1596.77, 1738.51, 2837.74, 2935.13, 3001.66, 3181.01, 3333.36. UHPLC-ESI-MS: ABS t<sub>R</sub> = 1.93 min, 92% pure (λ = 210 nm), *m/z* = 453.12 [*M* + H]<sup>+</sup>.

*N*-((3-Phenoxyphenyl)(tosyl)methyl)formamide (17p)

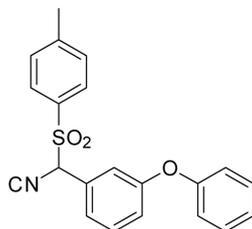


A solution of sodium *p*-toluenesulfinate (4.45 g, 25 mmol) in water (30 mL) was stirred for 30 minutes until a clear solution was obtained. Methyl tert-butyl ether (MTBE, 30 mL) was added, followed by the slow addition of HCl (37%, 2.1 mL, 25 mmol). The reaction mixture was stirred for an additional 30 minutes, and the aqueous layer was removed. The organic layer was diluted with toluene (30 mL) and concentrated under reduced pressure until approximately 80% of the solvent was removed. Heptane (10 mL) was added, and the precipitate white solid was collected by filtration, rinsed with heptane (10 mL) and dried *in vacuo* for 4 hours to give *p*-toluenesulfonic acid (**14**) (2.46 g, 63% yield).

A solution of 3-phenoxybenzaldehyde (**15**) (1.70 mL, 9.83 mmol), formamide (**16**) (1 mL, 25 mmol) and chlorotrimethylsilane (1.40 mL, 11.00 mmol), in acetonitrile (5 mL) and toluene (5 mL), was stirred under a nitrogen atmosphere at 50 °C for 5 hours. Then, *p*-toluenesulfonic acid (**14**) (2.46 g, 15.80 mmol) was added, and the reaction mixture was stirred at 50 °C for 16 hours. The reaction was cooled to rt, and then MTBE (5 mL) was added. After 5 minutes stirring, water (30 mL) was added. The collected organic phases were dried over dry MgSO<sub>4</sub>, filtered and concentrated *in vacuo*, affording the crude. MTBE (10 mL) was added, and the resulting mixture was cooled to 0 °C for 1 hour, and the precipitated white solid was collected by filtration, rinsed with MTBE (2 x 5 mL) and dried *in vacuo* at 60 °C for 5 hours. Compound **17p** was obtained as a white solid (2.33 g, 62% yield) and used in the next step without further

purification.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.12 (s, 1H), 7.71 (d,  $J = 7.8$ , 2H), 7.39-7.29 (m, 5H), 7.15 (d,  $J = 7.8$ , 2H), 7.07-6.95 (m, 4H), 6.25 (d,  $J = 10.5$ , 1H), 2.44 (s, 3H), 1.58 (s, 1H).

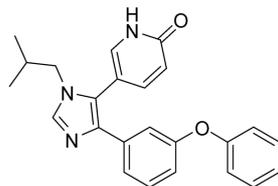
1-(Isocyano(tosyl)methyl)-3-phenoxybenzene (17)



A solution of intermediate 17p (2.00 mg, 5.25 mmol) and phosphorus oxychloride (1.1 mL, 10.74 mmol) in THF (20 mL) was stirred at rt for 5 minutes. After cooling the reaction mixture to  $0^\circ\text{C}$ , triethylamine (4.4 mL, 31.50 mmol) was added dropwise over 45 minutes whilst keeping the internal reaction temperature below  $10^\circ\text{C}$ . Afterwards, the reaction was warmed to  $10^\circ\text{C}$  for 45 minutes. EtOAc (15 mL) and water (15 mL) were added, and the reaction mixture was stirred at rt for 5 minutes. The organic layer was washed with water (2 x 15 mL), saturated  $\text{NaHCO}_3$  solution (15 mL) and a saturated  $\text{NaCl}$  solution (10 mL). The collected organic phases were dried over dry  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford compound 17 as a brown oil (1.62 g, 85%), which was used in the next step without further purification.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.64 (d,  $J = 8.2$ , 2H), 7.38–7.33 (m, 5H), 7.18–7.06 (m, 4H), 6.98–6.96 (m, 2H), 5.59 (s, 1H), 2.46 (s, 3H).

5-(1-Isobutyl-4-(3-phenoxyphenyl)-1H-imidazol-5-yl)106mylose106-2(1H)-one (3)

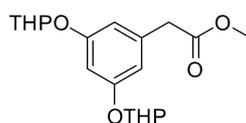


A solution of isobutyl amine (7) (0.33 mL, 3.30 mmol) and 6-oxo-1,6-dihydropyridine-3-carbaldehyde (18) (440 mg, 3.58 mmol) in anhydrous DMF (10 mL) was stirred under a nitrogen atmosphere at rt for 4 hours. Then, a solution of intermediate 17 (1.00 g, 2.75 mmol) in anhydrous DMF (5 mL) and  $\text{K}_2\text{CO}_3$  (760 mg, 5.50 mmol) were added. The reaction mixture was stirred at rt for 72 hours. EtOAc (20 mL) and water (20 mL) were added, and the mixture was stirred for 10 minutes and extracted with EtOAc (3 x 10 mL). The combined organic layers were washed with water (50 mL) and a saturated  $\text{NaCl}$  solution (3 x 50 mL). The collected

organic phase was dried over dry  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The crude was purified by silica gel column chromatography (DCM/MeOH, 96:4, v/v) giving **3** as a yellow solid (159 mg, 15%).

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.65 (s, 1H), 7.39–7.21 (m, 6H), 7.13–7.08 (m, 1H), 7.05 (t,  $J$  = 7.4, 1H), 6.95 (d,  $J$  = 7.8, 2H), 6.87 (dd,  $J$  = 8.0, 1.5, 1H), 6.62 (d,  $J$  = 9.8, 1H), 3.58 (d,  $J$  = 7.5, 2H), 1.84 (dp,  $J$  = 13.7, 6.7, 1H), 0.84 (d,  $J$  = 6.7, 6H).  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ):  $\delta$  164.65, 157.81, 156.90, 144.25, 139.00, 137.91, 136.52, 135.64, 129.95, 129.83, 123.69, 123.53, 121.45, 121.31, 119.47, 117.35, 116.49, 109.99, 52.79, 29.79, 20.02. IR ( $\nu_{\text{max}}/\text{cm}^{-1}$ ): 3060, 2960, 2926, 2871, 1678, 1658, 1623, 1581, 1542, 1488, 1434, 1332, 1312, 1231, 1163, 1126, 1102, 1073, 997, 885, 838, 812, 790, 757, 736, 694, 639, 485, 461, 440. HRMS (ESI): calculated for  $\text{C}_{24}\text{H}_{24}\text{N}_3\text{O}_2$   $[M+\text{H}]^+$  386.18685, found 386.18630.

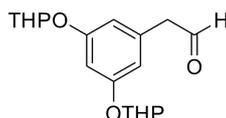
Methyl 2-(3,5-bis((tetrahydro-2H-pyran-2-yl)oxy)phenyl)acetate (**20p**)



To solution of methyl-2-(3,5-dihydroxyphenyl)acetate (**19**) (500 mg, 2.74 mmol) and pyridinium *p*-toluenesulfonate (22 mg, 0.087 mmol), in dry DCM (7 mL), was added dropwise dihydropyran (2.2 mL, 24.10 mmol) dissolved dry DCM (3 mL). The mixture became homogeneous at the end of the addition of the last reagent. After the end of reaction (3 h, TLC), the mixture was washed with water (2x), dried over  $\text{MgSO}_4$ , filtered and evaporated *in vacuo*. Compound **20p** was obtained as a yellow oil as a mixture of diastereoisomers and used in the next step without further purification (948 mg, 98%).

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  6.69 (q,  $J$  = 1.8 Hz, 1H), 6.61 (t,  $J$  = 2.3 Hz, 2H), 5.39 (dt,  $J$  = 6.1, 3.2 Hz, 2H), 3.89 (td,  $J$  = 11.3, 9.9, 3.0 Hz, 2H), 3.67 (s, 3H), 3.59 (dtd,  $J$  = 7.9, 4.0, 2.4 Hz, 2H), 3.54 (s, 2H), 2.05–1.52 (m, 12H).

2-(3,5-bis((tetrahydro-2H-pyran-2-yl)oxy)phenyl)acetaldehyde (**20**)

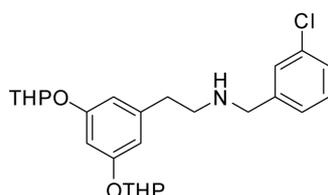


A solution of ester **20p** (900 mg, 2.57 mmol) in anhydrous DCM (10 mL) was stirred under a nitrogen atmosphere and cooled to  $-78^\circ\text{C}$ ; DIBAL-H dissolved in DCM (1 M, 4.13 mL, 3.86 mmol) was then added dropwise over 10 minutes with a syringe. After 30 minutes, the reaction

was quenched with a  $\text{NH}_4\text{Cl}$  aq solution (20 mL), followed by addition of water (20 mL) and extraction with EtOAc (3 x 20 mL). The combined organic layers were washed with water (50 mL) and an aqueous saturated NaCl solution (25 mL). The collected organic phases were dried over dry  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The crude so obtained was purified by FC (pentane/EtOAc, 8:2) affording aldehyde (**20**) as a mixture of diastereoisomers as a yellow oil (403 mg, 49% yield).

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.68 (t,  $J = 2.4$ , 1H), 6.71 (d,  $J = 2.1$ , 1H), 6.55 (t,  $J = 2.1$ , 2H), 5.38 (dt,  $J = 6.3, 3.1$ , 2H), 3.93–3.83 (m, 2H), 3.63–3.53 (m, 4H), 2.00–1.58 (m, 12H).

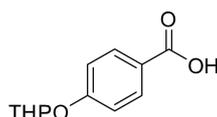
2-(3,5-Bis((tetrahydro-2H-pyran-2-yl)oxy)phenyl)-N-(3-chlorobenzyl)ethanamine (22)



Under a nitrogen atmosphere, a solution of aldehyde **20** (350 mg, 1.09 mmol) and 3-chlorobenzylamine (**21**) (134  $\mu\text{L}$ , 1.09 mmol) in anhydrous THF (20 mL) was refluxed for 15 minutes. After cooling to  $0^\circ\text{C}$ , sodium triacetoxyborohydride (693 mg, 3.27 mmol) was slowly added to prevent the formation of foam. After stirring at rt for 24-hour, the reaction was quenched with a saturated  $\text{NaHCO}_3$  solution (40 mL) and extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with water (20 mL), dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. Compound **22** was obtained as a mixture of diastereoisomers as a yellow oil (413 mg, 85% yield) and used in the next step without further purification.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.29 (s, 1H), 7.25–7.09 (m, 3H), 6.66–6.47 (m, 3H), 5.39–5.33 (m, 2H), 3.93–3.83 (m, 2H), 3.77 (s, 2H), 3.56 (d,  $J = 9.9$  Hz, 2H), 2.94–2.65 (m, 4H), 2.00–1.58 (m, 12H).

4-((Tetrahydro-2H-pyran-2-yl)oxy)benzoic acid (23)

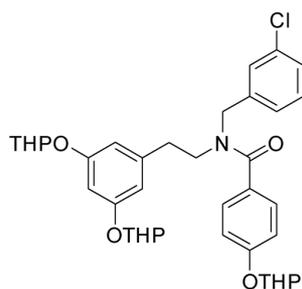


A solution of 4-hydroxybenzoic acid (1.00 g, 7.24 mmol) and pyridinium *p*-toluenesulfonate (64 mg, 0.25 mmol) in anhydrous DCM (12 mL) was treated dropwise with a solution of 3,4-dihydro-2H-pyran (0.80 mL, 8.77 mmol) in anhydrous DCM (3 mL). The reaction mixture was stirred at room temperature and after 1 hour and 30 minutes, the TLC indicated full consumption

of starting material. The organic layer was washed with water (2 x 15 mL), dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude was further purified by FC (DCM/MeOH, 99:1 to 96:4 v/v) afforded **23** as a white solid (441 mg, 27% yield).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.06 (d, J = 8.8, 2H), 7.10 (d, J = 8.8, 2H), 5.53 (t, J = 3.0, 1H), 3.86 (td, J = 10.9, 3.3, 1H), 3.63 (dt, J = 10.9, 3.3, 1H), 2.09–1.85 (m, 3H), 1.77–1.56 (m, 3H).

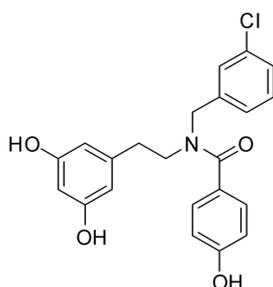
*N*-(3,5-Bis((tetrahydro-2H-pyran-2-yl)oxy)phenethyl)-N-(3-chlorobenzyl)-4-((tetrahydro-2H-pyran-2-yl)oxy)benzamide (4p)



A solution of carboxylic acid **23** (150 mg, 0.63 mmol) in anhydrous DMF (3.5 mL) was cooled to 0 °C, and HOBt (102 mg, 0.76 mmol), EDC·HCl (144 mg, 0.76 mmol) were subsequently added. The solution was stirred under a nitrogen atmosphere at rt for 30 minutes. Then, a solution of amine **22** (280 mg, 0.63 mmol) and DIPEA (275 μL, 1.58 mmol) in anhydrous DMF (3.5 mL) was added. The reaction was stirred at rt for 24 hours. Successively, the reaction was quenched with water (10 mL) and extracted with EtOAc (2 x 15 mL). The combined organic layers were washed with water (30 mL) and a saturated NaCl solution (30 mL). The collected organic phases were dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude was further purified by FC (pentane/EtOAc, 80:20 to 70:30, v/v), affording amide **4p** as a colorless oil (245 mg, 60% yield).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 7.40–7.28 (m, 3H), 7.24–6.90 (m, 5H), 6.71–6.49 (m, 2H), 6.30 (s, 1H), 5.43 (s, 1H), 5.35 (s, 2H), 4.71 (s, 1H), 4.28 (s, 1H), 3.95–3.80 (m, 3H), 3.67–3.54 (m, 5H), 3.03–2.60 (m, 2H), 2.02–1.65 (m, 18H).

*N*-(3-Chlorobenzyl)-*N*-(3,5-dihydroxyphenethyl)-4-hydroxybenzamide (**4**)



A solution of intermediate **4p** (100 mg, 0.15 mmol) and pyridinium *p*-toluenesulfonate (17.3 mg, 0.068 mmol) in EtOH (7 mL) was stirred at 55 °C for 16 hours. The reaction solvent was evaporated under reduced pressure, and the reaction mixture was dissolved in EtOAc (10 mL). The organic layer was washed with a saturated NaCl solution (2 x 10 mL) and water (10 mL). The collected organic phases were dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The obtained crude was purified by FC (DCM/MeOH, 97:3 to 96:4 v/v) yielding the final amide **4** as a white solid (45 mg, 75%).

<sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD at -50°C) [mixture of conformers]: δ 7.44–7.26 (m, 8H), 7.15–7.04 (m, 4H), 6.79 (d, *J*=8.4, 4H), 6.18–6.07 (m, 4H), 5.93 (d, *J*=1.6, 2H), 4.79 (s, 2H), 4.36 (s, 2H), 3.51 (t, *J*=6.7, 2H), 3.41 (t, *J*=6.7, 2H), 2.78 (t, *J*=6.7, 2H), 2.67 (t, *J*=6.7, 2H). <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD at -50°C) [mixture of conformers] : δ 175.15, 174.58, 160.80, 160.36, 159.79, 159.76, 159.48, 142.52, 141.47, 141.20, 140.71, 135.70, 135.41, 131.66, 131.44, 129.84, 129.59, 128.83, 128.64, 127.93, 127.35, 127.25, 127.11, 126.30, 116.13, 115.98, 108.09, 108.00, 101.36, 54.15, 51.75, 48.49, 47.92, 35.21, 34.14. IR (ν<sub>max</sub>/cm<sup>-1</sup>): 3307, 2927, 1600, 1516, 1471, 1426, 1345, 1274, 1160, 1079, 1043, 994, 842, 767, 691, 592, 444. HRMS (ESI): calculated for C<sub>22</sub>H<sub>21</sub>ClNO<sub>4</sub> [*M*+H]<sup>+</sup> 398.1200, found 398.1154.

***STD NMR: methods and spectra (compounds 1 and 3)***

**Methods**

All protein–ligand samples were prepared in a 1000:1 ligand/protein ratio. Typically, the final concentration of the samples was 400 μM of ligand and 0.4 μM of HuR, and the final volume was 200 μL. The buffer used is a 20 μM deuterated phosphate buffer pH 7.4.

Prior to STD-NMR experiments compounds **1** and **3** were converted to the corresponding hydrochloride salt according to the following procedures.

37% HCl (4 μL, 0.128 mmol) was added to a solution of (±)-**1** (30 mg, 0.064 mmol) in methanol (500 μL) under stirring. The solvent was then evaporated *in vacuo* obtaining final compound (±)-**1** HCl as a white solid. 37% HCl (4.2 μL, 0.132 mmol) was added to a solution of **3** (30 mg,

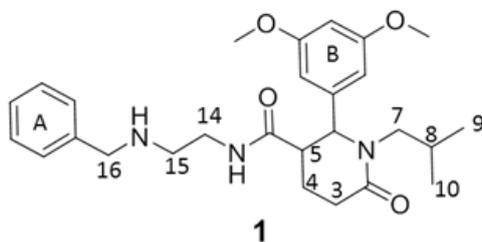
0.066 mmol) in methanol (500  $\mu$ L) under stirring. The solvent was then evaporated *in vacuo* obtaining final compound **3** HCl as a light-yellow solid.

$^1\text{H}$ -STD-NMR experiments were performed on a 600 MHz Bruker Avance spectrometer. The probe temperature was maintained at 283 K. In the STD experiments, water suppression was achieved by the WATERGATE 3-9-19 pulse sequence. The on-resonance irradiation of the protein was performed at  $-0.05$  ppm. Off-resonance irradiation was applied at 200 ppm, where no protein signals are visible. Selective presaturation of the protein was achieved by a train of Gauss-shaped pulses of 49 ms length each. The STD-NMR spectra were acquired with an optimized total length of saturation train of 2.94 s. Blank experiments were conducted in absence of protein in order to avoid artefacts.

Intensities of all STD effects (absolute STD) were calculated by division through integrals over the respective signals in STD-NMR reference spectra. The different signal intensities of the individual protons are best analyzed from the integral values in the reference and STD spectra, respectively.  $(I_0 - I_{\text{sat}})/I_0$  is the fractional STD effect, expressing the signal intensity in the STD spectrum as a fraction of the intensity of an unsaturated reference spectrum. In this equation,  $I_0$  is the intensity of one signal in the off-resonance or reference NMR spectrum,  $I_{\text{sat}}$  is the intensity of a signal in the on-resonance NMR spectrum, and  $I_0 - I_{\text{sat}}$  represents the intensity of the STD-NMR spectrum.

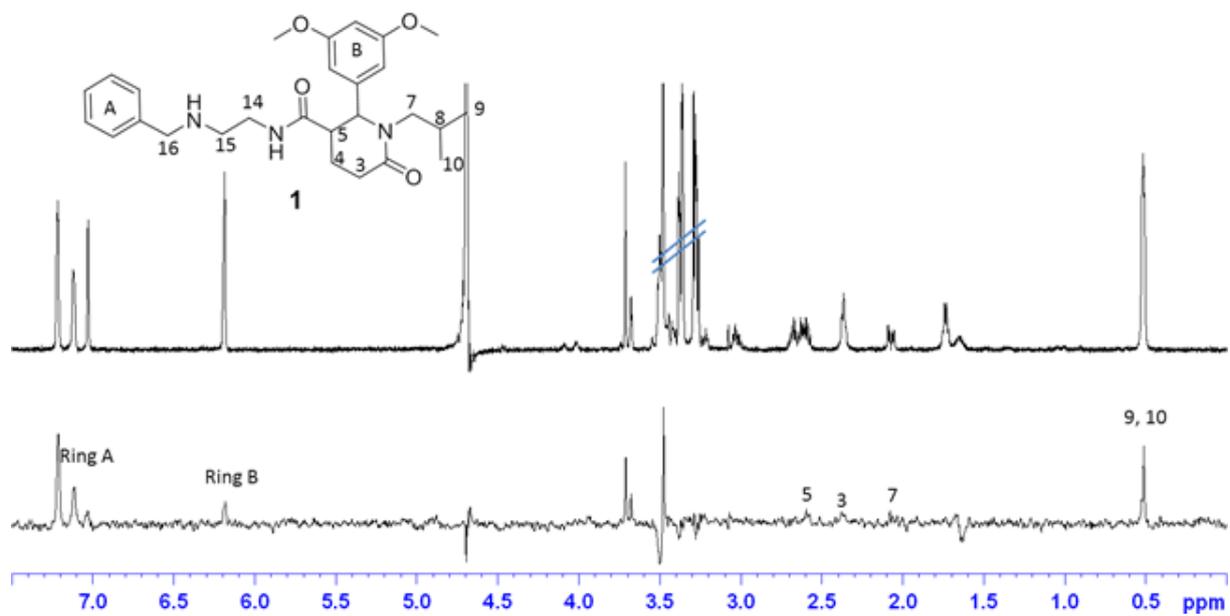
**Full NMR characterization in phosphate buffer (1, 3 and 4) and STD-NMR spectra (1 and 3)**

Chemical shift assignment of compound **1** in phosphate buffer 20 mM pH = 7.4 at 283 K.

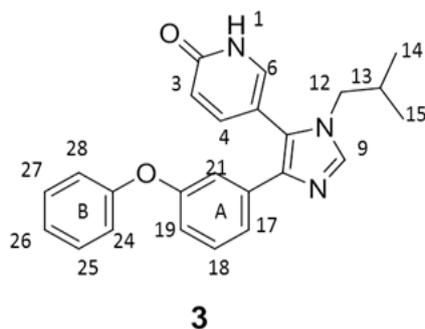


	$^1\text{H}$ (ppm)	$^{13}\text{C}$ (ppm)	Mult., $J$ (Hz)	<i>N.O.E.</i>	<i>STD</i> %
<i>Ring A</i>	7.12	129.28	bs	-	0.27
	7.25	129.42	bs	-	0.19
<i>Ring B</i>	6.20	99.93	d, $J = 1.8$	OCH <sub>3</sub>	0.03
	6.20	105.83	d, $J = 1.8$	OCH <sub>3</sub> -6-5-7°-8	
<i>OCH<sub>3</sub></i>	3.50	55.50	s	12-11-13	0.15
<b>3</b>	2.45	30.60	t, $J = 6.8$	4-5	0.20
<b>4</b>	1.70	22.00	m	3-5-6	
<b>5</b>	2.52	49.30	m	4-6-12-11	0.27
<b>6</b>	4.65	63.50	-	12-11-5-7°-8-9- 10-4	-
<b>7°</b>	2.05	51.40	dd, $J = 6.52$ ; 14.40	7b	0.33
<b>7b</b>	3.37	51.40	m	7°-9-10-8	-
<b>8</b>	1.62	25.40	m	9-10-6-7b	-
<b>9-10</b>	0.50	19.11	d, $J = 3.35$	8-7°-7b-6	0.40
<b>14</b>	3.10	36.20	m	15	-
<b>15</b>	2.6–2.55	45.80	m	14-16	-
<b>16</b>	3.69	51.10	s	15-ring A	-

STD spectrum (bottom) and  $^1\text{H-NMR}$  (top) of compound **1**.

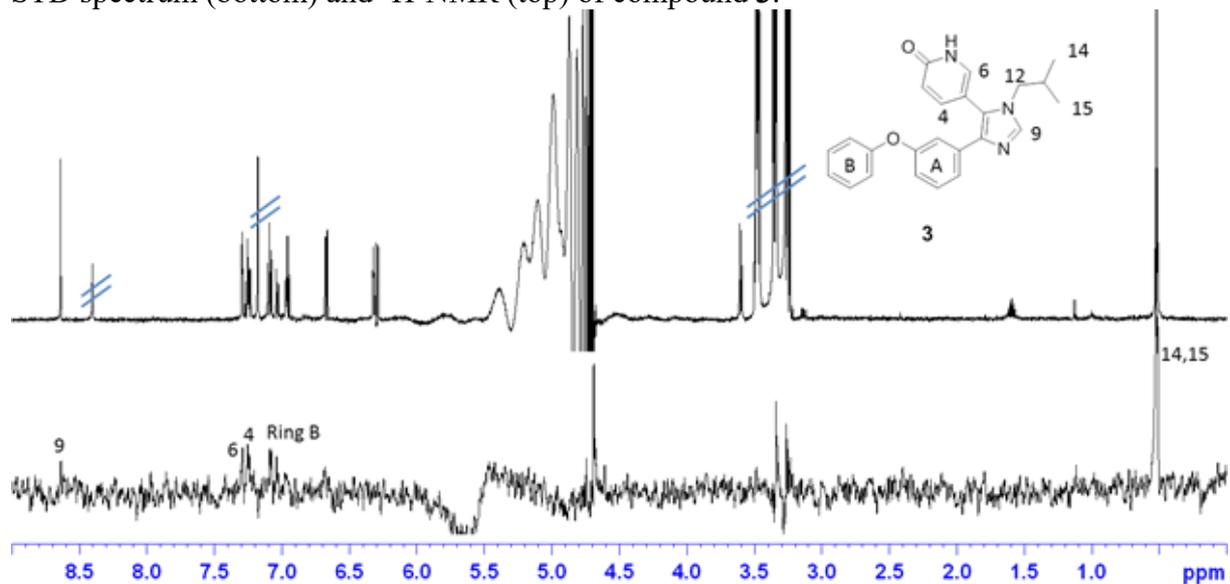


Chemical shift assignment of compound **3** in phosphate buffer 20 mM pH = 7.4 at 283 K.

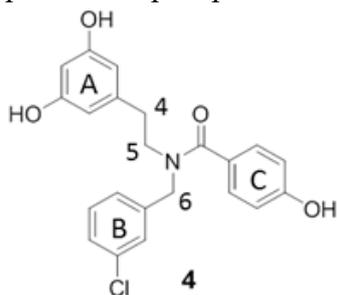


	$^1\text{H}$ (ppm)	$^{13}\text{C}$ (ppm)	Mult., J (Hz)	N.O.E.	STD %
<i>ring A</i>	7.26 ( <b>18</b> )	131.1	m		
	7.04 ( <b>17</b> )	121.3	d, J = 8.26		-
	6.96 ( <b>19</b> )	119.5	m		
<i>ring B</i>	6.67 ( <b>24-28</b> )	115.4	d, J = 8.6		
	7.09 ( <b>25-27</b> )	119.6	t, J = 8.6		0.30
	6.96 ( <b>26</b> )	129.9	m		
<b>3</b>	6.30	124.7	d, J = 9.4Hz		-
<b>4</b>	7.23	120.5	dd, J = 2.54; 9.4		1.20
<b>6</b>	7.29	144.1	d, J = 2.54		1.10
<b>9</b>	8.64	138.1	s		0.20
<b>12</b>	3.60	134.7	d, J = 7.8		-
<b>13</b>	1.59	54.0	m		-
<b>14-15</b>	0.53	28.3	d, J = 6.8		0.30
<b>21</b>	6.33	18.4	t, J = 2.01		-

STD spectrum (bottom) and  $^1\text{H-NMR}$  (top) of compound 3.



Chemical shift assignment of compound **4** in phosphate buffer 20 mM pH = 7.4 at 283 K.



	$^1H$ (ppm)	$^{13}C$ (ppm)	Mult., J (Hz)	N.O.E.	STD%
<b>Ring A</b>	6.10	108.11	d, J = 2.03	6, 5	
	5.98	100.86	s	4, 6, 5	0.40
	5.74	108.30	s	6, 5	
<b>Ring B</b>	7.16	127.40	s	5, 4, 6	
	7.13	126.32	m	5, 4, 6	0.36
	7.05	128.30	m	6	
<b>Ring C</b>	6.72	128.74	d, J = 8.42	6, 5, 4	
	6.58	115.24	d, J = 8.42	6, 5, 4	0.25
<b>4</b>	2.58	33.50	-	Ring B, ring C, ring A, 5, 6	-
<b>5</b>	3.57	50.70	bs	Ring B, ring C, ring A, 4 Ring C, 5,	-
<b>6</b>	4.30	52.90	-	ring A, ring B	-

## *In silico studies: Methods and docking*

### **Methods**

Modeling simulations were performed starting from the crystal structure of the two N-terminal RRM domains of HuR complexed with RNA (PDB code: 4ED5). The complex was prepared using the Protein Preparation Wizard implemented in Maestro, using OLPS-2005 as force field. The Prime module allowed us to add all missing side chains; residual crystallographic buffer components and water molecules were removed and hydrogen atoms were added; side chain protonation states at pH 7.4 were assigned.

The HuR-RNA complex was submitted to 10000 MacroModel minimization steps, using OPLS\_2005 as force field.

In order to consider the protein as ensembles of conformational states, we submitted the HuR-mRNA complex to 500 ns of molecular dynamics simulations (MDs). MDs were run using the Desmond package v. 3.8 at 300 K temperature and ensemble NPT class; the system was immersed in an orthorhombic box of TIP4P water molecules, extending at least 10 Å from the protein and counter ions were added to neutralize the system charge. The resulting trajectory was clustered with respect to RMSD (Root Mean Square Deviation), in order to explore all the structures obtained, getting ten representative structures, which were submitted to 10.000 MacroModel minimization steps, using OLPS-2005 as force field.

In detail, according to the Prime calculate Energy tool, the lowest- and highest-energy structures, were selected for molecular recognition studies. Furthermore, the two selected HuR states correspond to “open” and “closed” protein conformations. For our docking studies, we chose the “closed” HuR conformations.

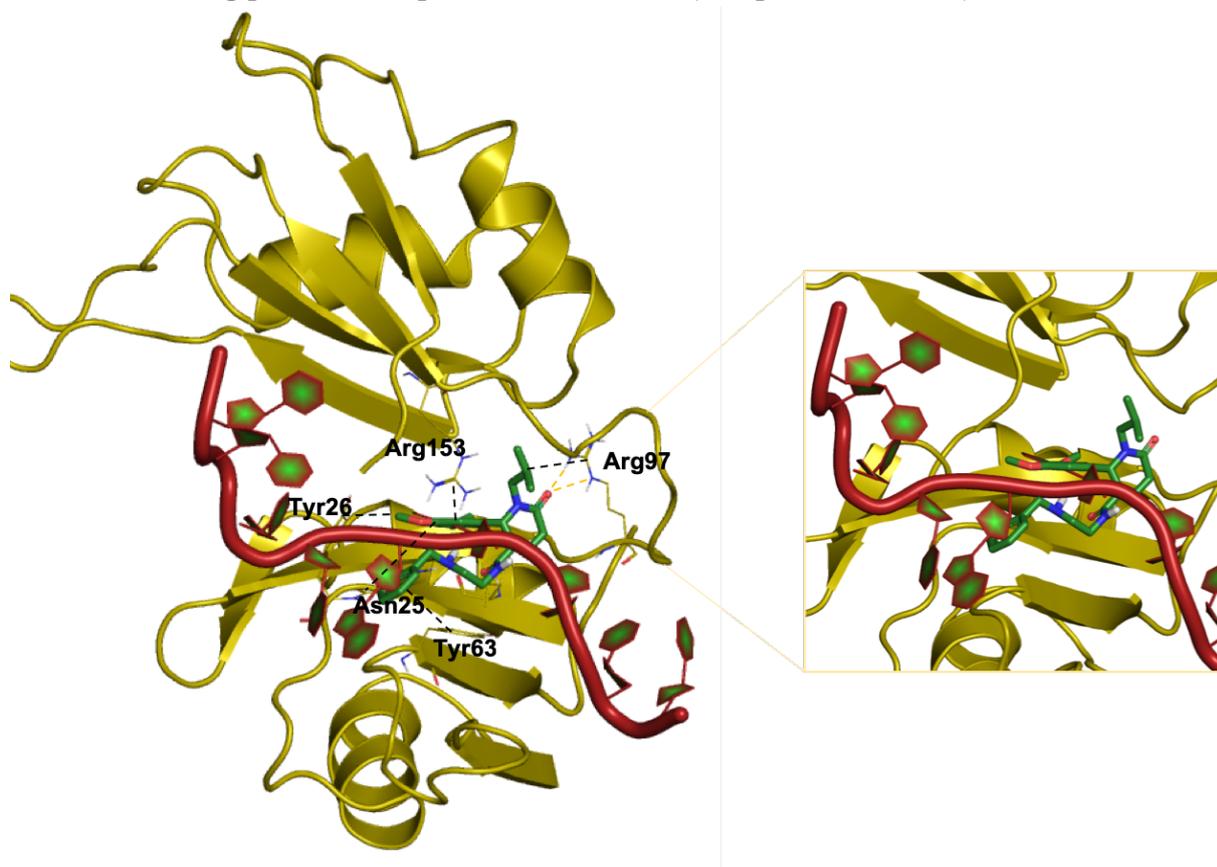
Compounds **1**, **3** and **4** were prepared by means of LigPrep Tools at pH 7.4 and were submitted to 10000 MacroModel minimization steps, using OPLS\_2005 as force field. Docking simulations were carried out exploiting the Glide software; by using SP v. 6.7 (standard precision) algorithm, 100 poses were generated for each ligand.

### **RMSD values for each analyzed compound with respect to U8**

<i>Compound</i>	<i>RMSD value</i>
<i>1</i>	1.8
<i>3</i>	1.9
<i>4</i>	0.8

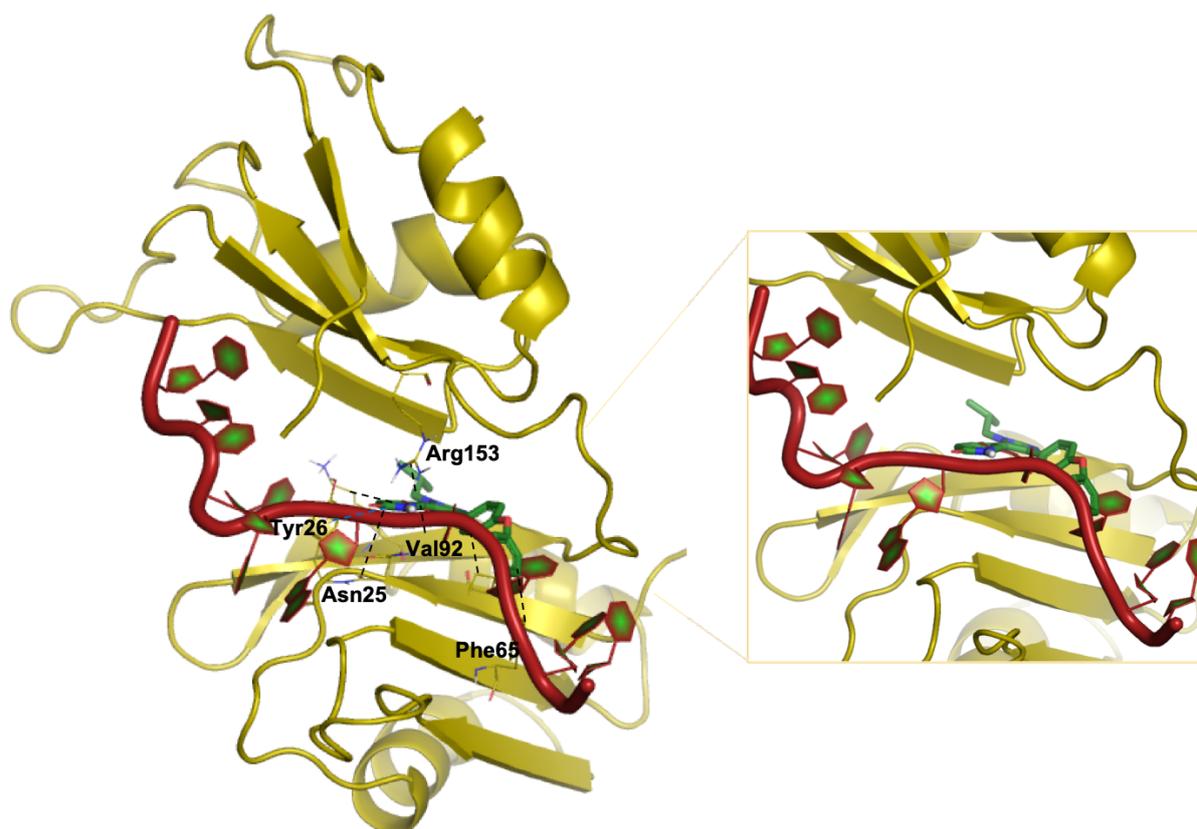
The RMSD value are reported in Å.

### Selected docking poses and reported interactions (compounds 1 and 3)



3D representation of HuR–1 interactions. Hydrogen bonds and hydrophobic interactions are displayed as yellow and dark dashed lines, respectively. RNA and HuR are represented as a brown and yellow cartoon, respectively. Compound 1 is shown in green sticks.

We report the main interactions as follows: the benzyl ring is involved in hydrophobic interactions with Tyr63, Ile23, Asn25 and Leu61 of HuR. The carboxyl group of the piperidone ring establishes a double hydrogen bond with Arg97 and hydrophobic interactions with the same residues. The isopropyl group (that has the strongest STD signal) establishes strong hydrophobic interactions with HuR (specifically, residues Ile103, Phe154, Asp155, Arg153 and Ile133). The anchor moiety establishes hydrophobic interactions with Tyr26, Asn25, Ile23, Ser94, Ile133 and Arg153; the zoom view highlights its full overlap with U8.



3D representation of HuR–**3** interactions. Hydrogen bonds and hydrophobic interactions are displayed as yellow and dark dashed lines, respectively; a  $\pi$ – $\pi$  stacking interaction is shown as blue dashed line. RNA and HuR are represented as a brown and yellow cartoon, respectively. Compound **3** is shown in green sticks.

We report the main interactions as follows: the external phenyl ring establishes a hydrophobic interaction with Phe65, while the phenyl directly bound to the imidazole ring establishes a  $\pi$ -cation interaction with Arg153 and hydrophobic interactions with Ile133, Asn25 and Ile23. The imidazole ring is involved in hydrophobic interactions with Ser94, Lys92, Val93, Ile23 and Ile133. The zoom view highlights the full overlap of U8 and the anchor motif of compound **3**.

## ***Chiral resolution via enantioselective HPLC***

### **General Methods**

HPLC analyses were performed on a JASCO system: pump (PU 1580), degasser (DG-2080-53) and mixer (LG-1580-02); PDA (photo array detector, MD 1510), injection system Rheodyne valve 7125 equipped with a 10  $\mu$ L loop. Semi-preparative methods were conducted on the same system equipped with a 1mL loop. Specific rotations were determined on a DIP 100 JASCO polarimeter, equipped with sodium ( $\lambda = 589$  nm) and mercury ( $\lambda = 405$  nm) lamps. A Laborota 4000 Efficient (Heidolph) rotary evaporator was exploited for solvent evaporation of each fraction.

HPLC solvents (*n*-hexane, 2-propanol, ethanol, methanol), were acquired from Sigma Aldrich and VWR.

Chromatographic analyses were performed on immobilized polysaccharide chiral columns: Regispack IA [tris-(3,5-dimethylphenyl carbamate) of 120mylose , 4.6 mm I.D. x 150 mm, s-5 $\mu$ m] *Regis Technologies*, and Chiralpak IC [tris-(3,5-dichlorophenyl carbamate) of cellulose, 4.6 mm I.D. x 250 mm, s-5 $\mu$ m].

Semi-preparative chromatography was carried out on Chiralpak IC (1 cm I.D. x 25 cm, s-5 $\mu$ m).

### **Screening protocol for eluent mixtures to separate ( $\pm$ ) 1 on REGISPACK IA and CHIRALPAK IC columns.**

Eluent	Composition (v/v)	Flow	DEA
1	<i>n</i> -Hex/EtOH (90:10)	1 mL/min	0.1%
2	<i>n</i> -Hex /EtOH (85:15)	1 mL/min	0.1%
3	<i>n</i> -Hex /EtOH (80:20)	1 mL/min	0.1%
4	<i>n</i> -Hex /IPA (90:10)	1 mL/min	0.1%
5	<i>n</i> -Hex /IPA (85:15)	1 mL/min	0.1%
6	<i>n</i> -Hex /IPA (80:20)	1 mL/min	0.1%
7	<i>n</i> -Hex /IPA (75:25)	1 mL/min	0.1%
8	<i>n</i> -Hex /IPA (70:30)	1 mL/min	0.1%
9	<i>n</i> -Hex /IPA (50:50)	1 mL/min	0.1%
10	IPA (100)	0.5 mL/min	0.1%

## Parameters

The retention factor **K** was calculated according to the following equation:

$$K = \frac{T_r - T_o}{T_o}$$

The selectivity **α** was calculated according to the following equation:

$$\alpha = \frac{K_2}{K_1}$$

The flow rate was determined as follows:

$$F_{sp} = F_a \times \frac{L_p}{L_a} \times \frac{d_{sp}^2}{d_a^2}$$

$F_{sp}$  = flow rate on semi-preparative scale

$F_a$  = flow rate on analytical scale

$L_{sp}$  and  $L_a$  = column length

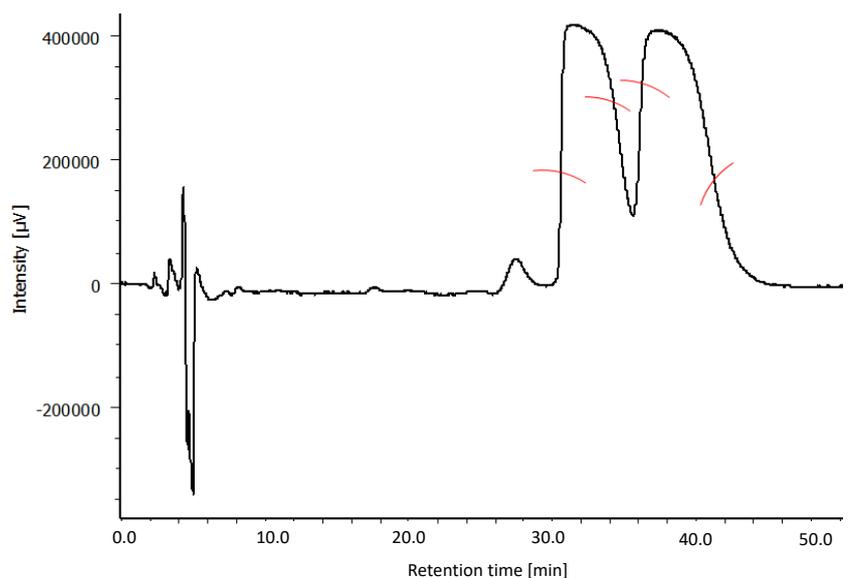
$d_{sp}$  and  $d_a$  = column inside diameter

$$F_{sp} = \frac{1 \text{ mL}}{\text{min}} \times \frac{25 \text{ cm}}{25 \text{ cm}} \times \frac{100 \text{ mm}^2}{21.16 \text{ mm}^2} = 4 \text{ mL/min}$$

The theoretical flow rate value was optimized according to the column operational pressure (flow = 2 mL/min).

**Results of the screening** *The most interesting conditions are highlighted in bold.*

Eluent	REGISPACK IA				CHIRALPAK IC			
	K1	K2	α	Rs	K1	K2	α	Rs
1	6.906	8.402	1.217	1.204				
2	1.859	2.107	1.133	0.776				
3					3.775	-	N/A	N/A
4	7.017	8.436	1.200	1.324				
5	3.209	3.825	1.192	0.762				
<b>6</b>	1.983	2.295	1.157	N/A	<b>8.786</b>	<b>10.818</b>	<b>1.231</b>	<b>2.530</b>
<b>7</b>					<b>4.974</b>	<b>5.890</b>	<b>1.184</b>	<b>1.896</b>
8	0.983	1.115	1.134	0.812	3.621	4.401	1.215	1.891
9	0.564	-	N/A	N/A	1.482	1.837	1.239	1.460
10	0.442	-	N/A	N/A				

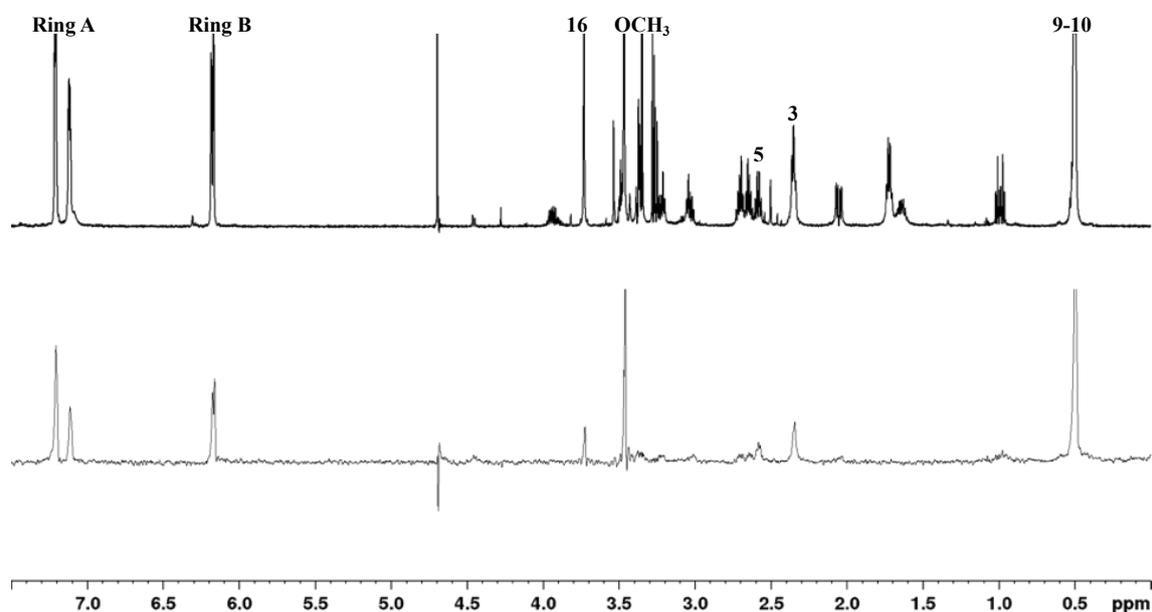


Semi-preparative chiral separation of ( $\pm$ ) **1** enantiomers performed using Chiralpak IC (1 cm I.D. x 25 cm, s-5 $\mu$ m), n-Hex:IPA:DEA 75:25:0.1 (v/v/v), flow: 2 mL/min, 25°C,  $\lambda$  = 220 nm. Injection volume 1 mL x 4 (10 mg/ml in n-Hex:IPA 75:25).

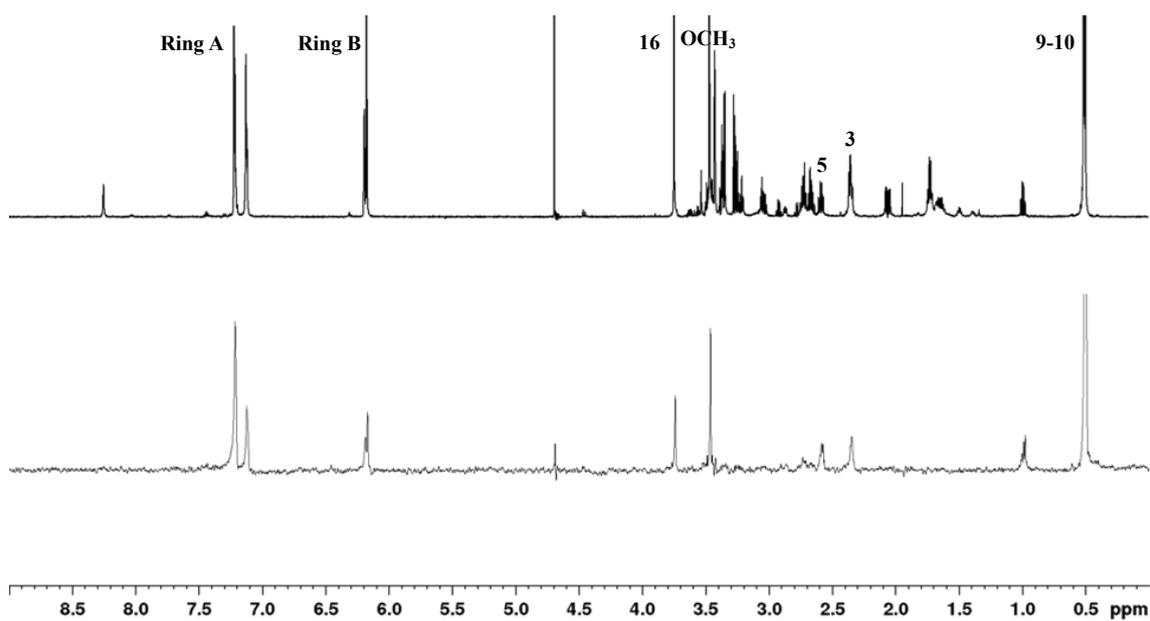
***Interaction studies with HuR of (+) 1 and (-) 1 by STD NMR, and solvent- DEEP-STD***

**STD NMR spectra**

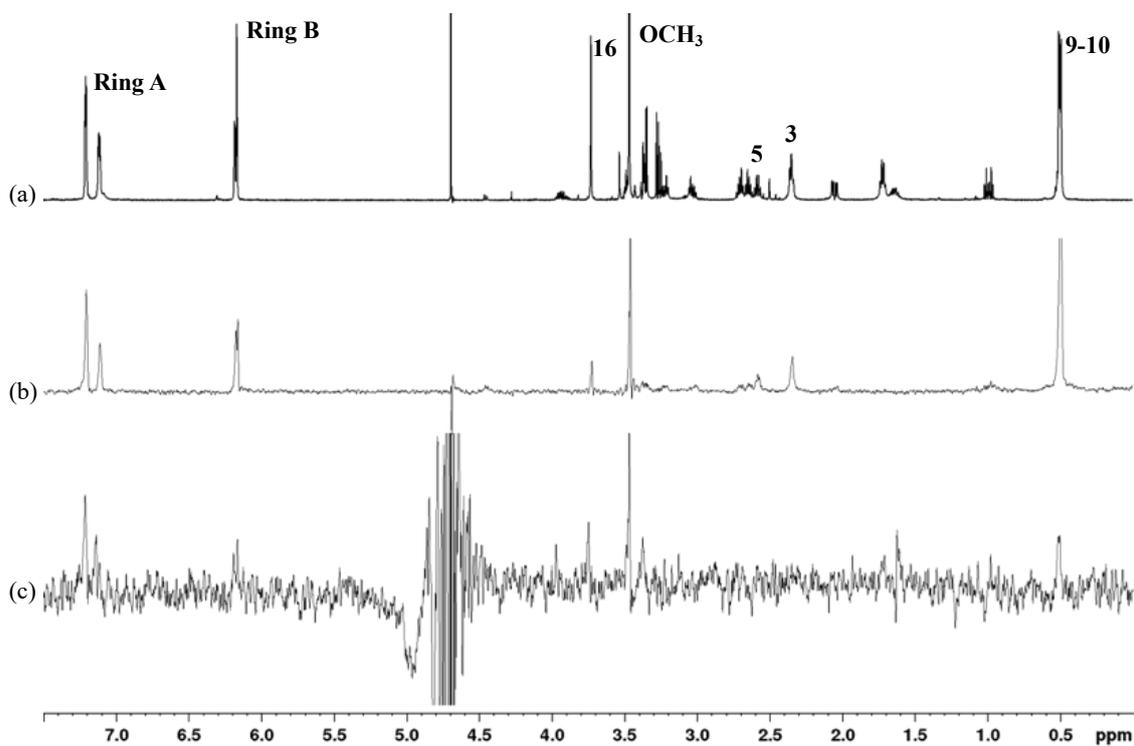
STD spectrum (bottom) and  $^1\text{H}$  NMR (top) of compound (+) **1**



STD spectrum (bottom) and  $^1\text{H}$  NMR (top) of compound (-) **1**



Solvent DEEP-STD spectra of of compound (+) **1**: (a) Reference spectra; (b) STD spectrum in D<sub>2</sub>O; (c) STD spectrum in H<sub>2</sub>O:D<sub>2</sub>O 90:10.



## ***Chapter 6: Design of new derivatives of synthetic hit 4 via fragment-based approach***

### ***Screening of halogen-enriched fragment library on native HuR***

#### **Screening of the 146-fragment library by SPR**

SPR experiments were performed using a Reichert SR7500DC surface plasmon resonance spectrometer (Reichert Technologies, Depew, NY, USA) equipped with a medium density carboxymethyl dextran hydrogel CMD500 sensor chip (Xantec Bioanalytics, Düsseldorf, Germany). HuR (37 kDa) was immobilized using amine coupling chemistry. The surface of the flow cells on the sensor chip were activated with a 100  $\mu$ l injection of a 1:1 mixture of 0.1 M NHS (N-hydroxysuccinimide) and 0.1 M EDC (3-(N,N-dimethylamino)propyl-N-ethylcarbodiimide) at a flow rate of 10  $\mu$ l/min. HuR protein was applied to the flow cells at a concentration of 50  $\mu$ g/ml in 10 mM acetate buffer pH 4.0 for immobilization at a flow rate of 5  $\mu$ l/min. The adjacent flow cell was left blank to use as reference channel. All surfaces were blocked with a 100  $\mu$ l injection of 1 M ethanolamine•HCl pH 8.5 at a flow rate of 25  $\mu$ l/min; the final signal (difference between protein channel and reference) corresponded to a density of 2000 RU (response units). Integrity of the binding site was confirmed by injecting epigallocatechin gallate (EGCG, see chapter 4, compound 1) as reference compound prior to the screening. To perform the screening and collect binding data, all fragments in 20 mM HEPES buffer pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% p20 (v/v), 5% DMSO (v/v), were injected over the flow cells at a flow rate of 50  $\mu$ l/min at 20°C. Association was allowed for 1 minute and dissociation for 2 minutes. A three-concentration-point screening was performed in duplicates (250–500–1000  $\mu$ M) allowing to select 10 fragments for  $K_D$  determination; the following curves were built with 10 concentration points (15  $\mu$ M to 2 mM). Ethylene glycol 80% in the running buffer was used for regeneration of the surface. Differences in the bulk refractive index due to DMSO were corrected by 7-concentration-point DMSO calibration (3.5%–6.5%) and blanks (running buffer) were injected for the following data fitting. Processing of the experimentally observed sensorgrams was obtained by successive subtraction of signals obtained for the protein and reference channels injected with the fragments under the same conditions, as well as DMSO and blank correction, exploiting Scrubber (Version 2.0c, 2008, BioLogic Software).

### **Confirmation of binding for the 3 hit fragments by STD NMR**

Hit fragments were characterized by  $^1\text{H}$  NMR and assigned prior to the STD NMR experiments.  $^1\text{H}$  NMR spectra were acquired on a 400 MHz Bruker Avance spectrometer using a 1 mM solution of each fragment in a 20 mM deuterated phosphate buffer pH 7.4, 5% DMSO- $d_6$  at 303K.

One sample of protein HuR–hit fragment mix (**a+b+c**) was prepared in a 1000:1 fragment/protein ratio. The final concentration of the samples was 800  $\mu\text{M}$  for the ligands and 0.8  $\mu\text{M}$  of HuR, and the final volume was 500  $\mu\text{L}$ . The buffer used is a 20 mM deuterated phosphate buffer pH 7.4.

$^1\text{H}$ -STD NMR experiments were performed on a 600 MHz Bruker Avance spectrometer. The probe temperature was maintained at 283 K. Water suppression was achieved by the WATERGATE 3-9-19 pulse sequence. The on-resonance irradiation of the protein was performed at  $-0.05$  ppm. Off-resonance irradiation was applied at 200 ppm, where no protein signals are visible. Selective presaturation of the protein was achieved by a train of Gauss-shaped pulses of 49 ms length each. The STD NMR spectra were acquired with an optimized total length of saturation train of 2.94 s. A blank experiment was conducted in absence of protein in order to avoid artefacts. Given the small size of the fragments, epitope mapping was not performed.

## Chapter 7: Fragment linking by protein templated dynamic combinatorial chemistry to identify new HuR ligands

### Thermal shift assay to study new protein constructs

Both constructs were tested in a thermal shift assay (TSA) in their native buffer (20 mM Phosphate buffer pH 7.0) at six protein concentrations (0.25–10  $\mu$ M). A 96-well plate was prepared for the experiment, each well was prepared in duplicate as reported in the table. The dye used is SyproOrange (Thermo Fischer Scientific) at the final concentration 4x.

The plate was subjected to a thermal cycle (21 to 95°C, incremental temperature ramp of 0.5°C per minute) on a real-time PCR instrument (Step one 2.3).

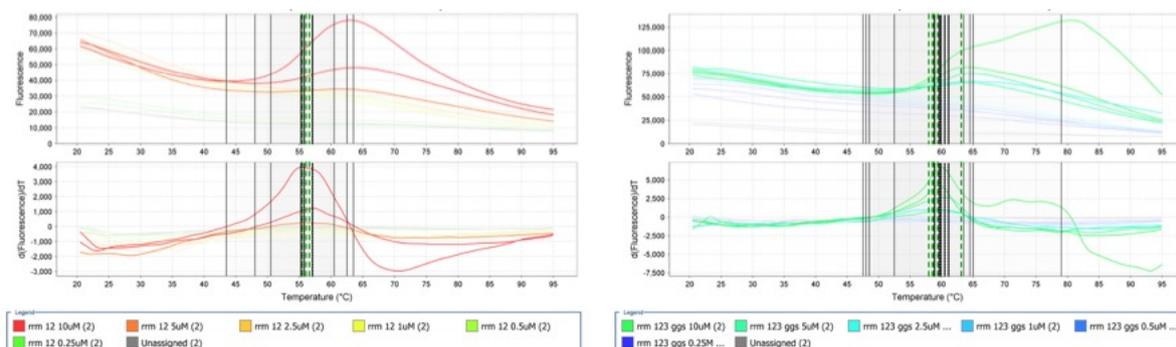
RRM1+2 (125  $\mu$ M)

final protein $\mu$ M	10	5	2.5	1	0.5	0.25
$\mu$ L protein	2	1	0.5	2.5	1.25	0.625
$\mu$ L dye	2.5	2.5	2.5	2.5	2.5	2.5
$\mu$ L buffer	19.25	20.25	20.75	18.75	20	20.625
$\mu$ L DMSO	1.25	1.25	1.25	1.25	1.25	1.25

RRM1+2+3\_GGS (27  $\mu$ M)

final protein $\mu$ M	10	5	2.5	1	0.5	0.25
$\mu$ L protein	9.26	4.63	2.32	2.5	1.25	0.625
$\mu$ L dye	2.5	2.5	2.5	2.5	2.5	2.5
$\mu$ L buffer	11.99	16.62	18.93	18.75	20	20.625
$\mu$ L DMSO	1.25	1.25	1.25	1.25	1.25	1.25

The melting curves were analyzed using the Protein Thermal Shift 1.3 software.



## ***Screening of halogen-enriched fragment library on native HuR***

### **Screening of the 146-fragment library by SPR**

SPR experiments were performed using a Reichert SR7500DC surface plasmon resonance spectrometer (Reichert Technologies, Depew, NY, USA) equipped with a medium density carboxymethyl dextran hydrogel CMD500 sensor chip (Xantec Bioanalytics, Düsseldorf, Germany). RRM1+2+3\_GGS (37 kDa) and RRM1+2 (20 kDa) were immobilized using amine coupling chemistry. The surface of the flow cells on the sensor chip were activated with a 100  $\mu$ L injection of a 1:1 mixture of 0.1 M NHS (N-hydroxysuccinimide) and 0.1 M EDC (3-(N,N-dimethylamino)propyl-N-ethylcarbodiimide) at a flow rate of 10  $\mu$ L/min. The constructs were applied to the flow cells at a concentration of 50  $\mu$ g/mL in 10 mM acetate buffer pH 4.5 for immobilization at a flow rate of 5  $\mu$ L/min. The adjacent flow cell was left blank to use as reference channel. All surfaces were blocked with a 100  $\mu$ L injection of 1 M ethanolamine•HCl pH 8.5 at a flow rate of 25  $\mu$ L/min; the final signal (difference between protein channel and reference) corresponded to a density of 3000 and 2000 RU respectively. Integrity of the binding site was confirmed by injecting epigallocatechin gallate (EGCG, see chapter 4, compound 1) as reference compound prior to the screening. To perform the screening and collect binding data, all fragments in 20 mM HEPES buffer pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% p20 (v/v), 5% DMSO (v/v), were injected over the flow cells at a flow rate of 50  $\mu$ L/min at 20°C. Association was allowed for 1 minute and dissociation for 2 minutes. A three-concentration-point screening was performed in duplicates (250–500–1000  $\mu$ M) allowing to select 20 fragments for  $K_D$  estimation; the following curves were built with 5 concentration points (125  $\mu$ M to 2  $\mu$ M). Ethylene glycol 80% in the running buffer was used for regeneration of the surface. Differences in the bulk refractive index due to DMSO were corrected by 7-concentration-point DMSO calibration (3.5%–6.5%) and blanks (running buffer) were injected for the following data fitting. Processing of the experimentally observed sensorgrams was obtained by successive subtraction of signals obtained for the protein and reference channels injected with the fragments under the same conditions, as well as DMSO and blank correction, exploiting Scrubber (Version 2.0c, 2008, BioLogic Software).

### ***Experimental conditions for the pt-DCC experiment***

#### **Buffer selection by TSA**

A 96-well plate, containing seventeen 10  $\mu$ M solutions of protein in sixteen buffers and H<sub>2</sub>O and 5% DMSO, (each solution prepared in duplicate) was stored at room temperature for four days, and sampled every day (day 0 to day 3) for TSA experiments. Blank and negative control wells were added to each TSA run. The solutions were sampled daily and transferred in a TSA 96-well plate prepared as described in the table.

The dye used is SyproOrange (Thermo Fischer Scientific) at the final concentration 4x.

The plate was subjected to a thermal cycle (21 to 95°C, incremental temperature ramp of 0.5°C per minute) on a real-time PCR instrument (Step one 2.3).

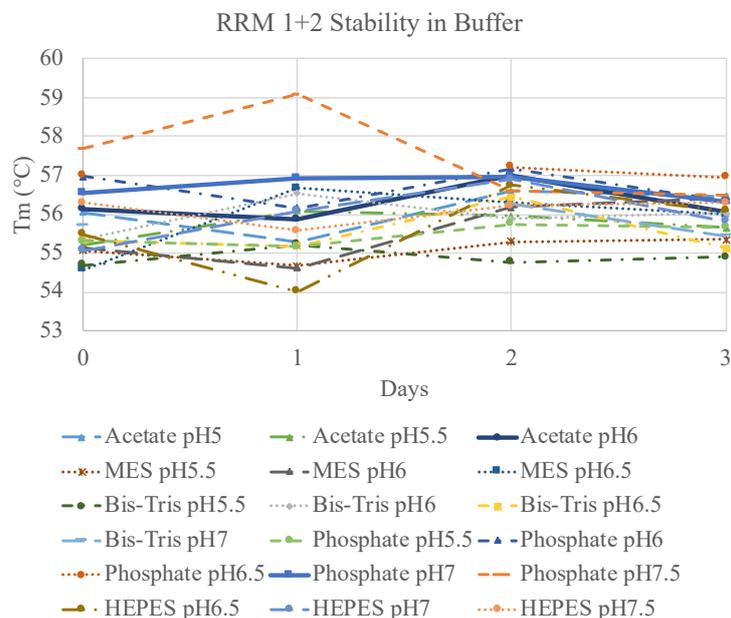
The buffers used are reported in the table

<b>Buffer Type</b>	<b>pH</b>
Acetate 10 mM	5–5.5–6
MES 20 mM	5.5–6–6.5
Bis-TRIS 20 mM	5.5–6–6.5–7
Phosphate 50 mM	5.5–6–6.5–7–7.5

The final content of each well is reported in the table

<b><math>\mu</math>L protein</b>	2
<b><math>\mu</math>L dye</b>	2.5
<b><math>\mu</math>L buffer</b>	19.25
<b><math>\mu</math>L DMSO</b>	1.25
<b>Total <math>\mu</math>L</b>	25

The melting curves were analyzed using the Protein Thermal Shift 1.3 software and the  $T_m$  recorded for each sample was plotted against time.

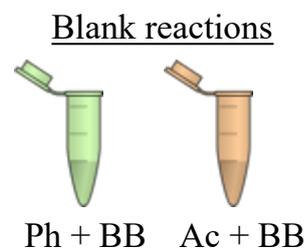


### *pt-DCC experiment*

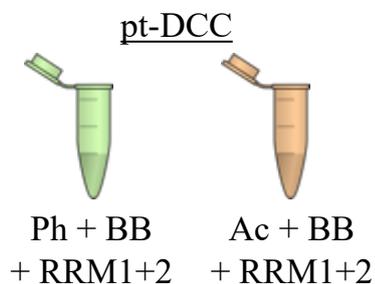
#### Experimental setup

The four reaction mixtures (two for each buffer type) were prepared as reported in the corresponding table.

Blank	Amount	Final concentration
Buffer	190 $\mu$ L	-
Aldehydes		100 $\mu$ M
Hydrazides	Premixed*	200 $\mu$ M
Aniline	10 $\mu$ L	1 mM
DMSO		5%
<i>Total</i>	<i>200 <math>\mu</math>L</i>	



pt-DCC	Amount	Final concentration
Buffer	126 $\mu$ L	-
Aldehydes		100 $\mu$ M
Hydrazides	Premixed*	200 $\mu$ M
Aniline	10 $\mu$ L	1 mM
DMSO		5%
RRM1+2 (125 $\mu$ M)	64 $\mu$ L	40 $\mu$ M
<i>Total</i>	<i>200 <math>\mu</math>L</i>	



Pre-mixed solution\*<sup>#,</sup>

<b>BB+aniline in DMSO</b>	<b>Amount</b>	<b>Final concentration</b>
Aldehydes (100 mM)	1 $\mu$ L x 2	2 mM
Hydrazides (100 mM)	2 $\mu$ L x 8	4 mM
Aniline (1 M)	1 $\mu$ L	20 mM
DMSO	31 $\mu$ L	-
<i>Total</i>	<i>50 <math>\mu</math>L</i>	

# The preparation of the pre-mixed solution followed a precise order to avoid triggering the reaction: 1) Hydrazides; 2) Aldehydes; 3) Aniline. The premixed solution was then added last to the reaction mixture.

The four reactions were then left rotating for 36h and sampled at regular intervals starting from hour 0. The samples for HPLC-UV-MS monitoring were prepared as follows: 5  $\mu$ L of each sample were added to a Eppendorf containing 44  $\mu$ L ACN analytical grade (precipitates protein) and 1  $\mu$ L NaOH 1 M (stops reaction reversibility). The vial was vortexed for 10 s and centrifuged for 5 min x 14000 RPM, 4°C, after which the supernatant was transferred to a HPLC vial for analysis. All analysis were performed on the same day at the end of the experiment. All samples previously collected were stored at 4°C.

### HPLC method

HPLC-UV-MS analyses were performed on a ThermoScientific Dionex Ultimate 3000 UHPLC System coupled to a ThermoScientific Q Exactive Focus with an electrospray ion source. An Acquity Waters Column (BEH, C8 1.7  $\mu$ m, 2.1 x 150 mm, Waters, Germany) equipped with a VanGuard Pre-Column (BEH C8, 5 x 2.1 mm, 1,7  $\mu$ m, Waters, Germany) was used for separation. The mass spectrum was measured in positive mode in a range from 100 – 700 m/z. At a flow rate of 0.250 mL/min, the elution was performed in a gradient as reported in the table: (A = ACN + 0.1% HCOOH; B = H<sub>2</sub>O + 0.1% HCOOH)

<b>min</b>	<b>%B</b>	<b>%A</b>
0	90	10
1	90	10
17	5	95
18.5	5	95
18.6	90	10
20.5	90	10

After each run, a 8 minute wash was performed on the column.

### **Determination of equilibrium and quantification of product amplification**

Products were identified univocally by high resolution mass and the area of the attributed peak was integrated through the Excalibur software at the selected wavelength of 272 nm.

To assess that equilibrium was reached, the area of each peak was plotted against time for each blank sample collected. Once the area reaches a value stable in time, the equilibrium is considered reached and the amplification can be investigated in pt-DCC samples.

Amplification is calculated for each peak by attributing the arbitrary value of 100% to the area in the blank. According to the following proportion for each peak:

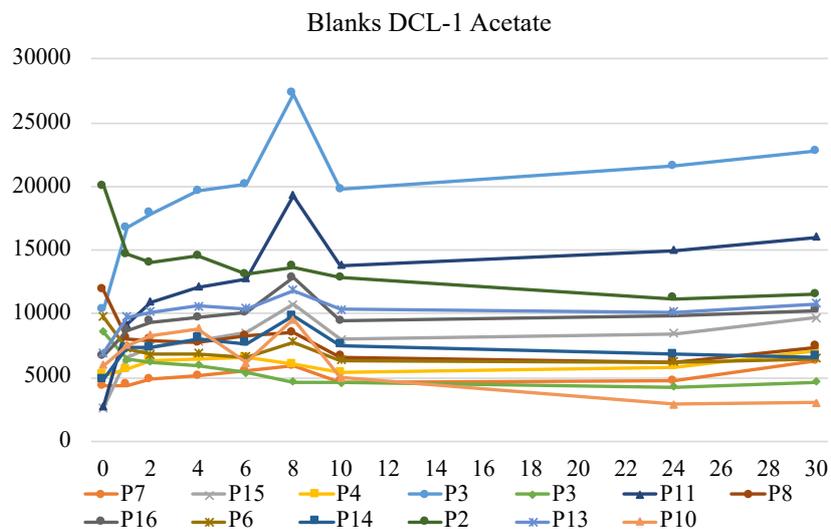
AreaDCC : x = Areablank : 100. Then,  $x = (\text{AreaDCC} * 100) / \text{Area blank}$ .

The amplification % is then calculated as  $A\% = x - 100$ .

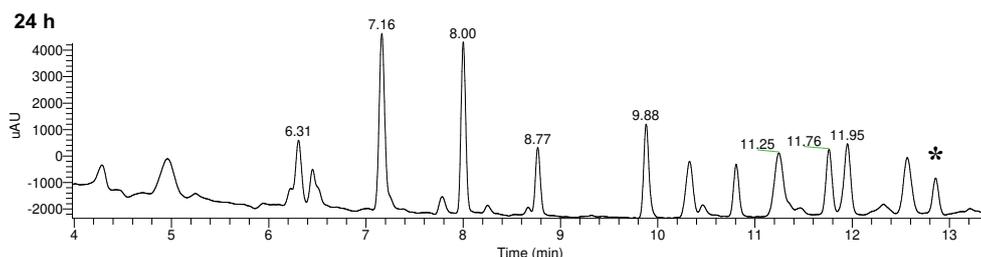
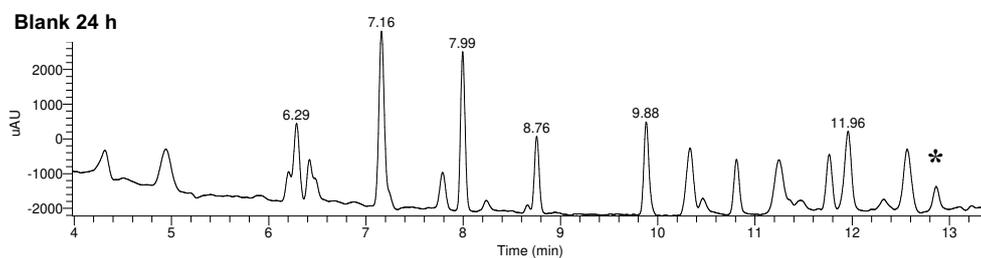
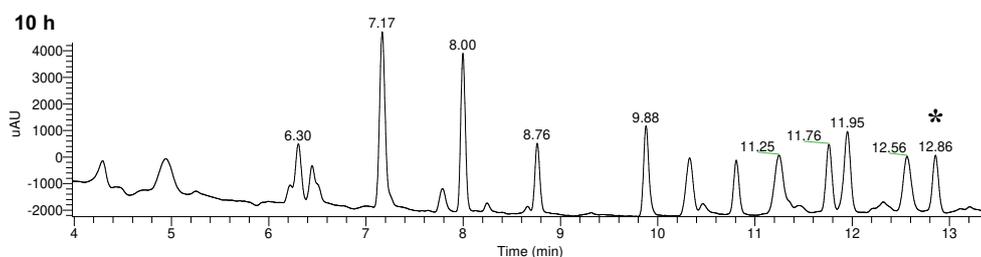
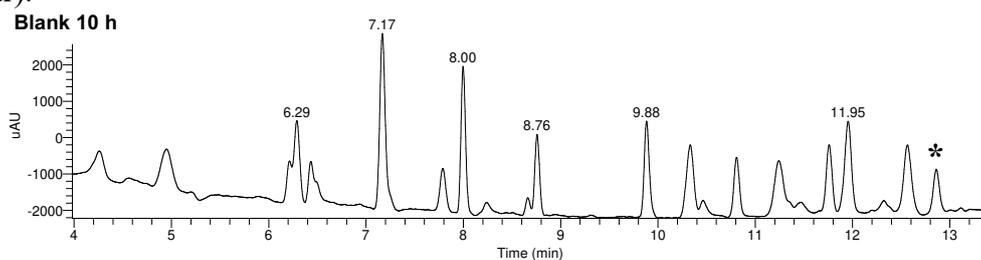
There is no standard cut-off for selecting amplified products, since this depends on the amount of protein used and on the product affinity for the template. Nonetheless, in general, low values of A% (calculated as reported) could be attributed to small integration differences rather than actual amplification. If low amplifications are seen, a systematic study on a single library to enhance amplification and understand the significance of low values can be performed.

## DCL1 – Acetate buffer

The equilibrium was reached after 10 h.

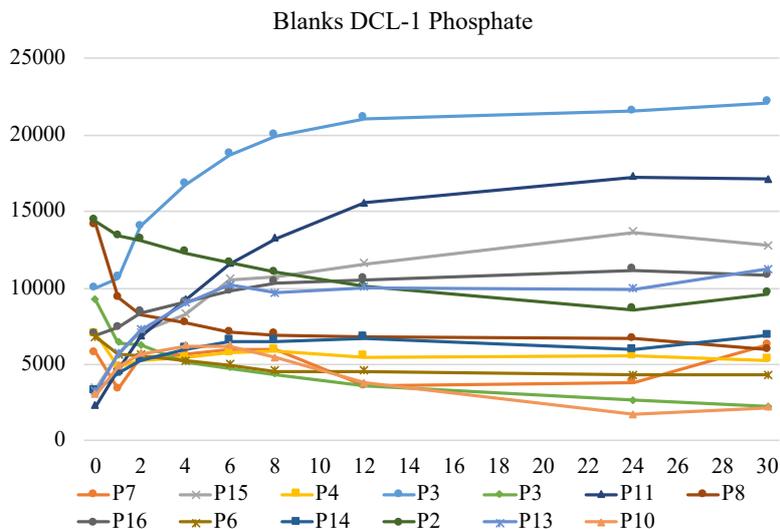


The amplification was measured at 10, 24 and 30 h, and significant values were found for 10 and 24 h for compound P10 (**1**): amplification 83% (10h), 107% (24h); Rt 12.86 min (indicated with a star).

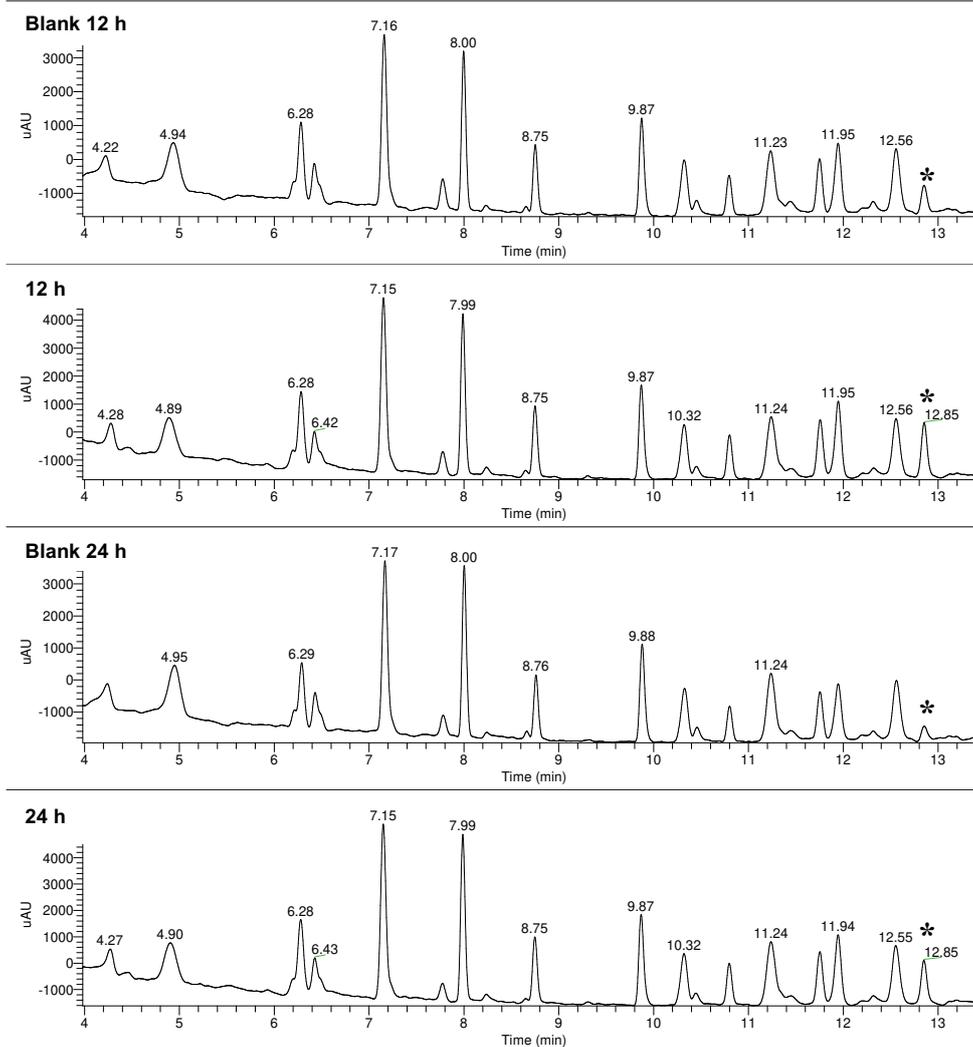


## DCL1 – Phosphate buffer

The equilibrium was reached after 12 h.

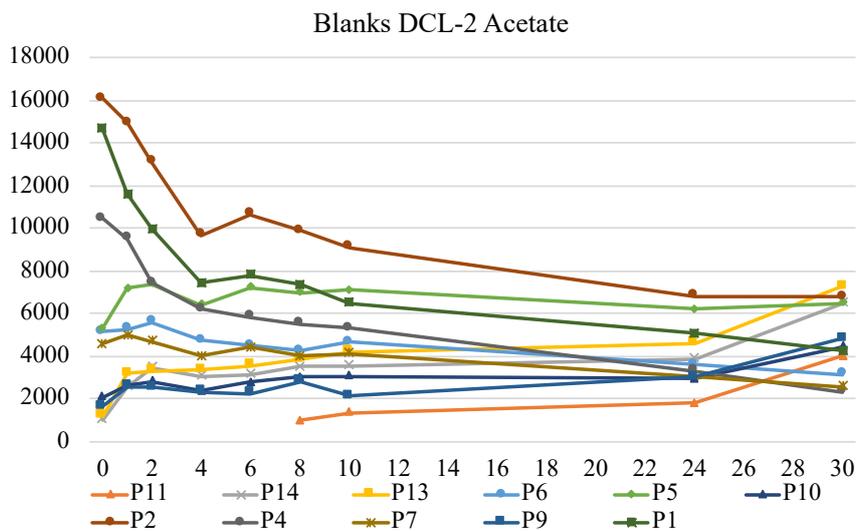


The amplification was measured at 12, 24 and 30 h, and significant values were found for 10 and 24 h for compound P10 (**1**): amplification 121% (12h), 328% (24h); Rt 12.85 min (indicated with a star).

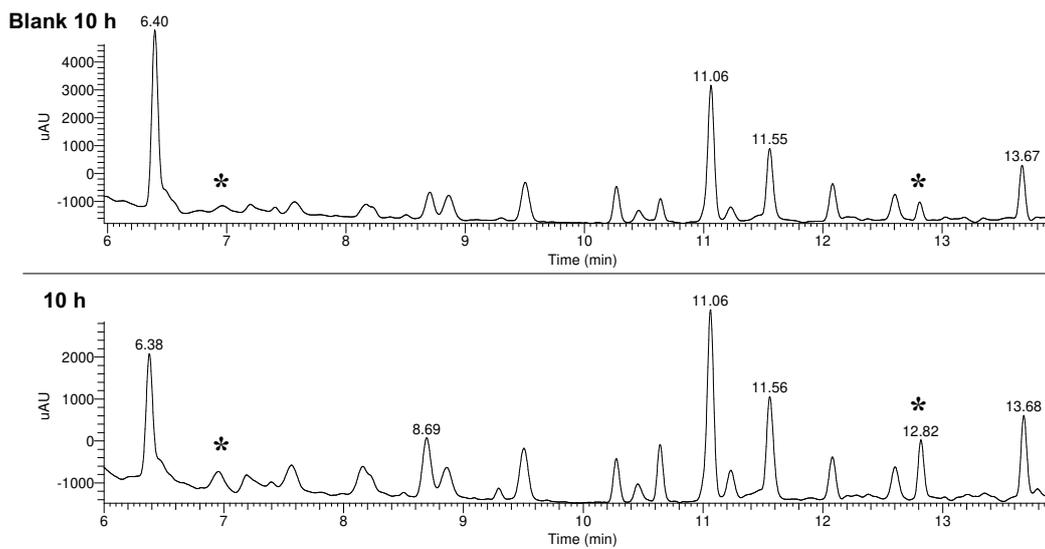


## DCL2 – Acetate Buffer

The equilibrium was reached after 10 h.

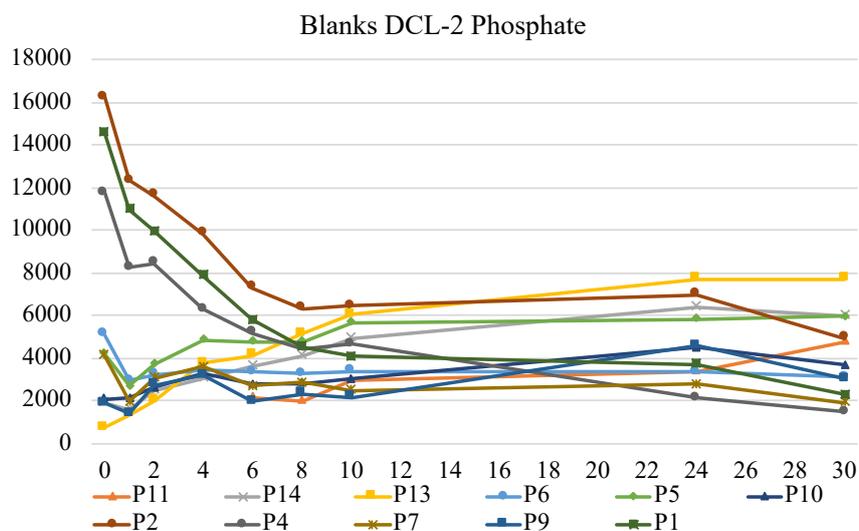


The amplification was measured at 10, 24 and 30 h, and significant values were found for 10 h for compounds P9 (**5**) and P11 (**6**): Hit P9 (**5**), amplification 121% (10h); Rt 12.82 min; Hit P11 (**6**), amplification 104% (10h); Rt 6.96 min (indicated with a star).

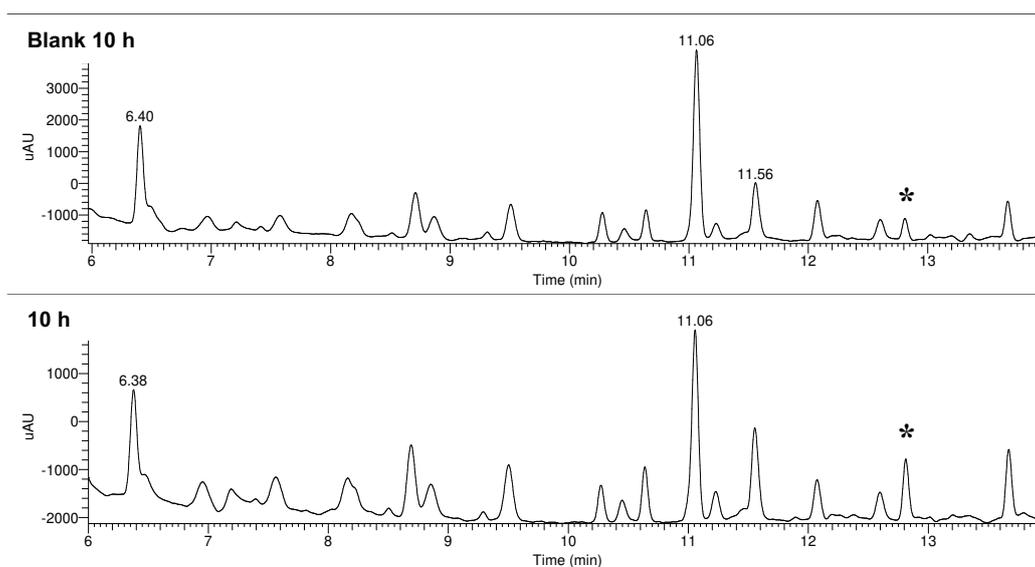


## DCL2 – Phosphate Buffer

The equilibrium was reached after 10 h.



The amplification was measured at 10, 24 and 30 h, and significant values were found for 10 h for compound P9 (**5**), amplification 89% (10h), Rt 12.82 min; (indicated with a star).

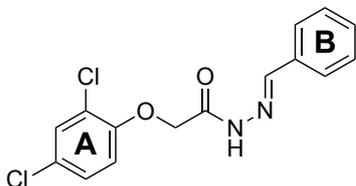
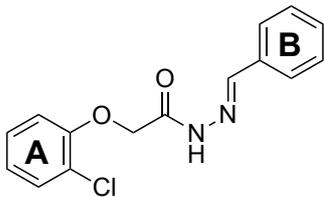


## General synthetic procedure

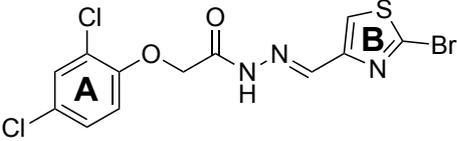
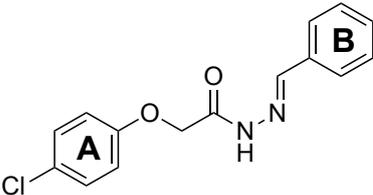
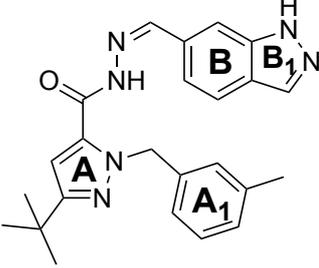
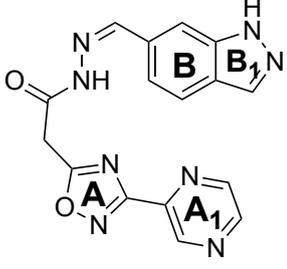
An equimolar amount of aldehyde (~50 mg) and hydrazide were dissolved in dry MeOH under a N<sub>2</sub> atmosphere, and the reaction was heated to reflux and stirred overnight. The final compounds often precipitated as white/yellowish solids. The solids were washed with cold methanol (3 x 3 mL) and centrifuged three times, finally the solvent was evaporated under reduced pressure. The purity and confirmation of Rt was assessed by HPLC, and the compounds were characterized by HRMS and <sup>1</sup>H NMR to confirm their identity (*Average yield 65-80%*).

## Characterization of synthesized compounds

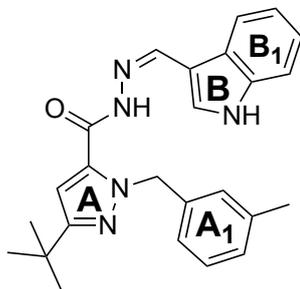
$^1\text{H}$  NMR spectra were measured in  $\text{DMSO-}d_6$  (D) or  $\text{CDCl}_3$  (Cl) on a Bruker Fourier 500 spectrometer (500 MHz). The chemical shifts are reported in parts per million (ppm) relative to the corresponding solvent peak. The coupling constants of the splitting patterns were reported in Hz and were indicated as singlet (s), doublet (d), triplet (t) and multiplet (m). Due to the presence of isomers for acylhydrazones, some of the signals are doubled (minor form in grey).

Compound	Structure	$^1\text{H}$ NMR (ppm) <sup>§§§§</sup>
1 (D)		11.69 (NH, s); 8.01 (CH, s); 7.72 (2H, 2,6 B, o/l); 7.59 (1H, 3 A s); 7.45 (3H, 3,4,5 B, d, J = 5.84); 7.34 (1H, 5 A, d, J = 9.35); 7.09 (1H, 6 A d, J = 9.35); 5.32 (2H, CH <sub>2</sub> , s).  11.65 (NH, s); 8.27 (CH, s); 7.71 (2H, 2,6 B, o/l); 7.63 (1H, 3 A, s); 7.45 (3H, 3,4,5 B, o/l); 7.40 (1H, 5 A, d, J = 9.35); 7.12 (1H, 6 A d, J = 9.35); 4.82 (2H, CH <sub>2</sub> , s). ratio ~3:1
2 (D)		11.64 (NH, s); 8.02 (CH, s); 7.72 (2H, 2,6 B, o/l); 7.45 (4H, 3,4,5 B, o/l; 3 A o/l); 7.27 (1H, 5 A, t, J = 7.72); 7.03 (1H, 6 A d, J = 8.37); 6.96 (1H, 4 A, t, J = 7.72); 5.32 (2H, CH <sub>2</sub> , s).  11.67 (NH, s); 8.28 (CH, s); 7.73 (2H, 2,6 B, o/l); 7.45 (4H, 3,4,5 B, o/l; 3 A o/l); 7.31 (1H, 5 A, t, J = 7.72); 7.09 (1H, 6 A d, J = 8.37); 7.01 (1H, 4 A, t, J = 7.72); 4.82 (2H, CH <sub>2</sub> , s). ratio ~2.5:1

<sup>§§§§</sup> o/l = overlap of peak chemical shift in the  $^1\text{H}$  NMR spectrum

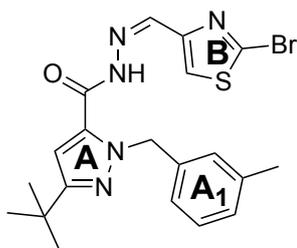
3 (D)		<p>11.78 (NH, s); 8.16 (5 B, s);        8.04(CH, s); 7.59 (1H, 3 A, d, J = 2.4);        7.33 (1H, 5 A, dd, J = 2.4; 8.7);        7.06 (1H, 6 A, d, J = 8.7);        5.27 (CH<sub>2</sub>, s).        11.75 (NH, s); 8.28 (5 B, s);        8.12 (CH, s); 7.62 (1H, 3 A, d, J = 2.4);        7.41 (1H, 5 A, dd, J = 2.4; 8.7);        7.1 (1H, 6 A, d, J = 8.7); 4.84 (CH<sub>2</sub>, s).        ratio ~3:1.</p>
4 (D)		<p>11.64 (NH, s); 8.02 (CH, s);        7.71-7.45 (5H, 3,4,5 B; 2,6 B, o/l);        7.33 (2H, 3,5 A, dt);        6.97 (2H, 2,6 A dt); 5.17 (2H, CH<sub>2</sub>, s).        11.59 (NH, s); 8.34 (CH, s);        7.71-7.45 (5H, 3,4,5 B; 2,6 B, o/l);        7.38 (2H, 3,5 A, dt);        7.03 (2H, 2, A dt); 4.69 (2H, CH<sub>2</sub>, s).        ratio ~1.8:1</p>
5 (D)		<p>13.28 (NH B<sub>1</sub>, s); 11.84 (NH, s);        8.50 (CH, s); 8.13 (1H, 2 B, d);        7.83 (1H, 5 B, d, J = 8.31);        7.79 (1H, 3 B<sub>1</sub>, s);        7.56 (1H, 6 B d, J = 8.31);        7.19 (1H, 5 A<sub>1</sub>, t, J = 7.72);        7.06 (1H, 6 A<sub>1</sub>, d, J = 7.72);        6.98 (1H, 2 A<sub>1</sub> s); 6.96 (1H, 5 A s);        6.90 (1H, 4 A<sub>1</sub>, d, J = 7.72);        5.68 (2H, CH<sub>2</sub>, s); 2.28 (3H, CH<sub>3</sub>, s),        1.31 (9H, (CH<sub>3</sub>)<sub>3</sub>, s).</p>
6 (D)		<p>13.26 (NH B<sub>1</sub>, s); 11.90 (NH, s);        9.29 (1H, s); 8.89 (2H, s); 8.15 (1H, s);        8.09 (1H, s); 7.74 (2H, m);        7.50 (1H, d);        6.96 (1H, 4 A, t, J = 7.72);        4.64 (2H, CH<sub>2</sub>, s).        13.28 (NH B<sub>1</sub>, s); 11.98 (NH, s);        8.90 (1H, s); 8.87 (2H, s); 8.36 (1H, s);        8.12 (1H, s); 7.85 (1H, s);        7.82 (1H, d); 7.57 (1H, d);        4.28 (2H, CH<sub>2</sub>, s). ratio ~3.3:1</p>

7 (D)



11.61 (NH, B, s); 11.48 (NH, s, o/l);  
8.55 (1H, s); 8.26 (1H, d);  
7.83 (1H, s); 7.44 (1H, d);  
7.23–7.13 (3H, m; o/l);  
7.06–6.98 (2H, m; o/l);  
6.94–6.90 (2H, m; o/l);  
5.70 (CH<sub>2</sub>, s); 2.26 (CH<sub>3</sub>, s);  
1.31 ((CH<sub>3</sub>)<sub>3</sub>, s);  
11.48 (NH, s, o/l); 8.07 (1H, d);  
7.81 (1H, s); 7.23–7.13 (3H, m; o/l);  
7.06–6.98 (2H, m; o/l);  
6.94–6.90 (2H, m; o/l);  
5.63 ((CH<sub>3</sub>)<sub>3</sub>, s); 2.17 (CH<sub>3</sub>, s);  
1.35 ((CH<sub>3</sub>)<sub>3</sub>, s);  
ratio ~5:1 (full assignment pending 2D  
spectra due to overlap)

8 (Cl)



13.19 (NH, s); 7.53 (1H, 5 B, s);  
7.32 (CH, s); 7.15 (1H, 2 A1, s);  
7.09 (2H, 4,6 A1, d, J = 6.93);  
6.96 (1H, d, 5 A1, J = 6.93);  
6.64 (1H, 5 A, s);  
5.74 (CH<sub>2</sub>, s); 2.23 (CH<sub>3</sub>, s);  
1.31 ((CH<sub>3</sub>)<sub>3</sub>, s).

## ***SECTION 2: Application of Dynamic Combinatorial Chemistry approach to obtain Bacterial Metalloprotease LasB ligands potentially acting as inhibitors***

### ***1. Introduction and Aim***

The present chapter describes the first application of Dynamic Combinatorial Chemistry performed during the research work of my thesis. Motivated to learn this strategy for ligand discovery, I participated in a project whose target is the extracellular metalloprotease (LasB), an anti-infective target studied in the group of Prof. Hirsch at the Helmholtz Institute for Pharmaceutical Research Saarland, where I spent three stays abroad. This research project aims at developing new compounds to address the increasing emergence of resistant bacteria, in particular Gram-negative species<sup>176,177</sup> such as *Pseudomonas aeruginosa*, one of the three most problematic pathogens on the WHO priority list.<sup>176</sup> In this study field, a promising new approach is to target bacterial virulence to disarm pathogens rather than killing them. *Pseudomonas aeruginosa* produces several virulence factors, which contribute to disease progression, some of which represent interesting anti-infective targets.<sup>178,179</sup> LasB is a major virulence factor, with a key role in the pathogenicity of *P. aeruginosa*. This enzyme is a zinc-metalloprotease with high structural similarity to thermolysin.<sup>180</sup> Among its main functions is the cleavage of connective tissue components such as elastin and collagen, thus allowing the bacteria to colonize in the host.<sup>181</sup> Moreover, LasB allows *P. aeruginosa* to evade the human immune response by cleaving selected immunoglobulins and cytokines.<sup>182</sup> For this reason, this protease represents an attractive anti-infective target, also due to its extracellular localization which facilitates drug discovery as permeation of the Gram-negative cell wall is not needed.<sup>183</sup> LasB belongs to the thermolysin (M4) family of enzymes and within this class, the rather open active site cleft of the protease can close through conformational changes upon binding of inhibitor.<sup>184</sup> To date, several zinc-chelating inhibitors of LasB have been described,<sup>185,186</sup> including compounds bearing a mercaptoacetamide, a thiol, or a hydroxamic acid motif attached either to small peptides<sup>187</sup> or to aniline.<sup>188</sup> Additional derivatives are currently under study in the group.

The application of a dynamic combinatorial chemistry approach could allow for the discovery of new scaffolds able to bind LasB which could be improved to act as enzyme inhibitors. Different groups and moieties with the potential for zinc-chelation were included in the building-block libraries. Moreover, given the similarity of binding site with other collagenases, the same approach was carried out in parallel with Dr Jelena Konstantinovic (in the same

research group) on the collagenase ColH from the Gram-positive pathogen *Clostridium histolyticum*.<sup>189</sup> The same four DCL libraries were utilized for experiments, and the hits are currently being evaluated for inhibition of both enzymes. In case of lack of inhibition, all hits will be tested to assess binding to both targets.

## **2. Protein-templated DCC to obtain new LasB ligands**

### ***Establishment of the experimental conditions***

For LasB, experimental conditions were established by applying the protocol of the LasB enzyme-activity assay, where the degradation of a fluorogenic substrate is monitored over the experimental timeframes. With high activity of the enzyme, the fluorescence increases over time, while if the enzyme is inhibited, or has lost activity due to a loss in stability, the fluorescence remains constant during the experiment; plotting the fluorescence against experimental time gives a measure of activity. This way, by testing the protein stored in different buffers, for different periods of time, it is possible to plot slopes against storage time, building a curve which shows how the activity changes during the period of storage.

The same 16 buffers plus water were exploited to establish the experimental conditions, and the activity was followed for 4 days for all buffers. Selected buffers were studied until day 8, finally allowing to choose phosphate buffer pH 6.

### ***Optimization of the pt-DCC experiment***

A total of six pt-DCC experiments were performed to progressively solve issues related to low amplification, slow equilibration and ill choice of building blocks. The final conditions chosen regarding protein concentration (to afford determination of significant amplification) and equilibrium (reached in reasonable timeframes) involve the use of 40% protein (over .

As for building blocks, the most sensitive component was identified in the aldehydes; in fact, the use of deactivated aldehydes (*i.e.*, in this specific case, few aldehydes were tried directly linked to a EDG) did not afford the corresponding acylhydrazones in the dynamic combinatorial library, supposedly due to their low reactivity, not compatible with the timeframes of the pt-DCC experiment. This should be taken into consideration when building a DCL, to avoid false negatives. Another issue was found with building blocks sensitive to the drastic pH changes applied to freeze the reaction at each sampling. Since detection is done by mass, if the final compounds have a different mass than expected, due to instability of one portion of the molecule (*i.e.*, in this specific case, a catechol-like portion), the impossibility to unequivocally identify the products and thus the related amplification, could once again lead to false negatives. This issue was identified (after noticing a whole series of undetected products) by subjecting

the suspected building blocks to the same treatment done at each sampling followed by injection in the HPLC-UV-MS, and comparison with the profile of the building blocks as such.

Such building blocks were thus removed from the libraries and replaced. The final compounds bearing the same portion could still be synthesized and tested, if considered necessary, without involving them in a pt-DCC experiment.

### **Dynamic combinatorial libraries**

Four dynamic combinatorial libraries were designed to exploit the chemistry of acylhydrazones, based on commercially available aldehydes and hydrazides (the structures are reported in figure 39). The blank and pt-DCC experiments were conducted in parallel and monitored *via* HPLC-UV-MS upon protein precipitation (by adding ACN) and pH-induced “freezing” of the reaction. As reported in paragraph 3.4.4, this is essential to separate, identify and quantify the amount of each DCL member formed unequivocally and without any artifacts to determine both equilibrium and pt-amplification.

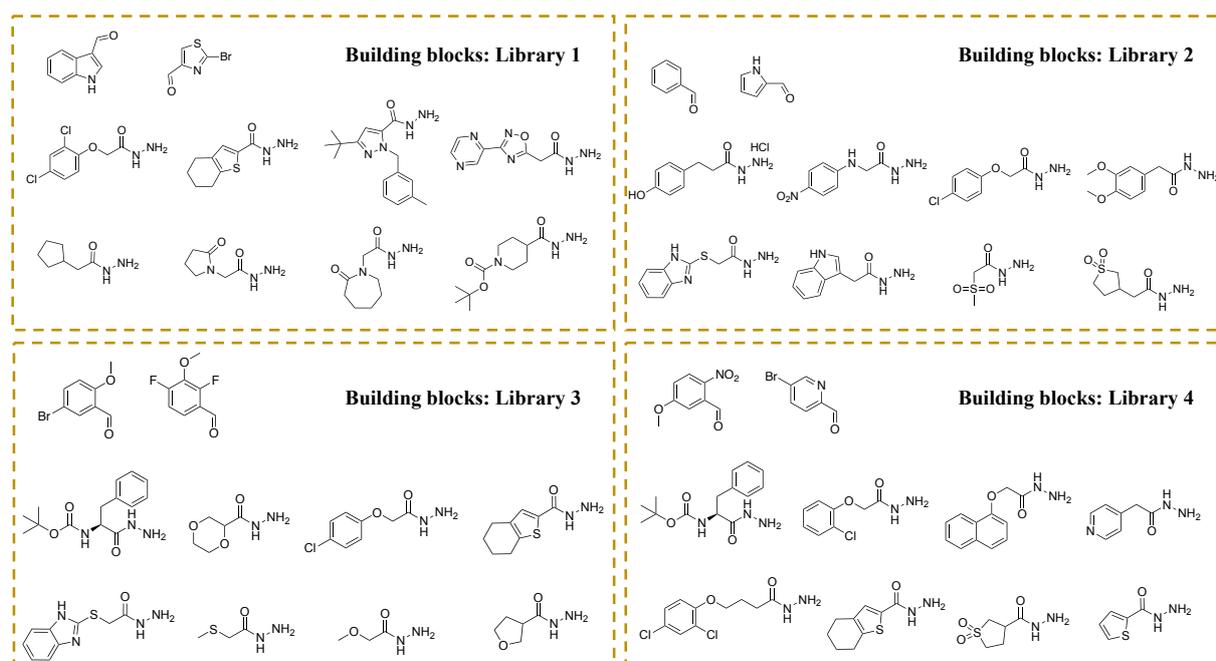


Figure 39. Building-block libraries subjected to the pt-DCC study for target protein LasB.

Within the four libraries, five hits were selected from the first three libraries, while no significant amplification was determined for library 4. A pool of derivatives were chosen for synthesis and testing (also after comparison with the results obtain by Dr Konstantinovic for the corresponding DCC experiments with ColH), for a total of ~15 final compounds (only the eleven structures of the LasB pt-DCC compounds are reported in figure 40).

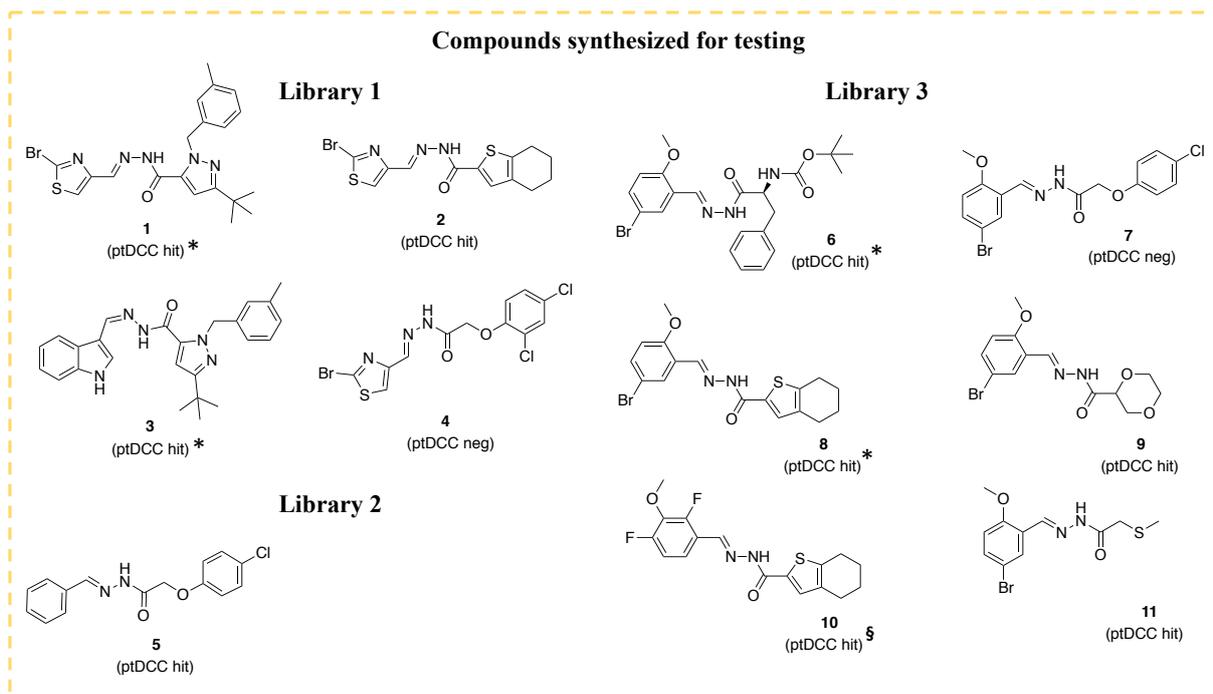


Figure 40. Structures of the compounds synthesized for evaluation against LasB. \*The correct amplification of compounds 1 and 3, and 6 and 8 could not be univocally calculated due to peak overlap within the same experiment; thus both couples were synthesized for testing. §Compound 10 showed lower amplification compared to pt-DCC hits 9 and 11.

### 3. Conclusion

A dynamic combinatorial chemistry approach was exploited to discover new scaffolds able to bind LasB, a zinc-metalloprotease involved in the virulence of Gram negative *P. aeruginosa*, and potentially act as enzyme inhibitors. Given the nature of the LasB active site which we are targeting, different groups and moieties with the potential for zinc-chelation were included in the building block libraries. Four DCL were designed and subjected to pt-DCC experiments, finally affording five pt-DCC hits that were synthesized for biological testing. Along with these, additional derivatives and compounds that were not amplified serving as negative controls were synthesized, leading to a total of ~15 final compounds. Given the similarity of binding site with other collagenases, the same approach was carried out in parallel by Dr Jelena Konstantinovic (in the same research group – DDOP at HIPS) on Gram-positive collagenase ColH.<sup>189</sup> The pooled hits, derivatives and negatives synthesized are currently being evaluated for inhibition of both enzymes. In case of lack of inhibition, all hits will be tested to assess binding to both targets, to elucidate the relevance of the pt-DCC approach in the discovery of new compounds as LasB ligands able to act as enzyme inhibitors. Positive results will open the way to the development of novel compounds useful as anti-infective agents according to a novel, straightforward strategy.

## ***Experimental section 2***

### ***Experimental conditions for the pt-DCC experiment***

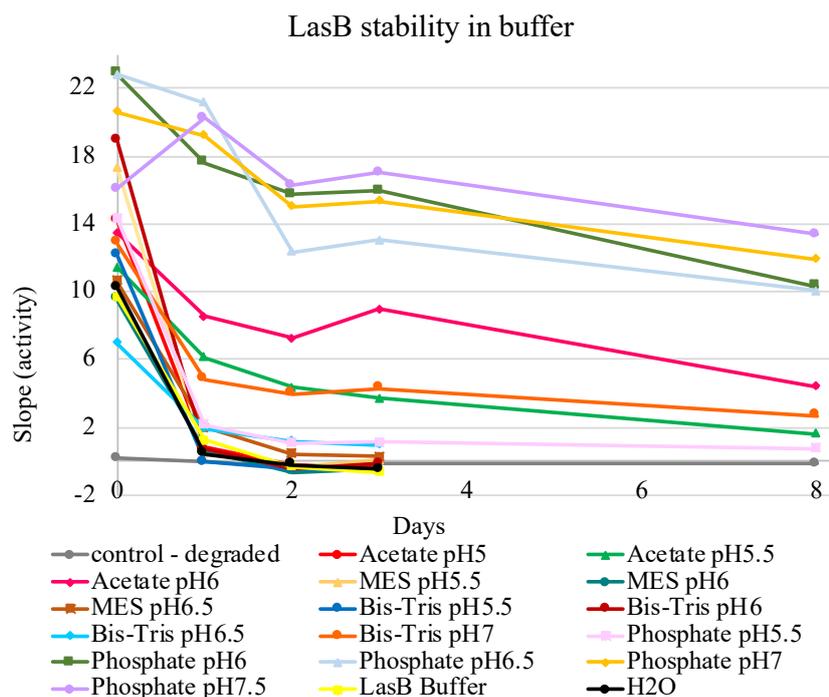
#### **Buffer selection by Enzyme Activity Assay**

A 96-well plate, containing seventeen 10  $\mu$ M solutions of protein in sixteen buffers and H<sub>2</sub>O and 5% DMSO, (each solution prepared in duplicate) was stored at room temperature for four days, and sampled every day (day 0 to day 3) for activity assays. Blank and negative control wells were added to each assay run. The solutions were sampled daily and transferred in a 384-well plate prepared as described in the procedure.

The buffers used (all containing 150 mM NaCl) are reported in the table:

<b>Buffer Type</b>	<b>pH</b>
Acetate 10 mM	5–5.5–6
MES 20 mM	5.5–6–6.5
Bis-TRIS 20 mM	5.5–6–6.5–7
Phosphate 50 mM	5.5–6–6.5–7–7.5

Purified elastase prepared according to the published procedure was used.<sup>181</sup> The fluorogenic substrate 2-aminobenzoyl-Ala-Gly-Leu-Ala-4-nitrobenzylamide was purchased from Peptides International (Louisville, KY, USA). Fluorescence intensity was measured for 60 min at 37 °C in black 384-well microtiter plates (Greiner BioOne, Kremsmünster, Austria) using a CLARIOstar microplate reader (BMG Labtech, Ortenberg, Germany) with an excitation wavelength of  $340 \pm 15$  nm and an emission wavelength of  $415 \pm 20$  nm. The assay was performed in a final volume of 50  $\mu$ L of assay buffer (50 mM Tris, pH 7.2, 2.5 mM CaCl<sub>2</sub>, 5% DMSO) containing LasB at a final concentration of 0.3 nM and the substrate at 150  $\mu$ M. Before substrate addition, the enzyme was preincubated with the buffers for 15 min at 37 °C. After blank subtraction the fluorescence slope recorded for each sample (activity) was plotted against storage time.



### *pt-DCC experiment*

#### Experimental setup

Experiment 0 was performed without aniline. All other components of the reaction mixture were added as for experiments 1 (including the use of the same DCL).

The reaction mixtures were prepared as reported in the following tables for experiments 1, 2 and 3a<sup>\*\*\*\*</sup> (on the respective DCLs)

Blank	Amount	Final concentration	Blank reaction	pt-DCC
Buffer	Premixed	-		
ZnCl <sub>2</sub>	190 μL	~10 μM		
Aldehydes		100 μM		
Hydrazides	Premixed*	200 μM		
Aniline	10 μL	1 mM		
DMSO		5%		
<i>Total</i>	<i>200 μL</i>			

\*\*\*\* DCL 3a was originally built in place of DCL 3, and the building blocks included a catechol-aldehyde (3,4-dihydroxybenzaldehyde). The latter was then removed (producing building-block library 3) to avoid NaOH-induced degradation of the catechol portion (noticed experimentally).

pt-DCC	Amount	Final concentration
Buffer	Premixed	-
ZnCl <sub>2</sub>	123.88 $\mu$ L	$\sim$ 10 $\mu$ M
Aldehydes		100 $\mu$ M
Hydrazides	Premixed*	200 $\mu$ M
Aniline	10 $\mu$ L	1 mM
DMSO		5%
LasB (121 $\mu$ M)	66.12 $\mu$ L	40 $\mu$ M
<i>Total</i>	<i>200 <math>\mu</math>L</i>	

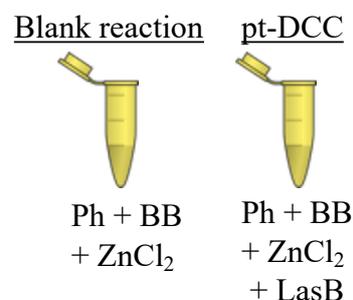
Pre-mixed solution\*.#

BB+aniline in DMSO	Amount	Final concentration
Aldehydes (100 mM)	1 $\mu$ L x 2	2 mM
Hydrazides (100 mM)	2 $\mu$ L x 8	4 mM
Aniline (1 M)	1 $\mu$ L	20 mM
DMSO	31 $\mu$ L	-
<i>Total</i>	<i>50 <math>\mu</math>L</i>	

# The preparation of the pre-mixed solution followed a precise order to avoid triggering the reaction: 1) Hydrazides; 2) Aldehydes; 3) Aniline. The premixed solution was then added last to the reaction mixture. The four reactions were then left rotating for 36h and sampled at regular intervals starting from hour 0. The samples for HPLC-UV-MS monitoring were prepared as follows: 5  $\mu$ L of each sample were added to a Eppendorf containing 44  $\mu$ L ACN analytical grade (precipitates protein) and 1  $\mu$ L NaOH 1 M (stops reaction reversibility). The vial was vortexed for 10 s and centrifuged for 5 min x 14000 RPM, 4°C, after which the supernatant was transferred to a HPLC vial for analysis. All analysis were performed on the same day at the end of the experiment. All samples previously collected were stored at 4°C.

The reaction mixtures were prepared as reported in the following tables for experiments 3 and 4 (on the respective DCLs)

Blank	Amount	Final concentration
Buffer	Premixed	-
ZnCl <sub>2</sub>	190 $\mu$ L	$\sim$ 10 $\mu$ M
Aldehydes		100 $\mu$ M
Hydrazides	Premixed*	200 $\mu$ M
Aniline	10 $\mu$ L	1 mM
DMSO		5%
<i>Total</i>	<i>200 <math>\mu</math>L</i>	



\*The preparation of the premixed solution remained unvaried

pt-DCC	Amount	Final concentration
Buffer	Premixed	-
ZnCl <sub>2</sub>	173.47 $\mu$ L	$\sim$ 10 $\mu$ M
Aldehydes		100 $\mu$ M
Hydrazides	Premixed*	200 $\mu$ M
Aniline	10 $\mu$ L	1 mM
DMSO		5%
LasB (121 $\mu$ M)	16.53 $\mu$ L	10 $\mu$ M
<i>Total</i>	<i>200 <math>\mu</math>L</i>	

### HPLC method

HPLC-UV-MS analyses were performed on a ThermoScientific Dionex Ultimate 3000 UHPLC System coupled to a ThermoScientific Q Exactive Focus with an electrospray ion source. An Acquity Waters Column (BEH, C8 1.7  $\mu$ m, 2.1 x 150 mm, Waters, Germany) equipped with a VanGuard Pre-Column (BEH C8, 5 x 2.1 mm, 1,7  $\mu$ m, Waters, Germany) was used for separation. The mass spectrum was measured in positive mode in a range from 100–700 m/z. At a flow rate of 0.250 mL/min, the elution was performed in a gradient as reported in the table: (A = ACN + 0.1% HCOOH; B = H<sub>2</sub>O + 0.1% HCOOH)

min	%B	%A
0	90	10
1	90	10
17	5	95
18.5	5	95
18.6	90	10
20.5	90	10

After each run, a 8-minute wash was performed on the column.

### Determination of equilibrium and quantification of product amplification

Products were identified univocally by high resolution mass and the area of the attributed peak was integrated through the Excalibur software at the selected wavelength of 272 nm.

To assess that equilibrium was reached, the area of each peak was plotted against time for each blank sample collected. Once the area reaches a value stable in time, the equilibrium is considered reached and the amplification can be investigated in pt-DCC samples.

Amplification is calculated for each peak by attributing the arbitrary value of 100% to the area in the blank. According to the following proportion for each peak:

Area<sub>DCC</sub> : x = Area<sub>blank</sub> : 100. Then,  $x = (\text{Area}_{\text{DCC}} * 100) / \text{Area}_{\text{blank}}$ .

The amplification % is then calculated as  $A\% = x - 100$ .

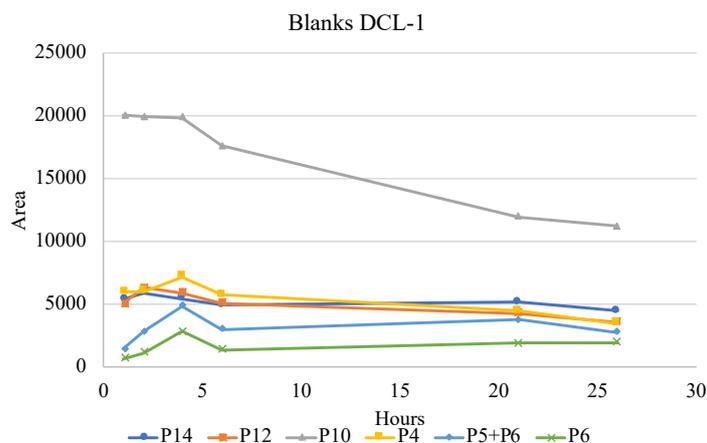
There is no standard cut-off for selecting amplified products, since this depends on the amount of protein used and on the product affinity for the template. Nonetheless, in general, low values of A% (calculated as reported) could be attributed to small integration differences rather than actual amplification. If low amplifications are seen, a systematic study on a single library to enhance amplification and understand the significance of low values can be performed.

### **Experiment 0 (DCL1)**

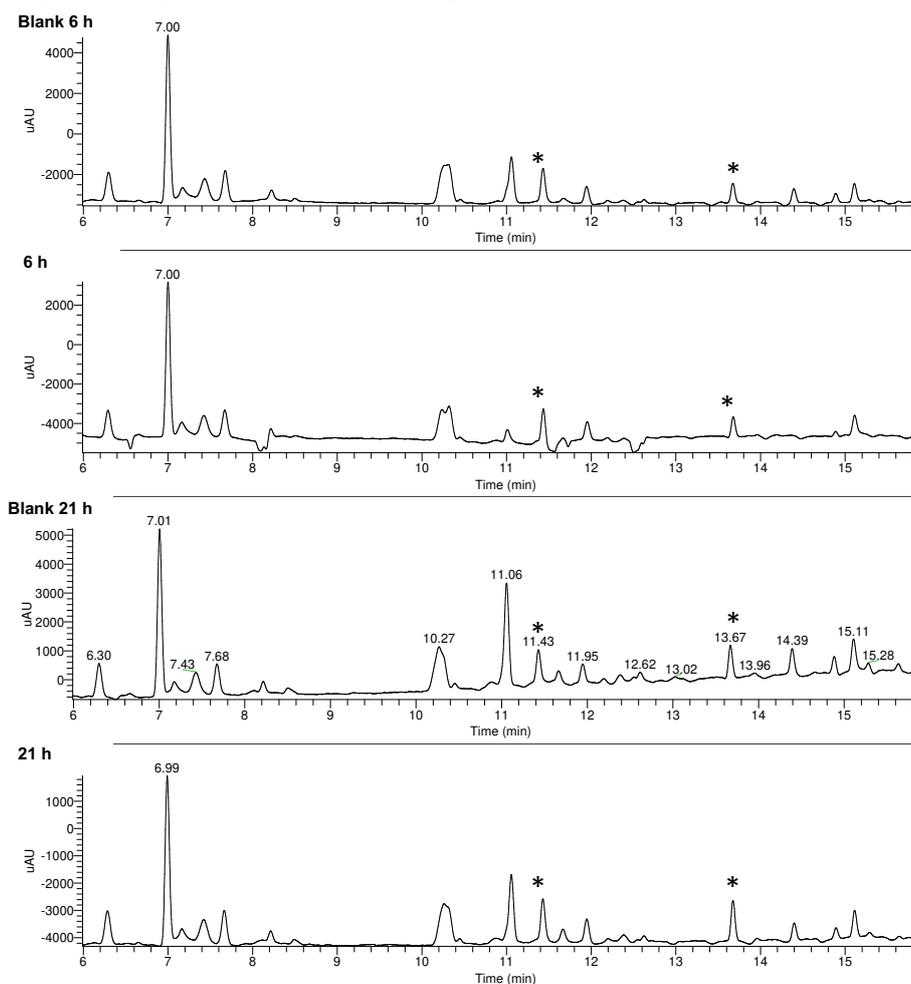
Given the demonstrated protein stability over 8 days in the chosen buffer, the first experiment was tried without aniline. The reactions (blank and pt-DCC) were monitored at time 0 and then once daily for 8 days. The area of each product peak was plotted against time to assess the moment equilibrium was reached. While for some products this appears to be after two days (~50 h), several others still showed variations after 3 days (~70 h) and equilibrium could not be ascertained by the end of the experiment. To reduce timeframes, all the following experiments included the use of 1 mM aniline.

## Experiment 1 (DCL1)

The equilibrium was reached after 6 h.



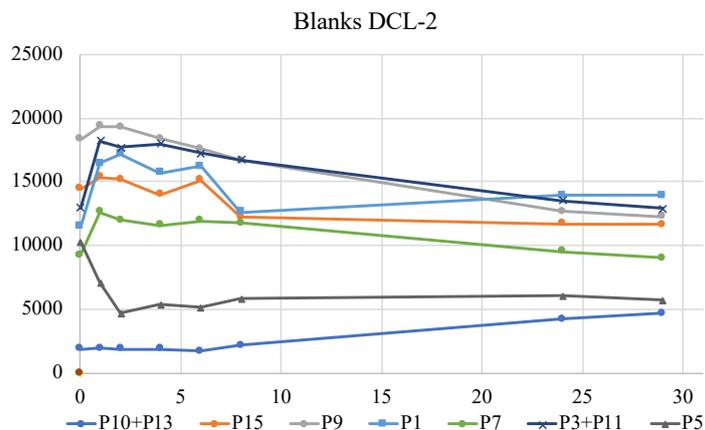
The amplification was measured at 6, 21 and 26 h, and significant values were found for for compound P4 (2): amplification 32% (6h), ~50% (21 and 26h) Rt 11.43 min (indicated with a star). Compounds P5 (3) and P6<sup>††††</sup> (1) showed peak overlap at Rt 13.67 min (indicated with a star); the amplification of the peak was 45% (6h) and ~35% (21 and 26h) both were synthesized as potential pt-DCC hits for testing.



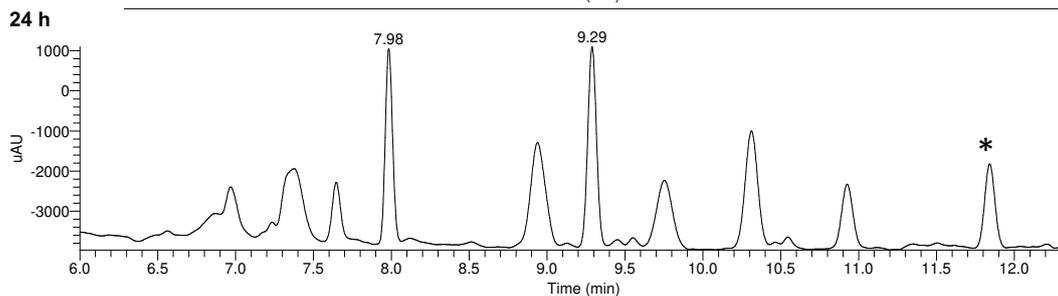
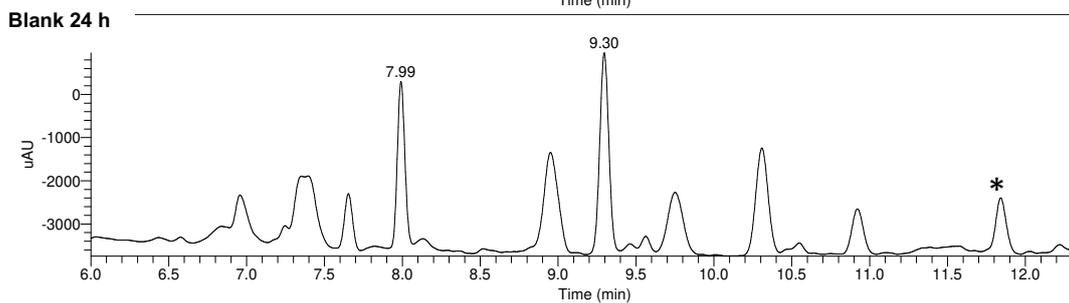
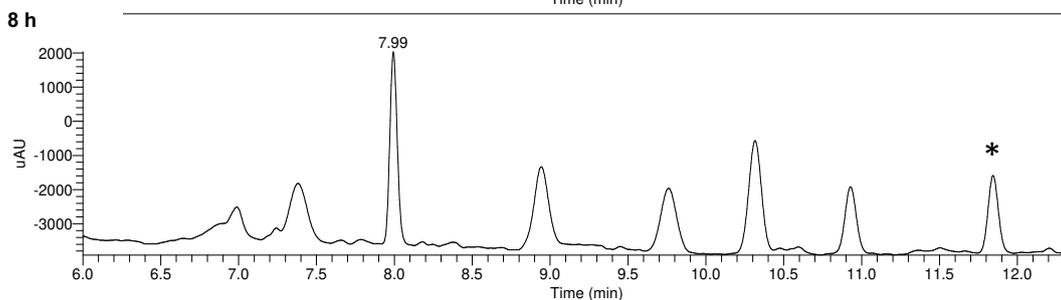
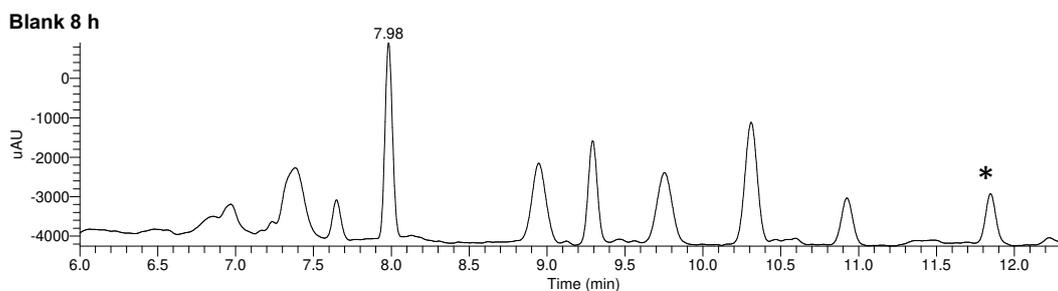
<sup>††††</sup> P6 shows a secondary peak (acylhydrazone isomer) at 14.88 min for which no amplification was observed.

## Experiment 2 (DCL2)

The equilibrium was reached after 8h.

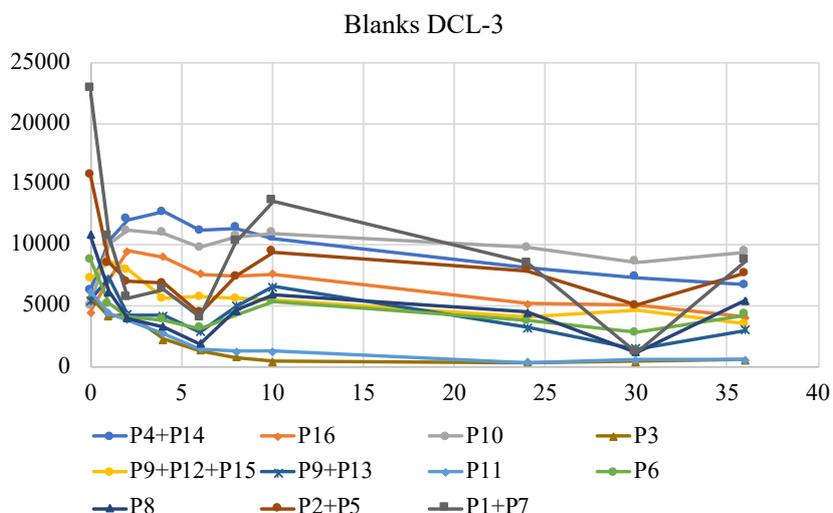


The amplification was measured at 8 and 24 h, and significant values were found for compound P5 (5): amplification 75% (8h), 46% (24h); Rt 11.84 min (indicated with a star).



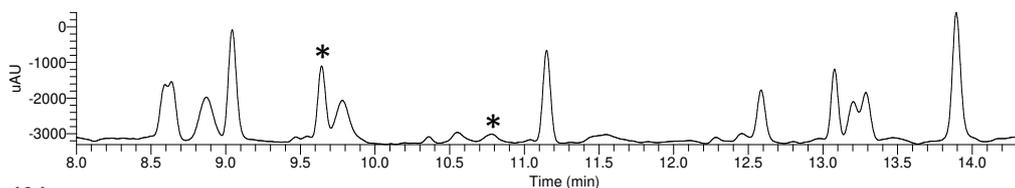
### Experiment 3 (DCL3)

The equilibrium was reached after 10 h.

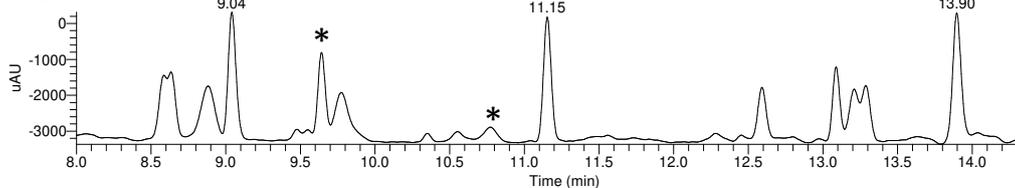


The amplification was measured at 10 and 24 h, and significant values were found for compound P3 (**9**) and P11 (**11**): amplification 211% (10h) and 81 and 179% (10 and 24h), respectively; Rt 9.50 min and 10.77 min, respectively (indicated with a star). Lower amplification was observed for compound P8 (**10**) 80% (24h), Rt 13.08 min (indicated with §). Compounds P1 and P7 (**6** and **8**) showed peak overlap at Rt 13.89 min, amplified 101% (24h, indicated with #); both compounds were synthesized for testing.

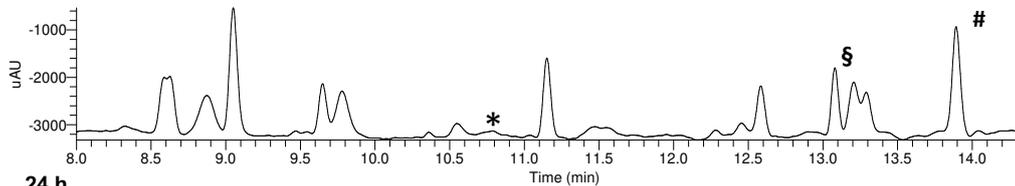
Blank 10 h



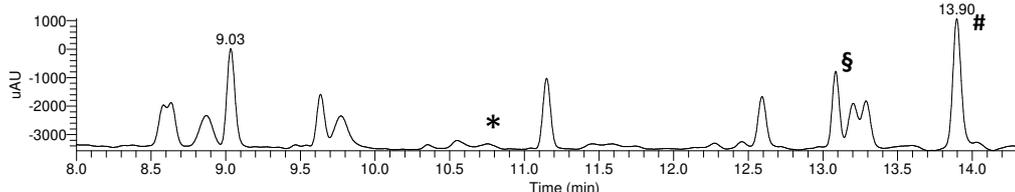
10 h



Blank 24 h



24 h



## Experiment 4 (DCL4)

No significant amplification was observed for this library.

### General synthetic procedure

An equimolar amount of aldehyde (~50 mg) and hydrazide were dissolved in dry MeOH under a N<sub>2</sub> atmosphere, and the reaction was heated to reflux and stirred overnight. The final compounds often precipitated as white/yellowish solids. The solids were washed with cold methanol (3 x 3 mL) and centrifuged three times, finally the solvent was evaporated under reduced pressure. The purity and confirmation of Rt was assessed by HPLC, and the compounds were characterized by HRMS and <sup>1</sup>H NMR to confirm their identity (*Average yield 65-80%*).

### Characterization of synthesized compounds

<sup>1</sup>H NMR spectra were measured in DMSO-*d*<sub>6</sub> (D) or CDCl<sub>3</sub> (Cl) on a Bruker Fourier 500 spectrometer (500 MHz). The chemical shifts are reported in parts per million (ppm) relative to the corresponding solvent peak. Due to the presence of isomers for acylhydrazones, some of the signals are doubled (minor form in grey) and/or partially overlapped (o/l).

**1** (Cl) 13.19 (1H, s); 7.53 (1H, s); 7.32 (1H, s); 7.15 (1H, s); 7.09 (2H, d); 6.96 (1H, s); 6.64 (1H, s); 5.74 (2H, s); 2.23 (3H, s); 1.31 (9H, s).

**2** (D) 11.82 (1H, s); 8.12 (1H, s); 7.90 (2H, s); 2.78 (2H, t); 2.61 (2H, t); 1.81–1.74 (4H, m).

**3** (D) 11.61 (NH, B, s); 11.48 (NH, s, o/l); 8.55 (1H, s); 8.26 (1H, d);

7.83 (1H, s); 7.44 (1H, d); 7.23–7.13 (3H, m; o/l); 7.06–6.98 (2H, m; o/l); 6.94–6.90 (2H, m; o/l); 5.70 (CH<sub>2</sub>, s); 2.26 (CH<sub>3</sub>, s); 1.31 ((CH<sub>3</sub>)<sub>3</sub>, s);

11.48 (NH, s, o/l); 8.07 (1H, d); 7.81 (1H, s); 7.23–7.13 (3H, m; o/l); 7.06–6.98 (2H, m; o/l); 6.94–6.90 (2H, m; o/l); 5.63 ((CH<sub>3</sub>)<sub>3</sub>, s); 2.17 (CH<sub>3</sub>, s); 1.35 ((CH<sub>3</sub>)<sub>3</sub>, s);

ratio ~5:1

**4** (D) 11.78 (1H, s); 8.16 (1H, s); 8.04(1H, s); 7.59 (1H, d); 7.33 (1H, dd); 7.06 (1H, d); 5.27 (2H, s).

11.75 (1H, s); 8.28 (1H, s); 8.12 (1H, s); 7.62 (1H, d); 7.41 (1H, dd); 7.1 (1H, d); 4.84 (2H, s).

ratio ~3:1.

**5** (D) 11.64 (1H, s); 8.02 (1H, s); 7.71–7.45 (5H, m); 7.33 (2H, dt); 6.97 (2H, dt); 5.17 (2H, s).

11.59 (1H, s); 8.34 (1H, s); 7.71–7.45 (5H, m); 7.38 (2H, dt); 7.03 (2H, dt); 4.69 (2H, s).

ratio ~1.8:1

**6** (D) 11.43 (1H, s); 8.25 (1H, s); 7.96 (1H, d); 7.58 (1H, m; o/l); 7.33–7.28 (4H m; o/l); 7.21–7.19 (2H, m); 7.1 (1H, dd, o/l); 4.92 (1H, td); 3.86 (3H, s); 2.97–2.75 (2H, m; o/l); 1.31 (9H, s).

11.68 (1H, s); 8.46 (1H, s); 7.87 (1H, d); 7.58 (1H, m; o/l); 7.33–7.28 (4H m; o/l); 7.21–7.19 (2H, m); 7.1 (1H, dd, o/l); 4.16 (1H, m) 3.87 (3H, s); 2.97–2.75 (2H, m; o/l); 1.32 (9H, s).

ratio ~1.7:1

**7 (D)** 11.66 (1H, s); 8.26 (1H, s); 7.95 (1H, d); 7.35 (3H, dd); 6.99 (3H, dd); 5.19 (2H, s); 3.86 (3H, s).

11.73 (1H, s); 8.61 (1H, s); 7.88 (1H, d); 7.57 (3H, td); 7.1 (3H, t); 4.67 (2H, s); 3.87 (3H, s).

ratio ~2:1

**8 (D)** 11.87 (1H, s); 8.67 (1H, s); 7.91 (1H, s); 7.58 (2H, dd); 7.12 (1H, d); 3.88 (3H, s); 2.77 (2H, s); 2.60 (2H, t); 1.8–1.74 (4H, m).

**9 (D)** 11.54 (1H, s); 8.69 (1H, s); 7.86 (1H, d); 7.58 (1H, dd); 7.09 (1H, d); 4.20 (1H, dd); 3.90 (1H, dd); 3.87 (3H, s); 3.72–3.66 (2H, m); 3.58–3.52 (2H, m).

**10 (D)** 11.86 (1H, s); 8.56 (1H, s); 7.67 (2H, d); 7.29 (1H, m); 3.96 (3H, s); 2.79 (2H, s); 2.61 (2H,t); 1.80–1.74 (4H, m).

**11 (D)** 11.46 (1H, s); 8.25 (1H, s); 7.87 (1H, m; o/l); 7.59–7.55 (1H, m; o/l); 7.09 (1H, t; o/l); 3.84 (3H, s); 3.60 (2H, s); 2.15 (3H, s).

11.65 (1H, s); 8.46 (1H, s); 7.87 (1H, m; o/l); 7.59–7.55 (1H, m; o/l); 7.09 (1H, t; o/l); 3.86 (3H, s); 3.17 (2H, s); 2.16 (3H, s).

ratio ~1.4:1.

## Appendix

### I. References

- (1) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discov.* **2006**. <https://doi.org/10.1038/nrd2199>.
- (2) Ivanov, A. A.; Khuri, F. R.; Fu, H. Targeting Protein-Protein Interactions as an Anticancer Strategy. *Trends in Pharmacological Sciences.* 2013. <https://doi.org/10.1016/j.tips.2013.04.007>.
- (3) Dunker, A. K.; Oldfield, C. J.; Meng, J.; Romero, P.; Yang, J. Y.; Chen, J. W.; Vacic, V.; Obradovic, Z.; Uversky, V. N. The Unfoldomics Decade: An Update on Intrinsically Disordered Proteins. In *BMC Genomics*; 2008. <https://doi.org/10.1186/1471-2164-9-S2-S1>.
- (4) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat. Rev. Drug Discov.* **2018**, *17* (8), 547–558. <https://doi.org/10.1038/nrd.2018.93>.
- (5) Disney, M. D. Targeting RNA with Small Molecules To Capture Opportunities at the Intersection of Chemistry, Biology, and Medicine. *Journal of the American Chemical Society.* 2019. <https://doi.org/10.1021/jacs.8b13419>.
- (6) Connelly, C. M.; Moon, M. H.; Schneekloth, J. S. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chemical Biology.* 2016. <https://doi.org/10.1016/j.chembiol.2016.05.021>.
- (7) Glisovic, T.; Bachorik, J. L.; Yong, J.; Dreyfuss, G. RNA-Binding Proteins and Post-Transcriptional Gene Regulation. *FEBS Lett.* **2008**, *582* (14), 1977–1986. <https://doi.org/10.1016/j.febslet.2008.03.004>.
- (8) Moore, M. J. From Birth to Death: The Complex Lives of Eukaryotic MRNAs. *Science.* 2005. <https://doi.org/10.1126/science.1111443>.
- (9) Zhou, H. L.; Mangelsdorf, M.; Liu, J. H.; Zhu, L.; Wu, J. Y. RNA-Binding Proteins in Neurological Diseases. *Science China Life Sciences.* 2014. <https://doi.org/10.1007/s11427-014-4647-9>.
- (10) Vanderweyde, T. E.; Wolozin, B. RNA Binding Proteins in Health and Disease. In *Neuroimmune Pharmacology*; 2016. [https://doi.org/10.1007/978-3-319-44022-4\\_19](https://doi.org/10.1007/978-3-319-44022-4_19).
- (11) Hentze, M. W.; Castello, A.; Schwarzl, T.; Preiss, T. A Brave New World of RNA-Binding Proteins. *Nature Reviews Molecular Cell Biology.* 2018. <https://doi.org/10.1038/nrm.2017.130>.
- (12) Pascale, A.; Govoni, S. The Complex World of Post-Transcriptional Mechanisms: Is Their Deregulation a Common Link for Diseases? Focus on ELAV-like RNA-Binding Proteins. *Cellular and Molecular Life Sciences.* 2012. <https://doi.org/10.1007/s00018-011-0810-7>.
- (13) Balcerak, A.; Trebinska-Stryjewska, A.; Konopinski, R.; Wakula, M.; Grzybowska, E. A. RNA-Protein Interactions: Disorder, Moonlighting and Junk Contribute to Eukaryotic Complexity. *Open Biol.* **2019**, *9* (6). <https://doi.org/10.1098/rsob.190096>.
- (14) Yeo, G.; Holste, D.; Kreiman, G.; Burge, C. B. Variation in Alternative Splicing across Human

- Tissues. *Genome Biol.* **2004**.
- (15) Darnell, R. B.; Posner, J. Paraneoplastic Syndromes Involving the Nervous System. *N Engl J Med* **2003**.
  - (16) Pignolet, B. S. L.; Gebauer, C. M. T.; Liblau, R. S. Immunopathogenesis of Paraneoplastic Neurological Syndromes Associated with Anti-Hu Antibodies: A Beneficial Antitumor Immune Response Going Awry. *OncImmunity*. 2013. <https://doi.org/10.4161/onci.27384>.
  - (17) Szabo, A.; Dalmau, J.; Manley, G.; Rosenfeld, M.; Wong, E.; Henson, J.; Posner, J. B.; Furneaux, H. M. HuD, a Paraneoplastic Encephalomyelitis Antigen, Contains RNA-Binding Domains and Is Homologous to Elav and Sex-Lethal. *Cell* **1991**. [https://doi.org/10.1016/0092-8674\(91\)90184-Z](https://doi.org/10.1016/0092-8674(91)90184-Z).
  - (18) Vasile, F.; Della Volpe, S.; Ambrosio, F. A.; Costa, G.; Unver, M. Y.; Zucal, C.; Rossi, D.; Martino, E.; Provenzani, A.; Hirsch, A. K. H.; et al. Exploration of Ligand Binding Modes towards the Identification of Compounds Targeting HuR: A Combined STD-NMR and Molecular Modelling Approach. *Sci. Rep.* **2018**, *8* (1). <https://doi.org/10.1038/s41598-018-32084-z>.
  - (19) Tang, A. Y. RNA Processing-Associated Molecular Mechanisms of Neurodegenerative Diseases. *Journal of Applied Genetics*. 2016. <https://doi.org/10.1007/s13353-015-0330-5>.
  - (20) Kotta-Loizou, I.; Giaginis, C.; Theocharis, S. Clinical Significance of HuR Expression in Human Malignancy. *Medical Oncology*. 2014. <https://doi.org/10.1007/s12032-014-0161-y>.
  - (21) Wang, J.; Guo, Y.; Chu, H.; Guan, Y.; Bi, J.; Wang, B. Multiple Functions of the RNA-Binding Protein HuR in Cancer Progression, Treatment Responses and Prognosis. *International Journal of Molecular Sciences*. 2013. <https://doi.org/10.3390/ijms140510015>.
  - (22) Campos-Melo, D.; Droppelmann, C. A.; Volkening, K.; Strong, M. J. RNA-Binding Proteins as Molecular Links between Cancer and Neurodegeneration. *Biogerontology*. 2014. <https://doi.org/10.1007/s10522-014-9531-2>.
  - (23) Wurth, L.; Gebauer, F. RNA-Binding Proteins, Multifaceted Translational Regulators in Cancer. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 2015. <https://doi.org/10.1016/j.bbagr.2014.10.001>.
  - (24) König, J.; Zarnack, K.; Luscombe, N. M.; Ule, J. Protein-RNA Interactions: New Genomic Technologies and Perspectives. *Nat. Rev. Genet.* **2012**. <https://doi.org/10.1038/nrg3141>.
  - (25) Doxakis, E. RNA Binding Proteins: A Common Denominator of Neuronal Function and Dysfunction. *Neurosci. Bull.* **2014**. <https://doi.org/10.1007/s12264-014-1443-7>.
  - (26) Nasti, R.; Rossi, D.; Amadio, M.; Pascale, A.; Unver, M. Y.; Hirsch, A. K. H.; Collina, S. Compounds Interfering with Embryonic Lethal Abnormal Vision (ELAV) Protein-RNA Complexes: An Avenue for Discovering New Drugs. *J. Med. Chem.* **2017**. <https://doi.org/10.1021/acs.jmedchem.6b01871>.

- (27) Dejong, E.; Luy, B.; Marino, J. RNA and RNA-Protein Complexes as Targets for Therapeutic Intervention. *Curr. Top. Med. Chem.* **2005**. <https://doi.org/10.2174/1568026023394245>.
- (28) Kim, D. H.; Behlke, M.; Rossi, J. J. Designing and Utilization of SiRNAs Targeting RNA Binding Proteins. *Methods Mol. Biol.* **2008**. [https://doi.org/10.1007/978-1-60327-475-3\\_24](https://doi.org/10.1007/978-1-60327-475-3_24).
- (29) Jacobsen, A.; Wen, J.; Marks, D. S.; Krogh, A. Signatures of RNA Binding Proteins Globally Coupled to Effective MicroRNA Target Sites. *Genome Res.* **2010**. <https://doi.org/10.1101/gr.103259.109>.
- (30) Guo, X.; Wu, Y.; Hartley, R. S. MicroRNA-125a Represses Cell Growth by Targeting HuR in Breast Cancer. *RNA Biol.* **2009**. <https://doi.org/10.4161/rna.6.5.10079>.
- (31) Makeyev, A. V.; Eastmond, D. L.; Liebhaber, S. A. Targeting a KH-Domain Protein with RNA Decoys. *RNA* **2002**. <https://doi.org/10.1017/S135583820202808X>.
- (32) Abil, Z.; Zhao, H. Engineering Reprogrammable RNA-Binding Proteins for Study and Manipulation of the Transcriptome. *Mol. Biosyst.* **2015**. <https://doi.org/10.1039/c5mb00289c>.
- (33) Buonamici, S.; Yoshimi, A.; Thomas, M.; Seiler, M.; Chan, B.; Caleb, B.; Darman, R.; Fekkes, P.; Karr, C.; Keaney, G. F.; et al. H3B-8800, an Orally Bioavailable Modulator of the SF3b Complex, Shows Efficacy in Spliceosome-Mutant Myeloid Malignancies. *Blood* **2016**.
- (34) Steensma, D. P.; Maris, M. B.; Yang, J.; Donnellan, W. B.; Brunner, A. M.; McMasters, M.; Greenberg, P.; Komrokji, R. S.; Klimek, V. M.; Goldberg, J. M.; et al. H3B-8800-G0001-101: A First in Human Phase I Study of a Splicing Modulator in Patients with Advanced Myeloid Malignancies. *J. Clin. Oncol.* **2017**. [https://doi.org/10.1200/jco.2017.35.15\\_suppl.tps7075](https://doi.org/10.1200/jco.2017.35.15_suppl.tps7075).
- (35) Seiler, M.; Yoshimi, A.; Darman, R.; Chan, B.; Keaney, G.; Thomas, M.; Agrawal, A. A.; Caleb, B.; Csibi, A.; Sean, E.; et al. H3B-8800, an Orally Available Small-Molecule Splicing Modulator, Induces Lethality in Spliceosome-Mutant Cancers. *Nat. Med.* **2018**. <https://doi.org/10.1038/nm.4493>.
- (36) Kotake, Y.; Sagane, K.; Owa, T.; Mimori-Kiyosue, Y.; Shimizu, H.; Uesugi, M.; Ishihama, Y.; Iwata, M.; Mizui, Y. Splicing Factor SF3b as a Target of the Antitumor Natural Product Pladienolide. *Nat. Chem. Biol.* **2007**. <https://doi.org/10.1038/nchembio.2007.16>.
- (37) Hinman, M. N.; Lou, H. Diverse Molecular Functions of Hu Proteins. *Cellular and Molecular Life Sciences.* 2008. <https://doi.org/10.1007/s00018-008-8252-6>.
- (38) Perrone-Bizzozero, N.; Bird, C. W. Role of HuD in Nervous System Function and Pathology. *Frontiers in Bioscience - Scholar.* 2013.
- (39) Antic, D.; Keene, J. D. Embryonic Lethal Abnormal Visual RNA-Binding Proteins Involved in Growth, Differentiation, and Posttranscriptional Gene Expression. *Am. J. Hum. Genet.* **1997**. <https://doi.org/10.1086/514866>.
- (40) Srikantan, S.; Gorospe, M. HuR Function in Disease. *Front. Biosci.* **2012**, *17* (1), 189–205. <https://doi.org/10.2741/3921>.

- (41) Nagai, K.; Oubridge, C.; Ito, N.; Avis, J.; Evans, P. The RNP Domain: A Sequence-Specific RNA-Binding Domain Involved in Processing and Transport of RNA. *Trends in Biochemical Sciences*. 1995. [https://doi.org/10.1016/S0968-0004\(00\)89024-6](https://doi.org/10.1016/S0968-0004(00)89024-6).
- (42) Fan, X. C.; Steitz, J. A. HNS, a Nuclear-Cytoplasmic Shuttling Sequence in HuR. *Proc. Natl. Acad. Sci. U. S. A.* **1998**. <https://doi.org/10.1073/pnas.95.26.15293>.
- (43) Eberhardt, W.; Doller, A.; Pfeilschifter, J. Regulation of the mRNA-Binding Protein HuR by Posttranslational Modification: Spotlight on Phosphorylation. *Curr. Protein Pept. Sci.* **2012**. <https://doi.org/10.2174/138920312801619439>.
- (44) Wang, X.; Hall, T. M. Structural Basis for Recognition of AU-Rich Element RNA by the HuD Protein. *Nat. Struct. Biol.* **2001**. <https://doi.org/10.1038/84131>.
- (45) Kasashima, K. Complex Formation of the Neuron-Specific ELAV-like Hu RNA-Binding Proteins. *Nucleic Acids Res.* **2002**. <https://doi.org/10.1093/nar/gkf567>.
- (46) Toba, G.; White, K. The Third RNA Recognition Motif of Drosophila ELAV Protein Has a Role in Multimerization. *Nucleic Acids Res.* **2008**. <https://doi.org/10.1093/nar/gkm1168>.
- (47) Díaz-Quintana, A.; García-Mauriño, S. M.; Díaz-Moreno, I. Dimerization Model of the C-Terminal RNA Recognition Motif of HuR. *FEBS Lett.* **2015**. <https://doi.org/10.1016/j.febslet.2015.03.013>.
- (48) Colombrita, C.; Silani, V.; Ratti, A. ELAV Proteins along Evolution: Back to the Nucleus? *Molecular and Cellular Neuroscience*. 2013. <https://doi.org/10.1016/j.mcn.2013.02.003>.
- (49) Talman, V.; Pascale, A.; Jääntti, M.; Amadio, M.; Tuominen, R. K. Protein Kinase C Activation as a Potential Therapeutic Strategy in Alzheimer's Disease: Is There a Role for Embryonic Lethal Abnormal Vision-like Proteins? *Basic and Clinical Pharmacology and Toxicology*. 2016. <https://doi.org/10.1111/bcpt.12581>.
- (50) Pascale, A.; Amadio, M.; Quattrone, A. Defining a Neuron: Neuronal ELAV Proteins. *Cellular and Molecular Life Sciences*. 2008. <https://doi.org/10.1007/s00018-007-7017-y>.
- (51) Scheckel, C.; Drapeau, E.; Frias, M. A.; Park, C. Y.; Fak, J.; Zucker-Scharff, I.; Kou, Y.; Haroutunian, V.; Ma'ayan, A.; Buxbaum, J. D.; et al. Regulatory Consequences of Neuronal ELAV-like Protein Binding to Coding and Non-Coding RNAs in Human Brain. *Elife* **2016**. <https://doi.org/10.7554/eLife.10421>.
- (52) Zucal, C.; D'Agostino, V.; Loffredo, R.; Mantelli, B.; NatthakanThongon; Lal, P.; Latorre, E.; Provenzani, A. Targeting the Multifaceted HuR Protein, Benefits and Caveats. *Curr. Drug Targets* **2015**. <https://doi.org/10.2174/1389450116666150223163632>.
- (53) Zaharieva, E.; Haussmann, I. U.; Bräuer, U.; Soller, M. Concentration and Localization of Co-Expressed ELAV/Hu Proteins Control Specificity of mRNA Processing. *Mol. Cell. Biol.* **2015**. <https://doi.org/10.1128/mcb.00473-15>.
- (54) Dong, R.; Yang, G.-D.; Luo, N.-A.; Qu, Y.-Q. HuR: A Promising Therapeutic Target for

- Angiogenesis. *Gland Surg.* **2014**. <https://doi.org/10.3978/j.issn.2227-684X.2014.03.02>.
- (55) Khabar, K. S. A. Post-Transcriptional Control during Chronic Inflammation and Cancer: A Focus on AU-Rich Elements. *Cellular and Molecular Life Sciences.* 2010. <https://doi.org/10.1007/s00018-010-0383-x>.
- (56) M., A.; A., P.; J., W.; L., H.; A., Q.; S., G.; V., H.; M., R.; G.M., P. NELAV Proteins Alteration in Alzheimer's Disease Brain: A Novel Putative Target for Amyloid-Beta Reverberating on AbetaPP Processing. *Journal of Alzheimer's Disease.* 2009.
- (57) Meisner, N. C.; Hintersteiner, M.; Mueller, K.; Bauer, R.; Seifert, J. M.; Naegeli, H. U.; Ottl, J.; Oberer, L.; Guenat, C.; Moss, S.; et al. Identification and Mechanistic Characterization of Low-Molecular-Weight Inhibitors for HuR. *Nat. Chem. Biol.* **2007**. <https://doi.org/10.1038/nchembio.2007.14>.
- (58) Kask, P.; Palo, K.; Ullmann, D.; Gall, K. Fluorescence-Intensity Distribution Analysis and Its Application in Biomolecular Detection Technology. *Proc. Natl. Acad. Sci. U. S. A.* **1999**. <https://doi.org/10.1073/pnas.96.24.13756>.
- (59) Kaur, K.; Wu, X.; Fields, J. K.; Johnson, D. K.; Lan, L.; Pratt, M.; Somoza, A. D.; Wang, C. C. C.; Karanicolas, J.; Oakley, B. R.; et al. The Fungal Natural Product Azaphilone-9 Binds to HuR and Inhibits HuR-RNA Interaction in Vitro. *PLoS One* **2017**. <https://doi.org/10.1371/journal.pone.0175471>.
- (60) Kwak, H.; Jeong, K. C.; Chae, M. J.; Kim, S. Y.; Park, W. Y. Flavonoids Inhibit the AU-Rich Element Binding of HuC. *BMB Rep.* **2009**. <https://doi.org/10.5483/BMBRep.2009.42.1.041>.
- (61) Fillebeen, C.; Wilkinson, N.; Pantopoulos, K. Electrophoretic Mobility Shift Assay (EMSA) for the Study of RNA-Protein Interactions: The IRE/IRP Example. *J. Vis. Exp.* **2014**. <https://doi.org/10.3791/52230>.
- (62) Chae, M. J.; Sung, H. Y.; Kim, E. H.; Lee, M.; Kwak, H.; Chae, C. H.; Kim, S.; Park, W. Y. Chemical Inhibitors Destabilize HuR Binding to the AU-Rich Element of TNF- $\alpha$  mRNA. *Exp. Mol. Med.* **2009**. <https://doi.org/10.3858/emm.2009.41.11.088>.
- (63) D'Agostino, V. G.; Adami, V.; Provenzani, A. A Novel High Throughput Biochemical Assay to Evaluate the HuR Protein-RNA Complex Formation. *PLoS One* **2013**. <https://doi.org/10.1371/journal.pone.0072426>.
- (64) D'Agostino, V. G.; Lal, P.; Mantelli, B.; Tiedje, C.; Zucal, C.; Thongon, N.; Gaestel, M.; Latorre, E.; Marinelli, L.; Seneci, P.; et al. Dihydrotanshinone-I Interferes with the RNA-Binding Activity of HuR Affecting Its Post-Transcriptional Function. *Sci. Rep.* **2015**. <https://doi.org/10.1038/srep16478>.
- (65) Lea, W. A.; Simeonov, A. Fluorescence Polarization Assays in Small Molecule Screening. *Expert Opinion on Drug Discovery.* January 2011, pp 17–32. <https://doi.org/10.1517/17460441.2011.537322>.

- (66) Wu, X.; Lan, L.; Wilson, D. M.; Marquez, R. T.; Tsao, W. C.; Gao, P.; Roy, A.; Turner, B. A.; McDonald, P.; Tunge, J. A.; et al. Identification and Validation of Novel Small Molecule Disruptors of HuR-MRNA Interaction. *ACS Chem. Biol.* **2015**. <https://doi.org/10.1021/cb500851u>.
- (67) Wang, Z.; Bhattacharya, A.; Ivanov, D. N. Identification of Small-Molecule Inhibitors of the HuR/RNA Interaction Using a Fluorescence Polarization Screening Assay Followed by NMR Validation. *PLoS One* **2015**. <https://doi.org/10.1371/journal.pone.0138780>.
- (68) Kakuguchi, W.; Nomura, T.; Kitamura, T.; Otsuguro, S.; Matsushita, K.; Sakaitani, M.; Maenaka, K.; Tei, K. Suramin, Screened from an Approved Drug Library, Inhibits HuR Functions and Attenuates Malignant Phenotype of Oral Cancer Cells. *Cancer Med.* **2018**. <https://doi.org/10.1002/cam4.1877>.
- (69) Hinman, M. N.; Zhou, H. L.; Sharma, A.; Lou, H. All Three RNA Recognition Motifs and the Hinge Region of HuC Play Distinct Roles in the Regulation of Alternative Splicing. *Nucleic Acids Res.* **2013**, *41* (9), 5049–5061. <https://doi.org/10.1093/nar/gkt166>.
- (70) Park, S.; Myszka, D. G.; Yu, M.; Littler, S. J.; Laird-Offringa, I. A. HuD RNA Recognition Motifs Play Distinct Roles in the Formation of a Stable Complex with AU-Rich RNA. *Mol. Cell. Biol.* **2000**, *20* (13), 4765.
- (71) Park-Lee, S.; Kim, S.; Laird-Offringa, I. A. Characterization of the Interaction between Neuronal RNA-Binding Protein HuD and AU-Rich RNA. *J. Biol. Chem.* **2003**, *278* (41), 39801–39808. <https://doi.org/10.1074/jbc.M307105200>.
- (72) Scheiba, R. M.; De Opakua, A. I.; Díaz-Quintana, A.; Cruz-Gallardo, I.; Martínez-Cruz, L. A.; Martínez-Chantar, M. L.; Blanco, F. J.; Díaz-Moreno, I. 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.* **2014**, *11* (10), 1250–1261. <https://doi.org/10.1080/15476286.2014.996069>.
- (73) Ripin, N.; Boudet, J.; Duszczuk, M. M.; Hinniger, A.; Faller, M.; Krepl, M.; Gadi, A.; Schneider, R. J.; Sponer, J.; Meisner-Kober, N. C.; et al. Molecular Basis for AU-Rich Element Recognition and Dimerization by the HuR C-Terminal RRM. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (8), 2935–2944. <https://doi.org/10.1073/pnas.1808696116>.
- (74) Pabis, M.; Popowicz, G. M.; Stehle, R.; Fernández-Ramos, D.; Asami, S.; Warner, L.; García-Mauriño, S. M.; Schlundt, A.; Martínez-Chantar, M. L.; Díaz-Moreno, I.; et al. HuR Biological Function Involves RRM3-Mediated Dimerization and RNA Binding by All Three RRMs. *Nucleic Acids Res.* **2019**, *47* (2), 1011–1029. <https://doi.org/10.1093/nar/gky1138>.
- (75) Wang, H.; Zeng, F.; Liu, Q.; Liu, H.; Liu, Z.; Niu, L.; Teng, M.; Li, X. The Structure of the ARE-Binding Domains of Hu Antigen R (HuR) Undergoes Conformational Changes during RNA Binding. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**.

- <https://doi.org/10.1107/S0907444912047828>.
- (76) Rossi, D.; Amadio, M.; Baraglia, A. C.; Azzolina, O.; Ratti, A.; Govoni, S.; Pascale, A.; Collina, S. Discovery of Small Peptides Derived from Embryonic Lethal Abnormal Vision Proteins Structure Showing RNA-Stabilizing Properties. *J. Med. Chem.* **2009**. <https://doi.org/10.1021/jm900741e>.
- (77) Amadio, M.; Pascale, A.; Govoni, S.; Laurini, E.; Pricl, S.; Gaggeri, R.; Rossi, D.; Collina, S. Identification of Peptides with ELAV-like mRNA-Stabilizing Effect: An Integrated in Vitro/in Silico Approach. *Chem. Biol. Drug Des.* **2013**. <https://doi.org/10.1111/cbdd.12117>.
- (78) Vasile, F.; Rossi, D.; Collina, S.; Potenza, D. Diffusion-Ordered Spectroscopy and Saturation Transfer Difference NMR Spectroscopy Studies of Selective Interactions between ELAV Protein Fragments and an mRNA Target. *European J. Org. Chem.* **2014**. <https://doi.org/10.1002/ejoc.201403014>.
- (79) Meyer, B.; Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chemie - Int. Ed.* **2003**, *42* (8), 864–890. <https://doi.org/10.1002/anie.200390233>.
- (80) Ciulli, A. Biophysical Screening for the Discovery of Small-Molecule Ligands. *Methods Mol. Biol.* **2013**. [https://doi.org/10.1007/978-1-62703-398-5\\_13](https://doi.org/10.1007/978-1-62703-398-5_13).
- (81) Mittermaier, A.; Meneses, E. *Analyzing Protein–Ligand Interactions by Dynamic NMR Spectroscopy*. In: *Protein-Ligand Interactions. Methods in Molecular Biology (Methods and Protocols)*; 2013. <https://doi.org/10.1007/978-1-62703-398-5>.
- (82) Li, Y.; Kang, C. Solution NMR Spectroscopy in Target-Based Drug Discovery. *Molecules*. 2017. <https://doi.org/10.3390/molecules22091399>.
- (83) Sugiki, T.; Furuita, K.; Fujiwara, T.; Kojima, C. *Molecules* Current NMR Techniques for Structure-Based Drug Discovery. <https://doi.org/10.3390/molecules23010148>.
- (84) Barile, E.; Pellicchia, M. NMR-Based Approaches for the Identification and Optimization of Inhibitors of Protein–Protein Interactions. **2014**. <https://doi.org/10.1021/cr500043b>.
- (85) Maity, S.; Gundampati, R. K.; Kumar, T. K. S. NMR Methods to Characterize Protein-Ligand Interactions. *Natural Product Communications*. Natural Product Incorporation May 1, 2019. <https://doi.org/10.1177/1934578X19849296>.
- (86) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chemie - Int. Ed.* **1999**, *38* (12), 1784–1788. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990614\)38:12<1784::AID-ANIE1784>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q).
- (87) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J. Chem. Educ.* **2011**, *88* (7), 990–994. <https://doi.org/10.1021/ed101169t>.

- (88) Mayer, M.; Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR to Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **2001**, *123* (25), 6108–6117. <https://doi.org/10.1021/ja0100120>.
- (89) Fielding, L.; Rutherford, S.; Fletcher, D. Determination of Protein-Ligand Binding Affinity by NMR: Observations from Serum Albumin Model Systems. *Magn. Reson. Chem.* **2005**, *43* (6), 463–470. <https://doi.org/10.1002/mrc.1574>.
- (90) Wang, Y. Sen; Liu, D.; Wyss, D. F. Competition STD NMR for the Detection of High-Affinity Ligands and NMR-Based Screening. *Magn. Reson. Chem.* **2004**, *42* (6), 485–489. <https://doi.org/10.1002/mrc.1381>.
- (91) Monaco, S.; Tailford, L. E.; Juge, N.; Angulo, J. Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein–Ligand Contacts. *Angew. Chemie - Int. Ed.* **2017**. <https://doi.org/10.1002/anie.201707682>.
- (92) Watt, J. E.; Hughes, G. R.; Walpole, S.; Monaco, S.; Stephenson, G. R.; Bulman Page, P. C.; Hemmings, A. M.; Angulo, J.; Chantry, A. Discovery of Small Molecule WWP2 Ubiquitin Ligase Inhibitors. *Chem. - A Eur. J.* **2018**. <https://doi.org/10.1002/chem.201804169>.
- (93) Nepravishita, R.; Walpole, S.; Tailford, L.; Juge, N.; Angulo, J. Deriving Ligand Orientation in Weak Protein–Ligand Complexes by DEEP-STD NMR Spectroscopy in the Absence of Protein Chemical-Shift Assignment. *ChemBioChem* **2019**. <https://doi.org/10.1002/cbic.201800568>.
- (94) Merwe, P. A. Van Der. *Surface Plasmon Resonance In: Harding S, Chowdhry PZ (Eds) Protein–Ligand Interactions: A Practical Approach*; 2000. <https://doi.org/10.1007/978-1-60761-670-2>.
- (95) Merwe, P. A. Van Der. *Surface Plasmon Resonance GENERAL PRINCIPLES OF BIACORE EXPERIMENTS. Physics (College. Park. Md).* **2010**. <https://doi.org/10.1007/978-1-60761-670-2>.
- (96) O’Shannessy, D. J.; Brigham-Burke, M.; Peck, K. Immobilization Chemistries Suitable for Use in the BIAcore Surface Plasmon Resonance Detector. *Anal. Biochem.* **1992**. [https://doi.org/10.1016/0003-2697\(92\)90589-Y](https://doi.org/10.1016/0003-2697(92)90589-Y).
- (97) Ezzati Nazhad Dolatabadi, J.; de la Guardia, M. Tips on Ligand Immobilization and Kinetic Study Using Surface Plasmon Resonance. *BioImpacts* **2016**. <https://doi.org/10.15171/bi.2016.17>.
- (98) Navratilova, I.; Hopkins, A. L. Fragment Screening by Surface Plasmon Resonance. *ACS Med. Chem. Lett.* **2010**. <https://doi.org/10.1021/ml900002k>.
- (99) Robson-Tull, J. Biophysical Screening in Fragment-Based Drug Design: A Brief Overview. *Biosci. Horizons* **2018**, *11*. <https://doi.org/10.1093/biohorizons/hzy015>.
- (100) Perspicace, S.; Banner, D.; Benz, J.; Müller, F.; Schlatter, D.; Huber, W. Fragment-Based Screening Using Surface Plasmon Resonance Technology. *J. Biomol. Screen.* **2009**. <https://doi.org/10.1177/1087057109332595>.

- (101) Nikolovska-Coleska, Z. Studying Protein-Protein Interactions Using Surface Plasmon Resonance. *Methods Mol. Biol.* **2015**, *1278*, 109–138. [https://doi.org/10.1007/978-1-4939-2425-7\\_7](https://doi.org/10.1007/978-1-4939-2425-7_7).
- (102) Douzi, B. Protein-Protein Interactions: Surface Plasmon Resonance. *Methods Mol. Biol.* **2017**, *1615*, 257–275. [https://doi.org/10.1007/978-1-4939-7033-9\\_21](https://doi.org/10.1007/978-1-4939-7033-9_21).
- (103) Pantoliano, M. W.; Petrella, E. C.; Kwasnoski, J. D.; Lobanov, V. S.; Myslik, J.; Graf, E.; Carver, T.; Asel, E.; Springer, B. A.; Lane, P.; et al. High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery. *J. Biomol. Screen.* **2001**, *6* (6), 429–440. <https://doi.org/10.1089/108705701753364922>.
- (104) Sun, C.; Li, Y.; Yates, E.; Fernig, D. SimpleDSFviewer: A Tool to Analyse and View Differential Scanning Fluorimetry Data for Characterising Protein Thermal Stability and Interactions. *PeerJ Prepr.* **2015**. <https://doi.org/10.7287/peerj.preprints.1555>.
- (105) Nguyen, C.; West, G. M.; Geoghegan, K. F. Emerging Methods in Chemoproteomics with Relevance to Drug Discovery. In *Methods in Molecular Biology*; 2017. [https://doi.org/10.1007/978-1-4939-6539-7\\_2](https://doi.org/10.1007/978-1-4939-6539-7_2).
- (106) Bruce, D.; Cardew, E.; Freitag-Pohl, S.; Pohl, E. How to Stabilize Protein: Stability Screens for Thermal Shift Assays and Nano Differential Scanning Fluorimetry in the Virus-X Project. *J. Vis. Exp.* **2019**. <https://doi.org/10.3791/58666>.
- (107) Vivoli, M.; Novak, H. R.; Littlechild, J. A.; Harmer, N. J. Determination of Protein-Ligand Interactions Using Differential Scanning Fluorimetry. *J. Vis. Exp.* **2014**, No. 91. <https://doi.org/10.3791/51809>.
- (108) Senisterra, G.; Chau, I.; Vedadi, M. Thermal Denaturation Assays in Chemical Biology. *Assay Drug Dev. Technol.* **2012**, *10* (2), 128–136. <https://doi.org/10.1089/adt.2011.0390>.
- (109) Ciulli, A.; Scott, D. E.; Ando, M.; Reyes, F.; Saldanha, S. A.; Tuck, K. L.; Chirgadze, D. Y.; Blundell, T. L.; Abell, C. Inhibition of Mycobacterium Tuberculosis Pantothenate Synthetase by Analogues of the Reaction Intermediate. *Chembiochem* **2008**. <https://doi.org/10.1002/cbic.200800437>.
- (110) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A Review of Protein-Small Molecule Docking Methods. *J. Comput. Aided. Mol. Des.* **2002**, *16* (3), 151–166. <https://doi.org/10.1023/A:1020155510718>.
- (111) Jorgensen, W. L. Rusting of the Lock and Key Model for Protein-Ligand Binding. *Science*. 1991. <https://doi.org/10.1126/science.1719636>.
- (112) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron*. 2018. <https://doi.org/10.1016/j.neuron.2018.08.011>.
- (113) Karplus, M.; McCammon, J. A. Molecular Dynamics Simulations of Biomolecules. *Nature Structural Biology*. 2002. <https://doi.org/10.1038/nsb0902-646>.

- (114) Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Frontiers in Pharmacology*. 2018. <https://doi.org/10.3389/fphar.2018.00923>.
- (115) Murray, C. W.; Rees, D. C. The Rise of Fragment-Based Drug Discovery. *Nat. Chem.* **2009**, *1* (3), 187–192. <https://doi.org/10.1038/nchem.217>.
- (116) Hopkins, A. L.; Groom, C. R. The Druggable Genome. *Nat. Rev. Drug Discov.* **2002**. <https://doi.org/10.1038/nrd892>.
- (117) Cavasotto, C.; W. Orry, A. Ligand Docking and Structure-Based Virtual Screening in Drug Discovery. *Curr. Top. Med. Chem.* **2007**. <https://doi.org/10.2174/156802607780906753>.
- (118) Hajduk, P. J.; Huth, J. R.; Tse, C. Predicting Protein Druggability. *Drug Discovery Today*. 2005. [https://doi.org/10.1016/S1359-6446\(05\)03624-X](https://doi.org/10.1016/S1359-6446(05)03624-X).
- (119) Andricopulo, A.; Salum, L.; Abraham, D. Structure-Based Drug Design Strategies in Medicinal Chemistry. *Curr. Top. Med. Chem.* **2009**, *9* (9), 771–790. <https://doi.org/10.2174/156802609789207127>.
- (120) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science*. 2004. <https://doi.org/10.1126/science.1096361>.
- (121) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nature Reviews Drug Discovery*. 2004. <https://doi.org/10.1038/nrd1549>.
- (122) Fattori, D. Molecular Recognition: The Fragment Approach in Lead Generation. *Drug Discovery Today*. March 1, 2004, pp 229–238. [https://doi.org/10.1016/S1359-6446\(03\)03007-1](https://doi.org/10.1016/S1359-6446(03)03007-1).
- (123) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent Developments in Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry*. July 10, 2008, pp 3661–3680. <https://doi.org/10.1021/jm8000373>.
- (124) Hajduk, P. J.; Greer, J. A Decade of Fragment-Based Drug Design: Strategic Advances and Lessons Learned. *Nature Reviews Drug Discovery*. 2007. <https://doi.org/10.1038/nrd2220>.
- (125) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A “Rule of Three” for Fragment-Based Lead Discovery? *Drug Discovery Today*. 2003. [https://doi.org/10.1016/S1359-6446\(03\)02831-9](https://doi.org/10.1016/S1359-6446(03)02831-9).
- (126) Zhu, J.; Bienaymé, H. *Multicomponent Reactions*; 2005. <https://doi.org/10.1002/3527605118>.
- (127) Weber, L.; Illgen, K.; Almstetter, M. Discovery of New Multi Component Reactions with Combinatorial Methods. *Synlett* **1999**. <https://doi.org/10.1055/s-1999-2612>.
- (128) Biggs-Houck, J. E.; Younai, A.; Shaw, J. T. Recent Advances in Multicomponent Reactions for Diversity-Oriented Synthesis. *Current Opinion in Chemical Biology*. June 2010, pp 371–382. <https://doi.org/10.1016/j.cbpa.2010.03.003>.
- (129) Dömling, A. Recent Developments in Isocyanide Based Multicomponent Reactions in Applied Chemistry. *Chemical Reviews*. 2006. <https://doi.org/10.1021/cr0505728>.

- (130) Hantzsch, A. Ueber Die Synthese Pyridinartiger Verbindungen Aus Acetessigäther Und Aldehydammoniak. *Justus Liebig's Ann. der Chemie* **1882**, 215 (1), 1–82. <https://doi.org/10.1002/jlac.18822150102>.
- (131) Biginelli, P. Aldureides of Ethylic Acetoacetate and Ethylic Oxalacetate. *Gazz. Chim. Ital* **1893**. <https://doi.org/10.1007/978-94-007-4698-5>.
- (132) Passerini, M. Isonitriles. I. Compound of p-Isonitrileazobenzene with Acetone and Acetic Acid. *Gazz. Chim. Ital.* **1921**.
- (133) Ugi, I.; Meyr, R.; Fetzer, U.; Steinbruckner, C. Studies on Isonitriles. *Angew. Chemie* **1959**, 71 (11), 386. <https://doi.org/10.1002/ange.19590711110>.
- (134) The Academic Pursuit of Screening. *Nature Chemical Biology*. Nature Publishing Group 2007, p 433. <https://doi.org/10.1038/nchembio0807-433>.
- (135) Screening We Can Believe In. *Nature Chemical Biology*. Nature Publishing Group 2009, p 127. <https://doi.org/10.1038/nchembio0309-127>.
- (136) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. Dynamic Covalent Chemistry. *Angew. Chemie Int. Ed.* **2002**. [https://doi.org/10.1002/1521-3773\(20020503\)41:9<1460::aid-anie11111460>3.0.co;2-n](https://doi.org/10.1002/1521-3773(20020503)41:9<1460::aid-anie11111460>3.0.co;2-n).
- (137) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.; Sanders, J. K. M.; Otto, S. Dynamic Combinatorial Chemistry Dynamic Combinatorial Chemistry. *Chem. Rev.* **2006**, 106 (August), 3652–3711. <https://doi.org/10.1021/cr020452p>.
- (138) Lehn, J. M. From Supramolecular Chemistry towards Constitutional Dynamic Chemistry and Adaptive Chemistry. *Chem. Soc. Rev.* **2007**, 36 (2), 151–160. <https://doi.org/10.1039/b616752g>.
- (139) Li, J.; Nowak, P.; Otto, S. Dynamic Combinatorial Libraries: From Exploring Molecular Recognition to Systems Chemistry. *Journal of the American Chemical Society*. June 26, 2013, pp 9222–9239. <https://doi.org/10.1021/ja402586c>.
- (140) Hartman, A. M.; Gierse, R. M.; Hirsch, A. K. H. Protein-Templated Dynamic Combinatorial Chemistry: Brief Overview and Experimental Protocol. *European Journal of Organic Chemistry*. 2019. <https://doi.org/10.1002/ejoc.201900327>.
- (141) Monjas, L.; Swier, L. J. Y. M.; Setyawati, I.; Slotboom, D. J.; Hirsch, A. K. H. Dynamic Combinatorial Chemistry to Identify Binders of ThiT, an S-Component of the Energy-Coupling Factor Transporter for Thiamine. *ChemMedChem* **2017**, 12 (20), 1693–1696. <https://doi.org/10.1002/cmdc.201700440>.
- (142) Das, M.; Yang, T.; Dong, J.; Prasetya, F.; Xie, Y.; Wong, K. H. Q.; Cheong, A.; Woon, E. C. Y. Multiprotein Dynamic Combinatorial Chemistry: A Strategy for the Simultaneous Discovery of Subfamily-Selective Inhibitors for Nucleic Acid Demethylases FTO and ALKBH3. *Chem. Asian J.* **2018**, 13 (19), 2854–2867. <https://doi.org/10.1002/asia.201800729>.
- (143) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F.

- Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry. *Chem. - A Eur. J.* **2012**, *18* (34), 10562–10570. <https://doi.org/10.1002/chem.201201507>.
- (144) Reddavid, F. V.; Lin, W.; Lehnert, S.; Zhang, Y. DNA-Encoded Dynamic Combinatorial Chemical Libraries. *Angew. Chemie - Int. Ed.* **2015**, *54* (27), 7924–7928. <https://doi.org/10.1002/anie.201501775>.
- (145) Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. Structure-Based Design of Inhibitors of the Aspartic Protease Endothiapepsin by Exploiting Dynamic Combinatorial Chemistry. *Angew. Chemie - Int. Ed.* **2014**. <https://doi.org/10.1002/anie.201309682>.
- (146) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. Nucleophilic Catalysis of Acylhydrazone Equilibration for Protein-Directed Dynamic Covalent Chemistry. *Nat. Chem.* **2010**, *2* (6), 490–497. <https://doi.org/10.1038/nchem.658>.
- (147) Crisalli, P.; Kool, E. T. Water-Soluble Organocatalysts for Hydrazone and Oxime Formation. *J. Org. Chem.* **2013**, *78* (3), 1184–1189. <https://doi.org/10.1021/jo302746p>.
- (148) Ramström, O.; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Discov.* **2002**, *1* (1), 26–36. <https://doi.org/10.1038/nrd704>.
- (149) Herrmann, A. Dynamic Combinatorial/Covalent Chemistry: A Tool to Read, Generate and Modulate the Bioactivity of Compounds and Compound Mixtures. *Chemical Society Reviews*. March 21, 2014, pp 1899–1933. <https://doi.org/10.1039/c3cs60336a>.
- (150) Huang, R.; Leung, I. K. H. Protein-Directed Dynamic Combinatorial Chemistry: A Guide to Protein Ligand and Inhibitor Discovery. *Molecules*. MDPI AG July 1, 2016. <https://doi.org/10.3390/molecules21070910>.
- (151) Mondal, M.; Hirsch, A. K. H. Dynamic Combinatorial Chemistry: A Tool to Facilitate the Identification of Inhibitors for Protein Targets. *Chemical Society Reviews*. 2015. <https://doi.org/10.1039/c4cs00493k>.
- (152) Tapas, A. R.; Magar, D. D.; Kawtikwar, P. S.; Sakarkar, D. M.; Kakde, R. B. Microwaves in Drug Discovery and Development: A Review. *International Journal of PharmTech Research*. 2009.
- (153) Tatke, P.; Jaiswal, Y. An Overview of Microwave Assisted Extraction and Its Applications in Herbal Drug Research. *Research Journal of Medicinal Plant*. 2011. <https://doi.org/10.3923/rjmp.2011.21.31>.
- (154) Santagada, V.; Frecentese, F.; Perissutti, E.; Favretto, L.; Caliendo, G. The Application of Microwaves in Combinatorial and High-Throughput Synthesis as New Synthetic Procedure in Drug Discovery. *QSAR and Combinatorial Science*. 2004. <https://doi.org/10.1002/qsar.200420039>.
- (155) Santagada, V.; Perissutti, E.; Caliendo, G. The Application of Microwave Irradiation as New

- Convenient Synthetic Procedure in Drug Discovery. *Curr. Med. Chem.* **2012**. <https://doi.org/10.2174/0929867023369989>.
- (156) Wathey, B.; Tierney, J.; Lidström, P.; Westman, J. The Impact of Microwave-Assisted Organic Chemistry on Drug Discovery. *Drug Discovery Today*. 2002. [https://doi.org/10.1016/S1359-6446\(02\)02178-5](https://doi.org/10.1016/S1359-6446(02)02178-5).
- (157) *Microwave Chemistry*; Cravotto, G., Carnaroglio, D., Eds.; De Gruyter: Berlin, Boston, 2017. <https://doi.org/10.1515/9783110479935>.
- (158) McCleverty, J. A.; Meyer, T. J. *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology*; Elsevier Science, 2003.
- (159) Hügél, H. M. Microwave Multicomponent Synthesis. *Molecules*. 2009. <https://doi.org/10.3390/molecules14124936>.
- (160) Cotterill, I. C.; Usyatinsky, A. Y.; Arnold, J. M.; Clark, D. S.; Dordick, J. S.; Michels, P. C.; Khmel'nitsky, Y. L. Microwave Assisted Combinatorial Chemistry. Synthesis of Substituted Pyridines. *Tetrahedron Lett.* **1998**. [https://doi.org/10.1016/S0040-4039\(97\)10796-1](https://doi.org/10.1016/S0040-4039(97)10796-1).
- (161) Larhed, M.; Hallberg, A. Microwave-Assisted High-Speed Chemistry: A New Technique in Drug Discovery. *Drug Discovery Today*. 2001. [https://doi.org/10.1016/S1359-6446\(01\)01735-4](https://doi.org/10.1016/S1359-6446(01)01735-4).
- (162) Wannberg, J.; Dallinger, D.; Kappe, C. O.; Larhed, M. Microwave-Enhanced and Metal-Catalyzed Functionalizations of the 4-Aryl-Dihydropyrimidone Template. *J. Comb. Chem.* **2005**. <https://doi.org/10.1021/cc049816c>.
- (163) Lal, P.; Cerofolini, L.; D'Agostino, V. G.; Zucal, C.; Fuccio, C.; Bonomo, I.; Dassi, E.; Giuntini, S.; Maio, D. Di; Vishwakarma, V.; et al. Regulation of HuR Structure and Function by Dihydro-tanshinone-I. *Nucleic Acids Res.* **2017**. <https://doi.org/10.1093/nar/gkx623>.
- (164) Manzoni, L.; Zucal, C.; Maio, D. Di; 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**. <https://doi.org/10.1021/acs.jmedchem.7b01176>.
- (165) Koes, D.; Khoury, K.; Huang, Y.; Wang, W.; Bista, M.; Popowicz, G. M.; Wolf, S.; Holak, T. A.; Dömling, A.; Camacho, C. J. Enabling Large-Scale Design, Synthesis and Validation of Small Molecule Protein-Protein Antagonists. *PLoS One* **2012**. <https://doi.org/10.1371/journal.pone.0032839>.
- (166) Koes, D. R.; Dömling, A.; Camacho, C. J. AnchorQuery: Rapid Online Virtual Screening for Small-Molecule Protein-Protein Interaction Inhibitors. *Protein Sci.* **2018**. <https://doi.org/10.1002/pro.3303>.
- (167) Kroon, E.; Schulze, J. O.; Süß, E.; Camacho, C. J.; Biondi, R. M.; Dömling, A. Discovery of a Potent Allosteric Kinase Modulator by Combining Computational and Synthetic Methods.

- Angew. Chemie - Int. Ed.* **2015**. <https://doi.org/10.1002/anie.201506310>.
- (168) Groebke, K.; Weber, L.; Mehlin, F. Synthesis of Imidazo[1,2-a] Annulated Pyridines, Pyrazines and Pyrimidines by a Novel Three-Component Condensation. *Synlett* **1998**, No. 6, 661–663. <https://doi.org/10.1055/s-1998-1721>.
- (169) Vijaya Gracias, \*; Alan F. Gasiiecki, and; Djuric, S. W. Synthesis of Fused Bicyclic Imidazoles by Sequential Van Leusen/Ring-Closing Metathesis Reactions. **2005**. <https://doi.org/10.1021/OL050852+>.
- (170) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. Reductive Amination of Aldehydes and Ketones with Sodium Triacetoxyborohydride. Studies on Direct and Indirect Reductive Amination Procedures. *J. Org. Chem.* **1996**.
- (171) Chupakhin, E.; Dar'in, D.; Krasavin, M. The Castagnoli-Cushman Reaction in a Three-Component Format. *Tetrahedron Lett.* **2018**, 59 (26), 2595–2599. <https://doi.org/10.1016/j.tetlet.2018.05.066>.
- (172) Nial J. Wheate; Craig R. Brodie; J. Grant Collins; Sharon Kemp; Janice R. Aldrich-Wright. DNA Intercalators in Cancer Therapy: Organic and Inorganic Drugs and Their Spectroscopic Tools of Analysis. *Mini-Reviews Med. Chem.* **2007**, 7 (6), 627–648. <https://doi.org/10.2174/138955707780859413>.
- (173) Brazil, M. Getting Stuck In. *Nat. Rev. Drug Discov.* **2002**, 1 (1), 9–9. <https://doi.org/10.1038/nrd717>.
- (174) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science* (80-. ). **1991**. <https://doi.org/10.1126/science.1962210>.
- (175) Ray, A.; Nordén, B. Peptide Nucleic Acid (PNA): Its Medical and Biotechnical Applications and Promise for the Future. *FASEB J.* **2000**. <https://doi.org/10.1096/fasebj.14.9.1041>.
- (176) WHO, W. H. O.; World Health Organization. *Antibacterial Agents in Clinical Development: An Analysis of the Antibacterial Clinical Development Pipeline, Including Tuberculosis*; 2017.
- (177) Taubes, G. The Bacteria Fight Back. *Science*. **2008**. <https://doi.org/10.1126/science.321.5887.356>.
- (178) Strateva, T.; Mitov, I. Contribution of an Arsenal of Virulence Factors to Pathogenesis of Pseudomonas Aeruginosa Infections. *Annals of Microbiology*. **2011**. <https://doi.org/10.1007/s13213-011-0273-y>.
- (179) Kamal, A. A. M.; Maurer, C. K.; Allegretta, G.; Haupenthal, J.; Empting, M.; Hartmann, R. W. Quorum Sensing Inhibitors as Pathoblockers for Pseudomonas Aeruginosa Infections: A New Concept in Anti-Infective Drug Discovery. In *Topics in Medicinal Chemistry*; **2018**. [https://doi.org/10.1007/7355\\_2017\\_17](https://doi.org/10.1007/7355_2017_17).
- (180) Thayer, M. M.; Flaherty, K. M.; McKay, D. B. Three-Dimensional Structure of the Elastase of

- Pseudomonas Aeruginosa at 1.5-Å Resolution. *J. Biol. Chem.* **1991**.
- (181) Kany, A. M.; Sikandar, A.; Hauptenthal, J.; Yahiaoui, S.; Maurer, C. K.; Proschak, E.; Köhnke, J.; Hartmann, R. W. Binding Mode Characterization and Early in Vivo Evaluation of Fragment-Like Thiols as Inhibitors of the Virulence Factor LasB from Pseudomonas Aeruginosa. *ACS Infect. Dis.* **2018**. <https://doi.org/10.1021/acsinfecdis.8b00010>.
- (182) Holder, I. A.; Wheeler, R. Experimental Studies of the Pathogenesis of Infections Owing to Pseudomonas Aeruginosa: Elastase, an IgG Protease. *Can. J. Microbiol.* **1984**. <https://doi.org/10.1139/m84-175>.
- (183) Graef, F.; Vukosavljevic, B.; Michel, J. P.; Wirth, M.; Ries, O.; De Rossi, C.; Windbergs, M.; Rosilio, V.; Ducho, C.; Gordon, S.; et al. The Bacterial Cell Envelope as Delimiter of Anti-Infective Bioavailability – An in Vitro Permeation Model of the Gram-Negative Bacterial Inner Membrane. *J. Control. Release* **2016**. <https://doi.org/10.1016/j.jconrel.2016.10.018>.
- (184) Holland, D. R.; Tronrud, D. E.; Matthews, B. W.; Pley, H. W.; Flaherty, K. M.; McKay, D. B.; Stark, W.; Jansonius, J. N. Structural Comparison Suggests That Thermolysin and Related Neutral Proteases Undergo Hinge-Bending Motion During Catalysis. *Biochemistry* **1992**. <https://doi.org/10.1021/bi00161a008>.
- (185) Kessler, E.; Israel, M.; Landshman, N.; Chechick, A.; Blumberg, S. In Vitro Inhibition of Pseudomonas Aeruginosa Elastase by Metal-Chelating Peptide Derivatives. *Infect. Immun.* **1982**.
- (186) Adekoya, O. A.; Sjøli, S.; Wuxiuer, Y.; Bילו, I.; Marques, S. M.; Santos, M. A.; Nuti, E.; Cercignani, G.; Rossello, A.; Winberg, J. O.; et al. Inhibition of Pseudolysin and Thermolysin by Hydroxamate-Based MMP Inhibitors. *Eur. J. Med. Chem.* **2015**. <https://doi.org/10.1016/j.ejmech.2014.10.009>.
- (187) Cathcart, G. R. A.; Quinn, D.; Greer, B.; Harriott, P.; Lynas, J. F.; Gilmore, B. F.; Walker, B. Novel Inhibitors of the Pseudomonas Aeruginosa Virulence Factor LasB: A Potential Therapeutic Approach for the Attenuation of Virulence Mechanisms in Pseudomonas Infection. *Antimicrob. Agents Chemother.* **2011**. <https://doi.org/10.1128/AAC.00776-10>.
- (188) Zhu, J.; Cai, X.; Harris, T. L.; Gooyit, M.; Wood, M.; Lardy, M.; Janda, K. D. Disarming Pseudomonas Aeruginosa Virulence Factor Lasb by Leveraging a Caenorhabditis Elegans Infection Model. *Chem. Biol.* **2015**. <https://doi.org/10.1016/j.chembiol.2015.03.012>.
- (189) Schönauer, E.; Kany, A. M.; Hauptenthal, J.; Hüsecken, K.; Hoppe, I. J.; Voos, K.; Yahiaoui, S.; Elsässer, B.; Ducho, C.; Brandstetter, H.; et al. Discovery of a Potent Inhibitor Class with High Selectivity toward Clostridial Collagenases. *J. Am. Chem. Soc.* **2017**. <https://doi.org/10.1021/jacs.7b06935>.
- (190) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *Journal of Medicinal Chemistry*. 2013. <https://doi.org/10.1021/jm3012068>.

- (191) Zimmermann, M. O.; Lange, A.; Boeckler, F. M. Evaluating the Potential of Halogen Bonding in Molecular Design: Automated Scaffold Decoration Using the New Scoring Function Xbscore. *J. Chem. Inf. Model.* **2015**. <https://doi.org/10.1021/ci5007118>.
- (192) Lange, A.; Heidrich, J.; Zimmermann, M. O.; Exner, T. E.; Boeckler, F. M. Scaffold Effects on Halogen Bonding Strength. *J. Chem. Inf. Model.* **2019**. <https://doi.org/10.1021/acs.jcim.8b00621>.

## II. Abbreviations

<b>3D</b>	Three-dimensional
<b>ACN</b>	Acetonitrile
<b>ADMET</b>	Absorption, distribution, metabolism, elimination, toxicity
<b>AlphaLISA</b>	Amplified Luminescent Proximity Homogeneous Assay Linked Immunosorbent Assay
<b>AlphaScreen</b>	Amplified Luminescent Proximity Homogeneous Assay Screen
<b>Bis-TRIS</b>	Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane
<b>CADD</b>	Computer-aided drug design
<b>CD</b>	Circular Dicroism
<b>cryoEM</b>	Cryogenic electron microscopy
<b>DCC</b>	Dynamic Combinatorial Chemistry
<b>DCL</b>	Dynamic combinatorial library
<b>DCM</b>	Dichloromethane
<b>DDOP</b>	Department of Drug Design and Optimization
<b>DEA</b>	Diethyl amine
<b>DEEP-STD NMR</b>	Differential Epitope Mapping by STD NMR
<b>DIBAL-H</b>	Diisobutyl alluminium hydride
<b>DIPEA</b>	Diisopropylethylamine
<b>DMF</b>	Dimethylformamide
<b>DMSO</b>	Dimethyl sulfoxide
<b>DOSY</b>	Diffusion-Ordered Spectroscopy
<b>EDC•HCl</b>	N-Ethyl-N'-(3-dimethylaminopropyl)carbodiimide hydrochloride
<b>EDG</b>	Electron-donor group
<b>EtOAc</b>	Ethylacetate
<b>FBDD</b>	Fragment-based drug design
<b>FBLD</b>	Fragment-based ligand design
<b>FC</b>	Flash chromatography
<b>FIDA</b>	Fluorescence Intensity Distribution Analysis
<b>FP</b>	Fluorescence Polarization
<i>ee</i>	Enantiomeric excess
<b>ELAV</b>	Embryonic Lethal Abnormal Vision
<b>HEFL</b>	Halogen-enriched Fragment Library
<b>Hex</b>	Hexane
<b>HIPS</b>	Helmholtz Institute for Pharmaceutical Research Saarland
<b>HPLC</b>	High performance liquid chromatography
<b>HRMS</b>	High resolution mass spectrometry
<b>HTS</b>	High Throughput Screening
<b>IPA</b>	Isopropyl alcohol
<b>ITC</b>	Isothermal Calorimetry
<b><sup>1</sup>H-<sup>15</sup>N HSQC</b>	<sup>1</sup> H- <sup>15</sup> N Heteronuclear Single Quantum Correlation
<b>MCR</b>	Multi-component Reaction
<b>MD</b>	Molecular Dynamics
<b>MedChem</b>	Medicinal Chemistry
<b>MeOH</b>	Methanol
<b>MES</b>	4-Morpholineethanesulfonic acid

<b>MM</b>	Molecular Modeling
<b>mRNA</b>	Messenger RNA
<b>MS</b>	Mass spectrometry
<b>MTBE</b>	<i>tert</i> -Butyl methyl ether
<b>mw</b>	Microwaves
<b>NOE</b>	Nuclear Overhauser Effect
<b>pt-DCC</b>	Protein templated DCC
<b>REMSA</b>	RNA Electrophoretic Mobility Shift Assay
<b>RBD</b>	RNA-binding domain
<b>RBP</b>	RNA-binding protein
<b>RNP</b>	Ribonucleoprotein
<b>rt</b>	Room temperature
<b>Rt</b>	Retention time
<b>rt-qPCR</b>	Quantitative reverse transcription-polymerase chain reaction
<b>SAR</b>	Structure-activity relationship
<b>SBDD</b>	Structure-based drug design
<b>SBLD</b>	Structure-based ligand design
<b>SBVS</b>	Structure-based virtual screening
<b>SDS-PAGE</b>	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
<b>SF3B1</b>	Splicing Factor 3B1
<b>siRNA</b>	Small <i>or</i> short interfering RNA
<b>SPR</b>	Surface Plasmon Resonance
<b>STD NMR</b>	Saturation Transfer Difference NMR
<b>TEA</b>	Triethylamine
<b>TBTU</b>	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylammonium tetrafluoroborate
<b>THF</b>	Tetrahydrofuran
<b>THP</b>	Tetrahydropyran
<b>TLC</b>	Thin-layer chromatography
<b><math>T_m</math></b>	Melting Point
<b>TNF</b>	Tumor Necrosis Factor
<b>TRIS</b>	Tris(hydroxymethyl)aminomethane
<b>TSA</b>	Thermal Shift Assay
<b>UV</b>	Ultraviolet
<b>VEGF</b>	Vascular Endothelial Growth Factor

### III. Publications related to the thesis

#### Article 1: Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD-NMR and Molecular Modelling approach

www.nature.com/scientificreports

# SCIENTIFIC REPORTS

OPEN

## Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD-NMR and Molecular Modelling approach

Received: 9 April 2018  
Accepted: 29 August 2018  
Published online: 13 September 2018

Francesca Vasile<sup>1</sup>, Serena Della Volpe<sup>1,2</sup>, Francesca Alessandra Ambrosio<sup>3</sup>, Giosuè Costa<sup>3</sup>, M. Yagiz Unver<sup>4</sup>, Chiara Zucal<sup>5</sup>, Daniela Rossi<sup>2</sup>, Emanuela Martino<sup>6</sup>, Alessandro Provenanzi<sup>5</sup>, Anna K. H. Hirsch<sup>4,7</sup>, Stefano Alcaro<sup>3</sup>, Donatella Potenza<sup>1</sup> & Simona Collina<sup>1,2</sup>

Post-transcriptional processes have been recognised as pivotal in the control of gene expression, and impairments in RNA processing are reported in several pathologies (i.e., cancer and neurodegeneration). Focusing on RNA-binding proteins (RBPs), the involvement of Embryonic Lethal Abnormal Vision (ELAV) or Hu proteins and their complexes with target mRNAs in the aetiology of various dysfunctions, has suggested the great potential of compounds able to interfere with the complex stability as an innovative pharmacological strategy for the treatment of numerous diseases. Here, we present a rational follow-up investigation of the interaction between ELAV isoform HuR and structurally-related compounds (i.e., flavonoids and coumarins), naturally decorated with different functional groups, by means of STD-NMR and Molecular Modelling. Our results represent the foundation for the development of potent and selective ligands able to interfere with ELAV–RNA complexes.

Post-transcriptional modifications have a crucial role in regulating gene expression by shaping the fate of RNA transcripts in their journey from the nucleus (i.e., alternative splicing, poly-adenylation, nuclear export) to the ribosome (i.e., cytoplasmic localisation, stability, translation rate). Dysfunctions within these routes may be directly involved in several pathologies, such as neurodegenerative diseases, inflammation and cancer<sup>1–5</sup>.

In this context, RNA-binding proteins (RBPs), have been recognised to be play a prominent role as they affect the fate of target messenger RNAs (mRNAs) coding for proteins pivotal in key cellular functions<sup>1–5</sup>.

Embryonic Lethal Abnormal Vision (ELAV) or Hu proteins are among the better characterised RBPs with 4 human mammalian isoforms: ubiquitous HuR, and HuB, HuC and HuD prevalently expressed in the nervous system (nELAVs)<sup>6</sup>. These proteins share a high degree of sequence homology (70–85%): they contain three RNA recognition motif-type (RRM) domains, each approximately 90 amino-acid in length<sup>7</sup>; the first two consecutive domains (RRM1 and RRM2) are near the N-terminus and link to the third domain (RRM3) by an unconserved hinge region, responsible for the nuclear/cytoplasmic shuttling occurring after protein activation<sup>8–10</sup>. RRM1 and RRM2 directly interact with target transcripts through highly conserved ribonucleoprotein (RNP) sequences 1

<sup>1</sup>Department of Chemistry, University of Milan, Via Golgi 19, 20133, Milano, Italy. <sup>2</sup>Department of Drug Sciences, Medicinal Chemistry and Technology Section, University of Pavia, Via Taramelli 12, 27100, Pavia, Italy.

<sup>3</sup>Department of Health Sciences, University "Magna Græcia" of Catanzaro, Viale Europa, 88100, Catanzaro, Italy.

<sup>4</sup>Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) - Helmholtz Centre for Infection Research (HZI),

Department of Drug Design and Optimization, Campus building E8.1, 66123, Saarbrücken, Germany. <sup>5</sup>Centre for

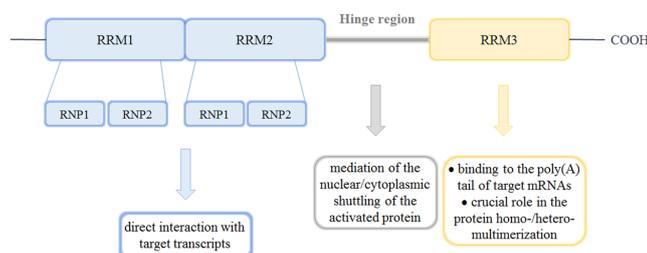
Integrative Biology, CIBIO, University of Trento, Via Sommarive 9, 38123, Povo, TN, Italy. <sup>6</sup>Department of Earth

and Environmental Sciences, University of Pavia, via S. Epifanio 14, 27100, Pavia, Italy. <sup>7</sup>Department of Pharmacy,

Medicinal Chemistry, Saarland University, Campus building E8.1, 66123, Saarbrücken, Germany. Francesca Vasile

and Serena Della Volpe contributed equally. Correspondence and requests for materials should be addressed to S.C.

(email: [simona.collina@unipv.it](mailto:simona.collina@unipv.it))



**Figure 1.** ELAV domains and their involvement in protein function.

and 2, formed by 8 and 6 amino acids, respectively<sup>11,12</sup> while RRM3, aside from binding the poly(A) tail of target mRNAs, has a crucial role in the homo- and heteromultimerisation of ELAV proteins (Fig. 1)<sup>13–15</sup>.

The interaction of the four mammalian ELAV proteins with many mRNAs, gives rise to various ELAV protein–RNA complexes, characterised by different physiological roles. As a consequence, ELAV proteins have a potential as pharmacological targets in several pathologies and compounds able to interfere with ELAV–RNA complexes may lead to different effects.

Particularly, it has been demonstrated that the RNA-binding protein HuR is highly abundant in many cancers and it may either be a marker for malignancy or have an oncogenic role in numerous tumor systems including breast, ovarian, and colon<sup>16–18</sup>. In fact, numerous HuR-regulated mRNAs encode proteins implicated in carcinogenesis<sup>19</sup>. For this reason, HuR can be considered a promising candidate target for governing gene regulatory mechanisms. Specifically, it was shown that HuR is upregulated and dysregulated in cancer cells, in part, through post-transcriptional gene regulation<sup>19,20</sup>. Among the small natural products that have been reported to inhibit HuR function, it is worth mentioning MS-444, since it has been shown to prevent translocation of HuR (and its associated mRNA cargos) to the cytoplasm. In a number of tumor cells, HuR inhibition by MS-444 leads to a dose-dependent reduction in cell proliferation by promoting apoptosis<sup>21</sup>. The majority of compounds interacting/interfering with HuR–RNA complexes have been discovered by screening a large number (or a library) of commercially available compounds with high-throughput screening (HTS) approaches and using various biological assays based on different detection technologies<sup>21,22</sup>. Among them, dihydrotanshinone (DHTS), which is a nanomolar disruptor of HuR–RNA binding<sup>23</sup>, has potent HuR-dependent antitumor activity *in vivo* and is able to inhibit HuR multimerization<sup>24,25</sup>. Starting from the DHTS scaffold, medicinal-chemistry efforts led to the discovery of a series of small molecules called Tanshinone Mimics (TMs) with improved affinity and potency compared to DHTS<sup>26</sup>. Summing up, and going beyond the state of the art of HuR modulators available so far, the identification of the key druggable pockets on the surface of HuR is the essential milestone for discovering new molecular scaffolds.

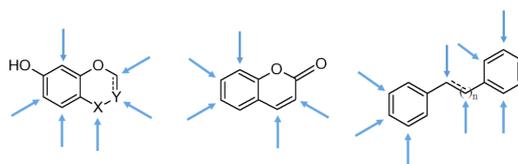
As a part of our ongoing efforts in this field, we had already demonstrated the importance of the four RNP sequences in binding and stabilising a target transcript RNA by means of real-time quantitative PCR<sup>27</sup>, molecular modelling<sup>28</sup> and advanced NMR techniques (Saturation-Transfer Difference, STD, and Diffusion-Ordered Spectroscopy, DOSY)<sup>29</sup>.

In the present work, we set out to study the interaction between HuR and a number of compounds performing a systematic study which combines a ligand-based NMR technique, namely STD-NMR, with a molecular modelling study. It is well known that STD-NMR can be used as an epitope mapping device to describe the target–ligand interactions<sup>30–32</sup> and it can be applied to weak and transient protein–ligand complexes that are difficult to study by other structural methods<sup>33,34</sup>. We compared results of STD-NMR with molecular dynamics and docking simulations, thus affording an improved understanding of ligand–HuR interactions. Despite the high protein mobility and width of the protein–RNA interface, the combination of NMR and *in silico* studies resulted to be a reliable tool for understanding HuR–ligand binding modes<sup>35,36</sup>. The exploitation of these interactions lays the foundation for the design of *ad hoc* molecules endowed with HuR–RNA complex interfering properties.

## Results

**Selection of compounds, solubility and stability assessment.** Our previous docking results had revealed that a series of natural products, *i.e.* epigallocatechin gallate, quercetin, okicenone, myricetyn, DHTS, MS-444 and others (described in SI), interact with HuR in the same regions as the target RNAs, more specifically with the RNP1 and RNP2 sequences of the RRM1 and RRM2 domains of the protein<sup>22</sup>. Based on these preliminary data, with the aim to study the ligand–HuR interaction, we collected a small series of compounds of natural origin, namely flavones, flavonols, flavan-3-ols and coumarins, naturally decorated with different functional groups, as well as some unrelated compounds with similar features and a high degree of structural diversity (Figs 2, 3). We then tested their solubility and stability in the buffer and time-ranges required by STD-NMR experiments (24–48 hours). In detail, a 1 mM solution of each compound was dissolved in a 20 mM deuterated phosphate buffer within a 5.0–7.4 pH range (where necessary, a DMSO-*d*<sub>6</sub> percentage ≤ 10% was added), at a 283–303 K temperature range and studied by preliminary <sup>1</sup>H-NMR experiments.

Among the tested molecules, we selected 13 soluble and stable compounds (Fig. 3) and used them in STD-NMR experiments to explore the ligand–protein interaction mode.



**Figure 2.** General structures of the naturally-occurring compounds selected; the arrows define the possible positions of decoration.

**NMR interaction studies with HuR.** Prior to the interaction study with HuR,  $^1\text{H}$ , COSY, TOCSY, HSQC and NOESY spectra of the selected compounds were recorded in the proper deuterated phosphate buffer solution (pH range 5.8–7.4) at 298 K and 283 K. The assignment of all compounds is reported in SI.

STD-NMR is one of the most widespread NMR techniques used to study the interactions between small ligands and macromolecules<sup>37,38</sup>. This method is based on the transfer of saturation from the protein to the bound ligand which in turn, by exchange, is moved into solution where it is detected. During the period of saturation (saturation time), the magnetisation gradually moves from the protein to the protons of the ligand when it binds to the target. The saturation time is chosen taking into account the efficiency of saturation transfer from protein protons during the bound state (intermolecular protein–ligand NOEs) and the rate of accumulation of saturated ligand molecules in the free state. Long saturation times permit to map all ligand contacts. This saturation process is very efficient, so the modulation of the ligand signal induced by the protein is readily detected, even in the presence of a large excess of ligand. The ligand protons nearest to the protein are most likely to be saturated to the highest degree, and therefore have the strongest signal in the one-dimensional STD spectrum. The ligand protons located further away are saturated to a lower degree, and their STD intensities are weaker. Therefore, the degree of saturation of individual ligand protons (expressed as absolute-STD %) reflects their proximity to the protein surface and can be used as an epitope-mapping method to describe the target–ligand interactions<sup>39,40</sup>. STD-NMR spectra were initially acquired with varying saturation times from 0.98 s to 2.94 s and using different ligand–protein ratios (from about 500:1 to 1000:1). This large molar excess of ligand was employed in order to preclude the perturbations of absolute STD intensities due to rebinding effects, which would impede the correct determination of the group epitope mapping. Additionally, negative controls were performed to avoid artefacts due to the presence of signals in the blank. The optimal conditions found entail the use of 2.94 s, with high ligand/protein ratios (around 1000:1) for all compounds and exploiting the Watergate sequence for water suppression.

The absolute STD percentages of each ligand protons were quantified in order to analyse their proximity to the protein surface (short protein–ligand distances produce a strong intensity of the corresponding STD signal). On the other hand, also relative STD percentages were calculated for each compound, by normalising all measured STD intensities against the most intense signal (which is arbitrarily assigned a value of 100%). The obtained group epitope mapping then illustrates which chemical moieties of the ligand are key for molecular recognition in the binding site.

STD-NMR spectra were acquired and processed accordingly for all 13 compounds and showed that 12 out of the 13 compounds studied interact with HuR displaying different intensities and epitopes of interaction. Given the difficulty of determining  $K_D$  values for each compound with STD-NMR experiments due to the compounds poor solubility, we analysed the absolute STD data following two different keys: (a) the intensity of the STD signals (indicative of the proximity to the protein) and (b) the number of interactions for each ligand.

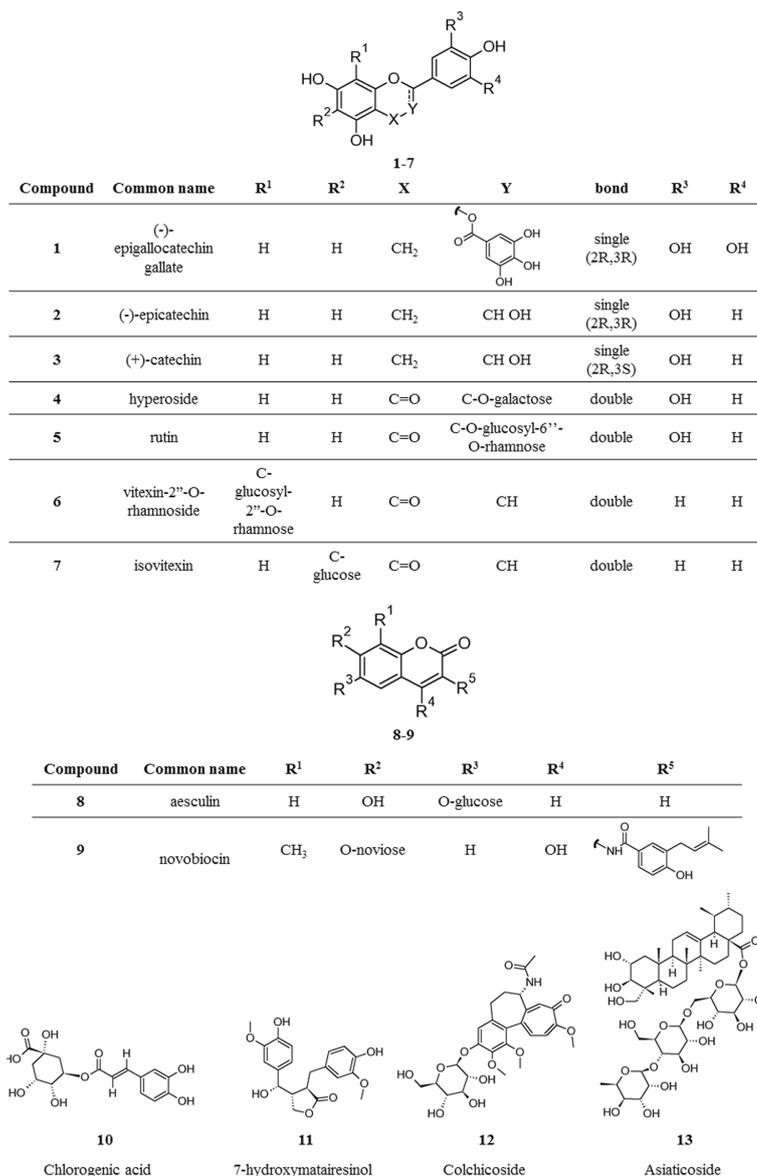
By reporting the number of the signals and related intensity of each compound, Table 1 gives a picture of their interaction with target HuR. Most of the compounds studied are characterised by a weak interaction with the protein within the absolute STD range of 0.4–0.1% and only a few ligands present STD signals exceeding the unit value (Table 1). Remarkably, compound 9 gives the highest number of interactions, while 6 is characterised by fewer and weaker interactions with the protein.

Moreover, compound 5 has the highest number of very strong interactions and several of medium and low intensity. On the other hand, the STD-NMR spectra for 13 and HuR showed absence of signal, indicative of lack of interaction between the two entities.

In addition to absolute STD% to evaluate the binding epitope of each compound, we calculated relative STD% as shown in Fig. 4 conveyed by colour code; black dots are used to indicate the most intense signal arbitrarily assigned the value of 100% relative STD. Subsequently, dark red dots represent relative STD over 80%, orange dots over 40%, and lime green dots under 40%; all are relative to the most intense STD signal. Additionally, in Fig. 5, we report the  $^1\text{H}$  and STD spectra for compound 5 and the corresponding coloured dots for relative STD% according to the colour code just described.

**In silico studies on HuR protein.** For the preliminary selection of the compounds to be used in our study, we utilised the rigid docking approach previously published<sup>22</sup>, which allowed us to quickly evaluate their potential. In order to better characterise the results obtained through STD-NMR, we decided to exploit a complementary *in silico* approach, which is computationally more demanding but enables a more detailed description of the ligand–protein behaviour in solution.

The protein crystal structure represents one of the many possible substates of the protein and, in most cases, the overall topology of the folded state is conserved, but a different orientation of even a single side chain in the binding site can significantly influence docking results<sup>41</sup>. Thus, in order to consider all conformational states of

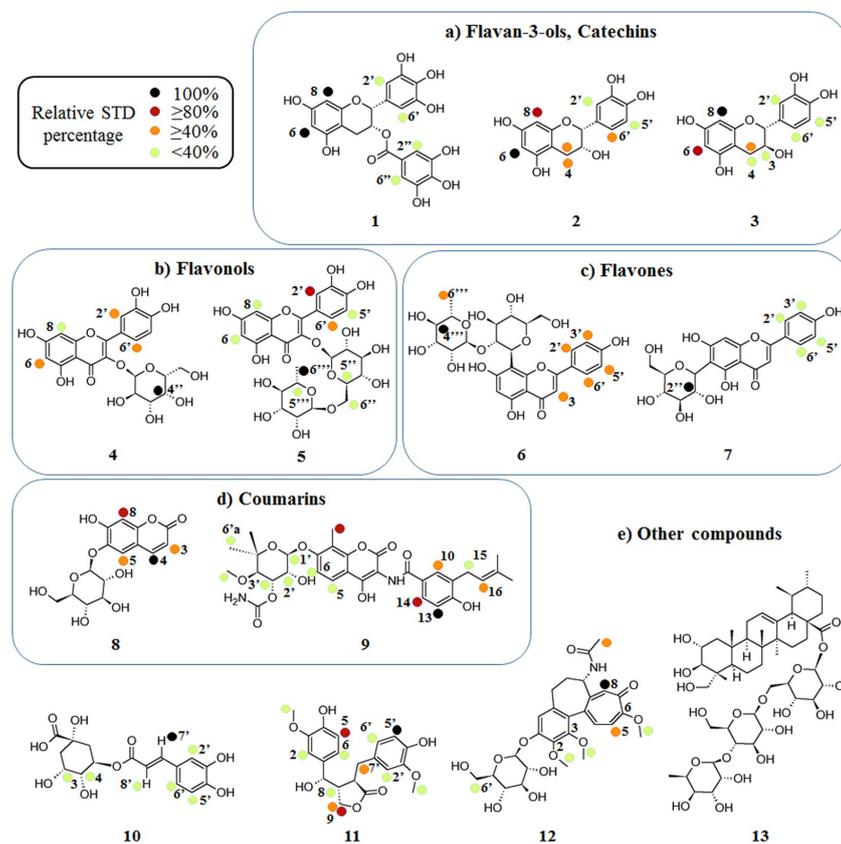


**Figure 3.** Structures of the naturally occurring products ultimately subjected to the study.

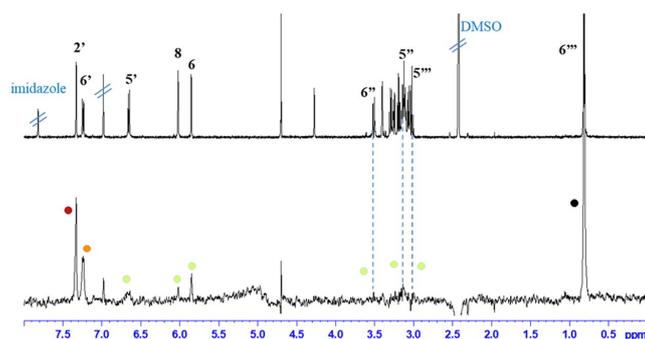
HuR and to investigate potential protein rearrangement, we submitted the HuR RRM domain–RNA<sup>c-fos</sup> complex (PDB code: 4ED5) to 500 ns of Molecular Dynamics simulations (MDs)<sup>42</sup>. The result trajectory was clustered with respect to the Root Mean Square Deviation (RMSD), affording ten representative structures, which we used for the following docking studies. By applying this approach, we were able to reproduce the results previously reported on the two main conformational states of the HuR RRM domains, “open” and “closed”<sup>24,43</sup>. We analysed how the ligands can bind the two states and we observed that the STD contacts better fit with the “closed” state of HuR. Here, we will describe the interaction of each ligand with a HuR snapshot corresponding to the “closed” conformation. Molecular-recognition studies revealed that 12 of the 13 considered compounds interact with the same region in the HuR interaction site (Figs 6b and 7), in a deep pocket of domains RRM1 and RRM2, while

Compound	Number of interacting protons					
	Absolute STD % ranges					
	>1	1.0-0.9	0.8-0.7	0.6-0.5	0.4-0.2	<0.2
1			2			4
2				2	3	2
3				2	1	5
4			1		3	1
5	3				5	1
6					1	6
7		1			2	2
8					2	2
9	2	2	1	1	7	
10	1			1	5	
11			1	2	5	5
12	1		2	1	3	
13	no signal shown					

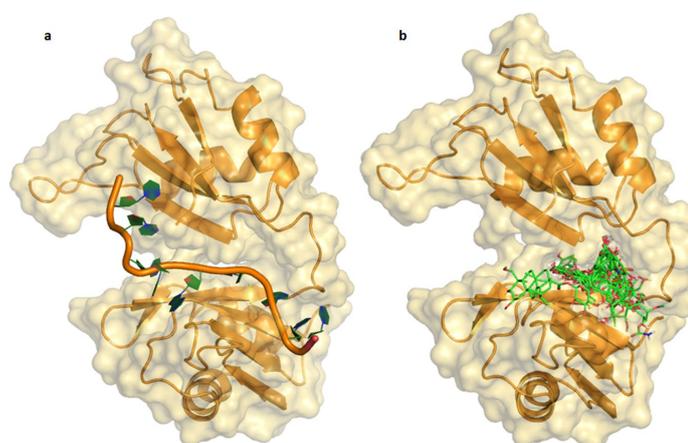
**Table 1.** Compound and related number of interacting protons within listed absolute STD range.



**Figure 4.** Group epitope mapping is highlighted for each compound. Relative STD percentages are conveyed by colour code: black dots indicate the most intense signal (100% relative STD), dark red dots over 80%, orange dots over 40%, and lime green dots under 40% relative to the most intense STD signal.



**Figure 5.**  $^1\text{H}$  and STD spectrum of compound 5; coloured dots convey relative STD % according to the following colour code, black dots indicate the most intense signal arbitrarily assigned the value of 100% relative STD; dark red dots represent relative STD over 80%, orange dots over 40%, and lime green dots under 40%; all relative to the most intense STD signal.



**Figure 6.** (a) 3D representation of HuR RRM1 and RRM2 domain–RNA<sup>c-fos</sup> complex in its “closed” conformation; (b) Superimposition of all studied compounds in the “closed” conformation of the HuR–RNA interaction site. HuR and RNA<sup>c-fos</sup> are represented as light-orange and orange cartoon, respectively; all compounds are depicted as green sticks.

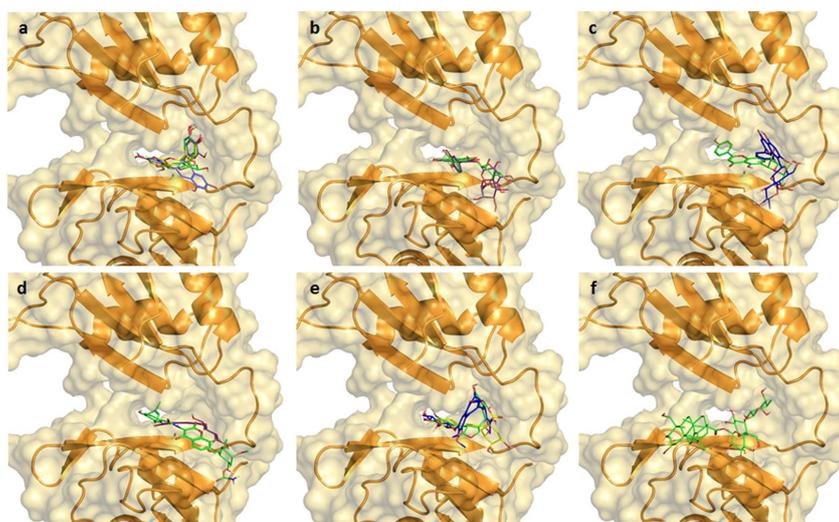
compound **13** appears to interact with their external surface (docking poses divided by class of compound are reported in Fig. 7, while poses for single compounds are reported in SI). All contacts between the HuR binding site and ligands will be identified and catalogued in terms of hydrophobic contacts, hydrogen bonds and cation- $\pi$  interactions.

Furthermore, it is worth noting that the HuR “closed” form thus explained could exceed the sole descriptive purpose and be exploited as a model for future structure-based virtual screening (SBVS) studies using different compound libraries, as already implemented on different targets in our previous investigations<sup>44,45</sup>, in order to identify new ligands interfering with the HuR–RNA complex.

## Discussion

STD-NMR and molecular modelling results have been compared to gain a more detailed picture of the ligand–protein interactions. As previously mentioned, MD simulations of HuR clearly showed that the protein can exist in two prevalent structural conformations: the “open” conformation and the “closed” one; the comparison of experimental STD-NMR and theoretical docking results, led us to hypothesise that the “closed” conformation is predominant in presence of interacting ligands.

For simplicity purposes, we present all the following considerations in the next section as a comparison of the results divided by class of compound. Combined NMR and docking data for all interacting compounds are



**Figure 7.** 3D representation of compounds divided by compound class in the HuR–RNA interaction site: (a) compounds 1, 2, 3 (blue, green and yellow sticks, respectively); (b) compounds 4 and 5 (green and violet sticks respectively); (c) compounds 6 and 7 (blue and green, respectively); (d) compounds 8 and 9 (violet and green sticks); (e) compounds 10, 11 and 12 (yellow, green, blue sticks, respectively); (f) compound 13 shown as green sticks; HuR protein shown as orange cartoon<sup>60</sup>.

reported in supporting information, where for each compound, the intensity of the STD signal of each proton is compared with the number of the interactions observed *in silico*.

Concerning Flavan-3-ols or catechins (1–3), the three compounds differ only in the substitution and stereochemistry of the carbon atom in position 3; in particular, 2 and 3 only differ in the configuration of carbon 3 (respectively, *R* and *S*), while 1 (*R*) hosts an additional aromatic ring. STD experiments showed that the protons in positions 6 and 8 of the chroman-3-ol nucleus (structures in Figs 4a and SI) are mainly involved in their interaction with HuR. For all compounds, the aromatic nucleus in position 2, and the additional aromatic ring (gallate ester) in position 3 for compound 1, participate in the interaction, even though to a lesser degree. Docking results show that all compounds interact in the same region of the HuR interaction site (Fig. 7a); in particular, we observed that all aromatic rings in position 2 overlap and are involved in a cation- $\pi$  interaction with Arg97. On the other hand, the chroman-3-ol nucleus of 1 does not show overlapping interactions with the same nucleus of compounds 2 and 3, as it projects towards Phe65, establishing hydrophobic interactions, whereas, the chroman-3-ol nucleus of compounds 2 and 3 overlaps with the gallate ester moiety of 1. Moreover, considering the evidence that 2 and 3 show very similar interactions, it is fair to assume that, within their scaffold, the absolute configuration at carbon 3 does not play a relevant role in the interaction with target HuR.

Within the flavonol compound class (4 and 5), the chromone-3-ol nucleus and the aromatic ring in position 2 are highlighted as interacting regions (Figs 4b and SI), in line with the STD results obtained with the previous class (Flavan-3-ols). Furthermore, in both cases, the additional sugar rings in position 3 reinforce the interaction with the protein. *In silico* results show compounds 4 and 5 bind to HuR by interacting with the same residues (Arg153, Tyr63, Arg97, Pro98). However, the second sugar moiety of compound 5 (rhamnose) establishes additional hydrophobic interactions with Lys55 and Ile52 and a hydrogen bond with Tyr63 of RRM1, justifying a slight deviation from the binding site for compound 5 (Fig. 7b); this is also in accordance with the strong STD signals seen for this sugar moiety.

Different to the first two classes of flavonoids, the chromone nucleus of Flavones (6 and 7) shows almost complete absence of STD peaks (Figs 4c and SI), which denotes a scarce contact with the protein, while, albeit weak, the interaction of the phenols in position 2 is still observed. The sugar moieties interact with the target protein HuR with different intensities, namely 1% absolute STD for the proton in position 2'' of 7, and 0.25% absolute STD for 4''' of 6. Accordingly, 6 and 7 show different binding poses and very limited common interactions (Fig. 7c), as compound 7 elongates in the site forming hydrogen bonds with Asn82, Asn 25, Arg97 and Pro98, while compound 6 adopts a rather closed conformation and interacts with Arg153 and Ile103 of RRM2 and Tyr63 and Lys55 of RRM1.

Focusing on coumarins (8 and 9), within STD experiments, compound 8 displayed few and weak interactions with HuR, presumably due to its small size. Furthermore, the sugar ring in position 6 does not improve its contact with HuR (Figs 4d and SI). In contrast, the STD spectrum of 9 showed a large number of intense peaks relative to both the coumarin nucleus and the conjugated moiety on the right-hand side (Figs 4d and SI). Moreover, the STD peaks relative to the noviose moiety, although weaker, convey its contribution to increasing the interaction with

the target HuR. Accordingly, docking studies reveal compounds **8** and **9** fit differently into the HuR interaction site, most likely due to their different size (Fig. 7d). In particular, the unsaturated conjugated regions of **9** protrude all along the considered interaction site, establishing hydrophobic interactions with Ile132 and Ile133 of RRM2 and Lys92, Val93 and Ser94 of RRM1 and hydrogen bonds with Arg153 of RRM2 and Asn25 of RRM1 with the sugar portion projecting towards the solvent, providing additional interactions with Lys55 of RRM1.

As mentioned above, compounds **10–13** (Fig. 4e) are structurally different to both the original ones and the others involved in the study. For all of them, STD-NMR spectra show that the main interactions are achieved through the presence of unsaturated conjugated regions. Related docking results show compounds **10–12** occupy the deep part of the interaction site (Fig. 7e), as they appear to be involved in hydrogen-bond interactions with amino acids lining the deep part of the pocket such as Arg97, Arg131, Arg153, Asp105.

Finally, the STD spectrum of compound **13** in presence of HuR shows no signals indicating the absence of interactions. Within docking studies against the “closed” conformation of the protein, compound **13** only appears to interact with the external surface of the HuR–RNA interacting site probably due to steric hindrance, and does not overlap with the same region interacting with all other molecules. As a consequence, the *in vitro* interaction may be too weak to fall in the detection range of STD-NMR.

In addition to the considerations reported and discussed inside each compound class, it is worth remarking that recurring patterns can be recognised within the entirety of the compound collection investigated or wider subgroups.

Particularly, among flavonoid derivatives, flavan-3-ols and flavonols (Figs 4a,b, 7a,b and SI), both STD-NMR epitopes and *in silico* data show similar data, while flavones (Figs 4c, 7c and SI) show different behaviours. This has been attributed to both the difference in substitution positions (position 3 for the first two classes, versus 6 and 8 for flavones) and in sugar linking (O- and C-glycoside, respectively), which increase the rigidity of the flavone scaffold.

Moving on to comment on the whole library, the importance of unsaturated conjugated regions as well as aromatic rings in the interaction with HuR has been noted throughout both STD-NMR and molecular modelling; additionally, these moieties seem to fit in repeated regions, and thus show hydrophobic interactions with repeated residues, namely Ile23, Asn25, Ser94, Ile133 and Arg153. Moreover, concerning the sugar moiety counterparts, we observed that the sugar rings directly attached to the aromatic ring of compounds **4**, **5**, **6**, **7** and **8** and the quinic acid moiety of **10**, overlap in the same region of the HuR binding site, establishing contacts with repeated residues (Arg97, Pro98, Ser99 and Ile 103).

To sum up, the results we obtained thus far evidenced that all interacting compounds interact with repeated residues of RNP1 and RNP2 of RRM1 domain; compounds **1–12** interact deep in the interaction site, while **13** protrudes more outwardly towards the solvent. Furthermore, our study corroborates the already reported observation that interacting compounds could stabilise a closed conformation of the HuR binding site<sup>22,26,46</sup>.

## Conclusions

The acknowledgement of the crucial role of post-transcriptional processes in the control of gene expression and their connection to various diseases has opened up a new fascinating route for discovering new drugs. RNA binding proteins, and particularly those belonging to the ELAV family, can affect the fate of target mRNAs whose coded proteins are fundamental for key cellular functions<sup>22,47</sup>. Interfering compounds may specifically affect the fate of ELAV protein–mRNA complexes at various levels, in both the nucleus and the cytoplasm, and also depending on the transcript targeted, as well as on other modulating factors (i.e., other RBPs and miRNAs). Since compounds interfering with all ELAV proteins may give rise to several effects, the tailored design and tissue-targeted delivery of selective compounds will be essential. To date, little structural information are available and a rational drug-design approach is difficult to employ at this stage and this is still considered a challenging research. In this manuscript, by focussing on HuR, owing to its high potential in cancer therapy and diagnosis<sup>48,49</sup> and in line with this observation, we studied the molecular recognition between natural ligands and the protein at the atomic level. We followed a systematic medicinal-chemistry approach based on structural biophysical studies by STD-NMR combined with molecular modelling. To properly explore the chemical space, we collected and selected a library of compounds of natural origin, applying structural diversity criteria.

The combined STD-NMR/molecular modelling approach allowed us to study the direct interactions between HuR and small molecules. By defining the ligand epitope of interaction and corroborating useful aspects about the protein conformational changes and HuR–ligand interactions, our findings represent a pivotal starting point to drive a drug discovery program. The preliminary SAR considerations drawn in this work will drive the design of a new *ad hoc* small focused compound library. In the near future, our efforts will be directed along this long and winding road.

## Methods

**Preliminary docking studies.** The preliminary docking investigation on all 28 selected compounds was carried out as already described<sup>22</sup>.

**Characterisation of the natural compounds, solubility and stability assessment.** To assess the compound stability and solubility, <sup>1</sup>H-NMR spectra were acquired on a 400 MHz Bruker Avance spectrometer using a 1 mM solution of each compound in a 20 mM deuterated phosphate buffer within a 5.0–7.4 pH range, and a 283–303 K temperature range; all occurring peak variations due to instability or solubility issues were monitored over time within a 48 h time period by acquiring <sup>1</sup>H spectra at regular intervals; a DMSO-d<sub>6</sub> percentage ≤ 10% was allowed to dissolve the less soluble compounds.

The full compound characterisation at the optimised conditions (pH 5.8 or 7.4, DMSO-d6% and 283 K) was performed for the final 13 selected compounds and required the further acquisition and assignment of COSY, TOCSY, HSQC and NOESY spectra.

**Protein expression and purification.** Protein expression, purification and purity assessment for HuR aliquots utilised in the STD-NMR study were performed as already described<sup>23,50</sup>.

**STD-NMR.** All protein/ligand samples were prepared in a 1000:1 ligand/protein ratio. Typically, the final concentration of the samples was 400  $\mu$ M of ligand and 0.4  $\mu$ M of HuR, and the final volume was 500  $\mu$ L. The buffer used is a 20  $\mu$ M deuterated phosphate buffer with 10% H<sub>2</sub>O, pH 5.8 for compounds 1–3, 9, 10; pH 7.4 for 4–8, 11–13; an additional amount of DMSO-d6  $\leq$  10% was used to aid the solubility of compounds 4, 5, 7, 11 and 13.

<sup>1</sup>H-STD-NMR experiments were performed at 600 MHz on a Bruker Avance spectrometer. The probe temperature was maintained at 283 K. In the STD experiments, water suppression was achieved by WATERGATE 3–9–19 pulse sequence. The on-resonance irradiation of the protein was performed at a chemical shift of –0.05 ppm for all compounds with the exception of compound 13 which was irradiated at –2.00 ppm to avoid artefacts. Off-resonance irradiation was applied at 200 ppm, where no protein signals are visible. Selective presaturation of the protein was achieved by a train of Gauss-shaped pulses of 49 ms length each. The STD-NMR spectra were acquired with varying saturation times from 0.98 s to 2.94 s; the optimised total length of the saturation train was 2.94 s for all compounds except for compound 1 for which we analysed the STD experiment at 0.98 s due to its instability observed after longer experimental timeframes.

Intensities of all STD effects (absolute STD) were calculated by division through integrals over the respective signals in STD-NMR reference spectra. The different signal intensities of the individual protons are best analysed from the integral values in the reference and STD spectra, respectively.  $(I_0 - I_{\text{sat}})/I_0$  is the fractional STD effect, expressing the signal intensity in the STD spectrum as a fraction of the intensity of an unsaturated reference spectrum. In this equation,  $I_0$  is the intensity of one signal in the off-resonance or reference NMR spectrum,  $I_{\text{sat}}$  is the intensity of a signal in the on-resonance NMR spectrum, and  $I_0 - I_{\text{sat}}$  represents the intensity of the STD-NMR spectrum.

**Molecular Dynamics and Docking simulations.** Starting from the crystal structure of the two N-terminal RRM domains of HuR complexed with RNA, deposited in the Protein Data Bank (PDB) with 4ED5 PDB code<sup>51</sup>, we performed our modelling simulations. The HuR-RNA complex was prepared through Protein Preparation Wizard implemented in Maestro using OLPS-2005 as force field<sup>52–54</sup>. Residual crystallographic buffer components and water molecules were removed, missing side chains were built using the Prime module<sup>55</sup>, hydrogen atoms were added, side chains protonation states at pH 7.4 were assigned. The structure was then submitted to 10000 of MacroModel minimisation steps using OPLS-2005 as force field<sup>54,56</sup>. Molecular Dynamics (MD) simulations were run using Desmond package v. 3.8 at 300 K temperature and ensemble NPT class<sup>57,58</sup>. The system was immersed in an orthorhombic box of TIP4P water molecules, extending at least 10 Å from the protein, and counter ions were added to neutralise the system charge. The resulting trajectory was clustered with respect to Root Mean Square Deviation (RMSD), in order to explore all the collection structures obtained, getting ten representative structures, which were submitted to 10.000 MacroModel minimisation steps, using OLPS-2005 as force field. Docking studies were carried out with Glide, software by using SP v. 6.7 (standard precision) algorithm and the binding pocket was identified by placing a cube centred on the mRNA, 10 poses for ligand were generated<sup>59</sup>.

## References

- Pascale, A. & Govoni, S. The complex world of posttranscriptional mechanisms: Is their deregulation a common link for diseases? Focus on ELAV-like RNA-binding proteins. *Cell. Mol. Life Sci.* **69**, 501–517 (2012).
- Talman, V. *et al.*, The C1 domain-targeted isophthalate derivative HMI-1b11 promotes neurite outgrowth and GAP-43 expression through PKC $\alpha$  activation in SH-SY5Y cells. *Pharmacol Res.* **73**, 44–54 (2013).
- Campos-Melo, D., Droppelmann, C. A., Volkening, K. & Strong, M. J. RNA-binding proteins as molecular links between cancer and neurodegeneration. *Biogerontology*. **15**, 587–610 (2014).
- König, J., Zarnack, K., Luscombe, N. M. & Ule, J. Protein-RNA interactions: new genomic technologies and perspectives. *Nat. Rev. Genet.* **18**, 77–83 (2012).
- Doxakis, E. RNA binding proteins: a common denominator of neuronal function and dysfunction. *Neurosci. Bull.* **1**, 610–626 (2014).
- Antic, D. & Keene, J. D. Embryonic lethal abnormal visual RNA-binding proteins involved in growth, differentiation, and posttranscriptional gene expression. *Am. J. Hum. Genet.* **61**, 273 (1997).
- Nagai, K., Oubridge, C., Ito, N., Avis, J. & Evans, P. The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *Trends Biochem. Sci.* **20**, 235–240 (1995).
- Fan, X. C. & Steitz, J. A. HNS, a nuclear-cytoplasmic shuttling sequence in HuR. *Proc. Natl. Acad. Sci. USA* **95**, 15293–15298 (1998).
- Hinman, M. N. & Lou, H. Diverse molecular functions of Hu proteins. *Cell Mol. Life Sci.* **65**, 3168–3181 (2008).
- Doller, A. & Pfeilschifter, J. Regulation of the mRNA-binding protein HuR by posttranslational modification: spotlight on phosphorylation. *Curr. Protein Pept. Sci.* **13**, 380–390 (2012).
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*. **348**, 515–520 (1990).
- Wang, X. & Tanaka Hall, T. M. Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat. Struct. Biol.* **8**, 141–145 (2001).
- Kasashima, K., Sakashita, E., Saito, K. & Sakamoto, H. Complex formation of the neuron-specific ELAV-like Hu RNA-binding proteins. *Nucleic Acids Res.* **30**, 4519–4526 (2002).
- Toba, G. & White, K. The third RNA recognition motif of Drosophila ELAV protein has a role in multimerization. *Nucleic Acids Res.* **36**, 1390–1399 (2008).
- Diaz-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).
- Heinonen, M. *et al.* Cytoplasmic HuR expression is a prognostic factor in invasive ductal breast carcinoma. *Cancer Res.* **65**, 2157–2161 (2005).

17. Huang, Y. H. *et al.* Insights from HuR biology point to potential improvement for second-line ovarian cancer therapy. *Oncotarget* **7**, 21812–21824 (2016).
18. Young, L. E. *et al.* The mRNA binding proteins HuR and tristetraprolin regulate cyclooxygenase2 expression during colon carcinogenesis. *Gastroenterology* **136**, 1669–1679 (2009).
19. Abdelmohsen, K. *et al.* miR-519 suppresses tumor growth by reducing HuR levels. *Cell Cycle* **9**, 1354–1359 (2010).
20. Abdelmohsen, K., Srikantan, S., Kuwano, Y. & Gorospe, M. miR-519 reduces cell proliferation by lowering RNA-binding protein HuR levels. *Proc. Natl. Acad. Sci. USA* **105**, 20297–20302 (2008).
21. Meisner, N. C. *et al.* Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat. Chem. Biol.* **3**, 508–515 (2007).
22. Nasti, R. *et al.* Compounds Interfering with Embryonic Lethal Abnormal Vision (ELAV) Protein-RNA Complexes: An Avenue for Discovering New Drugs. *J. Med. Chem.* **60**, 8257–8267 (2017).
23. D'Agostino, V. G. *et al.* Dihydrotanshinone-I interferes with the RNA-binding activity of HuR affecting its posttranscriptional function. *Sci. Rep.* **5**, 1–15 (2015).
24. Lal, P. *et al.* Regulation of HuR structure and function by dihydrotanshinone-I. *Nucleic Acids Res.* **45**, 9514–9527 (2017).
25. Filippova, N. *et al.* Hu antigen R (HuR) multimerization contributes to glioma disease progression. *J. Biol. Chem.* **292**, 16999–17010 (2017).
26. Manzoni, L. *et al.* Interfering with HuR-RNA interaction: Design, synthesis and biological characterization of Tanshinone mimics as novel, effective HuR inhibitors. *J. Med. Chem.* **61**, 1483–1498 (2018).
27. Rossi, D. *et al.* Discovery of small peptides derived from Embryonic Lethal Abnormal Vision proteins structure showing RNA-stabilizing properties. *J. Med. Chem.* **52**, 5017–5019 (2009).
28. Amadio, M. *et al.* Identification of peptides with ELAV-like mRNA-stabilizing effect: an integrated *in vitro/in silico* approach. *Chem. Biol. Drug. Des.* **707**–714 (2013).
29. Vasile, F., Rossi, D., Collina, S. & Potenza, D. Diffusion-Ordered Spectroscopy and Saturation Transfer Difference NMR spectroscopy studies of selective interactions between ELAV protein fragments and an mRNA target. *Eur. J. Org. Chem.* **6399**–6404 (2014).
30. Gatti, L. *et al.* Antitumor activity of a novel homodimeric SMAC mimetic in ovarian carcinoma. *Molecular Pharmacology*, **11**(1), 283–93 (2014).
31. Vasile, F. *et al.* Comprehensive analysis of blood group antigen binding to classical and El Tor cholera toxin B-pentamers by NMR. *Glycobiology* **24**, 766–778 (2014).
32. Heggelund, J. E. *et al.* Both El Tor and classical cholera toxin bind blood group determinants. *Biochem. Biophys. Res. Comm.* **418**, 731–735 (2012).
33. Vasile, F. *et al.* NMR interaction studies of Neu5Ac- $\alpha$ -(2,6)-Gal- $\beta$ -(1-4)-GlcNAc with influenza-virus Hemagglutinin expressed in transfected human cells. *Glycobiology* **28**(1), 42–49 (2018).
34. Guzzetti, I. *et al.* Insights into the binding of cyclic RGD peptidomimetics to  $\alpha 5 \beta 1$  integrin by live cell NMR and computational studies. *Chem. Open* **6**, 128–136 (2017).
35. Zega, A. NMR Methods for Identification of False Positives in Biochemical Screens. *J. Med. Chem.* **60**, 9437–9447 (2017).
36. Proudfoot, A., Bussiere, D. E. & Lingel, A. High-Confidence Protein-Ligand Complex Modeling by NMR-Guided Docking Enables Early Hit Optimization. *J. Am. Chem. Soc.* **139**, 17824–17833 (2017).
37. Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* **38**, 1784–1788 (1999).
38. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.* **42**, 864–890 (2003).
39. Angulo, J. *et al.* Blood Group B Galactosyltransferase: Insights into Substrate Binding from NMR Experiments. *J. Am. Chem. Soc.* **128**, 13529–13538 (2006).
40. Angulo, J., Enriquez-Navas, P. M. & Nieto, P. M. Ligand-receptor binding affinities from saturation transfer difference (STD) NMR spectroscopy: the binding isotherm of STD initial growth rates. *Chem. Eur. J.* **16**, 7803–7812 (2010).
41. Zhao, H. & Clafisch, A. Molecular dynamics in drug design. *Eur. J. Med. Chem.* **91**, 4–14 (2015).
42. DE Vivo, M., Masetti, M., Bottegoni, G. & Cavalli, A. Role of Molecular Dynamics and Related methods in Drug discovery. *J. Med. Chem.* **59**, 4035–4061 (2016).
43. Wang, H. *et al.* The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr.* **69**, 373–380 (2013).
44. Alcaro, S. *et al.* Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches. *J. Med. Chem.* **56**, 843–855 (2013).
45. Costa, G., Gidaro, M. C., Vullo, D., Supuran, C. & Alcaro, S. The Essential Oils as Resources of Anti-Obesity Potential Drugs investigated by *in silico* techniques. *J. Agricult. Food Chem.* **64**, 5295–5300 (2016).
46. Kaur, K. *et al.* The fungal natural product azaphilone-9 binds to HuR and inhibits HuR-RNA interaction *in vitro*. *PlosOne* **12** (2017).
47. Tang, A. Y. RNA processing-associated molecular mechanisms of neurodegenerative diseases. *J. Appl. Genetic* **57**, 323–333 (2016).
48. Zucal, C. *et al.* Targeting the multifaceted HuR protein, benefits and caveats. *Curr. Drug Targets* **16**, 499–515 (2015).
49. Chae, M. J. *et al.* Chemical inhibitors destabilize HuR binding to the AU-rich element of TNF- $\alpha$  mRNA. *Exp. Mol. Med.* **41**, 824–831 (2009).
50. 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.* **40**, 5088–100 (2012).
51. <https://www.rcsb.org/structure/4ED5>
52. Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2017.
53. Maestro, Schrödinger, LLC, New York, NY, 2017.
54. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
55. Prime, Schrödinger, LLC, New York, NY, 2017.
56. MacroModel, Schrödinger, LLC, New York, NY, 2017.
57. Desmond Molecular Dynamics System, D.E. Shaw Research, New York, NY, 2017.
58. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2017.
59. Glide, Schrödinger, LLC, New York, NY, 2017.
60. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

## Acknowledgements

The authors would like to mention Linnea SA, in appreciation of some compounds donated for the present study. SDV and SC thankfully recognise Scuola di Alta Formazione Dottorale of University of Pavia for the mobility research scholarship provided. AKHH gratefully acknowledges funding from the Helmholtz-Association's Initiative and Networking Fund. AP gratefully acknowledges funding from Associazione Italiana per la Ricerca sul Cancro (AIRC) [17153], CARITRO Riposizionamento Farmaci [40102838] and CARIPO, ricerca biomedica sulle malattie legate all'invecchiamento, 40102636].

### Author Contributions

S. Collina conceived the work. D. Potenza, S. Alcaro and S. Collina designed and organized this research. F. Vasile, S. Della Volpe were responsible for the NMR experiments. A. K. H. Hirsch was responsible for the initial docking studies and M. Y. Unver performed them. F. A. Ambrosio, G. Costa, performed the molecular modelling study. A. Provenzani was responsible for the HuR production and characterization. E. Martino was responsible for the natural product selection. C. Zucal prepared and purified the protein. F. Vasile, S. Della Volpe, G. Costa and D. Rossi analysed the obtained data and interpreted the results with the input from other co-authors. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-32084-z>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

## Article 2: Novel compounds targeting the RNA-binding protein HuR. Structure-based design, synthesis and interaction studies

Reprinted with permission from (*ACS Med. Chem. Lett.* 2019, 10, 4, 615-620;

<https://doi.org/10.1021/acsmchemlett.8b00600>). Copyright (2019) American Chemical Society

### Novel Compounds Targeting the RNA-Binding Protein HuR. Structure-Based Design, Synthesis, and Interaction Studies

Serena Della Volpe,<sup>†,‡</sup> Rita Nasti,<sup>†,‡</sup> Michele Queirolo,<sup>†</sup> M. Yagiz Unver,<sup>‡,§</sup> Varsha K. Jumde,<sup>‡,§</sup> Alexander Dömling,<sup>||,Ⓞ</sup> Francesca Vasile,<sup>\*,⊥</sup> Donatella Potenza,<sup>⊥</sup> Francesca Alessandra Ambrosio,<sup>◆</sup> Giosué Costa,<sup>◆,Ⓞ</sup> Stefano Alcaro,<sup>◆</sup> Chiara Zucal,<sup>Ⓞ</sup> Alessandro Provenzani,<sup>Ⓞ</sup> Marcello Di Giacomo,<sup>†</sup> Daniela Rossi,<sup>†</sup> Anna K. H. Hirsch,<sup>\*,‡,§</sup> and Simona Collina<sup>\*,†,Ⓞ</sup>

<sup>†</sup>Department of Drug Sciences, Medicinal Chemistry and Technology Section, University of Pavia, Via Taramelli 12, 27100 Pavia, Italy

<sup>‡</sup>Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) - Helmholtz Centre for Infection Research (HZI), Department of Drug Design and Optimization and Department of Pharmacy, Saarland University, Campus Building E8.1, 66123 Saarbrücken, Germany

<sup>§</sup>Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, NL-9747 AG, Groningen, The Netherlands

<sup>||</sup>Department of Drug Design, University of Groningen, A. Deusinglaan 1, Groningen, 9713 AV, The Netherlands

<sup>⊥</sup>Department of Chemistry, University of Milan, Via Golgi 19, 20133 Milano, Italy

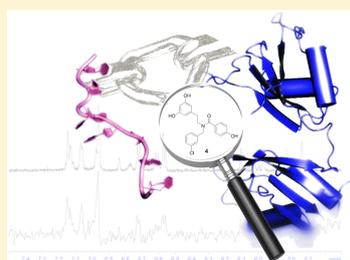
<sup>◆</sup>Department of Health Sciences, University "Magna Græcia" of Catanzaro, Viale Europa, 88100, Catanzaro, Italy

<sup>Ⓞ</sup>Department of CIBIO, University of Trento, Via Sommarive 9, 38123 Povo, TN, Italy

#### Supporting Information

**ABSTRACT:** The key role of RNA-binding proteins (RBPs) in regulating post-transcriptional processes and their involvement in several pathologies (*i.e.*, cancer and neurodegeneration) have highlighted their potential as therapeutic targets. In this scenario, Embryonic Lethal Abnormal Vision (ELAV) or Hu proteins and their complexes with target mRNAs have been gaining growing attention. Compounds able to modulate the complex stability could constitute an innovative pharmacological strategy for the treatment of numerous diseases. Nevertheless, medicinal-chemistry efforts aimed at developing such compounds are still at an early stage. As part of our ongoing research in this field, we hereby present the rational design and synthesis of structurally novel HuR ligands, potentially acting as HuR–RNA interferers. The following assessment of the structural features of their interaction with HuR, combining saturation-transfer difference NMR and *in silico* studies, provides a guide for further research on the development of new effective interfering compounds of the HuR–RNA complex.

**KEYWORDS:** RNA-binding protein, HuR–RNA complexes, STD-NMR, virtual screening, multicomponent reactions



RNA is an important regulatory element of many cellular processes, and thus, so-called RNA-binding proteins (RBPs), play a prominent role in affecting the fate of target mRNAs coding for proteins pivotal in key cellular functions.<sup>1–5</sup> Therefore, a dysregulation of RBPs may be related to the pathogenesis of several diseases.<sup>2,4,6</sup> In 2018, the first candidate drug targeting RBPs, called H3B-8800, reached the clinical phase for the treatment of acute myelogenous leukemia and chronic myelomonocytic leukemia.<sup>7</sup> Among RBPs, the family of ELAV (Embryonic Lethal Abnormal Vision) proteins is involved in controlling the functional activities of diverse RNA populations. In particular, HuR regulates splicing, stability, and translation of thousands of coding and noncoding RNAs and is therefore considered a valid drug target for anticancer therapy.<sup>8</sup> HuR is a nuclear protein, but upon cell stress such as DNA

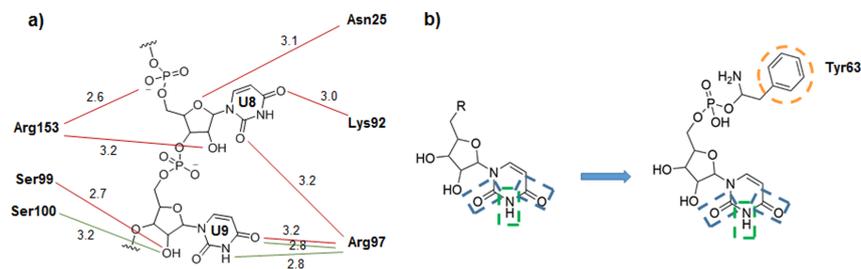
damage, it shuttles into the cytoplasm where it regulates the fate of cargo mRNAs and determines the abundance of the encoded proteins. Overexpression of HuR is associated with tumor progression and poor prognosis in various cancer types.<sup>8</sup> For this reason, compounds able to bind HuR and to inhibit the formation of the HuR–RNA complex may have anticancer properties. So far, various high-throughput screening campaigns afforded several natural products.<sup>9–12</sup>

**Special Issue:** Highlighting Medicinal Chemistry in Italy

**Received:** November 30, 2018

**Accepted:** January 21, 2019

**Published:** January 21, 2019



**Figure 1.** (a) HuR–RNA contacts (PDB code 4ED5) for U8–U9. Red lines show side-chain contacts; green lines show main-chain contacts. Distances are expressed in Å.<sup>20</sup> (b) Modification applied to the anchor to increase the predicted affinity. Dashed blue rectangles indicate hydrogen-bond acceptors and dashed green rectangles hydrogen-bond donors on both the original and modified anchor, while the dashed orange circle shows the anchor modification potentially allowing for additional hydrophobic interactions with the selected Tyr63 residue.

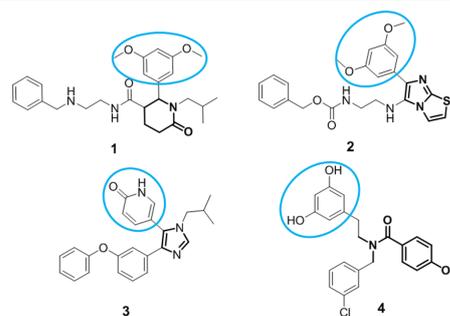
After studying the concept of druggability of ELAV proteins and related complexes with mRNA,<sup>13–15</sup> we analyzed the literature concerning the main findings on the ELAV–RNA complexes from a medicinal-chemistry standpoint and defined the interaction features of HuR and a small series of natural products.<sup>16,17</sup> In this Letter, we report on the structure-based design and the synthesis of compounds with different “core structures” and the investigation of their interactions with HuR using a combination of saturation-transfer difference (STD)-NMR and *in silico* studies. While on the RNA side the interaction of HuR and RNA is mediated by AU-rich elements (AREs), on the protein side two consecutive RNA recognition motif-type (RRM1 and RRM2) domains are involved; these interact directly with target RNAs through highly conserved ribonucleoprotein (RNP1 and RNP2) sequences.<sup>18,19</sup> For the design of new compounds, we focused on a pocket-like region hosting small HuR ligands, formed by the two asymmetric units of the protein, which belong to RNP1 and RNP2 of RRM1. Inspection of the cocrystal structure of the HuR RRM1 and RRM2 domains in complex with the ARE sequence of RNA<sup>c-fos</sup> (PDB code 4ED5) confirms that this region corresponds to binding site of RNA uridine residues 8 and 9 (U8–U9) (Figure 1a).

We based our search for new scaffolds for HuR ligands on the key interactions shown in Figure 1a and employed the free web-based virtual screening platform AnchorQuery.<sup>21–24</sup> This program is specifically designed for targeting protein–protein interactions (PPIs) with small molecules by combining the anchor concept with one-pot multicomponent reaction (MCR) chemistry. Briefly, PPI inhibitors are characterized by specific moieties able to mimic amino acid side chains of the donor protein, called “anchor motifs”. Since the contact surfaces involved in PPIs are typically large and flat, similar to protein–RNA interactions, we used NucleoQuery, a nucleoside derivative of the web application AnchorQuery, for the rational structure-based design of HuR–RNA targeting compounds. We selected U8 and U9 as anchors and investigated whether one of the two nucleotides may be a pharmacophore in terms of occupied position and interactions. To increase the probability of discovering new ligands, the anchors were modified so as to bear an additional phenyl group to establish a  $\pi$ – $\pi$ -stacking interaction with Tyr63 (Figure 1b). As a result of four runs in NucleoQuery, a large library (800 molecules) featuring a wide range of structurally diverse derivatives was obtained. To select the candidates for synthesis, we relied on visual inspection and molecular-

recognition studies. This way, we selected 17 compounds featuring piperidinones, aromatic heterocycles, *N,N*-disubstituted amides, and sulfonamides as scaffolds (see SI, S1.2). For all compounds, the anchor shows good overlap with the corresponding uracil ring, preserving the same hydrogen bonds with Arg97 and Lys92, while the additional aromatic rings should afford the desired interaction with Tyr63. Moreover, selected compounds are engaged in additional interactions with the protein, including hydrophobic interactions with residues Ile23, Asn25, Phe65, Ile103, and Ile133 (docking poses of exemplary compounds for each scaffold can be found in SI, S1.3). Taking into account the synthetic feasibility and commercial availability of the corresponding starting materials, we selected the most promising scaffolds and replaced the uracil moiety with an easy-to-handle starting material. We evaluated several hydrophobic/aromatic portions as anchors, potentially able to interact with HuR in the position occupied by U8 and preserving its main contacts. Thus, we designed and synthesized compounds 1–4 (Figure 2), characterized by different scaffolds.

The computationally predicted structures can be easily prepared by a short synthesis using MCR or equally efficient processes.

Thus, compound 1 was synthesized by a Castagnoli–Cushman reaction (CC-3CR)<sup>25</sup> followed by amidation; 2 by a Groebke–Blackburn–Bienaymé (GBB-2CR);<sup>26</sup> 3 by a Van



**Figure 2.** Structures of the four compounds selected for the experimental study. The modified anchors are highlighted with light blue circles. The different scaffolds (piperidinones, heterocycles, *N,N*-disubstituted amides) are highlighted in bold.

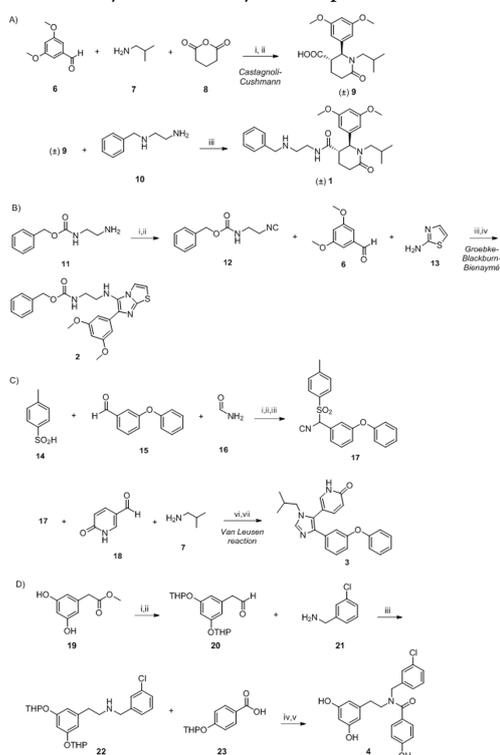
B

DOI: 10.1021/acsmedchemlett.8b00600  
ACS Med. Chem. Lett. XXXX, XXX, XXX–XXX

Leusen reaction (vL-3CR),<sup>27</sup> and **4** by a reductive amination<sup>28</sup> followed by acylation. Herein, we will briefly describe the synthetic strategies adopted to obtain compounds **1–4**. Further details on protocols employed and compound characterization are reported in the [Supporting Information](#) (SI, S2).

Amidation of the acid intermediate *trans* ( $\pm$ )-**9** with amine **10** afforded compound ( $\pm$ )-**1** as racemate.<sup>29</sup> The key intermediate *trans*-6-oxopiperidine-3-carboxylic acid (**9**) was obtained in good yield and diastereoselectivity, applying a microwave-aided one-pot CC-3CR on benzaldehyde (**6**), isobutylamine (**7**), and glutaric anhydride (**8**) (Scheme 1a).<sup>25</sup> A microwave-assisted GBB-2CR,<sup>26,30</sup> involving isocyanide **12**, aldehyde **6**, and amino-thiazole **13** produced compound **2**. Isocyanide **12** was obtained starting from amine **11**, which was converted into the corresponding

### Scheme 1. Synthetic Pathways of Compounds **1–4**<sup>a</sup>

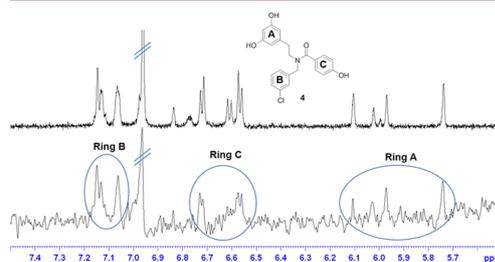


<sup>a</sup>Reagents and conditions: (a) Compound ( $\pm$ )-**1**: (i) **6**, **7**, HCOOH, ACN, mw 120 °C, 50 W, 30 min; (ii) **8**, *p*-xylene, reflux, 10 h, yield 62%; (iii) TBTU, DIPEA, THF, rt, 12 h, yield 71%. (b) Compound **2**: (i) Ethyl formate, 10 h; (ii) TEA, POCl<sub>3</sub>, DCM, rt, 4 h, yield 62%; (iii) **6**, **13**, HCOOH, ACN, mw 120 °C, 30 min; (iv) **12**, ZrCl<sub>4</sub>, 80 °C, 10 h, yield 21%. (c) Compound **3**: (i) H<sub>2</sub>O, 30 min; (ii) HCl, MTBE, 30 min; (iii) **15**, **16** Me<sub>2</sub>SiCl, ACN, toluene, N<sub>2</sub>, 50 °C, 5 h; (iv) **14**, 16 h, yield 62%; (v) POCl<sub>3</sub>, triethylamine, THF, 10 °C, 45 min, yield 85%; (vi) **18**, 7 dry DMF, rt, 4 h; (vii) **17**, K<sub>2</sub>CO<sub>3</sub>, rt, 72 h, yield 15%.

formamide and subsequently dehydrated (Scheme 1b). A vL-3CR<sup>27</sup> on commercially available aldehyde **18**, isobutylamine (**7**), and substituted TosMIC derivative **17** led to compound **3**. This MCR allows access to functionally rich imidazoles in a single pot via cycloaddition of TosMIC reagents on imines, generated *in situ* from an aldehyde and an amine under mildly basic conditions. We synthesized **17** starting from *p*-toluenesulfonic acid (**14**), which afforded the intermediate formamide derivative upon reaction with commercially available 3-phenoxybenzaldehyde (**15**), formamide (**16**), and chlorotrimethylsilane; subsequent dehydration furnished isocyanide **17** (Scheme 1c).<sup>31</sup> Compound **4** was obtained by reductive amination of aldehyde **20** followed by an amide coupling with acid **23** followed by removal of the THP protecting groups (Scheme 1d).<sup>28</sup> Prior to the interaction study, compounds **1–3** were converted into the corresponding hydrochloride salts, and their solubility in both the buffer and time-ranges required for STD-NMR experiments was evaluated. Under these conditions, compound **2** proved to be poorly soluble and had to be excluded from the corresponding interaction study.

The interaction of compounds **1**, **3**, and **4** with HuR was studied by STD-NMR spectroscopy. This technique, based on the nuclear Overhauser effect, is a well-established epitope-mapping methodology for studying the target–ligand interactions.<sup>32</sup> Briefly, the method relies on the selective irradiation of the protein, which allows magnetization to be transferred to the bound ligand; the observation of the ligand signals in the NMR spectrum provides an indication of the interaction. Those ligand protons that are nearest to the protein are more likely to become highly saturated and therefore show the strongest signal in the monodimensional STD spectrum. Therefore, the intensity of the STD signal (expressed as absolute STD percentage) reflects the proximity of the ligand to the protein surface.<sup>33–38</sup> The group epitope mapping illustrates which chemical moieties of the ligand are key for molecular recognition in the binding site. The analysis of STD data shows that compounds **1** (Figure S3.2.1 and Table S3.2.1), **3** (Figure S3.2.2 and Table S3.2.2), and **4** (Figure 3 and Table S3.2.3) interact with the protein.

For ligand **1**, the isopropyl group shows the strongest interaction (0.40 STD %), while the STD signal related to ring B is the least intense (0.03 STD %); ring A contributes to the interaction (0.20 STD %), while the protons of the amino chain (H14, H15, and H16) do not give STD signals. The binding of compound **3** to HuR is mediated by the pyridone



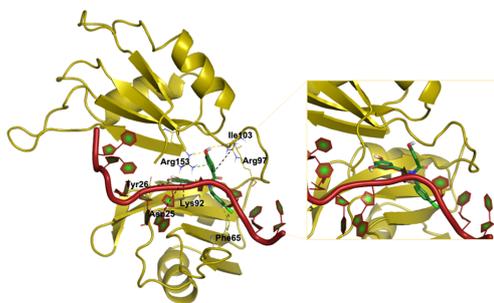
**Figure 3.** STD-NMR spectrum (bottom) and <sup>1</sup>H-NMR (top) of compound **4**. The strong interaction observed for aromatic rings A, B, and C is evidenced.

C

DOI: 10.1021/acsmchemlett.8b00600  
ACS Med. Chem. Lett. XXXX, XXX, XXX–XXX

anchor (1.20 STD %) and by the isopropyl moiety (0.30 STD %). For compound **4**, we observed STD signals of similar intensity for the three aromatic rings (0.40% for ring A, 0.36% for ring B, and 0.25% for ring C, respectively), suggesting that they are positioned within a protein pocket.

In parallel, we performed docking studies on the crystal structure of HuR RRM1-2 domains according to the approach we published.<sup>17</sup> We used STD-NMR in combination with *in silico* studies (applied to the “closed” HuR conformations, see SI) to enable a more detailed description of the ligand–protein behavior in solution. Molecular docking simulations were performed to elucidate the binding mode and the interactions between compounds **1**, **3**, and **4** with HuR protein, using Maestro tools.<sup>39</sup> We observed that the anchor moiety of each compound is predicted to be superimposed with U8, establishing pivotal interactions with the HuR protein. The interactions of compounds **1** and **3** with HuR is described in the SI (Section S4.3, Figures S4.3.1 and S4.3.2). As for compound **4**, we report strong hydrophobic interactions for all aromatic rings as evidenced by STD data. The amide carbonyl group establishes double hydrogen bonds with Arg97 and hydrophobic interactions with the same residue and Ile103 (Figure 4). Regarding ring B, it establishes a  $\pi$ – $\pi$ -stacking



**Figure 4.** 3D representation of HuR–**4** interactions. Hydrogen bonds and hydrophobic interactions are displayed as yellow and dark dashed lines, respectively; a  $\pi$ – $\pi$ -stacking interaction is shown as a blue dashed line. RNA and HuR are represented as a brown and yellow cartoon, respectively. Compound **4** is shown in green sticks.

interaction with Phe65 and hydrophobic interactions with Arg97, Pro98, and Phe65. Ring C is involved in a double hydrogen bond with Arg153 and Ile103 and different hydrophobic interactions. The zoomed-in view in Figure 4 highlights the full overlap of U8 and the anchor motif (ring A) of compound **4**.

In this Letter, we report for the first time on the structure-based design and synthesis of potential HuR ligands. The structure-based design strategy used in the present study enabled the identification of three novel compounds that interact with HuR, establishing contacts with the RNP regions of RRM1 and RRM2 domains of the protein. Targeting this region is pivotal for the inhibition of the protein activity, and so the identification of compounds able to bind to this site plays a critical role in the development of new potential HuR inhibitors. Biophysical experiments (STD-NMR) are in agreement with computational results, confirming that the designed compounds **1**, **3**, and **4** are indeed binders of HuR. Furthermore, theoretical experiments carried out by molecular

dynamics and docking studies confirm that compound **4** is the best binder of HuR protein, in terms of theoretical binding affinity (SI, S4.2). The identified hit compounds will be further optimized in forthcoming studies, building an *in silico* focused library. Our findings represent a step forward in discovering novel compounds potentially useful to counteract several pathologies such as cancer, inflammation, and neurodegeneration, in which ELAV–RNA complexes play a pivotal role.

## EXPERIMENTAL PROCEDURES

**Design of New Scaffolds.** We report all protocols in the SI, S1.

**Compound Synthesis.** Procedures are reported in the SI, S2.

**Protein Expression and Purification.** HuR aliquots utilized in the STD-NMR study were prepared as already described.<sup>10,40</sup>

**Interaction Study with HuR.** STD-NMR experiments, molecular dynamics simulations, and docking studies were carried out as reported in the SI (S3 and S4, respectively).

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsmedchemlett.8b00600.

Protocols (compound design, synthesis, characterization, and interaction studies) (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*S.C. e-mail: [simona.collina@unipv.it](mailto:simona.collina@unipv.it).

\*F.V. e-mail: [francesca.vasile@unimi.it](mailto:francesca.vasile@unimi.it).

\*A.K.H.H. e-mail: [Anna.Hirsch@helmholtz-hzi.de](mailto:Anna.Hirsch@helmholtz-hzi.de).

### ORCID

Alexander Dömling: 0000-0002-9923-8873

Giosué Costa: 0000-0003-0947-9479

Alessandro Provenzani: 0000-0003-1652-3415

Simona Collina: 0000-0002-2954-7558

### Present Address

R.N.: Department of Environmental Science and Policy (ESP), University of Milan, 20133 Milano, Italy. M.Q.: Agro Business Park 10, 6708 PW Wageningen, The Netherlands.

### Author Contributions

\*S.D.V. and R.N. contributed equally to the work. 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

F.V. and S.C. acknowledge Federico Sala for his collaboration in the NMR laboratory. S.D.V. and S.C. thankfully recognize Scuola di Alta Formazione Dottorale (University of Pavia) for the mobility research scholarship. R.N. and M.Q. acknowledge the Erasmus Traineeship program for the experience opportunity at the Stratingh Institute for Chemistry, University of Groningen. A.K.H.H. gratefully acknowledges funding from the Helmholtz-Association's Initiative and Networking Fund. A.P. acknowledges Associazione Italiana per la Ricerca sul Cancro (AIRC) [21548].

D

DOI: 10.1021/acsmedchemlett.8b00600  
ACS Med. Chem. Lett. XXXX, XXX, XXX–XXX

## ■ ABBREVIATIONS

ELAV, embryonic lethal abnormal vision; RBP, RNA-binding protein; STD-NMR, saturation transfer difference-NMR; RRM, RNA recognition motif-type; RNP, ribonucleoprotein; MCR, multicomponent reaction; CC-3CR, Castagnoli–Cushman reaction; GBB-2CR, Groebke–Blackburn–Bienaymé reaction; vL-3CR, Van Leusen reaction.

## ■ REFERENCES

- (1) Pascale, A.; Govoni, S. The complex world of posttranscriptional mechanisms: Is their deregulation a common link for diseases? Focus on ELAV-like RNA-binding proteins. *Cell. Mol. Life Sci.* **2012**, *69*, 501–517.
- (2) Talman, V.; Amadio, M.; Osera, C.; Sorvari, S.; Boije af Gennäs, G.; Yli-Kauhaluoma, J.; Rossi, D.; Govoni, S.; Collina, S.; Ekokoski, E.; Tuominen, R. K.; Pascale, A. The C1 domain-targeted isophthalate derivative HMI-1b11 promotes neurite outgrowth and GAP-43 expression through PKC $\alpha$  activation in SH-SY5Y cells. *Pharmacol. Res.* **2013**, *73*, 44–54.
- (3) Campos-Melo, D.; Droppelmann, C. A.; Volkening, K.; Strong, M. J. RNA-binding proteins as molecular links between cancer and neurodegeneration. *Biogerontology* **2014**, *15*, 587–610.
- (4) König, J.; Zarnack, K.; Luscombe, N. M.; Ule, J. Protein-RNA interactions: new genomic technologies and perspectives. *Nat. Rev. Genet.* **2012**, *18*, 77–83.
- (5) Doxakis, E. RNA binding proteins: a common denominator of neuronal function and dysfunction. *Neurosci. Bull.* **2014**, *1*, 610–626.
- (6) Hong, S. RNA Binding Protein as an Emerging Therapeutic Target for Cancer Prevention and Treatment. *J. Cancer. Prev.* **2017**, *22*, 203–210.
- (7) Seiler, M.; Yoshimi, A.; Darman, R.; Chan, B.; Keaney, G.; Thomas, M.; Agrawal, A. A.; Caleb, B.; Csibi, A.; Sean, E.; Fekkes, P.; Karr, C.; Klimek, V.; Lai, G.; Lee, L.; Kumar, P.; Lee, S. C.; Liu, X.; Mackenzie, C.; Meeske, C.; Mizui, Y.; Padron, E.; Park, E.; Pazolli, E.; Peng, S.; Prajapati, S.; Taylor, J.; Teng, T.; Wang, J.; Warmuth, M.; Yao, H.; Yu, L.; Zhu, P.; Abdel-Wahab, O.; Smith, P. G.; Buonamici, S. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **2018**, *24*, 497–504.
- (8) Filippova, N.; Yang, X.; Ananthan, S.; Sorochinsky, A.; Hackney, J.; Gentry, Z.; Bae, S.; King, P.; Nabors, L. B. Hu antigen R (HuR) multimerization contributes to glioma disease progression. *J. Biol. Chem.* **2017**, *292*, 16999–17010.
- (9) Meisner, N. C.; Hintersteiner, M.; Mueller, K.; Bauer, R.; Seifert, J. M.; Naegeli, H. U.; Ottl, J.; Oberer, L.; Guenat, C.; Moss, S.; Harrer, N.; Woisetschlaeger, M.; Buehler, C.; Uhl, V.; Auer, M. Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat. Chem. Biol.* **2007**, *3*, 508–515.
- (10) D'Agostino, V. G.; Lal, P.; Mantelli, B.; Tiedje, C.; Zucal, C.; Thongon, N.; Gaestel, M.; Latorre, E.; Marinelli, L.; Seneci, P.; Amadio, M.; Provenzani, A. Dihydro-tanshinone-I interferes with the RNA-binding activity of HuR affecting its posttranscriptional function. *Sci. Rep.* **2015**, *5*, 1–15.
- (11) Lal, P.; Cerofolini, L.; D'Agostino, V. G.; Zucal, C.; Fuccio, C.; Bonomo, I.; Dassi, E.; Giuntini, S.; Di Maio, D.; Vishwakarma, V.; Preet, R.; Williams, S. N.; Fairlamb, M. S.; Munk, R.; Lehrmann, E.; Abdelmohsen, K.; Elezgarai, S. R.; Luchinat, C.; Novellino, E.; Quattrone, A.; Biasini, E.; Manzoni, L.; Gorospe, M.; Dixon, D. A.; Seneci, P.; Marinelli, L.; Fragai, M.; Provenzani, A. Regulation of HuR structure and function by dihydro-tanshinone-I. *Nucleic Acids Res.* **2017**, *45*, 9514–9527.
- (12) Manzoni, L.; Zucal, C.; Di Maio, D.; D'Agostino, V. G.; Thongon, N.; Bonomo, I.; Lal, P.; Miceli, M.; Baj, V.; Brambilla, M.; Cerofolini, L.; Elezgarai, S.; Biasini, E.; Luchinat, C.; Novellino, E.; Fragai, M.; Marinelli, L.; Provenzani, A.; Seneci, P. 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.
- (13) Rossi, D.; Amadio, M.; Carnevale Baraglia, A.; Azzolina, O.; Ratti, A.; Govoni, S.; Pascale, A.; Collina, S. Discovery of small peptides derived from Embryonic Lethal Abnormal Vision proteins structure showing RNA stabilizing properties. *J. Med. Chem.* **2009**, *52*, 5017–5019.
- (14) Amadio, M.; Pascale, A.; Govoni, S.; Laurini, E.; Pricl, S.; Gaggeri, R.; Rossi, D.; Collina, S. Identification of peptides with ELAV-like mRNA-stabilizing effect: an integrated in vitro/in silico approach. *Chem. Biol. Drug Des.* **2013**, *81*, 707–714.
- (15) Vasile, F.; Rossi, D.; Collina, S.; Potenza, D. Diffusion-Ordered Spectroscopy and Saturation Transfer Difference NMR spectroscopy studies of selective interactions between ELAV protein fragments and an mRNA target. *Eur. J. Org. Chem.* **2014**, *2014*, 6399–6404.
- (16) Nasti, R.; Rossi, D.; Amadio, M.; Pascale, A.; Unver, M. Y.; Hirsch, A. K. H.; Collina, S. Compounds Interfering with Embryonic Lethal Abnormal Vision (ELAV) Protein-RNA Complexes: An Avenue for Discovering New Drugs. *J. Med. Chem.* **2017**, *60*, 8257–8267.
- (17) Vasile, F.; Della Volpe, S.; Ambrosio, F. A.; Costa, G.; Unver, M. Y.; Zucal, C.; Rossi, D.; Martino, E.; Provenzani, A.; Hirsch, A. K. H.; Alcaro, S.; Potenza, D.; Collina, S. Exploration of ligand binding modes towards the identification of compounds targeting HuR: a combined STD-NMR and Molecular Modelling approach. *Sci. Rep.*, published online September 23, **2018**, DOI: 10.1038/s41598-018-32084-z.
- (18) Nagai, K.; Oubridge, C.; Jessen, T. H.; Li, J.; Evans, P. R. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **1990**, *348*, 515–520.
- (19) Wang, X.; Tanaka Hall, T. M. Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat. Struct. Biol.* **2001**, *8*, 141–145.
- (20) Wang, H.; Zeng, F.; Liu, Q.; Liu, H.; Liu, Z.; Niu, L.; Teng, M.; Li, X. The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 373–380.
- (21) <http://anchorquery.csb.pitt.edu/>.
- (22) Koes, D.; Khoury, K.; Huang, Y.; Wang, W.; Bista, M.; Popowicz, G. M.; Wolf, S.; Holak, T. A.; Dömling, A.; Camacho, C. J. Enabling large-scale design, synthesis and validation of small molecule protein-protein antagonists. *PLoS One*, published online March 12, **2012**, *7*, e32839.
- (23) Abdelraheem, E. M. M.; Camacho, C. J.; Dömling, A. Focusing on shared subpockets - New developments in fragment-based drug discovery. *Expert Opin. Drug Discovery* **2015**, *10*, 1179–1187.
- (24) Koes, D. R.; Dömling, A.; Camacho, C. J. AnchorQuery: Rapid online virtual screening for small-molecule protein-protein interaction inhibitors. *Protein Sci.* **2018**, *27*, 229–232.
- (25) Kroon, E.; Schulze, J. O.; Süß, E.; Camacho, C. J.; Biondi, R. M.; Dömling, A. Discovery of a Potent Allosteric Kinase Modulator by Combining Computational and Synthetic Methods. *Angew. Chem., Int. Ed.* **2015**, *54*, 13933–13936.
- (26) Groebke, K.; Weber, L.; Mehlin, F. A fast heterocyclic three component synthesis of imidazo[1,2-a]annulated pyridines, pyrazines, pyrimidines and thiazoles under microwave conditions. *Synlett* **1998**, *47*, 661–663.
- (27) Gracias, V.; Gasielki, F. A.; Djuric, S. W. Synthesis of Fused Bicyclic Imidazoles by Sequential Van Leusen/Ring-Closing Metathesis Reactions. *Org. Lett.* **2005**, *7*, 3183–3186.
- (28) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. Reductive Amination of Aldehydes and Ketones with Sodium Triacetoxyborohydride. Studies on Direct and Indirect Reductive Amination Procedures. *J. Org. Chem.* **1996**, *61*, 3849–3862.
- (29) Rossi, D.; Marra, A.; Picconi, P.; Serra, M.; Catenacci, L.; Sorrenti, M.; Laurini, E.; Fermiglia, M.; Pricl, S.; Brambilla, S.; Almirante, N.; Peviani, M.; Curti, D.; Collina, S. Identification of RC-33 as a potent and selective  $\sigma_1$  receptor agonist potentiating NGF-induced neurite outgrowth in PC12 cells. Part 2: g-Scale synthesis, physicochemical characterization and in vitro metabolic stability. *Bioorg. Med. Chem.* **2013**, *21*, 2577–2586.

E

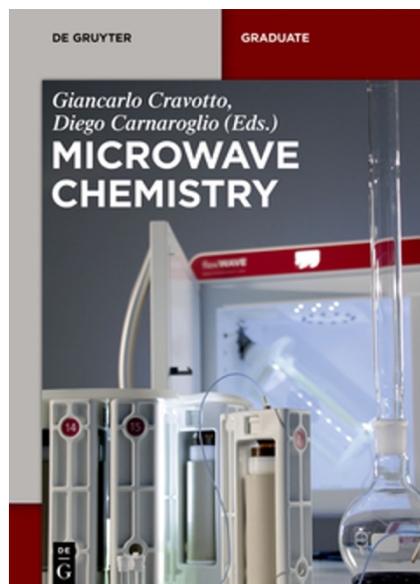
DOI: 10.1021/acsmedchemlett.8b00600  
ACS Med. Chem. Lett. XXXX, XXX, XXX–XXX

- (30) Shaaban, S.; Abdel-Wahab, B. F. Groebke–Blackburn–Bienaymé multicomponent reaction: emerging chemistry for drug discovery. *Mol. Diversity* **2016**, *20*, 233–254.
- (31) Sisko, J.; Mellinger, M.; Sheldrake, P. W.; Baine, N. H.  $\alpha$ -Tosylbenzyl isocyanide. *Organic Syntheses, Coll.* **2004**, *10*, 692 **2000**, 77, 198. .
- (32) Mayer, M.; Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784–1788.
- (33) Angulo, J.; Díaz, I.; Reina, J. J.; Tabarani, G.; Fieschi, F.; Rojo, J.; Nieto, P. M. Saturation transfer difference (STD) NMR spectroscopy characterization of dual binding mode of a mannose disaccharide to DC-SIGN. *ChemBioChem* **2008**, *9*, 2225–2227.
- (34) Monaco, S.; Tailford, L. E.; Juge, N.; Angulo, J. Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein–Ligand Contacts. *Angew. Chem., Int. Ed.* **2017**, *56*, 15289–15293.
- (35) Vasile, F.; Gubinelli, F.; Panigada, M.; Soprana, E.; Siccardi, A.; Potenza, D. NMR interaction studies of Neu5Ac- $\alpha$ -(2,6)-Gal- $\beta$ -(1–4)-GlcNAc with influenza-virus Hemagglutinin expressed in transfected human cells. *Glycobiology* **2018**, *28*, 42–49.
- (36) Dapiaggi, F.; Pieraccini, S.; Potenza, D.; Vasile, F.; Macut, H.; Pellegrino, S.; Aliverti, A.; Sironi, M. Computer aided design and NMR characterization of an oligopeptide targeting Ebola virus VP24 protein. *New J. Chem.* **2017**, *41*, 4308–4315.
- (37) Guzzetti, I.; Civera, M.; Vasile, F.; Arosio, D.; Tringali, C.; Piarulli, U.; Gennari, C.; Pignataro, L.; Belvisi, L.; Potenza, D. Insights into the binding of cyclic RGD peptidomimetics to  $\alpha 5 \beta 1$  integrin by live cell NMR and computational studies. *ChemistryOpen* **2017**, *6*, 128–136.
- (38) Vasile, F.; Menchi, G.; Lenci, E.; Guarna, A.; Potenza, D.; Trabocchi, A. Insight to the binding mode of triazole RGD-peptidomimetics to integrin-rich cancer cells by NMR and molecular modeling. *Bioorg. Med. Chem.* **2016**, *24*, 989–994.
- (39) *Maestro*; Schrödinger, LLC, New York, NY, 2018.
- (40) 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*, 5088–100.

*Publications related to microwave-assisted drug discovery*

*Book Chapter: 18. The use of microwaves in drug discovery*

†††††



---

††††† This content could not be reproduced either partially or entirely, due to copyright.

Collina, S.; Della Volpe, S. 18. The Use of Microwaves in Drug Discovery. In *Microwave Chemistry*; Cravotto, G., Carnaroglio, D., Eds.; De Gruyter: Berlin, Boston, 2017; pp 334–357.

<https://doi.org/10.1515/9783110479935-018>.

# Article 3: (R)-(-)-Aloesaponol III Methyl Ether from *Eremurus persicus*: A Novel Compound against Leishmaniosis



Article

## (R)-(-)-Aloesaponol III 8-Methyl Ether from *Eremurus persicus*: A Novel Compound against Leishmaniosis

Daniela Rossi <sup>1</sup>, Karzan Mahmood Ahmed <sup>1,2</sup>, Raffaella Gaggeri <sup>1,3</sup>, Serena Della Volpe <sup>1</sup>, Lauretta Maggi <sup>1</sup>, Giuseppe Mazzeo <sup>4</sup>, Giovanna Longhi <sup>4</sup>, Sergio Abbate <sup>4</sup>, Federica Corana <sup>5</sup>, Emanuela Martino <sup>6,\*</sup>, Marisa Machado <sup>7,8</sup>, Raquel Varandas <sup>9,10</sup>, Maria do Céu Sousa <sup>9,10</sup> and Simona Collina <sup>1,\*</sup>

<sup>1</sup> Department of Drug Sciences, Medicinal Chemistry and Pharmaceutical Technology Section, University of Pavia, Viale Taramelli 12, 27100 Pavia, Italy; daniela.rossi@unipv.it (D.R.); karzchem@yahoo.com (K.M.A.); raffaella.gaggeri@irst.emr.it (R.G.); serena.dellavolpe01@universitadipavia.it (S.D.V.); lauretta.maggi@unipv.it (L.M.)

<sup>2</sup> Department of Science-Chemistry, University of Garmian, Kalar 46021, Kurdistan Region, Iraq

<sup>3</sup> Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) Srl—IRCCS Via Piero Maroncelli, 40, 47014 Meldola (FC), Italy

<sup>4</sup> Dipartimento di Medicina Molecolare e Traslazionale, Università di Brescia, Viale Europa 11, 25123 Brescia, Italy; giuseppe.mazzeo@unibs.it (G.M.); giovanna.longhi@unibs.it (G.L.); smrmachado@gmail.com (S.A.)

<sup>5</sup> Centro Grandi Strumenti, University of Pavia, Via Bassi 21, 27100 Pavia, Italy; federica.corana@unipv.it

<sup>6</sup> Department of Earth and Environmental Sciences, University of Pavia, Via S. Epifanio 14, 27100 Pavia, Italy

<sup>7</sup> CESPU, Instituto de Investigação e Formação Avançada em Ciências e Tecnologias da Saúde, 4585-116 Gandra PRD, Portugal; smrmachado@gmail.com

<sup>8</sup> CIBIO-UP, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, InBIO, 4485-661 Vairão, Portugal

<sup>9</sup> Faculty of Pharmacy, University of Coimbra, Pólo das Ciências da Saúde, Azinhaga de Santa Comba, 3000-548 Coimbra, Portugal; raquel.varandas@ci.uc.pt (R.V.); mcsousa@ci.uc.pt (M.d.C.S.)

<sup>10</sup> CNC—Center for Neurosciences and Cell Biology, University of Coimbra, Rua Larga Faculty of Medicine, Pólo I, 3004-504 Coimbra, Portugal

\* Correspondences: emanuela.martino@unipv.it (E.M.); simona.collina@unipv.it (S.C.); Tel.: +39-0382-986-810 (E.M.); +39-0382-987-379 (S.C.); Fax: +39-0382-422-975 (E.M. & S.C.)

Academic Editors: Diego Muñoz-Torrero and Kelly Chibale

Received: 15 February 2017; Accepted: 20 March 2017; Published: 24 March 2017

**Abstract:** Leishmaniosis is a neglected tropical disease which affects several millions of people worldwide. The current drug therapies are expensive and often lack efficacy, mainly due to the development of parasite resistance. Hence, there is an urgent need for new drugs effective against *Leishmania* infections. As a part of our ongoing study on the phytochemical characterization and biological investigation of plants used in the traditional medicine of western and central Asia, in the present study, we focused on *Eremurus persicus* root extract in order to evaluate its potential in the treatment of leishmaniosis. As a result of our study, aloesaponol III 8-methyl ether (ASME) was isolated for the first time from *Eremurus persicus* root extract, its chemical structure elucidated by means of IR and NMR experiments and the (R) configuration assigned by optical activity measurements: chiroptical aspects were investigated with vibrational circular dichroism (VCD) and electronic circular dichroism (ECD) spectroscopies and DFT (density functional theory) quantum mechanical calculations. Concerning biological investigations, our results clearly proved that (R)-ASME inhibits *Leishmania infantum* promastigotes viability (IC<sub>50</sub> 73 µg/mL), inducing morphological alterations and mitochondrial potential deregulation. Moreover, it is not toxic on macrophages at the concentration tested, thus representing a promising molecule against *Leishmania* infections.

**Keywords:** leishmaniosis; drug identification; *Eremurus persicus*; plant extract; (R)-aloesaponol III-8 methyl ether

---

## 1. Introduction

Leishmaniosis, a neglected tropical disease (NTD), continues to be a major health problem, affecting 12 million people worldwide [1]. This disease is caused by the *Leishmania* species, and is generally classified into three different clinical forms: visceral leishmaniosis (VL), cutaneous leishmaniosis (CL) and mucocutaneous leishmaniosis (MCL). These forms differ in the pattern and clinical manifestations of the infection. VL can be fatal if left untreated, CL is localized and frequently self-heals within 3–18 months, while MCL leaves disfiguring scars. Although new drugs, such as paromycin, miltefosine and liposomal amphotericin B, are available as antileishmanial therapy in several countries, currently pentavalent antimonials and amphotericin B are the most used drugs for such purpose. Still, they induce toxic effects, produce development of parasite resistance and, no less important, they are expensive [2]. For these reasons, the need for new chemotherapeutic drugs is still strongly felt. Among different approaches proposed to fulfill this primary need, the investigation of medicinal plants with evidence of traditional use against *Leishmania* infections represents one of the most promising strategies. In fact, several plants and plant-derived natural products have been investigated so far as antileishmanial drug candidates [3,4].

Our nature-aided drug discovery research takes place as a response to this situation. As part of our ongoing study on the phytochemical characterization and biological investigation of plants used in western central Asia traditional medicine [5], we addressed our attention to plants belonging to genus *Eremurus* (Xanthorrhoeaceae), which includes forty species mainly found in western and central Asia. In the present work, we focused on *Eremurus persicus* (Jaub & Spach) Boiss. as source of potential antiprotozoal drugs. *Eremurus persicus* has a widespread usage in Kurdistan for both food purposes and to cure diseases related to inflammation [6] and infections; it has been demonstrated that aerial parts possess antibacterial and cytotoxic properties [7] and that root extracts show anti-inflammatory properties [8]. Nonetheless, so far, no scientific evaluation of the antileishmanial properties of *Eremurus persicus* had been carried out. Therefore, we herein report the preparation of *Eremurus persicus* roots ethanolic extract, the bioassay-guided fractionation of the extract in order to identify constituents with antileishmanial potential and the isolation, structural elucidation and biological investigation of its most abundant secondary metabolite.

## 2. Results and Discussion

Maceration using ethanol is a generally employed method in selective extraction of secondary metabolites from plant tissues, and we had already utilized such a procedure in our previous works on *Eremurus Persicus* [8]. Given the long timeframes (overnight maceration) and generally low and hardly reproducible yields (3.9%–7.6%) of maceration, on the basis of our previous experience, we studied the applicability of microwave-assisted solvent extraction (MASE) to the extraction of *E. Persicus* and compared it with dynamic maceration. Pilot extraction procedures with dynamic maceration and microwave-assisted extraction were investigated keeping the ratio sample weight/solvent volume constant with a value of about 0.05 g/mL.

The maceration procedure substantially followed the one described in our previous work [8], while for MASE, the extractions were effected in closed vessel system, under controlled pressure and temperature, to minimize the losses of components due to volatilization.

Through our study, we found that a rather short extraction time (20 min) afforded the ethanolic extract (EE) in higher yields (21.2%) compared to dynamic maceration. The EE obtained from both approaches showed a superimposable chromatographic profile at HPLC-UV performed according to our previous method [9].

Following the assessment of scale-up applicability, the root ethanolic extract of *E. Persicus* was then prepared according to the optimized MASE procedure (see Section 3.3).

To obtain a preliminary phytochemical fingerprint, the HPLC-UV/PAD method from [9] was exploited. The HPLC method was then slightly modified to afford faster analyses (see Section 3.4). The optimized method involves a switch in solvent from acetonitrile to methanol, which allowed to reduce chromatographic run times from 60 to 30 min without compromising the resulting information. The same method was consequently applied to both HPLC-UV/PAD and HPLC-ESI-MS analysis, with optimized MS experimental parameters for positive and negative ion modes. The HPLC-MS analysis unambiguously showed the presence of a preponderant secondary metabolite characterized by a molecular weight of 272 Da, given that the corresponding MS spectrum clearly displays an  $[M + H]^+$  at 273  $m/z$ ,  $[M + Na]^+$  at 295  $m/z$  and a dimeric ion  $[2M + Na]^+$  at 567  $m/z$  (Figure 1).

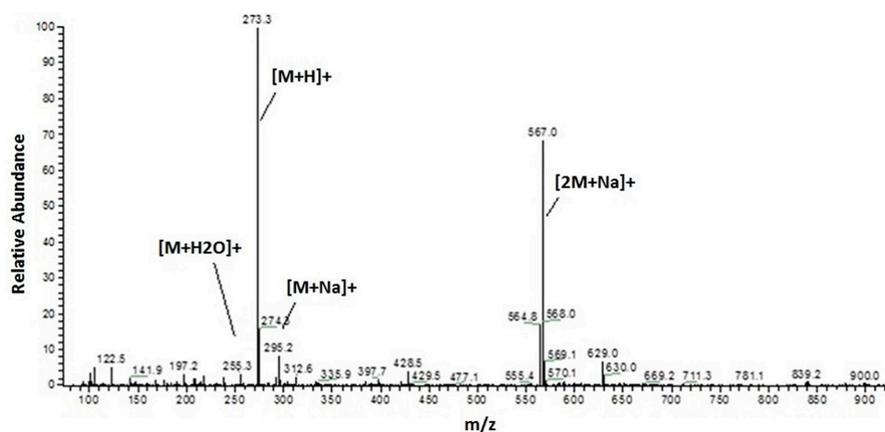


Figure 1. ESI-MS (positive ion mode, full scan) of the most abundant metabolite of EE.

With the aim to speed up the identification of new potential antiprotozoal compounds we adopted a bioassay-guided approach. The extract was fractionated by flash chromatography on silica gel (see Section 3.5) and analyzed applying the optimized HPLC method (see Section 3.4).

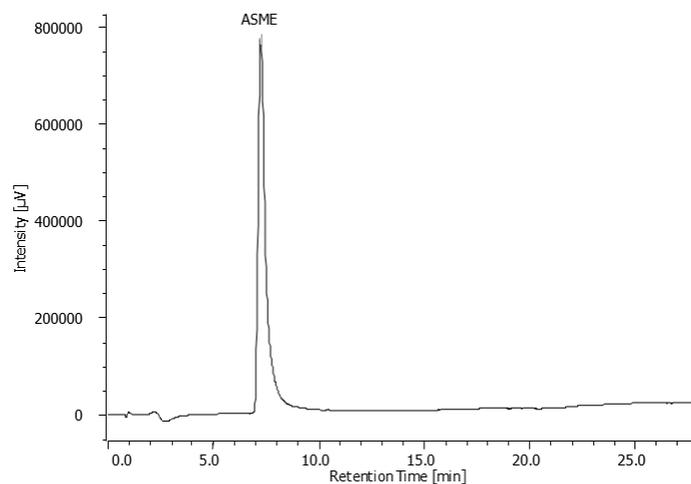
Flash chromatography allowed the isolation of six fractions which underwent a preliminary cytotoxicity screening. Only one fraction resulted active against *Leishmania infantum* in our initial screening whose structure matches, as shown by HPLC analysis, to the most abundant metabolite (Figure 2).

In order to obtain the compound of interest in the amount and chemical purity adequate for structural characterization as well as for in depth biological in vitro investigation, liquid–liquid extraction of the crude extract using water/dichloromethane (see Section 3.6) was then experimented. This procedure allowed for the isolation of the compound of interest with higher process yields in shorter timeframes.

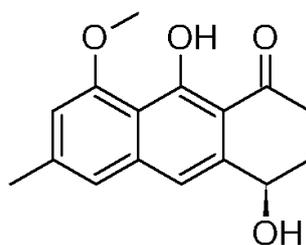
The structure of the isolated compound was elucidated by mono and bidimensional NMR ( $^1\text{H-NMR}$ ,  $^{13}\text{C-NMR}$ ,  $^1\text{H-}^1\text{H COSY}$ ,  $^1\text{H-}^{13}\text{C HSQC}$ ,  $^1\text{H-}^{13}\text{C HMBC}$ ), as well as IR spectroscopy and polarimetric analysis. Overall, spectroscopic data were consistent with those reported in literature for (–)-1-oxo-4,9-dihydroxy-8-methoxy-6-methyl-1,2,3,4-tetrahydroanthracene or aloesaponol III 8-methyl ether, from now on called (–)-ASME [10–12].

The absolute configuration of (–)-ASME (Figure 3) was unambiguously assigned through VCD (vibrational circular dichroism) and ECD (electronic circular dichroism) spectroscopies and DFT (density functional theory) calculations, through the approach reported in renowned literature [13–15]

and also already employed in instances relevant to the present case [16–19]. In Figure 4, we report the comparison of the spectra obtained with the computations performed for the assumed (*R*) configuration. The major features were correctly predicted by calculations conducted on the basis of the assumed configuration.



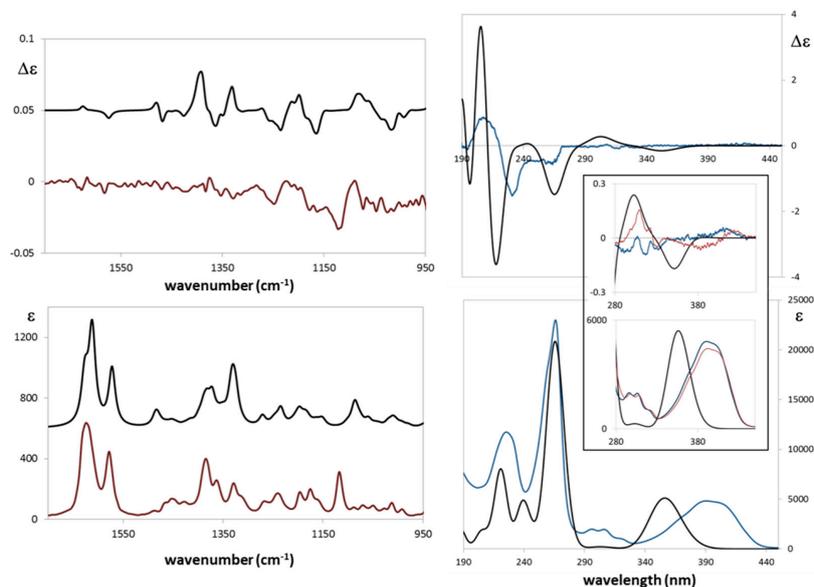
**Figure 2.** HPLC-UV chromatogram of the compound isolated through flash chromatography ( $\lambda$ : 270 nm).



**Figure 3.** Chemical structure of (*R*)-aloesaponol III-8 methyl ether ((*R*)-ASME).

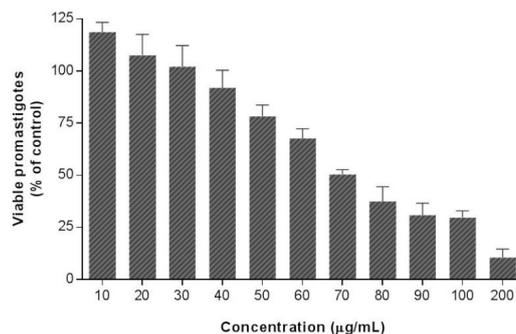
Even the weak features of the ECD spectra presented in the insets of Figure 4 are accounted for in the specific case of chloroform measurement: these weak features come from the subtle balance of the relative statistical weights of the ECD spectra associated to the various conformers (see Table S1 and Figure S9) and we verified experimentally that, switching to acetonitrile as solvent, a change in sign occurs as shown in Figure 4, inset. In passing, we note that, for the best prediction of most VCD and ECD data, it is crucial that the hydroxyl group on C4 be axial, as already noted in [16,17].

The minor  $n \rightarrow \pi^*$  feature at ca. 290 nm is crucially dependent on the population of conformers with the OH groups in the axial vs. equatorial position; in parallel, the  $^1L_b$  feature at 380 nm [20] depends on which fraction of the OCH<sub>3</sub> groups is either below or above the naphthalene plane (d instead of u), as defined in Figure S9; such conformations are overall approximately 58–56% vs. 42–44% in the first case or 51–50% vs. 49–50% in the second case and might reverse with solvent change (see Table 1). The specific OR might accordingly change and we do not consider them here, since the latter data are notoriously more difficult to theoretically predict [15].



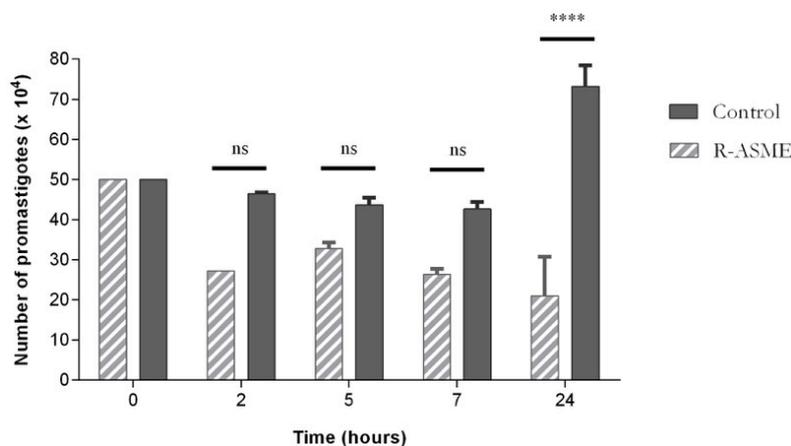
**Figure 4.** Left: Calculated (black trace) and experimental (red trace) VCD (top) and IR (bottom) spectra of **1**; Right: DFT Calculated (PCM/ACN, polarizable continuum model for acetonitrile-black trace) and experimental (blue trace) ECD (top) and UV (bottom) spectra of **1** in acetonitrile. In the inset, we repeat the magnified portions of spectra between 280 and 450 nm; in this region, we also report experimental ECD and UV data for chloroform solution (red trace). The calculations are in the PCM/CHCl<sub>3</sub> approximation (see text).

The antileishmanial potential of EE and (*R*)-ASME was finally examined through in vitro biological investigation, in particular their antiparasitic effects were evaluated on *L. infantum* promastigotes through MTT test. The results of our study clearly suggested that EE does not possess any effect on viability of *L. infantum* promastigotes, whereas (*R*)-ASME showed an interesting anti-*Leishmania* activity (Figure 5), its IC<sub>50</sub> value being equal to 73 µg/mL.



**Figure 5.** Effects of (*R*)-ASME on *L. infantum* promastigotes viability. Cultures of log-phase promastigotes ( $2 \times 10^6$  cells·mL<sup>-1</sup>) were incubated at 26 °C for 24 h at different drug concentrations. Values are expressed as means and SEM.

The treatment of *L. infantum* promastigotes with (R)-ASME at the IC<sub>50</sub> concentration induced a decrease of cell number, as clearly stated in Figure 6. This effect was particularly evident after 24 h of treatment, although evidence suggests that it occurred even earlier, i.e., after five and seven hours of incubation. The decrease of cell number was probably due to the inhibition of proliferation or/and lysis of promastigotes by (R)-ASME.



**Figure 6.** Effects of (R)-ASME on total cell number of *L. infantum* promastigotes, along the time of incubation. Each value represents the mean  $\pm$  SEM from three independent experiments (\*\*\*\*  $p < 0.0001$ , compared to control; ns, not significant).

To investigate the mechanisms of action responsible for the biological activity observed for (R)-ASME, its effects on (i) morphology; (ii) apoptosis/necrosis events; (iii) DNA integrity and (iv) cell cycle of *L. infantum* promastigotes were studied.

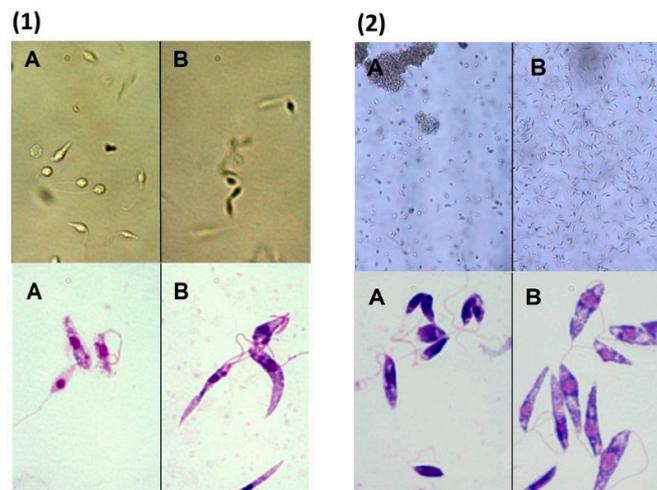
Concerning the effect of (R)-ASME on *L. infantum* promastigotes morphology, modifications were evaluated by optical microscopy using direct examination of live microorganism (hanging drop) and after Giemsa stain. The morphological observations pointed out a change in the shape and mobility of *L. infantum* cells treated with (R)-ASME at the IC<sub>50</sub> concentration (Figure 7). In details, control cells were very mobile with characteristic fusiform shape. After Giemsa staining, it was possible to observe the characteristic shape, long flagellum emerging from the anterior region of the parasite, nucleus and kinetoplast, posterior to the nucleus. In contrast, treated cells did not show the typical elongated shape since they became rounder and smaller, losing their mobility.

This evidence suggested that cytoskeletal disorganization and/or alterations on mitochondrial bioenergetics activity occur in treated cells.

**Table 1.** Effects of (R)-ASME (IC<sub>50</sub> concentration) on mitochondrial transmembrane potential of *L. infantum* promastigotes.

	<i>L. infantum</i> Intracellular Entities (% of Cells)											
	JC1Ag				JC1Mon				MIF <sup>a</sup> Agreg/MIFMon			
	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h
Control	98	88	91.6	93.7	2	11.2	8.2	6.2	4	7.4	2.7	2.2
(R)-ASME	97	81.4	88.2	90.5	2.9	18.5	11.4	9.5	5.6	6	2.2	1.5

<sup>a</sup> MIF: mean intensity fluorescence.



**Figure 7.** Optical microscopy observation of *L. infantum* promastigotes exposed to (*R*)-ASME (A); and to DMSO (control cells) (B); for 4 h (1); and 24 h (2). Hanging drop in phase contrast (magnification 100× and 200×) and Giemsa staining (magnification 1000×).

Alterations in the mitochondrial transmembrane potential were then studied by flow cytometry, using JC-1 as a probe. In healthy cells characterized by a high mitochondrial membrane potential, JC-1 tends to form aggregates; in impaired cells, with a low membrane potential, it is found as monomers. Cell treatment with (*R*)-ASME at  $IC_{50}$  concentration caused a loss of potential after four hours of incubation (Table 1), suggesting an early event of apoptosis mechanisms that could be associated with impaired cellular function.

To investigate the effect of (*R*)-ASME on apoptosis/necrosis events of *L. infantum* cells, the annexin-V/PI double staining method was employed. Annexin-V was used to mark early apoptotic cells, while PI, a non-permeable stain with affinity for nucleic acid, was used to distinguish necrotic cells. Our results revealed that apoptosis could be seen on treated cells after four hours of incubation and that necrosis mechanism did not occur at this time of incubation, being the percentage of cell positive to annexin-V and PI equal to 8.5 and 0.1, respectively (Table 2). After 24 h of treatment, neither annexin-V nor PI were positive, pointing out that the apoptosis/necrosis process was not maintained at this time of incubation.

**Table 2.** Flow cytometry analysis of *L. infantum* promastigotes treated with (*R*)-ASME showing the percentage of PI and annexin-V positive cells.

	<i>L. infantum</i> Intracellular Entities (% of Cells)											
	Annexin-V				PI				Annexin/PI			
	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h
Control	18.8	6.3	3.3	2.4	0.4	0.3	0	0.4	3.7	1.5	0.4	0.7
( <i>R</i> )-ASME	2.8	8.5	4.5	4.3	0.5	0.1	0.1	0.4	1.2	2	0.1	0.6

Since the degradation of nuclear DNA is one of the hallmarks of apoptotic cell death, the DNA integrity of *L. infantum* promastigotes treated with (*R*)-ASME at the  $IC_{50}$  concentration was then investigated by means of gel electrophoresis DNA analysis. Interestingly, similarly to control cells, the DNA of treated cells remained intact and devoid of any fragmentation (no smears, data not shown).

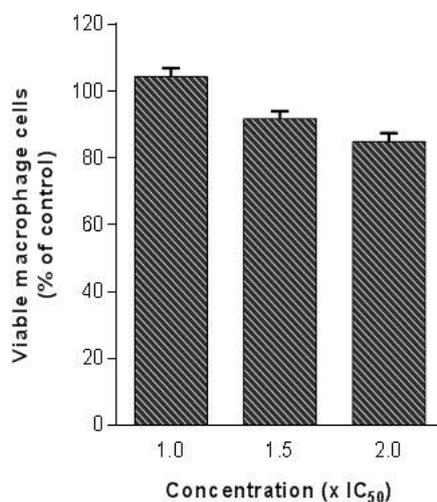
Finally, the effects of (*R*)-ASME ( $IC_{50}$  concentration) on the cell cycle of *L. infantum* promastigotes were investigated by flow cytometry. (Table 3). After six hours of incubation the treated parasite cells were on G0/G1 phase (88.6%) and S phase (11%) of the cell cycle. Comparatively, in non-treated cells the number of cells on G0/G1 and S phases was 71.4% and 29%, respectively. This means that, at this time of incubation, (*R*)-ASME promoted retention of *L. infantum* promastigotes in the G0/G1 phase of the cell cycle, thus suggesting an arrest of the cycle at this stage. Interestingly, this evidence seem to be correlated with the effects of (*R*)-ASME on cell number.

**Table 3.** Effects of (*R*)-ASME on the cellular cycle of *L. infantum* promastigotes.

	<i>L. infantum</i> Intracellular Entities (% of Cells)											
	Phase G0/G1				Phase S				Phase G2/M			
	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h	2 h	4 h	6 h	24 h
Control	83.3	87.1	71.4	47.7	16	13	29	36	0.3	0	0	16.4
( <i>R</i> )-ASME	83.9	87.1	88.6	44.9	15	13	11	38	1.1	0	0	17.6

Altogether, the results of our study focused on the mechanism of action of (*R*)-ASME suggest that initially (*R*)-ASME induced apoptosis events and arrest on G0/G1 phase cell cycle. Proliferation inhibition together with cell death were responsible for the rapid effects of (*R*)-ASME on cell number and morphological alterations (four to seven hours). After these events, the surviving cells are healthy; thus, at 24 h no differences in the apoptosis/necrosis parameters, DNA fragmentation and cell cycle were noticed.

Since the evaluation of cytotoxicity is a highly important issue in the drug discovery process, as the last step of our study we evaluated the toxicity of (*R*)-ASME in a macrophage cell line (RAW 264.7). Indeed, it is mandatory that a molecule with anti-*Leishmania* activity does not produce significant toxicity in the host cells. In detail, (*R*)-ASME was tested at high concentrations (until two folds the  $IC_{50}$  against *L. infantum*) on mammalian cells. Results showed that (*R*)-ASME seems to be safe for mammalian cells at tested concentrations, since only a slight decrease in macrophages viability was observed (Figure 8).



**Figure 8.** Effects of (*R*)-ASME on macrophage viability. Different concentrations ( $1.0 \times IC_{50}$ ,  $1.5 \times IC_{50}$  and  $2 \times IC_{50}$ ) were tested on macrophage cells (RAW 264.7) to evaluate cytotoxicity on mammalian cells.

### 3. Materials and Methods

#### 3.1. General

Solvents for extraction, purification (analytical grade), and high performance liquid chromatography analysis—HPLC-(HPLC grade) were supplied by Carlo Erba (Milan, Italy). Polyvinylpyrrolidone (PVP), gallic acid, deuterated solvent for NMR spectroscopy, RPMI-1640 Medium, Phosphate Buffered Saline (PBS), 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), dimethylsulfoxide (DMSO) were purchased from Sigma Aldrich (Milan, Italy). Fetal bovine serum (FBS; GibcoR) was purchased from ThermoFisher Scientific (Lisbon, Portugal). All solvents were evaporated under reduced pressure using a Heidolph Laborota 4000 instrument (Heidolph Instruments GmbH & Co., Schwabach, Germany).

Extractions were carried out a multimode Microwave apparatus using a closed-vessel system (MARSX press, CEM Corporation, Matthews, NC, USA).

Melting points were measured on an SMP3 Stuart Scientific apparatus and are uncorrected. Analytical thin-layer chromatography (TLC) was carried out on silica gel pre-coated glass-backed plates (Fluka Kieselgel 60 F254, Merck, Darmstadt, Germany) and visualized by UV light, acidic ammonium molybdate (IV), or potassium permanganate.

Flash chromatography was performed with silica gel 60 (particle size 230–400 mesh) purchased from Nova Chimica (Cinisello Balsamo, Italy).

<sup>1</sup>H-NMR, <sup>13</sup>C-NMR and bidimensional experiments were performed at 400.0 MHz and 100.6 MHz, respectively, on the Bruker Avance 400 MHz FT NMR spectrometer (Bruker, Leipzig, Germany) with a multinuclear BBO probe. CDCl<sub>3</sub> was used as solvent. <sup>1</sup>H chemical shift values were reported on the  $\delta$  scale in ppm, relative to TMS ( $\delta = 0.0$  ppm) and in <sup>13</sup>C-NMR, chemical shift values were reported posing CDCl<sub>3</sub> ( $\delta = 77.36$  ppm) as reference.

Optical rotation values were measured on a Jasco (Cremella, LC, Italy) photoelectric polarimeter DIP 1000 with a 1 dm cell at the sodium D line ( $\lambda = 589$  nm); sample concentration values ( $c$ ) are given in  $10^{-2}$  g·mL<sup>-1</sup>.

ECD spectra were run on a Jasco-815SE instrument (Jasco Corporation, Hachioji, Tokyo, Japan) using 0.1 to 1 mm cuvettes with chloroform and acetonitrile solutions of various concentrations (from 0.002 M to 0.004 M). VCD spectra were obtained on a Jasco FVS 6000: a 0.200 mm BaF<sub>2</sub> cell and 0.056 M/CDCl<sub>3</sub> solutions were employed. DFT calculations ancillary to chiroptical spectroscopies were run by use of Gaussian09 set of programs [21], preceded by Molecular Mechanics analysis of possible conformers of assumed configuration.

#### 3.2. Plant Material

*Eremurus persicus* (Jaub. & Spach) Boiss. was collected in a mountain area (Kūh-e Golestān, Golpayegan) located 120 km from Isfahan/Iran, at an altitude of 3000–3200 m. The collected plant materials were identified and classified by Dr. Abdulla Sa'ad at the Education Science Department, Faculty of Biology, Salahaddin University, Hawler/Iraq. The voucher specimen (No. 6856) was deposited at ESUH (Education Salahaddin University Herbarium), Hawler/Iraq. Freshly cut roots were dried in a drying room with active ventilation at room temperature (about 20–22 °C) until they showed no further weight loss. The roots were cut to small size and grounded with a blade mill (A10 IKA-Werke GmbH & Co. Staufen, Germany) to obtain a homogeneous fine powder. The plant material so treated was stored in dark conditions.

#### 3.3. Extraction Procedure

In our previous works [8,9], the ethanolic extract (EE) of *Eremurus persicus* was prepared through a dynamic maceration procedure (extraction yield 7.6%).

In the present work, the extraction process was optimized through the application of a MASE approach which gained higher yields in shorter timeframes. Briefly, the dried roots of *E. persicus* (25 g)

were extracted by using ethanol (500 mL) in a multimode microwave apparatus at 120 °C for 20 min, with a power of 800 W. The extract was separated by Buchner filtration and solvent was evaporated to dryness under vacuum, yielding a yellow oil (5.2 g, extraction yield 20.8%).

#### 3.4. High Performance Liquid Chromatography Analyses

High performance liquid chromatography-photodiode array (HPLC-UV/PAD) analyses were performed on a Jasco system (Cremella, LC, Italy) equipped with a Jasco AS-2055 plus autosampler, a PU-2089 plus pump and a MD-2010 plus multi-wavelength detector. Experimental data were acquired and processed by Jasco Borwin PDA and Borwin Chromatograph Software.

High performance liquid chromatography-electrospray-tandem mass spectrometry (HPLC-ESI-MS) analyses were carried out on Finnigan LCQ fleet ion trap system, controlled by Xcalibur software 1.4 (ThermoFinnigan, San Jose, CA, USA). Mass spectra were generated both in positive and negative ion mode under constant instrumental conditions. For positive ion mode: ion spray voltage 5 kV, capillary voltage 46 V, capillary temperature 220 °C, and tube lens voltage 120 V. For negative ion mode: ion spray voltage 5 kV, capillary voltage −35 V, capillary temperature 220 °C, and tube lens voltage −100 V.

Reverse phase chromatographic analyses were carried out at room temperature (RT) under gradient conditions, using a Chromolith SpeedROD RP-18 endcapped column (50 mm × 4.6 mm, ID 3 mm, macropore size 2 µm, mesopore size 13 nm, Merck, Darmstadt, Germany).

The HPLC analysis conditions were properly optimized to monitor the extract fractionation. The mobile phase was water containing 0.1% formic acid (A) and methanol containing 0.1% formic acid (B), and the composition gradient was: from 30% to 39% of B in 7 min, 42% B until 4 min, 50% B until 5 min followed by an isocratic elution for 2 min, 70% B until 5 min 30% B until 5 min, followed by a re-equilibration step of 2 min; total run time 30 min.

For all analyses the flow rate was set at 1 mL/min. Samples were dissolved in methanol (3 mg/mL) and filtered with a 0.45 µm GH Polypro (GHP) membrane before injection into the HPLC-system.

#### 3.5. Bio-Guided Fractionation

The EE (1 g) was subjected to Flash Chromatography on Silica gel under gradient conditions eluting with 50% ethyl acetate and 50% hexane to 70% and then 100% ethyl acetate. Six fractions were collected on the basis of the TLC profile, analyzed according to described HPLC method and evaluated with a preliminary biological assay (MTT test, Section 3.8.2). Fraction 2 (19 mg, yield 1.9%), characterized by the presence of a single peak (HPLC: RT = 6.94 min, purity > 99%) corresponding to the main peak of EE chromatogram, resulted the only effective.

#### 3.6. Isolation of the Main Secondary Metabolite

The EE (3 g) was dissolved in water (900 mL) and extracted with dichloromethane (DCM, 900 mL) under mechanical stirring for 3 h at RT. The organic fraction was then collected and the extraction procedure repeated for three times. The combined organic phases, dried over sodium sulphate and evaporated in vacuo furnished a yellow solid (210 mg, yield 7%, HPLC: RT 6.94 min, purity = 75%, TLC  $R_f$  0.87 EtAc/MeOH/H<sub>2</sub>O). Further purification of the solid via crystallization (acetone) allowed the isolation of 135 mg of a pure compound 210 mg, m.p. 174–175 °C, yield 4.5%, HPLC: RT 6.93 min, purity > 99%). Nuclear Magnetic Resonance (<sup>1</sup>H- and <sup>13</sup>C-NMR) and Mass Spectral (MS) techniques were employed for the structure elucidation of the isolated compound. The structure was determined as (−)-1-oxo-4,9-dihydroxy-8-methoxy-6-methyl-1,2,3,4-tetrahydroanthracene or aloesaponol III 8-methyl ether by comparison of the spectroscopic data with the previous reports.

$[\alpha]_D^{20}$  −37.9 (c 0.04, CHCl<sub>3</sub>),  $[\alpha]_D^{20}$  −18.2 (c 0.2, Acetone),  $[\alpha]_D^{20}$  −18.5 (c 0.3, CH<sub>3</sub>OH); ESI-MS:  $m/z$  273 [M + H]<sup>+</sup>, 295 [M + Na]<sup>+</sup>, 271 [M − H]<sup>−</sup>.

$^1\text{H-NMR}$  ( $\text{CDCl}_3$ , 300 MHz):  $\delta$  2.01 (1H, brs, OH), 2.28 (2H, m, H-3), 2.48 (3H, s, H-11), 2.69 (1H, m, H-2), 3.05 (1H, m, H-2), 4.02 (3H, s, OCH<sub>3</sub>), 4.94 (1H, dd,  $J = 5$  Hz, H-4), 6.69 (1H, d,  $J = 2$  Hz, H-7 or H-5), 7.06 (1H, s, H-10), 7.14 (1H, d,  $J = 2$  Hz, H-7 or H-5).

$^{13}\text{C-NMR}$  ( $\text{CDCl}_3$ ):  $\delta$  203.3 (C-1, s), 165.9 (C-9, s), 159.5 (C-8, s), 142.0 (C-6, s), 140.3 (C-5a, s), 139.9 (C-4a, s), 119.9 (C-5, d), 115.4 (C-10, d), 113.8 (C-8a, s), 109.2 (C-1a, s), 108.3 (C-7, d), 68.0 (C-4, d), 56.0 (OCH<sub>3</sub>, q), 34.0 (C-2, t), 30.7 (C-3, t), 22.1 (C-11, q).

### 3.7. ECD, VCD Spectra and DFT Calculations

ECD spectra were run using 0.1 to 1 mm cuvettes with chloroform and acetonitrile solutions at different concentrations (from 0.002 M to 0.004 M) to accommodate the various spectroscopic regions. Twenty scans were run to obtain data in the 500–180 nm range. The corresponding ECD spectra for the solvent in the same conditions were run and subtracted from those of the solution. VCD spectra were obtained employing a 0.200 mm BaF<sub>2</sub> cell and 0.056 M/ $\text{CDCl}_3$  solutions. Five thousand scans were taken and solvent subtraction was made with VCD spectra in the same condition. DFT calculations ancillary to chiroptical spectroscopies were run by use of Gaussian09 set of programs [21], preceded by Molecular Mechanics analysis of possible conformers of the assumed configuration (*R*): subsequent quantum mechanical DFT calculations were conducted at B3LYP/TZVP level of theory in the PCM-IEF modelization of solvent [22]. The computed VCD and IR spectra were obtained from calculated rotational strengths and dipole strengths and wavenumbers by assigning Lorentzian bandshapes with  $16\text{ cm}^{-1}$  bandwidth; spectra were scaled by 0.98 in frequency and divided by 3 in intensity to facilitate comparison with experimental spectra. The UV and ECD spectra were generated assigning Gaussian bandshapes with 0.2 bandwidth. Computed spectra were shifted by 20 nm and intensities were divided by 3.

### 3.8. Biological Evaluation

#### 3.8.1. Parasites and Cultures

Promastigote forms of *Leishmania infantum* Nicolle (zymodeme MON-1) were maintained at 26 °C with weekly transfers in HEPES (25 mM)-buffered RPMI 1640 medium enriched with 10% inactivated fetal bovine serum (FBS).

#### 3.8.2. Viability Assay

The antiparasitic effect of EE and (*R*)-ASME was studied through MTT test, a colorimetric assay of cell survival, following the method of [23]. Log phase *L. infantum* promastigotes were incubated with increasing concentration of EE and (*R*)-ASME, in fresh medium for 2 h. After the incubation, 25  $\mu\text{L}$  of MTT (5  $\text{mg}\cdot\text{mL}^{-1}$ ) was added to each well, incubated for 2 h at 37 °C and centrifuged at 3000 rpm for 5 min. The supernatant was removed, the cells were washed in PBS, and the precipitated formazan was dissolved in DMSO (250  $\mu\text{L}$ ). Cell viability was measured by absorbance at 530 nm on an ELISA plate reader (Synergy HT, Bio-TEK, Winooski, VT, USA), and calculated using the following formula:  $[(L2/L1) \times 100]$ , where L1 is the absorbance of control cells and L2 is the absorbance of treated cells. Three separate experiments were performed for each sample and the concentration that inhibited viability by 50% ( $\text{IC}_{50}$ ) was determined through dose-response regression analysis, plotted by GraphPad Prism 6 (GraphPad Software, Inc., La Jolla, CA, USA).

#### 3.8.3. Morphological Studies and Cell Counting

Parasites cells were exposed to (*R*)-ASME and morphological alterations were investigated by optical microscopy using direct examination of live microorganisms (hanging drop) and after Giemsa stain. Briefly, exponentially grown of *L. infantum* ( $2 \times 10^6$  cells  $\text{mL}^{-1}$ ) were treated with the compound at  $\text{IC}_{50}$  concentrations for 2 h, 4 h, 6 h and 24 h at 26 °C. After incubation, cells were pelleted by centrifugation at 3000 rpm for 5 min and the supernatant was discarded by aspiration. The cell pellet

was suspended in fresh medium and approximately 10  $\mu\text{L}$  were placed on a Koch slide and directly observed under the optical microscope phase contrast (Eclipse E400, Nikon coupled with a digital camera 165 DN100 Nikon, Nikon Europe B.V. Amsterdam, the Netherlands). In addition, a total cell counting was performed using a haemocytometer and a smear was made which was submitted to Giemsa stain. The smear was fixed with methanol for 5 min, stained with aqueous solution of Giemsa (1/10, *v/v*) for 10 min at room temperature and finally washed with water and air dried. The stained smear was observed under the microscope with a 100 $\times$  lens (Eclipse E400, coupled with Nikon digital camera, Nikon DN100 165).

#### 3.8.4. Cell Cycle Analysis

For the analysis of DNA content, exponentially grown *L. infantum* ( $2 \times 10^6$  cells  $\text{mL}^{-1}$ ) were treated with (R)-ASME at  $\text{IC}_{50}$  concentration for 2 h, 4 h, 6 h and 24 h at 26  $^{\circ}\text{C}$ . At each time point, cells were fixed in 200  $\mu\text{L}$  of 70% ethanol for 30 min. at 4  $^{\circ}\text{C}$ . After washing cells with 2 mL of PBS, enriched with 2% of bovine serum albumin (BSA), the pellets were suspended in 0.5 mL of propidium iodide (PI) solution (PI/Rnase, Immunostep, Salamanca, Spain) and incubated for 15 min at 37  $^{\circ}\text{C}$  [24]. Cells were then analyzed by flow cytometry (D Biosciences, San Jose, CA, USA). Results were treated using ModFit LT V 2.0 programme (D Biosciences).

#### 3.8.5. Phosphatidylserine Externalization

Double staining for annexin V-FITC and PI was performed as described previously [25]. Briefly, *L. infantum* promastigotes ( $2 \times 10^6$  cells  $\text{mL}^{-1}$ ) were exposed to (R)-ASME at  $\text{IC}_{50}$  concentrations for 2 h, 4 h, 6 h, and 24 h at 26  $^{\circ}\text{C}$ . Cells were then washed with PBS and re-suspended in binding buffer (10 mM HEPES-NaOH, pH 7.4, 140 nM NaCl, 2.5 mM  $\text{CaCl}_2$ ). To 100  $\mu\text{L}$  of this suspension were added 5  $\mu\text{L}$  of annexin V-FITC and 5  $\mu\text{L}$  of PI (AnnexinV-FITC Apoptosis Detection Kit, Immunostep). After 15 min incubation in the dark at room temperature, 400  $\mu\text{L}$  of binding buffer were further added and cells were analyzed by flow cytometry (FacsCalibur-Beckton-Dickinson). Data analysis was carried out using the program Paint-a-gate, and values are expressed as a percentage of positive cells for a given marker, relatively to the number of cells analyzed.

#### 3.8.6. Measurement of Mitochondrial Membrane Potential

To assess mitochondrial membrane potential ( $\Delta\psi_m$ ), a cell-permeable cationic and lipophilic dye, JC-1 (5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolcarbocyanine iodide), was used as previously described [26]. This probe aggregates within mitochondria and fluoresces red (590 nm) at higher  $\Delta\psi_m$ . However, at lower  $\Delta\psi_m$ , JC-1 cannot accumulate within the mitochondria and instead remains in the cytosol as monomers, which fluoresces green (490 nm). Therefore, the ratio of red to green fluorescence gives a measure of the transmembrane electrochemical gradient. *L. infantum* promastigotes ( $10^6$  cells) were exposed to (R)-ASME at the  $\text{IC}_{50}$  concentrations for 24 h at 26  $^{\circ}\text{C}$ . Promastigotes were then incubated with JC-1 (5  $\mu\text{g}\cdot\text{mL}^{-1}$ ) (Molecular Probes, Invitrogen) in the dark for 15 min at RT. Then, cells were washed in PBS, suspended in 400  $\mu\text{L}$  of PBS and analyzed by flow cytometry. Data analysis was carried out using the program Paint-a-gate.

#### 3.8.7. DNA Fragmentation Assay

Promastigotes of *L. infantum* ( $2 \times 10^6$  cells  $\text{mL}^{-1}$ ) were exposed to (R)-ASME at  $\text{IC}_{50}$  concentration or to dissolution vehicle (DMSO), and incubated at 26  $^{\circ}\text{C}$  for 24 h. The *Leishmania* DNA extraction was carried out according to the procedure in DNeasy Blood & Tissue (Qiagen, Hilden, Germany). DNA integrity analysis was done by electrophoresis, running DNA through an EtBr-treated agarose gel and visualizing it with UV light.

### 3.8.8. Mammalian Cell Cytotoxicity

For cytotoxicity assays on mammalian cells, log phase of macrophages (RAW 264.7) were trypsinized and incubated at 37 °C in 24-well tissue culture plates in Dulbecco's Modified Eagle Medium (DMEM), enriched with Glutamax and supplemented with 10% FBS, under microaerophilic conditions. As soon as the monolayers reached confluence, the medium was removed and cells were incubated at 37 °C for 24 h with fresh medium and (R)-ASME at IC<sub>50</sub> concentrations. After incubation, control and treated cells were washed with PBS, pH 7.2, and 450 µL of PBS and 50 µL MTT solution (5 mg·mL<sup>-1</sup>) were added to each well and incubated at 37 °C for 1 h. The cells were then washed with PBS, 500 µL of DMSO were added to the wells and absorbance was measured at 530 nm on an ELISA plate reader (Synergy HT, Bio-TEK). The percentage of viable cells was determined as described in the viability assay.

### 3.9. Statistical Analysis

The results are expressed as mean ± SEM from at least three independent experiments; they were analyzed by one-way analysis of variance (ANOVA), followed by Dunnett's test, using GraphPad Prism, version 6.0d (GraphPad Software, San Diego, CA, USA). The level of significance was \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$  when compared to control.

## 4. Conclusions

In this work, we used a bioassay-guided fractionation as a strategic approach to natural source product discovery and isolation. Liquid–liquid extraction of the ethanolic extract (EE) of *Eremurus persicus* roots combined with biological evaluation led to the identification of an active fraction with antileishmanial activity corresponding to the (R) enantiomer of Aloesaponol III 8-methyl ether ((R)-ASME). Interestingly, this is the first time that the (R)-ASME has been isolated from natural sources, while the (S) enantiomer had already been reported as found in *Kniphofia foliosa* [11], *Eremurus chinensis* [12] and *Asphodelus microcarpus* [27].

(R)-ASME possesses a remarkable antiprotozoal effect against *L. infantum*. We demonstrated that it is active with an IC<sub>50</sub> of 73 µg/mL, it is able to induce alterations on both morphology and mitochondrial potential of *L. infantum* promastigotes, and does not produce significant toxicity in a macrophage cell line. Results showed that (R)-ASME seems to be safe for mammalian cells at tested concentrations, since only a slight decrease in macrophages viability was observed. These results strongly suggest that (R)-ASME may represent a valuable lead against *Leishmania* infections. Although our preliminary results are encouraging, additional studies should be planned both in vitro and in vivo to get further insight into the mechanisms of action behind the antileishmanial activity of (R)-ASME.

**Supplementary Materials:** Supplementary materials are available online.

**Acknowledgments:** We gratefully acknowledge Giovanni Vidari from University of Pavia, Italy, for his helpful suggestions and Sara Ferrari for the experimental support in the biological investigation of EE and (R)-ASME. We thank Regione Lombardia (Astil-PROKURDUP Project) at University of Pavia as well as CARIPLO foundation and AGROFOOD LAB at University of Brescia for financial support.

**Author Contributions:** S.C. and E.M. conceived the work, contributed in reviewing the manuscript and were also responsible for the correctness of the whole study. D.R. wrote the manuscript. D.R. and R.G., were responsible for the design of experiments and for data analysis of the whole study. K.M.A., S.D.V., L.M. and F.C. performed the experiments (preparation and analysis of the extracts). G.M, S.A and G.L. were responsible for the absolute configuration assignment and carried out chiroptical spectroscopies characterization and DFT calculations. M.M., R.V. and M.d.C.S. were responsible for biological investigation and wrote the biological section.

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

## References

1. Alviano, D.S.; Barreto, A.L.S.; de Almeida Dias, F.; de Almeida Rodrigues, I.; dos Santos Rosa, M.D.S.; Alviano, C.S.; de Araújo Soares, R.M. Conventional therapy and promising plant-derived compounds against trypanosomatid parasites. *Front. Microbiol.* **2012**, *3*, 283. [[CrossRef](#)] [[PubMed](#)]
2. Singh, N.; Kumar, M.; Singh, R.K. Leishmaniasis: Current status of available drugs and new potential drug targets. *Asian Pac. J. Trop. Med.* **2012**, *5*, 485–489. [[CrossRef](#)]
3. Chouhan, G.; Islamuddin, M.; Sahal, D.; Afrin, F. Exploring the role of medicinal plant-based immunomodulators for effective therapy of leishmaniasis. *Front. Immunol.* **2014**, *5*, 193. [[CrossRef](#)] [[PubMed](#)]
4. Passero, L.F.D.; Laurenti, M.D.; Santos-Gomes, G.; Campos, B.L.S.; Sartorelli, P.; Lago, J.H.G. Plants Used in Traditional Medicine: Extracts and Secondary Metabolites Exhibiting Antileishmanial Activity. *Curr. Clin. Pharmacol.* **2014**, *9*, 187–204. [[CrossRef](#)] [[PubMed](#)]
5. Gaggeri, R.; Rossi, D.; Christodoulou, M.S.; Passarella, D.; Leoni, F.; Azzolina, O.; Collina, S. Chiral Flavanones from *Amygdalus lycioides* Spach: Structural Elucidation and Identification of TNF alpha Inhibitors by Bioactivity-guided Fractionation. *Molecules* **2012**, *17*, 1665–1674. [[CrossRef](#)] [[PubMed](#)]
6. Mati, E.; De Bore, H. Ethnobotany and trade of medicinal plants in the qeysare market, Erbil, Kurdish autonomy region Iraq. *J. Ethnopharmacol.* **2011**, *133*, 490–510. [[CrossRef](#)] [[PubMed](#)]
7. Vala, M.H.; Asgarpanah, J.; Hedayati, M.H.; Shirali, J.; Bejestani, F.B. Antibacterial and cytotoxic activity of *Eremurus persicus* (Jaub and Spach) Boiss. *Afr. J. Microbiol. Res.* **2011**, *16*, 2349–2352.
8. Gaggeri, R.; Rossi, D.; Azzolina, O.; Leoni, F.; Ahmed, K.M.; Avanzini, A.; Mantelli, M.; Gozzini, D.; Paolillo, M.; Collina, S. Anti-inflammatory properties of ethanolic root extract of *Eremurus persicus* (Jaub and Spach) Boiss, a Kurdish herbal remedy. *J. Med. Plants Res.* **2013**, *7*, 1730.
9. Gaggeri, R.; Rossi, D.; Mahmood, K.; Gozzini, D.; Mannucci, B.; Corana, F.; Daglia, M.; Avanzini, A.; Mantelli, M.; Martino, E.; et al. Towards elucidating *Eremurus* root remedy: Chemical profiling and preliminary biological investigations of *Eremurus persicus* and *Eremurus spectabilis* root ethanolic extracts. *J. Med. Plants Res.* **2015**, *8*, 1038–1048.
10. Yagi, A.; Makino, K.; Nishioka, I. Studies on constituents of *Aloe saponaria* Haw. II. The structures of tetrahydroanthracene derivatives, aloesaponol III and IV. *Chem. Pharm. Bull.* **1977**, *25*, 1764–1770. [[CrossRef](#)]
11. Yenesew, D.E.; Muller, M.; Steglich, W. An Anthraquinone and two oxanthrones from *Kniphonia foliosa*. *Phytochemistry* **1994**, *37*, 525–528. [[CrossRef](#)]
12. Li, C.; Shi, J.G.; Zhang, Y.P.; Zhang, C.Z. Constituents of *Eremurus chinensis*. *J. Nat. Prod.* **2000**, *63*, 653–656. [[CrossRef](#)] [[PubMed](#)]
13. Abbate, S.; Burgi, L.F.; Castiglioni, E.; Lebon, F.; Longhi, G.; Toscano, E.; Caccamese, S. Assessment of configurational and conformational properties of Naringenin by Vibrational Circular Dichroism. *Chirality* **2009**, *21*, 436–441. [[CrossRef](#)] [[PubMed](#)]
14. Nafie, L.A. *Vibrational Optical Activity, Principles and Applications*; John Wiley & Sons: New York, NY, USA, 2011.
15. Polavarapu, P.L. *Chiroptical Spectroscopy. Fundamentals and Applications*; CRC Press, Taylor & Francis Group, LLC: Oxford, UK, 2017.
16. Abbate, S.; Lebon, F.; Longhi, G.; Morelli, C.F.; Ubiali, D.; Speranza, G. Vibrational and electronic circular dichroism spectroscopies and DFT calculations for the assignment of the absolute configuration of hydroxy-substituted 2-tetralols. *RSC Adv.* **2012**, *2*, 10200–10208. [[CrossRef](#)]
17. Mazzeo, G.; Cimmino, A.; Andolfi, A.; Evidente, A.; Superchi, S. Computational ECD Spectrum Simulation of the Phytotoxin Scytalone: Importance of Solvent Effects on Conformer Populations. *Chirality* **2014**, *26*, 502–508. [[CrossRef](#)] [[PubMed](#)]
18. Rossi, D.; Nasti, R.; Marra, A.; Meneghini, S.; Mazzeo, G.; Longhi, G.; Memo, M.; Cosimelli, B.; Greco, G.; Novellino, E.; et al. Enantiomeric 4-Acylamino-6-alkoxy-2-Alkylthiopyrimidines as Potential A3 Adenosine Receptor Antagonists: HPLC Chiral Resolution and Absolute Configuration Assignment by a Full Set of Chiroptical Spectroscopy. *Chirality* **2016**, *28*, 434–440. [[CrossRef](#)] [[PubMed](#)]
19. Rossi, D.; Nasti, R.; Collina, S.; Mazzeo, G.; Ghidinelli, S.; Longhi, G.; Abbate, S. The Role of Chirality in a Set of Key Intermediates of Pharmaceutical Interest, 3-aryl-substituted- $\gamma$ -butyrolactones, evidenced by Chiral HPLC Separation and by Chiroptical Spectroscopies. *J. Pharm. Biomed. Appl.* **2017**. [[CrossRef](#)] [[PubMed](#)]

20. Kurtan, T.; Antus, S.; Pescitelli, G. Electronic CD of benzene and other aromatic chromophores for determination of absolute configuration. In *Comprehensive Chiroptical Spectroscopy*; Berova, N., Polavarapu, P.L., Nakanishi, K., Woody, R.W., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2012; Volume 2, pp. 101–104.
21. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian. Gaussian 09, Revision A.02*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
22. Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3094. [[CrossRef](#)] [[PubMed](#)]
23. Denizot, F.; Lang, R. Rapid colorimetric assay for cell growth and survival. Modifications to the tetrazolium dye procedure giving improved sensitivity and reliability. *J. Immunol. Methods* **1986**, *89*, 271–277. [[CrossRef](#)]
24. Darzynkiewicz, Z.; Juan, G.; Bedner, E. Chapter 8: Unit 8.4: Determining cell cycle stages by flow cytometry. In *Current Protocols in Cell Biology*; New York Press: New York, NY, USA, 2001.
25. Vermes, I.; Haanen, C.; Steffens-Nakken, H.; Reutelingsperger, C.A. A novel assay for apoptosis Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. *J. Immunol. Methods* **1995**, *184*, 39–51. [[CrossRef](#)]
26. Cossarizza, A.; Baccarani-Contri, M.; Kalashnikova, G.; Franceschi, C. A new method for the cytofluorimetric analysis of mitochondrial membrane potential using the J-aggregate forming lipophilic cation 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolcarbocyanine iodide (JC-1). *Biochem. Biophys. Res. Commun.* **1993**, *30*, 40–45. [[CrossRef](#)]
27. Ghoneim, M.M.; Ma, G.; El-Hela, A.A.; Mohammad, A.E.I.; Kottob, S.; El-Ghaly, S.; Cutler, S.J.; Ross, S.A. Biologically active secondary metabolites from *Aphodelus microcarpus*. *Nat. Prod. Commun.* **2013**, *8*, 1117–1119. [[PubMed](#)]

**Sample Availability:** Samples of the (R)-ASME available from the authors.



© 2017 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 (<http://creativecommons.org/licenses/by/4.0/>).

**Article 4: The use of a microwave assisted solvent extraction coupled with HPLC UV/PAD to assess the quality of *Marrubium Vulgare L.* (white horehound) herbal raw material**

Reprinted with permission from (*Phytochemical Analysis*. 2019; 30: 377– 384.; <https://doi.org/10.1002/pca.2820>). Copyright (2019) John Wiley & Sons, Ltd

Received: 7 September 2018 | Revised: 9 November 2018 | Accepted: 19 December 2018

DOI: 10.1002/pca.2820

RESEARCH ARTICLE

WILEY *Phytochemical Analysis*

## The use of a microwave-assisted solvent extraction coupled with HPLC-UV/PAD to assess the quality of *Marrubium vulgare L.* (white horehound) herbal raw material

Emanuela Martino<sup>1</sup> | Serena Della Volpe<sup>2</sup> | Valeria Cavalloro<sup>1</sup> | Bedis Amri<sup>3</sup> |  
Leila Bettaieb Been Kaab<sup>3</sup> | Giorgio Marrubini<sup>2</sup> | Daniela Rossi<sup>2</sup> | Simona Collina<sup>2</sup> 

<sup>1</sup>Department of Earth and Environmental Sciences, University of Pavia, Pavia, Italy

<sup>2</sup>Department of Drug Sciences, University of Pavia, Pavia, Italy

<sup>3</sup>Département de Biologie, Faculté des Sciences de Tunis el Manar, Unité de recherche "Nutrition et métabolismes azoté et protéines de stress", Tunis, Tunisia

Correspondence

Simona Collina, Department of Drug Sciences, University of Pavia, Viale Taramelli 12, 27100 Pavia, Italy.

Email: [simona.collina@unipv.it](mailto:simona.collina@unipv.it)

### Abstract

**Introduction:** *Marrubium vulgare* is a herbal remedy presents in several European Pharmacopoeias and commonly marketed as white horehound. The chemotaxonomic marker of *Marrubium* genus is marrubiin and its content may change in response to biotic and abiotic stress.

**Objective:** Development of a microwave-assisted solvent extraction (MASE) methodology suitable for exhaustively extracting marrubiin from *M. vulgare* leaves, easily applicable to large sets of samples. Evaluation of the influence of copper(II) on marrubiin production.

**Material and methods:** *M. vulgare* leaves were dried, extracted exploiting MASE and analysed via high-performance liquid chromatography ultraviolet photodiode array detection (HPLC-UV/PAD) system. A design of experiments approach was adopted to select the best extraction conditions. Extraction parameters (solvent composition, extraction time and temperature), were studied applying two full factorial experimental designs in a sequential approach. To analyse samples, a rapid HPLC-UV/PAD method was set up.

**Results:** The best results in terms of marrubiin extraction yield were obtained extracting samples at 120°C with 100% ethanol, for 15 min (3 × 5 min microwave cycles). The developed methodology was successfully applied to matrices grown in Greenhouse conditions and under stress induced by copper(II), selected as model agent for abiotic stress. Progressively decreasing production of marrubiin was evidenced in connection with treatment with 80, 200 and 300 mg/L copper sulphate.

**Conclusion:** An efficient methodology for the extraction and determination of the amount of marrubiin in large sets of samples of *M. vulgare* plants was developed. Results demonstrated that marrubiin is an easily detectable marker useful for evaluating *M. vulgare* reaction to stress.

### KEYWORDS

copper(II)-induced stress, DoE, HPLC-UV/PAD analysis, *Marrubium vulgare L.* leave extract, MASE

## 1 | INTRODUCTION

Nature contributes greatly to the history and landscape of drugs. Around half of the drugs currently in clinical use are of natural origin, about 30% of drugs approved by the FDA (US Food and Drug Administration) over the past three decades are nature-inspired and nearly one-quarter derive from the vegetal kingdom.<sup>1</sup> By their very nature, plants protect themselves producing compounds commonly known as secondary metabolites.<sup>2</sup> These compounds are characterised by a high degree of structural variability, including terpenes, phenols, alkaloids, also in their glycosidic forms. Secondary metabolites produced in nature have survival functions for producing organisms and are able to set off physiological and pharmacological effects within living cells. For this reason, on one hand, natural compounds play a relevant role in the nature-aided drug-discovery process and, on the other hand, given their involvement in the plant defence system, they are useful markers for evaluating the plant adaptation to environmental stress factors (i.e. growth, geographical position, climate change, waste deposition or agricultural practices).

The extensive use of plants for millennia for healthcare purposes has found renewed interest due to recent advances in modern biology, pharmacology, chemistry and related disciplines.<sup>3</sup> The integration of traditional and modern medicine has given birth to "Green Medicine"; which has been proven to be more effective with fewer side effects on human beings. Within this context, we can mention white horehound whose great interest as a herbal remedy is testified by its presence in Pharmacopoeias and the Merck Index of Phytotherapy since 1910. Recent literature has evidenced that *Marrubium vulgare* shows several *in vivo* and *in vitro* activities including antihypertensive, antioxidant, antiinflammatory, antidiabetic, effects on respiratory system, digestive stimulant, antiasthmatic, hypolipidemic, antibacterial, and antifungal effects. Within this context, and as a part of our current research in this field, in the present contribution we focused on *M. vulgare* L. (white/common horehound), an annual herbaceous plant belonging to Lamiaceae family, native to Eurasia and northern Africa zones.<sup>4-17</sup> Due to its biological activities, the European Medicines Agency has published a monograph for *M. vulgare* in which all the traditional, non-clinical and clinical data are reported.<sup>18</sup> The main secondary metabolite of the plant is marrubiin, a furane labdane diterpenoid, isolated for the first time from *M. vulgare* in 1930 and considered a chemotaxonomic marker of *Marrubium* genus.<sup>19</sup> During the last 20 years, a plethora of works has highlighted that marrubiin can be related to pharmacological effects such as antinociceptive, antioxidant,

antigenotoxic, cardioprotective, vasorelaxant, gastroprotective, antispasmodic, immunomodulating, antioedematogenic, analgesic, and antidiabetic properties (Figure 1).<sup>10,12,20-29</sup>

Aim of our general project is to understand whether culture and environmental conditions may affect the growth and quality of white horehound. We have already demonstrated that copper(II)-induced stress reduced the uptake and translocation of cationic elements, like iron(II), potassium(I) and calcium(II), whereas increased the content of total phenols and flavonoids. It is worth noting that in connection with waste deposition and agricultural practices, copper(II), commonly present at low concentrations in the soil, tends to accumulate to high and toxic concentrations, leading to an alteration of vital physiological or biochemical functions. These effects may have a great impact on human health, since *M. vulgare* is commonly used worldwide for its medicinal properties. As a continuation of our research, the aim of the herein presented work is to develop a rapid and reproducible microwave-assisted solvent extraction (MASE) methodology for exhaustively extracting marrubiin from *M. vulgare* leaves. In the last decades, MASE has gained increasing importance for isolation and extraction of metabolites from vegetal matrices, since it entails low solvent consumption, gives high and fast extraction performances and offers protection to thermo-labile constituents. To date, MASE is considered a robust alternative to traditional extraction techniques, especially for analytical purposes.<sup>4-7</sup>

In detail, we herein describe the development of a MASE methodology easily applicable to the extraction of marrubiin from *M. vulgare*, and the evaluation of the impact of copper(II) concentration in the soil on the production of marrubiin. It is worth noting that in connection with waste deposition and agricultural practices, copper(II), commonly present at low concentration in the soil, tends to accumulate to high and toxic concentrations, leading to an alteration of vital physiological or biochemical functions. These effects may have a great impact on human health, since *M. vulgare* is used worldwide for its medicinal properties.

To set up the experimental conditions suitable to ensure the exhaustive extraction of marrubiin from the vegetal matrix, a design of experiments (DoE) approach was adopted. Moreover, a proper high-performance liquid chromatography (HPLC) method allowed the rapid analysis of the samples. The proposed MASE methodology allowed the exhaustive extraction of marrubiin from *M. vulgare* leaves and may be successfully applied for evaluating the marrubiin content in a wide number of samples. Until now, such approach for evaluating the production of marrubiin in *M. vulgare* had not been reported.

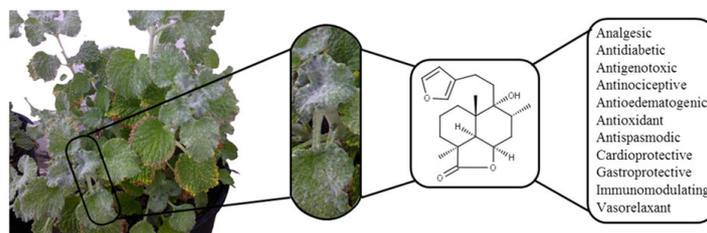


FIGURE 1 Marrubiin natural source, chemical structure and biological activities [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 2 | EXPERIMENTAL

### 2.1 | Plant material and copper treatment

Seeds of *M. vulgare* were issued from a non-contaminated wild population in the region of Béja (north-western Tunisia; latitude 36°43'30" (N), longitude 9°10'51" (E), altitude 255 m). Voucher specimen of *M. vulgare* was deposited in the herbarium unit for the research of Plant Ecology, Department Biology, Faculty of Sciences of Tunisia, under the registration number MV-N-20.

The experiments started in August 2014, and lasted four months. Initially, three replicates of 20 seeds for each treatment were sown at a depth of 2 cm in plastic pots (26 cm × 22 cm) fitted with commercial peat and sand (1:2, v/v) and maintained under Greenhouse conditions (naturally exposed to sunlight, with a temperature range 20–25°C, and relative humidity range 50–80%). Seedlings were separately transplanted into individual pots (18.2 cm × 14.2 cm) and treated with 0 (control), 80, 200 and 300 mg/L copper sulphate (CuSO<sub>4</sub>) for eight weeks. Nutrient solutions were renewed every 4 days.<sup>3</sup> At the end of the treatment (in December), the fresh leaves were separated, and dried in a drying room with active ventilation at room temperature (about 26–28°C) until constant weight. The aerial parts were stored in closed glass jars and preserved in dark conditions. The plant material was thus ground with a blade-mill (A10 IKA-Werke GmbH & Co., Staufen, Germany) to obtain a homogenous powder, just before performing the extractions.

### 2.2 | Chemicals and standards

Solvents for both HPLC (HPLC grade) and extraction procedures (analytical grade) were supplied by Carlo Erba (Milan, Italy). The reference standard of marrubiin (> 98% purity) was obtained from Carbosynth, while charcoal from Sigma Aldrich (Milan, Italy).

### 2.3 | Instruments and apparatus

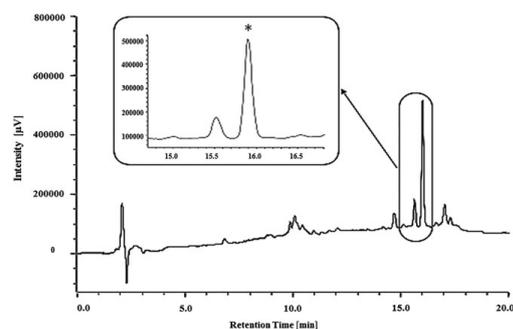
All extraction experiments were conducted under microwave irradiation in a microwave mono-mode oven (Discover® Lab-Mate instrument, CEM Corporate, Buckingham, UK) equipped with a power and temperature controller. Extract analyses were carried out on a Jasco (Tokyo, Japan) high-performance liquid chromatography ultraviolet photodiode array detection (HPLC-UV/PAD) system equipped with a Jasco AS-2055 plus autosampler, a PU-2089 plus pump and a -2010 plus multi-wavelength detector. Experimental data were acquired and interpreted with Borwin PDA and Borwin chromatograph software (Tokyo, Japan).

### 2.4 | HPLC-UV/PAD analysis

We prepared the hydroalcoholic leaves extract by applying the MASE methodology we had previously developed and reported<sup>3</sup> and analysed by HPLC-UV/PAD, at 1 mL/min flow rate, at room temperature using a Phenyl X-Bridge endcapped column (50 mm

× 4.6 mm, length × inner diameter, 3 μm particle size, macropore size 2 μm, mesopore size 13 nm, Merck, Darmstadt, Germany). The mobile phase consisted of water containing 0.1% (v/v) formic acid (A) and methanol containing 0.1% (v/v) formic acid (B). The eluent was applied onto the column in gradient mode from 10% to 90% B in 14 min, followed by an isocratic elution step for 6 min. Reconditioning of the column was accomplished by eluting the column with the mobile phase from 90% B to 10% B in 5 min, followed by a re-equilibration step of 5 min under isocratic conditions with 10% B. The analytical conditions reported earlier ensure a good resolution of marrubiin (unequivocally identified comparing its HPLC retention time and UV spectrum with those of the standard analysed in the same conditions, Figure 2 and Supporting Information Figure S1) with respect to the other analytes present in the extract. A representative chromatographic profile of the extract is reported in Figure 2. Quantitative determination of marrubiin was performed using the external standard procedure by means of a six-point calibration curve with three replicate measurements for each calibration point (Figure S1).

The limits of detection (LOD) and limits of quantification (LOQ) were estimated using the calibration curves calculated during the validation procedure, from which the average of slope (*S*) and the standard deviation of intercept (*δ*) were calculated. LOD and LOQ were obtained as follows: LOD = 3.3δ/*S*, LOQ = 10δ/*S*, thus the values were 0.016 mg/mL and 0.032 mg/mL, respectively. The calibration curve in the range 1.00–0.003 mg/mL was linear with a correlation coefficient (*R*) of 0.9878 ( $y = 7E + 06x - 434459$ ). The statistical control of the method was assessed by performing three injections of reference standard at the beginning of every measurement session on different days. The method accuracy was evaluated by assessing the marrubiin recovery by single-point standard addition method, spiking 0.30 g of samples with 3.2 mg of marrubiin. The recovery found by comparison of the spiked amount of marrubiin to the amount found in the non-spiked samples was as high as 99%.



**FIGURE 2** High-performance liquid chromatography ultraviolet photodiode array detection (HPLC-UV/PAD) profile at  $\lambda = 220$  nm of ethanol microwave-assisted solvent extraction (MASE) obtained from plant grown in Greenhouse conditions. The star \* identifies the marrubiin peak

## 2.5 | MASE parameters determination: DoE

The experimental extraction plan was designed to study three responses, i.e. total extraction yield percentage (computed as milligrams of dried extract per milligrams of dried plant material  $\times 100$ ,  $Y_1$ ), total area of the peaks recorded in the HPLC-UV/PAD chromatograms (mAU-s,  $Y_2$ ), and marrubiin yield (expressed in milligrams of marrubiin per gram of dried plant material,  $Y_3$ ). Three experimental parameters were evaluated in the first set of experiments; the parameters were solvent composition (%volume of solvent/volume of water,  $X_1$ ), time of extraction (in minutes,  $X_2$ ) and temperature (in Celsius degrees,  $X_3$ ). All other parameters including the operator were kept constant (e.g. amount of dried material, solvent volume, type of vessel, etc.). The selected parameters were varied within defined ranges according to Tables 1 and 2 for methanol and ethanol.

## 2.6 | Statistical analysis

All collected data were studied using Microsoft Excel 2010 and R for Microsoft Windows version 3.2.3, Copyright© 2014. The R Foundation for Statistical Computing R-based chemometric software routines were used for DoE calculations.

The R-based software has been developed by the Group of Chemometrics of the Italian Chemical Society (<http://gruppochemometria.it/gruppo-lavoro-r-inchemometria.html>).

## 2.7 | Sample extraction and analysis

### 2.7.1 | *Marrubium vulgare* matrices obtained from plant grown in Greenhouse conditions

Here 0.30 g of matrix were dispersed in 10 mL of solvent (Tables 1 and 2) under magnetic stirring and subjected to microwave heating (2 min ramping, maximum pressure 120 psi, maximum potency 100 W, temperature – Tables 1 and 2) for one to three cycles of 5 min with solvent renewal (Tables 1 and 2). The mixture was left to cool at room temperature, charcoal (0.090 g) was added, stirred for 5 min and filtered over paper; the last procedure was repeated once and the solvent was evaporated under reduced pressure to obtain a yellow-green oil (yield, Tables 1 and 2).

### 2.7.2 | *Marrubium vulgare* matrices from plants treated with copper(II)

Here 0.30 g of matrix in 10 mL of ethanol under magnetic stirring were subjected to microwave heating (2 min ramping, maximum pressure 120 psi, maximum potency 100 W, temperature 120°C) for three cycles of 5 min. The mixture was left to cool at room temperature, added with 0.090 g of charcoal, stirred for 5 min, and filtered over paper; the filtration was repeated once and the solvent was evaporated under reduced pressure to obtain a yellow-green oil (yield, Table 3).

All samples were analysed by HPLC-UV/PAD. Each extract was dissolved in a 90:10 methanol/water mixture (12.5 mg/mL) and filtered with a 0.45  $\mu$ m GH Polypro (GHP- PerkinElmer, China) membrane before injection into the HPLC-system. The different extracts were analysed exploiting the chromatographic conditions described earlier.

## 3 | RESULTS AND DISCUSSION

To establish a powerful approach for the determination of marrubiin in *M. vulgare* leaves, the research activity was organised as follows:

- a HPLC-UV/PAD methodology for marrubiin quantitation was properly set up, using marrubiin as external standard;
- the dried samples of *M. vulgare* leaves were ground and extracted using microwave irradiation; and a screening of the factors relevant to the MASE procedure was conducted;
- the best MASE parameters were determined by DoE, quickly allowing the identification of the best extraction conditions;
- samples of *M. vulgare* leaves grown under abiotic stress induced by  $\text{CuSO}_4$  (80–200–300 mg/L) were analysed;

every step is discussed hereinafter.

### 3.1 | Optimisation of the HPLC method

Relying on the results obtained in our previous work, a rapid and economic chromatographic method suitable for the complete resolution

**TABLE 1** Experimental parameters used on unstressed matrix to set up the extraction method

Entry number	Solvent composition $v_m:v_w$	Number of cycles $\times$ minutes	Temperature (°C)	$Y_{1m}$ Extraction yield (%)	$Y_{2m}$ Total area (mAU-s)	$Y_{3m}$ Marrubiin yield (mg/g dried plant)
1	80:20	1 $\times$ 5	40	16.29	7831107	2.20
2	20:80	1 $\times$ 5	40	13.86	7706809	0.15
3	80:20	3 $\times$ 5	40	14.60	10681435	2.03
4	20:80	3 $\times$ 5	40	16.84	9326463	0.23
5	80:20	1 $\times$ 5	80	20.82	16897393	5.17
6	20:80	1 $\times$ 5	80	22.29	14019358	3.39
7	80:20	3 $\times$ 5	80	25.22	22448046	8.01
8	20:80	3 $\times$ 5	80	17.31	13538365	2.52
9	50:50	2 $\times$ 5	60	23.66	13058322	1.22
10	50:50	2 $\times$ 5	60	24.00	14430454	1.40

Note: the solvents used are m, methanol and w, water.

**TABLE 2** Experimental parameters used on unstressed matrix to set up the extraction method

Entry number	Solvent composition $v_e:v_w$	Number of cycles $\times$ minutes	Temperature (°C)	$Y_{1e}$ Extraction yield (%)	$Y_{2e}$ Total area (mAU-s)	$Y_{3e}$ Marrubiin yield (mg/g dried plant)
1	80:20	1 $\times$ 5	40	13.84	10563133	2.92
2	20:80	1 $\times$ 5	40	18.77	6679179	1.61
3	80:20	3 $\times$ 5	40	17.85	9005601	5.25
4	20:80	3 $\times$ 5	40	23.05	7063361	2.25
5	80:20	1 $\times$ 5	80	19.27	10828100	5.39
6	20:80	1 $\times$ 5	80	20.87	7498999	2.78
7	80:20	3 $\times$ 5	80	17.35	13320100	6.24
8	20:80	3 $\times$ 5	80	22.37	6644688	4.82
9	50:50	2 $\times$ 5	60	21.45	7927552	3.49
10	50:50	2 $\times$ 5	60	19.40	12993662	4.12
11	75:25	3 $\times$ 5	80	21.27	10753422	5.89
12	100	3 $\times$ 5	120	14.53	14040263	9.20
13	100	3 $\times$ 5	80	10.10	9167573	4.98
14	80:20	3 $\times$ 5	120	21.42	11635041	8.47
15	90:10	3 $\times$ 5	100	15.47	13380524	6.45

Note: the solvents used are e, ethanol and w, water.

**TABLE 3** Experimental parameters used on stressed matrices

Copper sulphate (mg/L)	Solvent composition $v_e:v_w$	Number of cycles $\times$ minutes	Temperature (°C)	Extraction yield (%)	Extract total area (mAU-s)	Marrubiin yield (mg/g dried plant)
80	100	3 $\times$ 5	120	12.87	10200575	4.97
200	100	3 $\times$ 5	120	12.50	9509691	3.88
300	100	3 $\times$ 5	120	10.23	21595277	2.79

Note: the solvents used are e, ethanol and w, water.

of marrubiin was developed.<sup>3</sup> The optimised method allowed us to obtain a shorter time of analysis exploiting an XBridge Phenyl (5  $\mu$ m, 4.6 mm  $\times$  150 mm) column under gradient elution conditions, using a mobile phase composed of water and methanol both with 0.1% (v/v) formic acid. Marrubiin was identified comparing its HPLC retention time and UV spectrum with those of the standard analysed in the same conditions.

Quantitative determinations of this secondary metabolite were performed using external standard by means of a six-point calibration curve. The equation was obtained over a wide concentration range, in accordance with the level of marrubiin expected in the drug powder. The calibration equation thus attained was then used to compare the chromatographic marrubiin peak areas of the extracted samples with the external standard to achieve this secondary metabolite quantification. The chromatographic method response resulted a linear equation in the concentration range considered.

### 3.2 | Screening of the factors relevant for marrubiin extraction

The set up of the extraction method was conducted on the dry matrix of unstressed *M. vulgare* according to the experimental procedure described in sections 2.1–2.7. Based on our previous experience, we applied a DoE approach for identifying the best experimental

conditions.<sup>30</sup> The application of a DoE strategy for the development of optimised MASE procedures on natural matrices is a recent approach.<sup>31</sup>

Three responses were studied varying three different parameters. As responses, we considered the total extraction yield percentage ( $Y_1$ ), the total area of the peaks recorded in the HPLC-UV/PAD chromatograms ( $Y_2$ ), and the marrubiin yield ( $Y_3$ ). It is well known that the efficiency of MASE strongly relies on the selection of operating conditions and parameters affecting the extraction mechanisms and yield: the main factors affecting MASE performance are, together with sample characteristics, solvent nature and composition, extraction time and temperature (strictly related to microwave power). Accordingly, we studied the parameters, solvent composition ( $X_1$ ), extraction time ( $X_2$ ) and temperature ( $X_3$ ), applying two full factorial experimental designs in a sequential approach. The selected parameters were varied within defined ranges, as reported in Tables 1 and 2. The first eight experiments constituted the initial screening, followed by three test experiments in which the computed models were validated. The second set of four experiments was designed following the observations collected in the first phase of the study. Experiment numbers 7, 12, 13, and 14 of Table 2 were used to compute a 2<sup>2</sup> full factorial design, and experiment number 15 was used to validate this latter design. The detailed DoE experimental plan is reported in the Supporting Information.

Results of experiments with methanol (Table 1) showed that the proposed model is able to predict the trend of the expected extract

yield, total area of the chromatograms and marrubiin yield. Accordingly, we moved to the second set of experiments (Table 2), using ethanol, generally recognised as an eco-friendly and green solvent.

Results of the screening phase showed that:

- the extraction yield ( $Y_{1e}$ ) obtained using ethanol reaches higher values when factor  $X_1$  (solvent%) is kept at its minimum level (ethanol at the 80% level), while extraction time and temperature are not important to enhance the extraction yield;
- the total area of the peaks ( $Y_{2e}$ ) is greater when factors  $X_1$  and  $X_3$  at their highest levels (ethanol 100% and temperature 120°C, respectively), whereas the extraction time seems to have no effect on the response;
- the marrubiin yield ( $Y_{3e}$ ) seems higher when all three factors  $X_1$ ,  $X_2$ , and  $X_3$ , are at the highest levels simultaneously (i.e. ethanol 100%, time of extraction 15 min, and temperature 120°C).

To assess whether the responses could be improved, one additional set of four experiments was then performed at the extremes of the experimental domain. The results are reported in Table 2 (experiment numbers 7, 12, 13, 14 for model computation and 15 for model validation). The response  $Y_{1e}$ , extraction yield %, is modelled very accurately by the simplified linear model without interactions of the solvent composition and temperature, according to the equation

$$Y_{1e} = (15.85 \pm 0.09) - (3.54 \pm 0.09)X_1 + (2.12 \pm 0.09)X_2$$

which provides a good fit ( $r = 0.9993$ ) and accurate prediction of the experimental extraction yield percentage found (Table 2, experiment number 15, relative error percentage equal to 2.5%). The marrubiin

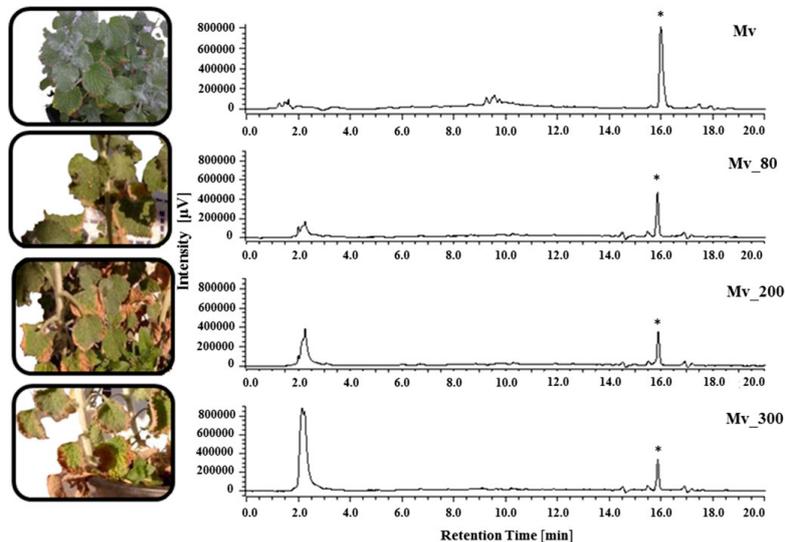
yield ( $Y_3$ ) results, turned out instead to be better modelled by the simplified linear equation

$$Y_{3e} = (7.2 \pm 0.4) + (1.6 \pm 0.4)X_3$$

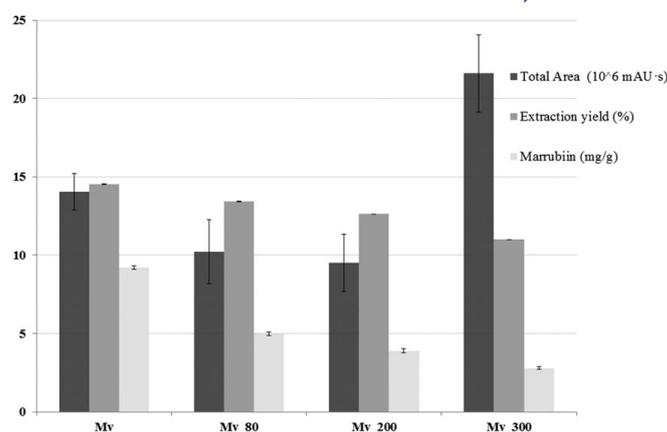
which showed only a simple dependence of the marrubiin yield from the extraction temperature and that the time of extraction and solvent composition do not actually influence the results. The model correctly predicts the marrubiin extraction yield found in the centre of the experimental domain (Table 2, experiment number 15, measured yield 6.45 mg/g, with a relative error of 12% versus predicted value of  $7.2 \pm 0.4$  mg/g). Conversely, the total area of the peak in the chromatograms ( $Y_{2e}$ ) is not described by a valid model and the mean value of the total area found,  $(1.20 \pm 0.1)107$  mAU-s, is independent of the variations made to the factors solvent composition and temperature studied (experiment numbers 7, 12, 13, and 14, Table 2).

Based on the results discussed earlier, we can consider validated the models for  $Y_{1e}$  and  $Y_{3e}$ , since the error is within the uncertainty limits associated with the used experimental procedure.

The best results in terms of marrubiin extraction yield ( $Y_3$ ), whose quantification is the main goal of the work, were obtained using 100% ethanol, applying three cycles of microwave heating of 5 min each at 120°C (experiment number 12, Table 2). To exclude thermal marrubiin degradation at the extraction conditions, an additional MASE experiment (experiment number 12, Table 2) was carried out, adding known amounts of marrubiin to vegetal matrix (obtained from plant grown in Greenhouse conditions). The good recoveries obtained (99%) confirmed the chemical stability of marrubiin under the experimental conditions. Accordingly, these conditions have been applied for evaluating the effect of the treatment with copper(II) to the marrubiin production.



**FIGURE 3** Chromatographic profile ( $\lambda = 220$  nm) and morphology of the aerial part of *Marrubium vulgare* grown in Greenhouse conditions (Mv) and subsequently to an increasing stress caused by 80 (Mv\_80), 200 (Mv\_200) and 300 (Mv\_300) mg/L of  $\text{CuSO}_4$ . The star \* identifies the marrubiin peak [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Total area, extraction yield and marrubiin concentration in *Marrubium vulgare* grown in Greenhouse conditions (Mv) and subsequently to an increasing stress caused by 80 (Mv\_80), 200 (Mv\_200) and 300 (Mv\_300) mg/L of CuSO<sub>4</sub>. Boxes and error bars represent the mean  $\pm$  standard deviation over repeated measurements

### 3.3 | MASE extraction of *M. vulgare* samples grown in Greenhouse conditions and treated with copper(II)

The optimised MASE protocol (experiment number 12, Table 2) was applied to *M. vulgare* leaves obtained from plants grown in Greenhouse conditions and treated with 80, 200 and 300 mg/L of CuSO<sub>4</sub> and the ethanolic extracts analysed using the HPLC-UV/PAD method developed here. The chromatographic profiles are shown in Figure 3, and the results are summarised in Figure 4.

Our results showed that raw materials originating from unstressed plant contained more marrubiin compared to leaves harvested from plants grown in stressed conditions. In detail, the plants were treated with CuSO<sub>4</sub> (80, 200 and 300 mg/L) for eight weeks according to the related literature.<sup>32-34</sup> As highlighted in Figure 4, there are significant differences ( $P < 0.05$ ) in the amount of marrubiin, depending on the growth conditions. Furthermore, the concentration of marrubiin ranged from a maximum amount of 9.20 mg/g of dried plant grown in Greenhouse conditions and progressively decreased to 4.97 mg/g, 3.88 mg/g and 2.79 mg/g subsequently to an increasing stress caused by 80, 200 and 300 mg/L of CuSO<sub>4</sub>, respectively. A similar trend was observed for total extraction yields (differences are significant at a statistical level,  $P < 0.05$ ), suggesting a modification of the composition of secondary metabolites in the raw materials. This hypothesis may be corroborated by the chromatographic profiles of the extracts: the peak area corresponding to the first compounds eluted, the most polar, increased with stress. Therefore, we can state that there is a negative relation between copper(II) concentration, selected as test model substance responsible for soil pollution, and marrubiin content.

To conclude, we developed a cost effective MASE methodology to exhaustively extract marrubiin from *M. vulgare* leaves. The HPLC-UV/PAD analysis of the raw extract allows a rapid and reliable quantification of marrubiin in *M. vulgare* leaves and can be applied to evaluate the quality of a wide range of samples of white horehound. To optimise the process, we used the statistical DoE approach. DoE findings indicated that the highest extraction efficiency of marrubiin with high repeatability were obtained using 100% ethanol at 120°C for 15 min, with significant benefits in terms of extraction times and

environmental impact, given that ethanol is completely biodegradable. The developed methodology was applied to evaluate the quality of white horehound, grown under Greenhouse as well as under copper(II)-mediated stress conditions. We demonstrated that growth conditions may affect the marrubiin content and therefore the quality of the herbal drug.

We propose this methodology for the characterisation of *M. vulgare* herbal drug samples, thus evaluating their exposure to abiotic stress, revealing their phytochemical status, and facilitating the identification of raw materials obtained from plant grown under stress conditions.

#### ORCID

Simona Collina  <https://orcid.org/0000-0002-2954-7558>

#### REFERENCES

- Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod*. 2012;75(3):311-335.
- Bartwal A, Mall R, Lohani P, Guru SK, Arora S. Role of secondary metabolites and Brassinosteroids in plant defense against environmental stresses. *J Plant Growth Regul*. 2013;32(1):216-232.
- Martino E, Casamassima G, Castiglione S, et al. Vinca alkaloids and analogues as anti-cancer agents: looking back, peering ahead. *Bioorg Med Chem Lett*. 2018;28(17):2816-2826.
- Amri B, Martino E, Vitulo F, et al. *Marrubium vulgare* L. leave extract: phytochemical composition, antioxidant and wound healing properties. *Molecules*. 2017;22(11):1851-1869.
- Martino E, Collina S, Rossi D, et al. Influence of the extraction mode on the yield of hyperoside, vitexin and vitexin-2-O-rhamnoside from *Crataegus monogyna* Jacq. (hawthorn). *Phytochem Anal*. 2008;19(6):534-540.
- Rossi D, Ahmed KM, Gaggeri R, et al. (R)-(-)-Aloesaponol III 8-methyl ether from *Eremurus persicus*: a novel compound against leishmaniasis. *Molecules*. 2017;22(4):519-534.
- Martino E, Ramaiola I, Urbano M, Bracco F, Collina S. Microwave-assisted extraction of coumarin and related compounds from *Melilotus officinalis* (L.) Pallas an alternative to Soxhlet and ultrasound-assisted extraction. *J Chromatogr A*. 2006;1125(2):147-151.
- Lucconi G, Chlapanidas T, Martino E, et al. Formulation of microspheres containing *Crataegus monogyna* Jacq. extract with free radical scavenging activity. *Pharm Dev Technol*. 2014;19(1):65-72.

9. Sahpaz S, Garbacki N, Tits M, Baillieu F. Isolation and pharmacological activity of phenylpropanoid esters from *Marrubium vulgare*. *J Ethnopharmacol*. 2002;79(3):389-392.
10. El Bardai S, Morel N, Wibo M, et al. The vasorelaxant activity of marrubenol and marrubiin from *Marrubium vulgare*. *Planta Med*. 2003;69(1):75-77.
11. El Bardai S, Lyoussi B, Wibo M, Morel N. Comparative study of the antihypertensive activity of *Marrubium vulgare* and of the dihydropyridine calcium antagonist amlodipine in spontaneously hypertensive rat. *Clin Exp Hypertens*. 2004;26(6):465-474.
12. Meyre-Silva C, Yunes RA, Schlemper V, Campos-Buzzi F, Cechinel-Filho V. Analgesic potential of marrubiin derivatives, a bioactive diterpene present in *Marrubium vulgare* (Lamiaceae). *IL Farmaco*. 2005;60(4):321-326.
13. Stulzer HK, Tagliari MP, Zampirolo JA, Cechinel-Filho V, Schlemper V. Antioedematogenic effect of marrubiin obtained from *Marrubium vulgare*. *J Ethnopharmacol*. 2006;108(3):379-384.
14. Masoodi MH, Bahar A, Zagar IM, Khan SA, Khan S, Singh P. Antibacterial activity of whole plant extract of *Marrubium vulgare*. *Afr J Biotechnol*. 2008;7:86-87.
15. Boudjelal A, Henchiri C, Siracusa L, Sari M, Ruberto G. Compositional analysis and *in vivo* anti-diabetic activity of wild Algerian *Marrubium vulgare* L. infusion. *Fitoterapia*. 2012;83(2):286-292.
16. Akther N, Shawl AS, Sultana S, Chandan BK, Akther M. Hepatoprotective activity of *Marrubium vulgare* against paracetamol induced toxicity. *J Pharm Res*. 2013;7(7):565-570.
17. Boullila A, Sanaa A, Ben Salem I, et al. Antioxidant properties and phenolic variation in wild populations of *Marrubium vulgare* L. (Lamiaceae). *Ind Crop Prod*. 2015;76:616-622.
18. 9 July 2013 EMA/HMPC/604273/2012 Committee on Herbal Medicinal Products (HMPC).
19. McCrea A. Comparative marrubiin content in *Marrubium vulgare* from European vs. American seed. *J Am Pharm Assoc*. 1930;19(3):231.
20. De Jesus RAP, Cechinel-Filho V, Oliveira AE, Schlemper V. Analysis of the antinociceptive properties of marrubiin isolated from *Marrubium vulgare*. *Phytomedicine*. 2000;7(2):111-115.
21. Mnonopi N, Levendal RA, Davies-Coleman RT, Frost CL. The cardioprotective effects of marrubiin, a diterpenoid found in *Leonotis leonurus* extracts. *J Ethnopharmacol*. 2011;138(1):67-75.
22. Laonigro G, Lanzetta R, Parrilli M, Adinolfi M, Mangoni L. The configuration of the diterpene spiro ethers from *Marrubium vulgare* and from *Leonotis leonurus*. *Gazz Chim Ital*. 1979;109:145-150.
23. Paula de Olivera A, Santin JR, Lemos M, et al. Gastroprotective activity of methanol extract and marrubiin obtained from leaves of *Marrubium vulgare* L. (Lamiaceae). *J Pharm Pharmacol*. 2011;63(9):1230-1237.
24. Zaabat N, Hay AE, Michalet S, et al. Antioxidant and antigenotoxic properties of compounds isolated from *Marrubium deserti* de Noe. *J Food Chem Toxicol*. 2011;49(12):3328-3335.
25. Karioti A, Skopeliti M, Tsitsilonis O, Heilmann J, Skaltsa H. Cytotoxicity and immunomodulating characteristics of labdane diterpenes from *Marrubium cylleneum* and *Marrubium velutinum*. *Phytochemistry*. 2007;68(11):1587-1594.
26. Hellen K, Stulzer HK, Tagliari MP, Zampirolo JA, Cechinel-Filho V, Schlemper V. Antioedematogenic effect of marrubiin obtained from *Marrubium vulgare*. *J Ethnopharmacol*. 2006;108:379-384.
27. De Souza MM, De Jesus RAP, Cechinel-Filho V, Schlemper V. Analgesic profile of hydroalcoholic extract obtained from *Marrubium vulgare*. *Phytomedicine*. 1998;5(2):103-107.
28. Mnonopi N, Levendal RA, Mzilikezi N, Frost CL. Marrubiin, a constituent of *Leonotis leonurus*, alleviates diabetic symptoms. *Phytomedicine*. 2012;19(6):488-493.
29. Popoola OK, Elbagory AM, Ameer F, Hussein AA. Marrubiin. *Molecules*. 2016;18:9049-9060.
30. Marrubini G, Fattorini P, Previderé C, et al. Experimental design applied to the optimization of microwave-assisted DNA hydrolysis. *J Chromatogr A*. 2012;1249:8-16.
31. Leone A, Tamborrino A, Zagaria R, Sabella E, Romaniello R. Plant innovation in the olive oil extraction process: a comparison of efficiency and energy consumption between microwave treatment and traditional malaxation of olive pastes. *J Food Eng*. 2015;146:44-52.
32. Amri B, Kaab SB, Gouia H, Martino E, Collina S, Ben-Kaâb LB. Copper-induced changes in nutrient uptake, enzymatic and non-enzymatic antioxidant systems in horehound (*Marrubium vulgare* L.). *Botan Sci*. 2017;95(3):565-575.
33. Gupta D, Abdullah. Toxicity of copper and cadmium on germination and seedling growth of maize (*Zea mays* L.) seeds. *Ind J Sci Res*. 2011;2:67-70.
34. Tóth G, Hermann T, Da Silva MR, Montanarella L. Heavy metals in agricultural soils of the European Union with implications for food safety. *Environ Int*. 2016;88:299-309.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Martino E, Della Volpe S, Cavalloro V, et al. The use of a microwave-assisted solvent extraction coupled with HPLC-UV/PAD to assess the quality of *Marrubium vulgare* L. (white horehound) herbal raw material. *Phytochemical Analysis*. 2019;1-8. <https://doi.org/10.1002/pca.2820>

## ***V. Contributions to the research***

No (wo)man and no scientist is an island.

Several people directly contributed to the work performed in this thesis. Here they are, listed by research group. Additional people in the working groups of Prof. Simona Collina and Prof. Anna Hirsch did not directly participate but were always available for advice and constructive discussions around science.

**University of Pavia** Prof. Simona Collina, Dr. Daniela Rossi, Dr. Marcello Di Giacomo, Dr. Rita Nasti, Ms. Roberta Listro, Ms. Michela Parafioriti, Mr. Michele Queirolo, Ms. Katia Russo and Mr. Federico Sala: MedChem.

**Helmholtz Institute for Pharmaceutical Research Saarland** Prof. Anna K.H. Hirsch, Dr. Walid Elgaher, Dr. Varsha Jumde and Dr. Yagiz Unver (HuR project); Dr. Jelena Konstantinovic, Ms. Alaa Alhayek and Ms. Cansu Kaya (LasB project): MedChem.

**University of Milano** Dr. Francesca Vasile and Dr. Donatella Potenza: ligand-observed NMR studies.

**University of Magna Graecia** Prof. Stefano Alcaro, Dr. Giosuè Costa and Dr. Francesca Alessandra Ambrosio: Molecular dynamics studies.

**University of Trento** Prof. Alessandro Provenzani and Dr. Chiara Zucal: Native protein expression and purification; *preliminary biological evaluations (tbp)*.

**Helmholtz Zentrum Munich** Prof. Michael Sattler, Dr. Marta Pabis and Dr. Grzegorz Popowicz: New protein constructs expression and purification; *protein-observed NMR studies (tbp)*.

The listed groups and scientists constitute an ever evolving research network whose members are welcome to exchange opinions (and sometimes location) to actively integrate several disciplines, all fundamental to the success of a complete Medicinal Chemistry Campaign.

The collaborations interwoven and strengthened during this 3-year experience, along with those yet to be consolidated, certainly represent an added value to our research.