



**HAL**  
open science

## The diversity of splicing modifiers acting on A -1 bulged 5 -splice sites reveals rules for rational drug design

Florian Malard, Antje C Wolter, Julien Marquevielle, Estelle Morvan, Agathe Ecoutin, Simon H Rüdisser, Frédéric H T Allain, Sebastien Campagne

### ► To cite this version:

Florian Malard, Antje C Wolter, Julien Marquevielle, Estelle Morvan, Agathe Ecoutin, et al.. The diversity of splicing modifiers acting on A -1 bulged 5 -splice sites reveals rules for rational drug design. Nucleic Acids Research, 2024, 52 (8), pp.4124-4136. 10.1093/nar/. hal-04576074

**HAL Id: hal-04576074**

**<https://hal.science/hal-04576074>**

Submitted on 15 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The diversity of splicing modifiers acting on A<sub>-1</sub> bulged 5'-splice sites reveals rules for rational drug design

Florian Malard<sup>1,2,†</sup>, Antje C. Wolter<sup>3,†</sup>, Julien Marquevielle<sup>1,2</sup>, Estelle Morvan<sup>4</sup>, Agathe Ecoutin<sup>1,2</sup>, Simon H. Rüdiger<sup>5</sup>, Frédéric H.T. Allain<sup>3,\*</sup> and Sebastien Campagne<sup>1,2,3,\*</sup>

<sup>1</sup>Université de Bordeaux, Inserm U1212, CNRS UMR5320, ARNA unit, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France

<sup>2</sup>Institut Européen de Chimie et Biologie, 2 rue Robert Escarpit, 33607 Pessac Cedex, France

<sup>3</sup>ETH Zürich, Department of Biology, Institute of Biochemistry, Hönggerberggring 64, 8093 Zürich, Switzerland

<sup>4</sup>Institut Européen de Chimie et Biologie, UAR3033 CNRS, Université de Bordeaux, INSERM US01, Pessac 33600, France

<sup>5</sup>ETH Zürich, Department of Biology, BioNMR platform, Hönggerberggring 64, 8093 Zürich, Switzerland

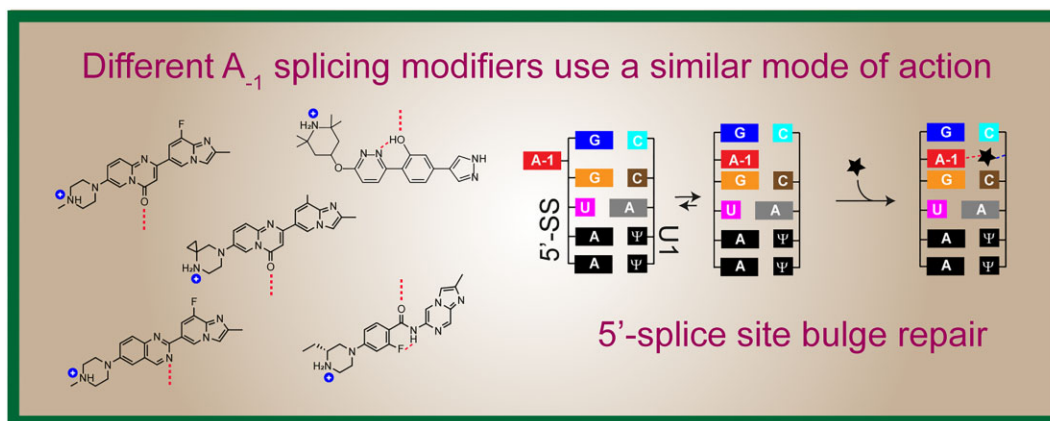
\*To whom correspondence should be addressed. Tel: +33 5 40 00 30 38; Fax: +33 5 40 00 30 68; Email: sebastien.campagne@inserm.fr  
Correspondence may also be addressed to Frédéric H.T. Allain. Email: allain@bc.biol.ethz.ch

†The first two authors should be regarded as Joint First Authors.

## Abstract

Pharmacological modulation of RNA splicing by small molecules is an emerging facet of drug discovery. In this context, the *SMN2* splicing modifier SMN-C5 was used as a prototype to understand the mode of action of small molecule splicing modifiers and propose the concept of 5'-splice site bulge repair. In this study, we combined *in vitro* binding assays and structure determination by NMR spectroscopy to identify the binding modes of four other small molecule splicing modifiers that switch the splicing of either the *SMN2* or the *HTT* gene. Here, we determined the solution structures of risdiplam, branaplam, SMN-CX and SMN-CY bound to the intermolecular RNA helix epitope containing an unpaired adenine within the G<sub>2</sub>A<sub>1</sub>G<sub>+1</sub>U<sub>+2</sub> motif of the 5'-splice site. Despite notable differences in their scaffolds, risdiplam, SMN-CX, SMN-CY and branaplam contact the RNA epitope similarly to SMN-C5, suggesting that the 5'-splice site bulge repair mechanism can be generalised. These findings not only deepen our understanding of the chemical diversity of splicing modifiers that target A<sub>-1</sub> bulged 5'-splice sites, but also identify common pharmacophores required for modulating 5'-splice site selection with small molecules.

## Graphical abstract



## Introduction

Alternative pre-mRNA splicing is a crucial step in gene expression that determines mRNA localization, translation, and decay, and thus plays a significant role in regulating gene expression. Abnormal alternative splicing patterns have been associated with various human diseases (1).

A major limiting step in the assembly of the spliceosome is the definition of the 5'-splice site by the U1 small nuclear

ribonucleoprotein (U1 snRNP), a seeding particle of the major spliceosome (2,3). The ribonucleoprotein uses the 5'-end of the U1 snRNA to base pair with the 5'-splice site and initiates spliceosome assembly (4). The recognition process is influenced by three factors: (i) the sequence of the 5'-splice site and its complementarity with the 5'-end of U1 snRNA (5,6), (ii) the accessibility of the 5'-splice site, which may be trapped by inhibitory secondary structures (7), and (iii) the network of

Received: April 28, 2023. Revised: December 7, 2023. Editorial Decision: February 23, 2024. Accepted: March 7, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

*trans*-acting splicing factors that can either enhance or inhibit U1 snRNP recruitment (8). Non-canonical base pair registers and bulged 5'-splice sites are often associated with alternative splicing patterns, and single nucleotide polymorphisms at the 5'-splice site have been linked to diseases (9–11). As a result, the definition of the 5'-splice site by U1 snRNP is a key step for splicing regulation and a major target for specific splicing correction. In recent decades, synthetic tools that enable the modulation of 5'-splice site selection and the correction of alternative splicing have emerged and have begun to offer therapeutics for inherited diseases (12–16). There are currently two main classes of synthetic splicing switches: the antisense oligonucleotides (17) and the small molecule splicing modifiers (16,18–20).

Antisense oligonucleotides modulate 5'-splice site selection indirectly, either by blocking key *cis*-regulatory RNA elements or disrupting the formation of inhibitory secondary structures (17), while small molecule splicing modifiers act directly at the interface between U1 snRNP and the 5'-splice site to modify 5'-splice site selection (19,21). There are two classes of small molecule splicing modifiers that have been described: the *SMN2* and the *HTT* splicing modifiers, both of which act on A<sub>-1</sub> bulged 5'-splice sites (16,18). The mode of action of the *SMN2* splicing modifier family was recently proposed based on studies with the prototypical member, SMN-C5 (22). Analysis of the sequences of the exons sensitive to SMN-C5 showed an enrichment for a G<sub>-2</sub>A<sub>-1</sub>G<sub>+1</sub>U<sub>+2</sub> motif at their exon-intron junctions. The solution structure of SMN-C5 bound to the RNA duplex formed by the U1 snRNA and the 5'-splice site (referred to as the RNA helix) revealed that the splicing modifier binds the RNA helix in the major groove at the exon-intron junction. Mechanistically, SMN-C5 acts as a molecular glue at the interface between U1 snRNP and the A<sub>-1</sub> bulged 5'-splice site. The carbonyl group of the central aromatic ring forms a direct hydrogen bond with the NH<sub>2</sub> group of the unpaired adenine -1, explaining the specificity for 5'-splice sites containing a G<sub>-2</sub>A<sub>-1</sub>G<sub>+1</sub>U<sub>+2</sub> motif. By stabilising the unpaired adenine in the RNA base stack, SMN-C5 stimulates and facilitates the binding of the U1 snRNP protein U1-C, a crucial step for 5' splice-site recognition as the precursor of spliceosome assembly. SMN-C5 is not only a molecular glue between the 5'-splice site of *SMN2* exon 7 and the U1 snRNA, but also an allosteric effector of 5'-splice site definition on *SMN2* exon 7. This mechanism of action, which has not been seen before, has been referred to as 5'-splice site bulge repair (or in short bulge repair), as the repair of the A<sub>-1</sub> bulge leads to the strengthening of the 5'-splice site (Figure 1). While SMN-C5 specifically corrects *SMN2* exon 7 splicing using the bulge repair mechanism, it is unclear whether the molecule can represent the entire *SMN2* splicing modifier family (23–25) due to the chemical diversity of scaffolds. This includes the FDA-approved risdiplam, the first orally available therapeutic for spinal muscular atrophy patients (23).

*SMN2* splicing modifiers increase the inclusion of *SMN2* exon 7, thereby restoring the expression of the functional SMN protein from the *SMN2* gene. In contrast, small molecule splicing modifiers that target the toxic huntingtin mRNA (mHTT) promote the inclusion of a premature termination codon (PTC) containing pseudo-exon that causes mRNA decay (18). These *HTT* splicing modifiers may represent potential therapeutics for Huntington's disease as they lower the levels of the toxic mHTT mRNA. Even though both classes of splicing modifiers identified so far promote

exon inclusion, they have different functional consequences: *SMN2* splicing modifiers restore essential cellular function, while *HTT* splicing modifiers help cells get rid of a toxic transcript (Figure 1). Like *SMN2* splicing modifiers, *HTT*-lowering molecules also act on an A<sub>-1</sub> bulged 5'-splice site and are most effective when a purine-rich element is located in the exon (18). Given that *HTT* splicing modifiers have different chemical scaffolds than SMN-C5, understanding the structural diversity of splicing modifiers acting at the interface with U1 snRNA and an A<sub>-1</sub> bulged 5'-splice site may also help to establish the rules governing their specificity for bulged splice sites, thereby enabling the rational design of small molecules that target other -1 bulged 5'-splice sites.

In order to understand how the *SMN2* and *HTT* splicing modifiers SMN-CX, SMN-CY, risdiplam and branaplam bind to the A<sub>-1</sub> bulge at the interface between U1 snRNP and the 5'-splice site, we used a combination of *in vitro* binding assays and structure determination by NMR spectroscopy. Our atomic pictures of splicing modifier/RNA complexes showed that splicing modifiers representing three different chemotypes bind to the same epitope in the RNA helix, near the unpaired adenine in position -1 of the 5'-splice site. Through this work, we have gained valuable insights into the relationship between chemical scaffold diversity and the ability of splicing modifiers to target A<sub>-1</sub> bulged 5'-splice sites.

## Materials and methods

### Sample preparation

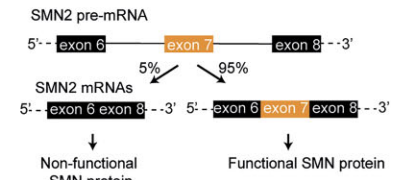
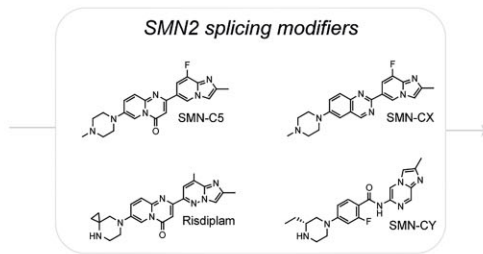
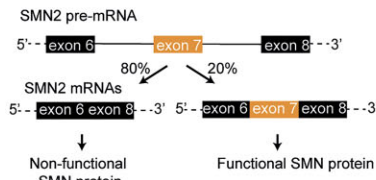
The 5'-splice site of *SMN2* exon 7 (5'-GGA<sub>-1</sub>GUAAGUCU-3') and the U1 snRNA 5'-end (5'-AUACΨΨACCUG-3') were purchased (Horizon), deprotected according to manufacturer instructions and lyophilized. To remove impurities, the RNAs were precipitated by butanol extraction, resuspended in water and lyophilized. The RNA helix was prepared by mixing equimolar amounts of each strand dissolved in the NMR buffer (MES-d<sub>13</sub> 5 mM pH 5.5, NaCl 50 mM). Risdiplam and branaplam were either provided by Skyhawk Therapeutics or purchased (MCE Inhibitors) while SMN-CX and SMN-CY were provided by Roche. Risdiplam, branaplam and SMN-CX were resuspended in DMSO-d<sub>6</sub> (Cambridge Isotopes) while SMN-CY was soluble at high concentration (5 mM) in the NMR buffer. Homonuclear bidimensional NMR experiments on Risdiplam, branaplam, SMN-CX and SMN-CY were collected with a 1 mM sample of small molecule in the NMR buffer containing 10% D<sub>2</sub>O. Using these data, we performed the proton assignment of the small molecules in their free state. To assign the carbon resonances of branaplam, we collected NMR experiments on a 10 mM sample prepared in 100% DMSO-d<sub>6</sub> (1D <sup>13</sup>C, 1D <sup>1</sup>H, 2D <sup>1</sup>H-<sup>13</sup>C HSQC and 2D <sup>1</sup>H-<sup>13</sup>C HMBC). For the titration with risdiplam, branaplam and SMN-CX, samples containing 0.15 mM of the RNA helix were prepared in the NMR buffer supplemented with 10% D<sub>2</sub>O and 2% DMSO-d<sub>6</sub>. These solutions of RNA helix were titrated with stock solutions (40 mM) of branaplam, risdiplam and SMN-CX dissolved in DMSO-d<sub>6</sub>. For the titration with SMN-CY, a sample containing 0.15 mM of intermolecular RNA helix was prepared in the NMR buffer supplemented with 10% D<sub>2</sub>O and titrated with a stock solution (5 mM) of SMN-CY dissolved in the NMR buffer. In order to record the data for determination of the dissociation constants, branaplam and risdiplam were dissolved at 50 μM in

**A**

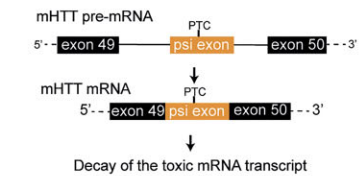
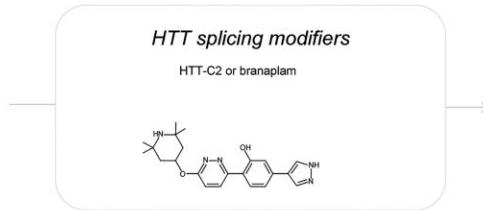
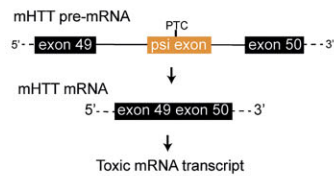
*Disease associated splicing pattern*

*Therapeutic splicing correction*

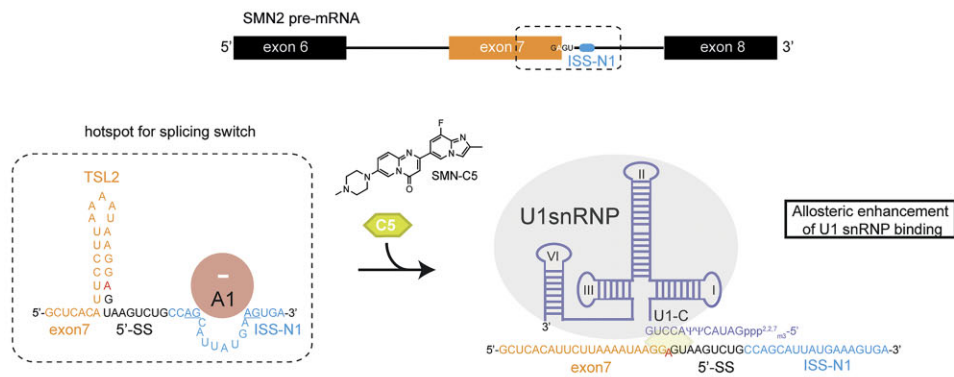
**Spinal Muscular Atrophy**



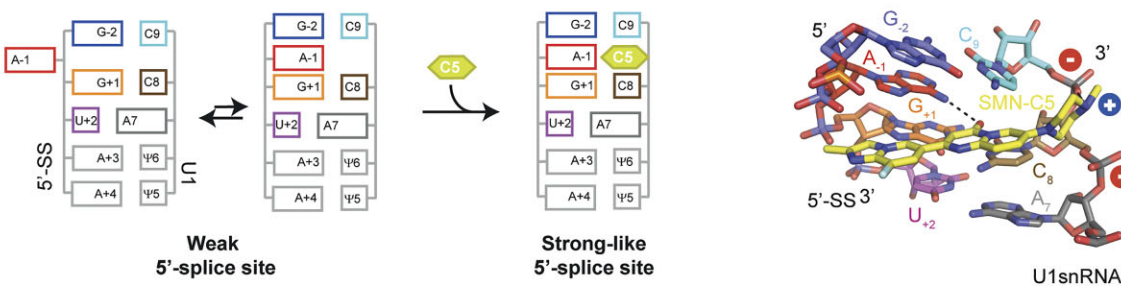
**Huntington disease**



**B**



**Specific splicing correction via 5'-splice site bulge repair**



**Figure 1.** Small molecule splicing modifiers acting on A<sub>1</sub> bulged 5'-splice sites: functions and mechanism of 5'-splice site bulge repair. **(A)** Schematic representation of the splicing correction induced by *HTT* and *SMN2* splicing modifiers. Both classes of splicing modifiers promote specific exon inclusion. *SMN2* splicing modifiers induce *SMN2* exon 7 inclusion and allow the production of functional SMN protein from the *SMN2* gene. *HTT* splicing modifiers induce the inclusion of a pseudo exon (psi exon) containing a premature stop codon (PTC) triggering the decay of the toxic *mHTT* mRNA. **(B)** Mode of action of the *SMN2* splicing modifier SMN-C5. The hotspot for *SMN2* splicing correction is the weak 5'-splice site of exon 7 that is sequestered in an inhibitory secondary structure (TSL2), flanked by a strong splicing repressor (ISS-N1) and has a non-optimal sequence with an adenine in position -1. SMN-C5 stimulates the binding of U1 snRNP on this weak 5'-splice site by binding the RNA helix formed by U1 snRNP and the 5'-splice site at the exon-intron junction. SMN-C5 contacts specifically the unpaired adenine in position -1 and transforms the weak 5'-splice site of *SMN2* exon 7 into a stronger one through the mechanism of 5'-splice site bulge repair.



the NMR buffer complemented with 10% D<sub>2</sub>O and a stock solution of RNA duplex at 1 mM in the NMR buffer was used to perform the titration. To collect the NMR data for structure determination of the small molecule/RNA helix complexes, samples contained 1 mM of the RNA helix and 1.5 molar equivalent of small molecule in the NMR buffer in 100% D<sub>2</sub>O.

### NMR spectroscopy

The proton resonance assignment of each small molecule was performed by combining 2D homonuclear experiments (2D <sup>1</sup>H–<sup>1</sup>H TOCSY and 2D <sup>1</sup>H–<sup>1</sup>H NOESY) in 90% H<sub>2</sub>O /10% D<sub>2</sub>O at 293 K. Data were recorded on an AVIII 750 MHz spectrometer or on an AVIII 400 MHz spectrometer (Bruker). The proton and carbon resonance assignment of branaplam was derived from the analysis of the 1D <sup>1</sup>H, 1D <sup>13</sup>C, 2D <sup>1</sup>H–<sup>1</sup>H TOCSY, 2D <sup>1</sup>H–<sup>1</sup>H NOESY, 2D <sup>1</sup>H–<sup>13</sup>C HSQC and 2D <sup>1</sup>H–<sup>13</sup>C HMBC experiments recorded in 100% DMSO using the AVIII 400 MHz spectrometer (Bruker) at 293 K. For SMN-CY, we also recorded a 2D <sup>1</sup>H–<sup>15</sup>N HSQC and a 2D <sup>1</sup>H–<sup>19</sup>F HOESY (mixing time 1000 ms) using a 5 mM sample dissolved in the NMR buffer complemented with 10% D<sub>2</sub>O. These experiments were recorded on an AVIII 400 MHz spectrometer (Bruker) at 293 K. Processing of the data was performed using Topspin 4.1 (Bruker) and the analysis of the data was done using CARA (26). To extract dissociation constants between risdiplam, branaplam and the RNA helix, we monitored the proton chemical shifts of the small molecule upon stepwise addition of the RNA duplex. We recorded a 1D <sup>1</sup>H of the small molecule at the following small molecule:RNA ratios : 1:0; 1:0.05; 1:0.1; 1:0.2; 1:0.3; 1:0.4; 1:0.5; 1:0.6; 1:0.8; 1:1; 1:1.4. These experiments were recorded at 293 K using an AVII 700 MHz. To map the small molecule binding pocket on the RNA duplex, we monitored the chemical shifts of the imino protons and of the H6-H5 couples of the pyrimidines (C and U) by recording 1D <sup>1</sup>H and 2D <sup>1</sup>H–<sup>1</sup>H-TOCSY experiments at 293 K upon stepwise addition of the small molecule splicing modifiers. These experiments have been recorded on an AVIII 600 MHz or on an AVIII 700 MHz spectrometers. Data were collected for the following RNA duplex : small molecule ratios: 1:0; 1:1; 1:2. In order to assign the RNA and the small molecule resonances in the complex to eventually identify intermolecular NOEs and NOE-derived distances, we recorded the following experiments: 1D <sup>1</sup>H, 2D <sup>1</sup>H–<sup>1</sup>H TOCSY, 2D <sup>1</sup>H–<sup>13</sup>C aliphatic HSQC, 2D <sup>1</sup>H–<sup>13</sup>C aromatic HSQC, 2D <sup>1</sup>H–<sup>1</sup>H NOESY. Several mixing times were used for the 2D <sup>1</sup>H–<sup>1</sup>H NOESY depending on the small molecule. For SMN-CY, we recorded 2D <sup>1</sup>H–<sup>1</sup>H NOESY with 120, 200 and 400 ms mixing times at 293K using an AVIII 900 MHz spectrometer (Bruker). For SMN-CX, risdiplam and branaplam, the 2D <sup>1</sup>H–<sup>1</sup>H NOESY were recorded with mixing times of 200 and 400 ms at 293K using an AVIII 900 MHz or an AVIII 800 MHz spectrometers (Bruker). The resonance assignments of the RNA duplex in complex with the different small molecules was performed with CARA (27) by comparison to the RNA resonance assignment of the RNA duplex in complex with SMN-C5 (BMRB code 34312).

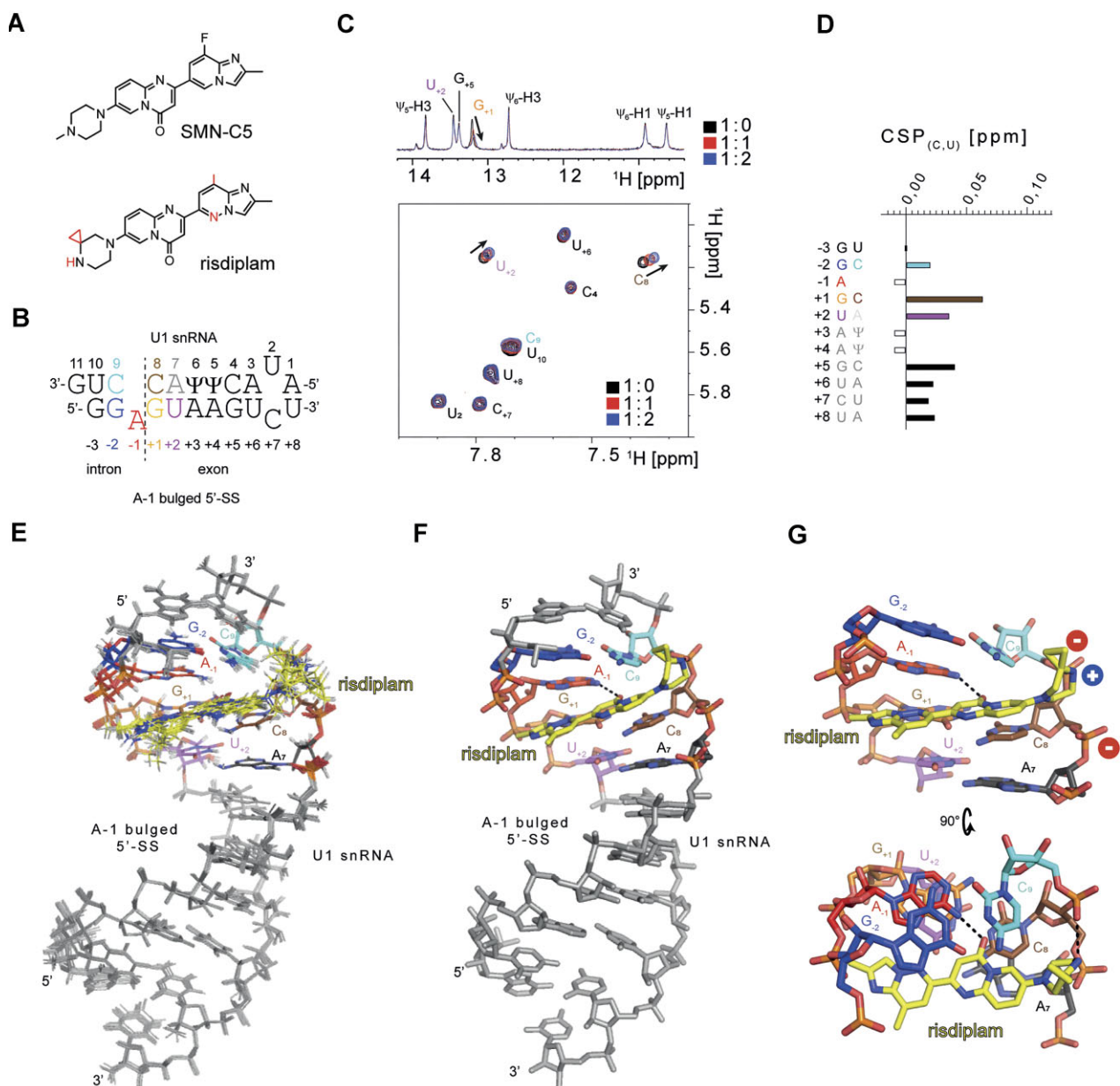
### Structure calculations

To determine the structures of the RNA duplex/small molecule complexes, we first determined the structure of the

RNA within the complex. The RNA base pairing was determined based on the 2D <sup>1</sup>H–<sup>1</sup>H NOESY recorded in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. Angular restraints for the sugar pucker were derived from the analysis of the 2D <sup>1</sup>H–<sup>1</sup>H TOCSY. Then, we combined this information with the chemical shifts of experimental NOESY spectra recorded in 100% D<sub>2</sub>O with a mixing time of 200 ms (2D <sup>1</sup>H–<sup>1</sup>H NOESY) for automatic NOE assignment and structure calculation using CYANA3.98.15 (28). A set of intramolecular NOE-derived distances was obtained from the automatic analysis and the peak lists were manually refined to obtain a list of NOE-derived distance restraints. For the identification of the intermolecular NOEs, we combined the assignment of the RNA and the small molecule in a single CARA file and performed the assignment manually. Due to the excess of small molecule compared to the RNA duplex, strong NMR signals of free state impaired the automatic analysis of the intermolecular NOEs. Therefore, intermolecular NOE correlations were identified, integrated and calibrated manually. To calibrate the intermolecular NOEs, we classified the NOEs according to their strength and stronger NOEs were set to 4 Å while the upper limit of weaker NOEs set to 5.5 Å. In case the NOE-derived distance set to 5.5 Å was violated, the upper limit was increased to 6Å. In order to include the small molecule in the CYANA calculation, a coordinate file together with a definition file for CYANA were generated using the Win!P software (<http://www.biochem-cafisch.uzh.ch/download>). After adding the CYANA ring closure restraints for the RNA helix, initial structure calculations of the RNA duplex/SMN-CY complex were performed using CYANA in the torsion-angle space. For each complex, we computed between 200 and 500 models and selected the 50 lowest target function models for refinement in Cartesian space using the sander module of the AmberTools23 package (29). The AMBER parameters for risdiplam, branaplam, SMN-CX and SMN-CY were generated using the Antechamber module of AmberTools23 package (29) and are available upon request. For the calculations of the complexes formed by the RNA helix and SMN-CX, risdiplam or branaplam, the sander protocol was performed in water and was composed of a minimisation followed by a simulated annealing protocol including a short run of molecular dynamics (30 ps). For the calculation of the complex formed the RNA helix and SMN-CY, the sander protocol was also performed in water and was composed of a minimisation followed by a simulated annealing protocol including a longer run of molecular dynamics (500 ps). In this latter case, we used a longer molecular dynamic step to explore the mobility of the ligand and the stability of the proposed intramolecular connection between the fluorine and the amide proton of SMN-CY. Analysis of the calculation was performed using the sviol script of AmberTools23 and lowest energy models were selected. Structures were aligned and visualised using PyMol (Schrodinger).

### Binding affinity determination

Binding affinities ( $K_D$ ) were derived by plotting the chemical shift perturbations (CSP) observed on 1D <sup>1</sup>H experiments of the titrations of risdiplam or branaplam by the RNA helix against the concentration of the RNA helix in μM. We applied a cumulative curve fitting with a single  $K_D$  value using `curve_fit` from the SciPy module and pandas for data handling in Python 3 (30,31).



**Figure 2.** Structural basis for risdiplam recognition of the RNA helix formed upon 5'-splice site recognition. **(A)** Planar structures of SMN-C5 and risdiplam. The differences between both *SMN2* splicing modifiers are highlighted in red. **(B)** Schematic representation of the RNA helix. **(C)** Assigned overlays of the 1D  $^1\text{H}$  (imino region, top) and 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (H6-H5 region, bottom) spectra of the RNA helix recorded upon successive additions of risdiplam. The spectra are coloured according to the ratio RNA helix: risdiplam. **(D)** Plot of the pyrimidine (H5-H6) chemical shift perturbations (CSP) as a function of the positions on the 5'-splice site. White bars correspond to position without values (-1, +3 and +4). In position 7, the contributions of both pyrimidines have been added. **(E)** Overlay of the 14 solution structures of risdiplam bound to the RNA helix. **(F)** Atomic representation of the lowest energy model of the NMR ensemble. **(G)** Closed up views of the intermolecular interface between risdiplam and the RNA helix epitope. Intermolecular hydrogen bonds are shown as dashed lines.

## Results

### Molecular basis for *SMN2* splicing correction by risdiplam

The first *SMN2* small molecule splicing modifier that was evaluated clinically was RG7800 (an homologue of SMN-C5). However, an unexpected eye finding was observed in an animal study evaluating the long term safety of RG7800 and the associated clinical study was discontinued (20). Accordingly, chemical modifications were performed on RG7800 to improve selectivity and reduce phototoxicity, metabolism, and basicity in the development of risdiplam (20). In com-

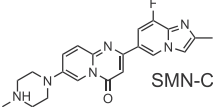
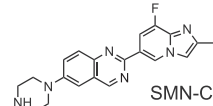
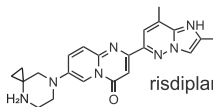
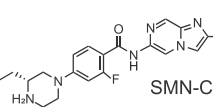
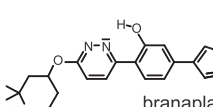
parison with SMN-C5, risdiplam differed in several ways: (i) the fluorine group on the first aromatic cycle was replaced with a methyl group; (ii) the methylated piperazine group was dealkylated to prevent N-dealkylation *in vivo* and (iii) the piperazine group was modified by adding a cyclopropyl group at position 3 (Figure 2A). The modifications made to RG7800 in the development of risdiplam had no effect on the mode of action compared to SMN-C5, risdiplam showed optimal biological activity ( $\text{IC}_{50} = 23 \text{ nM}$ , Table 1) on *SMN2* splicing (32). Here, we used nuclear magnetic resonance (NMR) spectroscopy to study the binding mode of risdiplam to the RNA

helix formed by the 5'-splice site of *SMN2* exon 7 and the 5' end of U1 snRNA (Figure 2B). When risdiplam was added to the RNA helix sample, we observed line broadening of the  $G_{+1}$  imino proton resonance and characteristic chemical shift perturbations of the H5-H6 TOCSY correlation peaks of  $C_8$  and  $U_{+2}$  (Figure 2C, D). In order to quantify the binding affinity of risdiplam for the RNA helix, we monitored the chemical shift perturbations (CSP) of risdiplam protons upon stepwise addition of the RNA helix. By fitting the  $^1\text{H}$  CSP as a function of the RNA concentration, we determined a dissociation constant of  $12 \pm 1 \mu\text{M}$  (Supplementary Figure S1 and Table 1), in agreement with the previously determined affinity of SMN-C5 for the same RNA (22). To get a more detailed understanding of how risdiplam binds to the  $A_{-1}$  bulge of the RNA helix, we attempted to determine the structure of risdiplam bound to the RNA helix using solution state NMR spectroscopy. The proton resonances of risdiplam were assigned using a combination of two dimensional homonuclear experiments (Supplementary Figure S2 and Supplementary Table S1) and we identified 20 intermolecular nuclear Overhauser effects (NOEs) that unambiguously defined the intermolecular interface (Supplementary Figure S3). However, a strong line broadening effect was observed on the aromatic protons of the small molecule which prevented the identification of intermolecular connection between the aromatic core of risdiplam and the RNA (Supplementary Figure S4 and Supplementary Table S1). The solution structure of risdiplam in complex with the RNA helix is composed of 14 models that overlay with an all-atom Root Mean Square Deviation (RMSD) of  $0.54 \pm 0.12 \text{ \AA}$  (Table 2). In the structure (Figure 2E–G), the central aromatic ring inserts between  $C_8$  and  $C_9$  of the U1 snRNA and the carbonyl group forms a direct hydrogen bond with the amino group of the unpaired adenine at position  $-1$ . The positively charged amine in the piperazine group forms a salt bridge with the negatively charged oxygen in the phosphate group of  $C_9$  and in addition, we observed a hydrogen bond from the HN to  $C_9$  phosphate oxygen that may explain the higher affinity of risdiplam over SMN-C5 (Figure 2G and Table 1). Like SMN-C5, risdiplam links the U1 snRNA and the 5'-splice site of *SMN2* exon 7 and serves as the minimal *trans*-splicing factor. Overall, our data suggest that risdiplam acts through the mechanism of 5'-splice site bulge repair.

### SMN2 splicing correction can be achieved using a splicing modifier lacking the carbonyl group of the central aromatic ring

The structures of SMN-C5 and risdiplam bound to the RNA helix highlight the important role of the carbonyl group on the central aromatic unit (22). The carbonyl group provides specificity by forming a direct hydrogen bond with the primary amine of the unpaired adenine at position  $-1$ . Unlike SMN-C5 and risdiplam, the *SMN2* splicing modifier called SMN-CX lacks this carbonyl group (Figure 3A) and binds to the RNA helix with a dissociation of  $70 \mu\text{M}$  (Table 1, 22). Like SMN-C5 and risdiplam, and despite the absence of the carbonyl group on the central aromatic ring, SMN-CX showed biological activity on the *SMN2* system (25). This finding was surprising and prompted us to verify whether SMN-CX binds to the RNA helix in the same way as SMN-C5 or risdiplam. Using NMR spectroscopy, we titrated the RNA helix with increasing amounts of SMN-CX and monitored the RNA resonances using 1D  $^1\text{H}$  and 2D  $^1\text{H}$ – $^1\text{H}$  TOCSY experiments (Fig-

**Table 1.** SMN2 splicing modifiers, biological activity and binding affinities for the RNA duplex

A <sub>-1</sub> splicing modifier	IC <sub>50</sub> [nM]	K <sub>D</sub> [μM]
 SMN-C5	30 nM <sup>a</sup>	$28 \pm 9 \mu\text{M}^f$
 SMN-CX	36 nM <sup>b</sup>	$74 \pm 17 \mu\text{M}^f$
 risdiplam	23 nM <sup>c</sup>	$12 \pm 1 \mu\text{M}^g$
 SMN-CY	29 nM <sup>d</sup>	$60 \pm 13 \mu\text{M}^f$
 branaplam	20 nM <sup>e</sup>	$18 \pm 1 \mu\text{M}^g$

<sup>a</sup>Determined in HEK293T cells transfected with pCI-SMN2 (22);

<sup>b</sup>Unpublished results.

<sup>c</sup>Determined in SMA type I fibroblasts (20).

<sup>d</sup>Determined in SMA type I fibroblasts (24).

<sup>e</sup>Determined by SMN ELISA (19).

<sup>f</sup>Determined by  $^{19}\text{F}$  NMR titrations (22).

<sup>g</sup>Determined in this study using  $^1\text{H}$  NMR titrations.

ure 3B–D). Upon adding SMN-CX to the RNA helix sample, we observed the characteristic line broadening of the imino proton of  $G_{+1}$  and changes in the chemical shifts of the signals of  $C_8$  and  $U_{+2}$  on the TOCSY spectrum. These data suggest that SMN-CX binds to the RNA in the major groove near the exon-intron junction, similar to SMN-C5 and risdiplam. To get a more detailed view of the interaction between SMN-CX and the RNA helix, we assigned the proton resonances of SMN-CX (Supplementary Table S1), identified 17 intermolecular NOEs between SMN-CX (Supplementary Figure S4 and Supplementary Table S2) and solved the structure of SMN-CX bound to the RNA helix (Table 2 and Figure 3E–G). Similarly than risdiplam, a strong line broadening of the SMN-CX aromatic resonances prevented the identification of intermolecular connections between the aromatic core of the small molecule and the RNA (Supplementary Figure S5). The structure reveals that SMN-CX binds to the RNA helix in the major groove at the exon-intron junction (Figure 3E). The positive charge on the piperazine forms an electrostatic interaction with the phosphate group of  $C_9$ , while the central aromatic ring inserts between  $C_8$  and  $C_9$  to position the nitrogen atom to accept a hydrogen from the amino group of the unpaired adenine (Figure 3E). Overall, the solution structure of SMN-CX bound to the RNA helix demonstrates that the carbonyl group on the central aromatic unit of SMN-C5 and risdiplam is not essential to target an  $A_{-1}$  bulged 5'-splice site and can be replaced by other hydrogen bond acceptors.



**Table 2.** NMR and refinement statistics

	RNA-risdiplam	RNA-SMNCX	RNA-SMN-CY	RNA-branaplam
<b>NMR distance and dihedral constraints</b>				
Distance restraints				
Total NOE	348	347	341	352
Intra-residue	255	254	246	259
Inter-residue	93	93	95	93
Sequential ( $ i - j  = 1$ )	87	87	89	87
Nonsequential ( $ i - j  > 1$ )	6	6	6	6
Intra-ligand	0	0	40	0
Hydrogen bonds	18	18	18	18
Ligand–nucleic acid intermolecular	20	22	17	18
Total dihedral angle restraints	170	167	146	167
<b>Structure statistics</b>				
Violations (mean and s.d.)				
Distance constraints (0.3 Å)	4.72 ± 0.87	0	0.20 ± 0.40	0
Dihedral angle constraints (>5°)	2.42 ± 1.22	10.60 ± 0.73	0	11.20 ± 1.12
Max. dihedral angle violation (°)	6.24 ± 0.77	11.44 ± 1.17	0.61 ± 0.26	10.96 ± 0.77
Max. distance constraint violation (Å)	0.34 ± 0.03	0.23 ± 0.01	0.28 ± 0.03	0.22 ± 0.01
Deviations from idealized geometry				
Bond lengths (Å)	0.0015 ± 0.0001	0.0014 ± 0.0001	0.0054 ± 0.0002	0.0014 ± 0.0001
Bond angles (°)	1.75 ± 0.01	1.73 ± 0.01	2.33 ± 0.01	1.76 ± 0.01
Average pairwise r.m.s. deviation (Å)				
Number of models	14	15	14	15
All atoms	0.54 ± 0.12	2.19 ± 0.49	2.35 ± 0.22	0.96 ± 0.80
All heavy atoms	0.35 ± 0.12	2.13 ± 0.44	2.30 ± 0.23	0.86 ± 0.81
All atoms binding site*	0.60 ± 0.13	0.35 ± 0.05	1.61 ± 0.39	0.56 ± 0.07
All heavy binding site*	0.38 ± 0.11	0.26 ± 0.07	1.59 ± 0.40	0.41 ± 0.04
<b>Data availability</b>				
PDB ID	8R62	8R8P	8CF2	8R63
BMRB ID	34878	34885	34784	34879

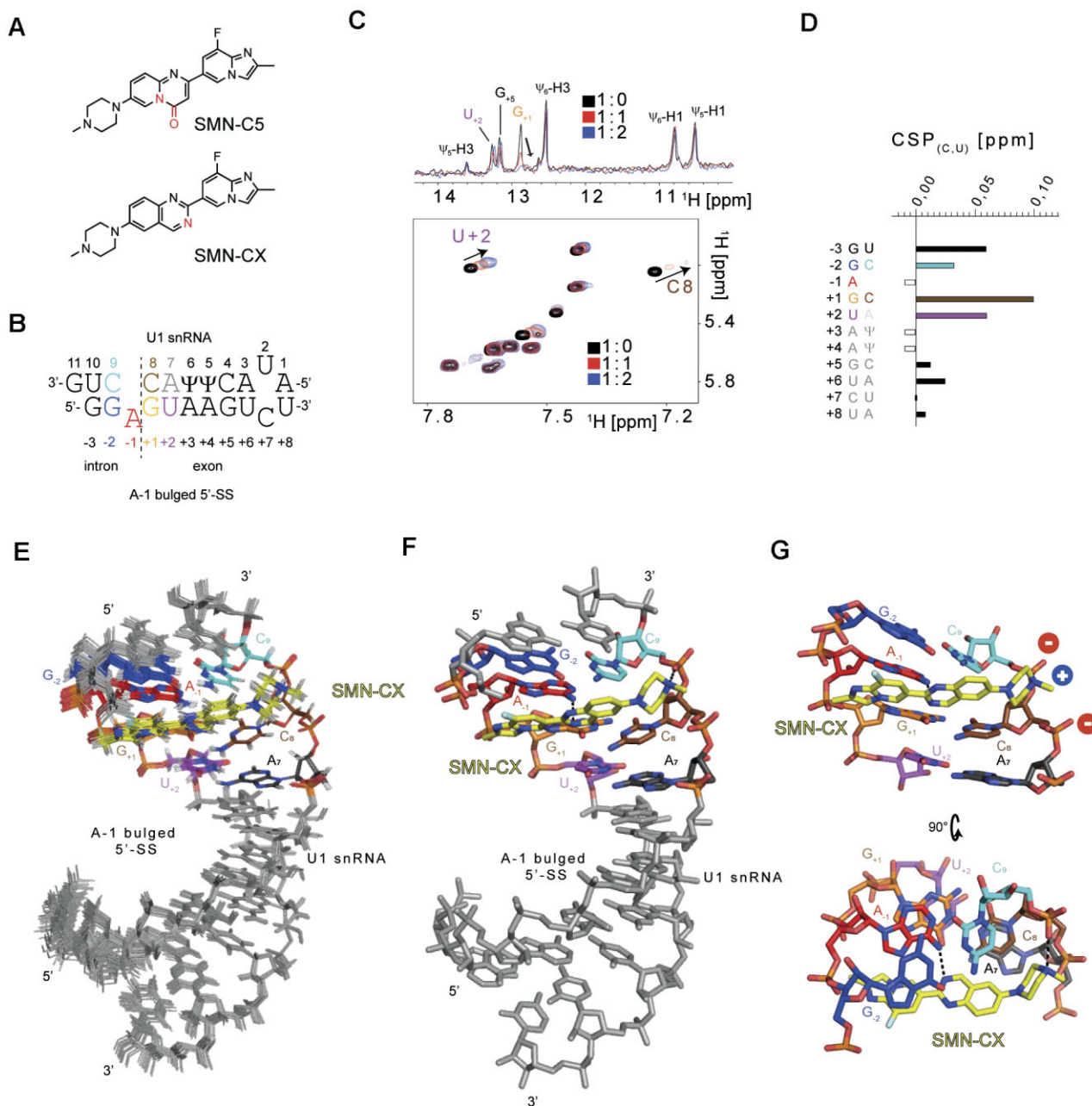
\* Pairwise r.m.s. deviation was calculated among all the refined structures on the region including nucleotides 4–10 and nucleotides –2 to +6 from the U1 snRNA and the 5'-SS and the small molecule.

### The size of the central aromatic unit can be reduced if the planarity is preserved

The clinical evaluation of RG7800 (a homolog of SMN-C5) revealed a long-term treatment safety issue. To address this problem, two approaches were taken: (i) optimization of RG7800, which led to the discovery and approval of risdiplam, and (ii) the search for other compounds with similar activity. This effort resulted in the identification, through rational drug design, of a new family of compounds with a benzamide core and a novel chemotype that corrects the splicing of *SMN2* exon 7 in a dose-dependent manner with an  $IC_{50}$  of 29 nM (Table 1,24). During the structure-activity relationship study, a fluorine atom was added to the phenyl ring to create SMN-CY, which has a drastic effect on the biological activity of the splicing modifier. Here, we assigned the  $^1H$  resonances of SMN-CY using a combination of homonuclear NMR experiments (Supplementary Figure S6 and Supplementary Table S1). The pyridyl analogue generates lone-pair repulsion between the pyridyl nitrogen and the carbonyl of the amide, as well as a favourable electrostatic interaction with the amide NH. In support of this, our NMR data showed that the amide proton of the molecule was strongly upfield shifted (Figure 4A). The presence of a weak interaction between the fluorine and the amide proton was validated experimentally by measuring a 2D  $^{15}N$ - $^1H$  HSQC and a 2D  $^1H$ - $^{19}F$  HOESY which confirmed the identity of the amide proton and its proximity to the fluorine. This interaction almost leads to a planar configuration of the carbonyl of the peptidyl-like bond and the benzamide. To understand how SMN-CY corrects *SMN2* splicing, we first determined whether SMN-CY interacts with the RNA helix (Figure 4B–D). Upon adding

SMN-CY to the RNA helix sample, we observed characteristic chemical shift perturbations of the  $U_{+2}$  and  $C_8$  resonances and perturbations on the signal corresponding to the imino proton of  $G_{+1}$ , suggesting that SMN-CY binds the same RNA pocket as SMN-C5 (Figure 4C). We also monitored changes on the RNA spectra using 2D  $^1H$ - $^{13}C$  HSQC experiments (Supplementary Figure S6). However, the perturbed RNA signals did not exhibit strong line broadening effects, as when the RNA helix was titrated with SMN-C5, risdiplam, or SMN-CX. This may be due to the reduction in size of the central aromatic ring and associated ring current effects. These data indicate that the size of the central aromatic unit can be reduced while still binding the RNA helix with a dissociation constant of 60  $\mu M$  (Table 1, 22) and achieving *SMN2* splicing correction (24). To gain structural insights into the interaction between SMN-CY and the RNA helix, we solved the structure of the RNA/small molecule complex in solution (Figure 4D–F). Using a mixing time of 200 ms, we identified 17 intermolecular NOEs between SMN-CY and the RNA helix that defined the intermolecular interface (Supplementary Figure S6). Both methyl groups of the molecule provided clear intermolecular NOEs that oriented the small molecule within the major groove (Supplementary Figure S7). On one side, Q28, which is the pseudo-atom representing the three magnetically equivalent methyl protons, was in close proximity to  $A_{-1}$  and  $U_{+2}$ , while the ethyl moiety on the piperazine was located closer to  $C_9$  and  $U_{10}$ . The solution structure of the RNA helix bound to SMN-CY was determined using NOE-derived distances, including 18 intermolecular distances (Table 2 and Supplementary Table S2). The 14 NMR structures of the NMR ensemble superimpose well on the SMN-CY bind-





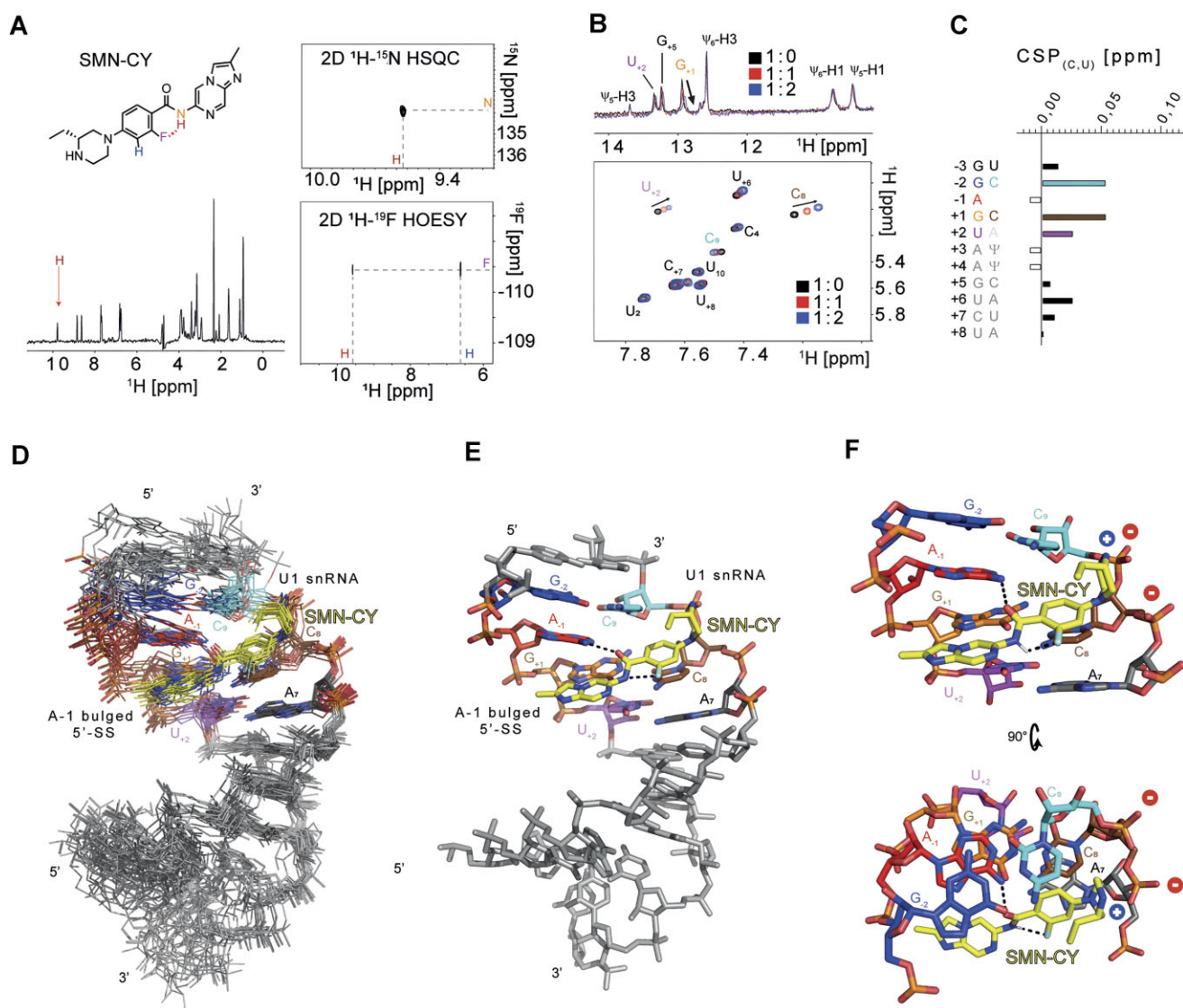
**Figure 3.** Structural basis for SMN-CX recognition of the RNA helix formed upon 5'-splice site recognition. **(A)** Planar structures of SMN-C5 and SMN-CX. The differences between the two splicing modifiers are highlighted in red. **(B)** Schematic representation of the RNA helix. **(C)** Overlay of the 1D  $^1\text{H}$  (imino region) and 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (H6-H5 region) spectra of the RNA recorded upon successive addition of SMN-CX. The spectra are coloured according to the ratio RNA helix : SMN-CX. **(D)** Plot of the pyrimidine (H5-H6) chemical shift perturbations (CSP) as a function of the positions on the 5'-splice site. White bars correspond to position without values (-1, +3 and +4). In position 7, the contributions of both pyrimidine have been added. **(E)** Overlay of the 15 solution structures of SMN-CX bound to the RNA helix. **(F)** Atomic representation of the lowest energy model of the NMR ensemble. **(G)** Closed up views of the intermolecular interface between SMN-CX and the RNA helix epitope. Intermolecular hydrogen bonds are shown as dashed lines.

ing site with an average pairwise root mean square deviation of  $1.61 \pm 0.39$  Å. The solution structure reveals that SMN-CY orients its piperazine positive partial charge near the phosphate group of  $C_9$  and inserts its benzamide between  $C_8$  and  $C_9$  aromatic rings. By interacting with the amide proton, the fluorine atom maintains planarity and fixes the orientation of the carbonyl group of the peptidyl bond placing it in an optimal position to form a direct hydrogen bond with the amino group of the unpaired  $A_1$ . While SMN-CY represents a different chemotype than SMN-C5, their modes of action are very

similar, and many common features were observed between both intermolecular interfaces, suggesting that these may be crucial for targeting the  $A_1$  bulged 5'-splice site epitope.

### Molecular basis for *HTT* splicing modifiers

Structural analysis of SMN-C/RNA complexes reveals that two different chemotypes can target a similar RNA epitope at the  $A_1$  5'-splice site. To increase our understanding of  $A_1$  5'-splice site targeting molecules, we decided to investigate a



**Figure 4.** Structural basis for SMN-CY recognition of the RNA helix formed upon 5'-splice site recognition. **(A)** Planar structures of SMN-CY and 1D  $^1\text{H}$  NMR spectra of SMN-CY recorded in 10%  $\text{D}_2\text{O}$ . The signal that was assigned to the amide proton is shown by an arrow. On the right are depicted a 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC showing the correlation between the amide proton and its covalently attached nitrogen and a 2D  $^1\text{H}$ - $^{19}\text{F}$  HOESY showing the proximity between the fluorine and the amide proton. **(B)** Overlay of the 1D  $^1\text{H}$  (imino region) and 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (H6-H5 region) spectra of the RNA recorded upon successive addition of SMN-CY. The spectra are coloured according to the ratio RNA helix : SMN-CY. **(C)** Plot of the pyrimidine (H5-H6) chemical shift perturbations (CSP) as a function of the positions on the 5'-splice site. White bars correspond to position without values (-1, +3 and +4). In position 7, the contributions of both pyrimidine has been added. **(D)** Overlay of 14 solution structures of SMN-CY bound to the RNA helix. **(E)** Atomic representation of the lowest energy model of the NMR ensemble. **(F)** Closed up views of the intermolecular interface between SMN-CX and the RNA helix epitope. Intermolecular hydrogen bonds are shown as dashed lines.

compound from a third chemotype, branaplam (also called HTT-C2). Originally developed as a SMN2 splicing modifier (19), clinical evaluation of branaplam was halted and it was recently repurposed as an efficient *HTT* splicing modifier (18,33). This compound promotes the inclusion of a PTC-containing pseudo exon in the toxic mRNA mHTT and is being investigated as a potential therapeutic solution for Huntington disease. Like the SMN-C compounds, branaplam targets an  $\text{A}_1$  bulged 5'-splice site and requires the presence of a purine-rich motif in the exon to achieve optimal biological activity (18). Branaplam is based on a different chemical scaffold than the SMN-C compounds, consisting of a pyridazine substituted with a tetramethylated piperidin cycle on one side and a cyclohexa-2,4-dien-1-one attached to a piperidinol.

From a theoretical perspective, the small molecule exists in a tautomeric equilibrium between two forms (Supplementary Figure S8). In tautomer 1, a hydroxyl group is present on the cyclohexane moiety and the pyridazine cycle is non-protonated. In tautomer 2, a carbonyl group is present on the cyclohexane and the pyridazine cycle is protonated. The proton is shared between both aromatic cycles, making them coplanar and similar to the central ring of chemotype 1 (SMN-C5, risdiplam, and SMN-CX). To better understand the tautomeric equilibrium, we conducted NMR spectroscopy experiments in deuterated DMSO to observe exchangeable protons and carbons (Supplementary Figure S8). We assigned the resonances of the small molecules (Supplementary Figure S8 and Supplementary Table S3) and used a 2D  $^1\text{H}$ - $^{13}\text{C}$  HMBC to ob-

serve scalar coupling-mediated correlation between the protons of the cyclohexane ring and the carbon holding the hydroxyl group (tautomer 1) or the carbonyl group (tautomer 2). On the HMBC spectrum, we observed a correlation between the protons of the cyclohexane ring and a quaternary carbon at 158.4 ppm, corresponding to the carbon holding the hydroxyl group (Supplementary Figure S8). We did not observe a correlation between the protons of the cyclohexane and a potential  $^{13}\text{C}$  signal around 200 ppm that could correspond to the carbonyl group. In deuterated DMSO, we only observed tautomer 1, indicating that the equilibrium is skewed towards tautomer 1 in these conditions. Unfortunately, similar experiments cannot be performed in aqueous solution or in the presence of RNA because we had no access to isotopically labelled small molecules. Therefore, we investigated how branaplam interacts with the RNA helix. To monitor the binding of branaplam on the RNA helix, we recorded 1D  $^1\text{H}$  and 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY spectra of the RNA helix upon the addition of branaplam (Figure 5A, B). We observed typical chemical shift perturbations of  $\text{C}_8$ ,  $\text{U}_{+2}$ , and the broadening of the  $\text{G}_{+1}$  imino proton, suggesting that branaplam targets the same RNA epitope as the SMN-C splicing modifiers (Figure 5C, D). The analysis of the CSP of the small molecule upon adding the RNA helix allowed us to determine a dissociation constant of  $18 \pm 1 \mu\text{M}$  (Table 1). As previously demonstrated in the context of *HTT* splicing correction, branaplam requires an adenine in position -1 to achieve optimal biological activity (18,32). To gain structural insights into the branaplam/RNA interaction, we solved the solution structure of branaplam bound to the RNA helix using 18 intermolecular NOEs (Supplementary Figures S9-S10 and Supplementary Table S2). The 15 structures of the NMR ensemble overlay with an all-atom RMSD of  $0.96 \pm 0.80 \text{ \AA}$  (Table 2). The solution structure of branaplam bound to the RNA helix showed that the tetramethylpiperidine interacts with the phosphate group of  $\text{C}_9$  of the U1 snRNA through a salt bridge and an hydrogen bond. At the other extremity, the 1,2-diazol moiety contacts the phosphodiester backbone of the 5'-splice site also via an hydrogen bond (Figure 5E-G). These two additional hydrogen bonds might explain the higher affinity of branaplam for the RNA over SMN-C5 (Table 1). As observed for other  $\text{A}_{-1}$  splicing modifiers, the central aromatic ring is inserted between  $\text{C}_8$  and  $\text{C}_9$ . By sharing its exchangeable proton with the adjacent nitrogen of the pyridazine, planarity is maintained allowing its oxygen atom to form a direct hydrogen bond with the amine group of the unpaired adenine. A similar model can be obtained when the calculation is performed with tautomer 2. Altogether, the structure explains how branaplam can correct the splicing of target exons using the 5'-splice site bulge repair mechanism. In conclusion, branaplam is a third chemotype that can correct the splicing of  $\text{A}_{-1}$  bulged 5'-splice sites by targeting the same RNA epitope as chemotypes 1 and 2.

## Discussion

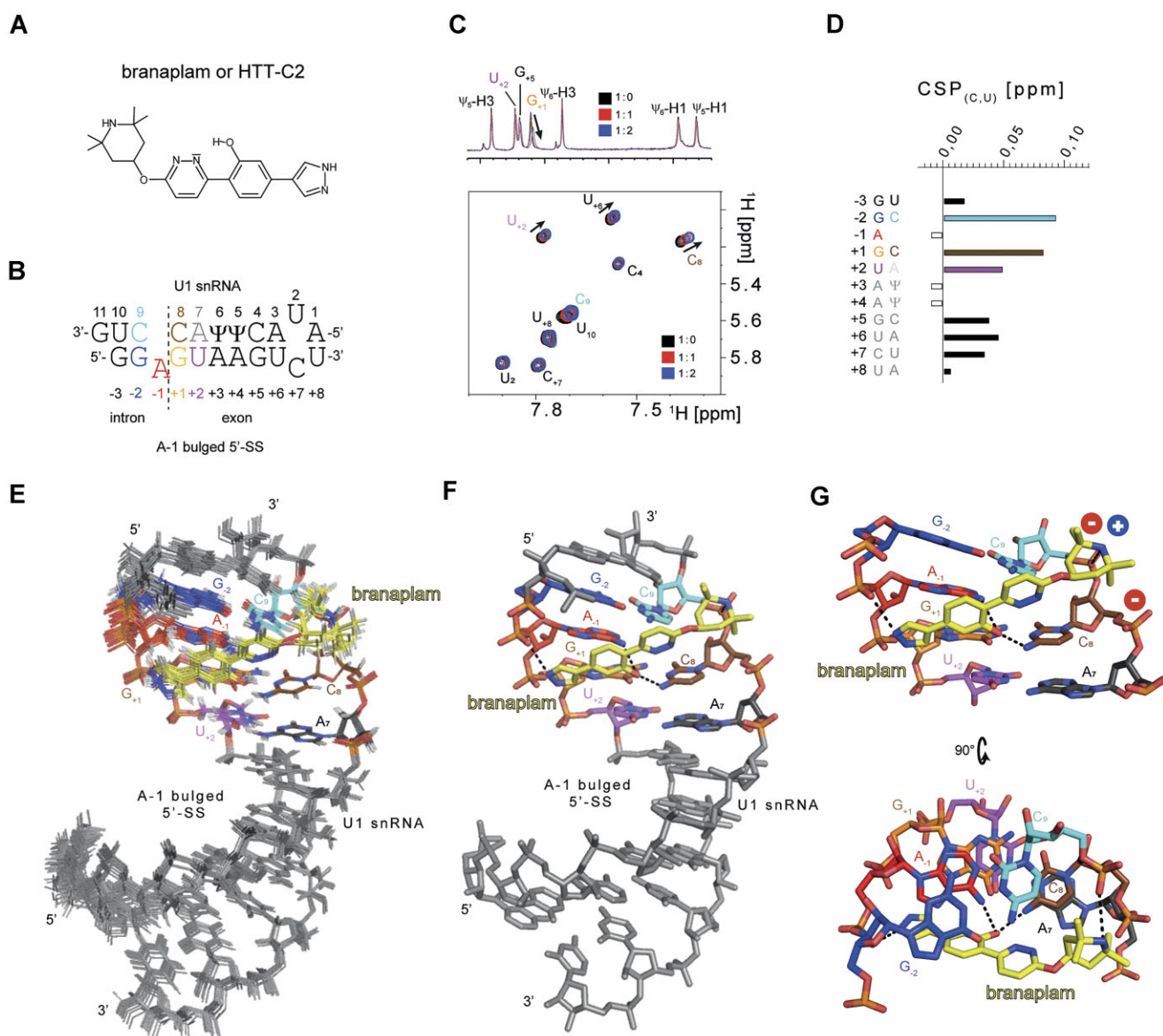
Single nucleotide polymorphisms (SNPs) at 5'-splice sites are implicated in many human diseases (9,11) and documented in the DBASS database (34). These diseases frequently arise from mutations at specific nucleotide positions, notably the invariant  $\text{G} + 1$  and at -1, +2, and + 5 positions. Our study focused on chemically diverse splicing modifiers that target  $\text{A}_{-1}$  bulged 5'-splice sites, representing a promising class of RNA thera-

peutics for precise splicing correction (2). The SMN2 splicing modifier, SMN-C5, served as a foundational model in this family, elucidating the '5'-splice site bulge repair' mechanism (22). However, we noted that various small molecule splicing modifiers with distinct chemical structures can induce similar biological effects (24,19,20). In this context, we solved the solution structures of four small molecules representing three different chemical classes (chemotypes) bound at the interface between the U1 snRNA and an  $\text{A}_{-1}$  bulged 5'-splice site.

The structures reveal that these molecules bind to the same RNA epitope, particularly near the unpaired adenosine at position -1. The atomic-level insights derived from the structures uncovered several shared features across the small molecules. They span the major groove of the RNA helix and comprise three distinct units: one interfacing with the U1 snRNA (Unit  $\text{R}_3$ ), a central aromatic unit (Unit  $\text{R}_2$ ), and one interacting with the 5'-splice site (Unit  $\text{R}_1$ , Figure 6). Unit  $\text{R}_3$  includes a piperazine or piperidine group, positively charged at physiological pH, forming a salt bridge with the phosphate backbone of U1 snRNA. Modifying this charge drastically affects biological activity (21). The importance of this unit is highlighted in the development of therapeutics like risdiplam, where a cyclopropyl substitution was used to modulate the basicity of the piperazine  $\text{R}_3$  unit (20). These moieties do not confer specificity but affinity may be enhanced as we observed an hydrogen bond to the phosphate backbone of U1 snRNA for both risdiplam and branaplam. Unit  $\text{R}_2$ , the central molecule component, specifically targets the unpaired adenine. It fits between the  $\text{C}_8$  and  $\text{C}_9$  C-H edges, a hydrophobic patch on the RNA epitope, and this explains why unit  $\text{R}_2$  must be aromatic and planar. Unit  $\text{R}_2$  contains a hydrogen bond acceptor, often a carbonyl group, that interacts with the adenine  $\text{A}_{-1}$ . Our models suggest the position of this group is crucial, as shown for various molecules like SMN-C5, risdiplam, SMN-CY, and branaplam. Interestingly, in SMN-CY and branaplam, weak interactions play a pivotal role in orienting the carbonyl group for optimal adenine interaction, a factor critical for biological activity. In SMN-CY, the carbonyl group is part of an amide bond linked to a fluoro-benzene ring. This structure is stabilised by a hydrogen bond between the fluorine on the benzene and the hydrogen of the amide, thus aligning the amide with the benzene ring. This alignment is crucial for positioning the carbonyl group to form a hydrogen bond with unpaired adenine. In branaplam, the small molecule undergoes tautomeric exchange. Here, the cyclohexa-2,4-dienol's hydroxyl group forms a hydrogen bond with a pyridazine nitrogen atom. In both cases, the result is a planarity of the central unit, securing the oxygen atom in a position enabling interaction with the unpaired adenine. Finally, unit  $\text{R}_1$  is often an aromatic ring with nitrogen atoms, but its role remains unclear. In our structures, unit  $\text{R}_1$  is close to the AGU motif and may provide additional stacking, polar interactions or water mediated contacts that are extremely difficult to observe using solution state NMR spectroscopy. Interestingly, in the structure with branaplam, the 1,2-diazol moiety (unit  $\text{R}_1$ ) is hydrogen bonded to the phosphodiester backbone of the 5'-splice site. This suggests that adding positively charged arms on unit  $\text{R}_3$  or  $\text{R}_1$  to contact the phosphate backbone by hydrogen bond appear possible to increase the affinity of the splicing modifier for the RNA helix and therefore its potency.

Overall, our study underscores the diversity and potential of splicing modifiers acting on  $\text{A}_{-1}$  bulged 5'-splice sites for rational drug design (Figure 6). Although our study tested a lim-





**Figure 5.** Structural basis for branaplaml recognition of the RNA helix formed upon 5'-splice site recognition. **(A)** Planar structures of branaplaml (or HTT-C2). **(B)** Schematic representation of the RNA helix. **(C)** Overlay of the 1D  $^1\text{H}$  (imino region) and 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (H6-H5 region) spectra of the RNA recorded upon successive addition of branaplaml. The spectra are coloured according to the ratio RNA helix : branaplaml. **(D)** Plot of the pyrimidine (H5-H6) chemical shift perturbations (CSP) as a function of the positions on the 5'-splice site. White bars correspond to position without values (-1, +3 and + 4). In position 7, the contributions of both pyrimidines have been added. **(E)** Overlay of the 15 solution structures of branaplaml bound to the RNA helix. **(F)** Atomic representation of the lowest energy model of the NMR ensemble. **(G)** Closed up views of the intermolecular interface between branaplaml and the RNA helix epitope. Intermolecular hydrogen bonds are shown as dashed lines.

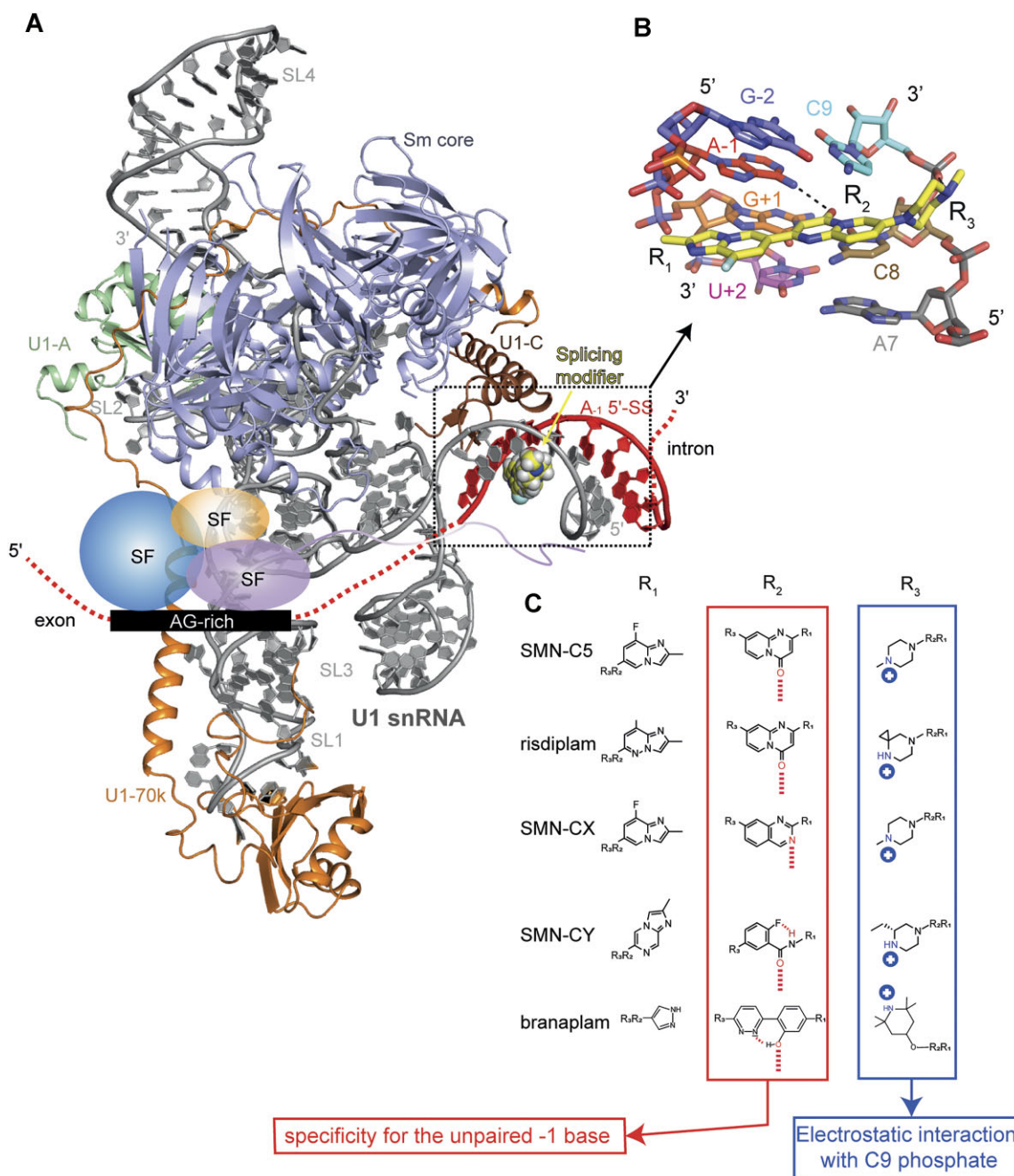
ited number of molecules, we carefully chose representatives from each relevant chemotype. By determining the experimental solution structures of the RNA helix when bound to each of these representative molecules, we successfully identified a unique binding mode applicable to all compounds tested. By elucidating their binding modes to the RNA helix, we offer valuable insights for designing new splicing modifiers. Despite these advancements, the high selectivity of *SMN2* or *HTT* splicing modifiers, which affect only a few exons out of the 200 000 A<sub>-1</sub> bulged 5'-splice sites in the genome (19), remains unexplained. These modifiers require both an A<sub>-1</sub> bulged site and a purine-rich element within the exon (21,19). While the role of A<sub>-1</sub> site role can be understood through the 5'-splice site bulge repair mechanism, the function of the purine-rich

element which is well known to recruit *trans*-splicing factors is less clear (22). It acts as a splicing enhancer and could potentially stabilise the interaction between the RNA helix and the small molecules, adding a layer of specificity toward these genes. Future design efforts should consider these factors for modifying base specificity at position -1 in small molecule splicing modifiers.

### Data availability

NMR chemical shifts and atomic coordinates for the four reported solution structures (RNA bound to risdiplaml, *SMN-CX*, *SMN-CY* and branaplaml) have been deposited with the BioMagResBank (BMRB) under the accession code 34878,





**Figure 6.** Chemical diversity of splicing modifiers acting on A<sub>1</sub> bulged 5'-splice sites. **(A)** Model of U1 snRNP bound to an A<sub>1</sub> bulged 5'-splice site in presence of SMN-C5. The exon position was extrapolated in order to depict the AG-rich motif and the splicing factors that are recruited. **(B)** Atomic details of the interface between the RNA helix epitope and SMN-C5. The splicing modifier has been split in three functional parts: R<sub>1</sub>, R<sub>2</sub> and R<sub>3</sub>. **(C)** Functional analysis of the composition and role of the five splicing modifiers acting on A<sub>1</sub> bulged 5'-splice sites.

34885, 34784, 34879 and the Protein Data bank under accession number 8R62, 8R8P, 8CF2, 8R63.

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We would like to thank the NMR platforms of ETH Zurich and of the European Institute of Chemistry and Biology (IECB), Roche and Skyhawk Therapeutics for providing the

compounds and Dr Julien Boudet (Skyhawk Therapeutics) for helpful discussion on risdiplam and branaplam.

## Funding

INSERM and the INSERM transfer office [COPOC2021 MAT-API-00785-A-01 to S.C.]; Federal Council of La Ligue contre le cancer [APPARN2021.LCC/SeC to S.C.]; Fédération de la Recherche Médicale [AJE202310017978 to S.C.]; FHFA acknowledges financial support from the NCCR RNA and Disease [51NF40-182880] and from Skyhawk Therapeutics. The open access publication charge for this paper has

been waived by Oxford University Press – NAR Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.

## Conflict of interest statement

None declared.

## References

- Scotti, M.M. and Swanson, M.S. (2016) RNA mis-splicing in disease. *Nat. Rev. Genet.*, **17**, 19–32.
- Malard, F., Mackereth, C.D. and Campagne, S. (2022) Principles and correction of 5'-splice site selection. *RNA Biol.*, **19**, 943–960.
- Roca, X., Krainer, A.R. and Eperon, I.C. (2013) Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev.*, **27**, 129–144.
- Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983) The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell*, **33**, 509–518.
- Roca, X., Sachidanandam, R. and Krainer, A.R. (2005) Determinants of the inherent strength of human 5' splice sites. *RNA*, **11**, 683–698.
- Wong, M.S., Kinney, J.B. and Krainer, A.R. (2018) Quantitative activity profile and context dependence of all human 5' splice sites. *Mol. Cell*, **71**, 1012–1026.
- Singh, N.N. and Singh, R.N. (2019) How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1862**, 194403.
- Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291–336.
- Buratti, E., Chivers, M., Kráľovičová, J., Romano, M., Baralle, M., Krainer, A.R. and Vořechovský, I. (2007) Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res.*, **35**, 4250–4263.
- Roca, X., Akerman, M., Gaus, H., Berdeja, A., Bennett, C.F. and Krainer, A.R. (2012) Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev.*, **26**, 1098–1109.
- Roca, X., Olson, A.J., Rao, A.R., Enerly, E., Kristensen, V.N., Børresen-Dale, A.-L., Andresen, B.S., Krainer, A.R. and Sachidanandam, R. (2008) Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. *Genome Res.*, **18**, 77–87.
- Michaels, W.E., Pena-Rasgado, C., Kotaria, R., Bridges, R.J. and Hastings, M.L. (2022) Open reading frame correction using splice-switching antisense oligonucleotides for the treatment of cystic fibrosis. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2114886119.
- Nakamura, A. and Takeda, S. (2011) Exon-skipping therapy for Duchenne muscular dystrophy. *Lancet North Am. Ed.*, **378**, 546–547.
- Kim, J., Hu, C., Moufawad El Achkar, C., Black, L.E., Douville, J., Larson, A., Pendergast, M.K., Goldkind, S.F., Lee, E.A., Kuniholm, A., et al. (2019) Patient-customized oligonucleotide therapy for a rare genetic disease. *N. Engl. J. Med.*, **381**, 1644–1652.
- Hua, Y., Sahashi, K., Rigo, F., Hung, G., Horev, G., Bennett, C.F. and Krainer, A.R. (2011) Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature*, **478**, 123–126.
- Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K.Y., Karp, G.M., Qi, H., Woll, M.G., et al. (2014) SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science*, **345**, 688–693.
- Crooke, S.T., Baker, B.F., Crooke, R.M. and Liang, X. (2021) Antisense technology: an overview and prospectus. *Nat. Rev. Drug Discov.*, **20**, 427–453.
- Bhattacharyya, A., Trotta, C.R., Narasimhan, J., Wiedinger, K.J., Li, W., Effenberger, K.A., Woll, M.G., Jani, M.B., Risher, N., Yeh, S., et al. (2021) Small molecule splicing modifiers with systemic HTT-lowering activity. *Nat. Commun.*, **12**, 7299.
- Palacino, J., Swalley, S.E., Song, C., Cheung, A.K., Shu, L., Zhang, X., van Hoesear, M., Shin, Y., Chin, D.N., Keller, C.G., et al. (2015) SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat. Chem. Biol.*, **11**, 511–517.
- Ratni, H., Ebeling, M., Baird, J., Bendels, S., Bylund, J., Chen, K.S., Denk, N., Feng, Z., Green, L., Guerard, M., et al. (2018) Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J. Med. Chem.*, **61**, 6501–6517.
- Sivaramakrishnan, M., McCarthy, K.D., Campagne, S., Huber, S., Meier, S., Augustin, A., Heckel, T., Meistermann, H., Hug, M.N., Birrer, P., et al. (2017) Binding to SMN2 pre-mRNA-protein complex elicits specificity for small molecule splicing modifiers. *Nat. Commun.*, **8**, 1476.
- Campagne, S., Boigner, S., Rüdiger, S., Moursy, A., Gillioz, L., Knörlein, A., Hall, J., Ratni, H., Cléry, A. and Allain, F.H.-T. (2019) Structural basis of a small molecule targeting RNA for a specific splicing correction. *Nat. Chem. Biol.*, **15**, 1191–1198.
- Ratni, H., Scalco, R.S. and Stephan, A.H. (2021) Risdiplam, the First Approved Small Molecule Splicing Modifier Drug as a Blueprint for Future Transformative Medicines. *ACS Med. Chem. Lett.*, **12**, 874–877.
- Pinard, E., Green, L., Reutlinger, M., Weetall, M., Naryshkin, N.A., Baird, J., Chen, K.S., Paushkin, S.V., Metzger, F. and Ratni, H. (2017) Discovery of a novel class of survival motor neuron 2 splicing modifiers for the treatment of spinal muscular atrophy. *J. Med. Chem.*, **60**, 4444–4457.
- Ratni, H., Karp, G.M., Weetall, M., Naryshkin, N.A., Paushkin, S.V., Chen, K.S., McCarthy, K.D., Qi, H., Turpoff, A., Woll, M.G., et al. (2016) Specific correction of alternative survival motor neuron 2 splicing by small molecules: discovery of a potential novel medicine to treat spinal muscular atrophy. *J. Med. Chem.*, **59**, 6086–6100.
- Van Rossum, G. and Drake, F.L. (2009) In: *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.
- Keller, R. (2004) The computer-aided resonance assignment tutorial. In: *Cantina Verlag*.
- Güntert, P. and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR*, **62**, 453–471.
- Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B. and Woods, R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272.
- McKinney, W. (2010) Data structures for statistical computing in python. In: *Proceedings of the 9th Python in Science Conference*. Austin, pp. 56–61.
- Ishigami, Y., Wong, M.S., Martí-Gómez, C., Ayaz, A., Kooshkbaghi, M., Hanson, S., McCandlish, D.M., Krainer, A.R. and Kinney, J.B. (2022) Specificity, synergy, and mechanisms of splice-modifying drugs. *Nat. Commun.*, **15**, 1880.
- Keller, C.G., Shin, Y., Monteys, A.M., Renaud, N., Beibel, M., Teider, N., Peters, T., Faller, T., St-Cyr, S., Knehr, J., et al. (2022) An orally available, brain penetrant, small molecule lowers huntingtin levels by enhancing pseudoexon inclusion. *Nat. Commun.*, **13**, 1150.
- Buratti, E., Chivers, M., Hwang, G. and Vorechovsky, I. (2011) DBASS3 and DBASS5: databases of aberrant 3'- and 5'-splice sites. *Nucleic Acids Res.*, **39**, D86–D91.

Received: April 28, 2023. Revised: December 7, 2023. Editorial Decision: February 23, 2024. Accepted: March 7, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)