

Identification of Ligands for the Tau Exon 10 Splicing Regulatory Element RNA Using Dynamic Combinatorial Chemistry

Paula López-Senín,^[a] Irene Gómez-Pinto,^[b] Anna Grandas,^[a] and Vicente Marchán*^[a]

[a] P. López-Senín, Prof. A. Grandas, Dr. V. Marchán
Departament de Química Orgànica and IBUB
Universitat de Barcelona
Martí i Franquès 1-11, E-08028 Barcelona (Spain)
Fax: + 34 93 339 7878
E-mail: vmarchan@ub.edu

[b] Dr. I. Gómez-Pinto
Instituto de Química-Física Rocasolano
CSIC
Serrano 119, E-28006 Madrid (Spain)

Abstract: We describe the use of dynamic combinatorial chemistry (DCC) to identify ligands for the stem-loop structure located at the exon 10-5'-intron junction of Tau pre-mRNA, which is involved in the onset of several tauopathies including frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17). A series of ligands that combine the small aminoglycoside neamine and heteroaromatic moieties (azaquinolone and two acridines) have been identified using DCC. These compounds effectively bind the stem-loop RNA target ($EC_{50} = 2\text{-}58\ \mu\text{M}$), as determined by fluorescence titration experiments. Importantly, most of them are able to stabilize both wild-type and the +3 and +14 mutated sequences associated with the development of FTDP-17 without producing a significant change in the overall structure of the RNA (CD spectroscopy), which is a key issue for recognition by the splicing regulatory machinery. A good correlation has been found between the ligands' affinity for the target and their capacity to stabilize the RNA secondary structure.

Introduction

RNA plays an essential role in many cellular processes, from the regulation of gene expression to protein synthesis, as well as in the progression of viral diseases. This macromolecule is able to adopt more complex three-dimensional structures than DNA is; they are similar to those observed in proteins, and their tertiary structure is as yet difficult to predict. Hence, discovering selective and, more importantly, specific ligands for this macromolecule represents a challenge for medicinal and bioorganic chemistry and therapy.^[1] However, most of the known RNA-binding ligands can each bind to several different RNA targets of unrelated sequence and structure,^[2] which compromises their biological applications. We hypothesize that higher affinity and better specificity could be achieved by combining two or more small molecules with different affinities and capacities to recognize some of the secondary structural motifs present in a particular RNA target.^[3] This fragment-based lead-discovery searching approach,^[4] in combination with dynamic combinatorial chemistry (DCC),^[5] offers new perspectives and opportunities for identifying RNA ligands, as well as for improving our understanding of RNA recognition mechanisms.

Over the past few years, DCC has demonstrated its potential as a tool for discovering ligands for biomolecules, mainly because both the library synthesis and the affinity screening steps are carried out simultaneously in a single process.^[5] In the field of nucleic acids DCC has been used to select a copper(II) coordination complex which binds to an RNA hairpin,^[6] as well as ligands that bind to DNA duplex or quadruplex structures,^[7] and also to stabilize oligonucleotide complexes by covalent linkage of small molecules.^[8] More recently, resin-bound DCC has been used to identify ligands for RNA structures involved in HIV-1 replication and in myotonic dystrophy type 1.^[9] Typically, most ligands for DNA and RNA targets that have been identified using DCC have high to medium affinities ($\sim 2\text{-}50\ \mu\text{M}$). In the study reported here, we use DCC to identify ligands for the Tau exon 10 splicing regulatory element RNA involved in frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17).^[10] Tau is a microtubule-associated protein that is required for the polymerization and stability of microtubules as well as for axonal transport in neurons. In a normal adult human brain, alternative splicing of exons 2, 3 and 10 produces six tau isoforms with either three (3R) or four (4R) repeat domains, with a 4R/3R ratio of approximately 1.^[11] However, mutations found in the tau gene of FTDP-17 patients, especially at the exon 10-5'-intron junction, alter pre-mRNA splicing and result in an increase in the inclusion of exon 10. The normal one-to-one isoform balance is thus perturbed, with the 4R isoform being overproduced in most cases. This alters the microtubule function and consequently leads to the development of the tauopathy.^[12] The stem-loop structure located at the exon 10-5'-intron

junction seems to be an important regulatory element in pre-mRNA splicing. Both *in vitro* and *in vivo* experiments have shown that the extent of exon 10 inclusion is inversely proportional to the stability of this structure; a stability which is considerably diminished by mutations (Figure 1a).^[13] Thus, small molecules that selectively bind to and stabilize this stem-loop structure, and in particular the mutated variants found in FTDP-17 patients, would allow the physiological balance between Tau isoforms to be restored and, consequently, the disease to be treated.^[14] Importantly, most of the ligands identified in this study using the DCC approach, in addition to binding with high to medium affinity and stabilizing wild-type (wt) RNA, are capable of stabilizing the two mutated sequences (+3 and +14) that cause the highest destabilization of the stem-loop structure.

Results and Discussion

Construction of the Dynamic Combinatorial Library. First, we synthesized several building blocks, including known RNA binding ligands together with some other compounds with a potential to bind to a particular RNA motif, that could be assembled modularly into new lead ligands for Tau stem-loop targets (the structures are shown in Figure 1b).^[15] Thiol derivatization was chosen because disulfide exchange is relatively fast in aqueous solutions of pH near 7-8 and fully compatible with RNA. In addition, this approach increases the diversity of the dynamic combinatorial library (DCL) in comparison with non-symmetrical bonds, since formation of disulfide bridges can afford both homo- and heterodimers.^[16]

Figure 1

Aminoglycoside antibiotics are possibly the most studied RNA ligands. They are known to be able to discriminate A-type from B-type duplexes and to have relatively high affinity for RNA structures.^[2b,17] However, since natural aminoglycosides do not exhibit inherent specificity for biological RNA sequences, several groups have synthesized aminoglycoside analogues to tune specificity without compromising affinity. In order to reduce the risk of non-specific electrostatic interactions, we decided to select a neamine derivative, **Nea**, as a building block. This small aminoglycoside incorporates rings I and II of neomycin B and still presents a high charge density. In addition, two acridine derivatives (**Acr1** and **Acr2**) that have the capacity either to intercalate in duplex regions or to stack with unpaired nucleobases were also included in the library of monomers.^[18] Although acridines are also promiscuous molecules, they have recently been used in the development of RNA ligands with high binding affinity and good specificity.^[19] We also selected a third heteroaromatic compound, an azaquinolone derivative, **Azq**, which could recognize the bulged adenine of the Tau stem-loop and provide higher specificity than the aminoglycosides or intercalators through the formation of complementary hydrogen bond acceptor-donors.^[20] Finally, two cationic tripeptides containing an aromatic residue, **TyrP** and **TrpP**, were assembled using solid-phase procedures, and included in the library of monomers.

Selection of Ligands using Dynamic Combinatorial Chemistry. In order to facilitate the identification of RNA-interacting molecules, DCC experiments were carried out in solution-phase in the presence of the biotinylated RNA target (Scheme 1). Previously, UV-monitored thermal denaturation experiments and circular dichroism had showed that the 5'-end biotin derivatization affects neither the overall stem-loop structure nor its stability (see the Supporting Information). In addition, the unique monophasic curve obtained at higher concentration (25 μM) was further evidence for the presence of a single intramolecular species (no transitions were detected at low temperatures) and the claim that no supramolecular

aggregates are formed during DCC experiments. After incubation of the thiol-derivatized monomers and the target, the use of streptavidin anchored to magnetic beads allowed the RNA and the interacting ligands to be separated easily from the other members of the DCL.^[7] After denaturation of RNA at 90°C to release any bound ligands, MS HPLC was used to identify and quantify the compounds. In all cases, a control experiment in the absence of the biotinylated RNA was performed in parallel to determine the effect of the target on the amplification of the ligands.

Scheme 1

The first DCC experiment was performed with wt RNA (25 μM) and the **Nea**, **Acr1**, **Azq** and **TyrP** monomers (100 μM each) in 50 mM Tris-HCl buffer, pH 7.7, 100 mM NaCl and 0.1 mM EDTA at room temperature, under an air atmosphere and without stirring. Comparison with the control experiment (in the absence of the RNA) indicated a clear amplification of two disulfide heterodimers, **Acr1-Nea** (~100%) and **Azq-Nea** (~75%).^[21] Equilibrium was reached within 48 h in both cases, and the composition of the mixture remained unchanged over the next 2 days (see the Supporting Information).^[22] Furthermore, when the **Acr1** monomer was replaced by its disulfide dimer, **Acr1-Acr1** (50 μM), a similar distribution of products was produced, both in the presence and in the absence of the RNA. This demonstrated that a true thermodynamic equilibrium had been reached and that the air-mediated oxidation process is sufficiently slow to allow for equilibration of the different species.

In subsequent DCC experiments we included two additional monomers, **Acr2** and **TrpP**, both to increase the diversity of the DCL (21 theoretical compounds may be formed) and to study the competition between the two acridines and the two peptides. As shown in Figure 2, a small amplification of **Acr1-Nea** (60%) and **Azq-Nea** (40%) was again observed. However, two new ligands were amplified in much higher proportion, **Acr2-Nea** (3600%) and **Acr2-Acr2** (2300%), and to a lesser extent **TyrP-Acr2** (140%), **TrpP-Acr2** (190%) and **Azq-Acr2** (200%).^[23] This result indicates that ligands incorporating the **Acr2** moiety have a higher affinity for the target RNA than those containing **Acr1** or **Azq**. Importantly, the same species were amplified when the DCC experiment was carried out in the presence of the biotinylated +3-mutated sequence instead of wt RNA (see the Supporting Information). This suggests that their binding site is not close to the mismatched base pair in the +3-mutated sequence.

Figure 2

A crucial parameter in our fragment-based approach is the selection of the appropriate distance between the building blocks to allow the simultaneous recognition of two or more structural motifs in a particular target RNA. This parameter may be used to tune both the binding affinity and specificity of the ligand, as found by Tor *et al.* in some neomycin-acridine conjugates.^[24] With these premises in mind, we decided to carry out a new DCC experiment with five monomers: **Acr2**, **TyrP**, **Azq**, **Nea** and a neamine derivative with a shorter spacer between the aminoglycoside core and the thiol group, **Nea2** (Figure 1b). As expected, the **Azq-Nea**, **Acr2-Nea** and **Acr2-Acr2** ligands were amplified in a relative ratio similar to that obtained in the previous experiments (see the Supporting Information). Interestingly, the amplification of the neamine-acridine ligand containing the shorter linker, **Acr2-Nea2**, was slightly higher than that of **Acr2-Nea**, suggesting a higher affinity.

Binding Affinities and Specificities of the Selected Ligands. In most cases, the selected dimeric ligands were synthesized on a larger scale by replacing the disulfide linkage by the

non-reversible thioether isoster (CH₂-S). Then, quantitative binding studies were carried out to determine their binding affinities. Titration experiments were performed by monitoring the quenching of the fluorescence intensity of the fluoresceine-labelled wt RNA as a function of the increasing the concentration of the ligand.^[25] In all cases, a characteristic dose-dependent, saturatable change in fluorescence was observed, which can be attributed to a slight conformational change in the RNA target induced by the ligand upon complexation (see below the CD results). EC₅₀ values (the effective ligand concentration required for 50 % RNA response) of 28.6 μM and 58 μM were obtained for **Acr1-Nea** and **Azq-Nea**, respectively, by fitting the data in a dose-response curve. As shown in Table 1, the overall results are consistent with DCC amplification data since the highest binding affinities are exhibited by the most amplified ligands; this confirms the power of this methodology for identifying ligands for RNA targets. Indeed, ligands containing **Acr2** showed higher binding affinities than those containing **Azq** or **Acr1** (5.9 μM and 2.9 μM for **Acr2-Nea** and **Acr2-Acr2**, respectively). Interestingly, the binding was observed to be approximately 3 times stronger for the ligand containing the shorter spacer between acridine and neamine, **Acr2-Nea2**.

Table 1

These results prompted us to evaluate the specificity of the ligands. Fluorescence binding assays were repeated in the presence of a 30-fold nucleotide excess of a commercially available tRNA (Table 1), which is a relevant competitor since it contains a mixture of both pre- and mature tRNAs.^[9b,24b,25b,26] Interestingly, the specificity (EC₅₀ in the presence of the competitor/EC₅₀ in the absence of the competitor) of the ligands was shown to be highly dependent on the nature of the acridine building block (the specificity of **Azq-Nea** could not be determined because the emission maximum of fluoresceine varied more than 3 nm at high ligand concentrations). In the presence of the competitor, the EC₅₀ values of **Acr2-Nea** and **Acr2-Acr2** for wt RNA were increased by 11-fold and 14-fold, respectively, whereas that of **Acr1-Nea** was only increased by 4-fold. Hence, replacement of the **Acr2** moiety by **Acr1** in the acridine-neamine ligands increases their specificity. On the other hand, as shown in Table 1, the specificity ratio for **Acr2-Nea2** was the highest. This result indicates that in the **Acr2**-neamine ligands a longer spacer confers higher specificity. This trend is different from that previously reported for some acridine-neomycin conjugates with another target, the HIV-1 RRE RNA.^[24b]

All these experiments with Tau RNA provide three important conclusions, some of which have also been previously observed in different RNA targets. First, the covalent attachment of an heteroaromatic moiety, either azaquinolone or acridine causes a substantial increase in the binding affinity of the aminoglycoside scaffold (see the Neamine entry in Table 1). Second, ligands containing **Acr2** have higher binding affinities than those containing **Acr1** and **Azq**, and it seems that a shorter spacer is preferred in the acridine-neamine ligands with Tau RNA. Third, there is an inverse correlation between affinity and specificity since the higher the affinity, the less specific the ligand seems to be.

Effect of the Selected Ligands on the Thermal Stability of Tau RNA Targets. As previously stated, besides binding with good affinity and specificity to the Tau RNA targets, effective ligands for these targets must stabilize them, in particular the mutated sequences. So, our next objective was to evaluate the capacity, if any, of the most amplified ligands to produce stability. The impact of the ligands on the thermal stability of the stem-loop structures (wt and mutated) was estimated by UV-monitored melting experiments (Table 2). In the presence of **Azq-Nea** or **Acr1-Nea** a slight increase in the T_m value was observed in all RNAs (ΔT_m ≈ +1°C), with the exception of **Acr1-Nea** and the +14 mutant, where no

significant shift occurred. This small stabilization, similar to that obtained with neamine alone, could suggest that the **Azq** or the **Acr1** fragments do not interact with RNA. However, binding affinities (Table 1) and spectroscopic data (see below) have demonstrated the active participation of the heteroaromatic moiety in the RNA binding. Replacement of **Acr1** by **Acr2** in the acridine-neamine ligand (**Acr2-Nea**) caused a greater increase in T_m values for both the mutated sequences ($\Delta T_m = +2.8^\circ\text{C}$ and $+2^\circ\text{C}$ for +3 and +14 mutants, respectively) and wt RNA ($\Delta T_m = +2.1^\circ\text{C}$). To our surprise, the T_m of the +3 mutant was clearly increased ($\Delta T_m = +5.7^\circ\text{C}$; Table 2) in the presence of the ligand with the shortest spacer, **Acr2-Nea2**. The effect of this ligand on wt RNA and the +14 mutant was smaller ($\Delta T_m = +2.4^\circ\text{C}$ and 3.2°C , respectively) but still higher than that induced by the ligand with the longest spacer, **Acr2-Nea**. It is particularly relevant that the degree of stabilization of the +3 mutant with **Acr2-Nea2** is higher than that produced by neomycin B alone, and of the same order as that produced by Mitoxantrone, a ligand for Tau RNA recently identified in a high-throughput fluorescent binding assay.^[14a,14b,27]

All together, these results show a direct correlation between the T_m of the RNA-ligand complex, the level of amplification observed in the DCC experiments and the binding affinities determined by fluorimetry, which is consistent with previous observations in different systems.^[7,8] The most amplified ligands (**Acr2-Nea** and **Acr2-Nea2**) have higher binding affinities and produce a higher degree of stabilization than the least amplified ligands (**Azq-Nea** and **Acr1-Nea**). No stabilization was observed upon incubation with **Acr1-Acr1** since this compound was not templated by the RNA. As expected, a significant increase in T_m also occurred with **Acr2-Acr2**. These results indicate that the 9-(alkylamino)acridine is more stabilizing than that bearing the 9-carboxamide group. Both the positive charge of the former ($\text{pK}_a \approx 5$ for **Acr1** vs $\text{pK}_a \approx 8$ for **Acr2**) and its ability to establish additional stabilizing interactions may account for this stabilizing effect and the high binding affinity.^[28]

Table 2

Spectroscopic Studies of the Complexes formed between Tau RNA and the Selected Ligands. Considering the fact that all of the best ligands included a heteroaromatic moiety, we wanted to gain some insight into its role upon complexation with RNA. When **Acr1-Nea**, **Acr2-Nea** and **Acr2-Nea2** were titrated with increasing amounts of wt RNA, a strong hypochromism was observed in their UV-Vis spectra (30-40% at a 1:1 ratio; Figure 3). In addition, the 423- and 444-nm bands of the free **Acr2-Nea** and **Acr2-Nea2** ligands exhibited a 7- and 9-nm red-shift, respectively. This shift to higher wavelengths was smaller for **Acr1-Nea** (2 nm for the 360-nm band), which is consistent with previously reported results.^[29] Although a similar hypochromism was observed in the case of **Azq-Nea**, the interference of the RNA absorbance near the azaquinolone band prevented the titration experiment from being completed (see the Supporting Information). Similar bathochromic effects were observed in fluorescence titration experiments, including reduction in the fluorescence intensity of the acridine moieties together with changes in the spectral profiles (see the Supporting Information). These effects are commonly observed when planar heteroaromatic compounds bind to nucleic acids via an intercalative mechanism, or at least when stacking occurs with base-pair nucleobases.

Figure 3

Finally, we used CD spectroscopy to check whether these ligands might substantially change the overall conformation of the target RNA structures upon binding. This is a key issue for the recognition of the ligand-stabilized stem-loop by the splicing regulatory machinery. Although

+3 and +14 mutations decrease the thermal stability of the stem-loop RNAs, they exhibit CD spectra that are very similar to that of wt RNA. This indicates that mutations do not produce a significant change in the structure (see the Supporting Information). When CD spectra were registered in the presence of the selected ligands, minimal alterations were observed (a decrease in the ellipticity of the positive band at 234 nm, and a concomitant increase in both the intensity and the wavelength shift of the positive band around 267 nm). These results indicate that the overall stem-loop structure in the Tau RNA target is maintained in all cases. These data, together with the UV-Vis and fluorescence spectroscopy results, suggest that the heteroaromatic moiety of the ligands might intercalate or stack around the bulged adenine, as described by Varani *et al.* for the binding of Mitoxantrone to the Tau-RNA.^[14c]

Conclusion

In conclusion, we have shown the usefulness of the dynamic combinatorial chemistry approach to identify ligands that bind with high to medium affinity ($EC_{50} = 2\text{-}58\ \mu\text{M}$), as determined by fluorimetry, to the Tau exon 10 splicing regulatory element RNA. This confirms the potential of this methodology for developing new RNA-binding compounds. Importantly, most of the selected compounds, which combine a small aminoglycoside and a heteroaromatic moiety, are able to stabilize the +3 and +14 mutated sequences associated with the development of FTDP-17 as well as the wt RNA. Furthermore, they bind the RNA target without producing a significant change in the overall structure of the stem-loop (CD), and some of them could have a preferred binding site near the bulged adenine of the stem-loop structure as suggested by the biophysical experiments. The overall results show a direct correlation between the affinity of the ligands and their stabilizing properties. It is important to point out that the most specific ligands, such as **Acr1-Nea**, have a moderate affinity for the RNA target whereas those with high affinities (compounds containing **Acr2**) are less specific as inferred from the competition studies performed in the presence of an excess of tRNA. The identification of ligands that stabilize the mutated Tau stem-loop may open the way for the generation of drugs that restore the physiological balance of Tau isoforms, thus allowing for the treatment of frontotemporal dementias such as FTDP-17 as well as other neurodegenerative tauopathies including Alzheimer's disease. Moreover, these ligands may become valuable tools for studying and understanding the alternative splicing of exon 10 and, for instance, how pre-mRNA secondary structures influence the regulation of alternative splicing.^[12b,30]

Efforts are underway to improve the affinity and especially the specificity of the most promising ligands, as well as to elucidate the structure by NMR of their complexes with both wt RNA and mutated sequences, using labelled oligoribonucleotides.

Experimental Section

General procedures for DCC experiments. In general, DCC experiments were repeated at least twice in order to confirm their reproducibility. A typical experiment includes the following steps:

A) *Quantitation of building blocks and target oligoribonucleotides.* In general, building blocks containing thiol were quantified by Ellman's test. *Procedure:* Add 50 μL of Ellman's reagent (2 mM dithio-bis-2-nitrobenzoic acid (DTNB) in 50 mM aqueous NaOAc), 100 μL of aqueous 2 M Tris-HCl pH 8 and H_2O (850 μL minus the sample volume) into a 1-cm path-length quartz cuvette, and mix thoroughly. Zero the UV spectrometer at 412 nm. Prepare the sample containing thiol by adding the appropriate volume (10 μL , 20 μL , 30 μL , etc.). Mix well and incubate at RT for 2 min. Measure the absorbance at 412 nm and calculate the

concentration using the extinction coefficient (ϵ_{412} : $13,600 \text{ M}^{-1} \text{ cm}^{-1}$). Although building blocks were dissolved in 0.1% TFA in H_2O to increase their solubility, the pH of the Ellman solution was not affected.

For the quantitation of the acridine derivatives **Acr1** and **Acr2**, the extinction coefficient of their amino precursors, ϵ_{343} : $1,979 \text{ M}^{-1} \text{ cm}^{-1}$ and ϵ_{361} : $9,910 \text{ M}^{-1} \text{ cm}^{-1}$ respectively, was used (determined in 0.1% TFA in H_2O).

The concentration of oligoribonucleotides was determined from the 260 nm absorbance value at 25°C. The extinction coefficient was calculated on the basis of dinucleotide frequencies and composition using the nearest-neighbor model. The following molar extinction coefficients of ribonucleosides were used for the UV-quantitation of oligoribonucleotides: ϵ_i (L/mol/cm, 260 nm): U, 9,900; A, 15,400; C, 7,200; and G, 11,500.

B). *RNA free exchange experiments.* 24 nmol of each building block (in 0.1% TFA in H_2O) were combined in an Eppendorf tube, and freeze dried or evaporated in a Speed-Vac. Then, 240 μL of buffer (50 mM Tris-HCl, pH 7.7, 100 mM NaCl and 0.1 mM Na_2EDTA) was added and the mixture was shaken gently. The DCL mixtures were left to stand at RT under air without stirring. After the desired time, an aliquot was taken (approximately 1/3 part) and the disulfide exchange was stopped by the addition of 100 μL of a 0.1% TFA solution in water (pH ~2-3), and analysed by UV-MS HPLC.

In some DCC experiments the building blocks were incorporated as the corresponding disulfide homodimers. In those cases, 12 nmol were used in order to keep the same final concentration of monomer.

C). *RNA templated exchange experiments.* 6 nmol of biotinylated RNA (wt or +3) were annealed in 240 μL of buffer (50 mM Tris-HCl, pH 7.7, 100 mM NaCl and 0.1 mM Na_2EDTA) by heating to 90°C for 5 min and then slowly cooling to room temperature. After overnight incubation at RT, the solutions were stored at 4°C. Then, annealed biotinylated RNA was added to the Eppendorf tube containing the evaporated building blocks, and the resulting mixture left to stand for 4 days at RT under air without stirring. At the desired time, an aliquot was withdrawn (approx. 1/3) and the disulfide exchange was stopped by the addition of 45-70 μL of 0.1% TFA solution in water (pH ~ 5-6).

Streptavidin-coated magnetic beads (Biomag Streptavidin, 5 mg/mL suspension, Qiagen) were used to isolate the biotinylated RNA and the binding ligands. In all washing procedures, a magnet was used to retain the beads in the tube while the supernatant was pipetted off. First, the beads (500 μL of suspension for each DCL aliquot) were separated from the commercial buffer solution and washed with an acidic buffer (3 x 500 μL of 50 mM Tris-HCl, pH 5.8, 100 mM NaCl and 0.1 mM Na_2EDTA). DCL aliquots were added to the washed beads and incubated at room temperature. After 20 min, the beads were retained in the vessel using the magnet and the supernatant solution was pipetted off again. Then, the beads were treated to remove the non-interacting ligands and building blocks (3 x 200 μL of 50 mM Tris-HCl, pH 5.8, 100 mM NaCl and 0.1 mM Na_2EDTA). Finally, the beads were washed with a hot solution of 0.1% TFA in H_2O in order to liberate RNA-binding ligands (3 x 200 μL , incubation at 90°C for 10 min). The solutions were combined and evaporated in a Speed-Vac. The final residue was dissolved in 0.1% TFA in H_2O and analysed by UV-MS HPLC.

D) *UV-MS HPLC analysis.* UV-MS HPLC analysis of DCC libraries was performed using a Micromass ZQ mass spectrometer equipped with an electrospray source and a single quadrupole detector coupled to a Waters 2695 HPLC (photodiode array detector). The detection wavelength was set to 260 nm. Elution was performed on a GraceSmart C_{18} column (150 x 2.1 mm, 5 μm , flow rate: 0.25 mL/min) with linear gradients of H_2O and ACN containing both solvents either 0.1% formic acid or 0.1 % formic acid and 0.01% TFA. Typical gradient: 0 to 35% B in 15 min and from 35% to 80 % B in 10 min. In some cases,

UV-MS HPLC analysis was carried out with both elution conditions to avoid the overlapping of some peaks in order to allow a more accurate integration.

All peak areas of the HPLC traces were integrated and normalized taking into account the extinction coefficient of each compound at the detection wavelength: ϵ_{260} : **Acr1** 37,090, **Acr2** 13,334, **Azq** 2,930, **TyrP** 596, **TrpP** 3,484 M⁻¹ cm⁻¹. Histograms showing the change in the mol percentage for all DCL members at different times were generated in order to verify that thermodynamic equilibrium had been reached. In addition, histograms showing the mol percentage changes in each species of the equilibrium mixture (amplification %) upon introduction of the target RNA were also generated.

Evaluation of the interaction between RNA and ligands. A. UV-monitored melting experiments. The impact of the ligands on the thermal stability of wt and mutated (+3 and +14) stem-loop RNA structures was estimated by cooling/heating experiments. Samples were placed in 1-cm path-length quartz cuvettes in a Jasco V-550 spectrophotometer equipped with a thermoregulated cell holder. Melting curves were recorded by cooling the samples from 90°C to 20°C at a constant rate of 0.5 °C/min and measuring the absorbance at 260 nm as a function of temperature. Then, the reverse denaturation curve (20°C to 90 °C) was recorded. This cooling/heating experiment ensures that the initial state corresponds to a thermodynamic equilibrium. In all cases the two curves were superimposable, which indicated that the transition was kinetically reversible. Unless otherwise indicated, the solutions were 1 μ M both in RNA and ligands, in 10 mM sodium phosphate buffer pH 6.8, 100 mM NaCl and 0.1 mM Na₂EDTA. The corresponding melting temperature values (T_m) were determined using the baseline method. All experiments were repeated at least three times until coincident T_m values were obtained. The error in T_m values was $\pm 0.2^\circ\text{C}$.

B. Circular dichroism. Samples were prepared as described in the UV-monitored melting experiments section (3 μ M both in RNA and ligands, in 10 mM sodium phosphate buffer pH 6.8, 100 mM NaCl and 0.1 mM Na₂EDTA). Spectra were recorded on a Jasco J-720 spectropolarimeter with a thermoregulated cell holder and interfaced with a Neslab RP-100 water bath, at 20°C. All CD spectra were baseline subtracted with a separately acquired buffer spectrum.

C. UV-Vis titration experiments. A 45-55 μ M solution of the ligand (**Acr1-Nea**, **Acr2-Nea/Nea2** or **Azq-Nea**) and the corresponding amount of wt RNA (0, 0.02, 0.05, 0.1, 0.2, 0.5, 1 and 2 eq) was prepared in 10 mM sodium phosphate buffer pH 6.8 containing 100 mM NaCl and 0.1 mM Na₂EDTA. The mixture was heated for 5 min to 90°C and left to slowly cool to RT. The absorption spectra were recorded at RT. The percentage of absorption quenching was determined at the following absorption bands: 360 nm (31%) for **Acr1-Nea** and 423 (40%) and 444 nm (35%) for **Acr2-Nea**.

D. Fluorescence titration experiments. A solution of the ligand (**Acr1-Nea** or **Acr2-Nea**) and the corresponding amount of wt RNA was prepared in 10 mM sodium phosphate buffer pH 6.8 containing 100 mM NaCl and 0.1 mM Na₂EDTA. The mixture was heated for 5 min to 90°C and left to slowly cool to RT. The fluorescence emission spectra were recorded at RT.

E. Fluorescence binding assays. Fluorescence measurements were performed in 1-cm path-length quartz cells on a QuantaMaster fluorometer (PTI) at 20°C, with an excitation slit width of 4.0 nm and an emission slit width of 5.2 nm. Upon excitation at 490 nm, the emission spectrum was recorded over a range between 500 and 550 nm until no changes in the

fluorescence intensity were detected. All binding assays were performed in 10 mM sodium phosphate buffer, pH 6.8, 100 mM NaCl and 0.1 mM Na₂EDTA.

For each experiment, the fluorescence spectrum of 120 μ L buffer solution without RNA or ligand was first taken, to be used as the baseline. Following this buffer blank, the spectrum of a 0.25 μ M solution of refolded RNA containing fluoresceine (120 μ L) was recorded, and the baseline blank subtracted. Subsequent aliquots of 1 μ L of an aqueous ligand solution (increasing in concentration from 0 to 0.75 mM, 0.0005-3000 eq, depending on the ligand affinity) were added to the solution containing RNA, and the fluorescence spectrum was recorded after addition of each aliquot until the fluoresceine fluorescence signal at 517 nm reached saturation (typically 5-10 min). Over the entire range of ligand concentrations, the emission maxima varied less than 1 nm. The total volume of the sample never changed more than 20%. The full titration was repeated in the absence of labelled RNA to correct for the presence of the ligand's fluorescence. These spectra were subtracted from each corresponding point of the labelled RNA titrations, and the resulting fluorescence intensity was corrected for dilution ($F \cdot V/V_0$).

The emission fluorescence at 517 nm was normalized by dividing the difference between the observed fluorescence, F , and the initial fluorescence, F_0 , by the difference between the final fluorescence, F_f , and the initial fluorescence, F_0 . This normalized fluorescence intensity was plotted as a function of the logarithm of the total ligand concentration. Finally, nonlinear regression using a sigmoidal dose-response curve was performed with the software package GraphPad Prism 4 (GraphPad Software, San Diego, CA) to calculate the EC₅₀ values.

Experimental errors were less than or equal to $\pm 25\%$ of each value.

For competitive experiments, a tRNA from baker's yeast (*S. cerevisiae*) was purchased from Sigma. Stock solutions of tRNA^{mix} were quantified using an average extinction coefficient of 9,640 cm⁻¹ per base.^[24b] The fluorescence binding assays were carried out as described above with the exception that a 30-fold excess (base) of the tRNA^{mix} was added to the refolded fluoresceine-labelled RNA (or to the buffer for the titration without target RNA).

Acknowledgements

The authors wish to thank Prof. Carlos González for fruitful discussions. This work was supported by funds from the Spanish *Ministerio de Educación y Ciencia* (grant CTQ 2007-68014-01/02), the *Generalitat de Catalunya* (2005SGR-693 and the *Xarxa de Referència de Biotecnologia*) and the *Programa d'Intensificació de la Recerca* (Universitat de Barcelona). Paula López-Senín received a fellowship from the *Ministerio de Educación y Ciencia*.

References

- [1] a) J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171-1224; b) F. Aboul-ela, *Future Med. Chem.* **2010**, *2*, 93-119.
- [2] a) J. Gallego, G. Varani, *Acc. Chem. Res.* **2001**, *34*, 836-843; b) Y. Tor, *ChemBioChem* **2003**, *4*, 998-1007; c) N. Foloppe, N. Matassova, F. Aboul-ela, *Drug Discov. Today* **2006**, *11*, 1019-1027; d) T. Hermann, *Biopolymers* **2003**, *70*, 4-18.
- [3] B. Carlson, O. M. Stephens, P. A. Beal, *Biopolymers* **2003**, *70*, 86-102.
- [4] a) S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science*, **1996**, *274*, 1531-1534; b) M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, *J. Med. Chem.* **2008**, *51*, 3661-3680; c) M. M. Lee, J. L. Childs-Disney, A. Pushechnikov, J. M. French, K. Sobczak, C. A. Thornton, M. D. Disney, *J. Am. Chem. Soc.* **2009**, *131*, 17464-17472.
- [5] a) J. Lehn, A. V. Eliseev, *Science*, **2001**, *291*, 2331-2332; b) P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652-

- 3711; c) *Dynamic Combinatorial Chemistry: In Drug Discovery, Bioorganic Chemistry, and Materials Science*, (Ed. B. Miller), John Wiley & Sons, Inc., Hoboken, N. J., **2010**.
- [6] C. Karan, B. L. Miller, *J. Am. Chem. Soc.* **2001**, *123*, 7455-7456.
- [7] a) A. M. Whitney, S. Ladame, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2004**, *43*, 1143-1146; b) S. Ladame, A. M. Whitney, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2005**, *44*, 5736-5739; c) A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2008**, *47*, 2677-2680.
- [8] a) A. Bugaut, J.-J. Toulme, B. Rayner, *Angew. Chem. Int. Ed.* **2004**, *43*, 3144-3147; b) A. Bugaut, J.-J. Toulme, B. Rayner, *Org. Biomol. Chem.* **2006**, *4*, 4082-4088.
- [9] a) B. R. McNaughton, P. C. Gareiss, B. L. Miller, *J. Am. Chem. Soc.* **2007**, *129*, 11306-11307; b) P. C. Gareiss, K. Sobczak, B. R. McNaughton, P. B. Palde, C. A. Thornton, B. L. Miller, *J. Am. Chem. Soc.* **2008**, *130*, 16254-16261.
- [10] a) M. Goedert, R. Jakes, *Biochim. Biophys. Acta* **2005**, *1739*, 240-250; b) C. Ballatore, V. M.-Y. Lee, J. Q. Trojanowski, *Nat. Rev. Neurosci.* **2007**, *8*, 663-672.
- [11] F. Liu; G. Cheng-Xin, *Mol. Degener.* **2008**, *3*:8.
- [12] a) K. R. Brunden, J. Q. Trojanowski, V. M.-Y. Lee, *J. Alzheimer's Dis.* **2008**, *14*, 393-399; b) M. S. Wolfe, *J. Biol. Chem.* **2009**, *284*, 6021-6025.
- [13] a) L. Varani, M. Hasegawa, M. G. Spillantini, M. J. Smith, J. R. Murrell, B. Ghetti, A. Klug, M. Goedert, G. Varani, *Proc. Nat. Acad. Sci. U. S. A.* **1999**, *96*, 8229-8234; b) C. P. Donahue, C. Muratore, J. Y. Wu, K. S. Kosik, M. S. Wolfe, *J. Biol. Chem.* **2006**, *281*, 23302-23306.
- [14] a) L. Varani, M. G. Spillantini, M. Goedert, G. Varani, *Nucleic Acids Res.* **2000**, *28*, 710-719; b) C. P. Donahue, J. Ni, E. Rozners, M. A. Glicksman, M. S. Wolfe, *J. Biomol. Screening* **2007**, *12*, 789-799; c) S. Zheng, Y. Chen, C. P. Donahue, M. S. Wolfe, G. Varani, *Chem. & Biol.*, **2009**, *16*, 557-566; d) Y. Liu, E. Peacey, J. Dickson, C. P. Donahue, S. Zheng, G. Varani, M. S. Wolfe, *J. Med. Chem.* **2009**, *52*, 6523-6526.
- [15] Experimental details are provided in the Supporting Information.
- [16] S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Drug Discovery Today* **2002**, *7*, 117-125.
- [17] a) D. P. Arya, L. Xue, B. Willis, *J. Am. Chem. Soc.* **2003**, *125*, 10148-10149; b) K. F. Blount, F. Zhao, T. Hermann, Y. Tor, *J. Am. Chem. Soc.* **2005**, *127*, 9818-9829.
- [18] R. Martínez, L. Chacón-García, *Curr. Med. Chem.* **2005**, *12*, 127-151.
- [19] a) J. F. Arambula, S. R. Ramisetty, A. M. Baranger, S. C. Zimmerman, *Proc. Nat. Acad. Sci.* **2009**, *106*, 16068-16073; b) Z. Yan, S. Sikri, D. L. Beveridge, A. M. Baranger, *J. Med. Chem.* **2007**, *50*, 4096-4104.
- [20] S. Hagihara, H. Kumasawa, Y. Goto, G. Hayashi, A. Kobori, I. Saito, K. Nakatani, *Nucleic Acids Res.* **2004**, *32*, 278-286.
- [21] Control experiments showed that the streptavidin-coated magnetic beads did not significantly bind to any of the members of the library or to the ligands, thereby confirming that amplification of **Acr1-Nea** and **Azq-Nea** was a consequence of the RNA templating.
- [22] Although the disulfide-mediated linkage of the four thiol-derivatized monomers may produce up to 10 theoretical dimers, the abundance of only 7 compounds was determined. The fact that **Nea-Nea**, **Nea-TyrP** and **TyrP-TyrP** were eluted very quickly under the conditions used for HPLC analysis prevented accurate quantification. However, it is noteworthy that close inspection of the MS traces in the presence of RNA revealed a small amplification of **Nea-Nea** and **Nea-TyrP**.
- [23] Control experiments in the presence of 0.01% Triton X-100 indicated that the amplification of **Acr2-Acr2** and **Acr2-Azq** was not a consequence of self-recognition of the heteroaromatic moieties (S. L. McGovern, B. T. Helfand, B. Feng, B. K. Shoichet, *J. Med. Chem.* **2003**, *46*, 4265-4272).

- [24] a) S. R. Kirk, N. W. Luedtke, Y. Tor, *J. Am. Chem. Soc.* **2000**, *122*, 980-981; b) N. W. Luedtke, Q. Liu, Y. Tor, *Biochemistry* **2003**, *42*, 11391-11403.
- [25] a) B. Llano-Sotelo, C. S. Chow, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 213-216; b) J. R. Thomas, X. Liu, P. J. Hergenrother, *Biochemistry*, **2006**, *45*, 10928-10938.
- [26] S. J. Lee, S. Hyun, J. S. Kieft, J. Yu, *J. Am. Chem. Soc.*, **2009**, *131*, 2224-2230.
- [27] Effect of Neomycin B (1 eq): $\Delta T_m = +3^\circ\text{C}$ (wt) and $+5^\circ\text{C}$ (+3 and +14 mutants). It is worth noting that, in our experimental conditions, these values are slightly lower than those previously reported in the literature with the same amount of neomycin.^[14a] Effect of Mitoxantrone (1 eq): $\Delta T_m = +1.9^\circ\text{C}$ (wt), $+6.1^\circ\text{C}$ (+3 mutant) and $+4.8^\circ\text{C}$ (+14 mutant).
- [28] a) S. E. Kitchen, Y.-H. Wang, A. L. Baumstark, W. D. Wilson, D. W. Boykin, *J. Med. Chem.* **1985**, *28*, 940-944; b) F. Hamy, V. Brondani, A. Florsheimer, W. Stark, M. J. Blommers, T. Klimkait, *Biochemistry* **1998**, *37*, 5086-5095.
- [29] S. Sánchez-Carrasco, J. G. Delcros, A. A. Moya-García, F. Sánchez-Jiménez, F. J. Ramírez, *Biophys. J.* **2008**, *133*, 54-65.
- [30] a) E. Buratti, F. E. Baralle, *Mol. Cell. Biol.* **2004**, *24*, 10505-10514; b) M. B. Warf, J. A. Berglund, *Trends Biochem Sci.*, **2010**, *35*, 169-178.

Scheme and Figure legends

Figure 1. a) Sequences and secondary structure of wild-type (wt) and mutated Tau stem-loop target RNAs. Exonic sequences are shown in capital letters and intronic sequences in lower case, with the +3 and +14 FTDP-17 mutations indicated. Nucleotides involved in base pairs identified previously by NMR are connected by a dash.^[13a] When required, biotin or fluoresceine derivatization was performed at the 5' end. The ends of the chains were modified with 2'-*O*-methylribonucleosides to increase stability with RNases (denoted by an asterisk). b) Structure and peptide sequences of the building blocks used in the DCC experiments.

Scheme 1. Schematic representation of a DCC process.

Figure 2. DCC experiments involving wt RNA and **Nea**, **Acr1**, **Acr2**, **Azq**, **TyrP** and **TrpP** monomers. a) HPLC traces showing the composition of the DCL in the absence (left) and in the presence (right) of wt RNA, after 48 h. b) Histograms showing the changes in the DCL composition (left) and the percentage changes (% amplification) of each species (right) upon introduction of the wt RNA.

Figure 3. UV-Vis titration of the **Acr1-Nea** (a) and **Acr2-Nea** (b) ligands (50 μM) with increasing amounts of wt RNA (0-1 eq) in a 10 mM sodium phosphate buffer, pH 6.8, containing 100 mM NaCl and 0.1 mM Na₂EDTA.

Tables

Table 1. Binding of the ligands to wt RNA in the absence or in the presence of a tRNA competitor.

Ligand	EC ₅₀ (μM) ^[a]	EC ₅₀ (μM) +tRNA ^[b]	EC ₅₀ +tRNA / EC ₅₀
Neamine	3300	nd	nd
Azq-Nea	58.0	nd	nd
Acr1-Nea	28.6	112.2	3.9
Acr2-Nea	5.9	63.1	10.7
Acr2-Nea2	2.1	47.0	22.4
Acr2-Acr2	2.9	41.0	14.0

[a] All fluorescence measurements (25 μM RNA) were performed in 10 mM sodium phosphate buffer pH 6.8, 100 mM NaCl and 0.1 mM Na₂EDTA. [b] Measured in the presence of a 30-fold nucleotide excess of a mixture of tRNA (tRNA^{mix}).

Table 2. Melting temperature (T_m) values for the complexation of ligands with target RNAs (1 μ M both in RNA and in ligands in 10-mM sodium phosphate buffer, pH 6.8, 100 mM NaCl and 0.1 mM Na₂EDTA).

	T_m wt ^[a]	ΔT_m	T_m+3	ΔT_m	T_m+14	ΔT_m
No ligand	66.4	-	50.8	-	54.0	-
Neamine	67.6	+1.2	51.8	+1.0	54.4	+0.4
Azq-Nea	67.5	+1.1	51.9	+1.1	54.9	+0.9
Acr1-Nea	67.1	+0.7	52.0	+1.2	54.1	+0.1
Acr2-Nea	68.5	+2.1	53.6	+2.8	56.0	+2.0
Acr2-Nea2	68.8	+2.4	56.5	+5.7	57.2	+3.2
Acr1-Acr1	65.4	-1.0	50.6	-0.2	53.8	-0.2
Acr2-Acr2	67.6	+1.2	53.4	+2.6	55.4	+1.4

[a] $\Delta T_m = (T_m \text{ of the RNA in the presence of ligand}) - (T_m \text{ of RNA alone})$.

Text for the Table of Contents

Dynamic combinatorial chemistry has allowed the identification of several ligands that bind with high to medium affinity and stabilize the stem-loop structure located at the exon 10-5'-intron junction of pre-mRNA which is involved in the onset of several tauopathies. The overall results show a good correlation between the ligand's affinity for the target and its stabilizing properties. These results contribute to enhancing our knowledge of how to design more specific RNA ligands.

Headline for the Table of Contents

Selecting ligands for Tau RNA by Dynamic Combinatorial Chemistry

Keywords: combinatorial chemistry • template synthesis • molecular recognition • RNA structure • RNA ligands