

ARTICLE

Received 24 Jun 2015 | Accepted 18 Apr 2016 | Published 1 Jun 2016

DOI: 10.1038/ncomms11647

OPEN

Design of a bioactive small molecule that targets r(AUUCU) repeats in spinocerebellar ataxia 10

Wang-Yong Yang¹, Rui Gao², Mark Southern³, Partha S. Sarkar² & Matthew D. Disney¹

RNA is an important target for chemical probes of function and lead therapeutics; however, it is difficult to target with small molecules. One approach to tackle this problem is to identify compounds that target RNA structures and utilize them to multivalently target RNA. Here we show that small molecules can be identified to selectively bind RNA base pairs by probing a library of RNA-focused small molecules. A small molecule that selectively binds AU base pairs informed design of a dimeric compound (**2AU-2**) that targets the pathogenic RNA, expanded r(AUUCU) repeats, that causes spinocerebellar ataxia type 10 (SCA10) in patient-derived cells. Indeed, **2AU-2** (50 nM) ameliorates various aspects of SCA10 pathology including improvement of mitochondrial dysfunction, reduced activation of caspase 3, and reduction of nuclear foci. These studies provide a first-in-class chemical probe to study SCA10 RNA toxicity and potentially define broadly applicable compounds targeting RNA AU base pairs in cells.

¹Departments of Chemistry and Neuroscience, The Scripps Research Institute, Scripps Florida, Jupiter, Florida 33458, USA. ²Mitchell Center for Neurodegenerative Disorders, Department of Neurology, Neuroscience and Cell Biology, University of Texas Medical Branch, Galveston, Texas 77555, USA. ³Informatics Core, The Scripps Research Institute, Scripps Florida, Jupiter, Florida 33458, USA. Correspondence and requests for materials should be addressed to M.D.D. (email: Disney@scripps.edu).

RNA has diverse cellular functions. For example, messenger RNAs (mRNAs) encode protein, microRNAs regulate the lifetime of mRNAs, and the ribosome translates mRNAs into proteins^{1,2}. In bacteria, riboswitches control the production of proteins by binding to small molecule metabolites^{3,4}. In fact, many non-coding RNAs have been found to play significant roles in cellular biology, and these discoveries expand even further the known functions of RNA⁵.

Because of the important cellular functions of RNA under normal conditions, it is not surprising that mutations in RNA can cause disease. Single-nucleotide polymorphisms can give rise to cryptic alternative pre-mRNA splicing sites, leading to production of aberrant, defective proteins, as is the case with β -thalassaemia^{6,7}. Expanded RNA repeats can also contribute to disease and can be present in 5' and 3' untranslated regions (UTRs; fragile X-associated tremor ataxia syndrome (FXTAS)⁸ and myotonic dystrophy type 1 (DM1)⁹), introns (spinocerebellar ataxia type 10 (SCA10)¹⁰ and myotonic dystrophy type 2 (DM2)¹¹) or coding regions (Huntington's disease (HD)¹²). Small molecules that target these RNA and inhibit its dysfunction are thus highly desirable.

The bacterial ribosome is the most widely exploited RNA target^{13,14}. Ribosomes and ribosomal RNA (rRNA) are privileged targets as (i) ribosomes play essential roles in cellular homeostasis and modulation of ribosome activity can have drastic cellular consequences, and (ii) rRNA comprises about 80–90% of the total RNA content of a cell¹⁵. Riboswitches are also emerging and established targets of RNA-directed small molecules. Compounds identified to bind to and modulate riboswitches generally mimic the structure of the RNA's natural metabolite^{3,16}, akin to substrate mimicry to design enzyme inhibitors¹⁷. Most RNA targets to which a small molecule binder is desired, however, are of low abundance and have no natural metabolite to inform drug design.

To aid RNA-targeting endeavours, our group has developed two-dimensional combinatorial screening (2DCS) to identify optimal (high affinity and selective) RNA motif–small molecule interactions. These interactions are deposited into a database and can be used to design small molecules to target RNAs by comparing motifs in the desired target to the database. This approach has been used to design small molecules targeting the RNAs that cause DM^{18–20}, FXTAS²¹ and HD²². All of the interactions that are presently in the RNA motif–small molecule database are between small molecules and RNA loops such as hairpins, bulges and internal loops. Herein, we report small molecules that bind selectively to RNA base pairs. Among a variety of compounds tested, small molecules with benzimidazole moieties were identified to bind selectively to AU base pairs. These data were leveraged to design the first bioactive small molecule to target the expanded r(AUUCU) repeat that causes SCA10, an incurable neuromuscular disorder. The compound targets the central AU base pairs in r(AUUCU)^{exp} and its dimeric compound (2AU-2) displaces sequestered proteins and improves defects in patient-derived cells. The observation that base pair-binding modules can provide bioactive compounds suggests that many other RNAs can be exploited as targets of small molecules.

Results and Discussion

Binding of RNA-focused small molecules to RNA base pairs.

By using chemical similarity searching, small molecules with features that should pre-dispose them for binding RNA^{23,24} were collected from both the National Cancer Institute's and The Scripps Research Institute's chemical libraries, including benzimidazole, benzamidine, aniline moieties. Compounds were

further restrained to be fluorescent to allow for easy screening of binding events, affording 104 small molecules (Fig. 1a and Supplementary Table 1).

Compounds were screened for binding to different base pairs using four model constructs with stretches of AU and GC base pairs embedded in the stem of a common hairpin loop (Fig. 1b). The four RNAs display AU or GC pairs with different nearest neighbours. AUAU and AAUU RNAs have 5'AU/3'UA or 5'AAUU/3'UUAU stretches, respectively, while GCGC and GGCC have 5'GC/3'CG or 5'GGCC/3'CCGG stretches, respectively. Due to significant differences in the thermodynamic stability of the hairpins, RNAs with AU pairs had 12 base pairs in the stem while RNAs with GC pairs had 8 (ref. 25). The ΔG_{37}° for AUAU and AAUU is -8.5 and -6.7 kcal mol⁻¹, respectively, while for GCGC and GGCC it is -17.5 and -19.1 kcal mol⁻¹, respectively. If only eight AU base pairs were present, then the free energy drops considerably to -3.7 and -2.5 kcal mol⁻¹, respectively.

In initial compound screens, each small molecule was incubated with the four RNAs and the change in emission was measured (Fig. 1c). Changes in emission were not only analysed for statistical significance for binding to these RNAs in general, but also for binding to the different RNA structures. For this compound collection, 29% of the compounds exhibited a

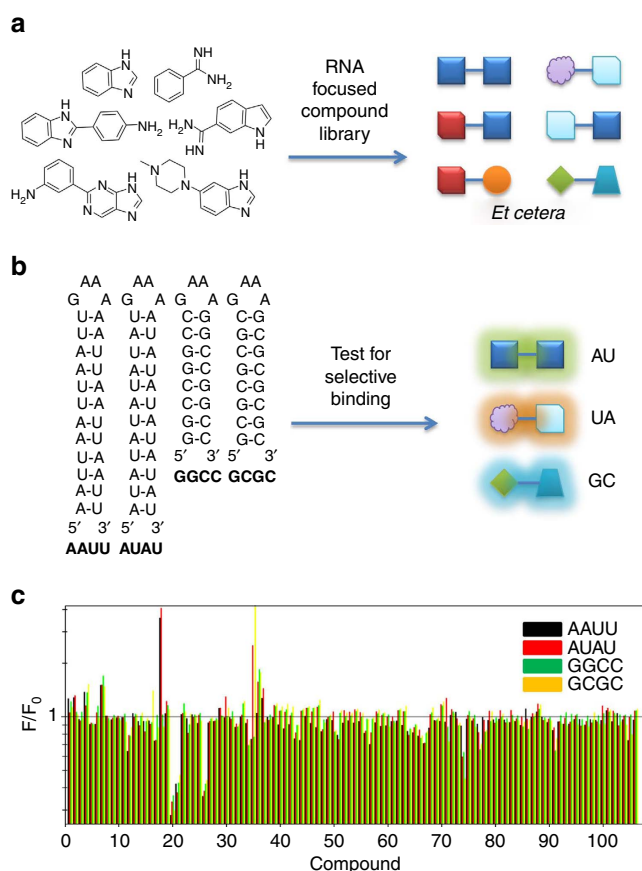


Figure 1 | Overall scheme to identify small molecules that bind to RNA base pairs.

(a) A small molecule library was collected in which the small molecules have chemotypes present in compounds known to bind RNA.

(b) The RNA-focused small molecule library was tested for binding to four different RNAs that have different orientations and identities of RNA base pairs.

(c) Compounds were tested in a fluorescence emission assay for selectively binding to RNAs with different paired elements. Compounds that were selective were further analysed.

change in emission upon binding to any of the RNAs used in this study, and 18% of the compounds showed preferential binding to AU or GC pair-containing duplexes (Supplementary Table 2).

On the basis of the substructures within compounds that bind, a Venn diagram was constructed to correlate chemotypes and binding to AU and GC base pairs (Fig. 2). For example, various functionalized purines bound to RNAs with GC base pairs. Previously, other compounds that bind RNA base pairs have been identified. For example, the Beal group has used threading intercalators with acridine or other aromatic functionalities to target RNA bulges with nearby GC pairs²⁶. Thus, it appears that acridine, and related compounds, can provide a GC base-pair-binding module. The group of Tor has shown that ethidium bromide is a 20-fold selective binder to poly r(A)-poly r(U) over poly r(G)-poly r(C)²⁷.

Each small molecule from the initial screen was tested for saturable binding when incubated with serially diluted RNA. The benzamidine compounds **1** and **2** provided the most robust and saturable change in emission and were found to bind selectively to AU base pairs (Supplementary Figs 1, 2 and 3). Due to the binding of these compounds to AU base pairs and their saturable emission properties, these compounds were selected for further study. This assay, by nature of using emission to study binding, could generate false negatives; however, it allowed us to further characterize positive hits and produce a first-in-class bioactive ligand that targets r(AUUCU), *vide infra*.

Compounds 1 and 2 bind selectively to AU base pairs. The selectivities of compounds **1** and **2** were further assessed by measuring EC_{50} s for all four RNAs. The EC_{50} s for **1** for binding to AAUU, AUAU, GGCC and GCGC are 170, 45, 390 and 240 nM, respectively, while the EC_{50} s for **2** are 170, 190, 8,550 and 6,950 nM, respectively (Table 1). On the basis of these data, **2** is much more selective for AU base pairs than **1**, with an AU pair selectivity of ~ 40 -fold (Fig. 3). Measurement of the K_d 's of **2** to AUAU and AAUU give values of 210 and 320 nM, respectively, and stoichiometries of 2.9 and 3.7, respectively (Supplementary Fig. 4). These data suggests that **2** is a high affinity binder to stretches of AU base pairs and that the compound interacts with between 3 and 4 base pairs in each of the model hairpins. Previous studies have suggested that the aminoglycoside tobramycin recognizes model poly(rI)·poly(rC) RNA duplexes and that the compound interacts with approximately 4 bp, or a similar number of interacting sites that are proposed for **2** (ref. 28). Diphenylfuran amidine was previously reported to bind poly(rA)·poly(rU) duplex via intercalation, as determined from viscosity and circular dichroism studies^{29,30}. These observations were further supported by docking studies in which the intercalation of **2** is stabilized through stacking interactions with the positively charged amidine residing in the major groove; minor groove binding could not be accommodated^{29,30}. Further, the substituents in this class of compounds affect the thermal stability of poly(rA)·poly(rU), with imidazoline providing the largest enhancement^{29,30}.

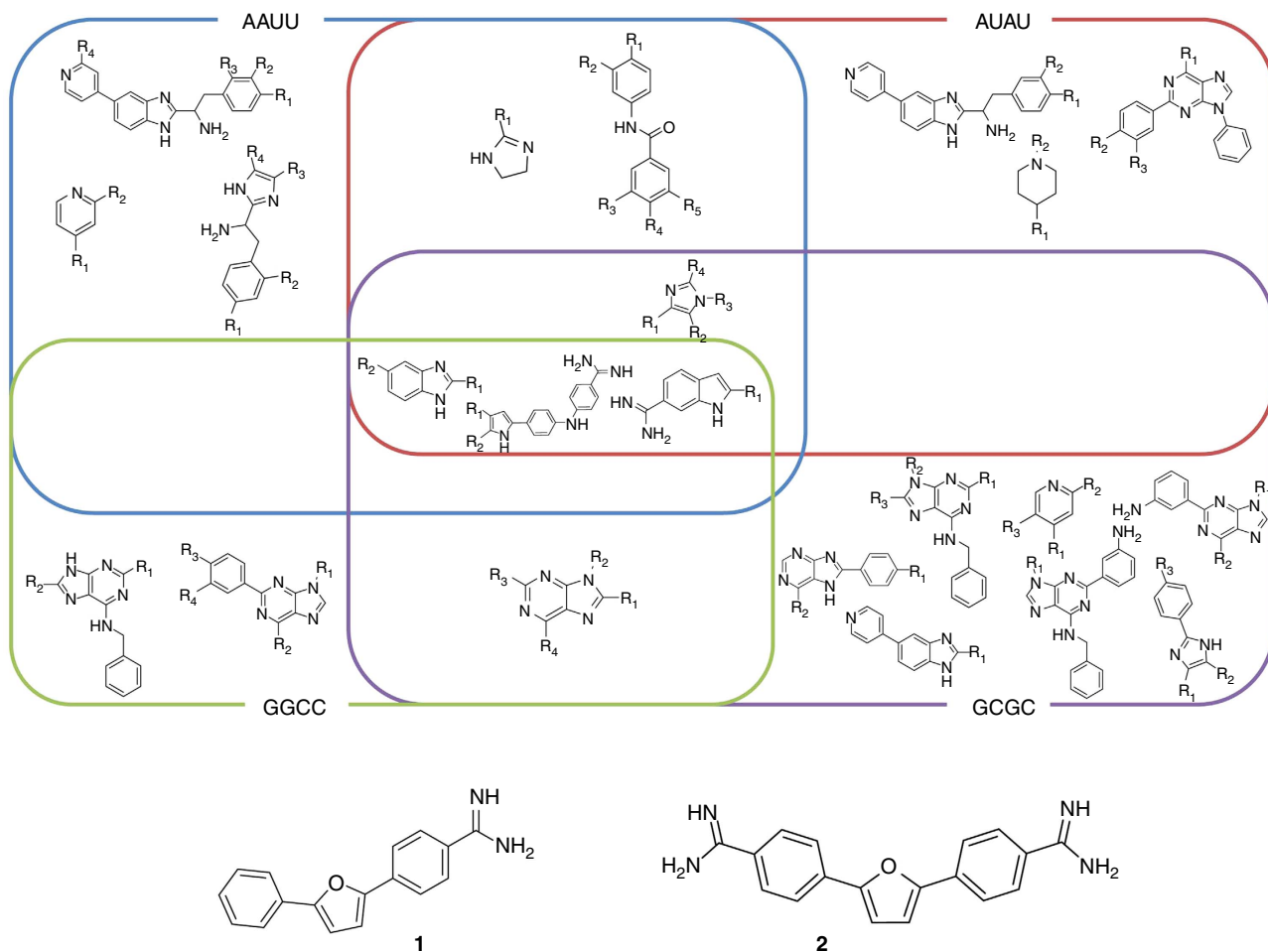
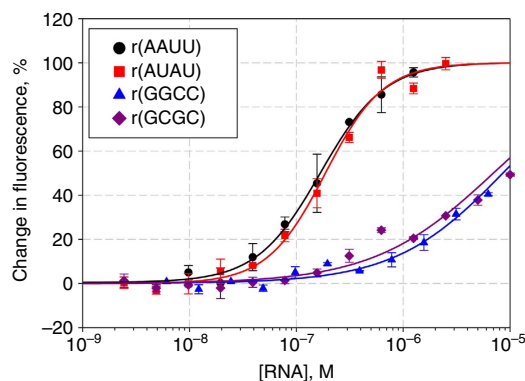


Figure 2 | Analysis of compounds identified to bind RNA base pairs. Top, Venn diagram of substructures in compounds that were found to bind to RNA from the fluorescence screening assay showed in Fig. 1c. Data were compiled by using compounds that had a P value of < 0.001 for binding to the RNA hairpins. Bottom, structures of compounds **1** and **2** that were the most avid for binding to AUAU and AAUU RNA hairpins.

Table 1 | EC₅₀s (nM) and Hill coefficients (shown in parentheses) of compounds binding to the base paired RNAs shown in Fig. 1.

Compound	AAUU	AUAU	GGCC	GCGC
1 *	170 ± 10 (3.1 ± 0.79)	45 ± 1 (1.5 ± 0.32)	390 ± 20 (1.4 ± 0.34)	240 ± 10 (0.92 ± 0.34)
2 †	170 ± 40 (1.5 ± 0.24)	190 ± 20 (1.6 ± 0.28)	8,550 ± 250 (0.85 ± 0.08)	6,950 ± 250 (0.77 ± 0.16)
AU-azide ‡	45 ± 6 (1.0 ± 0.17)	31 ± 4 (1.0 ± 0.13)	> 6,000	> 2,000

*[Compound] = 3 μM.
†[Compound] = 1 μM.
‡[Compound] = 1 μM.

**Figure 3 | Binding of compound 2 to four different RNAs** Fluorescence binding assays of compound 2 to RNAs with different base pairs show that 2 binds AU base pairs selectively over GC base pairs by >40-fold.

Compounds 1 and 2 bind r(AUUCU) repeats that cause SCA10. The ultimate goal of identifying small molecules that bind RNA motifs *in vitro* is to apply these identified interactions to target an RNA that causes disease. Fortuitously, the expanded repeating r(AUUCU) RNA (r(AUUCU)^{exp}) that causes spinocerebellar ataxia type 10 (SCA10) contains stretches of AU base pairs¹⁰, in particular repeating 5'AU/3'UA base pair steps (Fig. 4b)³¹. SCA10 is an incurable neuromuscular disorder that is mainly found in Latin America³². It is a slowly progressive disease that results in poor balance followed by loss of control over upper limbs. Previous experiments have shown that the disease is due to an RNA gain-of-function in which r(AUUCU)^{exp}, located within spliced intron 9 of the ataxin 10 (ATX10) mRNA, binds to and sequesters proteins involved in RNA biogenesis such as heterogeneous nuclear ribonucleoprotein K (hnRNP K; Fig. 4a)³³. Sequestration of hnRNP K causes a host of cellular defects that include the formation of RNA nuclear foci, translocation of protein kinase C-δ (PKCδ) in mitochondria resulting in mitochondrial dysfunction, and activation of caspase 3 and subsequent apoptosis³³. The binding of the repeats to RNA-binding proteins causes the transcript to be retained in nuclear foci in patient-derived cells and model cellular systems³³.

We assessed the selectivities of compounds 1 and 2 for binding r(AUUCU) repeats over other disease-associated RNA repeats including r(CUG)₁₂ (DM1)⁹, r(CGG)₁₂ (FXTAS and FXS)^{8,34}, and r(CCUG)₁₂ (DM2)¹¹. Compound 1 has a similar EC₅₀ for all RNA repeats (36–87 nM), indicating that 1 binds to RNA internal loops as well as to AU base pairs with similar binding affinity (Supplementary Fig. 5). Compound 2 selectively recognizes r(AUUCU)₁₁ with 5–15-fold selectivity over the other repeats (Table 2 and Supplementary Fig. 6). Compound 2 bound r(AUUCU)₁₁ with a K_d of 300 nM and a stoichiometry of ca. 11:1 (2:r(AUUCU)₁₁). Because r(AUUCU)₁₁ contains 10 AU base pairs, it indicates that 2 is binding to each AU base pair in the target. Indeed, a nuclease protection assay revealed that 2 protects each AU base pair in r(AUUCU)₁₁ from cleavage (Fig. 5).

Interestingly, calculated Hill coefficients indicate that 2 binds to r(AUUCU)₁₁ with positive cooperativity ($n=1.7$; Table 2), suggesting that 2 could be an ideal module to design a polyvalent compound to target multiple adjacent sites in r(AUUCU)^{exp} simultaneously.

An AU base pair-binding module to target r(AUUCU)^{exp}. Previously, our group has developed modularly assembled small molecules that target various repeating RNAs^{18,19,35–38}. In this approach, an RNA-binding module or modules that recognize different motifs in an RNA target are displayed on a single molecule. These polyvalent compounds allow for simultaneous recognition of multiple motifs in an RNA target, thereby increasing affinity and selectivity relative to monomeric binders (Fig. 6).

To enable this approach for 2 and r(AUUCU)^{exp}, a derivative of 2, AU-azide (Fig. 6), was synthesized to install an orthogonally reactive group. The azide is used for conjugation onto peptoid polyvalent scaffolds that display alkynes via a Cu-catalysed Huisgen cycloaddition reaction. Dibenzimidate (3) was synthesized from the furan via two reaction steps as reported previously^{39,40}, followed by two amidation reactions to obtain AU-azide (Supplementary Fig. 19; see Supplementary Information for chemical synthesis details and compound characterization). After obtaining the desired compound, it was tested for selective recognition of AU over GC base pairs (Table 1 and Supplementary Fig. 7). Interestingly, AU-azide has enhanced selectivity of AU base pair than 2, >45-fold selective for AU over GC base pairs (Table 1). The selectivity of AU-azide for r(AUUCU)₁₁ over other RNA repeats compared with 2 was also improved (Table 2 and Supplementary Fig. 8).

Development of dimeric compounds to target r(AUUCU)^{exp}.

We next synthesized a small library of AU dimers (Fig. 6) by using a previously published approach^{19,35–37}. A peptoid backbone was used as a polyvalent scaffold that contains two propargylamine submonomers separated by different distances afforded by varying the number of propylamine spacers inserted into the backbone. The AU-azides (RNA-binding modules) were then conjugated by using a Cu-catalysed click reaction (Supplementary Fig. 21). The nomenclature used for these compounds is 2AU-n where AU indicates the RNA-binding module; the number before AU indicates valency; and the number after the dash indicates the number of propylamine spacing modules between AU RNA-binding modules.

The library of dimeric compounds was then tested for binding to r(AUUCU)₁₁ by using a filter-binding assay (Supplementary Fig. 10). After incubation of radioactively labelled r(AUUCU)₁₁ with dimeric compound, the compound-bound RNA was separated from free RNA by using a Dot-Blot apparatus. These studies showed that 2AU-2 bound to the RNA to the greatest extent, 2–20-fold greater than the other dimers. Additional filter-binding assays were completed with 35-fold excess transfer RNA (tRNA) to gain insight into selectivity. Only a modest

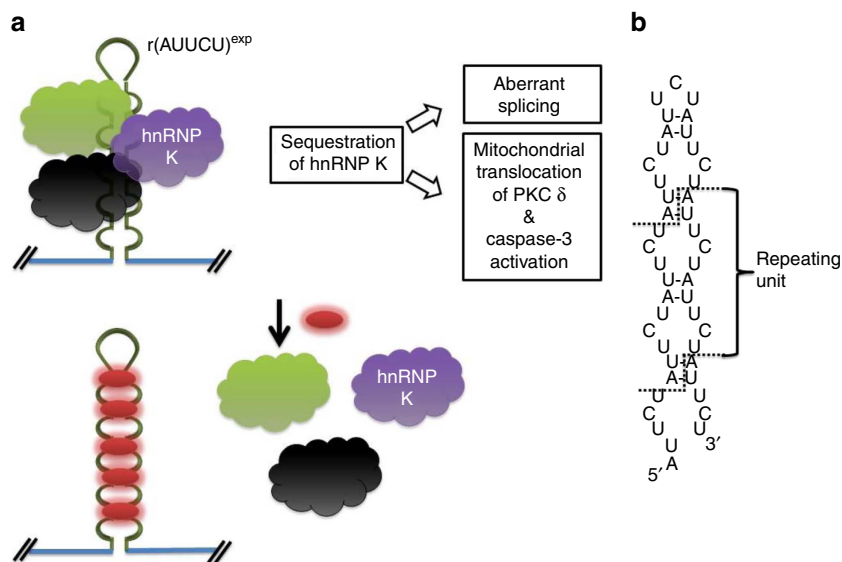


Figure 4 | Schematic of the pathogenic mechanism of SCA10 and a therapeutic approach by targeting the secondary structure of r(AUUCU) repeats. (a) r(AUUCU)^{exp}, located within intron 9 of ataxin 10 (*ATX10*) mRNA, sequesters proteins including hnRNP K. Sequestration of hnRNP K results in aberrant splicing of transcripts, mitochondrial translocation of PKC δ and caspase-3 activation, leading to apoptosis (top). A possible therapeutic approach is using small molecules that bind r(AUUCU) repeats and displace sequestered proteins (bottom). (b) secondary structure of r(AUUCU) repeats with periodically repeating 5'AU/3'UA base pair steps.

Table 2 | K_d s, Hill coefficients (shown in parentheses) and stoichiometries of 2, AU-azide and 2AU-2 binding to RNA repeats.

RNA repeats	r(AUUCU) ₁₁		r(CUG) ₁₂ [*]		r(CGG) ₁₂ [*]		r(CCUG) ₁₂ [*]	
	K_d , nM	Stoichiometry	K_d , nM	Stoichiometry	K_d , nM	Stoichiometry	K_d , nM	Stoichiometry
2	300 ± 47 (1.7 ± 0.2)	11 ± 0.5	~4,000	~7	~2,000	~6	~4,000	~11
AU-azide	375 ± 34 (1.3 ± 0.1)	13 ± 0.9	~5,000	~11	~3,000	~9	~5,000	~16
2AU-2	185 ± 1.0 (1.4 ± 0.4)	7 ± 0.1	~4,000	~9	~4,000	~6	~4,000	~6

^{*}The stoichiometry from unsaturated binding curves was estimated from data fit by equation (1). The K_d was calculated by using the estimated stoichiometry.

decrease (~20%) of 2AU-2 binding was observed, suggesting the compound is selective.

To further confirm the results of the filter binding assays, we developed an assay that evaluates the potency of small molecules for inhibiting protein loading onto r(AUUCU)₁₁. It is known that DiGeorge syndrome critical region 8 (DGCR8), a protein that is involved in microRNA biogenesis, binds a wide variety of RNAs⁴¹ and thus could be of potential use as the protein component in this assay. To establish that DGCR8 Δ binds r(AUUCU)₁₁, a gel mobility shift assay was completed, affording a dissociation constant of 1.6 μ M for the RNA–protein interaction (Supplementary Fig. 11). Such affinity is similar to those found between other RNA repeats and proteins, including r(CUG)^{exp}-MBNL1 (muscleblind-like 1), r(CAG)^{exp}-MBNL1, and r(CGG)^{exp}-DGCR8 (refs 20–22).

Time-resolved fluorescence resonance energy transfer (TR-FRET) assays have been developed to screen for inhibitors of each RNA–protein complexes mentioned above^{20–22}. Thus, we developed a TR-FRET assay for r(AUUCU)₁₁ and DGCR8 Δ analogously. Screening results showed that dimers are better inhibitors than the monomer and that 2AU-2 is the most potent inhibitor amongst them (Fig. 7a). At 1.5 μ M concentration, AU-azide only inhibits ~25% of r(AUUCU)₁₁-DGCR8 Δ complex formation while 2AU-2 inhibits ~70%. Further, the IC₅₀ of 2AU-2 was 2.7-fold less than that of AU-azide (Supplementary Fig. 12). Thus, both the TR-FRET and filter-binding assays establish that 2AU-2 is the most potent

binder of r(AUUCU)₁₁ and that there is considerable enhancement for the dimer relative to the monomer.

Consistent with results from filter binding and TR-FRET assays, selective binding of 2AU-2 to r(AUUCU)₁₁ over other RNA repeats was observed (Table 2 and Supplementary Fig. 9). The stoichiometry and the K_d of dimer 2AU-2 to r(AUUCU)₁₁ was measured. The number of 2AU-2 per r(AUUCU)₁₁ was approximately half that of AU-azide ($n=7$ and $n=13$, respectively), suggesting that each RNA-binding module interacts with each AU pair. The affinity of 2AU-2 for r(AUUCU)₁₁ is twofold greater than the monomer and AU-azide and 2AU-2 maintained positively cooperative binding to r(AUUCU)₁₁. Similar increases in affinity when increasing valency from $n=1$ to $n=2$ have been observed with other repeats^{19,35}. Furthermore, we studied the binding of 2AU-2 and r(AUUCU)₁₁ under molecular crowding conditions that mimic a cellular environment by adding 20% (w/v) PEG 8000. A ninefold enhancement in binding was observed (EC₅₀ of 16 ± 4 nM) relative to non-molecular crowding conditions (EC₅₀ of 146 ± 8 nM; Supplementary Fig. 13).

The binding of 2AU-2 to a DNA hairpin that contains a stretch of four consecutive AT pairs was studied. Saturable binding was not observed ($K_d > 10 \mu$ M; Supplementary Fig. 14). Interestingly, AU-azide binds to the DNA hairpin (EC₅₀ = 116 ± 15 nM) with similar affinity as the AU-rich RNAs (EC₅₀ = 111 ± 3 and 113 ± 8 nM for AAUU and AUAU, respectively; Supplementary Figs 14 and 15). Thus, a monomeric binding module with modest

selectivity can be reprogrammed to afford highly selective compounds for repetitive sequences by linking the binding modules together.

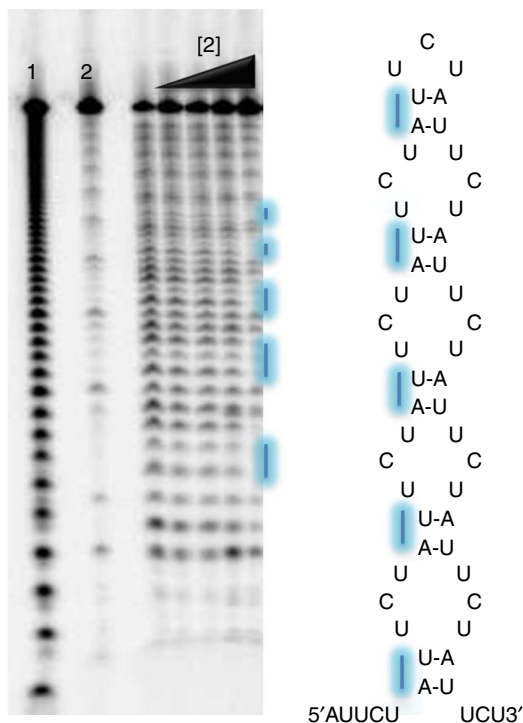


Figure 5 | Compound 2 protects r(AUUCU)₁₁ from cleavage by RNase V1.

Left, representative gel image of the RNase protection assay. Lane 1, alkaline hydrolysis of r(AUUCU)₁₁; lane 2, r(AUUCU)₁₁ + No RNase V1; lanes 3–7, r(AUUCU)₁₁ + RNase V1 + **2** (0, 0.02, 0.2, 2 and 20 μM). Lines indicate sites of protection. Right, secondary structure of r(AUUCU)₁₁. Protected nucleotides are indicated with blue lines.

Recognition of r(AUUCU)₅₀₀ by 2AU-2 in cellular lysates. To investigate whether the designer dimer binds to r(AUUCU)₅₀₀, we developed a method named Chem-Quant-Seq. To enable Chem-Quant-Seq a biotinylated derivative, **2AU-2-Biotin**, was synthesized (Fig. 7b and Supplementary Fig. 23). After incubation of 250 nM of **2AU-2-Biotin** with total RNA isolated from cells that express r(AUUCU)₅₀₀ (ref. 33), bound RNAs were isolated with streptavidin resin. After extensive washing and elution of bound RNAs, the amount of r(AUUCU)₅₀₀ bound to **2AU-2-Biotin** was quantified by using quantitative PCR with reverse transcription (qRT-PCR). Results show that there is a significant enrichment of the r(AUUCU)₅₀₀ in the pulled-down material, as compared to 18S rRNA, showing that the compound indeed recognizes that target in the presence of cellular RNAs (Fig. 7c). To further assess the ability of the compound to recognize r(AUUCU)₅₀₀ over shorter repeats of non-pathogenic length, various concentrations of r(AUUCU)₁₁ in excess of r(AUUCU)₅₀₀ were added to the lysate and enrichment of r(AUUCU)₅₀₀ in the pulled-down fraction was quantified. Results from this competition experiment showed that 100-fold excess of r(AUUCU)₁₁ was required to decrease enrichment of r(AUUCU)₅₀₀ in the pulled-down fraction (Fig. 7d). Thus, **2AU-2-Biotin** recognizes long repeats over shorter ones. Furthermore, these data also point to cooperative binding of **2AU-2** to r(AUUCU) repeats as being a manner in which longer repeats are preferred over shorter ones. Positive cooperativity for binding to **2AU-2** was also observed *in vitro* (Hill coefficients; Table 2).

We further profiled target selectivity amongst 93 highly abundant transcripts in the pulled-down fraction via qRT-PCR. The RNAs span the diverse biology in the transcriptome, including rRNAs, mRNAs, small RNAs (sRNAs) and tRNAs. The rRNAs included all four rRNA subunits (18S, 28S 5S and 5.8S) and 45S rRNA. The 50 mRNAs were chosen from the most abundant mRNAs in HeLa cells for which there are established qRT-PCR primers⁴². The 17 sRNAs were selected from different structural and functional classes⁴³: small nucleolar RNAs (HBII-85, HBII-420, U105 C/D Box snoRNAs, and ACA-16, ACA-44,

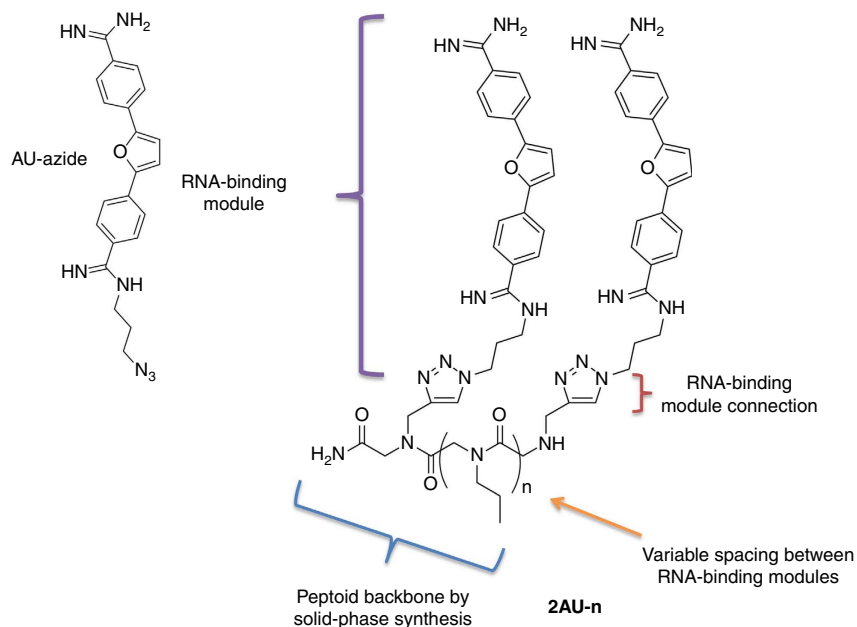
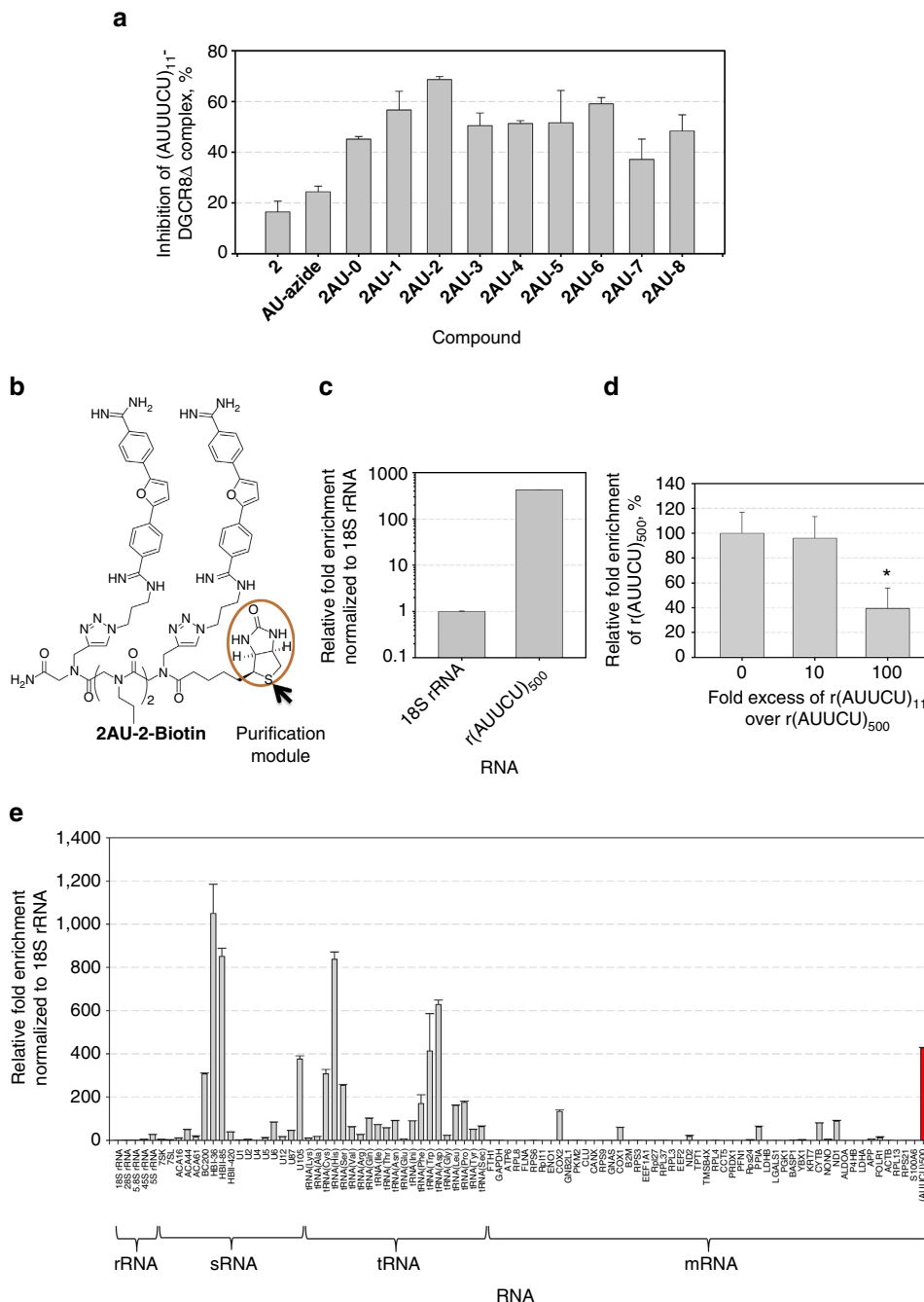


Figure 6 | Design of dimeric compounds to target r(AUUCU)^{exp}. The RNA-binding module, **AU-Azide** (left) was assembled onto various peptoid backbones to afford dimeric compounds (right). The **AU-azides** were separated by different distances by varying the number of spacing modules (*n*) to generate a series of compounds named **2AU-*n***, where **2AU** denotes two RNA-binding modules and *n* denotes the number of spacing modules.



and 100 nM, *vide infra*. Interestingly, 40% of the amino acids in GFP are encoded by seven of the tRNAs pulled down by **2AU-2**. Taken together with the lack of toxicity in healthy and SCA10 patient-derived cells (*vide infra*), these data suggest that binding to tRNAs in a pull-down is not sufficient to elicit a biological effect.

Bioactivity of **2AU-2** in SCA10 patient-derived fibroblasts.

Encouraged by these results, the bioactivity of **2AU-2** was assessed by measuring the ability of the compound to improve three downstream disease-associated defects in SCA10 patient-derived fibroblasts³³. As mentioned above, caspase 3 is abnormally activated in SCA10 fibroblasts by greater than twofold³³. Thus, the effect of **2AU-2** on the caspase-3 activity in control and SCA10 fibroblasts was measured. When SCA10-affected cells were treated with 50 and 100 nM of **2AU-2** for 48 h, caspase-3 activity was reduced to levels observed in healthy cells (Fig. 8a). For comparison, we also tested the monomer **AU-azide** and two dimers with reduced *in vitro* potencies, **2AU-3** and **2AU-4**. In contrast to **2AU-2**, treatment with 100 nM **AU-azide** reduced levels of over-activated caspase-3 activity by only ~30% while **2AU-3** and **2AU-4** were inactive up to 100 nM dosage despite showing activity *in vitro* (Supplementary Fig. 17). These results suggest that **2AU-3** and **2AU-4**, which have suboptimal distances between binding modules, may bind off-targets in cells. Thus, optimal bioactivity of **2AU-2** is a function of valency and the spacing between RNA-binding modules, which affords affinity and selectivity.

In agreement with the downregulated caspase-3 activity, **2AU-2** also reduced the mitochondrial abundance of PKC δ in SCA10 fibroblasts (Fig. 8c). PKC δ translocated to 70–80% of mitochondria in SCA10 cells, whereas only 10–15% mitochondria included PKC δ after treatment with 50 nM **2AU-2**, similar to levels observed in healthy fibroblasts. Next, the ability of **2AU-2** to disrupt formation of nuclear foci was measured. After 48 h incubation, **2AU-2** (50 nM) diminished ca. 70% of nuclear foci in SCA10 cells (Fig. 8b).

To confirm that improvement of SCA10-associated defects was due to binding of the compound to r(AAUCU)^{exp} and not toxicity, we studied the cytotoxicity of **2AU-2** in healthy and SCA10 fibroblasts by measuring released lactate dehydrogenase (LDH)⁴⁵. No significant toxicity of the compound was observed at its active concentration, 50 nM, in either healthy or SCA10 fibroblasts (Supplementary Fig. 18). Thus, the observed downregulation of caspase-3 activity and reduction of the mitochondrial abundance of PKC δ in SCA10 fibroblasts does not result from compound toxicity. Likewise, these results are consistent with our observation that caspase-3 activity is unchanged in healthy fibroblasts upon **2AU-2** treatment.

Further, the observed bioactivity of **2AU-2** is not due to reduced abundance of the mutant *ATXN10* transcript carrying expanded AUUCU repeats as determined by qRT-PCR analysis (Fig. 8d). That is, the compound works at the RNA level, not at the transcriptional level. Interestingly, previous studies have shown that silencing of the *ATXN10* transcript improves

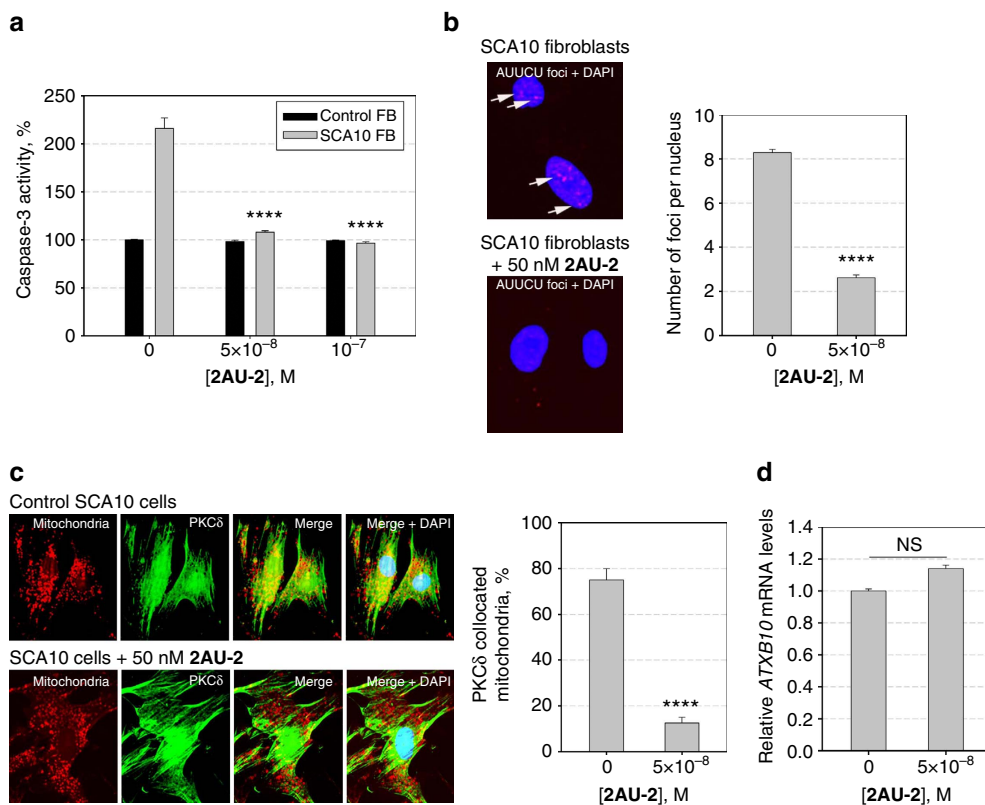


Figure 8 | Studying the bioactivity of **2AU-2 in SCA10 patient-derived fibroblasts.** (a) Relative caspase-3 activities in normal and SCA10 fibroblasts before and after treatment with compound **2AU-2**. Caspase-3 in SCA10 was reduced to normal levels when cells were treated with 50 nM compound. (b) Confocal images showing r(AAUCU)^{exp}-containing nuclear foci in SCA10 cells without treatment (left, top) and after treatment with 50 nM **2AU-2** (left, bottom). Quantification of the number of r(AAUCU)^{exp}-containing nuclear foci (right) in treated and untreated cells. (c) confocal images showing mitochondrial translocation of PKC δ after treatment with 50 nM **2AU-2** (left, bottom) compared with control, untreated SCA10 cells (left, top). Quantification of PKC δ collocated to mitochondria in treated and untreated cells (right). (d) Real-time RT-PCR analysis of *ATXN10* mRNA expression levels after treatment with **2AU-2**. The amount of mRNA was normalized relative to *GAPDH* mRNA. Values are reported as the mean \pm s.e. ($n = 6$). **** $P < 0.0001$ and 'ns' denotes $P > 0.05$, as compared with the untreated sample with a two-tailed Student *t*-test.

SCA10-associated defects³³ and some small molecules that improve microsatellite disease-associated defects work at the transcriptional level⁴⁶.

Taken together, **2AU-2** markedly improves defects from hnRNP K sequestration by r(AUUCU)^{exp} in SCA10 patient-derived cells, suggesting that the compound binds to the cellular target, r(AUUCU)^{exp}, and frees sequestered proteins. **2AU-2** is the most potent inhibitor known for a traditional non-covalent binder to repeat expansions.

Conclusion

To identify selective RNA base pair binders, we screened small molecules that have RNA-binding scaffolds. The *bis*-benzimidine compound **2** is a selective AU base pair binder, as it binds AU base pairs 40-fold more strongly than GC base pairs. This compound was applied to target the pentanucleotide r(AUUCU) expansion that causes SCA10. The repeat periodically displays 5'AUUA3' base pair steps in its secondary structure. To improve affinity and selectivity, we modularly assembled **2** and determined that the optimal distance between RNA-binding modules was afforded by two propylamine spacing modules, or **2AU-2**. **2AU-2** significantly improves SCA10-associated defects to wild-type levels when patient-derived fibroblasts are treated with 50 nM compound. It is the first bioactive small molecule targeting r(AUUCU)^{exp}. The potent bioactivity of **2AU-2** suggests that base pair-targeting RNA modules could have broad utility to provide bioactive compounds targeting other RNAs in the transcriptome.

Methods

Instrumentation. All pH measurements were performed at room temperature using a Mettler Toledo SG2 pH metre that was standardized at pH 4.0, 7.0, and 10.0. Absorption and emission spectra were measured using SpectraMax M5 plate reader (Molecular Devices, Inc.). Sigma Plot (version 11.0) was used for all curve fitting.

Small molecules. All small molecules were procured from the National Cancer Institute (NCI) or The Scripps Research Institute. Emission spectra (excitation: 300 nm, cutoff: 325 nm, emission: 330–600 nm) of all compounds were measured in a 384-well plate (Greiner Low-Volume 784076) to select fluorescent compounds for screening (50 μ M compound in 1 \times Screening Buffer (8 mM Na₂HPO₄, pH 7.0, 185 mM NaCl, 0.1 mM EDTA)).

Compound purification and analysis. Preparative HPLC was performed using a Waters 1525 Binary HPLC pump equipped with a Waters 2487 dual-absorbance detector system and a Waters Sunfire C18 OBD 5- μ m 19 \times 150 mm column. Absorbance was monitored at 220 and 345 nm. A gradient of 20–100% methanol in H₂O with 0.1% trifluoroacetic acid (TFA) over 60 min was used for compound purification. Analytical HPLC was performed using a Waters Symmetry C18 5 μ m 4.6 \times 150 mm column. Compounds were analysed using a gradient of 20–60% MeOH in H₂O with 0.1% TFA over 30 min. All compounds evaluated had \geq 95% purity as determined by analytical HPLC. Mass spectrometry was performed with an Applied Biosystems MALDI ToF/ToF Analyzer 4800 Plus and Microflex (Bruker) using an α -hydroxycinnamic acid matrix. See the Supplementary Information for details of compound synthesis and compound characterization.

Oligonucleotide preparation and purification. The RNAs used in fluorescence binding assays, nuclease mapping, and filter-binding assays were purchased from Dharmacon. The ACE protecting groups were cleaved by using Dharmacon's deprotection buffer by incubating at 60 °C for 30 min. The samples were lyophilized, resuspended in water and gel purified. Concentrations were determined by absorbance using a Beckman Coulter DU800 ultraviolet-visible spectrophotometer at 85 °C. Extinction coefficients (at 260 nm) were calculated using the HyTher server, which uses nearest-neighbour parameters^{47–49}.

Initial screen for small molecules that bind RNA base pairs. RNAs were folded in 1 \times Screening Buffer at 95 °C for 2 min followed by slowly cooling to room temperature on the bench top. A 10- μ l aliquot of a 1 μ M RNA solution was dispensed into each well of a black 384-well plate (Greiner Low-Volume 784076) using an Aurora Discovery FRD-1B liquid dispenser. A 10-nl aliquot of a 2.5 mM stock of small molecule was pinned into each well using Biomek NXP Laboratory Automation Workstation that was equipped with a 384-pin head. The solution was incubated at room temperature for 30 min. Fluorescence intensity was measured

using the maximum excitation/emission wavelength for each compound and the change in fluorescence was calculated by the ratio of F/F_0 where F is the fluorescence intensity in the presence of RNA and F_0 is the fluorescence intensity in the absence of RNA.

Compounds were scored as hits if a >20% change in emission (either enhancement or quenching) was observed upon incubation with RNA. The selectivity of a small molecule was computed by comparing the relative change in emission when incubated different RNAs. Statistically significant differences were calculated by using one-way analysis of variance function in Sigma Plot (version 11.0); compounds that had a P value of <0.05 (95% confidence) were chosen as selective binders.

Chemoinformatic analysis. To identify the chemical substructures that facilitate binding, hit compounds were tested by an automated R-group analysis (Tripod Development; Division of preclinical Innovation, National Center for Advancing Translational Sciences: <http://tripod.nih.gov/?p=46>). The functional groups that provided recognition for each type of RNA were then compared.

Fluorescence binding assays. Direct binding assays for all selective binders were performed. RNAs were folded as described above. Binding assays were performed with a constant compound concentration (1 or 3 μ M) and serial dilutions of RNA or DNA in 1 \times Screening Buffer. For molecular crowding experiments, PEG 8000 was added to a final concentration of 20% (w/v) to the folded RNA and to the solution used for serial dilutions. After a 20 min incubation, fluorescence intensity was measured. The resulting curves were fit to the following equation to determine EC₅₀ values:

$$y = B + \frac{A - B}{1 + \left(\frac{EC_{50}}{x}\right)^{\text{hillslope}}} \quad (1)$$

where y is fluorescence intensity, x is the concentration of RNA, B is the minimum fluorescence; A is the maximum fluorescence; and the EC₅₀ is the concentration of RNA where half of the compound is bound.

Two types of plots were constructed to determine stoichiometries and K_d 's: fluorescence versus [nucleic acid]/[ligand] to determine stoichiometry and fraction-bound/[nucleic acid] versus fraction bound to determine K_d 's. Stoichiometries were determined from the former plots by fitting each of the two slopes (pre-saturated and saturated portions of the curves) to a line⁵⁰. For unsaturated binding curves, the saturated portions of the curves were estimated by the fitted data from equation (1). The point at which the two equations intersect affords the stoichiometry. The K_d 's were determined by fitting fraction bound/[nucleic acid] versus fraction bound to equation (2):

$$\frac{v}{[L]} = \frac{N(1 - lv/N)}{k} \left(\frac{1 - lv/N}{1 - (l-1)v/N} \right)^{l-1} \quad (2)$$

where v is the moles of RNA lattice bound per moles of ligand, $[L]$ is the concentration of ligand, N is the number of repeating units on the RNA, l is the number of consecutive lattice units occupied by the ligand, and k is the microscopic dissociation constant.

Nuclease mapping of the small molecule-binding site. r(AUUCU)₁₁ was radioactively labelled at the 5' end with [γ -³²P] ATP (Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs) using standard methods and gel purified to homogeneity⁵¹. The RNA was folded by incubation at 60 °C for 5 min in 1 \times RNA Structure Buffer (Ambion) followed by slow cooling to room temperature. Serially diluted concentrations of the inhibitor was added to the RNA solution and incubated at room temperature for 15 min. RNase V1 (Ambion) was added to the RNA-inhibitor complex to a final concentration of 5 μ U μ l⁻¹ and the samples were incubated at room temperature for 60 min. RNase V1 was then inactivated by heating at 95 °C for 1 min, and cleavage products were separated on a denaturing 20% polyacrylamide gel. A hydrolysis ladder was prepared by using Alkaline Hydrolysis Buffer (Ambion) and the manufacturer's protocol.

Chemical syntheses of dimeric compounds. Details of compound syntheses and characterization are provided in Supplementary Figs 19–24 and Supplementary Table 4.

Screening of dimeric compounds for binding to r(AUUCU)₁₁ by filter binding.

To determine the optimal distance between RNA-binding modules, a library of dimers was synthesized and screened by using a filter-binding assay. Radioactively labelled r(AUUCU)₁₁ (100 nM) was folded in 1 \times PBS buffer (pH 7.4) containing 1 mM MgCl₂ by incubation at 60 °C for 5 min followed by slow cooling to room temperature. BSA was added to a final concentration of 50 μ g ml⁻¹ followed by addition of 1 μ M compound. The samples were incubated at room temperature for 15 min. Nitrocellulose and nylon membranes were incubated in 1 \times filter binding assay buffer (1 \times PBS buffer (pH 7.4) containing 1 mM MgCl₂ and 50 μ g ml⁻¹ BSA). Bound and unbound RNA were separated using a Dot-Blot apparatus followed by washing with 1 \times filter binding assay buffer. The membranes were

exposed to a phosphorimager screen and imaged using a Molecular Dynamics Typhoon phosphorimager. The amount of r(AUUCU)₁₁ bound to each membrane was quantified using QuantityOne software (BioRad).

Mobility shift assay of r(AUUCU)₁₁ with DGCR8Δ. r(AUUCU)₁₁ was 5'-end labelled as described above and folded in 1 × Folding Buffer (20 mM HEPES, pH 7.5, 110 mM KCl, and 110 mM NaCl) by incubation at 60 °C for 5 min followed by slow cooling to room temperature on the bench top. The buffer was adjusted to 1 × TR-FRET Assay Buffer (20 mM HEPES pH 7.5, 110 mM KCl, 110 mM NaCl, 0.1% BSA, 2 mM MgCl₂, 2 mM CaCl₂, 0.05% Tween-20 and 5 mM DTT) and various concentrations of DGCR8Δ were added. The samples were incubated at room temperature for 15 min and loaded onto a pre-chilled native 5% polyacrylamide gel. The gel was imaged and quantified as described above. The resulting curves were fit to equation (3):

$$y = \frac{x B_{\max}}{K_d + x} \quad (3)$$

where y is percentage of bounded DGCR8Δ, x is the concentration of protein, B_{\max} is maximum percentage of protein bound (restrained to equal 100%) and K_d is dissociation constant.

Determination of compound potency via a TR-FRET assay. TR-FRET assays were completed as previously described²¹ with the following modifications. After folding the RNA, compound was added and incubated for 15 min at room temperature followed by addition of DGCR8Δ. The final concentrations of r(AUUCU)₁₁ and DGCR8Δ were 60 and 40 nM, respectively. TR-FRET was measured after an additional 30-min incubation at room temperature. IC₅₀ values were calculated by curve fitting using equation (1).

Pull-down of 2AU-2's cellular targets. HeLa cells were maintained as monolayers in growth medium (1 × DMEM, 10% fetal bovine serum and 1 × GlutaMax (Invitrogen)). Cells were plated in 10-cm dishes and grown to ~90% confluency and then transfected with a plasmid encoding (AUUCU)₅₀₀ using Lipofectamine 2,000 (Invitrogen) per the manufacturer's recommended protocol. Cells were collected 18–24 h post-transfection and total RNA was extracted by using Trizol reagent (Ambion) according to the manufacturer's protocol. After RQ1 DNase (Promega) treatment, the DNase was removed by phenol:chloroform extraction and total RNA was ethanol precipitated. Next, 100 μg of total RNA was folded in 1 × Screening Buffer by heating at 75 °C for 5 min and cooling to room temperature slowly. The folded RNAs were incubated with 2AU-2-Biotin for 30 min at room temperature. The solution was then incubated with streptavidin beads (250 μl of slurry, Sigma-Aldrich) for 30 min at room temperature with gentle shaking. The supernatant (containing unbound RNAs) was removed, and the beads were washed with 250 μl 1 × Screening Buffer after gentle shaking for 5 min at room temperature. Bound RNA was released from the beads by heating the beads in 300 μl H₂O at 80 °C for 3 min twice. The solution containing bound RNAs was concentrated to 5–50 μl by vacuum concentration. Complementary DNA (cDNA) was generated from 40 ng of RNA using a qScript cDNA Synthesis Kit (Quanta Biosciences) per the manufacturer's protocol. qPCR was performed on an ABI 7900 HT Real-Time PCR System using the following primers to detect the r(AUUCU)₅₀₀-containing RNA: 5'-AGTCTCTCT ATGTTGCCAGG-3' and 5'-ACTTCCCGAAACACCGTCTC-3'. The relative fold enrichment of the RNA pulled by the compound was calculated by normalization to 18S rRNA.

Profiling of cellular RNAs pulled down by 2AU-2-Biotin. Profiling was completed by qRT-PCR as described above. The 93 highly abundant RNAs, including rRNAs, sRNA, tRNAs and mRNAs, were selected based on a previous report⁴². Primer sets for mRNAs were obtained from RTPrimerDB⁵² (www.rtpriimerdb.org) and qPrimerDepot (<http://primerdepot.nci.nih.gov>). The DNA sequences of tRNAs were obtained from tRNAdb⁵³ (<http://trna.bioinf.uni-leipzig.de/>) and primer sets were designed by using Primer 3 software (<http://frodo.wi.mit.edu/primer3/>). All sequences of primers are listed in Supplementary Table 3.

Cell culture of SCA10 fibroblasts. SCA10 fibroblasts³³ were cultured in MEM with Eagle-Earle salt and 2 mM L-glutamine containing 15% fetal bovine serum and antibiotic in 5% CO₂ at 37 °C in 75-cm² flasks. Compounds were dissolved in 50% DMSO and added to the SCA10 cells at the indicated concentrations. Fresh medium with drug was added to the cells after every 12 h and cells were collected after 48 h for the caspase-3 assay, detection of AUUCU RNA foci, or for analysing the subcellular translocation of protein kinase C δ (PKCδ).

Translocation of PKCδ after drug treatment. SCA10 fibroblasts (2 × 10⁴ cells) were seeded in chamber slides. When cells were 80–90% confluent, fresh cell culture medium containing compound was added to the cells and incubated for 48 h. The drug-treated and control cells were then incubated with mitotracker deep red 633 (Invitrogen, USA) at a concentration of 250 nM in cell culture medium and

incubated at 37 °C for 30 min. The cells were then washed three times with ice-cold 1 × PBS, fixed with 4% paraformaldehyde for 30 min at room temperature, washed three times with 1 × PBS and stored in 70% ethanol for up to 24 h. Cells were blocked with DAKO antibody blocking solution (serum free) and later double stained with anti-PKCδ 1:500 in DAKO antibody diluent. Goat anti-mouse 488 was used to identify PKCδ. Fluorescent photomicrographs were taken using an Hamamatsu Camera Controller using DP controller software.

Fluorescent in situ hybridization to detect AUUCU RNA foci. SCA10 and control fibroblasts (2 × 10⁴ cells) were seeded in chamber slides. When the cells were 60–70% confluent, cell culture medium containing compound was added to the cells and incubated for 48 h. After incubating with compound, the cells were fixed with 4% paraformaldehyde for 30 min at room temperature, and washed three to four times with ice-cold 1 × PBS. The AUUCU RNA foci were detected using a Cy3-labelled (AGAAU)₁₀ RNA oligonucleotide probe as described previously³³. In brief, the control and SCA10 cells were pre-hybridized at 65 °C in RNA Hybridization Buffer for 1.5 h, and hybridized overnight in hybridization buffer containing 250 ng of (AGAAU)₁₀ Cy3-labelled RNA oligo at 45 °C. Slides were rinsed with 1 × PBS three times and extensively washed 4 × 5 min to remove all non-specific binding of Cy3-labelled RNA probes. Slides were then mounted with DAPI mounting medium and fluorescent images were taken using a confocal microscope.

Cytotoxicity of 2AU-2. SCA10 and healthy fibroblasts were treated either with 2AU-2 (50, 100 and 500 nM) or vehicle for 24 h. Later, the culture medium containing LDH was collected and the amount of LDH in the medium was quantified using TOX7 cell toxicity assay kit (Sigma-Aldrich).

Data availability. Data supporting the findings of this study are available within the article and its supplementary information files and from the corresponding author upon request.

References

- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Tenson, T. & Mankin, A. Antibiotics and the ribosome. *Mol. Microbiol.* **59**, 1664–1677 (2006).
- Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956 (2002).
- Winkler, W. C., Cohen-Chalamish, S. & Breaker, R. R. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl Acad. Sci. USA* **99**, 15908–15913 (2002).
- Harries, L. W. Long non-coding RNAs and human disease. *Biochem. Soc. Trans.* **40**, 902–906 (2012).
- Shen, L. X., Basilion, J. P. & Stanton, Jr. V. P. Single nucleotide polymorphisms can cause different structural folds of mRNA. *Proc. Natl Acad. Sci. USA* **96**, 7871–7876 (1999).
- Sierakowska, H., Agrawal, S. & Kole, R. Antisense oligonucleotides as modulators of pre-mRNA splicing. *Methods Mol. Biol.* **133**, 223–233 (2000).
- Brunberg, J. A. *et al.* Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *Am. J. Neuroradiol.* **23**, 1757–1766 (2002).
- Brook, D. *et al.* Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 (1992).
- Matsuura, T. *et al.* Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat. Genet.* **26**, 191–194 (2000).
- Liquori, C. L. *et al.* Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867 (2001).
- The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–983 (1993).
- Gallego, J. & Varani, G. Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. *Acc. Chem. Res.* **34**, 836–843 (2001).
- Poehlsgaard, J. & Douthwaite, S. The bacterial ribosome as a target for antibiotics. *Nat. Rev. Microbiol.* **3**, 870–881 (2005).
- Johnson, L. F., Williams, J. G., Abelson, H. T., Green, H. & Penman, S. Changes in RNA in relation to growth of the fibroblast. III. Posttranscriptional regulation of mRNA formation in resting and growing cells. *Cell* **4**, 69–75 (1975).
- Blount, K. F., Wang, J. X., Lim, J., Sudarsan, N. & Breaker, R. R. Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.* **3**, 44–49 (2007).
- Schramm, V. L. Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes. *Annu. Rev. Biochem.* **80**, 703–732 (2011).

18. Childs-Disney, J. L., Hoskins, J., Rzuczek, S., Thornton, C. & Disney, M. D. Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol.* **7**, 856–862 (2012).
19. Lee, M. M., Pushechnikov, A. & Disney, M. D. Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* **4**, 345–355 (2009).
20. Parkesh, R. *et al.* Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database & chemical similarity searching. *J. Am. Chem. Soc.* **134**, 4731–4742 (2012).
21. Disney, M. D. *et al.* A small molecule that targets r(CGG)^{exp} and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem. Biol.* **7**, 1711–1718 (2012).
22. Kumar, A. *et al.* Chemical correction of pre-mRNA splicing defects associated with sequestration of muscleblind-like 1 protein by expanded r(CAG)-containing transcripts. *ACS Chem. Biol.* **7**, 496–505 (2012).
23. Thomas, J. R. & Hergenrother, P. J. Targeting RNA with small molecules. *Chem. Rev.* **108**, 1171–1224 (2008).
24. Guan, L. & Disney, M. D. Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* **7**, 73–86 (2012).
25. Mathews, D. H. *et al.* Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl Acad. Sci. USA* **101**, 7287–7292 (2004).
26. Carlson, C. B., Vuyisich, M., Gooch, B. D. & Beal, P. A. Preferred RNA binding sites for a threading intercalator revealed by in vitro evolution. *Chem. Biol.* **10**, 663–672 (2003).
27. Luedtke, N. W. *et al.* The DNA and RNA specificity of eilatin Ru(II) complexes as compared to eilatin and ethidium bromide. *Nucleic Acids Res.* **31**, 5732–5740 (2003).
28. Jin, E. *et al.* Aminoglycoside binding in the major groove of duplex RNA: the thermodynamic and electrostatic forces that govern recognition. *J. Mol. Biol.* **298**, 95–110 (2000).
29. Wilson, W. D., Ratmeyer, L., Zhao, M., Strekowski, L. & Boykin, D. The search for structure-specific nucleic acid-interactive drugs: effects of compound structure on RNA versus DNA interaction strength. *Biochemistry* **32**, 4098–4104 (1993).
30. Zhao, M. *et al.* Small changes in cationic substituents of diphenylfuran derivatives have major effects on the binding affinity and the binding mode with RNA helical duplexes. *Bioorg. Med. Chem.* **3**, 785–794 (1995).
31. Handa, V., Yeh, H. J. C., McPhie, P. & Usdin, K. The AUUCU repeats responsible for spinocerebellar ataxia type 10 form unusual RNA hairpins. *J. Biol. Chem.* **280**, 29340–29345 (2005).
32. Teive, H. A. G. *et al.* Spinocerebellar ataxia type 10 – A review. *Parkinsonism Relat. Disord.* **17**, 655–661 (2011).
33. White, M. C. *et al.* Inactivation of hnRNP K by expanded intronic AUUCU repeat induces apoptosis via translocation of PKC δ to mitochondria in spinocerebellar ataxia 10. *PLoS Genet.* **6**, e1000984 (2010).
34. Jin, P., Alisch, R. S. & Warren, S. T. RNA and microRNAs in fragile X mental retardation. *Nat. Cell Biol.* **6**, 1048–1053 (2004).
35. Pushechnikov, A. *et al.* Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3. *J. Am. Chem. Soc.* **131**, 9767–9779 (2009).
36. Childs-Disney, J. L., Tsitovich, P. B. & Disney, M. D. Using modularly assembled ligands to bind RNA internal loops separated by different distances. *Chembiochem* **12**, 2143–2146 (2011).
37. Lee, M. M. *et al.* Controlling the specificity of modularly assembled small molecules for RNA via ligand module spacing: targeting the RNAs that cause myotonic muscular dystrophy. *J. Am. Chem. Soc.* **131**, 17464–17472 (2009).
38. Rzuczek, S. G. *et al.* Features of modularly assembled compounds that impart bioactivity against an RNA target. *ACS Chem. Biol.* **8**, 2312–2321 (2013).
39. Fu, H. Y. & Doucet, H. Methyl 2-furoate: an alternative reagent to furan for palladium-catalysed direct arylation. *Eur. J. Org. Chem.* **2011**, 7163–7173 (2011).
40. Das, B. P. & Boykin, D. W. Synthesis and antiprotozoal activity of 2,5-bis(4-guanylphenyl)furans. *J. Med. Chem.* **20**, 531–536 (1977).
41. Roth, B. M., Ishimaru, D. & Hennig, M. The core microprocessor component DiGeorge syndrome critical region 8 (DGCR8) is a nonspecific RNA-binding protein. *J. Biol. Chem.* **288**, 26785–26799 (2013).
42. Nagaraj, N. *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **7**, 548 (2011).
43. Galiveti, C. R., Rozhdestvensky, T. S., Brosius, J., Lehrach, H. & Konthur, Z. Application of housekeeping npcRNAs for quantitative expression analysis of human transcriptome by real-time PCR. *RNA* **16**, 450–461 (2010).
44. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
45. Decker, T. & Lohmann-Matthes, M.-L. A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor (TNF) activity. *J. Immunol. Methods* **115**, 61–69 (1988).
46. Coonrod, L. A. *et al.* Reducing levels of toxic RNA with small molecules. *ACS Chem. Biol.* **8**, 2528–2537 (2013).
47. Peyret, N., Seneviratne, P. A., Allawi, H. T. & SantaLucia, Jr. J. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry* **38**, 3468–3477 (1999).
48. SantaLucia, Jr. J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA* **95**, 1460–1465 (1998).
49. Puglisi, J. D. & Tinoco, Jr. I. Absorbance melting curves of RNA. *Methods Enzymol.* **180**, 304–325 (1989).
50. Tse, W. C. & Boger, D. L. A fluorescent intercalator displacement assay for establishing DNA binding selectivity and affinity. *Acc. Chem. Res.* **37**, 61–69 (2004).
51. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* 2nd edn. (Cold Spring Harbor Laboratory, 1989).
52. Lefever, S., Vandesompele, J., Speleman, F. & Pattyn, F. RTPrimerDB: the portal for real-time PCR primers and probes. *Nucleic Acids Res.* **37**, D942–D945 (2009).
53. Juhling, F. *et al.* tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* **37**, D159–D162 (2009).

Acknowledgements

This work was funded by the National Institutes of Health (DP1 NS096898) to M.D.D., a John Sealy Grant to P.S.S., a FRAXA postdoctoral fellowship to W.-Y.Y. and The Scripps Research Institute. M.D.D. is a Camille & Henry Dreyfus Teacher-Scholar.

Author contributions

M.D.D. provided conceptual framework. M.D.D., W.-Y.Y., R.G. and P.S.S. designed the experiments. W.-Y.Y. and R.G. performed the experiments and statistical analysis. M.S. performed chemical similarity searching. M.D.D., W.-Y.Y., R.G. and P.S.S. wrote the manuscript. All authors reviewed the final version of the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Yang, W.-Y. *et al.* Design of a bioactive small molecule that targets r(AUUCU) repeats in spinocerebellar ataxia 10. *Nat. Commun.* **7**:11647 doi: 10.1038/ncomms11647 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>