High-throughput screening yields several small-molecule inhibitors of repeat-associated non-AUG translation

Katelyn M Green[1,2], Udit Sheth[1], Brittany N Flores[1,2], Shannon E Wright[1], Alexandra Sutter[1], Michael G Kearse[1,3], Sami Barmada[1,2], Magdalena I Ivanova[1,4], Peter K Todd[1,2,5*]

From the [1]Department of Neurology, University of Michigan; [2]Cellular and Molecular Biology Graduate Program, University of Michigan; [3]Department of Biological Chemistry and Pharmacology, Center for RNA Biology, The Ohio State University; [4]Biophysics Program, University of Michigan; [5]VA Ann Arbor Healthcare System, Ann Arbor, MI

Running title: *Small molecule inhibitors of RAN translation*

*please send requests to:
Peter K Todd, MD, PhD
Department of Neurology
University of Michigan
4005 BSRB
109 Zina Pitcher Place
Ann Arbor, MI 48109
petertod@umich.edu

_____

ABSTRACT

Repeat-associated non-AUG (RAN) translation is a non-canonical translation initiation event that occurs at nucleotide-repeat expansion mutations that are associated with several neurodegenerative diseases, including fragile X-associated tremor ataxia syndrome (FXTAS), amyotrophic lateral sclerosis (ALS), and frontotemporal dementia (FTD). Translation of expanded repeats produces toxic proteins that accumulate in human brains and contribute to disease pathogenesis. Consequently, RAN translation constitutes a potentially important therapeutic target for managing multiple neurodegenerative disorders. Here, we adapted a previously developed RAN translation assay to a high-throughput format to screen 3253 bioactive compounds for inhibition of RAN translation of expanded CGG repeats associated with FXTAS. We identified five diverse small molecules that dose-dependently inhibited CGG RAN translation, while relatively sparing canonical translation. All five compounds also inhibited RAN translation of expanded GGGGCC repeats associated with ALS and FTD. Using circular dichroism and native gel analyses, we found evidence that three of these compounds, BIX01294, CP-31398, and propidium iodide, bind directly to the repeat RNAs. These findings provide proof-of-principle supporting the development of selective small-molecule RAN translation inhibitors that act across multiple disease-causing repeats.

_____

Repeat-associated non-AUG (RAN) translation has recently emerged as a common pathogenic mechanism among nucleotide repeat expansion diseases. Through this process, the ribosome initiates translation upstream of or within stable secondary structures formed by repetitive RNA sequences, in the absence of an AUG start codon (1). This leads to production of RAN peptides; often aggregate-prone proteins with repetitive amino acid sequences resulting from translation of the repetitive RNA element.

To date, RAN translation is known to occur in at least nine different human diseases: fragile X-

associated tremor ataxia syndrome (FXTAS) (2,3), amyotrophic lateral sclerosis (ALS) (4-7), frontotemporal dementia (FTD) (4-7), spinocerebellar ataxia type 8 (1), myotonic dystrophy types 1 and 2 (1,8), Huntington's disease (9), fragile X-associated primary ovarian insufficiency syndrome (10), and Fuchs' endothelial corneal dystrophy (11). These diseases are caused by multiple different microsatellite repeat sequences located within various regions (e.g., coding sequence, UTRs, intron) of different genes.

FXTAS is a late-onset neurodegenerative disease characterized by difficulty walking, loss of fine motor skills, and progressive cognitive and behavioral changes (12), that affects up to 1 in 3,000 men over the age of 50 (13). FXTAS results from the expansion of CGG repeats within the 5′ leader of *FMR1*, from an average of 30 or fewer CGGs, to 55-200 (14). RAN translation of this repeat in the +1 (GGC) reading frame generates a polyglycine protein, FMRPolyG, that is found within inclusions in patient neurons (2,15,16). When overexpressed from a CGG repeat in *Drosophila* or mice, FMRPolyG is neurotoxic, causing motor deficits, neurodegeneration and reduced lifespans (2,15).

ALS and FTD are also neurodegenerative diseases. ALS is the most common form of motor neuron disease, and FTD the second most common form of early-onset dementia. The most common genetic cause of both ALS and FTD in an expansion of GGGGCC repeats within the first intron of *C9orf72* (17-19). RAN translation in all three reading frames of the expanded GGGGCC repeat produces three different dipeptide repeat proteins (DPRs); glycine-alanine (GA: GGG-GCC), glycine-proline (GP: GGG-CCG), and glycine-arginine (GR: GGC-CGG). All three DPRs are detected in patient neurons (4,5,7), and the GR product has repeatedly been shown to be highly toxic across model systems (20-22), with more moderate toxicity from the GA product also reported (20,23-25).

Despite differences in repeat sequence and context, our lab and others have shown that RAN translation from reporter constructs with FXTAS-associated CGG repeats (CGG RAN) and C9ALS/FTD-associated GGGGCC repeats (C9RAN) share common mechanistic features (26-29). RAN translation in all three reading frames

across both repeats occurs most efficiently when the ribosome is able to bind at the 5′ cap and scan in a 5′ to 3′ direction (26-29). Additionally, for both repeats, RAN translation in the reading frame that generates the most abundant RAN peptide in patient tissue (FMRPolyG and GA) (2,30), predominantly utilizes a near-AUG codon upstream of the repeat for initiation (26-28,31). However, evidence also suggests that initiation can occur, at lower levels, within the repeat sequence itself and in a cap-independent manner (1,29,31). Intriguingly, activation of the integrated stress response and phosphorylation of the initiation factor eIF2α, enhances RAN translation of both the CGG and GGGGCC repeats in model systems (27,29,31,32)

To better define the mechanism of RAN translation and identify potential inhibitors of this process, we developed an *in vitro*, reporter-based small molecule screen for bioactive compounds that selectively and dose-dependently inhibit RAN translation of CGG repeats. From this screen, we identified five novel CGG RAN translation inhibitors, and found that each compound also inhibits RAN translation of C9ALS/FTD-associated GGGGCC repeats, in multiple sense-strand reading frames. Using circular dichroism and native gel analysis, we show that three of these compounds bind the repeats and alter their secondary structure in a manner similar to previously developed small molecule inhibitors (33-35). These studies establish that RAN translation from multiple repeat expansions, and across multiple reading frames, can be selectively targeted by small molecule inhibitors, providing insights into the mechanisms by which RAN translation occurs and establishing a framework for future therapeutic development in nucleotide repeat expansion disorders.

RESULTS
Primary screen of 3253 bioactive compounds for inhibitors of CGG RAN translation

To identify small molecule inhibitors of RAN translation of FXTAS-associated CGG repeats, we adapted a previously developed *in vitro* RAN translation assay (26) to a 384-well format (Fig. 1A). This assay utilized a +1CGGx100 RAN translation nanoluciferase (NLuc) reporter mRNA, which was added to wells containing rabbit reticulocyte lysate (RRL) treated with one of 3253

2

bioactive compounds at 20 µM (Fig. 1A). After a 30-minute incubation at 30°C, luminescence was measured as a readout for production of the neurotoxic polyglycine RAN protein (FMRpolyG) in each well (Fig. 1A). We confirmed that under these conditions, the reporter mRNA produced robust RAN-specific luminescent signal within the dynamic linear range of detection (Fig. 1B & C, Supplementary Fig. 1A).

The 3253 compounds screened with this assay came from the Pilot LOPAC, Pilot Prestwick, Pilot NCC-focused and Navigator Pathways small molecule libraries. These libraries contain pharmacologically active compounds, small molecules previously tested in clinical trials, and FDA-approved drugs, all in DMSO. As a positive control for translation inhibition, the last two columns of each plate were treated with 3 µM cycloheximide, and as an internal negative control, the first two columns were treated with DMSO (Fig. 1A). The average Z-factor and coefficient of variability (CV) value, calculated per plate, was 0.79 and 6.93%, respectively (Table 1), indicating that the screen was of high quality. This allowed us to identify 289 "hits," which we defined as compounds that reduced +1CGG RAN translation by greater than 20% or 3 standard deviations relative to vehicle-treated controls (Fig. 1D, Table 2, and Supplementary Table 1). These hits had diverse structures and annotated biological targets. While 14 known ribosomal inhibitors and 17 DNA/RNA intercalators were identified, the majority of hits represented other classes of compounds, acting through less obvious modes of translational inhibition (Fig. 1E).

Concentration response and counter screen

From 289 primary hits, we selected 110 inhibitors to advance to a secondary concentration response curve analysis (Fig. 2A). Compounds were selected primarily based on their percent inhibition and how many standard deviations their effect differed from vehicle-only controls. Preference was given to compounds that were FDA-approved or that were hits multiple times when present in multiple libraries. Compounds that were flagged for concerns about toxicity, active in greater than 90% of luciferase assays previously performed at our screening facility or known translation inhibitors of a mechanism already represented among selected hits, were excluded

unless of particular interest. Additionally, a standard deviation cut-off of 2.5 was applied to compounds added to wells at the edge of the plate, where signal was more variable.

Selected compounds were added to the RRL at 8 concentrations from 50 to 8 µM, in duplicate. To obtain a 50 µM dose, the volume of compound added to each well increased the final percentage of DMSO in the reaction mixture from 1% in the primary screen, to 2%. We confirmed that at this level, DMSO did not significantly inhibit +1CGG RAN translation on its own (Supplementary Fig. 2A). Of the 110 compounds selected, 77 exhibited dose-dependent inhibition of +1CGG RAN translation, while 33 did not (Fig. 2A and Supplementary Table 2). Among the 33 compounds that failed to validate, only three reduced signal in the primary screen by both more than 20% and 4 standard deviations relative to controls (Supplementary Table 2).

To eliminate compounds that non-specifically inhibited all translation, with no specificity for RAN translation, the concentration response curve assay was simultaneously performed with an NLuc reporter mRNA lacking any repeat element and translated using a canonical AUG start codon (AUG-NLuc) (Supplementary Fig. 2B and C). Of the 77 validated hits, 54 also inhibited translation of the AUG-initiated canonical reporter in a dose-dependent manner. Consequently, 23 compounds exclusively and dose-dependently inhibited +1CGG RAN translation, which were called "RAN specific inhibitors" (Fig. 2B and Supplementary Table 2). None of these compounds were of high potency; the IC50 of the most potent, amlexanox, was 26.3 µM (Supplementary Table 2). An additional 30 compounds were more potent inhibitors of +1CGG RAN translation than canonical translation (Fig. 2C-F and Supplementary Table 2). These were called "RAN selective inhibitors." As a group, the RAN selective inhibitors were more potent than the RAN specific inhibitors, with IC50s ranging from 0.72 to 81.3 µM (Supplementary Table 2).

Of the remaining 24 compounds that reduced both +1CGG RAN and AUG-NLuc translation, 16 had similar activity with both, while 8 were more potent inhibitors of canonical translation (Supplementary Table 2).

3

Five bioactive compounds selectively inhibit CGG RAN translation in multiple reading frames

Based on results from the concentration response and counter screen, 22 compounds were selected for independent validation, including 13 RAN specific and nine RAN selective inhibitors (Fig. 2A and Table 3). Each compound was added at 4-5 doses to a 10 µL RRL *in vitro* translation assay using the same reporter mRNAs as in the primary and counter screens. Of these compounds, only one RAN specific and four RAN selective inhibitors significantly and dose-dependently inhibited +1CGG RAN translation, while leaving translation of the canonical AUG-NLuc reporter relatively spared (Table 3). These compounds included cephalothin, a discontinued beta-lactam antibiotic; BIX01294, a histone lysine methyltransferase inhibitor; anthralin, an FDA-approved drug used in topical treatments for psoriasis; propidium iodide, a fluorescent nucleic acid intercalating agent; and CP-31398, a p53 stabilizer (Fig. 3A-F). As a group, these compounds have diverse biological targets and structures. However, interestingly, BIX01294 and CP-31398 share a similar functional group (Fig. 3A).

The remaining 17 compounds either had no effect on RAN translation or also inhibited canonical translation to a similar degree during independent validation (Supplementary Fig. 3A-D, Table 3). To assess the possibility that results observed in the primary screen for these 17 compounds came from a by-product produced through extended storage durations, we incubated a subset at 37°C for 1 week. However, this did not substantially affect their inhibitory properties (data not shown).

RAN translation of the expanded CGG repeat also occurs in the +2 reading frame (GCG), generating a polyalanine protein (FMRPolyA) (2). However, unlike initiation in the +1 reading frame, initiation in the +2 frame does not utilize a near-AUG codon, and likely occurs within the repeat sequence itself (26). We were therefore interested to see if the five inhibitors of +1CGG RAN translation also reduced +2CGG RAN translation. Cephalothin, BIX01294, anthralin, and CP-31398 all significantly inhibited translation from a +2CGGx100 RAN translation reporter mRNA (26), relative to the canonical AUG-NLuc control (Fig. 3B-D and F). However, at the doses tested, propidium iodide did not (Fig. 3E). Consequently,

RAN translation resulting in the synthesis of distinct polypeptides, from initiation at different codons in different reading frames, can be inhibited simultaneously by the same small molecule, although this effect cannot be generalized to all inhibitors.

We also assessed the inhibitory activity of two additional small molecules, protoporphyrin IX (PPIX) and TMPyP4. PPIX is a natural compound that consists of porphyrin ring. It is known to bind g-quadruplex structures and increases levels of FMRP in fragile-X patient cells (36). TMPyP4 is structurally similar to PPIX, also consisting of a porphyrin ring, and known to bind CGG and GGGGCC repeat RNA (37,38). When tested, both compounds selectively inhibited CGG RAN translation, relative to the AUG-NLuc control (Fig. 3G and H). However, unlike PPIX, TMPyP4 only had relative selectivity for CGG RAN translation in the +1, but not +2, reading frame for the doses tested, much like what we observed with propidium iodide (Fig. 3H).

Inhibitors of CGG RAN translation also inhibit C9RAN translation

Recent work suggests mechanistic similarity between CGG RAN translation and C9ALS/FTD-associated GGGGCC RAN translation (C9RAN translation) (27-29). We therefore tested if these small molecule inhibitors of CGG RAN translation also impacted C9RAN translation. To accomplish this, we repeated the RRL *in vitro* translation assays using previously developed C9RAN translation-specific NLuc reporter mRNAs that contained 70 GGGGCC repeats (27). All five compounds identified in the CGG RAN translation screen, as well as PPIX, selectively and dose-dependently decreased C9RAN translation in multiple reading frames (GA, GP, and GR), relative to the AUG-NLuc control (Fig. 4A-C, Supplementary Fig. 4A-C). We further confirmed the decrease in luminescence from the CGG and C9RAN reporter mRNAs was due to a decrease in RAN polypeptide synthesis by western blot. BIX01294, anthralin, and PPIX all decreased poly-GA and FMRPolyG synthesis in RRL, in a dose-dependent manner, while leaving canonical NLuc synthesis relatively spared (Fig. 4D-F).

4

Inhibitors target different aspects of RAN translation

To determine if these small molecules inhibit RAN translation by targeting its non-AUG initiation, we repeated the *in vitro* translation reactions using reporters with AUG start codons inserted upstream of the expanded CGG repeats in either the +1 or +2 reading frames (26) (Fig. 5A and Supplementary Fig. 5A). Anthralin, TMPyP4, and PPIX all inhibited AUG +1CGGx100 and AUG +2CGGx100 translation to a similar or greater extent than +1CGG and +2CGG RAN translation (Fig. 5B-D). Consequently, the selectivity of these compounds for translation of expanded CGG repeats, relative to AUG-NLuc translation, was maintained regardless of whether repeat translation initiated with an AUG or non-AUG codon (Fig. 5B-D). Conversely, BIX01294 and CP-31398 inhibition was markedly decreased when translation of the expanded repeats initiated with an AUG start codon (Fig. 5E and F), indicating that their activity is dependent upon a non-AUG initiation event. Therefore, these five small molecules fall within two distinct groups that produce similar inhibitory effects by targeting different aspects of RAN translation.

To assess whether the inhibitory activity of BIX01294 and CP-31398 on non-AUG initiated-translation was specific to mRNAs containing a repeat, we utilized two additional NLuc reporters; one that initiates at an ACG and the other at a CUG start codon (ACG-NLuc and CUG-NLuc), both lacking a repetitive element (39) (Supplementary Fig. 5B). We specifically chose these two non-AUG start codons as they were previously shown to be used for FMRpolyG and poly-GA synthesis, respectively (26-28). Interestingly, relative to the canonical AUG-NLuc control, BIX01294 and CP-31398 more significantly inhibited translation from both the ACG-NLuc and CUG-NLuc reporters (Supplementary Fig. 5C and D). Therefore, a repeat is not required for translation inhibition by either compound. This supports the idea that non-AUG-initiated translation can be sufficiently distinct from canonical AUG-initiated translation to allow for selective targeting (40).

In contrast, anthralin, PPIX and TMPyP4 all failed to inhibit translation of ACG-NLuc and CUG-NLuc reporter mRNAs, relative to the AUG-NLuc control (Supplementary Fig. 5E-G), indicating that their inhibitory effects are repeat-

dependent. Furthermore, as expression of the near-AUG initiated reporters is significantly lower than the AUG-initiated control (Supplementary Fig. 5B), this suggests the relatively low expression of the repeat-containing reporters does not alone account for their selective inhibition by these compounds.

BIX01294 interacts with repeat RNAs

To test if the RAN translation inhibitors directly interact with the repeat RNAs, as a possible mechanism for their inhibition, we utilized circular dichroism (CD). First, we attained the CD spectrum of CGGx16 repeat RNA folded in the presence of 100 mM KCl (the concentration present in the RRL reactions) by heating to 95ºC and returning to room temperature. UUUx16 RNA, folded under the same conditions, was used as a control lacking both sequence and structural similarity. We then incubated each RNA with 25 and 50 µM TMPyP4, as a positive control for CGG repeat RNA binding. Consistent with previous findings (37,41), TMPyP4 caused a dose-dependent shift in the CGGx16 RNA CD spectrum, while a higher dose of was required to alter UUUx16's CD spectrum (Fig. 6A and B).

We next incubated the RNAs with increasing amounts of BIX01294 and anthralin. At both 50 and 100 µM, BIX01294 shifted CGGx16's CD spectrum, but had little effect on UUUx16's (Fig. 6C and D). Conversely, incubation with 100 or 200 µM anthralin had no effect on the CD spectra of either CGGx16 or UUUx16. Higher doses of anthralin were not tested by CD, as increasing concentrations of DMSO interfered with the absorbance spectra.

To determine if these compounds interact similarly with the *C9ORF72*-associated GGGGCC repeat, we performed CD using a GGGGCCx8 repeat RNA. This RNA was folded at room temperature following heating to 95°C in the presence of either 100 mM KCl, to promote the formation of a g-quadruplex structure, or in the presence of 100 mM NaCl, to promote hairpin formation (33,38). Regardless of the cation used, TMPyP4 and BIX01294 shifted the GGGGCCx8 RNA CD spectra, while anthralin did not (Fig. 7), mirroring their interactions with the CGGx16 RNA.

As a secondary measure of compound interaction with CGG repeat RNA, we utilized native gel electrophoresis to assay for an electrophoretic shift indicative of changes in the

5

RNA structure due to compound binding. For this, we used 5′ IRdye 800CW-labeled CGGx16 and UUUx16 RNAs folded in the presence of 100 mM KCl. 50 nM of each RNA was incubated with increasing concentrations of each inhibitor and run on a 12% native polyacrylamide gel. The labeled RNAs were then visualized at 800nm.

TMPyP4 was again used as a positive control for direct binding to the CGG repeat RNA (37). Increasing concentration of TMPyP4 led to a dose-dependent decrease in CGGx16's band intensity, despite equal RNA loading (Supplementary Fig. 6A). This differs from a previous report using a radio-labeled CGG RNA with extensive sequence 5′ to the repeat (41), and may be the result of TMPyP4 binding interfering with detection of the RNA's IRdye. A similar result was obtained when TMPyP4 was incubated with the UUUx16 RNA (Supplementary Fig. 6A).

BIX01294, propidium iodide, and CP-31398 all had a similar effect on the CGGx16 RNA as TMPyP4, causing a dose-dependent decrease in band intensity (Supplementary Fig. 6B-D). Additionally, they showed a preferential effect, as lower doses of each compound were required to reduce the intensity of the CGGx16 band relative to those needed to reduce intensity of the UUUx16 band (Supplementary Fig. 6B-D). However, only incubation with increasing concentrations of CP-31398 resulted in a clear upshift indicative of altered RNA secondary structure (Supplementary Fig. 6D).

PPIX also modestly, but preferentially decreased the CGGx16 RNA band intensity in a dose-dependent manner (Supplementary Fig. 6E), while anthralin and cephalothin showed little evidence of interacting with either CGGx16 or UUUx16 RNA, even at excessively high doses (Supplementary Fig. 6F and G).

Therefore, by CD and/or native gel analysis, BIX01294 and CP-31398 exhibit preferential interaction with CG-rich repeats RNAs. Native gel results suggest a possible interaction between propidium iodide and CGGx16 RNA, while anthralin and cephalothin show no evidence of interaction. This may indicate that these sets of compounds utilize different mechanisms for inhibiting CGG RAN translation.

BIX01294 inhibits CGG RAN in cultured cells, but is toxic

We next asked whether these small molecule inhibitors of RAN translation were active in cultured cells. To do so, we applied BIX01294 to HEK293 cells transfected with plasmids expressing either AUG-NLuc, +1CGGx100-NLuc, or GA70-NLuc reporters. As an internal control for transfection efficiency and overall cellular health, all wells were co-transfected with a firefly luciferase reporter. Twenty-four hours post treatment, luciferase activity was measured as a readout for translation of each reporter.

Relative to vehicle-treated controls, 25 µM BIX01294 significantly reduced +1CGG RAN translation, without affecting AUG-NLuc expression (Fig. 8). However, as assessed by changes in cellular morphology (data not shown), and a reduction in firefly luciferase signal (Supplementary Fig. 7A), BIX01294 treatment at this dose was moderately toxic to cells. At doses above 25 µM, BIX01294 caused cellular death, with complete cellular detachment from plates (data not shown) and loss of firefly luciferase expression (Supplementary Fig. 7A). In contrast, 25 µM BIX01294 treatment unexpectedly increased poly-GA expression (Supplementary Fig. 7B), perhaps indirectly through changes in gene expression related to its primary biological activity as a histone lysine methyltransferase inhibitor.

We also applied BIX01294, anthralin, propidium iodide, and PPIX to primary rodent neurons at 1 µM, and tracked the survival of these neurons over the course of ten days using automated longitudinal fluorescence microscopy. Even at a dose well below the small molecules' *in vitro* IC50s (Fig. 3), treatment with each increased the neurons' cumulative risk of death (Supplementary Fig. 7C). Consequently, none of these small molecules are likely to represent good therapeutic candidates.

DISCUSSION

In recent years, an emerging body of research has shown RAN translation to be a common factor in numerous repeat-expansion diseases. Although much work has established that RAN peptides are sufficient to cause neurotoxicity in a broad range of disease models, our understanding of the mechanisms of RAN translation, as well as how to selectively target it, is still incomplete.

Here we used an *in vitro* RAN translation reporter assay to perform a screen of bioactive compounds and identified five novel small molecule inhibitors of this non-canonical translation event. Although none of these compounds are bioavailable, by interrogating their activity, they have revealed new insights into the mechanism of RAN translation.

We performed the initial screens using only a +1CGGx100 RAN reporter mRNA. However, each compound identified in the screen had similar activity against GGGGCCx70 RAN translation reporter mRNAs. Therefore, despite being different repeats, located within different sequence contexts, that produce different RAN peptide products, this finding supports mechanistic overlap between these translation events, and suggests that future strategies to modify RAN translation in one disease could be more broadly applicable to the growing class of diseases associated with this event.

Previous studies to develop inhibitors of RAN translation have specifically focused on identifying small molecules that bind to the repeat RNAs (33-35,42). For instance, compound 1a was identified based on its ability to bind to loops produced by the unpaired G-G nucleotides in the hairpin stem structures formed by both CGG and GGGGCC repeats (35). Subsequent studies established that 1a, and structurally similar compounds, reduce RAN translation of both RNAs in cultured cells, including *C9ORF72*-patient-derived iNeurons (33,42). Additionally, a recent screen for molecules that bind G-quadruplexes successfully identified two such molecules that reduced C9RAN translation in GGGGCCx36-expressing *Drosophila* and *C9ORF72*-patient-derived motor neurons (34).

Consequently, small molecule binding to repeat RNAs has proven to be a successful strategy at inhibiting RAN translation. This is consistent with the findings we report here, as three of the small molecule inhibitors we identified in the screen, BIX01294, CP-31398, and propidium iodide, show evidence of interacting with CGG and/or GGGGCC repeat RNAs by circular dichroism and/or native gel analysis. Future work to determine if the functional group shared by BIX01294 and CP-31398 is important for their interaction with the repeat RNAs and/or selective inhibition of RAN translation, could provide insight into strategies for optimizing this effect. However,

at higher doses, these compounds also interacted with the control UUUx16 RNA. Therefore, we have not excluded the possibility that interactions of these compounds with other RNAs or proteins in our translation system, such as ribosomal RNAs, contribute to their relatively selective inhibition of RAN translation.

Unlike previous studies, by using a luciferase-based RAN translation reporter assay to directly measure RAN translation, we were able to unbiasedly screen a diverse library of bioactive compounds, and potentially identify small molecules that inhibit RAN translation through novel mechanisms not related to repeat RNA binding. In this context, our finding that neither anthralin nor cephalothin directly interact with the CGG and/or GGGGCC repeat RNAs as measured by circular dichroism or native gel analysis, is quite intriguing. This suggests that these compounds inhibit RAN translation through novel means, potentially by interacting with or inhibiting non-RNA-based factors that are selectively required for RAN translation.

Our findings also suggest that relatively selective inhibition of RAN translation can occur by targeting different aspects of the translation process. Relative to +1 and +2CGGx100 RAN translation, anthralin, TMPyP4 and PPIX inhibited AUG-initiated repeat-translation to a similar or greater extent. This indicates that these compounds do not target the non-AUG initiation event of CGG RAN translation but may instead impair translation elongation through the repetitive RNA elements. Alternatively, BIX01294 and CP-31398 do specifically target the non-AUG initiation event; they were less effective at inhibiting AUG-initiated repeat-translation compared to RAN translation, and selectively inhibited translation from CUG-NLuc and ACG-NLuc reporters, relative to an AUG-NLuc control.

In conclusion, this study serves as a proof-of-principle that small, bioactive compounds can selectively inhibit RAN translation across multiple disease-causing repeat expansion mutations. It has also revealed new lead compounds that could be used in future work to identify novel, targetable binding pockets in the repeat RNAs, as well as new factors involved in RAN translation.

EXPERIMENTAL PROCEDURES
Reporter RNA in vitro transcription

pcDNA3.1(+) NLuc reporter plasmids (26,27,36,39) were linearized with PspOMI. Capped and polyadenylated mRNAs were synthesized as previously described, using HiScribe T7 High Yield RNA Synthesis Kit (NEB, Catalog No. E2040S), 3′-O-Me-m⁷GpppG anti-reverse cap analog (ARCA) (NEB, Catalog No. S1411S), RNAse-free DNaseI (NEB, Catalog No. M0303S), and *E. coli* Poly-A Polymerase (NEB, Catalog No. M0276S) (26). Following synthesis, mRNAs were purified with RNA Clean and Concentrator-25 Kit from Zymo Research (Catalog No. R1017). The integrity and size of all *in vitro* transcribed mRNAs was verified on a denaturing formaldehyde RNA gel.

Primary and secondary screens

Rabbit reticulocyte lysate (RRL) *in vitro* translation reactions were performed in white, flat bottom, polypropylene 384-well plates from Greiner (Catalog No. 784075). RRL reaction mixture was prepared using Promega's Flexi System (Catalog No. L4540), and consisted of 30% RRL, 10 µM minus leucine amino acid mix, 10 µM minus methionine amino acid mix, 0.5 mM magnesium acetate, 100 mM potassium chloride, 4U murine RNAse inhibitor (NEB, Catalog No. M0314L), and 0.05% Tween-20. The addition of Tween-20 was necessary to minimize adhesion of the reaction mixture to the multidrop tubing and even dispensing between wells, and did not alter the levels of translation from the +1 CGG RAN or AUG-NLuc reporter RNAs (Supplementary Fig. 1B). 4 µL of RRL mixture was dispensed per well using Thermo Labsystems multidrop plate dispenser, at medium speed. 50 nL of DMSO was then multidropped into columns 1 and 2, and 50 nL cycloheximide was added to columns 23 and 24, for a final concentration of 3 µM. To the remaining wells, 50 nL of 3253 compounds, at 2 mM, from the Pilot LOPAC, Pilot Prestwick, Pilot NCC Focused, and Navigators Pathways Libraries were pintooled using a Sciclone ALH 300 advanced liquid handling system, for a final concentration of 20 µM. 1 µL (3 fmol +1 CGG RAN, 1 fmol AUG-NLuc) *in vitro* transcribed reporter RNA was then multidropped into each well.

Plates were incubated for 30 minutes at 30ºC. Translation reactions were then terminated by addition of 50 nL, 10 µM cycloheximide to each well, using a multidrop dispenser. 10 µL of room temperature Glo Lysis Buffer (Promega, Catalog No. E2661) was dispensed into each well with the multidrop, at low speed, which was essential for maintaining the stability of NLuc luminescence. To this, 5 µL of room temperature NanoGlo Substrate diluted 1:50 in NanoGlo Buffer (Promega, Catalog No. N1120) was dispensed via multidrop at low speed. Plates were then covered, mixed gently, and incubated at room temperature for 5 minutes. Luminescence was measured with a Perkin Elmer EnVision plate reader.

The concentration response curve and counter screen were performed using all the same reagents, materials, and equipment as in the primary screen. The only difference was that selected compounds were added to 384-well plates at 8 doses from 50-8 µM, in duplicate, using TTP Labtech Mosquito X1 Hit-Picking Liquid Handler, prior to addition of RRL.

After measuring luminescence, pIC50s were calculated through the University of Michigan's Center for Chemical Genomics' MScreen software, using a 4 parameter logistic equation (43). Initial minimum limits for curve fits were informed by the minimum value of each plate, determined by the cycloheximide-treated internal positive controls. Regression analysis then derived a curve fit, with a calculated minimum and maximum limit, for each compound. IC50s greater than the maximum dose in the assay (50 µM), were extrapolated from the curve fit.

Independent validation of hits in RRL system

The following compounds were purchased through Sigma Marketsite: cephalothin sodium (Sigma-Aldrich; C4520), rufloxacin HCl (Vitas-M Lab., Ltd.; STK711124), rofecoxib (Sigma-Aldrich; MFCD00935806), alfuzosin HCl (Sigma-Aldrich; A0232), olanzapine (Sigma-Aldrich; O1141), protoporphyrin IX (Chem-Impex Int, Inc; 21661), pefloxacine mesylate (Selleck Chemicals, S1855), reserpine (Sigma-Aldrich; MFCD00005091), isoxicam (Sigma-Aldrich; MFCD000079374), phenazopyridine HCl (Sigma-Aldrich; MFCD00035347), parbendazole (Sigma-Aldrich; MFCD018910864), balsalazide disodium salt dihydrate (Key Organics/BIONET; KS-5216), efavirenz (Sigma-Aldrich; SML0536), oxiconazole nitrate (Key Organics/BIONET; KS-5288),

8

sulfamethazine sodium salt (Sigma; S5637-25G), kenpaullone (Adooq Bioscience; A11220), propidium iodide (Sigma; P4170), BIX01294 HCl hydrate (Cayman Chemical; 13124), anthralin (Key Organics/BIONET; KS-5183), CP-31398 dihydrochloride hydrate (Tocris Bioscience; 3023), X80 (Sigma; X3629), amlexanox (Sigma; 68302-57-8).

Compounds were all dissolved to 10-100 mM stocks in DMSO and stored in single-use aliquots at -20ºC. For *in vitro* translation assays, compounds were diluted in DMSO and added to RRL for a final DMSO concentration of 1% total reaction volume. mRNAs were *in vitro* translated with Flexi Rabbit Reticulocyte Lysate System from Promega, as performed for the primary screen, with a few slight modifications. First, 0.05% Tween-20 was excluded from reactions, as regular pipetting eliminated the need to prevent RRL adhesion to multidrop dispenser tubing. Additionally, the reaction volume was increased to 10 µL, with an RNA concentration of 0.6 nM. Incubations were performed in polypropylene PCR tubes, reactions were terminated on ice, and then transferred to black 96-well plates. Samples were diluted 1:7 in Glo Lysis Buffer (Promega) prior to a 5 minute in the dark with NanoGlo Substrate freshly diluted 1:50 in NanoGlo Buffer (Promega). Luminescence was measured on a GloMax 96 Microplate Luminometer.

Reactions for western blot assays were performed as above, except 50 ng mRNA was used, as previously described (27). Membranes were probed with the following antibodies: 1:1,000 FLAG-M2 (mouse, Sigma F1804), 1:1,000 GAPDH 6C5 (mouse, Santa Cruz sc32233).

Circular dichroism

CGGx16, UUUx16, and GGGGCCx8 RNAs were purchased from IDT, with RNAse-free HPLC purification.

CGGx16 and UUUx16 RNAs were diluted to 5 µM in 300 µL RNA folding buffer consisting of 100 mM KCl, 10 mM Tris-HCl, and 0.1 mM EDTA (38). GGGGCCx8 RNA was diluted to 2.5 µM in 250 µL RNA folding buffer, consisting of either 100 mM KCl or 100 mM NaCl, with 10 mM Tris-HCl, and 0.1 mM EDTA. Diluted RNAs were heated to 95ºC for 1 minute and cooled to room temperature.

CD was performed using a Jasco J-1500 Circular Dichroism Spectrophotometer. Sample were scanned from 320-220 nm at 25ºC (38). Spectra were collected with scanning speed of 20nm/min (38), data interval of 0.1nm, response time of 1 sec and 1nm bandwidth. Spectra are the average of six accumulations, and only data with a heat tension less than 500 is shown. Increasing volumes of compounds were progressively added to the same RNA sample, to obtain increasing doses. The resulting decrease in RNA concentration due to increasing volume was accounted for in the molar ellipticity calculations. Spectra collected on the same day are compared to the same untreated RNA spectra, for each RNA and cation tested. All spectra were corrected for the background, by subtracting the buffer spectra using Jasco software, and baselined by setting the first value of the RNA-only spectra to zero.

Native gel analysis

5′ IRDye-800CW-modified CGGx16 and UUUx16 RNAs were purchased from IDT, with RNAse-free HPLC purification. RNAs was diluted to 50 nM in buffer containing 100 mM KCl, 10 mM Tris-HCl, and 0.1 mM EDTA, heated to 95ºC for 1 minute, and cooled to room temperature.

Folded RNAs were then incubated with the indicated compounds and at the indicated doses for 30 minutes at room temperature. 4 µL of Orange-G loading Dye (40% glycerol, 0.1% Orange-G, 1x TBE, and sterile MQ H_2O) was added to each reaction, and entire reactions were loaded into 12% native polyacrylamide gels (12% acrylamide:bis acrylamide 29:1, 1X TBE, 0.1% APS, TEMED, and sterile MQ H_2O), with 4% stacks. Prior to loading, gels were pre-ran in 1x TBE at 100V for approximately 45 minutes. Samples were run on ice at 70V for approximately 95 minutes, and entire gels within glass cartridges were imaged with LI-COR Odyssey CLx Imaging System.

Compound activity in HEK293 cells

HEK293 cells (ATCC CRL-1573) were maintained at 37ºC and 5% CO_2 in 10 cm dishes containing 10 mL DMEM + high glucose (GE Healthcare Bio-Science, SH30022FS) supplemented with 9.09% fetal bovine serum (50 mL added to 500 mL DMEM; Bio-Techne, S11150). 100 µL cells were plated at $2.2x10^5$ cells/mL in 96-well plates (Fisher, FB012931).

9

Approximately 24 hours post plating, at 50-60% confluency, cells were transfected with pcDNA3.1+ NLuc reporter plasmids and control pGL4.13 Firefly Luciferase plasmid (26,27). 50 ng of each plasmid was transfected per well, using FuGene HD (Promega, E2312) at a 3:1 ratio of reagent to total DNA, according to manufacturer's recommended protocol. Approximately 24 hours post transfection, at approximately 90% confluency, cells were treated with BIX01294 at indicated doses. 24 hours post treatment, cells were lysed in 60 µL Glo Lysis Buffer per well, for 5 minutes at room temperature with rocking. In opaque, black 96-well plates (Grenier Bio-One, 655076), 25 µL lysate was then separately incubated for 5 minutes in the dark at room temperature with 25 µL NanoGlo Substrate freshly diluted 1:50 in NanoGlo Buffer or 25 µL ONE-Glo Buffer (Promega, E6110). NLuc and Firefly luminescence were measured on a GloMax 96 Microplate Luminometer.

Primary neuron survival experiments

Primary cortical neurons from embryonic day 19-20 rats were harvested and cultured at 0.6 x 10^6 cells/mL in 96 well cell culture plates, as previously described (44,45). Neurons were cultured in NEUMO photostable medium containing SOS supplement (Cell Guidance Systems) at 37ºC in 5% CO_2. Primary neurons were co-transfected with 0.1 µg of the survival marker, pGW1-mApple, and 0.1 µg of pGW1-EGFP on DIV4 using Lipofectamine 2000 (Invitrogen). Twenty-four hours post-transfection, compounds or DMSO were added to neurons, immediately following the first imaging run. Images were taken for 10 consecutive days with an automated fluorescent microscopy platform, as previously described (44-47). Image processing and survival analysis were acquired for each neuron at each timepoint using custom code written in Python or the ImageJ macro language. For survival analyses, differences among populations through Cox proportional hazards analysis was determined with the publicly available R survival package.

10

Acknowledgment

Conflict of interest

PKT serves as a consultant with Denali Therapeutics, and he, KMG, and MGK licensed technology through the University of Michigan to Denali Therapeutics that is based on the work presented here.

REFERENCES

1. Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., Margolis, J., Peterson, M., Markowski, T. W., Ingram, M. A., Nan, Z., Forster, C., Low, W. C., Schoser, B., Somia, N. V., Clark, H. B., Schmechel, S., Bitterman, P. B., Gourdon, G., Swanson, M. S., Moseley, M., and Ranum, L. P. (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* **108**, 260-265

2. Todd, P. K., Oh, S. Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A. J., Chen, K. C., Scaglione, K. M., Basrur, V., Elenitoba-Johnson, K., Vonsattel, J. P., Louis, E. D., Sutton, M. A., Taylor, J. P., Mills, R. E., Charlet-Berguerand, N., and Paulson, H. L. (2013) CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **78**, 440-455

3. Krans, A., Kearse, M. G., and Todd, P. K. (2016) Repeat-associated non-AUG translation from antisense CCG repeats in fragile X tremor/ataxia syndrome. *Ann Neurol* **80**, 871-881

4. Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., Dejesus-Hernandez, M., van Blitterswijk, M. M., Jansen-West, K., Paul, J. W., Rademakers, R., Boylan, K. B., Dickson, D. W., and Petrucelli, L. (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639-646

5. Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H. A., Cruts, M., Van Broeckhoven, C., Haass, C., and Edbauer, D. (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335-1338

6. Mori, K., Arzberger, T., Grässer, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S. M., Schludi, M. H., van der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E., Kretzschmar, H. A., Haass, C., and Edbauer, D. (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol* **126**, 881-893

7. Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H., Ostrow, L. W., Rothstein, J. D., Troncoso, J. C., and Ranum, L. P. (2013) RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* **110**, E4968-4977

8. Zu, T., Cleary, J. D., Liu, Y., Bañez-Coronel, M., Bubenik, J. L., Ayhan, F., Ashizawa, T., Xia, G., Clark, H. B., Yachnis, A. T., Swanson, M. S., and Ranum, L. P. W. (2017) RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron* **95**, 1292-1305.e1295

9. Bañez-Coronel, M., Ayhan, F., Tarabochia, A. D., Zu, T., Perez, B. A., Tusi, S. K., Pletnikova, O., Borchelt, D. R., Ross, C. A., Margolis, R. L., Yachnis, A. T., Troncoso, J. C., and Ranum, L. P. (2015) RAN Translation in Huntington Disease. *Neuron* **88**, 667-677

10. Buijsen, R. A., Visser, J. A., Kramer, P., Severijnen, E. A., Gearing, M., Charlet-Berguerand, N., Sherman, S. L., Berman, R. F., Willemsen, R., and Hukema, R. K. (2016) Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency. *Hum Reprod* **31**, 158-168

11. Soragni, E., Petrosyan, L., Rinkoski, T. A., Wieben, E. D., Baratz, K. H., Fautsch, M. P., and Gottesfeld, J. M. (2018) Repeat-Associated Non-ATG (RAN) Translation in Fuchs' Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci* **59**, 1888-1896

12. Jacquemont, S., Hagerman, R. J., Leehey, M., Grigsby, J., Zhang, L., Brunberg, J. A., Greco, C., Des Portes, V., Jardini, T., Levine, R., Berry-Kravis, E., Brown, W. T., Schaeffer, S., Kissel, J., Tassone, F., and Hagerman, P. J. (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am J Hum Genet* **72**, 869-878

13. Jacquemont, S., Hagerman, R. J., Leehey, M. A., Hall, D. A., Levine, R. A., Brunberg, J. A., Zhang, L., Jardini, T., Gane, L. W., Harris, S. W., Herman, K., Grigsby, J., Greco, C. M., Berry-

Kravis, E., Tassone, F., and Hagerman, P. J. (2004) Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA* **291**, 460-469

14. Hagerman, R. J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B., and Hagerman, P. J. (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* **57**, 127-130

15. Sellier, C., Buijsen, R. A., He, F., Natla, S., Jung, L., Tropel, P., Gaucherot, A., Jacobs, H., Meziane, H., Vincent, A., Champy, M. F., Sorg, T., Pavlovic, G., Wattenhofer-Donze, M., Birling, M. C., Oulad-Abdelghani, M., Eberling, P., Ruffenach, F., Joint, M., Anheim, M., Martinez-Cerdeno, V., Tassone, F., Willemsen, R., Hukema, R. K., Viville, S., Martinat, C., Todd, P. K., and Charlet-Berguerand, N. (2017) Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. *Neuron* **93**, 331-347

16. Buijsen, R. A., Sellier, C., Severijnen, L. A., Oulad-Abdelghani, M., Verhagen, R. F., Berman, R. F., Charlet-Berguerand, N., Willemsen, R., and Hukema, R. K. (2014) FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. *Acta Neuropathol Commun* **2**, 162

17. DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G. Y., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D. S., Petersen, R. C., Miller, B. L., Dickson, D. W., Boylan, K. B., Graff-Radford, N. R., and Rademakers, R. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245-256

18. Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., Schymick, J. C., Laaksovirta, H., van Swieten, J. C., Myllykangas, L., Kalimo, H., Paetau, A., Abramzon, Y., Remes, A. M., Kaganovich, A., Scholz, S. W., Duckworth, J., Ding, J., Harmer, D. W., Hernandez, D. G., Johnson, J. O., Mok, K., Ryten, M., Trabzuni, D., Guerreiro, R. J., Orrell, R. W., Neal, J., Murray, A., Pearson, J., Jansen, I. E., Sondervan, D., Seelaar, H., Blake, D., Young, K., Halliwell, N., Callister, J. B., Toulson, G., Richardson, A., Gerhard, A., Snowden, J., Mann, D., Neary, D., Nalls, M. A., Peuralinna, T., Jansson, L., Isoviita, V. M., Kaivorinne, A. L., Hölttä-Vuori, M., Ikonen, E., Sulkava, R., Benatar, M., Wuu, J., Chiò, A., Restagno, G., Borghero, G., Sabatelli, M., Heckerman, D., Rogaeva, E., Zinman, L., Rothstein, J. D., Sendtner, M., Drepper, C., Eichler, E. E., Alkan, C., Abdullaev, Z., Pack, S. D., Dutra, A., Pak, E., Hardy, J., Singleton, A., Williams, N. M., Heutink, P., Pickering-Brown, S., Morris, H. R., Tienari, P. J., Traynor, B. J., and Consortium, I. (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268

19. Sareen, D., O'Rourke, J. G., Meera, P., Muhammad, A. K., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A., Gendron, T., Petrucelli, L., Baughn, M., Ravits, J., Harms, M. B., Rigo, F., Bennett, C. F., Otis, T. S., Svendsen, C. N., and Baloh, R. H. (2013) Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* **5**, 208ra149

20. Mizielinska, S., Grönke, S., Niccoli, T., Ridler, C. E., Clayton, E. L., Devoy, A., Moens, T., Norona, F. E., Woollacott, I. O., Pietrzyk, J., Cleverley, K., Nicoll, A. J., Pickering-Brown, S., Dols, J., Cabecinha, M., Hendrich, O., Fratta, P., Fisher, E. M., Partridge, L., and Isaacs, A. M. (2014) C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* **345**, 1192-1194

21. Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S. S., Shi, Y., Lin, S., Shneider, N. A., Monaghan, J., Pandey, U. B., Pasinelli, P., Ichida, J. K., and Trotti, D. (2014) Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* **84**, 1213-1225

22. Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y., and McKnight, S. L. (2014) Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139-1145

23. Zhang, Y. J., Jansen-West, K., Xu, Y. F., Gendron, T. F., Bieniek, K. F., Lin, W. L., Sasaguri, H., Caulfield, T., Hubbard, J., Daughrity, L., Chew, J., Belzil, V. V., Prudencio, M., Stankowski, J. N., Castanedes-Casey, M., Whitelaw, E., Ash, P. E., DeTure, M., Rademakers, R., Boylan, K. B., Dickson, D. W., and Petrucelli, L. (2014) Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol* **128**, 505-524

24. May, S., Hornburg, D., Schludi, M. H., Arzberger, T., Rentzsch, K., Schwenk, B. M., Grässer, F. A., Mori, K., Kremmer, E., Banzhaf-Strathmann, J., Mann, M., Meissner, F., and Edbauer, D. (2014) C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol* **128**, 485-503

25. Flores, B. N., Dulchavsky, M. E., Krans, A., Sawaya, M. R., Paulson, H. L., Todd, P. K., Barmada, S. J., and Ivanova, M. I. (2016) Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS One* **11**, e0165084

26. Kearse, M. G., Green, K. M., Krans, A., Rodriguez, C. M., Linsalata, A. E., Goldstrohm, A. C., and Todd, P. K. (2016) CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol Cell* **62**, 314-322

27. Green, K. M., Glineburg, M. R., Kearse, M. G., Flores, B. N., Linsalata, A. E., Fedak, S. J., Goldstrohm, A. C., Barmada, S. J., and Todd, P. K. (2017) RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nature Communications* **8**, 2005

28. Tabet, R., Schaeffer, L., Freyermuth, F., Jambeau, M., Workman, M., Lee, C. Z., Lin, C. C., Jiang, J., Jansen-West, K., Abou-Hamdan, H., Désaubry, L., Gendron, T., Petrucelli, L., Martin, F., and Lagier-Tourenne, C. (2018) CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat Commun* **9**, 152

29. Cheng, W., Wang, S., Mestre, A. A., Fu, C., Makarem, A., Xian, F., Hayes, L. R., Lopez-Gonzalez, R., Drenner, K., Jiang, J., Cleveland, D. W., and Sun, S. (2018) C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2α phosphorylation. *Nat Commun* **9**, 51

30. Mackenzie, I. R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S. M., Haass, C., Kretzschmar, H. A., Edbauer, D., and Neumann, M. (2013) Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol* **126**, 859-879

31. Sonobe, Y., Ghadge, G., Masaki, K., Sendoel, A., Fuchs, E., and Roos, R. P. (2018) Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. *Neurobiol Dis* **116**, 155-165

32. Westergard, T., McAvoy, K., Russell, K., Wen, X., Pang, Y., Morris, B., Pasinelli, P., Trotti, D., and Haeusler, A. (2019) Repeat-associated non-AUG translation in C9orf72-ALS/FTD is driven by neuronal excitation and stress. *EMBO Mol Med* **11**

33. Su, Z., Zhang, Y., Gendron, T. F., Bauer, P. O., Chew, J., Yang, W. Y., Fostvedt, E., Jansen-West, K., Belzil, V. V., Desaro, P., Johnston, A., Overstreet, K., Oh, S. Y., Todd, P. K., Berry, J. D., Cudkowicz, M. E., Boeve, B. F., Dickson, D., Floeter, M. K., Traynor, B. J., Morelli, C., Ratti, A., Silani, V., Rademakers, R., Brown, R. H., Rothstein, J. D., Boylan, K. B., Petrucelli, L., and Disney, M. D. (2014) Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* **83**, 1043-1050

34. Simone, R., Balendra, R., Moens, T. G., Preza, E., Wilson, K. M., Heslegrave, A., Woodling, N. S., Niccoli, T., Gilbert-Jaramillo, J., Abdelkarim, S., Clayton, E. L., Clarke, M., Konrad, M. T., Nicoll, A. J., Mitchell, J. S., Calvo, A., Chio, A., Houlden, H., Polke, J. M., Ismail, M. A., Stephens, C. E., Vo, T., Farahat, A. A., Wilson, W. D., Boykin, D. W., Zetterberg, H., Partridge, L., Wray, S., Parkinson, G., Neidle, S., Patani, R., Fratta, P., and Isaacs, A. M. (2018) G-

quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Mol Med* **10**, 22-31

35. Disney, M. D., Liu, B., Yang, W. Y., Sellier, C., Tran, T., Charlet-Berguerand, N., and Childs-Disney, J. L. (2012) A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem Biol* **7**, 1711-1718

36. Kumari, D., Swaroop, M., Southall, N., Huang, W., Zheng, W., and Usdin, K. (2015) High-Throughput Screening to Identify Compounds That Increase Fragile X Mental Retardation Protein Expression in Neural Stem Cells Differentiated From Fragile X Syndrome Patient-Derived Induced Pluripotent Stem Cells. *Stem Cells Transl Med* **4**, 800-808

37. Weisman-Shomer, P., Cohen, E., Hershco, I., Khateb, S., Wolfovitz-Barchad, O., Hurley, L. H., and Fry, M. (2003) The cationic porphyrin TMPyP4 destabilizes the tetraplex form of the fragile X syndrome expanded sequence d(CGG)n. *Nucleic Acids Res* **31**, 3963-3970

38. Zamiri, B., Reddy, K., Macgregor, R. B., and Pearson, C. E. (2014) TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNA-binding proteins. *J Biol Chem* **289**, 4653-4659

39. Kearse, M. G., Goldman, D. H., Choi, J., Nwaezeapu, C., Liang, D., Green, K. M., Goldstrohm, A. C., Todd, P. K., Green, R., and Wilusz, J. E. (2019) Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. *Genes Dev* **33**, 871-885

40. Starck, S. R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T., and Shastri, N. (2012) Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**, 1719-1723

41. Ofer, N., Weisman-Shomer, P., Shklover, J., and Fry, M. (2009) The quadruplex r(CGG)n destabilizing cationic porphyrin TMPyP4 cooperates with hnRNPs to increase the translation efficiency of fragile X premutation mRNA. *Nucleic Acids Res* **37**, 2712-2722

42. Yang, W. Y., Wilson, H. D., Velagapudi, S. P., and Disney, M. D. (2015) Inhibition of Non-ATG Translational Events in Cells via Covalent Small Molecules Targeting RNA. *J Am Chem Soc* **137**, 5336-5345

43. Sebaugh, J. L. (2011) Guidelines for accurate EC50/IC50 estimation. *Pharm Stat* **10**, 128-134

44. Barmada, S. J., Skibinski, G., Korb, E., Rao, E. J., Wu, J. Y., and Finkbeiner, S. (2010) Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci* **30**, 639-649

45. Barmada, S. J., Ju, S., Arjun, A., Batarse, A., Archbold, H. C., Peisach, D., Li, X., Zhang, Y., Tank, E. M., Qiu, H., Huang, E. J., Ringe, D., Petsko, G. A., and Finkbeiner, S. (2015) Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc Natl Acad Sci U S A* **112**, 7821-7826

46. Flores, B. N., Li, X., Malik, A. M., Martinez, J., Beg, A. A., and Barmada, S. J. (2019) An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep* **27**, 1133-1150.e1138

47. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805-810

SUPPLEMENTARY INFORMATION

Supplementary Figures 1-7

Supplementary Table 1: Results from primary screen

Supplementary Table 2: Results from secondary screen

Supplementary Table 3: NLuc reporter mRNA sequences

Table 1: Primary Screen Statistics

Z′ Scores per Plate	
Average	0.79
Minimum	0.70
Maximum	0.88
Coefficients of Variance per Plate	
Average	6.93%
Minimum	3.88%
Maximum	9.82%

Table 2: Summary of Primary Screen Hits

Signal Reduction Range	Number of Compounds
20-39%	154
40-59%	25
60-79%	9
80-99%	8
>3 SD Below Mean	211
Total (>20% or >3 SD)	274

Table 3: Activity of Compounds in Independent Validation

| Compound Name	ID Number	Active against +1CGGx100?	Active against AUG?
Kenpaullone	CCG-35778	No	No
BIX01294 (hydrochloride hydrate)	**CCG-208677**	**Yes**	**Less**
Propidium Iodide	**CCG-220792**	**Yes**	**Less**
Amlexanox	CCG-100953	No	No
Anthralin	**CCG-38920**	**Yes**	**Less**
Rosiglitazone Maleate	CCG-100943	No	No
X80	CCG-222138	Yes	Yes
Balsalazide Disodium	CCG-213078	No	No
CP-31398 dihydrochloride hydrate	**CGG-222578**	**Yes**	**Less**
Phenazopyridine hydrochloride	CGG-39935	No	No
Rufloxacin hyrdochloride	CGG-100908	No	No
Reserpine	CGG-204168	No	No
Efavirenz	CGG-101011	No	No
Alfuzosin hydrochloride	CGG-213373	No	No
Cephalothin Sodium	**CCG-38923**	**Yes**	**Less**
Rofecoxib	CGG-40253	No	No
Isoxicam	CCG-40307	Yes	Yes
Pefloxacine Mesylate	CCG-39994	No	No
Parbendazole	CCG-221110	No	No
Sulfamethazine sodium salt	CCG-220775	No	No
Olanzapine	CCG-100922	No	No
Oxiconazole Nitrate	CCG-40020	No	No

Figure 1
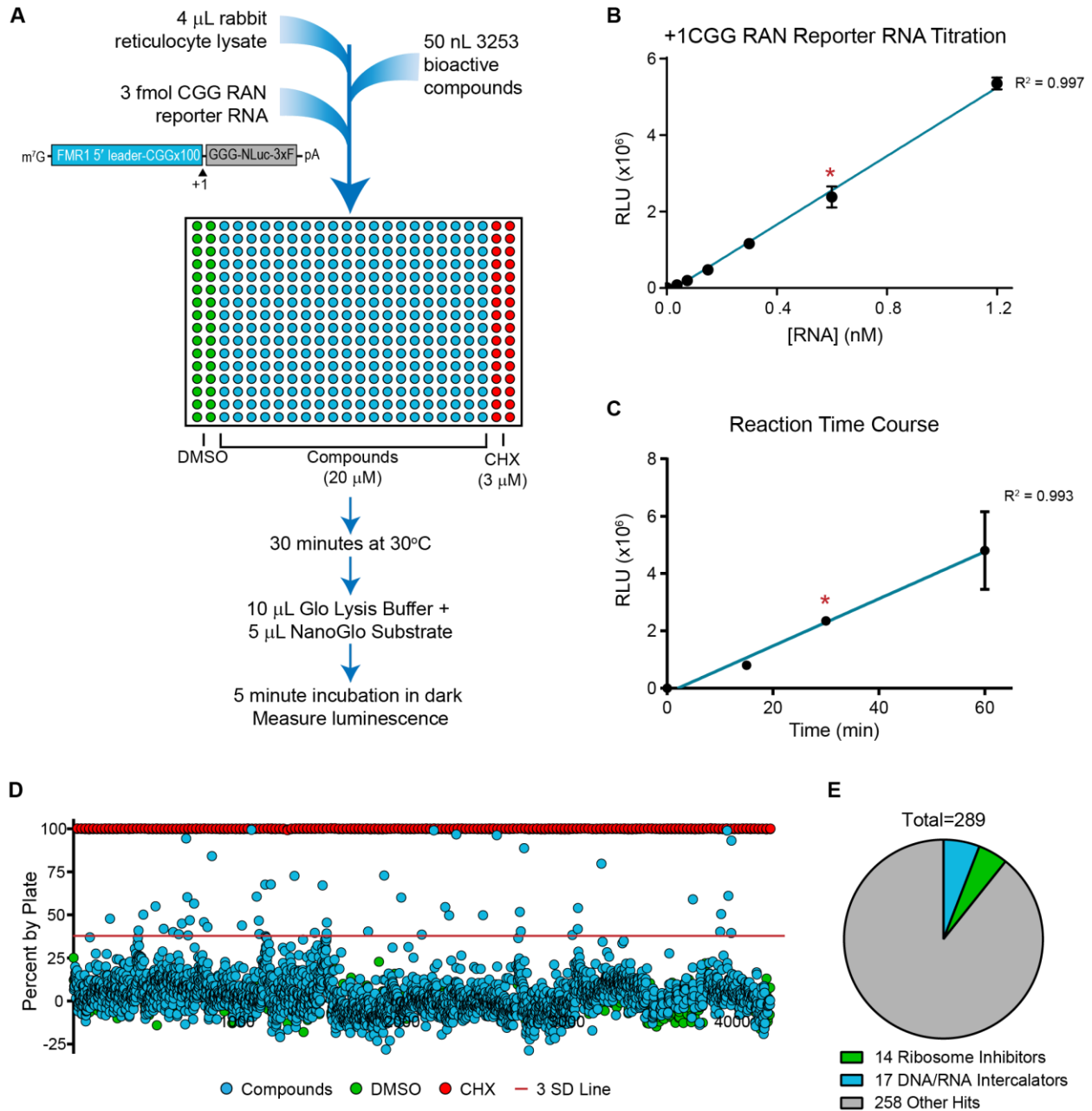
Figure 1. Screen of 3253 bioactive compounds for inhibitors of RAN translation

(A) Schematic of primary screen design. In each plate, DMSO (columns 1 and 2) served as an internal negative control and 3 µM cycloheximide (CHX, columns 23 and 24) as an internal positive control for translation inhibition. Compounds screened were from the Pilot LOPAC, Pilot Prestwick, Pilot NCC-Focused, and Navigator Pathways libraries. 3xF = 3xFLAG tag (B) Linear relationship between +1CGG RAN reporter mRNA concentration and luminescence for *in vitro* translation assay under conditions used

for primary screen. (C) Linear relationship between reaction time and luminescence for *in vitro* translation assay with 3fmol +1CGG nanoluciferase (NLuc) RAN reporter mRNA under conditions used for primary screen. For (B) and (C), teal lines represent linear regression fit, with red asterisk indicating conditions used in screen. Each point represents the mean of n=3, and error bars represent +/- standard deviation. (D) Percent change in NLuc signal for all 3,253 compounds relative to the DMSO vehicle controls. Red line approximately represents 3 standard deviations from vehicle controls. (E) Pie chart of screen hits, indicating the proportion of hits that are known ribosome inhibitors or DNA/RNA intercalators.

Figure 2
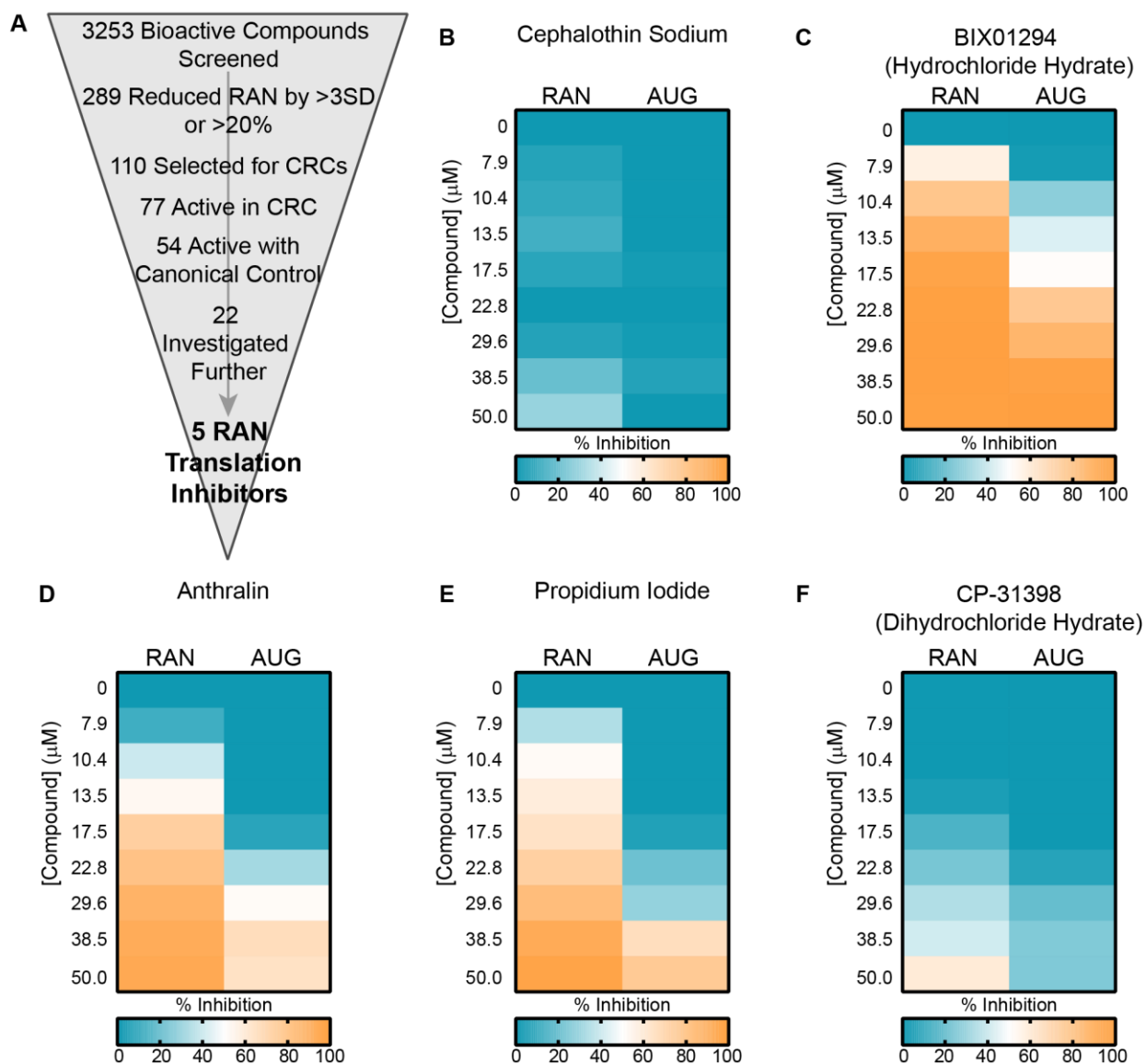
Figure 2. Concentration response curves and counter screen

(A) Schematic of screen work flow, with number of hits at each step indicated. SD = standard deviation, CRC = concentration response curve. (B) Representative heat map of a "RAN specific inhibitor," cephalothin, showing percent inhibition of +1CGG RAN translation and canonical AUG translation at indicated concentrations. (C-F) Representative heat maps of "RAN selective inhibitors," showing percent inhibition of +1CGG RAN translation and canonical AUG translation at indicated concentrations. For (B-F), heat map values are the average of duplicate reactions. For visual clarity, effects of <0% inhibition (increases) were set to zero.

Figure 3
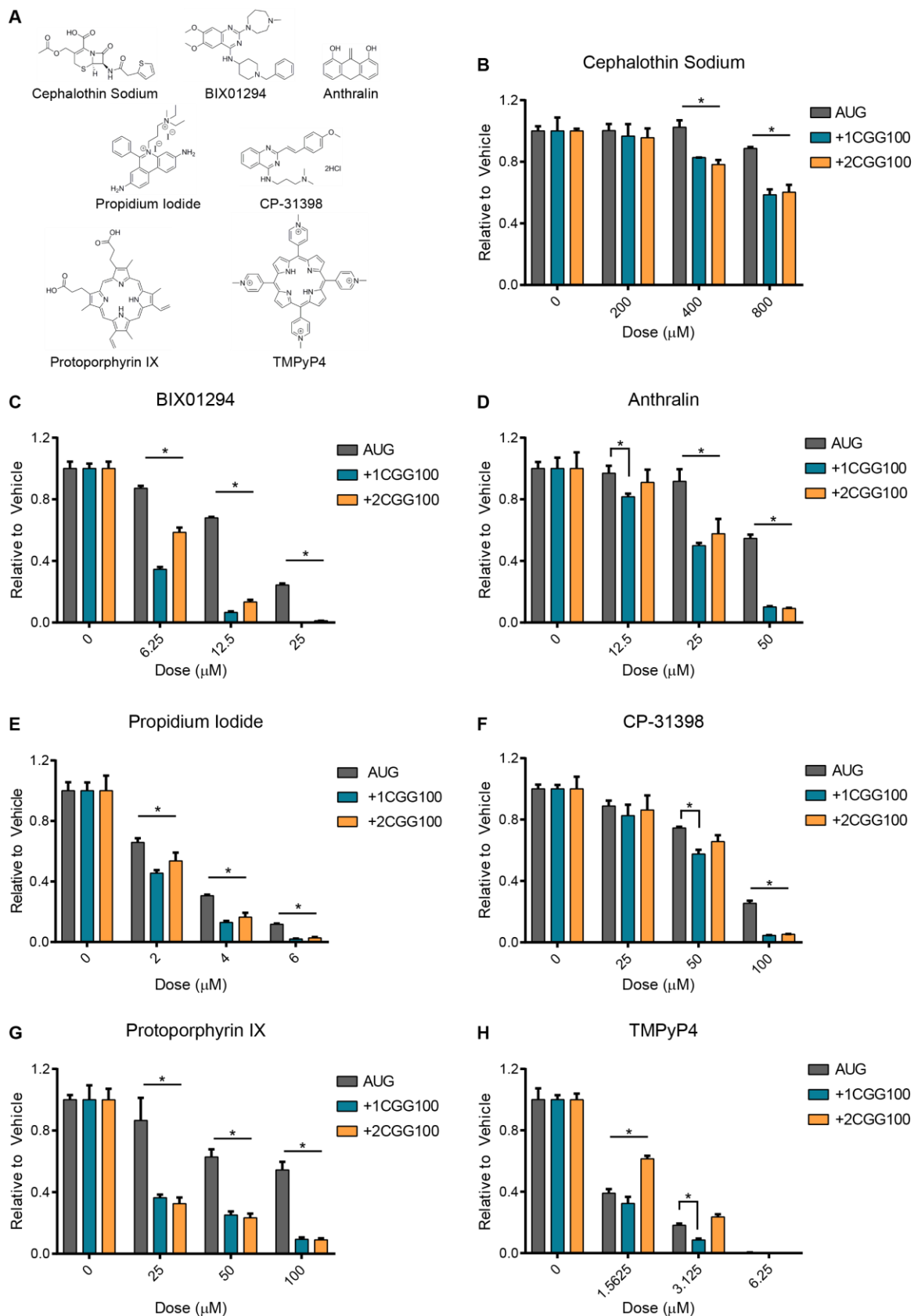
Figure 3. Small molecules inhibit CGG RAN translation in multiple reading frames

(A) Structures of compounds used in panels B-H. (B-F) Compounds identified from screen as RAN translation inhibitors were independently re-assessed for their activity in RRL *in vitro* translation assays. Indicated compounds were added at increasing doses to 30% RRL reactions with AUG, +1CGGx100, or +2CGGx100-NLuc reporter mRNAs and luminescence measured relative to vehicle (DMSO) treatment. (G-H) Candidates compounds protoporphyrin IX and TMPyP4 were added at increasing concentration to RRL reactions with AUG, +1CGGx100, or +2CGGx100-NLuc reporter mRNAs. Luminescence was measured relative to vehicle (DMSO) treatment. All graphs represent n=3 +/- standard deviation. *p < 0.05, 2-way ANOVA with Dunnett's multiple comparison test.
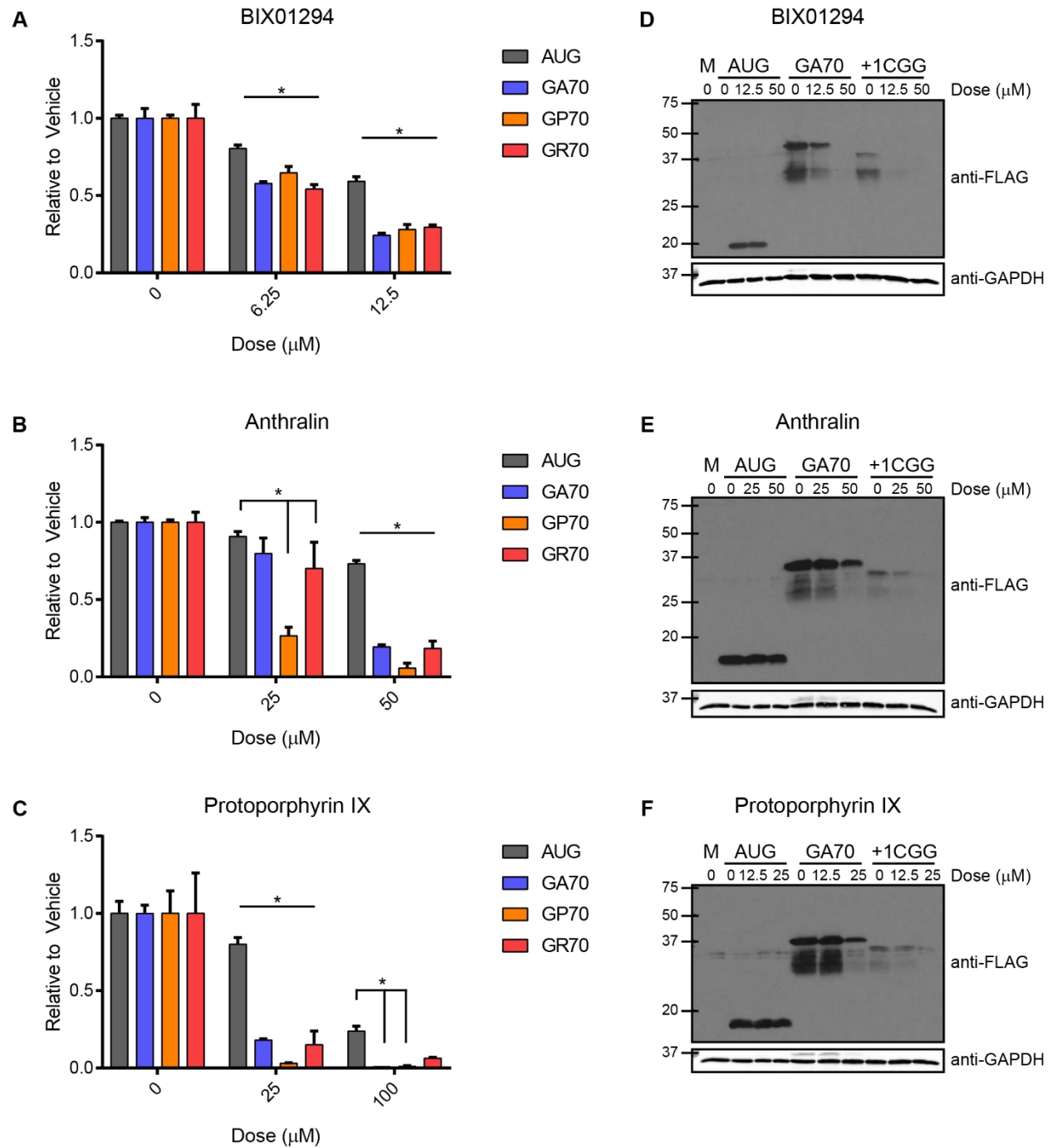
Figure 4

Figure 4. Small molecules inhibit C9RAN translation in all three reading frames

(A) BIX01294, (B) anthralin, and (C) protoporphyrin IX were added at increasing concentrations to 30% RRL *in vitro* translation reactions with AUG-NLuc or GGGGCCx70 reporter mRNAs for all three reading frames, and luminescence was measured relative to vehicle (DMSO) control. (D-F) Selective inhibition of RAN translation of GGGGCCx70 and CGGx100 NLuc reporter mRNAs, relative to AUG-NLuc control, was confirmed by western blot against a c-terminal FLAG tag for the glycine-alanine (GA)

and FMRPolyG (+1CGG) products, following incubation with BIX01294, anthralin, or protoporphyrin IX at indicated doses. GAPDH was used a loading control. For (A-C), all graphs represent mean of n=3 +/- standard deviation. *p < 0.05, 2-way ANOVA with Dunnett's multiple comparison test. M=mock, GA=glycine-alanine, GP=glycine-proline, GR=glycine-arginine.
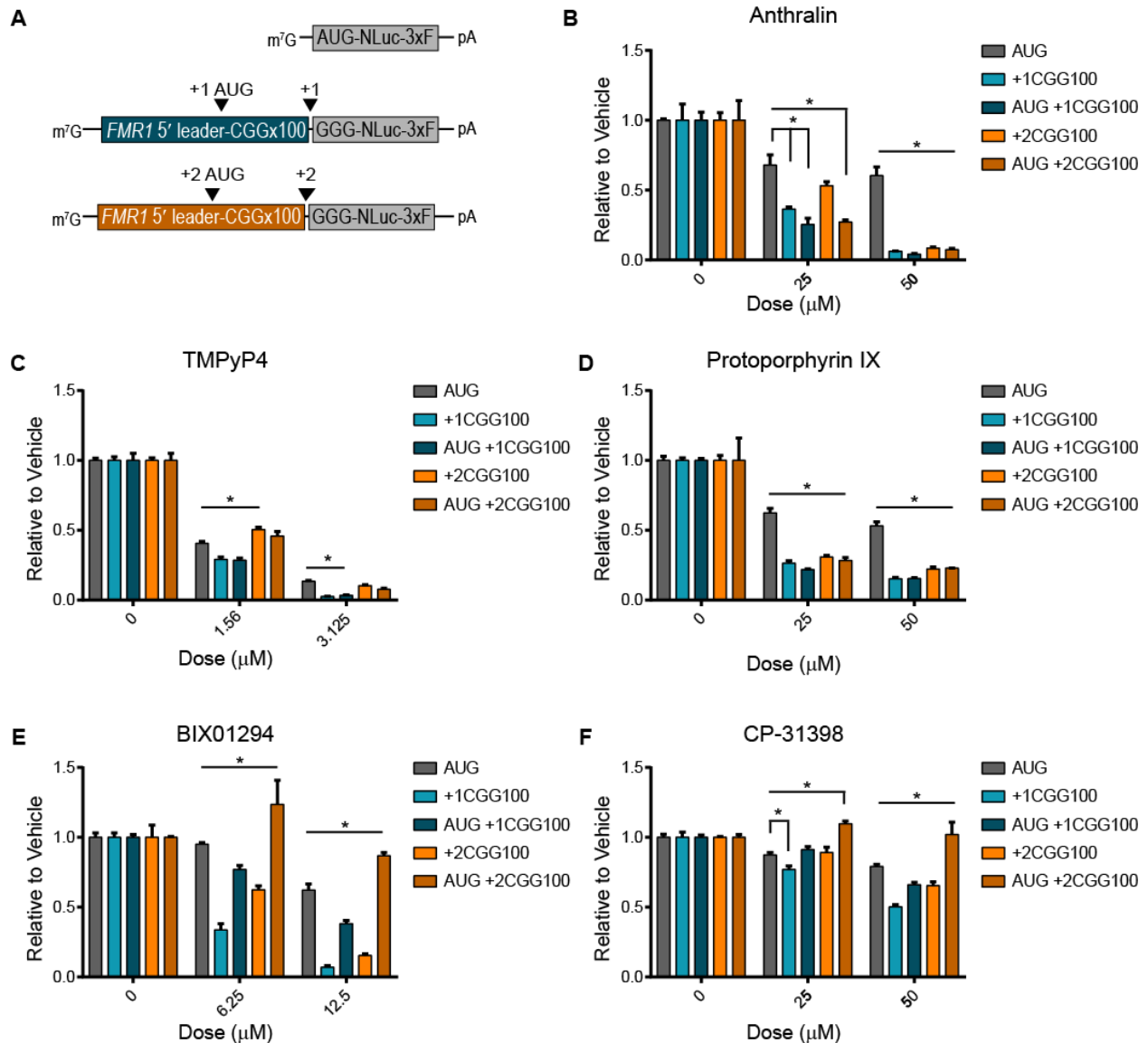
Figure 5

Figure 5. RAN translation inhibitors have varying effects on AUG-initiated repeat translation

(A) Schematic of nanoluciferase (NLuc) reporter mRNAs used in 30% RRL *in vitro* translation assays, illustrating the position of the AUG start codons inserted upstream of the repeats, to drive translation in the indicated reading frames. The effect of (B) anthralin, (C) TMPyP4, (D) protoporphyrin IX, (E) BIX01294, and (F) CP-31398 on +1 and +2CGGx100 RAN translation, compared to AUG-initiated CGGx100 translation, measured in the presence of increasing concentrations of each small molecule. For (B-F), values are relative to DMSO controls, and graphs represent the mean of n=3 +/- standard deviation. *$p < 0.05$, 2-way ANOVA with Dunnett's multiple comparison test. 3xF = 3xFLAG tag.
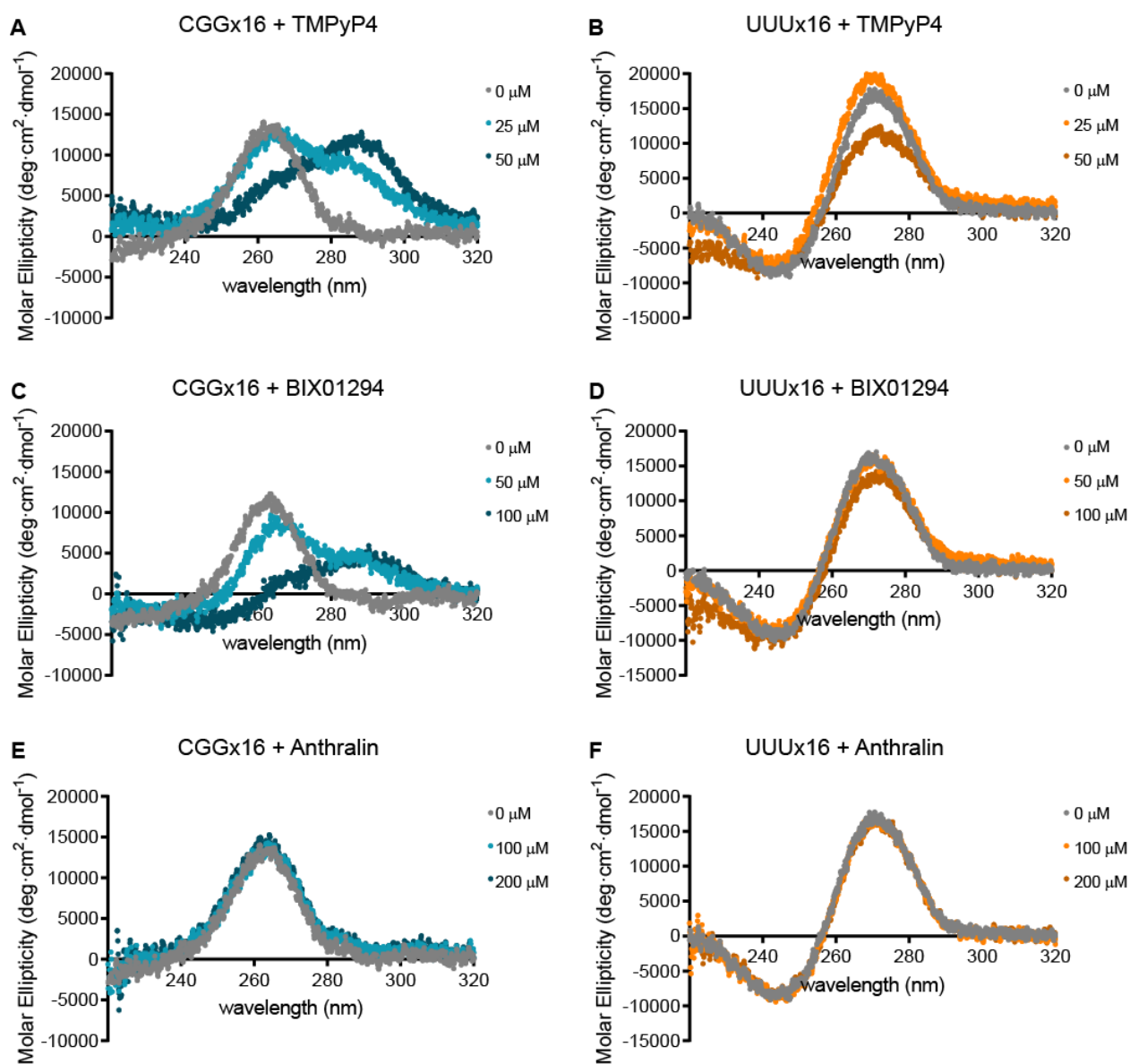
Figure 6

Figure 6. Circular dichroism analysis for small molecule interactions with CGG repeat RNA

Molar ellipticity of 5 µM CGGx16 or UUUx16 RNA, measured between 320-220 nm with circular dichroism (CD), following incubation with TMPyP4 (A-B), BIX01294 (C-D), or anthralin (E-F) at the indicated concentrations. All CD spectra were acquired at 25ºC, with 300 µL RNA samples folded in the presence of 100 mM KCl, and represent the average of six accumulations collected at 20 nm/minute.

Figure 7

A

GGGGCCx8 RNA

B **TMPyP4 - KCl**

C **TMPyP4 - NaCl**

D **BIX01294 - KCl**

E **BIX01294 - NaCl**
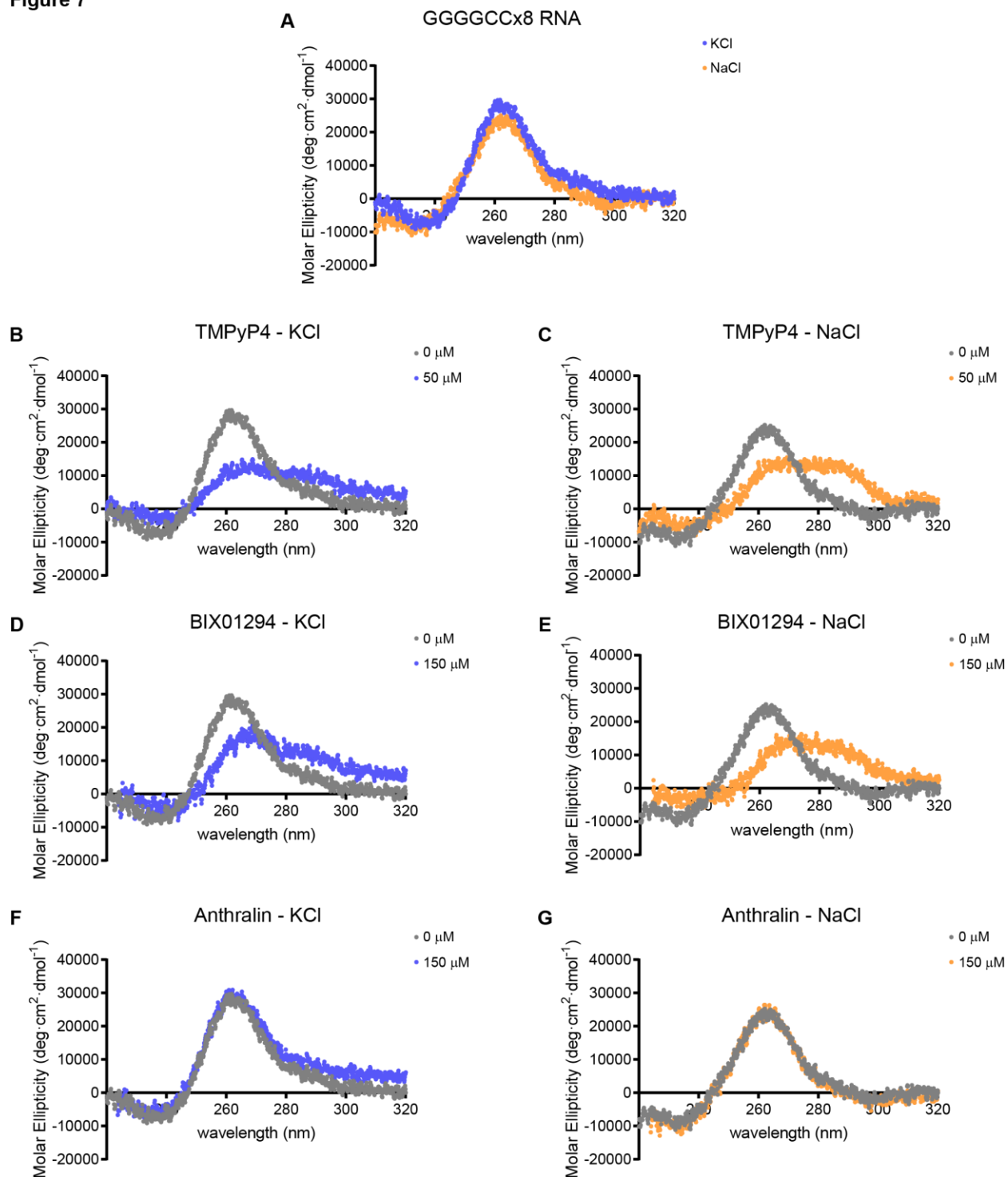
F **Anthralin - KCl**

G **Anthralin - NaCl**

Figure 7. Interaction of small molecule RAN translation inhibitors with GGGGCC repeat RNA by circular dichroism

(A) Comparison of the molar ellipticity of 2.5 µM GGGGCCx8 RNA, measured between 320-220 nm with circular dichroism (CD), following folding in the presence of 100 mM KCl (blue) or 100 mM NaCl

(orange). (B-G) CD spectra of GGGGCCx8 RNA folded in the presence of either 100 mM KCl (blue) or 100 mM NaCl (orange), following incubation with the indicated concentrations of TMPyP4 (B-C), BIX01294 (D-E), or anthralin (F-G). All CD spectra were acquired at 25ºC, with 250 µL RNA samples, and represent the average of six accumulations collected at 20 nm/minute. The untreated RNA spectra in (A) are used for comparison to the RNA+compound spectra in (B-G).
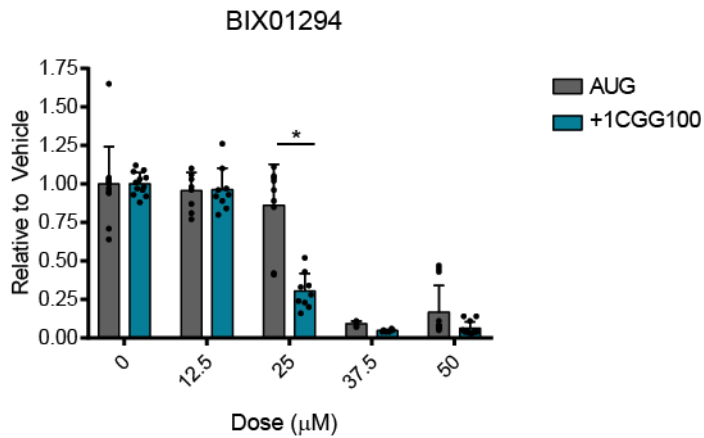
Figure 8

Figure 8. Effect of BIX01294 on +1CGG RAN translation in HEK293 cells

NLuc levels from HEK293 cells transfected with AUG-NLuc or +1CGGx100-NLuc expressing plasmids and treated with BIX01294 at the indicated doses for 24 hours. Luminescence values are expressed relative to DMSO treated cells (0 µM). Graphs represent mean +/- standard deviation, with each n represented by a dot. n=12 for 0 and 50 µM doses, n =9 for 12.5 and 25 µM doses, and n=6 for 37.5 µM, from 2-4 independent experiments with n=3 each. *p < 0.0001, 2-way ANOVA with Sidak's multiple comparison test.

High-throughput screening yields several small-molecule inhibitors of repeat-associated non-AUG translation

Katelyn M Green, Udit Sheth, Brittany N Flores, Shannon E Wright, Alexandra Sutter, Michael G Kearse, Sami Barmada, Magdalena I Ivanova and Peter K Todd

Access the most updated version of this article at doi: 10.1074/jbc.RA119.009951

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts