An antisense noncoding RNA enhances translation via localised structural rearrangements of its cognate mRNA

Rodrigo S. Reis1*, Jules Deforges1, Romy R. Schmidt2, Jos H.M. Schippers3 & Yves Poirier*1

1 Department of Plant Molecular Biology, University of Lausanne, 1015 Lausanne, Switzerland
2 Institute of Biology I,RWTH Aachen University, 52074 Aachen, Germany
3 Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany

*Authors for correspondence: rodrigo.siqueirareis@unil.ch, yves.poirier@unil.ch

SHORT TITLE: A mechanism for translational enhancement via RNA-RNA interaction

One-sentence summary: Interaction between PHO1.2 mRNA and its cis-natural antisense transcript enhances translation via a mechanism involving a local conformational shift and disruption of a key inhibitory region.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Yves Poirier (yves.poirier@unil.ch).

ABSTRACT
A large portion of eukaryotic genes are associated with noncoding, natural antisense transcripts (NATs). Despite sharing extensive sequence complementarity with their sense mRNAs, mRNA-NAT pairs elusively often evade dsRNA-cleavage and siRNA-triggered silencing. More surprisingly, some NATs enhance translation of their sense mRNAs by yet unknown mechanism(s). Here we show that translation enhancement of the rice (Oryza sativa) PHOSPHATE1.2 (PHO1.2) mRNA is enabled by specific structural rearrangements guided by its noncoding antisense RNA (cis-NATpho1.2). Their interaction in vitro revealed no evidence of widespread intermolecular dsRNA formation, but rather specific local changes in nucleotide base-pairing, leading to higher flexibility of PHO1.2 mRNA at a key high GC regulatory region inhibiting translation, approximately 350 nucleotides downstream of the start codon. Sense-antisense RNA interaction increased formation of the 80S complex in PHO1.2, possibly by inducing structural rearrangement within this inhibitory region, thus making this mRNA more accessible to 60S. This work presents a framework for nucleotide-resolution studies of functional mRNA-antisense pairs.

© The Author(s) (2021) . Published by Oxford University Press on behalf of American Society of Plant Biologists. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
INTRODUCTION

RNA is a unique informational molecule. In addition to transmitting the genetic code into proteins, RNAs fold into structures that far expand their functional repertoire. Their abilities to fold and interact with another RNA are critical features in gene regulation and are often determined by sequence complementarity (Wahl et al., 2009; Sharma et al., 2016; Mustoe et al., 2018). A large proportion of plant and animal transcripts are noncoding antisense transcripts to mRNAs, and these natural antisense transcripts (NATs) have been characterized primarily for their role in transcription regulation (Faghihi and Wahlestedt, 2009; Wang et al., 2014a). Surprisingly, much less is known about their contribution to the regulation of mRNA translation and other processes in the cytoplasm (Faghihi and Wahlestedt, 2009; Wang et al., 2005). A NAT can be transcribed from the opposing DNA strand at the same genomic locus as its pair transcript (hence, cis-NAT), or it can be transcribed from a separate locus (trans-NAT). The most obvious difference between cis-NATs and trans-NATs is the extent of complementarity to their cognate sense RNA, as cis-NAT pairs display perfect sequence complementarity over large regions, while trans-NAT pairs almost always display imperfect complementarity (Faghihi and Wahlestedt, 2009). NATs, therefore, provide a complex and poorly understood range of possible RNA-RNA interactions, and it is assumed that most NAT-mRNA pairs are likely to form duplexes throughout their overlapping regions. In plants, initial works on NAT-mRNA pairs corroborated this assumption of perfect double stranded RNA (dsRNA) formation resulting in gene silencing by small interfering RNAs (siRNAs) and, hence, NATs were thought to act predominantly as repressors of gene expression (Borsani et al., 2005; Lapidot and Pilpel, 2006; Zhou et al., 2009). However, recent works without a priori bias revealed that NAT expression can either positively or negatively correlate with their mRNA pair, regardless of predicted dsRNA length (Deforges et al., 2019a; Wang et al., 2014b; Bazin et al., 2017). This suggests that many NAT-mRNA pairs must interact in such a way that permits them to avoid silencing and positively regulate gene expression.

Although the known impact of cis-NATs on gene expression is primarily on transcription, some reports have demonstrated their activity as translation regulators, particularly as enhancers (Jabnoune et al., 2013; Carrieri et al., 2012; Tran et al., 2016; Deforges et al., 2019a; Bazin et al., 2017). In rice (Oryza sativa), the phosphate transporter PHOSPHATE1.2 (PHO1.2) mRNA is poorly translated under standard conditions. However, increased expression of its cis-NAT (cis-NATpho1.2) enhances PHO1.2 translation and protein accumulation, with positive consequences for phosphate homeostasis and plant fitness (Jabnoune et al., 2013). Similarly, in mice, ubiquitin carboxy-terminal hydrolase L1 (Uchl1) mRNA translation is enhanced when its cis-NAT (AS Uchl1) is expressed under the control of stress signalling pathways, resulting in higher UCHL1 protein levels (Carrieri et al., 2012). Also, translation of the human RNA-binding motif protein-15 (RBM15) is enhanced by the
expression of its cis-NAT (AS-RBM15) during megakaryocyte differentiation (Tran et al., 2016). Interestingly, these three cis-NATs seem unrelated regarding their sequences and regulatory motifs, and truncation analyses suggest that they have divergent modes of action, i.e. AS Uchl1 requires a small antisense sequence to Uchl1 5' UTR and a non-antisense inverted SINEB2 element, while AS-RBM15 only requires an antisense sequence to RBM15 5' UTR. Although the region of cis-NATpho1.2 essential for enhancing PHO1.2 translation is unknown, this cis-NAT lacks conserved elements such as SINEB2. A large-scale search for cis-NATs that can potentially regulate translation further evidenced an apparent lack of sequence conservation, or common complementarity profile, to explain their activity in modulating mRNA translation (Bazin et al., 2017; Deforges et al., 2019a). Furthermore, no cis-NAT pair has been mechanistically dissected to provide a reference for how an antisense noncoding RNA might enhance translation of a complementary mRNA.

Structures for several noncoding RNAs have been described, including the cis-NAT COOLAIR (Hawkes et al., 2016) associated with the Arabidopsis thaliana FLOWERING LOCUS FLC, as well as the mammalian paraspeckle-related NEAT1 (Lin et al., 2018), cancer-associated HOTAIR (Somarowthu et al., 2015) and X chromosome inactivation Xist (Maenner et al., 2010). However, these structures were reported for noncoding RNAs alone and structures involving interaction between long noncoding RNAs, such as cis-NATs, and their cognate sense transcripts are elusive. It is reasonable to expect that cis-NAT pairs might result in significant dsRNA formation and changes in mRNA structures; however, this poses a difficult question, because perfect dsRNA formation should result in highly constrained structures within a mRNA and, thus, lead to poor translation instead of enhancement (Burkhardt et al., 2017; Mustoe et al., 2018). Structural analysis of cis-NAT and mRNA interactions are, therefore, likely to reveal novel, perhaps non-intuitive, mechanism(s) for translation enhancement.

Here, we determined the structural architecture of PHO1.2 mRNA and cis-NATpho1.2 interaction, and proposed a mechanism for the translation enhancement activity of the cis-NAT. We found that cis-NATpho1.2 has a relatively flexible structure as it undergoes considerable structural changes in the presence of PHO1.2 mRNA. Surprisingly, PHO1.2 retained most of its structural elements in the presence of cis-NATpho1.2, and the changes were specifically restricted to particular regions of its complementarity with cis-NATpho1.2. We found no clear evidence for formation of long stretches of dsRNA between PHO1.2 and cis-NATpho1.2, perhaps explaining how their interaction evades the silencing machinery. Analysis of a short cis-NATpho1.2 truncation that retained translation enhancement activity revealed a critical structural rearrangement of PHO1.2 outside its binding region. Such structural rearrangement occurs in a region of unusually high GC content, and we further showed that this region acts as a strong translation inhibitor. In addition, we found that cis-
NAT$_{\text{pho1.2}}$ enhances translation of PHO1.2 by promoting 80S formation, possibly by removing structural elements that inhibit access of the 60S subunit to the 40S.

RESULTS

Interaction between cis-NAT$_{\text{pho1.2}}$ and PHO1.2 RNAs does not form extensive sense-antisense dsRNA

PHO1.2 mRNA and cis-NAT$_{\text{pho1.2}}$ have perfect complementary sequences over 1193 nucleotides, comprising ~40% of PHO1.2 and ~60% of cis-NAT$_{\text{pho1.2}}$ RNAs (Fig. 1A). Thus, they can potentially base pair extensively with each other. To investigate the interaction between PHO1.2 and cis-NAT$_{\text{pho1.2}}$ RNAs and uncover how this interaction could contribute to cis-NAT$_{\text{pho1.2}}$’s translation enhancement activity, we used the method selective 2′-hydroxyl acylation analysed by primer extension and mutational profiling (SHAPE-MaP) (Siegfried et al., 2014; Smola et al., 2015) to obtain quantitative, single-nucleotide resolution measurements of their RNA structures. SHAPE reactivities are measures of local nucleotide flexibility, which is higher in unpaired nucleotides and lower in more structured regions. 1-Methyl-7-nitroisatoic anhydride (1M7) modifies unpaired nucleotides leading to mutations during reverse transcription and, thus, a mutational profiling is obtained by massive parallel sequencing. PHO1.2 and cis-NAT$_{\text{pho1.2}}$ are expressed specifically in some of the innermost cells of the roots and leaves, namely the vascular tissues (Jabnoune et al., 2013). It is thus challenging to probe PHO1.2 and cis-NAT$_{\text{pho1.2}}$ in vivo without disrupting tissues and enriching for such cells. Indeed, a recent study analysing the rice mRNA structurome (i.e. transcriptome-wide in vivo structural probing) using DMS probing showed PHO1.2 data below threshold for almost every nucleotide (Deng et al., 2018), thus, hindering structural analysis in vivo (Supplemental Fig. 1). Hence, we applied SHAPE-MaP analysis on in vitro folded PHO1.2 mRNA and cis-NAT$_{\text{pho1.2}}$ alone or mixed in an equimolar ratio (Fig. 1B).

The cis-NAT$_{\text{pho1.2}}$ sequence is antisense to PHO1.2 exons 1 to 4, including the 5′UTR, as well as to introns 1 to 3, partially to intron 4, and ~300 nucleotides (nt) of the PHO1.2 promoter region (Fig. 1A). SHAPE-MaP analysis revealed localised changes in cis-NAT$_{\text{pho1.2}}$ structure in the presence of PHO1.2, as compared to its structure without intermolecular interactions with PHO1.2 (Fig. 1C). A perfect dsRNA formation is expected to result in an overall increase in paired nucleotides and, thus, decrease in SHAPE reactivity. Interestingly, there was little correlation between the profile of reactivity changes and whether the cis-NAT was complementary to the sense mRNA. For example, the structure of some regions of cis-NAT$_{\text{pho1.2}}$ that had perfect complementarity to PHO1.2 mRNA, such as exon 4 (position 100-499 nt in the cis-NAT$_{\text{pho1.2}}$), changed little in the presence of PHO1.2.
mRNA, while regions complementary to exon 3 (position 618-908 nt) and exon 1 (position 1151-1597 nt) showed substantial decreased and increased SHAPE reactivity (Fig. 1C, lower panel). Similarly, regions of cis-NAT\textsubscript{pho1.2} that had no complementary sequences in PHO1.2 mRNA, such as the first 100 and last 320 bases, showed little changes in reactivity, while the region complementary to PHO1.2 intron 3, which should be unavailable to form intermolecular interactions with the spliced PHO1.2 mRNA used for probing, showed a particularly marked reduction in reactivity (position 500-617 nt) compared to other regions with complementarity to introns. Complementary sequence to PHO1.2 intron 3 differs to other introns in GC content (~50% in intron 3 and 20-40% in other introns), potentially affecting the strength of intramolecular interactions of cis-NAT\textsubscript{pho1.2} intron-overlapping regions. Interestingly, interaction between cis-NAT\textsubscript{pho1.2} and PHO1.2 resulted in an overall similar entropic state for cis-NAT\textsubscript{pho1.2}, except for a marked increase in entropy in the region centred around position 1200 (Fig. 1C, middle panel).

Structural analysis of cis-NAT\textsubscript{pho1.2} with and without PHO1.2 mRNA indicated that the interaction between these RNAs results in an overall rearrangement of cis-NAT\textsubscript{pho1.2} structure. It appears that certain intramolecular structures are maintained, others are rearranged, and only part of its RNA follow dsRNA formation upon PHO1.2 addition. Consistent with this hypothesis, base pairing probability for cis-NAT\textsubscript{pho1.2} in the absence of PHO1.2 showed probable formation of intramolecular interaction stretches (e.g. stems) between RNA regions hundreds of nucleotides apart in the cis-NAT\textsubscript{pho1.2} linear sequence (Fig. 1C, upper panel), indicating a range of possibilities for structural rearrangements. This supports a model in which cis-NAT\textsubscript{pho1.2} partially interacts with PHO1.2, thus causing disruption of stems and other intramolecular rearrangements. These rearrangements might unpair nucleotides engaged in interactions within stems, hence increasing their reactivity, as well as pair nucleotides that were forming loops, reducing their reactivity.

Next, we applied SHAPE-MaP to PHO1.2 mRNA in the presence and absence of cis-NAT\textsubscript{pho1.2}. PHO1.2 showed a pairing probability pattern with higher complexity within the cis-NAT\textsubscript{pho1.2}-overlapping region (1 to ~1200 nt), as compared with other parts of the RNA (1500 to 3000 nt) (Fig. 1D, upper panel), indicating higher probability for diversity of RNA conformations within the complementary region. The SHAPE reactivity profile of PHO1.2 mRNA alone or with cis-NAT\textsubscript{pho1.2} was very similar throughout the 3’ non-overlapping region starting at approximately 1200 nt (Fig. 1D, lowest panel). In the region of overlap, changes in PHO1.2 mRNA reactivity in the presence of cis-NAT\textsubscript{pho1.2} occurred mostly in limited regions between 530 and 800 nt and between 1100 and 1200 nt (Fig. 1D, lowest panel). These altered regions can also be identified using a smoothing sliding window length ranging from 30 to 100 nt (Fig. S2), while a 55 nt sliding window is shown in Fig. 1. Importantly, the reactivity profile obtained for both PHO1.2 and cis-NAT\textsubscript{pho1.2} was highly
reproducible across biological replicates (Fig. S3). Functional elements are overrepresented in regions with both low SHAPE reactivity (indicating high degree of structure) and low Shannon entropy (indicating well-defined structure) (Siegfried et al., 2014). In the overlapping region with cis-NAT\textsubscript{pho1.2}, \textit{PHO1.2} exhibited low entropy in its 5' UTR (1-174 nt), without striking reactivity differences in the presence of cis-NAT\textsubscript{pho1.2}, and in a region centred around position 800 nt, in which changes in reactivity were observed in the presence of cis-NAT\textsubscript{pho1.2} (Fig. 1D, middle panel).

cis-NAT\textsubscript{pho1.2} and \textit{PHO1.2} interaction induces structural rearrangements

The structural differences upon interaction between cis-NAT\textsubscript{pho1.2} and \textit{PHO1} mRNA were further evaluated using diffBUM-HMM (Selega et al., 2017) (Fig. 2A, B). Note that the reactivity profiles shown in Fig. 1 use smoothed data, using a 55-nucleotide window, to globally show the effect of their intermolecular interactions, and cannot be directly compared to the diffBUM-HMM analysis, which is a statistical analysis at nucleotide-resolution. In cis-NAT\textsubscript{pho1.2}, diffBUM-HMM analysis revealed contrasting changes in reactivity upon interaction with \textit{PHO1} mRNA (Fig. 2A). While the region of cis-NAT\textsubscript{pho1.2} with complementarity to exon 4 showed relatively little overall change in reactivity, nucleotides complementary to other exons showed a greater level of changes. For example, nucleotides complementary to exon 2 showed predominantly a decrease in flexibility upon interaction, whereas those complementary to exon 1 showed regions with an increase and decrease in flexibility.

In \textit{PHO1.2}, the addition of cis-NAT\textsubscript{pho1.2} resulted in mostly increased reactivity in the first ~800 nucleotides corresponding to the first three exons. In sharp contrast, nucleotides matching exon 4 (between ~800 to ~1200 nt) resulted mostly in decreased reactivity (Fig. 2B, C). Beyond the region of complementarity, only scattered changes in nucleotide reactivity occurred. The sharp decrease in SHAPE reactivity occurring throughout \textit{PHO1} exon 4 is striking in view of the weak changes in reactivity that were observed in the complementary region of cis-NAT\textsubscript{pho1.2}, indicating that such changes are unlikely the results of the formation of dsRNA between sense and antisense RNA in this region. Together, these results indicate that \textit{PHO1.2} and cis-NAT\textsubscript{pho1.2} do not extensively form dsRNA over their ~1kb complementary nucleotides. Instead, their interaction induces changes that are consistent with a model in which intramolecular interactions are hierarchically dominant over intermolecular RNA-RNA interactions.

The cis-NAT\textsubscript{pho1.2} 3' end is sufficient for its translation enhancement activity
To identify the most relevant sequences within cis-NAT\textsubscript{pho1.2} for its activity, we studied the effects of various truncations on translation enhancement of \textit{PHO1.2} \textit{in vivo} (Fig. 3). Truncations were produced and co-expressed with a dual luciferase system expressing \textit{PHO1.2} exon 1 to 4, encoding the hydrophilic SPX domain (PHO1\textsubscript{spx}), fused to a foot-and-mouth disease virus 2A (F2A) peptide (Szymczak et al., 2004) and the firefly luciferase (Fig. 3A and B). We have previously demonstrated that this dual luciferase system is a powerful tool to study sense-antisense interaction and translational control (Deforges et al., 2019a, 2019b). We found that transfected rice protoplasts recapitulate cis-NAT\textsubscript{pho1.2} translation enhancement activity on \textit{PHO1.2}, as evidenced by increased relative luminescence in the presence of a full length cis-NAT\textsubscript{pho1.2} (cis-NAT FL), relative to co-expression with an empty vector control (Fig. 3C). We then tested several cis-NAT\textsubscript{pho1.2} truncations and found that \(-900 \text{ nt ("cis-NAT 2") at its 3’ end is sufficient for its translational enhancement activity (Fig. 3C), which is consistent with the higher level of structural changes in this region of cis-NAT\textsubscript{pho1.2} when in presence with \textit{PHO1.2} (Fig. 2A). Interestingly, “cis-NAT 2” retained activity upon further trimming of its 5’ or 3’ end (“cis-NAT 2A”, “2B” and “2C”). Although cis-NAT 2A (996-1597 nt, 602 nt in length) showed translation activity slightly lower than its longer counterpart, cis-NAT 2, it was selected for further experiments because it was the shortest truncation retaining activity (Fig. 3C).

To validate the translation enhancement activity of cis-NAT 2A in transgenic plants, we transformed wild type rice with constructs expressing a similar dual luciferase system as that used for protoplast transfection, with \textit{PHO1.2}\textsubscript{spx} expression driven by the native \textit{PHO1.2} promoter (Jabnoune et al., 2013), as well as a noncoding RNA (i.e. unrelated RNA derived from the vector as “no cis-NAT” control, full length cis-NAT\textsubscript{pho1.2} or cis-NAT 2A truncation) driven by the constitutive CaMV 35S promoter (Fig. 3D). Corroborating previous findings using transgenic rice (Jabnoune et al., 2013), constitutive expression of full length cis-NAT\textsubscript{pho1.2} resulted in increased \textit{PHO1.2} protein levels with no significant changes in mRNA levels, evidenced by increased relative luminescence in several independent transgenic lines, relative to the negative control (Fig. 3E), and further confirmed by immunoblot (Fig. 3F). Similarly, expression of cis-NAT 2A truncation increased \textit{PHO1.2} protein levels, without changes in mRNA, but to a much higher degree (>10 fold) than full length cis-NAT\textsubscript{pho1.2}. Consistent with an increase in mRNA translation, \textit{PHO1.2} mRNA was more strongly associated with ribosomes in the presence of full length cis-NAT\textsubscript{pho1.2} and cis-NAT 2A truncation, as compared with control, and this enrichment was more evident in the presence of cis-NAT 2A (Fig. 3G and 3H). Differences in activity between cis-NAT\textsubscript{pho1.2} and cis-NAT 2A are unlikely to be caused by differences in expression levels, because both noncoding RNAs were expressed to similar levels among transgenic lines (Fig. S4). Furthermore, lower cis-NAT 2A activity in protoplasts (Fig. 3C) than in transgenic lines expressing \textit{PHO1.2} under its native promoter (Fig. 3E), might indicate that
endogenous cell-specific factors play a role in cis-NAT_{pho1.2} activity, given that PHO1.2's expression is restricted to certain cell-types in transgenic plants (i.e. the vascular parenchyma cells (Jabnoune et al., 2013) which are not enriched for in protoplast preparations. Taken together, these results demonstrated that a short portion of the cis-NAT_{pho1.2} 3' end is enough to enhance translation of PHO1.2 in vivo and this effect is stronger than that observed with wild type cis-NAT_{pho1.2} expression.

**Structural rearrangement of PHO1.2 triggered by cis-NAT_{pho1.2} can be recapitulated by the cis-NAT_{pho1.2} 3' end**

Although the cis-NAT 2A truncation enhances PHO1.2 translation, the PHO1.2 sequence complementary to this cis-NAT_{pho1.2} region, essentially exons 1 and 2, shows relatively mild SHAPE reactivity changes in the presence of full length cis-NAT_{pho1.2} (Fig. 2C). Indeed, the effect of cis-NAT_{pho1.2} on PHO1.2 SHAPE reactivity is only pronounced downstream of the overlapping region with cis-NAT 2A, starting at around nucleotide position 650 in exon 3. However, it is possible that cis-NAT 2A truncation alone interacts differently with PHO1.2, as compared with wild type cis-NAT_{pho1.2}, because important structural elements might be missing in this truncation. Thus, we analysed the interaction of the cis-NAT 2A truncation with PHO1.2 using SHAPE-Map and diffBUM-HMM analysis (Fig. 2D). Remarkably, the overall reactivity profile was very similar for PHO1.2 interacting with wild type cis-NAT_{pho1.2} and cis-NAT 2A truncation. This was evident for the changes in reactivity within PHO1.2 exon 3, immediately downstream of the cis-NAT 2A interacting region, as well as within exon 4, extending over ~300 nt downstream of their interacting region. These results suggest that such reactivity changes are not a direct effect of base pairing between cis-NAT_{pho1.2} and PHO1.2. Instead, these changes induced by both cis-NAT_{pho1.2} and cis-NAT 2A truncation might be caused by a specific structural rearrangement of PHO1.2. It can thus be speculated that a more specific effect of cis-NAT 2A truncation on PHO1.2 structure might contribute to its stronger impact on PHO1.2 mRNA translation in transgenic plants.

**cis-NAT_{pho1.2} acts upon a high GC region within PHO1.2 to mediate translation enhancement**

PHO1.2 mRNA GC content is much higher within the overlap with cis-NAT_{pho1.2} than outside it, and this bias is particularly clear within nucleotides that showed differential SHAPE reactivity in the presence of cis-NAT_{pho1.2} or cis-NAT 2A (Fig. 4A, upper panel). GC content greatly influences the translational rate (Qu et al., 2011) and stability of secondary structures (Xia et al., 1998). We thus altered the GC content of PHO1.2 specifically within nucleotides that showed highest reactivity
changes as well as highest GC levels (Fig. 4A, lower panel). We selected a region with an overall increased flexibility in the presence of cis-NAT_{pho1.2} and cis-NAT 2A (475 to 771 nt; R1, for simplicity) and another with an overall decreased flexibility in the presence of cis-NAT_{pho1} and cis-NAT 2A, named R2 (898 to 1194 nt). Both R1 and R2 regions have a similar GC content of ~70%. The codon usage of R1 and R2 were altered to produce an identical PHO1.2_spx protein sequence using DNA with moderate (50%) and high (70%) GC content, yielding four synthetic sequences that replaced the wild type PHO1.2_spx counterpart in the dual-luciferase construct for protoplast transfection. The modified sequences in R1 and R2 have thus poor sequence complementarity to the cis-NAT_{pho1.2}. We found that, in the absence of cis-NAT_{pho1.2}, decreasing the GC content within region R1 had the most pronounced effect on protein levels, indicating enhancement of translation, while changes in GC content in the R2 region had only mild effects (Fig. 4B). Of note, modifications in the R1 and R2 regions to obtain mutants with 50% GC content resulted in an increased use of less frequent codons, while mutants with 70% GC content had similar codon preference (Fig. S5). These results suggest that the high GC content and its likely effect on RNA structure in region R1, and not potential effects of codon preference, play a predominant role in translation control.

Next, we tested the translation enhancement activity of cis-NAT_{pho1.2} and cis-NAT 2A to the various constructs with modified GC content, by co-transfecting protoplasts with a dual-luciferase and either cis-NAT_{pho1.2} or cis-NAT 2A plasmids (Fig. 4C-G). Co-transfection of the PHO1.2_spx wild type sequence with either antisense construct showed an approximate two-fold protein accumulation increase, compared to control without cis-NATs (Fig. 4C). In contrast, co-transfecting cis-NAT_{pho1.2} or cis-NAT 2A with construct expressing PHO1.2_spx R1 with moderate GC content did not further enhance protein accumulation beyond the high level already obtained in the absence of cis-NAT (Fig. 4D). Co-transfecting either cis-NATs with PHO1.2_spx R1 with a modified high GC content had only weak effects on protein accumulation compared to control (Fig. 4E), likely because this sequence alteration disrupts the RNA-RNA interactions necessary for cis-NAT_{pho1.2} activity to counteract the inhibitory high GC content within R1 of PHO1.2. Interestingly, protein production from PHO1.2_spx R2 constructs with either moderate or high GC content were similar to WT PHO1.2_spx, showing robust increase in protein accumulation when co-transfected with either cis-NAT (Fig. 4F and 4G). The observation that cis-NAT 2A and cis-NAT_{pho1.2} had similar effects in these assays further confirmed that the cis-NAT_{pho1.2} 3' region is sufficient to enhance translation of PHO1.2, and that both truncated and full length cis-NATs likely have a similar mode of action. Altogether, these results revealed that cis-NAT_{pho1.2} enhances translation of PHO1.2 via structural changes within a high GC region of PHO1.2 located in a region approximately 350 bp downstream of the ATG start codon. Although the present analysis does not allow us to experimentally identify the specific structural features present in this high GC region that is responsible for translation inhibition,
modelling was used to try to gain insights into the possible structural changes occurring in the various constructs. We folded each mutated PHO1.2 sequence constrained with covariation consensus, and compared them against the folded wild type sequence (Fig. S6). The positive predictive value (PPV) and sensitivity for the R1 GC-50, R1 GC-70, R2 GC-50 and R2 GC-70 structure against the structure for wild type PHO1.2 were 0.88/0.85, 0.88/0.87, 0.84/0.84 and 0.61/0.60, respectively, suggesting that these mutants exhibit and overall similar fold to the wild type PHO1.2, except for the R2 GC-70 mutant. These structural comparisons further showed that the R1 GC-50 and R1 GC-70 mutants showed greater structural changes in the region overlapping with cis-NAT\textsubscript{pho1.2} (i.e. first 1200 nucleotides), as compared to the R2 GC-50 and R2 GC-70 mutants. These structural models indicate that short-range stems located in the 500-800 nt region are more disrupted in R1 GC-50 than in R1 GC-70, suggesting that translation enhancement in these PHO1.2 mutants may involve local, rather than long-distance base pairing rearrangements.

cis-NAT\textsubscript{pho1.2} enhances 80S formation on PHO1.2 mRNA

Canonical eukaryotic translation begins with cap binding by the preinitiation complex, including the 40S ribosomal subunit, followed by mRNA scanning and start codon (AUG) recognition. Recognition of the AUG triggers modifications that arrest the preinitiation complex and promote joining of the 60S subunit to form the 80S initiation complex and, then, elongation starts (Fig. 5A). To study which step(s) of the translation pathway is affected in PHO1.2 by cis-NAT\textsubscript{pho1.2}, we in vitro translated PHO1.2 using wheat germ extract in the presence or absence of cis-NAT\textsubscript{pho1.2}, supplemented with inhibitors of specific translation steps, and separated the free RNA, ribosomes and polysomes on a sucrose density gradient (Fig 5B-D). In the control without inhibitor supplementation, a higher peak eluted in a fraction coinciding with the 80S peak (UV profile) in samples of PHO1.2 mixed with cis-NAT\textsubscript{pho1.2} truncation 2A, relative to PHO1.2 alone (Fig. 5B). The non-hydrolysable GTP analog guanylyl imidodiphosphate (GMP-PNP) was used to abolish the formation of the 80S complex, and cycloheximide (CHX) to inhibit translation elongation. The addition of GMP-PNP resulted in no difference in 40S accumulation in PHO1.2 translated in the presence or absence of cis-NAT\textsubscript{pho1.2} (Fig. 5C), suggesting that cis-NAT\textsubscript{pho1.2} has no effect on translation steps upstream of the formation of the 80S initiation complex. However, translation inhibition by CHX showed a significant increase in the 80S peak in the presence of cis-NAT\textsubscript{pho1.2} (Fig. 5D), suggesting that cis-NAT\textsubscript{pho1.2} enhances the formation of the 80S complex in PHO1.2 mRNA.

To validate this observation, we analysed the PHO1.2 mutants R1 and R2 (see Fig. 4) treated with CHX in the presence or absence of cis-NAT 2A (Fig. 5E and F). In accordance with our previous
results, R1 GC-50 and R1 GC-70 showed a higher 80S peak relative to other mutants and wild type PHO1.2 (Fig. 5E). Upon mixture with cis-NAT\textsubscript{pho1.2}, the higher 80S peak was observed for R1 GC-50 and wild type, relative to PHO1.2 alone, but similar to R1 GC-50 alone (Fig. 5F). These results further evidence that the R1 GC-50 mutant has enhanced translation, probably through a similar mechanism as that employed by cis-NAT\textsubscript{pho1.2} on PHO1.2. To obtain in vivo validation of these findings, we treated protoplasts with CHX 1 h prior to transfection and during incubation after transfection, thus stalling ribosomes as mRNA-80S complexes (Fig. 5G). We observed that both cis-NAT\textsubscript{pho1.2} and its truncation 2A enhance formation of 80S in PHO1.2 mRNA, similar to that found with our in vitro analysis using wheat germ extract. These results show that cis-NAT\textsubscript{pho1.2} enhances translation of PHO1.2 by promoting 80S formation; however, it is possible that it may, in addition, affect translation elongation. We quantified the rate of PHO1.2\textsubscript{spx} synthesis over an 8-hour window in the presence or absence of cis-NAT\textsubscript{pho1.2} and its truncation 2A using our dual luciferase system. The rate of synthesis of the firefly luciferase (Fluc) translationally fused to PHO1.2\textsubscript{spx} was normalised to the rate of synthesis of the control renilla luciferase (Rluc) located on the same plasmid (Fig. 5H). The rate of increase in luminescence produced by the firefly luciferase (Fluc) translationally fused to PHO1.2 was higher in the presence of cis-NAT\textsubscript{pho1.2} and its truncation 2A, relative to PHO1.2 alone (Fig. 5H and Fig. S7). Therefore, these data are consistent with an effect of cis-NAT\textsubscript{pho1.2} on translation elongation; however, this might be either a direct effect or consequence of enhanced 80S formation.

DISCUSSION

An important and largely unanswered question in biology is how NATs regulate translation of their cognate sense mRNAs. In this study, we showed that RNA structures play a key role in the interaction and biological function of a particular sense-antisense pair. Although the cis-NAT\textsubscript{pho1.2}-PHO1.2 pair displays over 1 kb of perfect nucleotide complementarity, there is no evidence for formation of long sense-antisense dsRNA stretches. While PHO1.2 retains most of its structural elements outside its complementarity with cis-NAT\textsubscript{pho1.2}, it shows specific and only local changes within some regions of complementarity. Similarly, cis-NAT\textsubscript{pho1.2} showed only local structural changes in the presence of PHO1.2, with some regions more constrained and others more flexible. The PHO1.2-cis-NAT\textsubscript{pho1.2} interaction, therefore, seems to primarily result in structural rearrangements within each transcript without major formation of intermolecular dsRNA stretches. Unfortunately, the current approaches to model RNA structures cannot provide a biologically meaningful co-folding of two long sense-antisense pairs of RNAs, because the complementary
nucleotides as paired are energetically favoured and, thus, aberrantly placed in long stretches of perfect dsRNA, as observed for PHO1.2 and cis-NAT_{pho1.2} (Fig. S8).

Full-length cis-NAT_{pho1.2} and the truncated cis-NAT 2A have similar effects on PHO1.2 structural changes, particularly between nucleotides ~500 and ~1000 (exon 3 and part of 4). Since cis-NAT 2A is not complementary to PHO1.2 exon 3 and 4, the structural changes within these regions are unlikely to be caused by direct base pairing between PHO1.2 and cis-NAT_{pho,2}, but rather result from structural rearrangements. Apparently, the critical structural rearrangement caused by the presence of cis-NAT_{pho1.2} in PHO1.2 involves a high GC content region of approximately 300 nt within its exon 3 (R1 region), located ~350 bases downstream of its start codon. PHO1.2 has a highly complex pairing probability that coincides with its high GC content region (Fig. 1D), suggesting that this region is particularly amendable to conformational changes. Indeed, translation enhancement could be recapitulated by mutating this high GC content in region R1. Hence, cis-NAT_{pho1.2} appears to enable specific structural rearrangements involving primarily the PHO1.2 R1 region, possibly via disruption of long-distance interactions, thereby releasing some of the negative effect of the high GC in this structured region on translation efficiency. This hypothesis is also supported by the observed enhancement effect of 80S formation by cis-NAT_{pho1.2} on PHO1.2, indicating that in the absence of cis-NAT_{pho1.2}, PHO1.2 RNA is less accessible to the 60S ribosomal subunit, since association with the 40S subunit is unaffected by cis-NAT_{pho1.2}. The relatively weak increase in 80S formation in wheat germ extract and in protoplast might indicate a limitation of these systems to fully recapitulate the effect of cis-NAT_{pho1.2} on PHO1.2 translation, since a stronger effect was observed in transgenic plants where both PHO1.2 and cis-NAT_{pho1.2} mRNAs were expressed in vivo in their native cell-type context. This observation indicates that cell-specific factors could potentially play a role in modulating the activity of cis-NAT_{pho1.2} on PHO1.2 translation. A proposed model for how cis-NAT_{pho1.2} enhances translation of PHO1.2 is illustrated in Figure 6. PHO1.2 is poorly translated possibly because its mRNA folds into structures that inhibit the binding of the 60S subunit to the 40S, with a key inhibitory region with high GC content being located downstream of its start codon. In the presence of cis-NAT_{pho1.2}, sense and antisense RNAs interact in such a way that leads to structural rearrangements in PHO1.2 mRNA, resulting in the release of the inhibitory effect of this high GC region. Thus, PHO1.2 three-dimensional conformation is specifically altered by the cis-NAT_{pho1.2}, resulting in increased 80S initiation complex formation and translation initiation, as well as potentially increasing translation elongation.

Our working model shares commonalities with other RNA-based regulations, such as riboswitches and small RNA activators. Riboswitches are highly structured regulatory elements controlling gene expression by inducing RNA structural changes upon metabolite or nucleic acid binding (Serganov...
and Nudler, 2013). For instance, in the cobalamin riboswitch, binding of adenosylcobalamin (AdoCbl) to btuB mRNA triggers a specific structural rearrangement, precluding binding to the ribosome binding site (RBS) sequence, thus inhibiting translation initiation (Nou and Kadner, 2000). AdoCbl binds the btuB 5' UTR at a position ~200 nt upstream of the AUG, inducing inhibitory structural changes involving formation of loop-loop interaction (kissing-loop) and substantial rearrangement of nucleotides within stems and loops. Such complex structures and their rearrangement are increasingly being identified as new technologies have allowed deeper investigation of RNA structures (Breaker, 2018; Bocobza et al., 2013). Similar involvement of small RNAs in structural rearrangement of inhibitory stem-loop in mRNA has also been described in the rpoS gene in Gram-negative bacteria (Majdalani et al., 2002). Given the large number of mRNA-NAT pairs across genomes, we envisage that the work and tools presented here will provide a framework for exploring the regulatory mechanisms governing sense and antisense RNA-RNA interactions, including in translation modulation.

METHODS

RNA Synthesis, Chemical Probing, and SHAPE-Map Analysis

PHO1.2 (AK100323) and cis-NATpho1.2 (AK071338) full-length RNAs, as well as cis-NATpho1.2 2A truncation, were synthesised using the HiScribe T7 RNA synthesis kit (NEB, E2040S). RNA samples were folded, probed and sequenced as previously described (Smola et al., 2015) with minor modifications. Five pmol of either PHO1.2 or cis-NATpho1.2 RNA were melted at 95°C in 0.5X standard TE buffer for 3 min, cooled on ice for 2 min, and folding buffer (at final concentration of 40 mM HEPES KOH pH 8.0, 10 mM MgCl2 and 100 mM KCl) was added. RNAs were folded at 22°C for 30 min before probing. For RNA-RNA interaction, 2.5 pmol of each RNA was folded separately, as described above, but for 15 min instead, mixed to a final RNA concentration of five pmol, and incubated at 22°C for a further 15 min before probing. Nucleotide probing was performed immediately after RNA folding by adding 1M7 (MedChemExpress, HY-D0913) to a final concentration of 10 mM, or equal volume of DMSO to untreated control. In all RNA probing assays, the final RNA concentration was 170 nM. Folded RNA samples were probed at 22°C for 3 min, fragmented by incubation at 94°C for 4 min, and desalted using G-50 columns. Reverse transcription of fragmented RNA was performed using 200 ng of RNA, 200 ng of random primer (N9), MaP buffer (250 mM Tris pH 8.0, 375 mM KCl, 50 mM DTT, 2.5 mM dNTP mix, and 15 mM MnCl2) and SuperScript II at 42 °C for 3 hours. Samples were desalted using G-50, and, to ensure strand-specificity of our library, the second-strand was produced using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760). Strand-specific RNA sequencing was provided by the Lausanne Genomic Technologies Facility (LGTF), at the University of Lausanne, using Illumina’s
HiSeq 4000 SR. The final dataset includes two biological replicates for the combination of PHO1.2 and cis-NAT\textsubscript{pho1.2}.

**Strand-specific SHAPE-Map data analysis**

The standard SHAPE-Map data analysis using ShapeMapper2 is not suitable for analysis of strand-specific sequencing data. Thus, we used an in-house developed script to split our sequencing results into plus and minus strands, and analysed these single-stranded datasets separately using ShapeMapper2 without further modifications. The script is publicly available at Github (https://github.com/jdeforges/filter_fastq_strand.py). Detailed analysis produced by ShapeMapper2, including mutation rate, read depth, and statistical output for all samples are presented in Supplemental Data Set 1. ShapeMapper2 outputs normalized reactivity profile in a standard scale, in which reactivity values <0.4 correspond to unreactive or poorly reactive nucleotides, 0.4-0.5 correspond to moderately reactive nucleotides, and >0.85 correspond to highly reactive nucleotides (Hajdin et al., 2013). Secondary structures and Shannon Entropy were calculated using SuperFold (Siegfried et al., 2014) and structures were visualized using RNAstructure Structure Editor (Bellaousov et al., 2013).

Statistical diffBUM-HMM analysis was performed using BUM-HMM (Selega et al., 2017) modified as described elsewhere (https://doi.org/10.1101/2020.07.30.229179). The scripts are available at https://git.ecdf.ed.ac.uk/sgrannem/diffbum-hmm. The analysis was performed without modifications to the original scripts.

**Pairwise covariation analysis and RNA structure modelling**

\textit{PHO1.2} genomic sequences in rice accessions were obtained from the Rice SNP-seek database (Mansueto et al., 2017). Genomic sequences were aligned using Clustal Omega (Sievers et al., 2011) and transcript sequences extracted from the alignment using EMBOSS (Rice et al., 2000). Statistical evaluation of pairwise covariation of structural models was performed using R-Scape (Rivas et al., 2016) against aligned transcript sequences from all available rice accessions from the Rice SNP-seek database (accessed in August 2020). CaCoFold (Rivas, 2020), with the --fold argument in R-Scape, was used to calculate the structure that includes the largest possible number of significantly covarying pairs, i.e. the maximum-covariation optimal consensus structure. R-chie was used to generate all other arc-plotted structures (Lai et al., 2012). RNA structure models were calculated using Superfold (Siegfried et al., 2014) or RNAfold (Hofacker, 2003). For co-folding
analysis, PHO1.2 and cis-NAT$_{pho1.2}$ RNAs were inputs to RNAcofold (Bernhart et al., 2006). Structural models were compared using the scorer utility in RNAstructure to calculate PPV and sensitivity (Seetin and Mathews, 2012).

Translation assay in transfected rice protoplasts

Plasmids used for protoplast transfection have been previously described and validated for translation assay (Deforges et al., 2019a, 2019b), and genes of interest for this work were cloned using Gateway® technology. Rice protoplasts were isolated and transformed as previously described (Zhang et al., 2011). Briefly, rice seeds were surface-sterilized with 0.05 % (v/v) Tween-20 and 50 % (v/v) hypochlorite solution for 3 min, washed with sterile water, and germinated and cultivated on water-soaked paper in a Magenta GA-7 box (Sigma-Aldrich V8505) at 28°C, for seven days in the dark, and then three days under a 12-h light /12-h dark cycle (Philips warm white fluorescent tubes with light intensity of 75 μmol m$^{-2}$ s$^{-1}$). The stems and sheathes of seedlings were cut into small strips (~0.5 mm) and incubated in solution containing cell wall digesting enzymes (1.5% cellulose RS, 0.75% macerozyme R-10, 0.6 M mannitol, 20 mM MES pH 5.7, 20 mM KCl and 10 mM CaCl$_2$) for 3–5 h in the dark with gentle shaking. The released protoplasts were collected and washed with W5 solution (2 mM MES pH 5.7, 154 mM NaCl, 125 mM CaCl$_2$ and 5 mM KCl), pelleted again and resuspended in MMG solution (4 mM MES pH 5.7, 0.4 M mannitol and 15 mM MgCl$_2$) before PEG-mediated transfection. Protoplast transfection was performed by combining ~10$^5$ protoplasts, 3-5 μg of dual luciferase plasmid and corresponding equimolar amounts of co-transfected plasmids, and PEG solution (40% PEG 4000, 0.2 M mannitol and 100 mM CaCl$_2$). After replacing the PEG solution with W5 solution by five consecutive washings, protoplasts were kept in the dark for approximately 16 h at 21 °C. For time-lapse analysis (16h and 24h), about a ten-fold higher amount of protoplast (~10$^6$ protoplasts) was transfected and transformed as above, then split (equal volumes) into ten new tubes and incubated as above. Transformed protoplasts were harvested by centrifugation at 6000 xg and 4°C for 1 min, and resuspended in 1X Passive Lysis buffer (Promega, E1941). The lysate was cleared by centrifugation at 6000 xg and 4°C for 1 min, and used for luminescence quantification using a Dual-Glo Luciferase Assay System (Promega, E2920). Luminescence values for firefly luciferase (Fluc) were normalized against renilla luciferase (Rluc). Statistically significant differences (one-way ANOVA, p-value<0.05) in relative luminescence (ratio Fluc:Rluc) were used to assess co-transfection effects.
Production and analysis of transgenic rice plants

Constructs used for transgenic rice transformation were assembled using a BsaI-based Golden Gate approach (Engler et al., 2008) and their sequences were deposited in GenBank (MN519522). Rice transformation was performed by the plant transformation facility of Iowa State University (http://www.biotech.iastate.edu/ptf/). Several independent hygromycin-resistant rooted plantlets were received and analysed. For relative luminescence quantification, leaf samples were ground in liquid nitrogen and powder was resuspended in 1X Passive Lysis buffer (Promega, E1941). The lysate was cleared by centrifugation at 6000 xg and 4°C for 1 min, and used for luminescence quantification using a Nano-Glo Dual-Glo Luciferase Assay System (Promega, N1610). Luminescence values for nano luciferase (nLuc) were normalized against firefly luciferase (Fluc). Statistically significant differences (Student’s t test, p-value<0.05) in relative luminescence (ratio Fluc:Rluc) were used to assess co-transfection effects. For immunoblot analysis, aliquots of protein extracted for luminescence analysis were separated by SDS-PAGE electrophoresis, transferred to a PVDF membrane and incubated with an anti-HA (Roche, 11867423001, at 1:7,000 dilution), followed by incubation with an anti-rat secondary antibody (Sigma, AP136A, at 1:25,000 dilution) before image capture. For transcript relative quantification, total RNA was extracted from rice leaves using an RNA purification kit as described by the manufacture (Jena Bioscience, PP-210), and reverse transcribed using the M-MLV Reverse Transcriptase (Promega, M3681) and oligo d(T)15 as primer following the manufacturer’s instructions. qPCR analysis was performed on a QuantStudio 3 real-time PCR instrument (ThermoFisher) using SYBR select Master Mix (Applied Biosystems, 4472908) with primer pairs specific to genes of interest, as well as to ACT2 used as reference. Log2 fold changes were calculated by the ΔΔCt method. Primer sequences are listed in Supplemental Table 1.

Polysomal RNA extraction and analysis

Rice leaves were frozen, ground and the polysomes were extracted essentially as previously described (Mustroph et al., 2009) with minor modifications. Briefly, the powder was resuspended in 2 volumes of polysome extraction buffer [200 mM Tris pH 9.0, 200 mM KCl, 1 % (w/v) deoxycholate, 25mM EGTA, 1 % (v/v) detergent mix containing equal proportion of Brij-35, Triton X-100, octylphenyl-polyethylene glycol and Tween-20, 1 % (w/v) polyoxyethylene 10 tridecyl ether, 35 mM MgCl2, 1 mM DTT and 100 ug/ml cyclohexomide], and incubated on ice for 15 min. The mixture was then centrifuged at 4°C for 15 min at 20,000 g, and the supernatant was layered on top of a 10 mL sucrose cushion [60 % (w/v) sucrose, 400 mM Tris pH 9.0, 200 mM KCl, 25 mM EGTA, 35 mM MgCl2, 1 mM DTT and 100 ug/mL cyclohexomide] and centrifuged for 3 h at 170,000 xg and 4°C.
The pellet was resuspended in 100 μL of the same cushion buffer without sucrose, and loaded on top of 5 mL 10-50 % continuous sucrose gradients. The gradients were centrifuged for 75 min at 237,000 xg and 4°C in swinging buckets and fractionated using a gradient holder apparatus (Brandel) into 12 fractions. During gradient fractionation, the UV absorbance was continuously monitored to detect the position of the different complexes within the gradient. RNA was extracted from each collected fraction using the RNA Clean and Concentrator-25 kit following the manufacture’s instructions (Zymo Research, R1017). Reverse transcription and qPCR were performed as described above. Polysome association was calculated as the relative proportion of mRNA in each fraction of the gradient, as previously described (Faye et al., 2014).

**In vitro translation and polysome association**

Capped and polyadenylated PHO1.2 full length mRNA was synthesized using the mMESSAGE mMACHINE T7 transcription kit (ThermoFisher, AM1344) and Poly(A) Polymerase Tailing Kit (Cambio, C-PAP5104H). [α-³²P]UTP labelling was performed during in vitro transcription and *in vitro* translation was performed using the RTS 100 Wheat Germ CECF (Biotechrabbit, BR1401001), according to the manufactures’ protocol. Unlabelled, capped and polyadenylated cis-NAT 2A RNA was synthesized as above. In vitro translation was performed with wheat germ lysate supplemented with labelled and unlabelled (or negative control) RNA was incubated at 22°C, with shaking at 250 rpm for 16 h. For in vitro translation reactions performed in the presence of translation inhibitors, wheat germ reactions were supplemented with either 1 mM GMP-PNP and 1.5 mM cycloheximide (GMP-PNP treatment) or 1.5 mM cycloheximide alone (cycloheximide treatment) before the start of incubation. The final reaction (50 μL) was directly loaded onto a 10-50% sucrose gradient, supplemented with the same amounts of translation inhibitors to prevent loss of ribosome binding, and centrifuged for 90 min at 237,000 xg and 4°C in swinging buckets. Polysome fractionation and sample collection were performed as described above for rice leaf material, with the exception that the time of centrifugation was adjusted to ensure that the 40S subunit for the for GMP-PNP treatment (5 hours of centrifugation) and the 80S for the CHX treatment (3 hours of centrifugation) were eluted in the same fraction. Collected sample fractions were spotted on a nitrocellulose membrane and revealed using a Typhoon phosphorimager. Ribosome (40S or 80S) association was calculated as the relative proportion of labelled mRNA in each fraction of the gradient by quantifying the signal intensity of each fraction on scanned membrane using Fiji (ImageJ).
Mutants and codon usage frequency

PHO1.2 sequence mutated for GC content was synthetized by GeneScript. Codon usage frequency was calculated based on reported frequencies for nucleotide triplets in *Oryza sativa* (Nakamura, 2000).

ACCESSION NUMBERS
Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers AK100323 and AK071338 for PHO1.2 mRNA and the cisNAT<sub>pho1.2</sub>, respectively. Shape-Map fastq files are available from NCBI under the BioSample accession number PRJNA575693.

Supplemental Data

Supplemental Figure 1. DMS *in vivo* probing data for PHO1.2.

Supplemental Figure 2. PHO1.2 SHAPE reactivity smoothed using different sliding windows.

Supplemental Figure 3. DeltaSHAPE analysis using individual biological replicates.

Supplemental Figure 4. Expression of cis-NAT<sub>pho1.2</sub> in transgenic lines.

Supplemental Figure 5. Codon usage frequency in PHO1.2 mutated sequences.

Supplemental Figure 6. Model for structural effect of mutations on high GC content region on PHO1.2.

Supplemental Figure 7. Time-lapse using dual luciferase in rice protoplast transfected with PHO1.2 alone or co-transfected with cis-NAT<sub>pho1.2</sub> or 2A truncation.

Supplemental Figure 8. Co-folding of PHO1.2 and cis-NAT<sub>pho1.2</sub>.

Supplemental Table 1. Primer list.

Supplemental Data Set 1. ShapeMapper2 output and statistical analysis.

AUTHOR CONTRIBUTIONS

RSR, JD and YP conceived the project. RSR performed all experiments. JD developed script for strand-specific Shape-Map analysis and contributed to bioinformatic analysis. JHMS and RRS helped establish the rice protoplast assays. RSR, JD and YP analysed the data. RSR and YP wrote the manuscript, and JD, JHMS and RRS read and provided feedback on it.
ACKNOWLEDGMENTS

The authors thank Ivo Hofacker for critical discussion, Yiliang Ding for providing DMS data for PHO1.2, and Joaquin Clua for reading the manuscript. The authors are also grateful to Sander Granneman (University of Edinburgh) for help with the diffBUM-HMM analysis. The expertise of the Genomic Technologies Facility of the University of Lausanne is also acknowledged. This work was supported by grants from the Swiss National Foundation to YP (Sinergia grant CRSII3_154471 and 31003A-159998).

COMPETING INTERESTS

The authors declare no competing interest.

REFERENCES

Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M., and Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc. Natl. Acad. Sci. 114: E10018–E10027.

Bellaousov, S., Reuter, J.S., Seetin, M.G., and Mathews, D.H. (2013). RNAstructure: Web servers for RNA secondary structure prediction and analysis. Nucleic Acids Res. 41.

Bernhart, S.H., Tafer, H., Mückstein, U., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2006). Partition function and base pairing probabilities of RNA heterodimers. Algorithms Mol. Biol. 1: 3.

Bocobza, S.E., Malitsky, S., Araújo, W.L., Nunes-Nesi, A., Meir, S., Shapira, M., Fernie, A.R., and Aharoni, A. (2013). Orchestration of Thiamin Biosynthesis and Central Metabolism by Combined Action of the Thiamin Pyrophosphate Riboswitch and the Circadian Clock in Arabidopsis. Plant Cell 25: 288–307.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.-K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell 123: 1279–91.

Breaker, R.R. (2018). Riboswitches and translation control. Cold Spring Harb. Perspect. Biol. 10.

Burkhardt, D.H., Rouskin, S., Zhang, Y., Li, G.W., Weissman, J.S., and Gross, C.A. (2017). Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. Elife 6: 1–23.

Carrié, C. et al. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 491: 454–457.
Deforges, J., Reis, R.S., Jacquet, P., Sheppard, S., Gadekar, V.P., Hart-Smith, G., Tanzer, A., Hofacker, I.L., Iseli, C., Xenarios, I., and Poirier, Y. (2019a). Control of Cognate Sense mRNA Translation by cis-Natural Antisense RNAs. Plant Physiol. 180: 305–322.

Deforges, J., Reis, R.S., Jacquet, P., Vuarambon, D.J., and Poirier, Y. (2019b). Prediction of regulatory long intergenic non-coding RNAs acting in trans through base-pairing interactions. BMC Genomics 20: 601.

Deng, H., Cheema, J., Zhang, H., Woelfenden, H., Norris, M., Liu, Z., Liu, Q., Yang, X., Yang, M., Deng, X., Cao, X., and Ding, Y. (2018). Rice in vivo RNA structurome reveals RNA secondary structure conservation and divergence in plants. Mol. Plant 11: 607–622.

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. PLoS One 3.

Faghihi, M.A. and Wahlestedt, C. (2009). Regulatory roles of natural antisense transcripts. Nat. Rev. Mol. Cell Biol. 10: 637–43.

Faye, M.D., Graber, T.E., and Holcik, M. (2014). Assessment of selective mRNA translation in mammalian cells by polysome profiling. J. Vis. Exp.: 1–8.

Hajdin, C.E., Bellaousov, S., Huggins, W., Leonard, C.W., Mathews, D.H., and Weeks, K.M. (2013). Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. Proc. Natl. Acad. Sci. U. S. A. 110: 5498–5503.

Hawkes, E.J., Hennelly, S.P., Novikova, I. V., Irwin, J.A., Dean, C., and Sanbonmatsu, K.Y. (2016). COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. Cell Rep. 16: 3087–3096.

Hofacker, I.L. (2003). Vienna RNA secondary structure server. Nucleic Acids Res. 31: 3429–3431.

Jabnoune, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q., and Poirier, Y. (2013). A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. Plant Cell 25: 4166–4182.

Lai, D., Proctor, J.R., Zhu, J.Y.A., and Meyer, I.M. (2012). R-CHIE: A web server and R package for visualizing RNA secondary structures. Nucleic Acids Res. 40.

Lapidot, M. and Pilpel, Y. (2006). Genome-wide natural antisense transcription: Coupling its regulation to its different regulatory mechanisms. EMBO Rep. 7: 1216–1222.

Lin, Y., Schmidt, B.F., Bruchez, M.P., and McManus, C.J. (2018). Structural analyses of NEAT1 IncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture. Nucleic Acids Res. 46: 3742–3752.

Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., Sanglier-Cianférani, S., Van Dorsselaer, A., Clerc, P., Avner, P., Visvikis, A., and Branlant, C. (2010). 2-D structure of the a region of Xist RNA and its implication for PRC2 association.
PLoS Biol. 8: 1–16.

Majdalani, N., Hernandez, D., and Gottesman, S. (2002). Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. Mol. Microbiol. 46: 813–826.

Mansueto, L. et al. (2017). Rice SNP-seek database update: New SNPs, indels, and queries. Nucleic Acids Res. 45: D1075–D1081.

Mustoe, A.M., Busan, S., Rice, G.M., Hajdin, C.E., Peterson, B.K., Ruda, V.M., Kubica, N., Nutiu, R., Baryza, J.L., and Weeks, K.M. (2018). Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. Cell 173: 181-195.e18.

Mustroph, A., Zanetti, M.E., Jang, C.J.H.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J. (2009). Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 106: 18843–18848.

Nakamura, Y. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28: 292–292.

Nou, X. and Kadner, R.J. (2000). Adenosylcobalamin inhibits ribosome binding to btuB RNA. Proc. Natl. Acad. Sci. U. S. A. 97: 7190–7195.

Qu, X., Wen, J.-D., Lancaster, L., Noller, H.F., Bustamante, C., and Tinoco, I. (2011). The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature 475: 118–121.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 16: 276–277.

Rivas, E. (2020). RNA structure prediction using positive and negative evolutionary information. bioRxiv: 2020.02.04.933952.

Rivas, E., Clements, J., and Eddy, S.R. (2016). A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. Nat. Methods 14: 45–48.

Seetin, M.G. and Mathews, D.H. (2012). RNA Structure Prediction: An Overview of Methods. In Bacterial Regulatory RNA, K.C. Keiler, ed (Humana Press: Totowa, NJ), pp. 99–122.

Selega, A., Sirocchi, C., Iosub, I., Granneman, S., and Sanguinetti, G. (2017). Robust statistical modeling improves sensitivity of high-throughput RNA structure probing experiments. Nat. Methods 14: 83–89.

Serganov, A. and Nudler, E. (2013). A decade of riboswitches. Cell 152: 17–24.

Sharma, E., Sterne-Weiler, T., O’Hanlon, D., and Blencowe, B.J. (2016). Global Mapping of Human RNA-RNA Interactions. Mol. Cell 62: 618–626.

Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A.E.E., and Weeks, K.M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat. Methods 11: 959–965.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H.,
Remmert, M., Söding, J., Thompson, J.D., and Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7: 539.

Smola, M.J., Rice, G.M., Busan, S., Siegfried, N.A., and Weeks, K.M. (2015). Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. Nat. Protoc. 10: 1643–1669.

Somarowthu, S., Legiewicz, M., Chillón, I., Marcia, M., Liu, F., and Pyle, A.M. (2015). HOTAIR Forms an Intricate and Modular Secondary Structure. Mol. Cell 58: 353–361.

Szymczak, A.L., Workman, C.J., Wang, Y., Vignali, K.M., Dilioglou, S., Vanin, E.F., and Vignali, D.A.A. (2004). Correction of multi-gene deficiency in vivo using a single “self-cleaving” 2A peptide–based retroviral vector. Nat. Biotechnol. 22: 589–594.

Tran, N., Su, H., Khodadadi-Jamayran, A., Lin, S., Zhang, L., Zhou, D., Pawlik, K.M., Townes, T.M., Chen, Y., Mulloy, J.C., and Zhao, X. (2016). The AS-RBM15 IncRNA enhances RBM15 protein translation during megakaryocyte differentiation. EMBO Rep. 17: 887–900.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell 136: 701–718.

Wang, H., Chung, P.J., Liu, J., Jang, I.-C.C., Kean, M.J., Xu, J., and Chua, N.-H.H. (2014a). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. Genome Res. 24: 3.

Wang, H., Chung, P.J., Liu, J., Jang, I.C., Kean, M.J., Xu, J., and Chua, N.H. (2014b). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. Genome Res. 24: 3.

Wang, X.-J., Gaasterland, T., and Chua, N.-H. (2005). Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana. Genome Biol. 6: R30.

Xia, T., and M. E. Burkard, Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C., and Turner, D.H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry 37: 14719–14735.

Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., Wang, P., Li, Y., Liu, B., Feng, D., Wang, J., and Wang, H. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. Plant Methods 7: 30.

Zhou, X., Sunkar, R., Jin, H., Zhu, J.K., and Zhang, W. (2009). Genome-wide identification and analysis of small RNAs originated from natural antisense transcripts in Oryza sativa. Genome Res. 19: 70–78.
Figure 1. *PHO1.2* and cis-NAT*pho1.2* structural architecture. (A) Genomic view of the *PHO1.2* locus. Black boxes and lines denote exons and introns, respectively, and arrow above *PHO1.2* schematic denotes start site for translation and square denotes end of translation. (B) Experimental strategy. *PHO1.2* and cis-NAT*pho1.2* RNAs were linearized and folded independently, and an aliquot of each folded RNA was mixed at equimolar ratio. Structures for RNAs folded alone and upon interaction were analysed using the SHAPE-MaP approach. (C-D) Pairing probability (top panel), Shannon entropy (middle panel) and SHAPE reactivity (lower panel). Pairing probability given for cis-NAT*pho1.2* or *PHO1.2* alone with arcs representing base pairs color-coded by pairing probability. Shannon entropy and SHAPE reactivity shown as the median reactivity over 55-nt sliding windows relative to the global median (log2 scale for SHAPE). Data shown are the mean values for two biological replicates. The nucleotide positions for *PHO1.2* exons overlapping with cis-NAT*pho1.2*, and cis-NAT*pho1.2* overlapping with *PHO1.2*, are highlighted at the bottom (C and D, respectively).
Figure 2. Differential reactivity for \textit{PHO1.2} and cis-NAT\textsuperscript{PHO1.2} interaction. (A) SHAPE reactivities of cis-NAT\textsuperscript{PHO1.2} (top panel) represent the average values (two biological replicates), with positive values (blue) assigned to cis-NAT\textsuperscript{PHO1.2} mixed with \textit{PHO1.2} mRNA while negative values (red) are assigned for cis-NAT\textsuperscript{PHO1.2} alone. The dataset was analysed using differential BUM-HMM. For these panels, positive (blue) and negative (red) values indicate nucleotide positions more and less reactive (flexible), respectively, in the mix of cis-NAT\textsuperscript{PHO1.2} with \textit{PHO1.2} as compared to cis-NAT\textsuperscript{PHO1.2} alone. \textit{PHO1.2} overlapping exons and introns are shown (lower panel, aligned with nucleotide position) (B) Genomic view of \textit{PHO1.2} locus. Arrow above \textit{PHO1.2} schematic denotes start site for translation and square denotes end of translation. (C-D) SHAPE reactivities of \textit{PHO1.2} represent the average values (two biological replicates), with positive values (blue) assigned to \textit{PHO1.2} mRNA mixed with cis-NAT\textsuperscript{PHO1.2} full-length (C) or cis-NAT 2A (D) while negative values (red) are assigned to \textit{PHO1.2} mRNA alone. The dataset was analysed using differential BUM-HMM. For these panels, positive (blue) and negative (red) bars indicate nucleotide positions more and less reactive (flexible), respectively, in the mix of \textit{PHO1.2} with cis-NAT\textsuperscript{PHO1.2} (C) or cis-NAT 2A (D) as compared to \textit{PHO1.2} alone. The nucleotide position for \textit{PHO1.2} overlap with cis-NAT\textsuperscript{PHO1.2} full-length (FL) and cis-NAT 2A truncation is depicted between panels C and D.
Figure 3. Translation enhancement activity of cis-NAT_{pho1.2} truncations. (A) Overview of cis-NAT_{pho1.2} truncations in the context of their overlap with genomic PHO1.2. (B) Schematic of constructs used in protoplast transfection. The CaMV35S promoter was used to drive expression of PHO1.2 SPX domain (PHO1.2_{spx}) cDNA fused to F2A peptide and firefly luciferase (F2A-Fluc), as well as to drive expression of renilla luciferase (Rluc) as loading control from the same plasmid (dual luciferase plasmid). cis-NAT_{pho1.2} truncations and controls (ncRNA) were expressed under the CaMV35S promoter from a separate construct. The F2A peptide is “self-cleaved” during translation; hence, firefly luciferase is found unfused to PHO1.2_{spx} in the cytoplasm, together with free renilla luciferase. (C) Relative luminescence of PHO1.2_{spx} translational fusion in the presence of cis-NAT_{pho1.2} truncations. Dual luciferase (coding) and ncRNA (noncoding) plasmids were combined at equimolar ratios using at least 0.5 pmol of dual luciferase plasmid per protoplast transfection. Relative luminescence is the Fluc:Rluc luminescence ratio. Individual data points represent biological replicates (t-test, p-value < 0.001***, error bar = SD, n ≥ 3). (D-H) Translation enhancement activity of cis-NAT_{pho1.2} truncations in transgenic rice. (D)
Schematic of construct used in rice transformation. PHO1.2 promoter (~2kb upstream of the PHO1.2 transcription start site) was used to drive expression of genomic PHO1.2 SPX domain (PHO1.2_{spx}) fused to nano luciferase (nLuc) and HA tag. The rice ubiquitin promoter (OsUBQ) controls firefly luciferase expression, used as loading control, and the CaMV35S promoter controls expression of ncRNA, i.e. cis-NAT_{pho1.2} full length (FL), 2A truncation and empty vector (no cis-NAT negative control). (E) Relative luminescence and transcript levels of the PHO1.2_{spx} translational fusion. Relative luminescence is the nLuc:Fluc luminescence ratio. Individual data points represent mean values for independent transgenic lines (t-test, p-value < 0.001***, n ≥ 4). Box plot defined by minimum, 25th percentile, median, 75th percentile, and maximum values. (F) Immunoblot validation of protein levels. Two independent lines were used for each construct. Anti-HA was used to detect PHO1.2_{spx}-nLuc-HA fusion and anti-Fluc against firefly luciferase, as loading control. Two times less protein was loaded for lines expressing the cis-NAT 2A truncation. (G) Polysome association of PHO1.2_{spx} fusion mRNA. Line thickness represents the distance between mean values for three independent transgenic plants. (H) Representative UV profiles of polysome association analysis. Highest peak represents 80S ribosomal subunit. UV profiles were manually stacked disregarding actual absorbance values.
Figure 4. Effect of a high GC content region on PHO1.2 translation efficiency. (A) Differential BUM-HMM (diffBUM-HMM) for PHO1.2 in the presence of cis-NAT_{pho1.2} (same as in Fig. 2C) overlaid with PHO1.2 GC content. GC percentage (%GC) shown for a 55-nt sliding windows. Schematic of constructs designed over two regions of PHO1.2 (R1 and R2) with different GC content and identity percentage, relative to respective wild type sequence (lower panel). (B-G) Relative luminescence of PHO1.2px (wild type and mutants modified for GC content) translational fusion in the absence or presence of cis-NAT_{pho1.2} or the cis-NAT 2A truncation. Each point corresponds to a biological replicate (n=6) and letters above boxes refer to statistical similarity (two-sided t-test, p-value < 0.05). Box plot defined by minimum, 25th percentile, median, 75th percentile, and maximum values.
Figure 5. Ribosomal association of PHO1.2 in the presence of cis-NAT<sub>pho1.2</sub>. (A) Schematic of the translation pathway including steps inhibited by the GTP analog 5'-guanylyl imidodiphosphate (GMP-PNP) and elongation inhibitor cycloheximide (CHX). (B-D) In vitro translated PHO1.2 supplemented with GMP-PNP, CHX or control without drug. Radioactively-labelled PHO1.2 mRNA was in vitro translated using a wheat germ extract in the presence and absence of cis-NAT 2A, followed by separation in a sucrose gradient for 3 hours (CHX and control) or 5 hours (GMP-PNP). Peaks corresponding to the 40S or 80S were analysed using a t-test (p-values on the panels, n=3, error bar = SD). (E) In vitro translated PHO1.2 mutants supplemented with CHX and analysed as above. (F) In vitro translated PHO1.2 mutant R1 GC-50 and wild type supplemented with CHX and with or without cis-NAT 2A, analysed as above. (G) PHO1.2 alone or co-transfected with cis-NAT<sub>pho1.2</sub> or cis-NAT 2A truncation in protoplasts supplemented with CHX 1h prior to transfection and during incubation for 16 h. PHO1.2 mRNA was analysed as above, except that it was quantified by RT-qPCR. (B-G) Each point corresponds to a radioactive PHO1.2 signal, or mRNA level by RT-qPCR in (G), in a given fraction, as a percentage of the total level in the corresponding sucrose gradient (n = 3 biological replicates). mRNA percentage over collected fractions (top) (±SD) and UV profile (bottom) are shown. (E-G) Inset graph depicts fraction number 6 only (80S peak) and was analysed using a t-test (p-value < 0.01**, <0.05*, n=3, error bar = SD). (H) Schematic of the dual luciferase construct used in protoplast transfection to express PHO1.2-F2A-Fluc translational fusion and free Rluc. Fluc, firefly luciferase, and Rluc, renilla luciferase. Separate constructs were used for co-transfection with cis-NAT (not shown). (I) Time-lapse using dual luciferase in rice protoplasts transfected with PHO1.2 alone or co-transfected with cis-NAT<sub>pho1.2</sub> or 2A truncation. The Fluc/Rluc (y-axis label for Fluc and Rluc coloured matching the schematic in H) luminescence ratio 16 or 24 hours after protoplast transfection is shown. Curve slopes were analysed using a t-test (p-value < 0.05*, ± SD, n = 4 biological replicates). See Supplemental Data Set1 for statistical analysis.
Figure 6. Proposed model for the cis-NAT_{pho1.2} translation enhancement of PHO1.2. A region in PHO1.2 with a high GC content, located downstream of its start codon, creates a structure inhibiting the binding of the 60S subunit to the 40S. In the presence of cis-NAT_{pho1.2}, a localised sense-antisense inter-molecular interaction leads to a modification of this inhibitory structure, resulting in increased 80S initiation complex formation and enhanced PHO1.2 translation.