RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer

Andrew L. Wolfe, Kamini Singh, Yi Zhong, Philipp Drewes, Vinagolu K. Rajasekhar, Viraj R. Sanghvi, Konstantinos J. Mayrakis, Man Jiang, Justine E. Roderick, Joni Van der Meulen, Jonathan H. Schatz, Christina M. Rodrigo, Chunying Zhao, Pieter Rondoud, Elisa de Stanchina, Julie Teruya-Feldstein, Michelle A. Kelliher, Frank Speleman, John A. Porco Jr, Jerry Pelletier, Gunnar Rätsch & Hans-Guido Wendel

The translational control of oncoprotein expression is implicated in many cancers. Here we report an eIF4A RNA helicase–dependent mechanism of translational control that contributes to oncogenesis and underlies the anticancer effects of silvestrol and related compounds. For example, eIF4A promotes T-cell acute lymphoblastic leukaemia development in vivo and is required for leukaemia maintenance. Accordingly, inhibition of eIF4A with silvestrol has powerful therapeutic effects against murine and human leukaemic cells in vitro and in vivo. We use transcriptome–scale ribosome footprinting to identify the hallmarks of eIF4A–dependent transcripts. These include 5′ untranslated region (UTR) sequences such as the 12–nucleotide guanine quartet (CGG)₄ motif that can form RNA G–quadruplex structures. Notably, among the most eIF4A-dependent and silvestrol–sensitive transcripts are a number of oncogenes, superenhancer–associated transcription factors, and epigenetic regulators. Hence, the 5′ UTRs of select cancer genes harbour a targetable requirement for the eIF4A RNA helicase.

The activation of translation contributes to malignant transformation and is an emerging target for cancer therapies. For example, oncogenic RAS, ERK and AKT stimulate cap-dependent translation via mTORC1 and the eIF4E–binding protein (4E-BP). The initiation of cap-dependent translation involves a tightly controlled multi-protein initiation complex. This includes the cap-binding protein eIF4E, which is abundantly expressed in cancer and can transform fibroblasts and promote tumour development in vivo. Another key factor is the eIF4A RNA helicase, which has a role in scanning the messenger RNA 5′ UTR for a translation start site and has been implicated in the translation of mRNAs with long and complex 5′ UTRs. eIF4A is the molecular target for three distinct natural compounds—silvestrol, hippuristanol and patacamine—and these compounds show promising anticancer activity. Recent studies have defined the helicase action of eIF4A, as well as its protein interactions, and revealed eIF4A to have a role in microRNA action. Exactly how eIF4A affects translation and which specific mRNA features necessitate its helicase activity have not been defined.

Ribosome footprinting provides a snapshot of translation across the transcriptome. It uses deep sequencing to identify ribosome–protected RNA fragments and compare them to total transcript levels. Among the various applications of this technology, the most relevant to this study concern the translational effects of mTORC1 inhibition. We use ribosome footprinting to identify eIF4A-dependent transcripts and their hallmark features.

eIF4A is an oncogene in T-ALL

NOTCH–driven T-cell acute lymphoblastic leukaemia (T-ALL) exemplifies the activation of AKT/mTORC1 and cap-dependent translation in cancer. In 36 paediatric T-ALL samples we find PTEN mutations (14%) and deletions (11%) (Extended Data Fig. 1a–c and Supplementary Table 1). These genetic lesions promote T-ALL development in a murine T-ALL model (Fig. 1a). NOTCH causes T-ALL with a mean latency of 91.5 days (n = 14), knockdown of Pten or expression of the mutant IL7R p.Leu243IleAsnProCys accelerates disease onset (Pten, 47.1 days, n = 10, P < 0.0001; IL7R, 35.5 days, n = 4, P < 0.0001). Remarkably, expression of eIF4E or eIF4A1 similarly accelerates leukaemia development (eIF4E, 30.75 days, n = 4, P < 0.0001; eIF4A1, 33.8 days, n = 5, P < 0.0001) (Fig. 1b and Extended Data Fig. 1d). All T-ALLs are CD4/CD8 double positive, and increased ribosomal 5′ phosphorylation indicates mTORC1 activation in Pten-deficient and IL7R-expressing T-ALLs (Extended Data Fig. 1d, f–i). EIF4E and eIF4A1 are required to maintain T-ALL and cells expressing a constitutive 4E-BP1-encoding allele (4E-BP1(4A)) or an eIF4A1 knockdown construct are rapidly eliminated from mixed populations (Fig. 1c, d and Extended Data Fig. 1e) (P = 0.000002, vector versus 4E-BP1(4A); P = 0.000008, vector versus shRNA (shRNA)-targeted eIF4A (shEIF4A)).

Silvestrol and a synthetic analogue, (+)-CR-31-B, inhibit eIF4A1/2 (Extended Data Fig. 2a, b) and silvestrol induces cell death in primary human T-ALL samples, cell lines, and murine T-ALLs at nanomolar half-maximum inhibitory concentration (IC₅₀) values (Fig. 2b and Extended Data Fig. 2d, e). In vivo, silvestrol is effective against murine or xenografted T-ALLs.

©2014 Macmillan Publishers Limited. All rights reserved
Figure 1 | eIF4A promotes T-ALL development in vivo. a, Diagram of the intracellular NOTCH1 (ICN)-driven murine T-ALL model. HPCs, hematopoietic progenitor cells. b, Kaplan–Meier analysis showing time to leukaemia development after transplantation of HPCs transduced with NOTCH1-ICN and empty vector (black, n = 14), eIF4E (green, n = 4), eIF4A1 (red, n = 5), IL7r p.Leu242_Leu243insAsnProCys (blue, n = 4), shPten (orange, n = 10). c, Experimental design of competition experiments. d, Results as percentage of each starting GFP-positive population of murine T-ALL cells partially transduced with vector/GFP, eIF4A or 4E-BP1(4A). CR, (±)-CR-31-B. Mean and standard deviation (s.d.) are shown, n = 3 biological replicates. (Fig. 2c and Extended Data Fig. 2f–h). In KOPT-K1 tumour-bearing (∼1 cm³) NOD/SCID mice, treatment with silvestrol (0.5 mg kg⁻¹, intraperitoneal (i.p.), days 0–6, n = 7, P < 0.001) or (±)-CR-31-B (0.2 mg kg⁻¹, i.p., days 0–6, n = 8, P < 0.001) delays tumour growth, and causes apoptosis and cell cycle arrest (Fig. 2c, d and Extended Data Fig. 2g, h). Detailed toxicology shows that this treatment is well-tolerated in mice (Extended Data Fig. 3a–j and Supplementary Table 2). Rapamycin induces an S6 kinase-dependent feedback activation of AKT (T308)51; by contrast, silvestrol or (±)-CR-31-B does not trigger this response in KOPT-K1 cells (Fig. 2e, f). This result implies that inhibition of eIF4A is effective without having an effect on S6 kinase.

Ribosome footprinting

For footprinting studies, we treated KOPT-K1 cells with 25 nM of silvestrol (0.5 mg kg⁻¹) or vehicle (dimethylsulphoxide (DMSO)), silvestrol, or (±)-CR-31-B treatment. Mean and s.d. is shown; ND, not done. c. Tumour size of KOPT-K1 xenografts treated with (±)-CR-31-B (0.2 mg kg⁻¹) or vehicle. Mean and s.d. of 5 tumours is shown. d, Immunohistochemical analysis of (±)-CR-31-B-treated KOPT-K1 tumours. H&E, haematoxylin and eosin; TUNEL, TdT-mediated dUTP nick end labelling. e, Diagram of drug targets. f, Lysates of KOPT-K1 cells treated with vehicle (Veh), rapamycin (Rapa; 25 nM), (±)-CR-31-B (CR; 25 nM), or silvestrol (Silv; 25 nM) for 48 h and probed as indicated.

Figure 2 | Silvestrol has single-agent activity against T-ALL. a, Reporter system with capped Renilla luciferase (red) and Firefly luciferase under the HCV IRES (black). Bottom, relative levels of Renilla luciferase (red) and Firefly (black) luciferase upon vehicle (dimethylsulphoxide (DMSO)), silvestrol, or (±)-CR-31-B treatment. Mean and s.d. is shown; n = 3 biological replicates. b, Viability of primary patient T-ALL samples treated with silvestrol (48 h; mean and s.d. of 4 replicates is shown). ND, not done. c, Tumour size of KOPT-K1 xenografts treated with (±)-CR-31-B (0.2 mg kg⁻¹) or vehicle. Mean and s.d. of 5 tumours is shown. d, Immunohistochemical analysis of (±)-CR-31-B-treated KOPT-K1 tumours. H&E, haematoxylin and eosin; TUNEL, TdT-mediated dUTP nick end labelling. e, Diagram of drug targets. f, Lysates of KOPT-K1 cells treated with vehicle (Veh), rapamycin (Rapa; 25 nM), (±)-CR-31-B (CR; 25 nM), or silvestrol (Silv; 25 nM) for 48 h and probed as indicated.
was the most (TE down; red) or least (TE up; blue) affected by silvestrol compared to background (grey) (Fig. 3b and Supplementary Table 3a–c). The TE down group includes 281 mRNAs (220 with annotated 5′ UTRs), TE up includes 190 mRNAs, and the background list includes 2,243 mRNAs. The footprinting methodology also provides positional information and the rDiff algorithm identifies mRNAs with significant shifts in ribosome footprint distribution25. For example, silvestrol caused an accumulation of ribosome footprints in the 5′ UTR of 847 protein-coding transcripts (rDiff positive genes; 641 with annotated 5′ UTRs, \( P < 0.001 \)) (Supplementary Table 3d). Sixty-two transcripts showed both decreased TE and altered ribosome footprint distribution, whereas we observed no change in distribution among TE up genes (Fig. 3c, Extended Data Fig. 5a–c and Supplementary Table 3e).

**The (CGG)\textsubscript{4} motif marks elf4A-dependent mRNAs**

We compared the TE up, TE down and background groups (Supplementary Table 3a–c), and confirm that mRNAs with longer 5′ UTRs were significantly enriched in the TE down group (mean UTR length, 368 nucleotides in TE down; 250 nucleotides in background, \( P = 7.6 \times 10^{-12} \), silvestrol versus control, by two-sample Kolmogorov–Smirnov) (Fig. 3d). The TE up group showed no difference in 5′ UTR length (265 nucleotides; \( P = 0.165 \), silvestrol versus control) (Extended Data Fig. 4j). We found no predilection for known regulatory elements in the TE down group; these included TOP\textsuperscript{26}, TOP-like sequences\textsuperscript{21}, IRES\textsuperscript{27} and pyrimidine-rich translational elements (PRTEs)\textsuperscript{26} (Fig. 3e). The TE up group showed the expected enrichment in IRES elements (Extended Data Fig. 4k). Using the DREME algorithm\textsuperscript{28}, we found a striking enrichment of a 12-nucleotide (CGG)\textsubscript{4} motif and related but shorter 9-nucleotide motifs among the 220 TE down genes (Fig. 3f, g and Supplementary Table 4a, b). Both the 12-nucleotide and 9-nucleotide motifs were significantly enriched over background and their frequency was significantly higher than expected merely by the longer UTR length (\( P < 2.2 \times 10^{-16} \), one-sided binomial test) (Fig. 3h). We found no enriched motif in the TE up group. Transcripts with an altered ribosome footprint distribution (rDiff positive) showed similar features. These included longer 5′ UTRs (rDiff, mean 271 nucleotides versus background: mean 230 nucleotides; \( P = 0.004 \)) (Extended Data Fig. 5d) and absence of enrichment for TOP, TOP-like, PTRE or IRES elements (Extended Data Fig. 5e). Notably, we again identified the 12-nucleotide motif seen in the TE analysis but shifted by one nucleotide (Extended Data Fig. 5f and Supplementary Table 4c), along with 9-nucleotide variants, highly significantly enriched in the rDiff positive set compared to background (\( P = 2.2 \times 10^{-16} \)) (Extended Data Fig. 5g and Supplementary Table 4d).

**The (CGG)\textsubscript{4} motif corresponds to G-quadruplexes**

We noticed that in many instances the 5′ UTR motifs coincided with computationally predicted G-quadruplex structures\textsuperscript{29}. For example, 51% of the 12-nucleotide (CGG)\textsubscript{4} sequences and 43% of the most common 9-nucleotide motifs localized precisely to the G-quadruplex structures—the ADAM10 5′ UTR provides an example (Fig. 4a, b, Extended Data Fig. 6a and Supplementary Table 4e–k). G-quadruplex structures form by non-Watson–Crick interactions between paired guanine quartets connected by at least one linker nucleotide (A or C) in a parallel or anti-parallel orientation. Accordingly, G-quadruplex structures were significantly enriched among TE down genes and 79 of 220 TE down transcripts harboured at least one G-quadruplex (\( P = 2 \times 10^{-11} \)) (Fig. 4a–c and Supplementary Table 4e–k).

We confirmed by circular dichroism that the 12-nucleotide and extended 9-nucleotide sequence motifs form G-quadruplexes. Briefly, we compared the molar ellipticity of RNA oligomers encoding the 12-nucleotide and the 9-nucleotide motifs to a known G-quadruplex element in human telomeric RNA\textsuperscript{30} and a randomly organized oligomer with equal GC content and length. The 9-nucleotide sequences included either two (oligonucleotide 1) or five (oligonucleotide 2) flanking nucleotides as seen in the 5′ UTRs of MTA2, TGFB1, MAPKAP1 and
ADAM10 (Supplementary Table 5). The human telomeric sequence, the 12-nucleotide motif, and the 9-nucleotide motifs showed typical positive and negative molar ellipticity peaks at 264 nm and 240 nm, indicating parallel G-quadruplex structures. By contrast, the random oligomer had a shift in peak wavelengths (270 and 233 nm) (Fig. 4d, e and Extended Data Fig. 6b). This effect was dependent on the presence of potassium and was abrogated in sodium phosphate buffer without potassium (Extended Data Fig. 6c). Circular dichroism combined with thermal unfolding revealed that under potassium conditions the melting temperature for the 12-nucleotide motif was higher (56 °C) than the random oligomer (49 °C), corresponding to a free energy difference of −32 kcal mol⁻¹ (Fig. 4f). Similarly, computational modelling of complete 5’ UTRs with increasing numbers of motifs showed a decrease of the predicted minimum free energy when the 5’ UTR is allowed to fold into G-quadruplexes (Fig. 4g). Hence, the 12-nucleotide and some of the 9-nucleotide motifs can form energetically favourable G-quadruplex structures (Fig. 4h).

Next, we tested the 12-nucleotide (CGG)₉ motif in a translation assay. We constructed a reporter system to compare four 12-nucleotide motifs in tandem (G-quadruplex construct) to a random sequence of equal length and GC content (control construct), using IRES firefly luciferase as an internal control (Fig. 4i). Treatment with silvestrol (25 nM) reduced the translation of the G-quadruplex construct and did not affect the control construct. Hippuristanol and patamine A had identical effects (Extended Data Fig. 6d), whereas cycloheximide (20 nM) suppressed both reporters (Fig. 4j, k). Other RNA helicases (DHX9, DHX36) may resolve G-quadruplex structures. However, we find predominant expression of eIF4A2 in T-ALL, and knockdown confirms an eIF4A-dependent effect on the G-quadruplex reporter (Fig. 4l and Extended Data Fig. 6e). Conversely, increased expression of eIF4A or a silvestrol-binding-site mutant of eIF4A (P159Q) renders the G-quadruplex reporter insensitive to silvestrol (Extended Data Fig. 6f). Hence, eIF4A is limiting for the translation of mRNAs harbouring G-quadruplexes in their 5’ UTRs.

Transcripts affected by silvestrol

The silvestrol-sensitive transcripts (TE down and rDiff gene lists) include many genes with known roles in T-ALL, indicating that expression of eIF4A is pivotal for their translation. Conversely, increased expression of eIF4A or a silvestrol-binding-site mutant of eIF4A (P159Q) renders the G-quadruplex reporter insensitive to silvestrol (Extended Data Fig. 6f). Hence, eIF4A is limiting for the translation of mRNAs harbouring G-quadruplexes in their 5’ UTRs.
The corresponding mRNAs were unchanged, and the increase in MYC kills cancer cells by disrupting the translation of several critical factors. (Fig. 5g and Extended Data Fig. 8d–g). This result indicates that silvestrol expression of IRES-MYC and IRES-BCL2 afforded significant protection in an eIF4A-independent manner from the HCV IRES. In mixed populations of KOPT-K1 cells treated with silvestrol (Silv. (25 nM)), G-quadruplex structures that correspond to the motif. RNA G-quadruplex structures are energetically favourable/stable RNA structures made from at least two stacks of four guanines with non-Watson–Crick interactions (Hoogsteen hydrogen bonds) with intervening linker nucleotides. The minimum number is 12 nucleotides—for example, the 12-nucleotide motif (CGG)4—but shorter 9-nucleotide motifs can include neighbouring nucleotides to complete the structure. Frequently, more than 12 nucleotides contribute to G-quadruplexes, and typically silvestrol-sensitive transcripts harbour several G-quadruplexes in their 5′ UTRs. These features define a specific subset of genes whose translation requires eIF4A and that is clearly distinct from mTORC1-dependent translation and transcription is controlled by superenhancers 34.

Many transcription factors and oncogenes require eIF4A for translation. L-graffit control of these genes. In this regard, we note that a recent study identified a role for eIF4A2 in microRNA function and reports that eIF4A-sensitive transcripts are probably controlled by microRNAs 39. We further find a significant overlap with genes whose transcription is controlled by superenhancers 34.

A selective mechanism of translational control

We report an eIF4A-dependent mechanism of translational control that is encoded in the 5′ UTR of susceptible transcripts (Fig. 5h). The eIF4A RNA helicase has been implicated in the translation of mRNAs with long and complex 5′ UTRs. Ribosome footprinting technology allowed us to identify silvestrol-sensitive mRNAs and analyse their hallmark features. Besides 5′ UTR length, these included a 12-nucleotide (CGG)4 motif, additional 9-nucleotide motifs, and computationally predicted G-quadruplex structures that correspond to the motif. RNA G-quadruplex structures are energetically favourable/stable RNA structures made from at least two stacks of four guanines with non-Watson–Crick interactions (Hoogsteen hydrogen bonds) with intervening linker nucleotides. The minimum number is 12 nucleotides—for example, the 12-nucleotide motif (CGG)4—but shorter 9-nucleotide motifs can include neighbouring nucleotides to complete the structure. Frequently, more than 12 nucleotides contribute to G-quadruplexes, and typically silvestrol-sensitive transcripts harbour several G-quadruplexes in their 5′ UTRs. These features define a specific subset of genes whose translation requires eIF4A and that is clearly distinct from mTORC1-dependent translation.

Many transcription factors and oncogenes require eIF4A for translation. L-graffit control of these genes. In this regard, we note that a recent study identified a role for eIF4A2 in microRNA function and reports that eIF4A-sensitive transcripts are probably controlled by microRNAs 39. We further find a significant overlap with genes whose transcription is controlled by superenhancers 34.

Our study provides new insight into the striking anticancer effect of silvestrol and related compounds. Several strategies are available to target oncoprotein translation (for example, mTORC inhibitors 15), the eIF4G inhibitory peptide 4EGI-1 (ref. 40), or ribavirin 39. We find that eIF4A has an oncogenic activity that is required for the expression of key oncogenes. Accordingly, silvestrol can obliterate oncoproteins including MYC, NOTCH, BCL2, and others. This finding suggests potential activity in other cancers as a single agent or in combination with other drugs.

METHODS SUMMARY

Ribosome footprinting. KOPT-K1 cells were treated with silvestrol (25 nM) or DMSO for 45 min, followed by cycloheximide treatment for 10 min. Total RNA and ribosome-protected fragments were isolated following a published protocol 14. Deep sequencing libraries were generated from these fragments and sequenced on the HiSeq 2000 platform. Genome annotation was from the GENCODE project (http://www.gencodegenes.org/releases/14.html). Ribosome footprint intensity (RPM) and the expression value (RPKM) were measured from total mRNA-sequencing data and translation values were measured from ribosome footprint data. To evaluate the TE change between silvestrol- and vehicle-treated samples, we then calculated TE = RPKMfootprint/RPKmmRNA (ref. 15). Changes in ribosome footprint profiles were determined using the DEXseq algorithm 16.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 December 2013; accepted 1 May 2014.

Published online 27 July 2014.

1. Wendel, H. G. et al. Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. Nature 428, 332–337 (2004).
2. Mamane, Y., Petroulakis, E., LeBacquer, O. & Sonenberg, N. mTOR, translation initiation and cancer. Oncogene 25, 6416–6422 (2006).
3. Rajasekhar, V. K. et al. Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. Mol. Cell 12, 889–901 (2003).
...
METHODS
Ribosome footprinting. KOPT-K1 cells were treated with silvestrol (25 nm) or DMSO for 45 min, followed by cycloheximide treatment for 10 min. Total RNA and ribosome-protected fragments were isolated following a published protocol.14 Deep sequencing libraries were generated from these fragments and sequenced on the HiSeq 2000 platform. Genome annotation was from the Gencode project (http://www.gencodegenes.org/releases/14.html). Ribosome footprint intensity (RPKM) and the expression value (RPKMmRNA) were measured from total mRNA-seq data and translation values were measured from ribosome footprint data. To evaluate the TE change between silvestrol- and vehicle-treated samples, we then calculated TE = RPKMfootprint/RPKMmRNA (ref. 15). Changes in ribosome footprint profiles were determined using the DEXseq algorithm.44 Metabolic labelling of nascent protein. KOPT-K1 cells were labelled for nascent protein synthesis using Click-iT AHA metabolic labelling reagent obtained from Invitrogen (catalogue no. C10102) as per the manufacturer’s instructions. Briefly, following silvestrol, cycloheximide or DMSO treatment, cells were incubated in methionine-free medium for 30 min before AHA labelling for 1 h. Cells were fixed with 4% paraformaldehyde in PBS for 15 min, permeabilized with 0.25% Triton X-100 in PBS for 15 min followed by one wash with 3% BSA. Cells were then stained using Alexa Fluor 488 Alkyne (Invitrogen catalogue no. A10267) with Click-iT Cell reaction Buffer Kit (Invitrogen catalogue no. C10269). Changes in mean fluorescence intensity as a measure of newly synthesized protein were detected by flow cytometric analysis.

Polysome profiling. KOPT-K1 cells were treated with silvestrol or DMSO for 45 min, followed by cycloheximide treatment for 10 min. Cell pellet was lysed in polysome lysis buffer (300 mM NaCl, 15 mM Tris-HCl (pH 7.5), 15 mM MgCl2, 1% Triton X-100, 0.1 mM mg l−1 cycloheximide, 1 mg l−1 heparin). Polysome fractions were isolated using 4 ml 10–50% sucrose density gradients (300 mM NaCl, 100 mM MgCl2, 15 mM Tris-HCl (pH 7.5), 1 mg ml−1 cycloheximide, 10 mg l−1 heparin). Gradients were centrifuged in an SW-40Ti rotor at 35,000 rpm for 2 h. Fractions of 100 pl were collected manually from the top, and optical density (OD) at 254 nm was measured.

Sequence alignment. The human genome sequence hg19 was downloaded from the UCSC public database (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/chromosomes). Ribosome footprint (RF) reads were aligned to reference genome hg19 using PALMapper.43 PALMapper clips the linker sequence (5′-CTCTAGAG CACCACCATAA-3′), which is technically introduced during RF library construction, and trims the remaining sequence from the 3′ end while aligning the reads to reference sequence. Briefly, we set the parameters for PALMapper as follows: maximum number of mismatches: 2; maximum number of gaps: 0; minimum alignment length: 15; maximum intron length (splice alignment): 10,000; minimum length of a splicing read aligned to either side of the intron boundary: 10. We only use the uniquely aligned reads for further analysis.

To remove ribosome RNA contamination, the footprint reads were also aligned to a ribosome database using PALMapper with the same parameters except allowing splice alignment. We retrieved the human ribosome sequences from BioMart Ensembl and SILVA databases and merged the results into a single FASTA file, which was used as reference sequence to align against. The RNA-aligned reads were filtered out from hg19-aligned reads.

After removing the RNA contamination, we still observed a portion of reads that were dominated by linker sequence and Illumina P7 adaptor. These reads can also be trimmed during mapping and cause false alignment. Therefore, we searched a string of 1 to ~8 nucleotides from the linker sequence around the trimming site (±2 bp) allowing a 1 nucleotide mismatch. We removed the read if there was no such linker sequence. Finally, we filtered out reads ≤24 bp and ≥36 bp, and the remaining reads with an aligned length from 25 to 35 bp were used to analyse the translational effects of silvestrol.

Total mRNA sequencing reads were aligned to the hg19 reference using STAR.45 We performed the splice alignment and only use the uniquely aligned reads with a maximum of three mismatches. rRNA containing reads were also filtered out using the same strategy described earlier. Out of six measurements we removed two outliers and the remaining two biological replicates showed excellent consistency (control, R2 = 0.90; silvestrol, R2 = 0.88; data not shown).

Footprint profile analysis. For each gene, we only counted the number of aligned reads that were mapped within exonic regions. Genome annotation was downloaded from the GENCODE project (http://www.gencodegenes.org/releases/14.html). Ribosome footprint intensity (RPKM) was calculated as RPKM = C/N(106), where C is the read count for gene i, and N is the library size of silvestrol- or vehicle-treated samples. In order to eliminate the influence of rRNA contamination, the rRNA library size was calculated after read filtering. Similarly, the expression value measured from the total mRNA-seq data and the translation value measured from ribosome footprint data (both in RPKM) were calculated as RPKM = C/(K × N(106)), where K is the non-overlapped exonic region of each gene. To evaluate the TE change between silvestrol- and vehicle-treated samples, we then calculated TE = RPKMfootprint/RPKMmRNA, as done recently.15 To detect the genes for which the ribosome footprint profiles were significantly changed between the silvestrol-treated sample and control, we used DEXseq44 to perform the statistical test. DEXseq accounts for the discrete nature of the read counts and it also models the biological variability, which has been demonstrated in other applications to be crucial to avoid a great number of false positives. Here, DEXseq was used in a specific way: we fitted the footprint and mRNA-seq read counts into the DEXseq framework, in which silvestrol and control are two biological conditions, and we tested whether the footprint (consisting of two replicates for each condition) and mRNA-seq (we split the three replicates and recombined them into two combinations such that each of them consists of two replicates) read counts were significantly different in the two conditions. The log-ratio of normalized read counts of the silvestrol-treated sample to control indicated whether the ribosome footprint profile was increased or decreased. In the end, we plotted the ratio of TEsilvestrol/TEcontrol of all the genes and colour-highlighted them according to the statistical significance of the DEXseq test.

In addition to studying the TE, we also evaluated the ribosomal distribution change between the silvestrol-treated sample and control. First, a BED file containing all non-overlapping exonic regions was generated based on genome annotation. Then the BED file and footprint BAM files were given as an input to SAMTOOLS46 to generate new BAM files only including exonic alignment. We input the exonic BAM files of two conditions to rDiff20 to identify genes that had a significant change in ribosomal distribution. In detail, we performed a nonparametric test implemented in rDiff to detect differential read densities. rDiff takes relevant read information, such as the mapping location and the read length, to measure the significance of changes in the read density within a given gene between two conditions. The minimal read length was set to 25 bp, and number of permutations was set to 10,000. To plot the ribosomal distribution curves for multiple genes, we normalized read coverage of each transcript by the mean coverage value of that particular transcript. Then the UTR and coding exon length were normalized in proportion to the overall average length of corresponding regions of a group of genes. Finally all the normalized transcripts were averaged together in a vectored way to plot the coverage distribution. The ribosomal distribution curves for a single gene were plotted in a similar way but without normalizing the read coverage, and the coverage was smoothed using ‘moving average’ smoothing algorithm.

Motif analysis. The transcripts of each gene were quantified based on the total mRNA-seq data using MISO.39 The 5′ UTR of the most abundant transcript was collected for predicting motifs. Both the significant genes with increased or decreased TE and altered ribosomal distribution and the corresponding background gene sets were predicted by DREME22. Over- and underrepresented motifs were determined with three different settings: searching for motifs with a length greater than or equal to 6, 9 and 12 bp. We considered the predicted consensus sequences with P < 1 × 10−4 as significant motifs. Motif occurrences were called using FIMO48 with default parameters for strand specific prediction. P values for the enrichment of motifs in G-quadruplexes (GQs) were calculated using a one-sided binomial test and account for UTR length. The secondary structure of different gene sets was predicted using RNAfold49 based on the same 5′ UTR region. For each motif, the UTR sequences for each respective group of targets were subjected to motif prediction using the program RegRNA (http://regrna.mbc.nctu.edu.tw/html/prediction.html) and we looked specifically for motifs that occur in 5′ UTRs. Statistical significance for the results obtained was calculated using Fisher’s exact test for count data.

 Luciferase assays. Four tandem repeats of the (CGG), 12-nucleotide motif (GQs) or random sequence matched for length and GC content (random) were cloned into the 5′ UTR of Renilla luciferase plasmid pGL4.73. Empty Firefly luciferase plasmid pGL4.13 or HCV-IREs Firefly were used as internal controls. Luciferase assays were performed using Dual-Luciferase Reporter Assay System (Promega E1960) following the manufacturer’s instructions.

QQ sequences were as follows (in bold): CGATGGTGAAGAAGCTTCTGGAC GGCAGCGCGGCGTCACCTCTACGCCGGCGCGAGCATGATACCGCGCGCG CGGGCTGTTAAACTAGCAGCGCGGGGGGTTAGATTAAGTATTAA Random sequence matched for GC content: CTACGAGGCGCTTCCGACTCGACAGAABTGGTCG CAGTCGACCGGGCTTCAAGGCTCAGCGGTACGCTCAGCGGTTTCTCG CGTCGACCGGTACCGGTTCCGCGGTGAC.

Circulardichroism and thermal denaturation analysis. RNA oligonucleotides (see Supplementary Table 5) were used at 10 μM concentration. The buffer used for QQ studies contained 40 mM KCl, 10 mM potassium phosphate, pH 7.0 (ref. 50). To plot the ribosomal distribution curves for multiple genes, we normalized read coverage of each transcript by the mean coverage value of that particular transcript.

Circular dichroism and thermal denaturation analysis. RNA oligonucleotides (see Supplementary Table 5) were used at 10 μM concentration. The buffer used for QQ studies contained 40 mM KCl, 10 mM potassium phosphate, pH 7.0 (ref. 50). To plot the ribosomal distribution curves for multiple genes, we normalized read coverage of each transcript by the mean coverage value of that particular transcript.

©2014 Macmillan Publishers Limited. All rights reserved
range of 15 °C to 95 °C, with a 1 °C per minute temperature gradient. At least five scans for each sample were collected and were corrected with respective buffer controls. Calculations for $T_{m}$ and AG were performed as described previously. Smoothing of the circular dichroism curves was performed with Prism software by cumulating five neighbouring data points.

**T-ALL samples.** Thirty-six bone marrow biopsies were historically collected from patients with T-ALL at multiple organizations (Universitair Ziekenhuis (UZ) Ghent; UZ Leuven; Hôpital Purpan; Centre Hospitalier Universitaire de Nancy-Brabois). The QuantiAmp DNA Mini kit was used to obtain genomic DNA (Qiagen 51304). The Medical Ethical Commission of Ghent University Hospital (B67020084745) approved this study.

**Mutation analysis.** NOTCH1 (exons 26, 27, and 34), FBXW7 (exons 7, 8, 9, 10 and 11), PTEN (exons 1 till 9), and IL7R (exon 6) were amplified and sequenced using primers as reported previously. FBXW7, PTEN, and IL7R amplification were performed using 20 ng of genomic DNA, 1× KapaTaq reaction buffer (Kapa-Biosystems), 1 U KapaTaq DNA polymerase, 0.2 mM dNTP, 2.5 μM MgCl₂, 0.2 mM forward and reverse primer in a 25 μl PCR reaction. For NOTCH1 amplification, we used the PCRx enhancer system (Invitrogen) for the PCR reaction. Reactions contained 20 ng of DNA, 2.5 U KapaTaq DNA polymerase, 1× PCRx Amplification Buffer, 2× PCRx Enhancer Solution, 0.2 mM dNTP, 1.5 mM MgSO₄ and 0.2 mM of each primer. The PCR steps were as follows: 95 °C for 10 min (96 °C for 15 s, 57 °C for 1 min, then 72 °C for 1 min) for 40 cycles, and then 72 °C for 10 min. Purified PCR products were analysed using the Applied Biosystems 3730XL DNA Analyser.

**Array complete genomic hybridization.** PTEN deletions were detected by array complete genomic hybridization (CGH) analysis using SurePrint G3 Human 4x180K CGH Microarrays (Agilent Technologies). First, random prime labelling of the T-ALL DNA sample and a control human reference DNA was performed with Cy3 and Cy5 dyes (Perkin Elmer), respectively. The subsequent hybridization protocol was performed according to the manufacturer’s instructions (Agilent Technologies). The data were analysed using arrayCGHbase.

**Generation of mice.** The ICN-driven mouse T-ALL model has been previously reported. Mice were C57BL/6J females between 6 and 10 weeks of age. Treatment conditions and HPC genotypes were randomized and non-blinded; n was determined empirically based on previous mouse experiments. Data were analysed in Kaplan–Meier format using the log-rank (Mantel–Cox) test for statistical significance. The surface marker analysis was as described previously. shRNAs against PTEN and elf4A1 (ref. S6) have been reported. For xenograft studies, 5,000,000 KOPT-K1 cells in 30% matrigel (BD 354234) were injected subcutaneously into C.B-17 SCID mice. When tumours were readily visible, the mice were injected on 7 consecutive days with 0.5 mg kg⁻¹ silvestrol, 0.2 mg kg⁻¹ (-) CR-31-B, or vehicle control. Tumour size was measured daily by calliper. $p$ values were calculated using two-way repeated measures ANOVA. All animal experiments were performed in accordance with regulations from Memorial Sloan-Kettering Cancer Center’s Institutional Animal Care and Use Committee.

**T-ALL cell lines.** T-ALL cell lines were obtained from ATCC and all tested negative for mycoplasma contamination. Lines were cultured in RPMI-1640 (Invitrogen), 20% fetal calf serum, 1% penicillin/streptomycin, and 1% l-glutamine. The MOHTO line was supplemented with 5 ng ml⁻¹ IL2 (Fitzgerald 30R-A01202) and 10 ng ml⁻¹ of IL7 (Fitzgerald 30R-A0180X). To generate IgC471 cells, T-ALL cell lines and samples were treated with silvestrol or (-)-CR-31-B for 48 h. Viable cells were measured using ATP quantification via the CellTiter-Glo Luminescent Cell Viability Assay (Promega G7571).

**Immunoblots.** Lysates were made using Laemli lysis buffer. Thirty micrograms of protein was loaded onto SDS–PAGE gels then transferred onto Immobilon-FL Transfer Membranes (Millipore, IPVH00100). The antibodies used were α-tubulin (Sigma, TS168), β-actin (Sigma, A3516), MYC (Santa Cruz Biotechnology, sc-40), p-AKT 308 (Cell Signaling, 9275), AKT (Cell Signaling, 9272), S6 (Cell Signaling, 2317), p-S6 (Cell Signaling, 2215), NOTCH1 (Cell Signaling, 3608), MYB (Santa Cruz Biotechnology, sc-517), CDK6 (Cell Signaling, 3136), EZH2 (Cell Signaling, 5246), MDM2 (Santa Cruz Biotechnology, sc-965), CCND3 (Cell Signaling, 2936), BCL2 (Santa Cruz Biotechnology, sc-509), RUNX1 (Cell Signaling, 4336), and GAPDH (Cell Signaling, 5174).

**Statistical analysis.** All Kaplan–Meier curves were analysed using the Mantel–Cox test. The significance of xenografted tumour size differences was determined using two-way repeated measures ANOVA tests. The significance of motif enrichments was determined using a one-tailed binomial test with correction for differences in 5′ UTR lengths. The $P$ value for this test is defined as

$$P = \sum_{i=0}^{T_{n2u}} B\left(\frac{T_{n1u}}{T_{n2u}}, \frac{L_{n1u}}{L_{n2u}}, \frac{H_{n1u}}{H_{n2u}}\right)$$

where $T_{n1u}$ and $T_{n2u}$ are the number of genes in the positive and negative (background) set respectively, $L_{n1u}$ and $L_{n2u}$ are the average length of the 5′ UTR of the positive and negative set of genes respectively, $H_{n1u}$ and $H_{n2u}$ are the number of genes with the motif under consideration, and $B(n,p)(i)$ is the probability of observing $i$ events drawn from the distribution $B(n,p)$.

**Compounds.** Silvestrol was a gift from J.P.’s laboratory and was subsequently purchased from ChemScore (CS-0543). (-)-CR-31-B was a gift from J.A.P. Jr’s laboratory. Each was suspended in DMSO for in vitro experiments and in 5.2% Tween 80 and 5.2% PEG 400 for in vivo experiments. Cycloheximide (C7698) and rapamycin (R8781) were purchased from Sigma.

**Toxicity studies.** Eight-week-old C57BL/6NTac female mice were randomly assigned to either control or treatment groups. Each treatment group received one daily dose of test article through i.p. injections over 5 consecutive days. Toxicity was monitored by weight loss and daily clinical observation for the 14 days following test article administration. Twenty-four hours after the last test article administration, four mice in each group were killed and clinical chemistry, haematology and tissue specific histopathology were done at autopsy. The remaining mice (n = 2 per group) were kept under observation for an additional 13 days; at that point all mice were killed and clinical chemistry, haematology and tissue-specific histopathology were done at time of autopsy.

42. Jean, G., Kahles, A., Sreedharan, V. T., De Bona, F. & Ratsch, G. RNA-Seq read alignments with PALMapper. Curr. Protoc. Bioinformatics 32, 11.6.1–11.6.37 (2010).
43. Flicek, P. et al. Ensembl 2013. Nucleic Acids Res. 41, D48–D55 (2013).
44. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
45. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
46. Katz, Y., Wang, E. T., Airoldi, E. M. & Burge, C. B. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nature Methods 7, 1009–1015 (2010).
47. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018 (2011).
48. Hortacu, I. L. Vienna RNA secondary structure server. Nucleic Acids Res. 31, 4329–4331 (2003).
49. Fratta, P. et al. CSR0r72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci. Rep. 2, 1016 (2012).
50. Greenfield, N. J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nature Protocols 1, 2527–2535 (2007).
51. Mavrakis, K. J. et al. A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). Nature Genet. 43, 673–678 (2011).
52. Shocat, C. et al. Gain-of-function mutations in interieur7 receptor-alpha (IL7R) in childhood acute lymphoblastic leukemias. J. Exp. Med. 208, 901–908 (2011).
53. Zuurier, L. et al. The significance of PTEN and AKT alterations in pediatric T-cell acute lymphoblastic leukemia. Haematologica 97, 1405–1413 (2012).
54. Menten, B. et al. arrayCGHbase: an analysis platform for comparative genomic hybridization microarrays. BMC Bioinformatics 6, 124 (2005).
55. Galicia-Vázquez, G., Cencic, R., Robert, F., Agenor, A. Q. & Pelletier, J. A. cellular response linking eIF4AI activity to eIF4AII transcription. RNA 18, 1373–1384 (2012).
Extended Data Figure 1 | Translational activation in T-ALL. a–c, Diagram of mutations in human T-ALL affecting PTEN (a), IL7R (b) and NOTCH1 (c). d, Immunoblots of lysates from ICN-driven murine leukaemia with the additional indicated construct, probed as indicated. e, Immunoblots of lysates from 3T3 cells with empty vector or shIF4A and probed as indicated. f, Representative fluorescence-activated cell sorting (FACS) profiles measuring levels of the indicated markers in murine leukaemia. g, Surface marker expression on murine leukaemic cells of the indicated genotype (plus and minus signs indicate <50% or ≥50% positive cells). h, Lysates of murine leukaemia expressing ICN and either empty vector or elf4A1 and probed as indicated. i, Representative histology detailing the pathological appearance of murine T-ALLs harbouring the indicated genes and stained as indicated.
Extended Data Figure 2 | Silvestrol and the synthetic analogue (±)-CR-31-B are effective against T-ALL. a, b, Chemical structure of silvestrol (a) and (±)-CR-31-B (b). c, Dual luciferase reporter assay, shown are relative levels of each firefly (cap-dependent) and Renilla (IRES-dependent) luciferase upon treatment with silvestrol or (±)-CR-31-B. Mean and s.d. are shown, n = 3 biological replicates. d, IC_{50} values for silvestrol and (±)-CR-31-B in a panel of human T-ALL primary patient samples and cell lines. Mean and s.d. are shown, n = 4 biological replicates. e, Silvestrol’s effect on murine T-ALLs with the indicated genetic lesions; curves are mean of triplicates and differences between the genotypes did not reach significance. f, Kaplan–Meier analysis showing time to leukaemia development after systemic transplantation of MOHTO cells in Balb/c mice followed by treatment on 7 consecutive days (treatments are indicated by red arrows) with either silvestrol (0.5 mg kg⁻¹, red line, n = 5) or vehicle (black line, n = 5). g, KOPT-K1 xenograft studies. Shown is the tumour volume during and after systemic treatment with (±)-CR-31-B or vehicle (i.p. injection, 0.2 mg kg⁻¹ on days indicated by red arrows). Mean and s.d. are shown, n = 6 biological replicates. h, Tumour volume upon i.p. treatment with vehicle or silvestrol (0.5 mg kg⁻¹ on days indicated by red arrows). Mean and s.d. are shown, n = 3 biological replicates.
Extended Data Figure 3 | Toxicity studies with (±)-CR-31-B. a–j, Mean and standard deviation are shown, n = 2 biological replicates. a, Animal weights during and after (±)-CR-31-B treatment (i.p. injection, 0.2 mg kg⁻¹ on days indicated by red arrows). Red, CR; black, vehicle. b–d, Counts of white blood cells (b), red cells (c) and platelets (d) 14 days after cessation of (±)-CR-31-B treatment. Blue lines indicate the species- and strain-specific reference range. NS, not significant, n = 2 biological replicates. e, Representative histology of gastrointestinal tract (small intestine) on the indicated days during (n = 4) and after (n = 2) (±)-CR-31-B treatment. f–j, Serum levels of alanine aminotransferase (ALT) (f), aspartate transaminase (AST) (g), albumin (h), total bilirubin (i) and creatinine (j) 2 weeks after cessation of treatment with (±)-CR-31-B or vehicle. Blue lines indicate the species- and strain-specific reference range. NS, not significant.
Extended Data Figure 4 | Ribosome profiling quality control data and effects on translation.

a, b, Read counts by length of mapped sequence before and after filtering rRNA, linker reads, non-coding RNAs, short mapped sequences (‘noisy’ reads; see text and Methods for details). n = 2 biological replicates. c, d, Read length frequency histograms and mapping analysis of ribosome footprint data after quality control filtering for vehicle-treated cells (c) or silvestrol-treated cells (d). n = 2 biological replicates. e, Silvestrol-induced changes in total RNA (log₂ fold change RPKM) and ribosome-protected RNA (RF). n = 2 biological replicates. f, Histogram of the ribosome footprint intensity of all genes (measured as unique RPM) for silvestrol- and vehicle-treated cells, indicating that silvestrol-affected mRNAs were broadly distributed (see text for details). n = 2 biological replicates.

g, Mean fluorescence intensity of incorporated AHA in newly synthesized proteins in KOPT-K1 cells treated with vehicle (DMSO), silvestrol (Silv.; 25 nM), or cycloheximide (CHX; 100 nM) for the indicated time period. n = 3 biological replicates.

h, Polyribosome profiles of silvestrol- (25 nM) or vehicle (DMSO)-treated KOPT-K1 cells showing OD₂₅₄ nm absorption across the ribosome-containing fractions. n = 3 biological replicates. i, Ribosome density for transcripts across control and silvestrol samples (ribosomal footprint (RF) reads per RPKM). n = 2 biological replicates. The correlation (R² = 0.94) indicates a broad effect on translation and transcripts with significantly differential changes in ribosome density are indicated as red and blue dots. j, Length comparison of 5’ UTRs of TE up genes and a background gene set. Asterisks indicate mean, n = 2 biological replicates. k, Percentage of TE up genes and background genes containing the indicated sequence motifs. *P < 0.001, n = 2 biological replicates.
Extended Data Figure 5 | Analysis of genes with differential ribosomal distribution rDiff. a, Representation of ribosome coverage for 826 transcripts with significant changes in distribution between silvestrol (red) and vehicle (black)—corresponding to the rDiff positive gene list after filtering out genes with 5'UTR length >20 nucleotides. Both ribosome footprint coverage and transcript length are normalized for comparison. Translation start and stop sites are indicated by blue lines, n = 826. b, c, Ribosomal distribution plots, as in a, showing how silvestrol affects ribosome distribution in all TE up genes (b) (n = 182 after filtering out genes with 5'UTR length <20 nucleotides) and all TE down genes (c) (n = 276 after filtering out genes with 5'UTR length <20 nucleotides). d, Length comparison of 5'UTRs of genes with significantly altered ribosomal distribution (rDiff positive, red) and background genes (black). Asterisks indicate mean value, n = 826. e, Percentage of rDiff-positive genes and background genes containing the indicated sequence motifs. *P < 0.05, n = 2 biological replicates. f, g, The rDiff-positive genes are enriched for the indicated 12-nucleotide (f) and 9-nucleotide (g) consensus motifs.
Extended Data Figure 6 | Circular dichroism and characterization of eIF4A.
a, Bar graph indicating the prevalence of each sequence motif from the rDiff data set and its predicted likelihood of forming G-quadruplex (GQ) structures (red). b, Circular dichroism spectra scan of 9-nucleotide (mer) motif with a 5-nucleotide flank taken from the actual 5' UTR of the indicated genes, folded in KCl. c, Circular dichroism spectra scan of 12-nucleotide motif and mutant folded in sodium phosphate buffer without KCl; note the y-axis scale. d, Relative amounts of Renilla luciferase (normalized to firefly), expressed from the G-quadruplexes (red bars) or control construct (black bars), treated with 8 nM pateamine A (Pat. A) or 50 nM hippuristanol (Hipp.) for 24 h. *P < 0.05, n = 3 biological replicates and n = 2 technical replicates. e, Analysis of mRNA expression of the indicated RNA helicases in normal T cells and T-ALL cells. *P < 0.05, n = 57 biological replicates. f, Relative amounts of Renilla luciferase expressed from the G-quadruplex construct in 3T3 cells and normalized to IRES/firefly with either empty vector or the indicated genes, treated with silvestrol (25 nM) for 24 h. Mean and s.d. are shown, n = 3 biological replicates, n = 2 technical replicates. NS, not significant.
Extended Data Figure 7 | Silvestrol-sensitive transcripts. 

**a**, Distribution of ribosomal footprints for the indicated genes, n = 2 biological replicates. Red, silvestrol; black, vehicle; purple dots, 9-nucleotide (mer) motifs; blue dots, 12-nucleotide motif. 

**b**, Gene ontology classification for genes in the TE down group with G-quadruplex (GQ), 12-nucleotide and 9-nucleotide motifs. 

**c**, Venn diagram illustrating the overlap between TE and/or rDiff genes and reported superenhancers in T-ALL cell lines.14.
Extended Data Figure 8 | Immunoblots and mRNA expression. a, Lysates from human T-ALL lines treated with (±)-CR-31-B (25 nM, 24 h) and probed as indicated. b, Lysates from Jurkat cells treated with escalating doses of silvestrol and probed as indicated. c, mRNA levels for the indicated genes treated with vehicle (DMSO, black) or silvestrol (red, 25 nM) for 45 min. Mean ± s.d. are shown, n = 2 biological replicates. d–g, Immunoblots of lysates from murine T-ALL cells expressing either vector control or IRES-MYC (d), IRES-CCND3 T283A (e), IRES-ICN (f) or IRES-BCL2 (g) and probed as indicated.