Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supporting Information

Mapping platinum adducts on yeast ribosomal RNA using high-throughput sequencing

K. Plakos[a, b] and V. J. DeRose[a]

[a] Department of Chemistry and Biochemistry and Institute of Molecular Biology, University of Oregon, Eugene, OR 97403

[b] Current address: Voxtel Corp., Eugene, OR

Email: derose@uoregon.edu

Table of Contents:

Materials and Methods

The following synthetic DNA oligonucleotides were purchased from IDT:

Mod-seq RT primer:

AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC/iSp18/CACTCA/iSp18/TT
CAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG

Mod-seq PCR FWD primer:

AATGATACGGCGACCACCGAGATCTACAC.

Mod-seq Indexed PCR Rev Primer:

CAAGCAGAAGACGGCATACGAGATxxxxxxGTGACTGGAGTTCAGACGTGTGCTCTTCCG

(where xxxxxx is a 6 nt Illumina Sequencing index.)

Yeast preparation

BY4741 yeast cells were grown from culture overnight at 30 °C in SD broth supplemented with 1x amino acids. Immediately before use, cisplatin stock was prepared by dissolving cisplatin (Sigma-Aldrich) to final concentration of 5 mM in SD broth. After overnight growth, yeast was diluted to OD 0.2 in SD broth supplemented with 1x amino acids and cisplatin was added to a final concentration of either 100 or 200 µM, in a final volume of 10 mL. Simultaneously, mock treatments were prepared by adding SD at the same volume as 100 uM cisplatin treatment. Each biological sample was prepared in duplicate. Samples were treated for six hours at 37 °C while shaking, then spun at 4 °C, 4K RPM for 10 minutes to pellet (Osborn et al. 2014).

RNA Extraction and fragmentation

After pelleting and decanting supernatant, remaining media was aspirated and cells were subject to lysis and RNA extraction using Qiagen's RNEasy kit, following their mechanical lysis protocol. Extracted RNA was quantified using a nanodrop, genomic DNA was removed using TurboDNAse (Thermo Fischer), and RNA was cleaned up via phenol-chloroform extraction followed by ethanol precipitation in the presence of glycogen coprecipitant.

Purified RNA was fragmented using RNA Fragmentation Reagents (Thermo Fisher) following the manufacturer's protocol for 100 nt fragments. After fragmentation, samples were ethanol precipitated and subjected to fragment analysis (UO GC3F Advanced Analytical Fragment Analyzer) to verify size distribution around 100 nt.

Sequencing Preparation

Samples were prepared for sequencing as described by Talkish et al. (2014) with some modifications. 3 µg of fragmented RNA was treated with T4 PNK (Thermo) in the absence of ATP to clean up 3' ends. Samples were phenol-chloroform extracted and ethanol precipitated in the presence of glycogen, then re-suspended in 4 µL nuclease-free water and 1 µL of 0.5 µg/µL miRNA cloning linker (NEB). Re-suspended samples were heated at 80 °C for 2 minutes then placed on ice. 5 µL of T4 RNA ligation mixture (1 µL T4 RNA ligase buffer, 2 µL Peg8000, 1 µL SUPERase In (Thermo Fisher), 1 µL T4 RNA Ligase, truncated KQ (NEB)) was added to each for a total of 10 µL. Ligation was allowed to proceed overnight at 16 °C.

Following ligation, samples were phenol-chloroform extracted followed by ethanol precipitation. Ethanol precipitated samples were re-dissolved in 11 µL nuclease free water and 1 µL 2.5 µM Mod-seq RT primer (IDT DNA). Samples were heated to 80 °C for 2 minutes and cooled on ice to anneal the RT primer. During this step, RT reaction mixture was prepared containing 4 µL 5x First Strand buffer (Thermo Fisher), 1 µL 10 mM DNTP's (AMRESCO), 1 µL 100 mM dTT (Thermo Fisher), and 1 µL SUPERase In (Thermo Fisher). After cooling, 7 µL RT reaction mixture was added to each sample. In prior experiments we noted that different reverse trancriptase enzymes result in differing levels of truncation at Pt adducts (not shown). For this work, we have chosen Superscript II because it is known to be stopped by nucleotide modifications, including DMS and SHAPE reagents (Talkish et al. 2014, Siegfried et al., 2014). Our own observations with targeted primer extension (see below) have shown that this enzyme also stops readily at cisplatin adducts, although we have not fully characterized its bypass characteristics. Reverse transcription was initiated by adding 1 µL SuperScript II reverse transcriptase (Thermo Fisher) then heating to 42 °C for 50 minutes. Reaction was stopped by heating to 70 °C for 15 minutes, then placed on ice. RNA template was destroyed by adding 2 µL 1 M NaOH and heating to 98 °C for 20 minutes.

Samples were ethanol precipitated to remove NaOH. RT products were run out on an 8% 19:1 mono:bis acrylamide gel containing 8M urea, 1xTBE. 15 µL nuclease free water and 5 µL loading buffer was added to each sample, then samples were heated to 80 °C for 3 minutes to re-dissolve, then immediately placed on ice. Gel lanes were flanked with Low MW DNA Ladder (NEB) to use as size markers for excising bands.

The acrylamide gel was run at 170V for approximately 2 hours, until the bromophenol blue band was approximately 1 cm from bottom. Bands were stained with SYBR Gold (Thermo Fisher) and visualized using blue light transillumination. Bands between 150 and 300 nt were excised from each lane and pulverized, 1 mL of 300 mM sodium acetate was added to each pulverized gel, and samples were rotated for 36 hours at 4 °C to allow RT products to diffuse into solution.

Pulverized gel pieces were filtered from supernatant using an empty spin column and samples were placed in a speed vacuum until volume was below 300 µL then ethanol precipitated in the presence of glycogen.

After ethanol precipitation, samples were treated with T4 PNK (Thermo Fisher) to phosphorylate their 5' ends, in preparation for circularization. After phosphorylation, samples were phenol chloroform extracted and ethanol precipitated.

To generate templates suitable for PCR amplification, samples were circularized using Circligase (Epicentre). Samples were resuspended in 15 µL nuclease free water, then 7.5 µL was added to 2 µL of Circligase reaction mixture (1 µL 10x Circligase buffer, 0.5 µL 1 mM ATP, 0.5 µL 50 mM $MnCl_2$) in a new PCR tube. 0.5 µL of Circligase was added, and samples were incubated at 60 °C for 2 hours, the reaction stopped by heating to 80 °C for 10 minutes, and samples were frozen until ready for use.

1 µL of circularized samples were amplified using Phusion Polymerase (Thermo Fisher) for 16 cycles using Mod-seq PCR FWD primer and one of ten balanced Mod-seq Indexed PCR Rev Primers. Reverse primers contain a unique 6 nt indexing region so each sample can be uniquely identified from the rest after pooling for sequencing. Together, these two primers build in necessary Illumina adapters, sequencing tags, and indexing tags. After PCR, samples were cleaned up using MagBind DNA extraction

beads following the manufacturer's instructions (Omega Bio-tek), nanodropped, and analyzed using a fragment analyzer (Advanced Analytical).

Illumina Sequencing

Samples were pooled to equivalent concentrations and submitted to the University of Oregon's Genomics and Cell Characterization Core Facility (GC3F) for Illumina Sequencing. Samples were sequenced on a NextSeq 500, 75 nt single-end read. Approximately 58 million reads were returned across 10 samples, ranging from 3.8-7.3 million reads per sample with a median of 5.7 million reads per sample.

Data Processing

Sequencing adapters were trimmed using CutAdapt (Martin, 2011), and trimmed files were aligned to the *Saccharomyces cerevisiae* genome using Bowtie (bowtie -v 0 -a --best --strata --chunkmbs 500 -p 16 -t –S) (v. 1.1.2, Langmead et al., 2009), constrained to allow either 0 or 1 mismatch in alignment. Aligned reads were analyzed for differential expression using the DESEQ2 R package to ensure correct matching of control treatments. Alignment files were fed into the Mod-seq pipeline (Talkish et al. 2014), bypassing their trim and align steps, to generate predicted platination locations based on the frequency of RT termination 5' to a given stop site. Stops were considered if they met a minimum threshold of 1.5-fold enrichment over background. The resulting data was parsed in R and plotted using the ggplot2 R package, and three-dimensional images were created in PyMol by noting our cisplatin stop sites on PDB 4V7R. All R scripts will be available online.

Reverse transcription primer extension analysis

Reverse transcription primers (CTCTTCCAGCCATAAGACCCC, which annealed near helix 80 and CACCGGAGCCAGCAAAGG, which annealed near C1711) were 5'end labeled with ^{32}P using PNK (Thermo Fisher). Yeast was grown as described and treated with either mock or 200 μM cisplatin for six hours, then RNA was extracted as described above. After extraction, 3 μg of purified RNA (+ or – cisplatin) was reverse transcribed using Superscript II and one of the two radiolabeled primers following manufacturer's protocol (20 μL). Simultaneously, dideoxy sequencing ladders were generated by reverse transcribing 3 μg of control RNA in the presence of 1 μL of 1 μM ddA, ddT, ddC, or ddG (Affymetrix). Post-RT, samples were cleaned up as described above, then run on a 20 x 40 cm sequencing gel (8 M urea, 1x TBE, 8% 19:1 mono:bis) at 30W for 2 hours. Gels were dried and exposed on a phosphorimager. Gel images were imported into ImageJ and contrast adjusted, then plots of line intensity were taken down each lane. Stop site peaks were picked out of control vs background and used to generate a line plot for overlay with Pt-seq data.

References:

B. Langmead, C. Trapnell, M. Pop, and S.L. Salzberg, *Genome Biol.*, 2009, **10**, 25.1-25.10.

J. Talkish, G. May, G., Y. Lin, J. L. Woolford Jr., and C. J. McManus, *RNA*, 2014, **20**, 713–720.

N. A. Siegfried, S. Busan, G. M. Rice, J. A. E. Nelson, K. M. Weeks, *Nat. Methods,* 2014, ***11***, 959–965.
M. Martin, *EMBnet J.*, 2011, **17**, 10–12.
M. F. Osborn, J. D. White, M. M. Haley, and V. J. DeRose, *ACS Chem. Biol.*, 2014, **9**, 2404–2411.

Supplementary Table 1. Counts of stops observed on all *S. cerevisiae* RNA at 100 μM cisplatin

100 μM Cisplatin

GeneName	Stops	GeneName	Stops	GeneName	Stops	GeneName	Stops
15S_rRNA	23	tD(GUC)J2	1	tG(GCC)D1	1	YDR261C-D	1
21S_rRNA	16	tD(GUC)J3	1	tG(GCC)D2	1	YDR316W-B	1
RDN18-1	74	tD(GUC)J4	1	tG(GCC)E	1	YDR365W-B	1
RDN18-2	74	tD(GUC)K	1	tG(GCC)F1	1	YER138C	1
RDN25-1	141	tD(GUC)L1	1	tG(GCC)F2	1	YER160C	1
RDN25-2	141	tD(GUC)L2	1	tG(GCC)G1	1	YGR038C-B	1
RDN37-1	214	tD(GUC)M	1	tG(GCC)G2	1	YGR161C-D	1
RDN37-2	214	tD(GUC)N	1	tG(GCC)J1	1	YHR214C-B	1
RDN5-1	3	tD(GUC)O	1	tG(GCC)J2	1	YJR027W	1
RDN5-2	2	tE(CUC)D	1	tG(GCC)M	1	YJR029W	1
RDN5-3	2	tE(CUC)I	1	tG(GCC)O1	1	YLR035C-A	1
RDN5-4	2	tE(UUC)B	1	tG(GCC)O2	1	YLR154C-G	4
RDN5-5	2	tE(UUC)C	1	tG(GCC)P1	1	YLR157C-B	1
RDN5-6	3	tE(UUC)E1	1	tG(GCC)P2	1	YML039W	1
RDN58-1	1	tE(UUC)E2	1	tL(UAA)B1	1	YML045W	1
RDN58-2	1	tE(UUC)E3	1	tL(UAA)B2	1	YMR045C	1
SCR1	13	tE(UUC)G1	1	tL(UAA)D	1	YMR050C	1
snR17a	1	tE(UUC)G2	1	tL(UAA)J	1	YNL054W-B	1
snR56	1	tE(UUC)G3	1	tL(UAA)K	1	YNL284C-B	1
snR9	1	tE(UUC)I	1	tL(UAA)L	1	YOL103W-B	1
tD(GUC)B	1	tE(UUC)J	1	tL(UAA)N	1	YOR142W-B	1
tD(GUC)D	1	tE(UUC)K	1	YAR009C	1	YPL257W-B	1
tD(GUC)G1	1	tE(UUC)L	1	YBL005W-B	1	YPR080W	1
tD(GUC)G2	1	tE(UUC)M	2	YBR012W-B	1	YPR137C-B	1
tD(GUC)I1	1	tE(UUC)P	1	YBR118W	1	YPR158W-B	1
tD(GUC)I2	1	tG(GCC)B	1	YDR098C-B	1		
tD(GUC)J1	1	tG(GCC)C	1	YDR210C-D	1		

Supplementary Table 2. Counts of stops observed on all *S. cerevisiae* RNA at 200 μM cisplatin

200 μM Cisplatin

GeneName	Stops
15S_rRNA	16
21S_rRNA	24
LSR1	9
RDN18-1	171
RDN18-2	171
RDN25-1	304
RDN25-2	304
RDN37-1	522
RDN37-2	522
RDN5-1	4
RDN5-2	2
RDN5-3	2
RDN5-4	2
RDN5-5	2
RDN5-6	4
RDN58-1	3
RDN58-2	3
SCR1	16
snR10	1
snR128	5
snR13	1
snR19	2
snR3	2
snR30	3
snR32	1
snR49	1
snR7-L	3
snR7-S	3
snR9	3
tA(UGC)A	1
tA(UGC)E	1
tA(UGC)G	1

GeneName	Stops
tA(UGC)L	1
tA(UGC)O	1
tE(CUC)D	1
tE(CUC)I	1
tE(UUC)B	1
tE(UUC)C	1
tE(UUC)E1	1
tE(UUC)E2	1
tE(UUC)E3	1
tE(UUC)G1	1
tE(UUC)G2	1
tE(UUC)G3	1
tE(UUC)I	1
tE(UUC)J	1
tE(UUC)K	1
tE(UUC)L	1
tE(UUC)M	1
tE(UUC)P	1
tG(GCC)B	1
tG(GCC)C	1
tG(GCC)D1	1
tG(GCC)D2	1
tG(GCC)E	1
tG(GCC)F1	1
tG(GCC)F2	1
tG(GCC)G1	1
tG(GCC)G2	1
tG(GCC)J1	1
tG(GCC)J2	1
tG(GCC)M	1
tG(GCC)O1	1
tG(GCC)O2	1

GeneName	Stops
tG(GCC)P1	1
tG(GCC)P2	1
YAR010C	2
YBR012W-A	2
YBR012W-B	2
YDR098C-A	2
YDR098C-B	2
YDR210C-C	2
YDR210C-D	2
YDR261C-C	2
YDR261C-D	2
YDR278C	4
YDR316W-A	2
YDR316W-B	2
YDR365W-A	2
YDR365W-B	2
YER137C-A	2
YER138C	2
YER159C-A	2
YER160C	2
YGL008C	1
YGR027W-A	2
YGR027W-B	2
YGR038C-A	2
YGR038C-B	2
YGR161C-C	2
YGR161C-D	2
YGR192C	1
YHR214C-B	2
YHR214C-C	2
YJR026W	2
YJR027W	2

GeneName	Stops
YJR028W	2
YJR029W	2
YLR154C-G	7
YLR157C-A	2
YLR157C-B	2
YLR227W-A	2
YLR227W-B	2
YLR256W-A	2
YML039W	2
YML040W	2
YML045W	2
YML045W-A	2
YMR050C	2
YMR051C	2
YNL054W-A	1
YNL054W-B	1
YOL103W-A	2
YOL103W-B	2
YOR142W-A	2
YOR142W-B	2
YPL257W-A	2
YPL257W-B	2
YPR137C-A	2
YPR137C-B	2
YPR158C-C	2
YPR158C-D	2
YPR158W-A	2
YPR158W-B	2