PLATINUM-SEQ: HIGH-THROUGHPUT MAPPING OF SMALL-MOLECULE PLATINUM ADDUCTS ON CELLULAR RNA

by

KORY JOHN IVERSON PLAKOS

A DISSERTATION

Presented to the Department of Chemistry and Biochemistry
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2016

DISSERTATION APPROVAL PAGE

Student: Kory John Iverson Plakos

Title: Platinum-Seq: High-Throughput Mapping of Small-Molecule Platinum Adducts on Cellular RNA

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Chemistry and Biochemistry by:

Diane Hawley	Chairperson
Victoria J. DeRose	Advisor
Michael M. Haley	Core Member
Alice Barkan	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2016

DISSERTATION ABSTRACT

Kory John Iverson Plakos

Doctor of Philosophy

Department of Chemistry and Biochemistry

December 2016

Title: Platinum-Seq: High-Throughput Mapping of Small-Molecule Platinum Adducts on Cellular RNA

Methods to map small-molecule interactions with cellular RNAs are important for understanding endogenous activation, such as in riboswitches, as well as the potential for exogenous compounds to target RNA. Cisplatin is one of the most widely used of the platinum anticancer drugs that are prescribed in approximately 40-50% of all chemotherapy treatments (Dyson and Sava, 2006; Harper et al., 2010). Despite nearly 40 years of experience with this class of drugs, we still lack a comprehensive understanding of the targets of Pt compounds and their effects on cells. Pt(II) compounds are well-known DNA and RNA crosslinking agents, but the latter area is under-studied. In order to better understand the impacts of cisplatin and other platinum(II)-derived small molecules on cellular RNA, we have developed a technique we call "Platinum-seq," which couples reverse transcription mapping of platinated RNAs to high-throughput sequencing. Chapter 1 is a study of cisplatin and a novel click-functionalized platinum compound (2-ADAP Pt) binding to the HDV ribozyme, a small catalytic RNA. Chapter 2 moves our platinum mapping approaches from low-throughput, sequencing gel based methods into next-generation sequencing for high-throughput analysis of all platinum sites in cellular RNA, a method we have named "Platinum-seq." Chapter 3 is a study of differential gene

expression of *Saccharomyces cerevisiae* treated with cisplatin and a second novel platinum(II) compound (azaplatin), using data acquired from the work in Chapter 2. Chapter 4 describes recent efforts to implement pre-enrichment of sequencing targets using click chemistry followed by DNA hybridization, in order to enrich for platinated fragments before sequencing library construction. Together, this work represents a significant step forward in advancing analysis of Pt(II) binding to cellular RNA, a potentially important target for this widely used class of anticancer compounds. Methods developed here are broadly applicable to genome-wide identification of platinum accumulation on DNA as well, which has not been pursued despite the extensive use of these compounds.

This dissertation contains unpublished, co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Kory John Iverson Plakos

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
University of California, Santa Barbara, Santa Barbara, CA

DEGREES AWARDED:

Doctor of Philosophy, Chemistry, 2016, University of Oregon
Bachelor of Science, Biochemistry, 2007, University of California, Santa Barbara

AREAS OF SPECIAL INTEREST:

RNA biochemistry
High-throughput sequencing
Structure-function relationships of biomolecules
Applications of small molecule/nucleic acid chemistry

PROFESSIONAL EXPERIENCE:

Laboratory Technician, Soh Lab, University of California, Santa Barbara
January, 2008 – June, 2009

Researcher, Special Technologies Lab, US Department of Energy, Operated by National Security Technologies, LLC. May, 2007 – October, 2007

Researcher, Special Technologies Lab, US Department of Energy, Operated by National Security Technologies, LLC. July, 2006 – September, 2006

Student Professional, Los Angeles Department of Water and Power, Aqueduct Division. June, 2005 – September, 2005

GRANTS, AWARDS, AND HONORS:

Fellowship, Content in Context Super Lessons, University of Oregon, 2015

Fellowship, Science Literacy Program, University of Oregon, 2015

National Academies Education Fellowship, University of Oregon, 2015

Fellowship, STEM CORE, University of Oregon, 2015

Ruth L. Kirschstein NRSA Trainee, Institute of Molecular Biology, University of Oregon, 2013

Travel Award, 17th Annual RNA Society Meeting, 2013

PUBLICATIONS:

Ward, W. L., Plakos, K. & DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chem. Rev.* **2014,** *114*, 4318–4342.

Oh, S., Plakos, K., Xiao, Y., Eisenstein, M., Soh., H. T. In Vitro Selection of Shape-Changing DNA Nanostructures Capable of Binding-Induced Cargo Release. *ACS Nano.* **2013,** *11*, 9675 – 9683.

Oh, S., Plakos, K., Lou, X., Xiao, Y., Soh., H. T. In Vitro Selection of Structure-Switching, Self- Reporting Aptamers. *Proc. Natl. Acad. Sci. USA.* **2010,** *107*, 14053-14058.

Xiao, Y., Lou, X. H., Uzawa, T., Plakos, K. J. I., Plaxco, K.W., H Soh., H. T. An Electrochemical Sensor for Single Nucleotide Polymorphism Detection in Serum Based on a Triple-Stem DNA Probe. *J. Am. Chem. Soc.* **2009,** *131*, 15311- 15316.

Xiao, Y., Plakos, K. J. I., Lou, X. H., Qian, J., White, R.J., Plaxco, K.W., Soh, H. T. Fluorescence Detection of Single Nucleotide Polymorphism via a Single Self-Complementary, Triple-Stem DNA Probe. *Angew. Chem. Int. Ed.* **2009,** *48*, 4354-4358.

L. Chandos, K. R. Kyle, A. Moy, K. Plakos. DNA Capture Material. *Nevada Test Site-Directed Research and Development, FY 2007 Report*. [Online] **2008**, 139–145. http://www.osti.gov/bridge/servlets/purl/927363-TWKqdW/927363.pdf (accessed September 21, 2016).

ACKNOWLEDGMENTS

I would like to sincerely thank my mentor Vickie DeRose for all of her support and guidance throughout my graduate career, and for her hard work assisting in preparing this manuscript. I would also like to thank my committee, Diane Hawley, Alice Barkan, Michael Haley, and formerly Andy Berglund. It has been an absolute pleasure getting to know you over the years. Thank you to Bryan Rebar and Elly Vandegrift for supporting me through STEM-CORE, C2SL, and the Science Literacy Program.

Alan Moghaddam, we unexpectedly became partners in crime these last few months. Long afternoons writing at Roma were a highlight, and something I will truly miss. Rachael Cunningham, Emily Reister-Morris, Emily Sutton, and Adam Saunders, you make the DeRose Lab what it is. Your partnership, friendship, and commiseration got me through graduate school. I would also like to thank former lab members Erich Chapman, who led me down this rabbit hole, Maire Osborn, who was my rock for most of my time here, Luke Ward, and Jon White. I've had the pleasure of working with some wonderful undergraduate students over the years, including Amanda Brenner, Susanna Mirashrafi, and most recently Anna Hickey. I'm so excited to see where your careers take you.

I made some wonderful friends here, especially Jon Marshall and M Jackson who, supported me through hard times and good. I would also like to thank my IMB and Chemistry cohort for being wonderful friends and peers. Matt Bailey, Sarah Casper, Max Rodnick Smith, Andrew Wagner, Gabe Rudebusch, you will always be Eugene for me.

Thank you to my family, who held me up when I needed it, and kept me grounded when I needed it more. And thank you to my dog Commander William T. Riker. Because

of him I always had a good reason to go home at the end of the day, and an excuse to go outside each morning. And I owe a very special, sincere thank you to Katie Bodie. You came into my life toward the end of this and went with it. You stuck by me when things got hard. You put up with so much. Your patience and understanding means the world to me. Thank you.

The investigation was supported in part by National Institutes of Health grant (GM058096), NIH T32 Grant GM007759-29, and National Science Foundation grant CHE-1153147, as well as the University of Oregon's Science Literacy Program, STEM-CORE, and C2SL programs.

This work is dedicated to my family, for their unending support and encouragement.
Without you, I wouldn't be here.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

Figure Page

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

In the early 1960's, Barnett Rosenberg and colleagues were exploring the effect of electric current on *E. coli* cells when they noticed cells grown in the presence of platinum electrodes ceased dividing (Rosenberg et al. 1965). Originally suspecting that the electric field was somehow causing the phenotypic response, they soon discovered it was a platinum compound itself being released from the electrodes that was responsible for this effect. Careful work narrowed the causative agent down to salts of Pt(II) and Pt(IV) (Rosenberg et al. 1969). Rosenberg and colleagues had serendipitously discovered cisplatin (*cis*-$Pt[NH_3]_2Cl_2$), arguably the most important anticancer compound to date (Alderden et al., 2006; Dyson et al., 2006).

Cisplatin (Figure 1.1 A) is comprised of a Pt(II) core with two exchange-inert NH_3 ligands and two labile chloride ligands. In its di-chloro form it is a water soluble, neutral compound. Administration into the high-chloride environment (approximately 100 mM) of a patient's bloodstream favors the uncharged di-chloro state, but upon entering the low chloride environment of a cell, either through passive diffusion (Basu and Krishnamurthy, 2005) or active transport (Safaei and Howell, 2005), one or more of the chloride ligands are exchanged for water, conferring a positive charge upon the compound (Figure 1.1D). Aquated cisplatin is highly reactive and binds to a wide variety of biomolecules, including protein, RNA, DNA, membrane phospholipids, and cytoskeleton components (Jamieson and Lippard, 1999). Phenotypic similarities to DNA damage suggested cisplatin's anticancer activity is attributed to coordination with nuclear

DNA (Jamieson and Lippard, 1999), where it forms covalent bonds with the N7 position of purine nucleobases, most often guanine (Fichtinger-Schepman et al., 1985). A prevailing theory suggests cisplatin-induced DNA lesions induce apoptosis in sensitive cells (Basu and Krishnamurthy, 2005). However that theory is challenged by several observations: cisplatin accumulates more on protein and RNA than DNA (Akaboshi et al., 1992), more cisplatin accumulates on total RNA than total DNA (Hostetter et al., 2012), and cisplatin can induce apoptotic signaling in a nucleus-independent manner (Mandic et al., 2003). These observations indicate a need for a more thorough understanding of cisplatin's effect on biomolecules, particularly RNA.

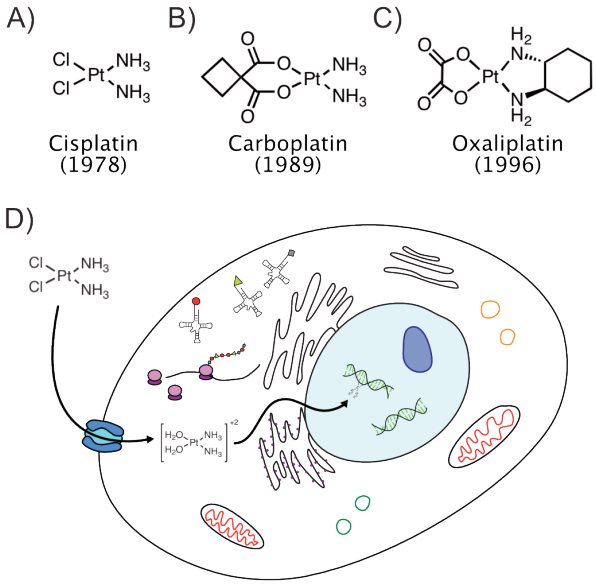


Figure 1.1 – Cisplatin and currently approved derivatives. A) Cisplatin. B) Carboplatin. C) Oxaliplatin. D) Cisplatin cell distribution. Upon entering the cell through passive diffusion (not shown) or active transport (shown), cisplatin becomes aquated and gains a positive charge. It is then attracted to numerous biomolecules, including DNA, RNA, and proteins.

The emerging role of RNA

One of the most unexpected results of the Human Genome Project was that the human genome consists of less than 20,000 protein coding regions, representing less than 3% of our genetic information (International Human Genome Sequencing Consortium, 2001). This remaining material, at least part of which was once termed "junk DNA," (Ohno, 1972), was quickly found to contain regions that in fact are transcribed into noncoding RNA (Maddick, 2004; Ulitsky and Bartel, 2013), constituting some 98% of human genetic output (Mattick, 2001). Long thought to be a passive information carrier shuttling genetic information from DNA in the nucleus to ribosomes in the cytoplasm for protein synthesis, RNA is now recognized to have important roles in peptide bond formation itself, regulating gene expression, and catalyzing chemical reactions, among many other things (Morris and Mattick, 2014).

Recognition of these key roles of cellular RNA has elucidated a whole new category of drug targets. Proteins are still considered the predominant class of drug targets, but estimates of the number of druggable proteins is surprisingly small – just over 200 by one estimate (Overington et al., 2006), most of which are G-protein coupled receptors, ion channels, or other enzymes (Thomas and Hergenrother, 2007; Makley and Gestwicki, 2013). Given the vast output of RNA in a cell, there is a crucial need for tools to better understand the diverse landscape of RNA binding sites and RNA binding events for small molecule compounds such as cisplatin.

Cisplatin accumulation on cellular RNA

Soon after cisplatin was approved for use in cancer therapy, research results began to indicate that RNA played a role in cytotoxicity. Aside from broadly demonstrating platinum accumulation in cells (reviewed in Chapman et al., 2011), studies have suggested that cisplatin interferes with translation (Rosenberg and Sato, 1988) and the RNA-dependent telomere maintenance enzyme telomerase (Burger et al., 1997). Recent studies have shown Pt accumulation on ribosomal RNA (Hostetter et al., 2012; Melnikov et al., 2016). A comprehensive molecular understanding of platinum accumulation on cellular RNA is lacking, however, due to inherent difficulties in identifying platinum adducts (Chapman et al., 2011). Numerous methods have been employed in attempts to infer platinum accumulation on RNA.

Methods of mapping cisplatin RNA binding sites

Covalently modify platinum compounds appended with fluorophores have enabled colocalization studies to interrogate the compartmentalization of cisplatin in treated cells. Fluorophores based on fluorescein (Molenaar et al., 2010), NBD (Benedetti et al., 2011), BODIPY (Jagodinsky et al., 2015), and acridine (Ding et al., 2013), have been appended to cisplatin-like scaffolds and studied for their localization in cells. Accumulation of these various groups was mostly diffuse; cytoplasmic, vesicular, golgi, lysosomal, and nucleolar accumulation were all seen, supporting no specific sites of platinum accumulation. Given the bulky, lipophilic nature of many of these fluorophores, the question remains whether the appended fluorescent moieties or the platinum core itself was directing localization.

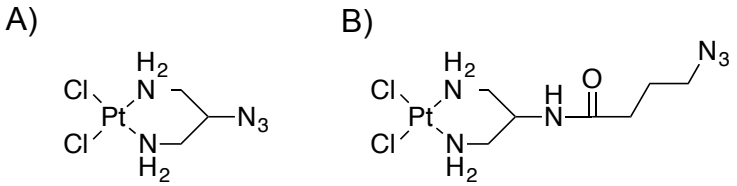


Figure 1.2 – Click-functionalized platinum compounds. A) 2-ADAP Pt (Moghaddam et al., 2015). B) Azaplatin (Wirth et al., 2015). Both compounds were synthesized in our lab, and feature in work presented later.

To overcome these concerns, the DeRose group (Moghaddam et al., 2015; Wirth et al., 2015), and the Bierbach (Ding et al., 2013) group have synthesized several click-ready platinum compounds for post-treatment fluorescent tagging, two of which are featured later in this work (Figure 1.2). As compared with the above Pt-fluorophore conjugates, click-ready compounds are minimally modified before reacting with cellular components. Cells are treated with a one of these click-ready compounds, fixed, permeabilized, and post-treatment modified using the azide-alkyne cycloaddition reaction with their fluorescent partner, avoiding unwanted interactions of fluorophores with cellular components (White et al., 2016). While post-platination click chemistry is not the focus of this work, RNA binding sites of the compounds in Figure 1.2 are mapped to single-nucleotide resolution and compared with those from cisplatin.

Though techniques are maturing rapidly, RNA crystallography is still a challenging area. To date, there are only two molecules crystallized with cisplatin on RNA: phenylalanine tRNA (Rhodes et al., 1974; Sundaralingam et al., 1985) and, most recently, the *T. thermophilus* ribosome, *in vitro* (Melnikov et al., 2016). In the ribosome crystallography experiment, the authors found that treatment with cisplatin *in vitro* resulted in nine major mainly monofunctional Pt adducts; only one adduct bonded to

adjacent guanines. The authors also noted a correlation between cisplatin-modified guanine bases and Mg^{2+} coordination at that site (Melnikov et al., 2016).

Analytical and spectroscopic techniques have also been applied to the study of cisplatin accumulation on RNA. Atomic absorption spectroscopy was used in a study by Roberts and colleagues to quantify platinum accumulation on biomolecules in HeLa cells. Their results suggest that more cisplatin is bound to RNA than DNA or proteins (Pascoe and Roberts, 1974). Likewise, Hostetter et al. used inductively coupled plasma mass spectrometry (ICP-MS) to quantify cisplatin accumulation in mRNA, rRNA, total RNA and DNA in *S. cerevisiae* (2012). They found greater accumulation of platinum on DNA than RNA on a per-nucleotide basis, but when the ratio of DNA to RNA is considered, platinum is approximately 4-20-fold more populated on total RNA than DNA (Hostetter et al., 2012). Matrix-assisted laser desorption ionization (MALDI) mass spectrometry has also been used to study inhibition of RNA enzymes *in vitro* by platinum adducts on RNA (Chapman and DeRose, 2010), and cisplatin accumulation on model ribosomal RNA hairpins (Dedduwa-Mudalige and Chow, 2015). In this work, highly conserved hairpins from *E. coli* (helix 69 from 23S rRNA and helix 24 from 16S rRNA) were modeled, platinated, and probed using a combination of RNAse T1, MALDI mass spectrometry, and dimethyl sulfate mapping to explore sites of platinum accumulation. The authors found platination at several GpG sites and noted similarities to antibiotic binding sites (Dedduwa-Mudalige and Chow, 2015). Nuclear magnetic resonance (NMR) spectroscopy has been applied to platinated DNA duplexes (Gelasco and Lippard 1998) but to the best of our knowledge, has not been applied to the question of cisplatin-derived adducts on RNA.

Mass spectrometry and crystallography-based techniques are well-suited for mapping RNA modification sites to single nucleotide resolution but these techniques are labor intensive, cost prohibitive, and require homogenous purified RNA, often in large amounts. Cisplatin adducts on DNA or RNA interfere with processive enzymes such as RNA polymerase II (Damsma et al., 2007) and reverse transcriptase (Rijal and Chow, 2009; Chapman and DeRose, 2010), allowing researchers to use reverse transcription (RT) primer extension analysis techniques to map cisplatin RNA binding (Stern et al., 1988; Chapman and DeRose, 2010). RT primer extension analysis creates cDNA copies of RNA templates, truncated 3' to the site of RNA modification (Stern et al., 1988). Gel-based readouts correlate fragment length with sequencing reference lanes in order to identify the modified nucleotide. RT primer extension experiments do not require as much RNA, work well in heterogenous RNA samples, and are inherently less labor-intensive than other analytical methods. Numerous RNA-platinum adducts have been studied using RT primer extension, both in synthetic constructs modified *in vitro* (Chapman and DeRose, 2010; Osborn et al., 2014), purified *E.* coli ribosomal RNA treated *in vitro* with cisplatin (Rijal and Chow, 2009), and ribosomal RNA isolated from *S. cerevisiae* treated with cisplatin (Hostetter et al., 2012; Osborn et al., 2014). The ribosome has been the focus of most of these mapping studies, discussed in detail below.

The ribosome is the large ribonucleoprotein complex responsible for protein synthesis. Peptide bond formation is catalyzed by the RNA components of the ribosome, making the ribosome a ribozyme (Moore and Steitz, 2003). Proper ribosomal function is absolutely crucial to cell survival. Single point mutations to ribosomal RNA can cripple the complex entirely, leading to inviable phenotypes (Rakauskaite and Dinman 2008).

Numerous antibiotics are known to disrupt protein synthesis by targeting both bacterial (Wilson 2014) and eukaryotic (Garreau de Loubresse et al., 2014) ribosomes, especially at the peptidyl transferase center (PTC). The toxins sarcin and ricin each act at the PTC as well, both damaging the sarcin ricin loop (SRL, 25S helix 95), the longest universally conserved sequence in the ribosome. Ricin depurinates A4324 in the rat 23S rRNA and sarcin cleaves the phosphodiester bond between G4325 and A4326. Both alterations halt protein synthesis, further demonstrating the ribosome's sensitivity to single nucleotide modifications (Tumer and Li, 2011).

The ribosome is a particularly attractive target for cisplatin, as it is located primarily in the cytoplasm and highly abundant (some 80% of yeast RNA is ribosomal RNA (Warner, 1999)). The Chow group explored cisplatin binding to bacterial ribosomal RNA hairpins using targeted RT primer extension analysis (2009). *E. coli* ribosomes were extracted and purified, then treated with cisplatin and analyzed using RT primer extension with a primer designed to anneal upstream of 16S helix 24. They found context-dependent platinum adducts on purine-rich regions, showing that helix 24 is clearly accessible to cisplatin (Rijal and Chow, 2009). Expanding upon this work, the DeRose group has mapped several platinum binding sites in ribosomal RNA isolated from cisplatin-treated yeast. Yeast cells were treated with cisplatin, the RNA extracted and purified, and primer extension performed across helix 18. This experiment showed five main dose-dependent stops along yeast helix 18, as compared with two reported by Rijal and Chow (2009).

The DeRose group also performed RT primer extension across the sarcin ricin loop on RNA extracted from yeast treated with cisplatin. Primer extension analysis

reveals platinum accumulation to the purine-rich terminal loop, near the substrates for sarcin and ricin (Osborn et al., 2014). Helix 90, another component of the PCT, was also explored and revealed several concentration-dependent cisplatin binding sites. Primer extension analysis suggests that cisplatin adducts are surprisingly frequent in ribosomal RNA.

Traditional RT primer extension approaches use targeted primers to analyze specific RNA strands, and due to limitations in sequencing gel resolution have a cleanly resolvable window of approximately 30-60 nucleotides. Combined, the ribosome mapping efforts described above for *S. cerevisiae* represent approximately 130 mapped nucleotides out of a total of 5196 nucleotides across the large and small ribosomal subunits (3396 nt in yeast 25S plus 1800 nt in yeast 18S), or just 2.5% of the yeast ribosome (Engel et al., 2013). Considering the myriad of other RNAs in the cell, using the techniques outlined above to gain a comprehensive picture of RNA binding is impractical. There is a strong need for more high-throughput methods of identifying cisplatin adduct sites.

High throughput sequencing of modified RNAs

High-throughput sequencing is well-suited for shotgun-style approaches to large-scale sequencing problems (Reuter et al., 2015). Techniques such as Mod-seq (Talkish et al., 2014) and StructureFold (Tang et al., 2015) couple DMS probing, a well-established technique for analyzing RNA structure, to high-throughput sequencing for secondary structure analysis of large RNAs. DMS is a nucleobase-modifying reagent that methylates the N1 of adenosine and N3 of cytosine (Lawley et al., 1963). As with

cisplatin adducts, DMS can be mapped in a targeted manner by RT primer extension analysis, because these modifications cause reverse transcriptase to stall as described above (Tijerina et al., 2007). Mod-seq and StructureFold each take shotgun-styled approaches to sequencing rRNA modifications by fragmenting treated RNA, reverse transcribing the resulting fragments, and sequencing cDNAs. Sequenced libraries are aligned to a source genome and modification sites in treated samples are reconstructed by identifying 3' ends enriched over background, since these represent RT products truncated by DMS (Talkish et al., 2014; Tang et al., 2015). Mod-seq differs from StructureFold in that the technique employs an enrichment step to preferentially sequence cDNA's resulting from DMS-modification stops (Talkish et al., 2014; Tang et al., 2015).

Modifications to the ribose backbone have also been mapped using RT-based approaches coupled to high-throughput sequencing. SHAPE chemistry uses alkylating agents to covalently modify accessible 2' hydroxyls on the RNA ribose backbone. Coupled to high-throughput sequencing, methods have been developed to map SHAPE modifications by both primer termination, similar to ModSeq and StructureFold (SHAPE-seq, Waters et al., 2015) and through the detection of mutations conferred by RT improperly processing modified nucleotides (SHAPE-MaP, Siegfried et al., 2014). Using SHAPE-MaP, the authors were able to develop secondary structure models for the entire HIV genome (Siegfried et al., 2014). Endogenous RNA 2'O-methylation can also be detected using RiboMethSeq, which uses small amounts of starting material (1 ng of total RNA) to detect all methylation sites in rRNA (Marchand et al., 2016).

Target enrichment is an important consideration when sequencing nucleotide modifications, as low-frequency modifications might be lost to background noise or

require numerous PCR cycles during library preparation, artificially enhancing signal (Mamanova et al., 2016). ModSeq accomplishes this through subtractive hybridization of fully extended cDNA's, preferentially enriching for truncations at modified nucleotides (Talkish et al., 2014). Another approach uses antibodies specific for the modification of interest. A recent work by Dominissini and coworkers used methylated RNA andibodies to selectively enrich and sequence the location of N^1 methyladenosine modifications, which had up to this time only been observed on tRNA and rRNA, on mRNA (Dominissini et al., 2016). This technique is reminiscent of Chem-seq, a sequencing technique that aims to sequence genomic binding sites of DNA-binding small molecules (Kapoor et al., 2016).

Chem-Seq

Understanding the genomic targets of small molecules is of critical importance for drug design and discovery. Dozens of FDA approved and clinical drug candidates interact directly with DNA and with DNA binding proteins (Rodriguez and Miller, 2014). A method for unbiased genome-wide identification of DNA-binding small molecules and proteins has its roots in ChIP-seq (Chromatin Immuno Precipitation followed by high-throughput sequencing), which broadly describes techniques of profiling the location of chromosome-binding proteins through affinity pulldown of fragmented DNA followed by high-throughput sequencing (Johnson et al., 2007). This technique has enhanced, among others, understanding of transcription factors (Valouev et al., 2008) and histone modifications (Barski et al., 2007). Recently, ChIP-seq principles have been applied to

studying the genomic binding sites of small molecules in a technique called Chem-seq (Anders et al., 2014; Rodriguez and Miller, 2014).

In one of the first examples of Chem-seq, a collaborative effort between the Bradner and Young groups (Anders et al. 2014), a small-molecule inhibitor (JQ1) of BRD-family transcription factor proteins was covalently modified with a biotin functional group (bio-JQ1). MM1.S cells were treated with bio-JQ1, DNA was harvested and fragmented, and fragments were enriched using streptavidin pulldown and sequenced. These data were compared with ChIP-seq on BRD2, BRD3, and BRD4, four proteins known to be inhibited by JQ1. ChIP-seq and Chem-Seq profiles correlated well and followed a pattern predicted by known affinities of JQ1 for these three proteins *in vitro*. Expanding upon this Chem-Seq technique, the authors performed Chem-Seq on biotinylated psoralen, a DNA intercalator recently shown to preferentially intercalate at transcription start sites. Chem-Seq confirmed this finding by comparison with ChIP-seq identification of RNA Pol II, establishing Chem-Seq as a method for directly mapping small molecule binding regions on genomic DNA (Anders et al., 2014). Chem-Seq shows promise for greatly enhancing our understanding of how small molecules interact with DNA.

Recent work has made great contributions toward better understanding the extent to which cisplatin interacts with cellular components, and this body of work represents a significant step toward comprehensive, high-throughput analysis of cisplatin binding sites in cellular RNA.

This dissertation includes unpublished co-authored material. Chapter II will be published with co-authors Barbara Golden (Purdue University) and Victoria DeRose.

Chapter III will be published with Victoria DeRose as a co-author. Chapter IV was co-authored with Emily Reister and Victoria DeRose. Chapter V was co-authored with Anna Hickey and Victoria DeRose.

CHAPTER II

RNA CROSSLINKING USING PT(II) COMPLEXES TO PROBE THE TERTIARY STRUCTURE OF THE HDV RIBOZYME

This chapter has contributions from Professor Barbara Golden (Purdue University) and Professor Victoria DeRose. I performed and analyzed the experiments and wrote the text. Professor Golden provided several of the RNA strands used. Professor DeRose provided direction and advice for the project and edited the manuscript. Additionally, small but valuable contributions were made by Adam Unger, for assisting with control experiments, and Amanda Brenner for her work on crosslinking kinetics.

Introduction

Understanding the tertiary structure of complex RNA is critical to understanding its function. Structured RNA plays roles in diseases, such as in subgenomic flaviviral RNA which adopts a particular tertiary structure that prevents XRN1 degradation by the host cell (Chapman et al., 2014). Additionally, it has been postulated that some single nucleotide polymorphisms lead to a disease state by altering the structure of certain regulatory RNAs (Halvorsen et al., 2010). Beyond disease states, higher-order RNA structure is important for protein recognition (Li et al., 2014) and the proper function of autocatalytic RNA sequences such as the hammerhead ribozyme, hepatitis delta virus (HDV) ribozyme, and group II intron (Butcher and Pyle, 2011; Ward et al., 2014).

Numerous methods are used to analyze RNA structure. X-ray crystallography, when possible, is an extremely important technique (Holbrook and Kim, 1997). X-ray

structures reflect a single crystallized population, however, and the relationship between these models and RNAs in solution require verification by other means (Golden and Kundrot, 2003). There are numerous chemical approaches used to predict RNA tertiary structure in solution conditions. SHAPE (selective 2' hydroxyl acylation analyzed by primer extension, Deigan et al., 2009), hydroxyl radical cleavage (Hampel and Burke, 2001), lead probing (Lindell et al., 2002), and in-line cleavage (Beaudoin et al., 2013) all report on conformationally flexible regions and backbone solvent accessibility and can be used to predict globally folded structures. They do not, however, report directly on tertiary contacts within a folded RNA.

Tertiary contacts in a complex RNA can be identified through the use of chemical crosslinking techniques that promote the formation of covalent bonds between nucleotides brought into close proximity (Harris and Christian, 2009). Extensive work has been done using UV crosslinking of native and UV-active modified nucleobases such as 4-thiouridine (Favre et al., 1998; Graifer and Karpova, 2013). The modified bases must be incorporated into the RNA being studied, either randomly *in vivo* or specifically through synthetic methods *in vitro* (Favre et al., 1998). UV-independent crosslinking reagents include formaldehyde (McGhee and von Hippel, 1975), a very short-range crosslinker. Other organic small molecule crosslinking reagents have been developed (reviewed in Harris and Christian, 2009).

While they have been effective, chemical crosslinking methods often suffer from poor yields (for example, 1-3% for UV-induced crosslinking, Zhirnov and Wollenzien, 2003) and require enrichment and purification steps (Harris and Christian, 2009). There is a need for reagents that can both crosslink with high yield and also affinity tag the RNA

for enrichment and structural analysis. We are developing a new method of RNA crosslinking using the coordination properties of Pt(II) compounds and based on the well-studied anticancer drug cisplatin and cisplatin-derived small molecules (Figure 2.1). These compounds exhibit interstrand RNA crosslinking yields as high as 40%, are specific for certain canonical nucleobases, and can be kinetically recruited to the backbone through addition of synthetically derived phosphorothioate substituted RNA, allowing us to examine interactions between specific backbone locations and nearby nucleophiles (Chapman and DeRose, 2012).

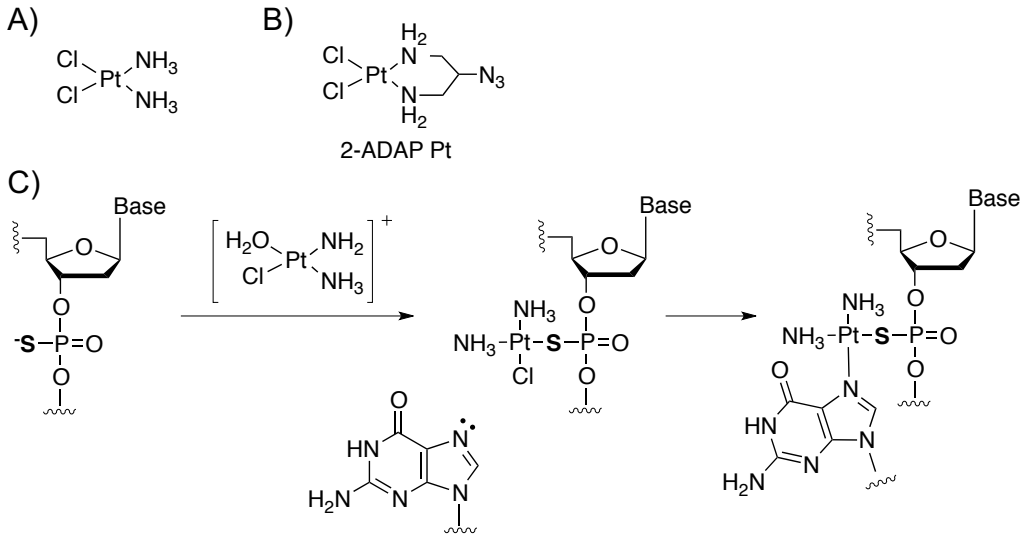


Figure 2.1. Platinum complexes and model of phosphorothioate-directed crosslinking. A) Cisplatin. B) 2-ADAP Pt (2-azido 1,3 diaminopropane Pt(II). C) Mechanism of phosphorothioatedirected cisplatin crosslinking (Chapman and DeRose 2012).

Cisplatin (Figure 2.1A) is a well-studied molecule and widely used as an anticancer compound (Jamieson and Lippard, 1999). Substitution at the Pt(II) metal center is known to form coordinate covalent bonds with nucleophilic nitrogen donor ligands in both DNA and RNA (Chapman et al., 2011). Additionally, Pt(II) adducts exhibit 3-4 fold kinetic preference with thiol residues (Elmroth and Lippard, 1994).

Incorporation of phosphorothioate-substituted RNA into folded complexes results in a kinetically-preferred binding site. Pt(II) complexes recruited to this site then sample nearby nucleophiles, creating crosslinks based on distance and orientation (Figure 2.1C, Chapman and DeRose, 2012). This phosphorothioate-directed crosslinking approach provides the ability to recruit cisplatin to specific sites along an RNA backbone to study interactions between the backbone and nucleobases. Such interactions are sometimes mediated by metal ions, as in the active sites of ribozymes including the hammerhead and the hepatitis delta virus (HDV) ribozyme (Golden, 2011). We have previously demonstrated the use of cisplatin crosslinking in probing the active site of the hammerhead ribozyme, where cisplatin crosslinks were specific to certain nucleophiles within the active site and dependent on proper ribozyme folding (Chapman and DeRose, 2012). Here we use this technique to explore the in-solution tertiary structure of the HDV ribozyme (Figure 2).

The Hepatitis Delta Virus (HDV) ribozyme (Figure 2.2A and B) was originally found in the RNA genome of the Hepatitis D virus, where it catalyzes self-cleavage as part of processing genomic units during rolling circle replication (Lai, 1995; Golden, 2011). Numerous HDV-like ribozymes have recently been discovered, including the HDV-like CPEB3 ribozyme, highly conserved in mammals (Salehi-Ashtiani et al., 2006), and the R2 retrotransposon, found in most arthropods (Eickbush et al., 2013).

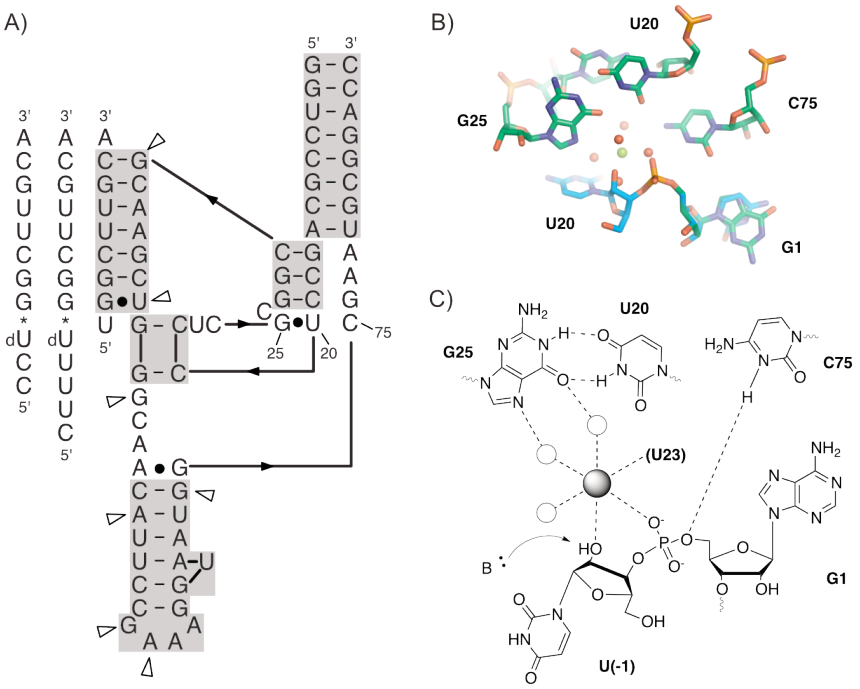


Figure 2.2. HDV ribozyme secondary and tertiary structure, based on the crystal structure by Chen et al. (2011). A) The HDV secondary structure, annealed to a cleavable substrate strand. White arrows represent observed platination sites (see text). Two substrate strands (11 and 13 nt) used in these experiments are shown, containing both 2'-deoxy and phosphorothioate substitutions at the cleavage site. Sulfur substitutions at the substrate cleavage site are indicated by * (phosphorothioate substitution). B) Model of the HDV ribozyme active site constructed as in Chen et al. (2010). Shown are key nucleotides implicated in catalysis from the enzyme strand (G25, U20, C75, green, from PDB ID 3NBK) and the modeled-in substrate strand (blue, from PDB ID 2OEU, Martick et al., 2008). C) Mechanistic model for the HDV, reviewed in (Golden, 2011). The active-site metal is in grey with predicted aqua ligands (white circles).

The HDV ribozyme utilizes a divalent metal ion for catalysis that is competitively inhibited by exchange inert cobalt hexamine, a hydrated magnesium ion mimic, indicating a requirement for inner-sphere coordination (Nakano et al., 2000). The latest crystal structure of the pre-cleavage HDV ribozyme shows a partially hydrated magnesium ion located in the active site near the scissile phosphate (Figure 2.2B), interacting through outer sphere coordination with two water molecules with a G25•U20 reverse wobble and a single inner-sphere coordination to the pro-Sp oxygen on U23

(Chen et al., 2010) (Figure 2.2C). On the inhibitor substrate strand, the authors were able to resolve G(1); however, due to poor electron density upstream of G(1), a clear picture of the scissile phosphate and U(-1) was not possible with this crystal. Given that cleavage of the HDV ribozyme requires that the 2' hydroxyl of U(-1) be aligned for in-line attack, the authors superimposed the structure of the nucleotide downstream of the hammerhead ribozyme cleavage site (PDB 2OEU (Martick et al., 2008)) with G1 from their crystal structure and found that the resulting structure is consistent with the crystallographic data (Figure 2.2B) (Chen et al., 2010).

Based on this model, Chen and colleagues hypothesized that the catalytic magnesium ion interacts through inner-sphere coordination with the 2'-OH of U(-1), lowering its pKa and activating it for nucleophilic attack on the scissile phosphate (Figure 2.2C) (Chen et al., 2010). This metal ion is also hypothesized to interact with the pro-Rp oxygen at the scissile phosphate, in agreement with phosphorothioate substitution studies (Fauzi et al., 1997; Wrzesinski et al., 2010). The contacts that orient the metal ion in the active site include an outer-sphere coordination with the N7 position on G25 of the enzyme strand (Figure 2.2C). This presents a good candidate for our technique of phosphorothioate-directed cisplatin crosslinking, and we hypothesize that a thiol substitution to a nonbridging oxygen at the scissile phosphate will attract our crosslinking platinum complexes, which should then crosslink with G25.

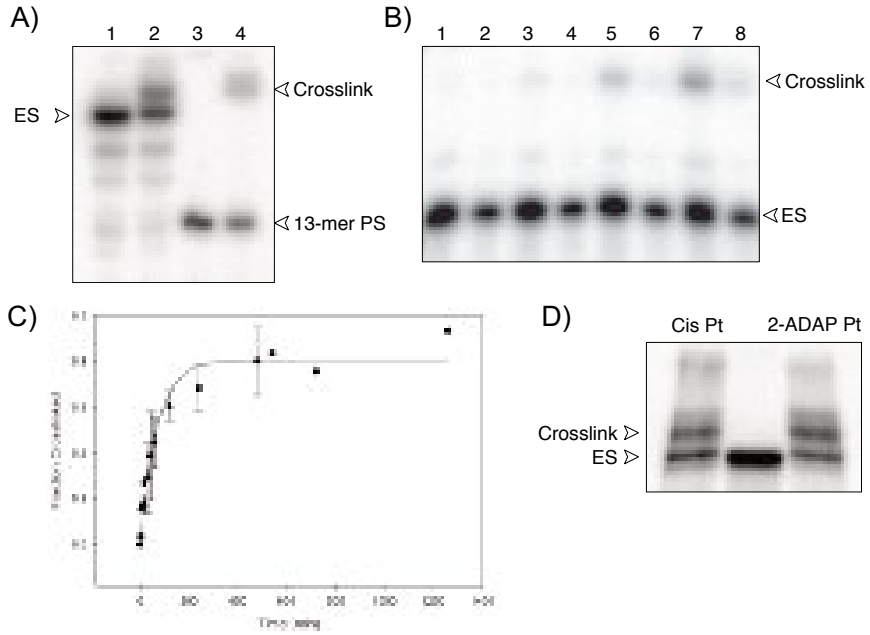


Figure 2.3. Cisplatin induces interstrand crosslinks between the HDV ribozyme enzyme strand (ES) and a phosphorothioate substituted substrate (SS). A) ES and SS (13-mer,PS,dU1) crosslinks observed after three hours using either 32P 5' end-labeled enzyme (lanes 1 and 2) or substrate strand (lanes 3 and 4). Odd numbered lanes are quenched at T=0, even numbered lanes at T=3 hours. B) Crosslinks are dependent on phosphorothioate substitution (11-mer,PS,dU1), and increase with increasing cisplatin concentration. In odd-numbered lanes, substrate strand is phosphorothioate-substituted 11-mer,PS,dU1. In even-numbered lanes, substrate is unsubstituted 11-mer,dU1. Compared with ES concentration, lanes 1 and 2 contain no Cis Pt; lanes 2 and 4, 0.1x Cis Pt; lanes 5 and 6, 1x Cis Pt; lanes 7 and 8, 3x Cis Pt. C) Time-dependent crosslinks visualized using 13-mer 32P 5' end-labeled substrate strand, with an observed rate constant of kobs = .014 ± 0.002 min-1, and a pseudo second-order rate constant calculated to be k2, calc = 3.97 ± 0.47 M-1 s-1 (Chapman and DeRose, 2012). D) Cisplatin and 2-ADAP Pt each induce approximately 40% crosslinking between ES and PS-11, as compared with no platinum control (quantification not shown). Briefly, HDV ribozyme ES and SS strands were folded by heating and slow cooling under conditions described in appendix A. $Mg(NO_3)_2$ and either cisplatin or 1,3-platin were added to initiate crosslinking. Reactions were quenched in formamide loading buffer and run on denaturing 1X TBE 8 M urea polyacrylamide gels.

In this paper, we demonstrate platinum(II) crosslinking of HDV ribozyme constructs with a phosphorothioate substitution at the scissile phosphate and map where these crosslinks are occurring. We obtained results consistent with geometric predictions from the current HDV active site models. Major sites of Pt(II) crosslinking throughout

the ribozyme are also identified, providing benchmark information on the utility of these compounds for RNA structure analysis. Additionally, we demonstrate the first crosslinking use of 2-ADAP Pt (Figure 2.1B), a cisplatin derivative capable of participating in click reactions, which are a class of fast, high-yielding, bio-orthogonal reactions that occur under mild conditions (Kolb and Sharpless, 2003) and are of great interest to biochemists for use in attaching affinity tags (Speers and Cravatt, 2004).

Results and Discussion

Treatment of pre-folded HDV ribozyme complexes containing a phosphorothioate-substituted substrate strand results in Pt-RNA crosslinks with yields of up to 40% (Figure 2.3). Cisplatin-induced crosslinks are visible using both 32P 5' end-labeled enzyme strand and substrate strand after three hours, indicating that both strands are present in the final crosslinked product (Figure 2.3A). Crosslinks were found to be phosphorothioate-dependent, with very low yields in a non-phosphorothioate SS, and increase with increasing cisplatin concentration (Figure 2.3B). Furthermore, crosslinked complexes collapse to a single band corresponding to an uncrosslinked enzyme strand after treatment with thiourea at 37 °C, consistent with reversal of cisplatin- induced crosslinks (Figure A.2). The observed crosslinking rate was kobs = .0143 ± 0.0017 min-1, with a pseudo second-order rate constant calculated to be k2, calc = 3.97 ± 0.47 M-1 s-1 (Figure 2.3C). This is an order of magnitude slower than cisplatin crosslinking observed for the hammerhead ribozyme (kobs = 0.14 ± 0.01 min-1, k2,calc = 19 ± 1 M-1 s-1 under 10 mM $Mg(NO_3)_2$, the closest conditions tested (Chapman and DeRose, 2012). This likely reflects differences in the organization of these active sites. The HDV ribozyme

enzyme strand lacks guide nucleotides that interact upstream of the scission site on the substrate strand, so a certain degree of disorder is inherent in the active site (Chen et al., 2010; Golden, 2011). This is reflected in the fact that the highest-resolution crystal structure of a pre-cleavage HDV ribozyme is lacking density for nucleotides immediately surrounding the scissile phosphate (Chen et al., 2010). The hammerhead ribozyme, on the other hand, has a more structured active site, with a number of base pairs between the substrate strand and the enzyme strand both upstream and downstream of the scissile phosphate (Martick et al., 2008). To test whether our trans-acting system is folding correctly, we substituted our inhibited 11-mer substrate strand for a cleavable version (11-mer,PO) and found that it cleaves at rate comparable to published rates for a trans-acting HDV ribozyme (Ananvoranich and Perreault, 1998), suggesting our deoxy-inhibited and phosphorothioate substituted complex represents an active complex (Figure A.3).

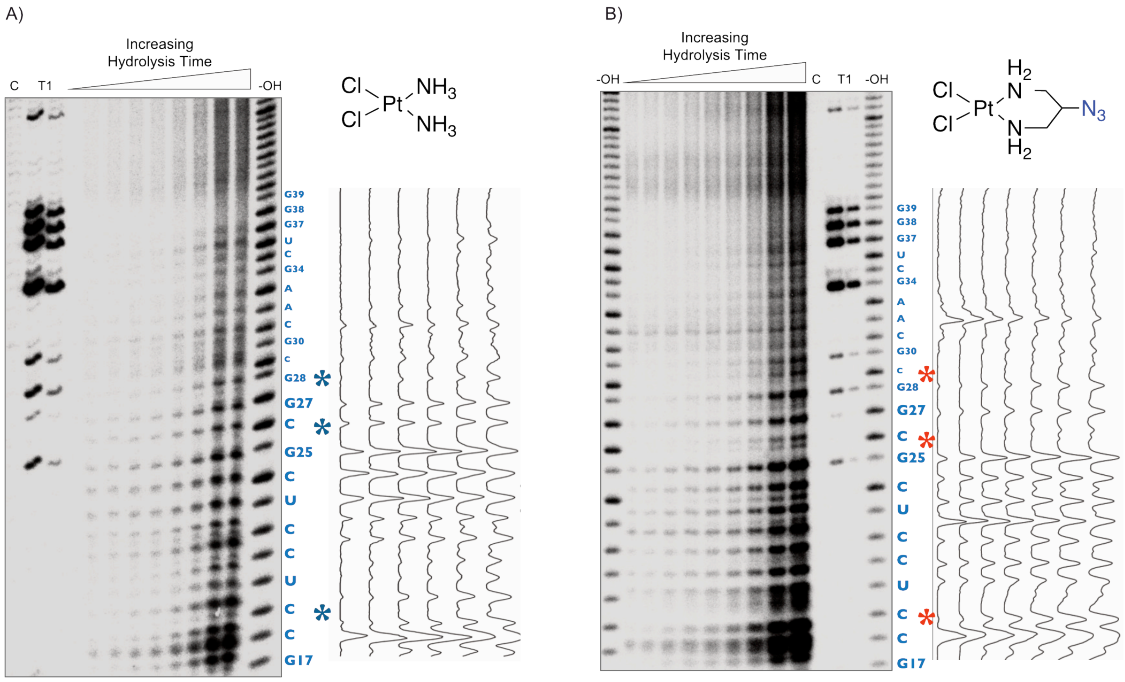


Figure 2.4. Alkali hydrolysis mapping of Pt(II)-crosslinked HDV ribozyme enzyme strand and 11mer phosphorothioate substituted substrate strand. Cleavage products generated by alkali hydrolysis of folded and A) cisplatin or B) 2-ADAP Pt crosslinked ES-SS(11-mer,PS,dU1) complex are shown using 5'-end-labeled HDV ribozyme ES. Band intensity down each lane is plotted next to each experiment. Lanes as labeled represent, in the order seen in A): Control (C); two lanes of partial digestion by two concentrations of RNase T1 (T1); Pt(II) crosslinked complex hydrolyzed for increasing time (overline); fully hydrolyzed reference ladder (-OH); asterisks designate predicted crosslink partners based on these results.

Cisplatin and the more sterically hindered 2-ADAP Pt are both capable of approximately 40% crosslinking (Figure 2.3D), which is consistent with accessibility around the active site of the HDV. The phosphorothioate-dependent crosslinks formed in Figure 2.3 are predicted to occur due to Pt(II) coordination to the phosphorothioate sulfur of the substrate strand and a second coordination to a nucleobase of the HDV enzyme strand. This is further supported by thiourea reversal of crosslinked complex (Figure A.1).

Crosslinking partners on the HDV enzyme strand were identified by hydrolysis mapping following isolation of cisplatin- and 2-ADAP Pt crosslinks (Figure 2.3D). Hydrolysis yields a ladder of the 5'-end labeled HDV enzyme sequence that is interrupted by the crosslinked partner(s) (Juzumiene et al., 2001; Chapman and DeRose, 2012). Hydrolysis mapping of cisplatin and 2-ADAP Pt crosslinked ES PS-11 suggests the two platinum(II) compounds react very similarly (Figure 2.4). Distinct changes in relative peak intensities at specific sites indicates crosslinks to both C19 and C26 for both compounds. Crosslinks to a cytosine, presumably the C3 imino nitrogen, are surprising because it has been demonstrated that the preferred Pt(II) binding site is the N7 position on guanine. Platination at cytosine is not unprecedented (Sherman and Lippard, 1987), however, and cisplatin and its structural isomer transplatin have both been observed to form interstrand GC crosslinks in vitro that include the C3 imino nitrogen of cytosine (Dalbiès et al., 1994).

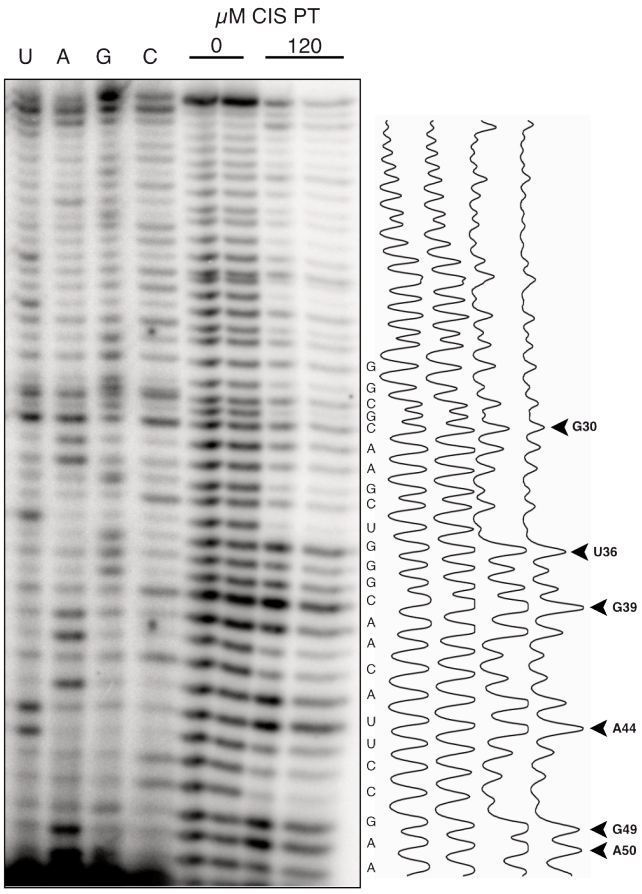


Figure 2.5. Reverse transcription primer extension analysis of cisplatin crosslinked HDV enzyme strand. Reverse transcriptase is stalled by platinum adducts on RNA, which results in a buildup of cDNA transcripts 3' of the platinum site. We found high concentrations of platinum in purine rich regions of the HDV ribozyme, as well as several candidate interstrand and sequentially distant intrastrand crosslinks.

In the HDV crystal structure by Chen et al. (2010), C26 is close to the active site but flipped outward away from the predicted phosphorothioate sulfur position (Figure 2.6). In this static model, C26 would be prohibitively far away for a Pt(II) crosslink, but the nature of this ribozyme suggests the 5'-end of the substrate strand is very flexible (Chen et al., 2010; Golden, 2011) which may make C26 more accessible to the phosphorothioate substitution in solution. C19 is located near the active site, base paired with G27, and the C19 C3 imino nitrogen is pointed toward the active site magnesium ion, with distances to the active-site scissile bond of 8.6 Å from the pro-Rp and 9.3 Å

from the pro-Sp position (not shown). These observations are consistent with a phosphorothioate-directed crosslink between an active-site phosphorothioate and C19 for both platinum compounds.

In addition to the crosslinks described above, cisplatin and 2-ADAP Pt each exhibit a unique crosslink, with the difference likely due to the steric difference between the two molecules. Cisplatin appears to additionally crosslink G28 (Figures 4A, 6A). G28 is located in the active site but not predicted to participate in catalysis (Chen et al., 2010). Furthermore, in the crystal structure G28 is oriented so its N7 position is pointed away from the scissile phosphate, but it is still within crosslinking distance (8.4 Å, Figure A.6). 2-ADAP Pt appears to crosslink C29 (Figure 2.4B, 6B). The N3 position on C29 also faces away from the active site in the HDV crystal structure. An interesting possibility that agrees with crosslinks to C19, G28, and C29 is shown in Figure 2.7. C19, observed as a crosslinking partner for both cisplatin and 2-ADAP Pt via hydrolysis mapping (Figure 2.4), is proximal to both G28 and C29 and oriented so that its nucleophilic N3 position points toward the active site. This puts its N3 nucleophile within reach of not only the scissile phosphate but also the N7 position on G28 and the N3 position on C29 (measured in Figure 2.7). Thus, all three of these hydrolysis-mapped crosslinks could be explained by intrastrand crosslinks between C19-G28 for cisplatin and C19-C29 for 2-ADAP Pt. Though these would not be phosphorothioate-directed crosslinks, they report on the arrangement of sequentially distant nucleobases near the active site and represent structurally relevant crosslinks. To our knowledge, cytosine has only been observed crosslinking to a base-paired guanine residue in vitro (Dalbiès et al., 1994). Because Pt(II) would be expected to first bind to a more reactive site, such as a purine N7 or a

phosphorothioate, a direct C-C crosslink is less likely to occur. All observed crosslinks are consistent with distances measured from the crystal structure (Figure 2.8C). The hydrolysis patterns were distinct from a mock-crosslinked and identically hydrolyzed control (Figure A.5).

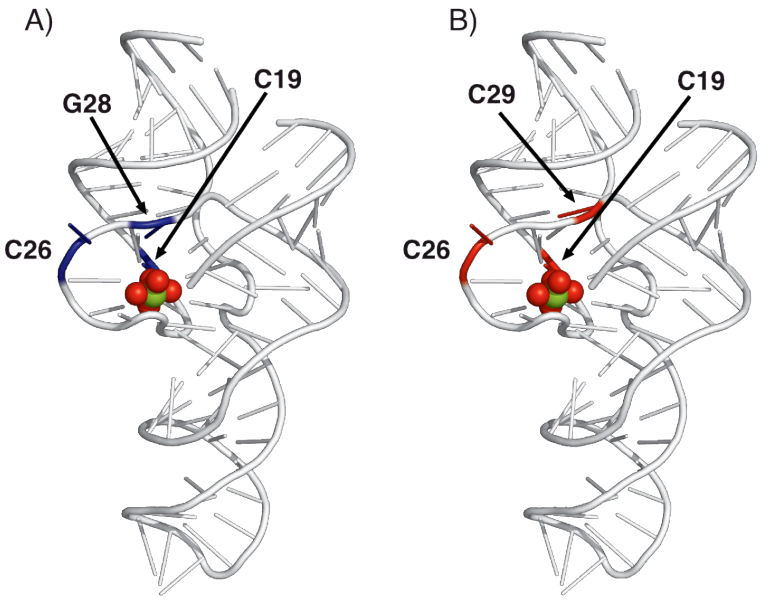


Figure 2.6. Observed Pt(II) crosslink sites between an active-site phosphorothioate and the HDV ES based on hydrolysis experiments. A) cisplatin-induced crosslinks are colored in blue and numbered. B) 2-ADAP Pt-induced crosslinks are colored in red and numbered. The hydrated magnesium ion resolved in the crystal structure is shown as spheres. Figures derived from PDB 3NKB.

In order to identify additional Pt binding sites on the HDV enzyme strand, the crosslinked complex was subject to reverse transcription using a 32P 5'-end labeled RT primer complementary to the 3'-end of the HDV ES (Figure 2.5). Platinum adducts stall reverse transcriptase, causing a build up of cDNA transcripts 3' to the adduct site (Chapman and DeRose, 2010). Several of these 'stop sites' are observed to occur with increasing intensity in samples treated with increasing concentrations of cisplatin (Figure 2.5). The majority of sites indicate Pt adduct formation on a purine of the HDV ES. Since

the most stable Pt adducts are expected to form with two RNA ligands, comparison of RT results with the HDV crystal structure further allows us to predict whether a particular stop site is due to a Pt adduct on neighboring purines, or due to a crosslink within the complex (Figure 2.8). It is believed that purine residues within 10 Å of a platination site may be bridged by cisplatin (Chapman and DeRose, 2012). Shown in Figure 2.8 are all possible partners for each Pt adduct site identified by RT.

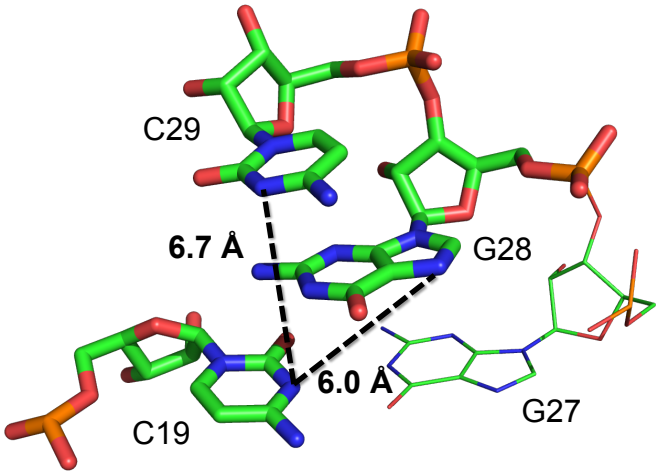


Figure 2.7. Distance measurements between N3 on C19 and N7 on G28 and N3 on C29. G27, which base pairs with C19, is shown as lines. From PDB 3NKB (Chen et al., 2010).

From prior work we expect platinum(II) crosslinking reagents to bind to sequential purine residues (Hostetter et al., 2012). In some cases our observed stop sites lack sequential purines, but a closer look at the three-dimensional crystal structure reveals purines in close proximity that could represent long-distance crosslink partners (Figure 2.8B). We observe several potential long-distance intrastrand crosslinks through three-dimensional space. The first, identified as a stop on U45 suggesting platination at A44 (Figure 2.5), is likely a sequentially distant intrastrand crosslink, and the only guanine N7 within reach of A44 is G59, on the complementary strand of the P4 stem (Figure 2.8B). G30 is likely a non-phosphorothioate directed interstrand crosslink between the ES and

SS strand to either G-6 or A-8 on the substrate strand, because in the crystal structure there are no other nearby nucleophiles (Figure 2.8B). G39 has sequential purine residues and could be a 1,2 adduct to G38, a 1,3 adduct to G37, or a sequentially distant intrastrand crosslink to either G60 or G59, since these purines are close in space to the platinated nucleotide. A mix of stop sites was observed in the P4 region GAAA tetraloop (Figures 5, 8B). This region is very purine rich and results in a complicated RT mapping profile, suggesting platinum likely forms numerous 1,2 and higher-order crosslinks. The final stop site observed is U36 (Figures 5, 8A). This is not likely a crosslink, as uracil does not have a nucleophilic imine and hasn't been observed as a platination site. Rather, it appears to be a stop caused by the 5' end of our substrate strand still annealed to the template enzyme strand.

Conclusions

Toward the goal of developing platinum(II)-derived compounds as RNA crosslinking reagents, we have investigated Pt-derived crosslinking in the HDV ribozyme, a highly structured RNA. We explored phosphorothioate-directed crosslinking using a substrate strand containing active-site phosphorothioate substitutions, and using two Pt(II) crosslinking complexes. Pt-directed adducts throughout the HDV were also identified using RT extension. The results are compared to crystallographic models of the HDV. As a divalent metal ion, we predict that aquated cisplatin will be drawn toward metal ion binding sites. In these experiments, we observe the majority of our platination sites correlate with locations where magnesium ions were co-crystallized with the HDV ribozyme (Figure 2.8A, (Chen et al., 2010)). This trait is useful because divalent metal

ions often mediate RNA interactions that form interesting tertiary structures (Ward et al., 2014).

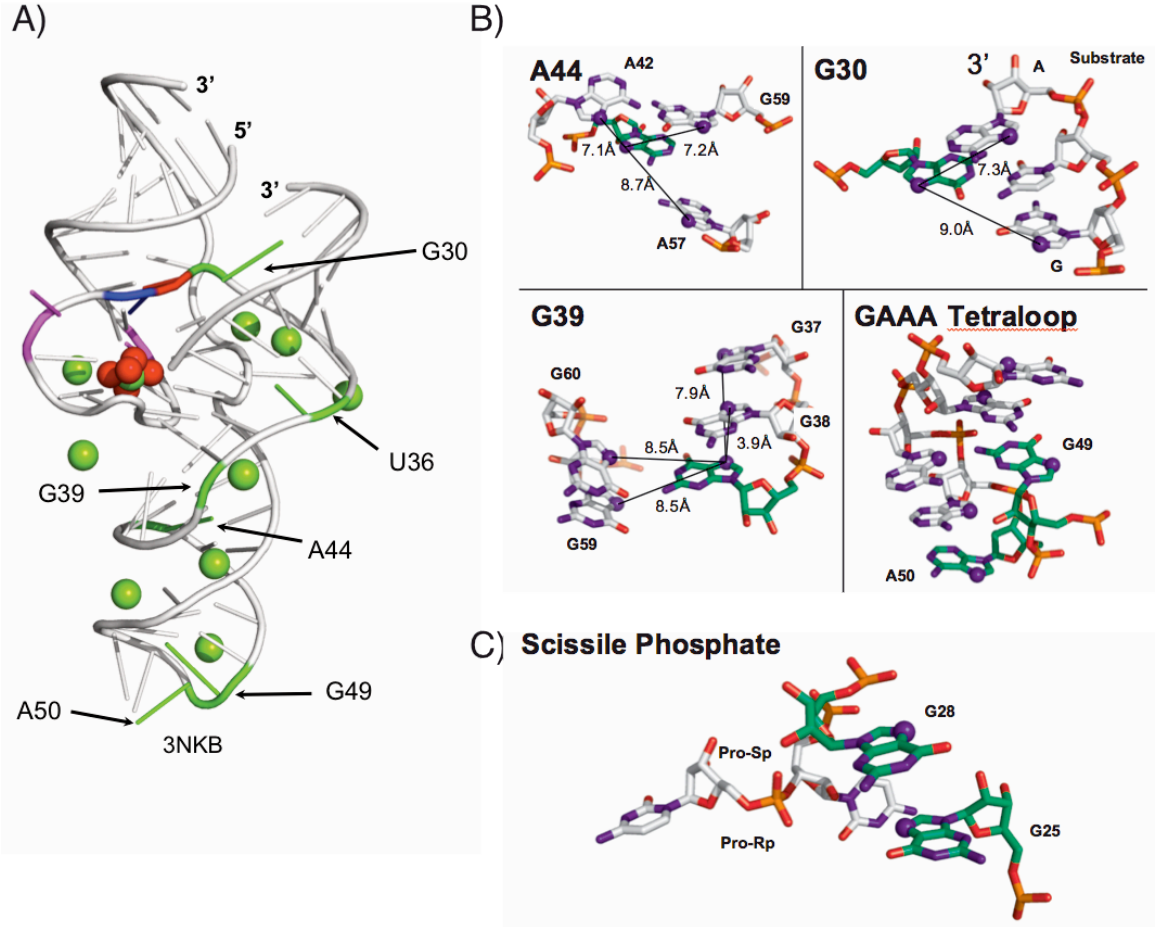


Figure 2.8. Potential cisplatin crosslink partners. A) Tertiary structure of the HDV ribozyme with observed RT stop sites indicated in green. B) Relevant platination sites as discerned by RT experiments are shown. Nucleotides observed to be platinated are in green, purine residues within 10 Å are in white. N7 positions are shown as spheres, and measured distances are presented. C) G28 was an observed crosslinking partner to the phosphorothioate-substituted scissile phosphate, its orientation with respect to the modeled in hammerhead ribozyme cleavage site is depicted here, next to the expected crosslinking partner G25. Measurements from the pro-Rp and pro-Sp positions are presented in Figure A.6. Figures derived from the HDV ribozyme crystal structure (PDB ID 3NKB, Chen et al., 2010) and the hammerhead ribozyme crystal structure (PDB ID 2OEU, Martick et al., 2008).

Crosslinks observed are consistent with known reactivity of platinum(II) compounds and fit crystal structures of the HDV ribozyme. This work demonstrates the

utility of platinum(II) compounds as crosslinking reagents. Platinum(II) compounds are interesting multifunctional crosslinking tools, capable of both directed recruitment to synthetically added nucleophilic positions as well as nonspecific recruitment to endogenous nucleophiles, allowing specific crosslinks (such as PS-G28, Figure 2.6A) and general crosslinks (G39-nearby nucleophiles, Figure 2.8B) to be resolved in a single experiment. Furthermore, as a divalent metal ion, cisplatin-derived crosslinking reagents are well-poised for interacting with interesting tertiary structures in RNA. The demonstrated reactivity of our click-functionalized crosslinker 2-ADAP Pt shows promise for adding the ability to purify using the same crosslinking tool, making functionalized platinum reagents a unique class of all-in-one affinity tagging, specific, and general crosslinking reagents. Given that cisplatin and 2-ADAP Pt crosslink very similarly (Figures 4 and 6), we believe that 2-ADAP Pt shows promise as a functionalized cisplatin mimic which will be useful for studying cisplatin's reactivity both in-vitro and in-vivo.

Summary of Chapter II and bridge to Chapter III

In this chapter, we describe the use of platinum complexes for studying metal ion coordination and RNA tertiary structure in the hepatitis delta virus (HDV) ribozyme. Crosslinked complexes were isolated and mapped using primer extension and hydrolysis, and compared to a current crystal structure. We show that cisplatin crosslinking in the HDV ribozyme reaches efficiency as high as ~40%, and demonstrate that cisplatin and a functionalized cisplatin derivative, 2-azido 1,3 diaminopropane Pt(II) (2-ADAP Pt), result in similar crosslinks and report structurally relevant information. Chapter II

introduces a novel high-throughput sequencing method to map all sites of Pt(II) small molecule binding in cellular RNA.

CHAPTER III

PT-SEQ: HIGH-THROUGHPUT MAPPING OF SMALL-MOLECULE PLATINUM ADDUCTS ON CELLULAR RNA

This chapter contains contributions from Victoria DeRose in the form of advice and direction throughout the project, as well as editing of the final manuscript. I performed and analyzed the experiments and wrote the text.

Introduction

Noncoding RNA makes up a significant portion of the eukaryotic genome (Mattick, 2001). Work has attributed regulatory roles to much of this noncoding RNA (Mattick and Makunin, 2006), highlighting RNA's importance in the functional landscape of cells (Sharp, 2009). Recent work has focused on the potential for RNA targeting small molecules as potential drug candidates (Thomas and Hergenrother, 2008). Cisplatin (Figure 3.1A), one of the most widely use anticancer drugs and the focus of this work, is a known RNA binder (Rijal and Chow, 2008; Hostetter et al., 2012; Dedduwa-Mudalige et al., 2015). Many RNA-based processes are disrupted by Pt compounds (reviewed in Chapman et al., 2011), but there is little information on global accumulation of Pt in cells.

Cisplatin is a widely prescribed anticancer drug which forms Pt(II)-coordinate bonds with a wide range of biomolecules (Harper et al., 2010). First approved by the FDA in 1978 (Sherman and Lippard, 1987), cisplatin and related Pt therapeutics are used in 40-50% of all chemotherapy treatments (Dyson and Sava, 2006; Harper et al., 2010).

Cisplatin is administered as an uncharged dichloro salt, cis-$Pt(NH_3)_2Cl_2$ which is favored in the high chloride concentrations of the bloodstream. When cisplatin enters the cell, chloride ligands are exchanged for water resulting in positively charged mono- and di-aquated species. It is generally accepted that a major target for cisplatin is cellular DNA, primarily the nucleophilic N7 position of purine residues, and DNA-Pt adducts induce downstream apoptosis (Siddik, 2003).

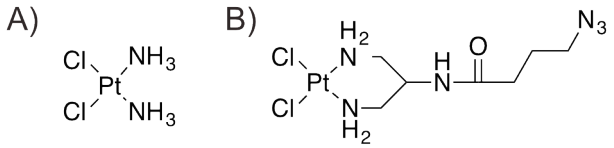


Figure 3.1. Platinum compounds used in this work. A) Cisplatin. B) Azaplatin (Wirth et al., 2015).

Much work has supported the hypothesis that cellular RNA is also a major target of cisplatin (Rijal and Chow 2008; Hostetter et al., 2012; Dedduwa-Mudalige et al., 2015). Research has suggested that Pt adducts from cisplatin treatment accumulate to a greater degree on RNA than DNA (Akaboshi and Miyahara, 1992; Hostetter et al., 2012), and that most Pt-RNA accumulation is on ribosomal RNA (Hostetter et al., 2012). Ribosomes present an attractive target for cisplatin as they are large RNA structures, crucial for cellular reproduction, and unlike DNA, present in the cytoplasm. Pt adduct accumulation on ribosomal RNA is particularly interesting because the ribosome is required for protein synthesis and is, in fact, a ribozyme (Moore and Steitz, 2003). Toxic agents such as sarcin and ricin act on the ribosome (Tumer and Li, 2012), as well as numerous ribosome targeting antibiotics, demonstrating its susceptibility to small molecule modification (Garreau de Loubresse, 2014). Platinum adducts from cisplatin

treatment have been observed on ribosomal RNA from *E. coli* and *S. cerevisiae* and have been mapped to single-nucleotide resolution using targeted primer extension (Rijal and Chow 2009; Hostetter et al., 2012; Osborn et al., 2014; Dedduwa-Mudalige and Chow, 2015). Recent crystallography work has also identified platinum binding sites on purified *T. thermophilus* ribosomes complexed *in vitro* with mRNA and tRNA, then treated briefly with cisplatin (Melnikov et al., 2016). This work provides an interesting molecular view of potential adduct types, but its relation to adducts found from in-cell treatment is not established.

A comprehensive picture of platinum accumulation on cellular RNA resulting from drug treatment is still missing. Current methods of mapping platinum adducts on RNA are largely confined to reverse transcription primer extension analysis (Rijal and Chow, 2009; Hostetter et al., 2012; Osborn et al., 2014) or mass spectrometry (Dedduwa-Mudalige and Chow, 2015). While well-represented in the literature, gel-based primer extension is time consuming, limited in scope, and poorly suited for high-throughput analysis.

Recent work has combined primer extension principles with high-throughput sequencing library generation in order to identify RNA modifications on a genome-wide scale. Notable examples are StructureFold and Mod-seq, in which cellular RNA is treated with DMS to methylate cytosine and adenine residues (Talkish et al., 2014; Tang et al., 2015), and methods based on SHAPE (selective 2'-hydroxyl acylation followed by primer extension) chemistry, which uses a library of alkylating reagents to modify the 2' hydroxide of RNA (Incarnato et al., 2014; Spitale et al., 2015). Both approaches initially used gel-based primer extension but have since been adopted for high-throughput

sequencing for comprehensive analysis of all reactive sites in cellular RNA. Toward a similar goal, we have developed Pt-seq, which harnesses the power of high-throughput sequencing to map all platinum adducts on cellular RNA in a single sequencing experiment (Figure 3.2).

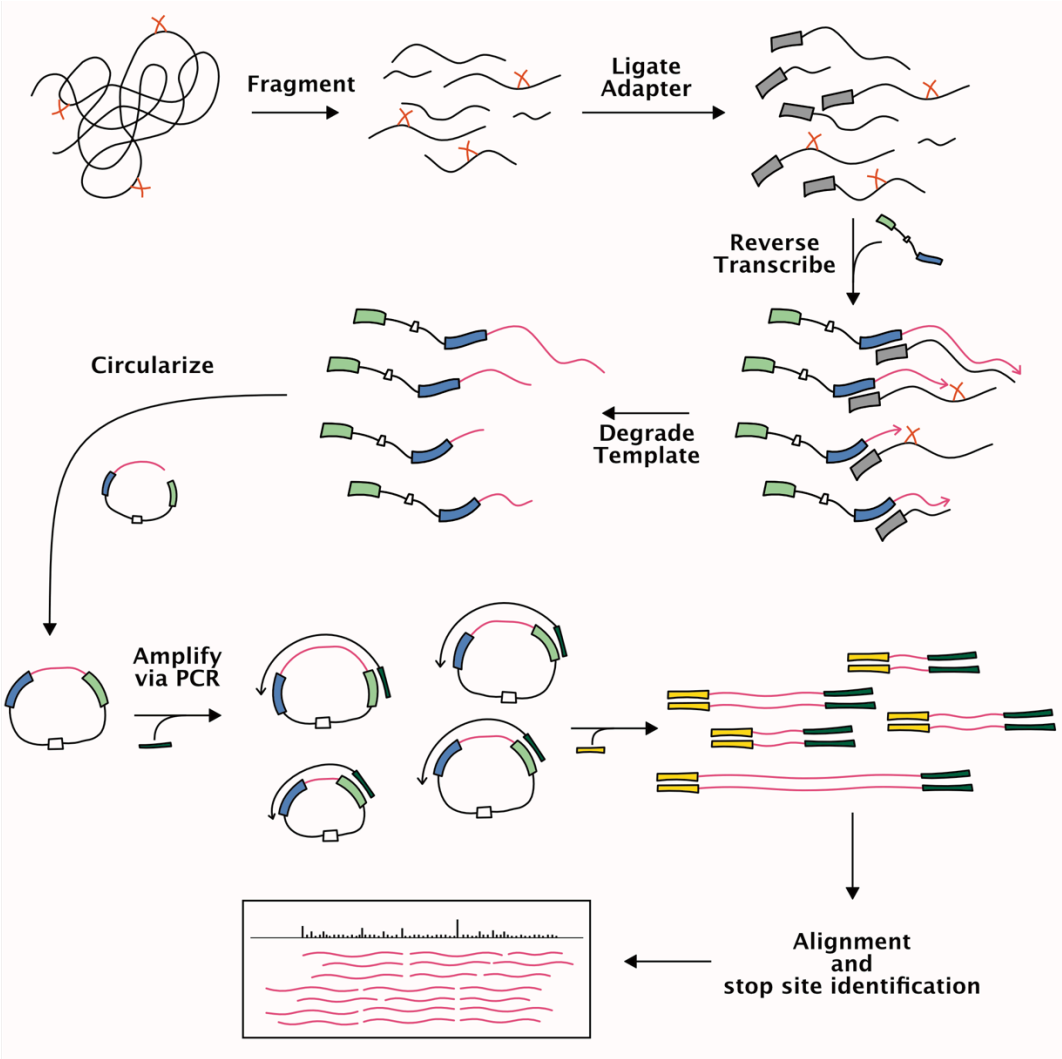


Figure 3.2. Pt-seq workflow. Pt-seq uses a modified Mod-seq workflow (Talkish et al.) to identify sites of platinum adducts on cellular RNA following cisplatin treatment.

Results and Discussion

To establish this method, yeast cultures were treated with cisplatin and a novel azide-modified compound azaplatin (Figure 3.1A and B). *S. cerevisiae* was chosen for this initial study based on our prior work using *S. cerevisiae* to explore cisplatin accumulation on cellular RNA (Hostetter et al., 2012). Yeast share many conserved cellular processes with mammals and like all eukaryotic organisms express large amounts of noncoding RNAs (Wu et al., 2012). Numerous cisplatin adducts have already been mapped on *S. cerevisiae* rRNA using reverse transcription primer extension analysis (Hostetter et al., 2012; Osborn et al., 2014). Because of this, *S. cerevisiae* are a good model system for developing our high-throughput platinum mapping technique.

The $[Pt(NH_3)_2]$-RNA adduct formed by cisplatin treatment is known to cause stalling of processive enzymes such as reverse transcriptases (Rijal and Chow, 2009). Current methods of mapping cisplatin adducts that are based on extending radiolabeled RT primers rely on bands of increased intensity relative to background under treatment conditions on sequencing gels. While effective, these techniques are poorly suited for high-throughput analysis of large structured RNAs as the resolvable window is limited to approximately 100 nucleotides per experiment and each primer needs to be individually optimized and labeled. Pt-seq is based on work by Talkish et al. (2014) and adapts RT primer extension for high-throughput sequencing (Figure 3.2). Briefly, RNA is extracted from treated and untreated cells, fragmented, and an adapter annealed to its 3' end. This adapter serves as a priming site for the Mod-seq primer designed by Ingola et al. (2012). This primer is extended using reverse transcriptase, the resulting cDNA fragments are size-selected and 5' phosphorylated, and then the construct is circularized using

Circligase (Epicentre). After circularization, PCR is performed using indexed primers containing Illumina adapters to generate sequencing libraries. After sequencing, reads are adapter trimmed, aligned to the yeast genome, and fed into the Mod-seq pipeline (Talkish et al., 2014) to identify platinum adduct sites (Figure 3.3).

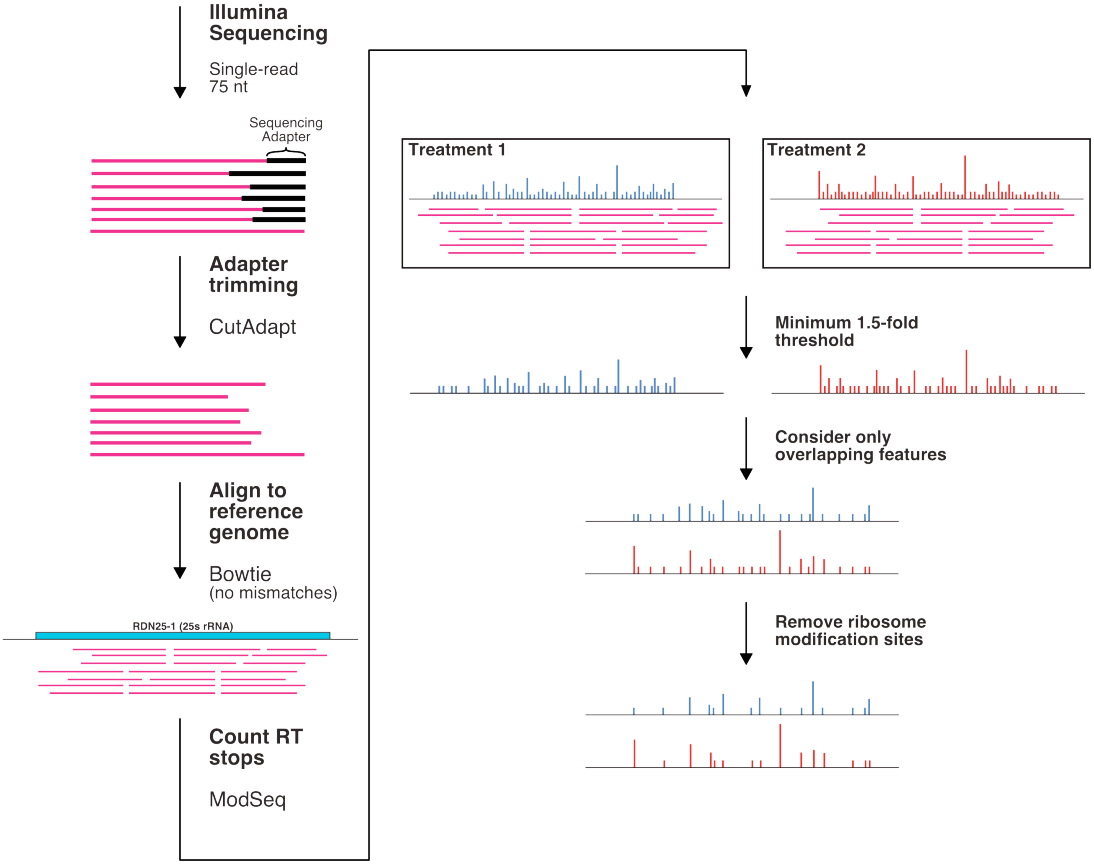


Figure 3.3 Post-sequencing data analysis workflow. Reads are adapter trimmed using CutAdapt (Martin et al., 2011), aligned to a reference genome using Bowtie (Langmead et al., 2009), and each treatment is fed separately into the Mod-seq pipeline with appropriate control reads to generate stop sites (Talkish et al., 2014). From these, we set a minimum 1.5-fold enrichment over background, merge the two datasets to consider only overlapping features, and remove modified nucleotides, if applicable.

A)

Treatment	Reads	No mismatches	
		Aligned reads	Percent Aligned
Control-1	4,113,364	2,287,036	56
Control-2	7,218,695	4,432,691	61
Cis100-1	4,851,790	3,008,974	62
Cis100-2	7,294,189	4,499,760	62
Cis200-1	5,925,462	3,292,636	56
Cis200-2	3,843,239	1,901,225	49
Rg100-1	5,152,589	3,811,213	74
Rg100-2	7,164,539	4,988,488	70
Rg200-1	5,549,454	4,106,606	74
Rg200-2	6,598,879	4,864,861	74

TOTAL READS 57,712,200

B)

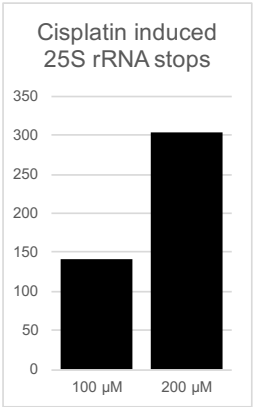
Cisplatin

	Count	Min	Median	Max
100 µM	141	1.5	1.8	13.9
200 µM	304	1.5	1.9	17.4

Azaplatin

	Count	Min	Median	Max
100 µM	835	1.5	2	30.8
200 µM	1188	1.5	2.2	114

C)



D)

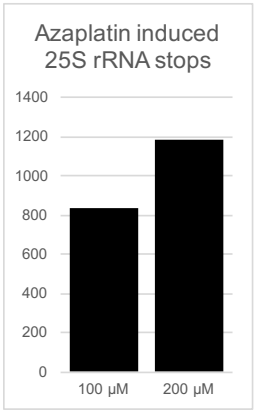


Table 1.1. Pt-seq sequencing results and 25S rRNA Pt-RNA adduct statistics. A) Number of total and aligned reads for each sequenced treatment. B) Count of significant stops with greater than 1.5-fold increased threshold for 100 µM and 200 µM cisplatin and azaplatin treatment. Median and max stop intensities shown. C) Total adduct counts on the yeast large ribosomal subunit for each cisplatin treatment. Doubling cisplatin treatment concentrations doubles the number of observed platinum adducts. D) Total adduct counts on the yeast large ribosomal subunit for each azaplatin treatment.

For this study, five treatments were chosen: mock treatment control, 100 µM cisplatin, 200 µM cisplatin, 100 µM azaplatin, and 200 µM azaplatin. Two biological replicates were performed for each condition. Approximately 58 million reads were obtained across the five treatments. Reads per condition vary from 3.8-7.3 million (Table 1.1). Between 50 – 74% of reads aligned to the yeast genome with no mismatches, and

80-90% align with one mismatch. Estimated distribution across genes was generated using HTSeq-count (Anders et al., 2005) which counts reads within gene features. Read distribution was predominantly on ribosomal RNA, as expected (Tables B.1-B.4).

We chose to use only reads with no mismatches to explore platinum accumulation on ribosomal RNA. We fed our trimmed and aligned reads (see Methods) into the Mod-seq script by Talkish et al. (2014), to identify sites of RT stalling, setting a minumum fold-enrichment threshold of 1.5, and compared cisplatin and azaplatin treatments vs. control to derive the location of platinum adducts to single-nucleotide resolution. Stops were seen on numerous cellular RNA's, including tRNA's, rRNA's, snoRNA's, and cytoplasmic RNA's (Table B.3 through B.6). We chose to focus on the ribosomal component of the large ribosomal subunit, 25S rRNA. Pt-seq results demonstrate a treatment effect by cisplatin. For cisplatin, as drug concentration is increased 2-fold, the number of sites demonstrating >1.5-fold stop enrichment increase 2.2-fold for 25S rRNA (Table 1.1 C). Median and maximum stop intensity also increases as drug concentration increases (Table 1.1 B). Cisplatin-induced stops are scattered across the 25S rRNA (Figure 3.4). Higher treatment concentration results in some overlapping and many new stop sites. Interestingly, for overlapping sites, increasing cisplatin treatment concentration from 100 to 200 μM did not always result in an increase in relative stop intensity, and occasionally higher concentrations of cisplatin result in decreased relative stop intensity (Figure 3.4A and B). Azaplatin treatment likewise exhibits a dose-dependent effect, though not as pronounced (Table 1.1B, D).

We were interested in whether the predicted platinum adducts exhibit preference toward solvent accessible platinum binging sites. Cisplatin is known to form adducts with

the N7 position of purine residues as well as the N3 position of cytosine (Sherman and Lippard 1987). Using CCP4's AreaImol software (Winn et al., 2011) we calculated the solvent accessible surface area across all ribosomal RNA atoms in the yeast 80S ribosome (25S, 18S, 5S, and 5.8S rRNA, PDB 4V7R) to a cisplatin-sized sphere with a radius of 2.2 Å (Melnikov et al., 2016), then plotted the natural log of solvent accessibility of each N7 and cytosine N3 against the location and intensity of each cisplatin-induced stop at both treatment concentrations (Figure 3.5). We found, surprisingly, that though the mean solvent accessibility was 3.4 $Å^2$ for 100 μM stops and 3.6 $Å^2$ for 200 μM cisplatin stops, we see a number of stops mapped to positions with 0 $Å^2$ calculated solvent accessibility, and no strong correlation between stop location or intensity and solvent accessibility. This is consistent with recent results from x-ray crystallography experiments by Melnikov et al. (2016) who found that Pt adducts formed on *T. thermophilus* 70s ribosomes following *in vitro* cisplatin treatment are not solely predicted by solvent accessibility.

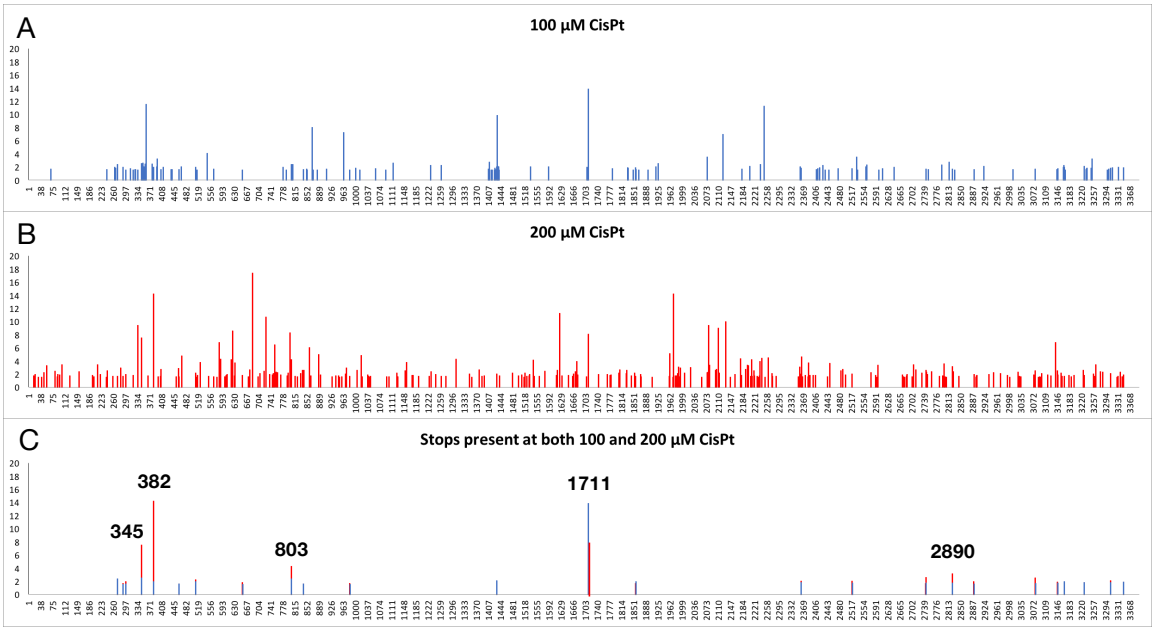


Figure 3.4. Cisplatin treatment effect on yeast large ribosomal subunit (25S). A) 100 µM cisplatin adducts. B) 200 µM cisplatin adducts. Increasing cisplatin concentration results in increased stop site observation. C) Stops present in both treatment concentrations, with a cutoff of 1.5-fold enrichment as described in text.

A)

B)

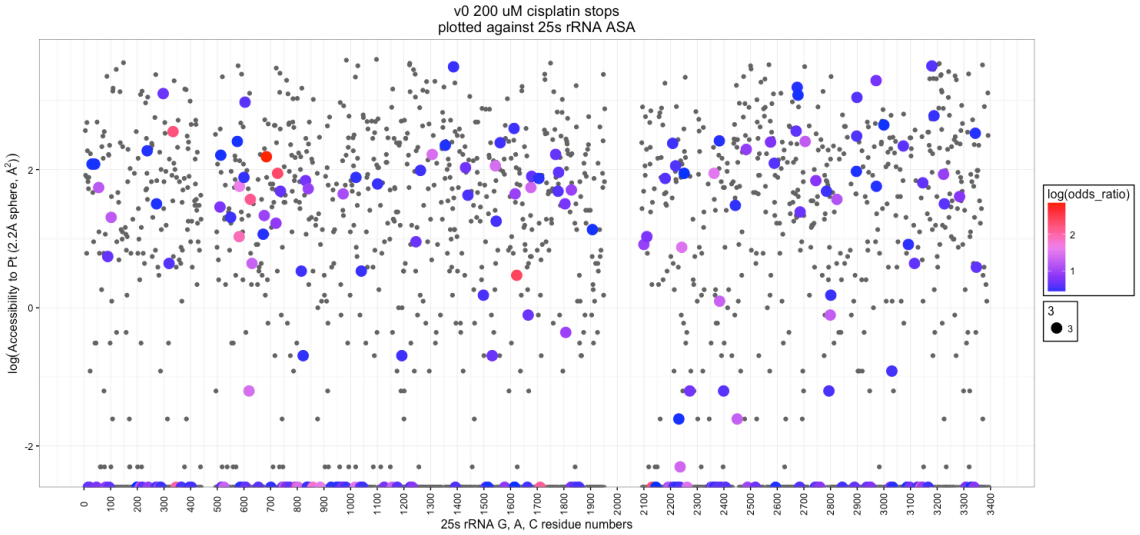


Figure 3.5. Calculated solvent accessible surface area for all 25S ribosomal subunit N7 and cytosine N3 positions in context of the 25S ribosome structure to a platinum-sized sphere of radius 2.2 Å. A) 100 µM cisplatin. B) 200 µM cisplatin. Platinum binding positions across the 25S sequence are plotted against the natural log of each solvent-accessible surface area, as calculated using CCP4's Area Imol package (Winn et al., 2011) (gray dots). Large dots are platinum binding sites as determined using our Pt-seq approach and are colored according to the natural log of their fold-enrichment as reported by the Mod-seq script (Talkish et al., 2014). Each treatment exhibits a number of binding sites in regions of no calculated solvent accessibility, represented as large half-dots along the x-axis. Gaps in coverage are due to missing nucleotides in the crystal structure. Only stops with greater than 1.5-fold increase over control are considered.

To narrow our focus for closer analysis we initially looked at the 25S large ribosomal subunit, compared stops across 100 and 200 μM cisplatin treatment conditions that are greater than 1.5-fold enriched over background, and then considered only those which were present in both treatments. Stops which were present in both conditions were then compared with a list of known modified nucleotides (Piekna-Przybylska et al., 2007) and any matches removed to ensure no overlap (Figure 3.6). Out of the 141 and 304 stop sites observed at 100 and 200 μM cisplatin, 26 are shared by both treatments (Figure 3.6). The median stop intensity for these paired stops was 1.8 at 100 μM cisplatin and 1.95 at 200 μM cisplatin, further exhibiting a treatment effect.

For this subset of 26 sites, the five sites with highest fold-enrichment are at positions G345, U382, C803, C1171, and G2824 (Figure 3.6). C1711 is located on helix 58, which is the binding site for ribosomal protein L30e, an indispensable ribosomal protein. Solvent accessible surface area calculations suggest that the C1711 N3 position is inaccessible to cisplatin, but that as a whole helix 58 is highly exposed (Figure 3.5, residues 1698-1748). Helix 58's interaction with protein L30e primarily relies on the unusual tertiary structure of helix 58, a kink-turn motif (Jenner et al., 2012; Bifano et al., 2013). C1711 is located just before this motif. Together, H58 and L30e form an intersubunit bridge, one of several contacts between the large and small ribosomal subunits (Halic et al., 2005; Jenner et al., 2012).

Figure 3.6. Stop fold-enrichment for cisplatin induced stops on the yeast large ribosomal subunit (25S rRNA). Shown are stops shared between the top-50 strongest stop sites after 200 µM cisplatin treatment and all stops observed in 100 µM cisplatin. Asterisks mark stops discussed further in the text.

G2824 is located at the base of helix 89, in the "accommodation corridor," a passageway located along the interface between H89 and the H90-92 structure of the large ribosomal subunit (Petrov et al., 2008). It is immediately adjacent to the peptidyl transfer center (Rakauskaite and Dinman, 2008), and close to the tRNA A-site. Interestingly, it is close to the binding sites of anisomycin and narciclasine, eukaryotic ribosome-targeting antibiotics that inhibit peptide bond formation (Garreau de Loubresse et al., 2014). Furthermore, mutations to nearby residues (2820, A2829) result in non-viable ribosomes (Rakauskaitė and Dinman, 2008), suggesting this is a region

particularly sensitive to modification. This implies a previously unexplored model for cisplatin inhibition of the ribosome. G345 and U382 are located near helix 23 and 24, respectively. U382 is highly solvent exposed, located near the surface of the large ribosomal subunit opposite the small subunit (Figure 3.10). G345 is near the surface of the large ribosomal subunit but behind the 5.8S ribosomal RNA (Figure 3.10).

In addition to cisplatin, we have also applied Pt-seq to map RNA accumulation of azaplatin, a novel click-functionalized platinum compound developed in our lab (Wirth et al., 2015). Pt-seq suggests that azaplatin either binds to more sites than cisplatin or that it causes a stronger stalling effect, as evidenced by increased stop site locations and higher intensities (Figure B.1 as compared with Figure 4). This could be due to structural differences, as azaplatin is a significantly bulkier molecule than cisplatin (Figure 3.1). As with cisplatin, we have plotted stop intensity vs solvent accessibility of all azaplatin binding sites and find no correlation between calculated solvent accessibility of Pt(II) binding sites and stop intensity (Figure 3.7).

Azaplatin induces 835 stops at 100 μM treatment and 1188 stops at 200 μM treatment with greater than 1.5-fold enrichment over control (Figure 3.4). Of these, 513 are shared stops between the two treatment levels (Figure B.1). We decided to consider only the top 50 strongest stops at 200 μM azaplatin treatment, then matched those with 100 μM azaplatin treatment to narrow down a smaller subset to investigate further. This results in 16 shared stop sites (Figure 3.8). These stops are more intense than similarly selected cisplatin stops (Figure 3.6), and only one site is shared among the two datasets (C1711, Figures 6 and 8).

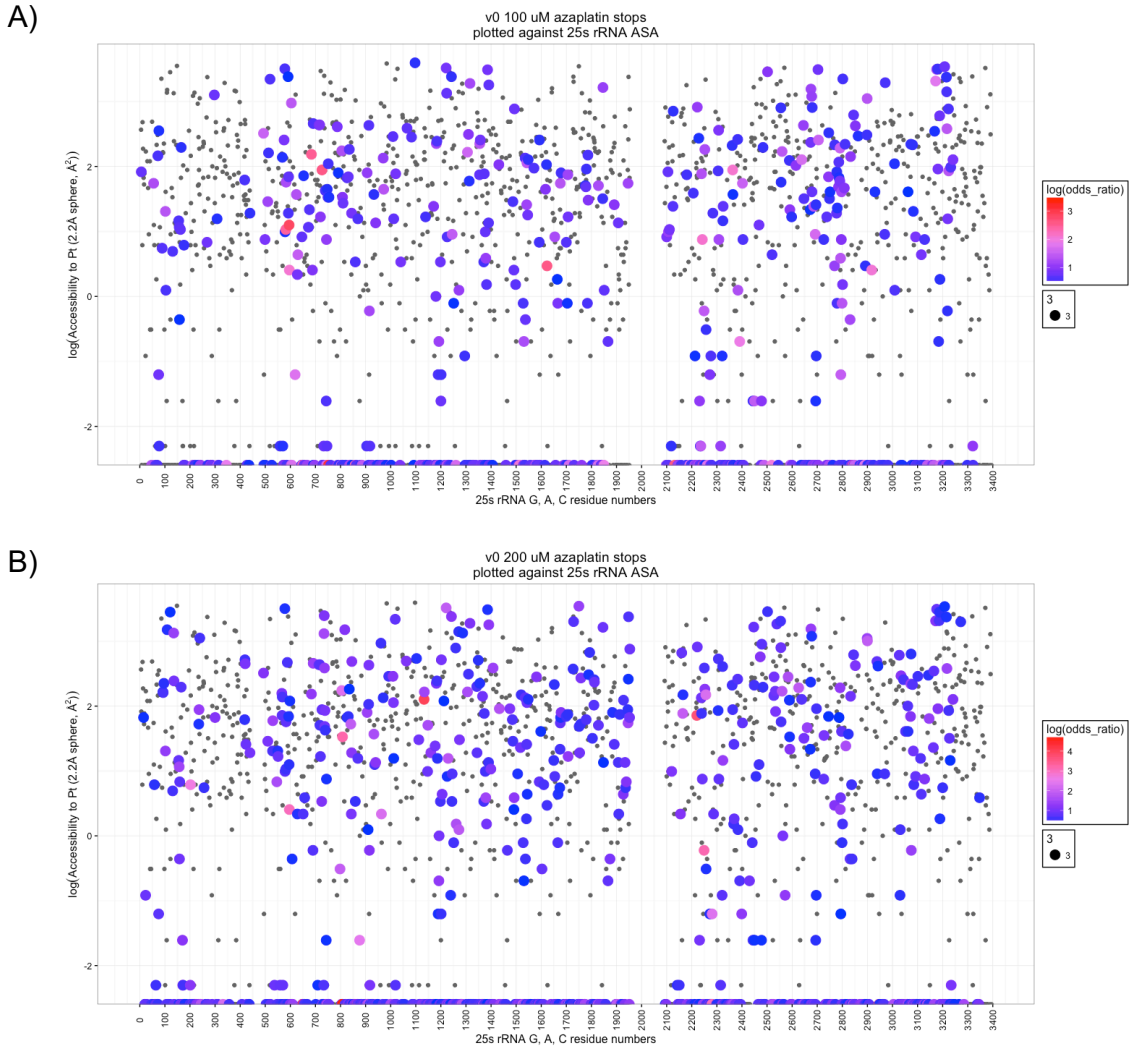


Figure 3.7. Calculated solvent accessible surface area for all 25S ribosomal subunit N7 and cytosine N3 positions in context of the 25S ribosome structure to a platinum-sized sphere of radius 2.2 Å. A) 100 µM azaplatin. B) 200 µM azaplatin. Platinum binding positions across the 25S sequence are plotted against the natural log of each solvent-accessible surface area, as calculated using CCP4's Area Imol package (Winn et al., 2011) (gray dots). Large dots are platinum binding sites as determined using our Pt-seq approach and are colored according to the natural log of their fold-enrichment as reported by the Mod-seq script (Talkish et al., 2014). Each treatment exhibits a number of binding sites in regions of no calculated solvent accessibility, represented as large half-dots along the X axis. Gaps in coverage are due to missing nucleotides in the crystal structure. Only stops with greater than 1.5-fold increase over control are considered.

Three azaplatin stops, at G805, A807, and A808, are close to an observed cisplatin stop (C803, Figure 3.6). Each of these sites is located on helix H32, an internal loop which supports the PTC wall. In a random mutagenesis study searching for structurally important residues in *E. coli* ribosomes, Yassin and Mankin found a cluster of deleterious mutations at this internal loop and suggest it plays an important structural role in the ribosome, suggesting it could be a target for ribosome inactivating antibiotic development (Yassin and Mankin, 2007). The azaplatin stop at position 596 is located on an expansion segment, the third arm of 25ES7c, and the stop at 1972 (in which 100 μM induced a stronger stop than 200 μM) is on expansion segment ES27a. Expansion segments are poorly understood domains found in eukaryotic ribosomes, located outside the conserved core of the ribosome and responsible for a majority of the increased size of eukaryotic ribosomes relative to prokaryotic ribosomes (Ramesh and Woolford, 2016). Deletion of ES7 in *S. cerevisiae* results in defects in ribosome biosynthesis and inhibits *S. cerevisiae* growth (Jeeninga et al., 1997; Ramesh and Woolford, 2016). Deletion of ES27 in *Tetrahymena* also prevents growth, demonstrating an important yet still poorly understood role (Sweeny et al., 1994). The azaplatin stop at G2249 is located at the base of helix H69, which itself is at the subunit interface and plays a key role in translation. This helix has been assessed for cisplatin binding in an *in vitro* modeled *E. coli* H69 by Dedduwa-Mudalige and Chow (2015) using a combination of mass spectrometry and RNAse T1 digestion, and cisplatin was found to accumulate to the equivalent G1906 (*E. coli* numbering) and its upstream neighbor G1907 (*E. coli* numbering).

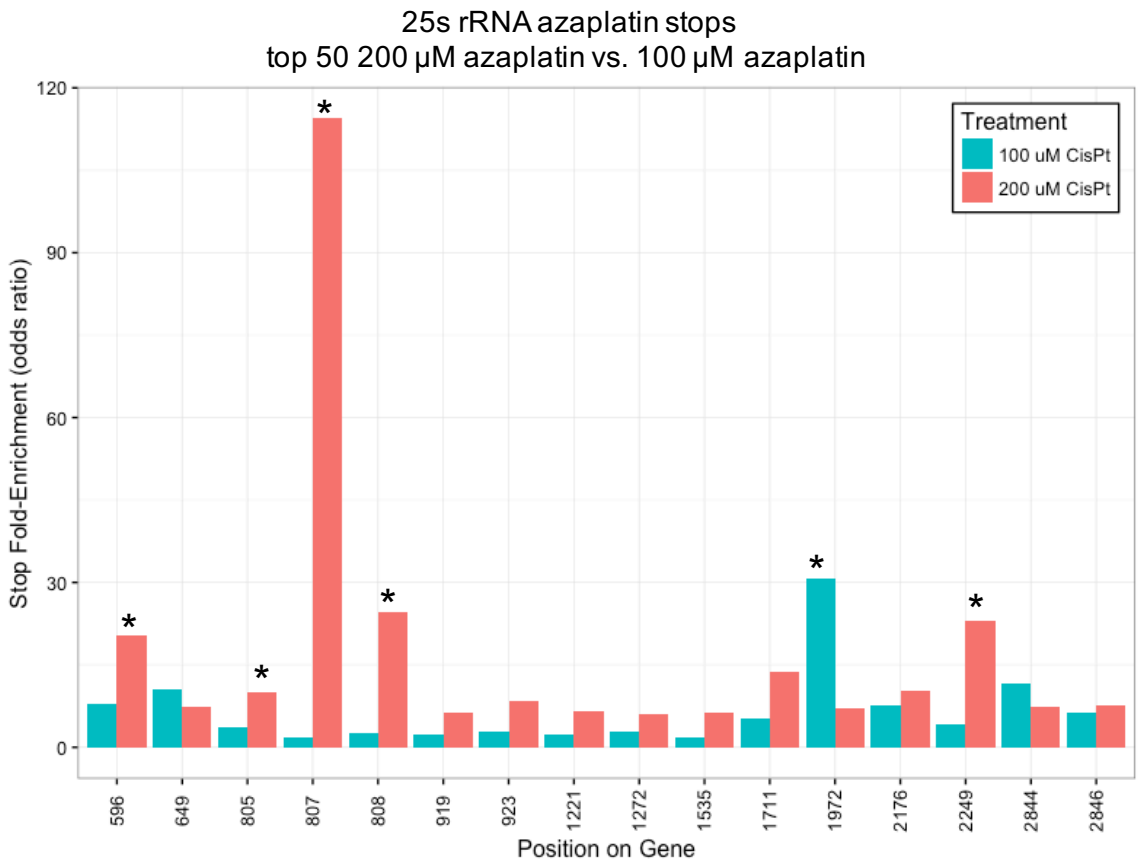


Figure 3.8. Stop fold-enrichment for azaplatin induced stops on the yeast large ribosomal subunit (25S rRNA). Shown are stops shared between the top-50 strongest stop sites after 200 µM azaplatin treatment and all stops observed in 100 µM azaplatin. Asterisks mark stops discussed further in the text.

We utilized reverse transcription primer extension analysis to verify our high-throughput sequencing approach. Yeast cultures were grown overnight and then treated with 200 µM cisplatin for six hours. RNA was isolated and reverse transcribed (see Methods) and results were visualized via autoradiography. We chose to analyze the region near C1711 since that was our strongest stop and helix 58 is highly solvent exposed, making it a good candidate for RT primer extension. Using a RT primer spanning 1761-1779 we were unable to verify a C1711 stop (Figure B.2), however a number of stops are observed upstream of C1711 in the region spanning G1599 through G1646 (Figure 3.9A).

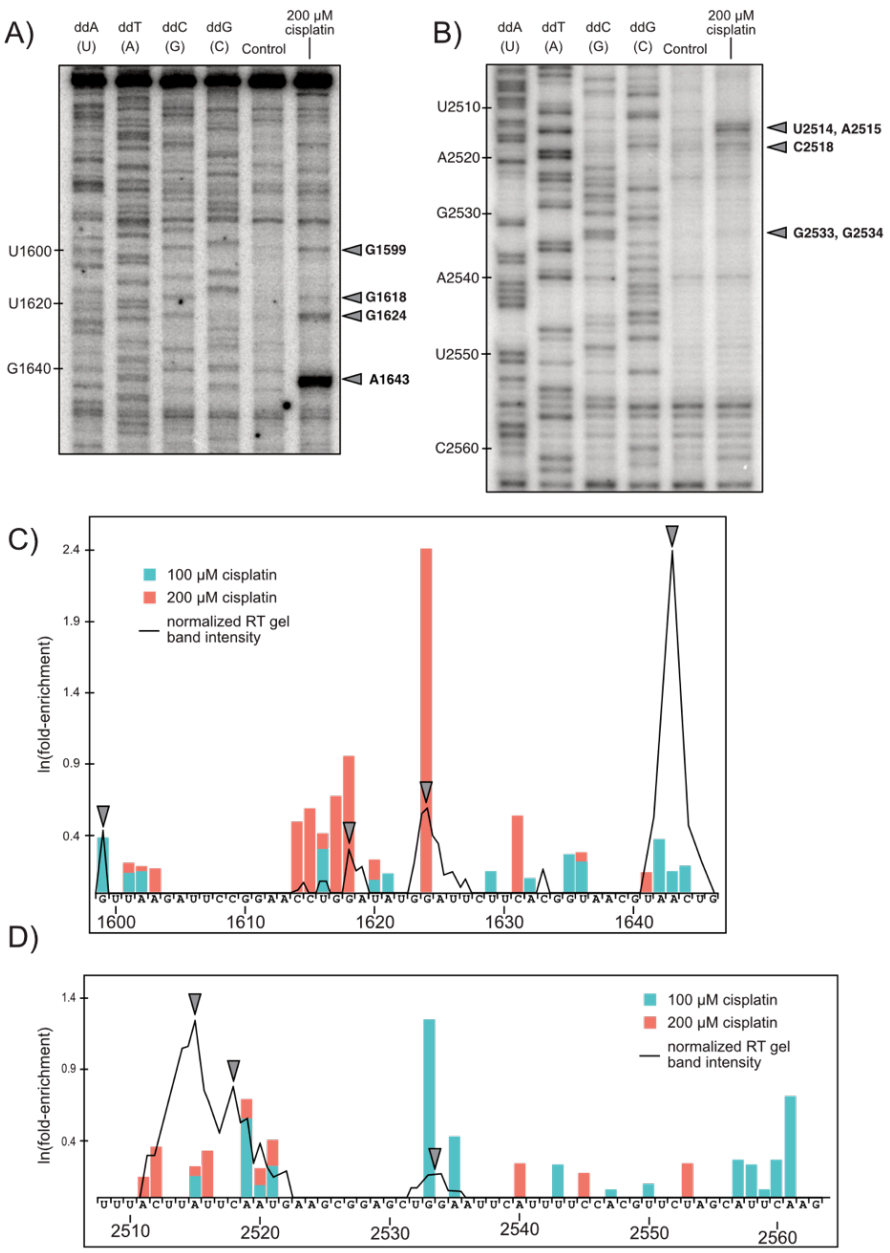


Figure 3.9. Comparison of gel-based reverse transcription primer extension analysis with results predicted by Pt-seq. Two regions of the yeast 25S large ribosomal subunit were subject to traditional gel-based primer extension analysis. A) positions 1600-1650 and B) positions 2510-2560 are shown. For both gels, lanes 1-4 are dideoxy sequencing ladders, lane 5 is 0 µM platinum control, and lane 6 is 200 µM cisplatin (6 hours treatment). Stop sites are marked with arrows. Band intensity down each lane was plotted using ImageJ and overlaid on a bar plot of Pt-seq results for both 100 and 200 µM cisplatin, shifted one nucleotide 3' to illustrate RT stop sites. C) 1600-1646 and D) 2508-2563 are shown.

We used ImageJ to generate an intensity plot of control and 200 μM cisplatin treatment lanes, then picked peaks (stop sites) unique to 200 μM cisplatin treatment and plotted these on a barplot of Pt-seq stop site intensities for both 100 and 200 μM cisplatin treatment (Figure 3.9C). The results suggest overlap of Pt-seq stops with those generated by traditional primer extension analysis, comparable to a similar analysis of DMS modification by (Tang et al., 2015). We see stops at G1599, G1618, G1624, and A1643 (Figure 3.8 A and C, gray arrows). The strongest stop observed on our RT gel (1643, 200 μM cisplatin) is represented by a cluster of Pt-seq 100 μM cisplatin stops but does not match Pt-seq 200 μM cisplatin stops.

A second primer intended to anneal near C1711 (1733-1753) instead annealed near helix 80 (3' end 2605) due to its high GC content and the highly single stranded nature of the region to which it bound. Repeated experiments verified that this primer reliably extended through U2508 and generated strong dideoxy sequencing ladders covering G2563 through U2510 (Fig. 3.9B). Under 200 μM cisplatin treatment this primer generates a strong stop at U2514 / A2515, a stop at C2518, and a weak stop at G2533/G2534 (Fig. 3.9B, D). As described, we quantified these lanes and plotted stop site intensities compared with sequenced stops and found a correlation between the two. Pt-seq determined stops from 2511-2521 reflect the broad set of stops from 2511-2522 observed via primer extension and suggest this region is particularly prone to modification.

Replicate experiments are important to the Platinum-seq approach. In the preceding analysis, biological replicates were combined to reduce experimental background (Talkish et al. 2014, Tang et al. 2015). In order to assess reproducibility of

our approach, we also performed analysis by treating each biological replicate as an individual experiment. Original alignments were fed into the Mod-Seq script (Talkish et al. 2014) and stop sites for each replicate were found by comparing against the same control to generate fold-enrichments by replicate. We isolated stops on just the 25S rRNA, then merged fold-enrichments for each replicate pair by position, and plotted pairs by stop intensity (replicate 2 vs. replicate 1 for each treatment, Figure B.4). These plots demonstrate a rough correlation between replicates for stop site position and intensity, especially with azaplatin treatment (R^2 values approximately 0.8 for both, figure B.4 C and D). Numerous stops were observed in one replicate but not the other (best demonstrated by Figure B.4 A, points at X=0 and Y=0). Since we are interested in "significant stops," which we define as stops present in two treatment concentrations above a certain threshold, we decided to remove stops from further analysis that were not present in both replicates. In total, 7% of the total stops with ≥1.5-fold enrichment observed in 100 μM cisplatin were reproduced in both replicates, 43% were reproduced by both 200 μM cisplatin replicates, 58% were reproduced by both 100 μM azaplatin replicates, and 64% were reproduced by both 200 μM azaplatin replicates. Considering only stops present in both replicates reveals a much stronger correlation between stop site position and strength by replicate, with R^2 values closer to 1 (Figure B.5). Natural-log transformation of this data further reveals this correlation (Figure B.6).

These analyses combined with primer extension results (Figure 3.9) provide early indications that our approach is reproducible for strong stop sites. C1711, for example, was the strongest significant stop for cisplatin treatment, and is reproduced across all replicates here (Table B.1). Interestingly, strong stops often show a clustering effect

across replicates (Table B.2). For example, the strong stop at 807 under azaplatin treatment, discussed above, is reproduced across all azaplatin replicates, as well as 200 μM cisplatin treatment and one of the two 100 μM cisplatin replicates. A tight cluster of highly replicated stops is seen from 804 through 808 in a purine-rich, single-stranded region (5'-G_{800}ACCCGAAAGA$_{810}$). This suggests there may be heavy accumulation of our platinum reagents to this region or a distribution of platinum adducts across this purine-rich stretch.

In this study, we have performed a high-throughput sequencing-based survey of Pt-RNA adducts formed following treatment of *S. cerevisiae* with cisplatin. 'Platinum-seq' identifies modification sites based on truncated cDNAs formed from random RT primer extension of RNA isolated from cells after treatment. A majority of predicted Pt-RNA sites were located on ribosomal RNA which comprises the vast majority of yeast RNA (>80% by nucleotide) (Warner 1999). Platinum-induced stops are also observed on tRNAs and cytoplasmic RNAs such as SCR1 as well as transposable elements (Tables B.1-B.4). Since most research on RNA-associated platinum adducts has been performed on ribosomal RNA, we decided to focus on the large ribosomal subunit (in yeast, 25S rRNA). Our high-throughput sequencing method found numerous platinum-induced RT stops on the 25S ribosomal subunit. We see a clear dose-dependent response as we increase cisplatin treatment concentration, with an increase in the number, median, and maximum stop intensity seen with doubling concentration of cisplatin (Table 1.1). Our Pt-seq results compare well with gel-based RT mapping methods and infer regions highly prone to modification (Figure 3.9).

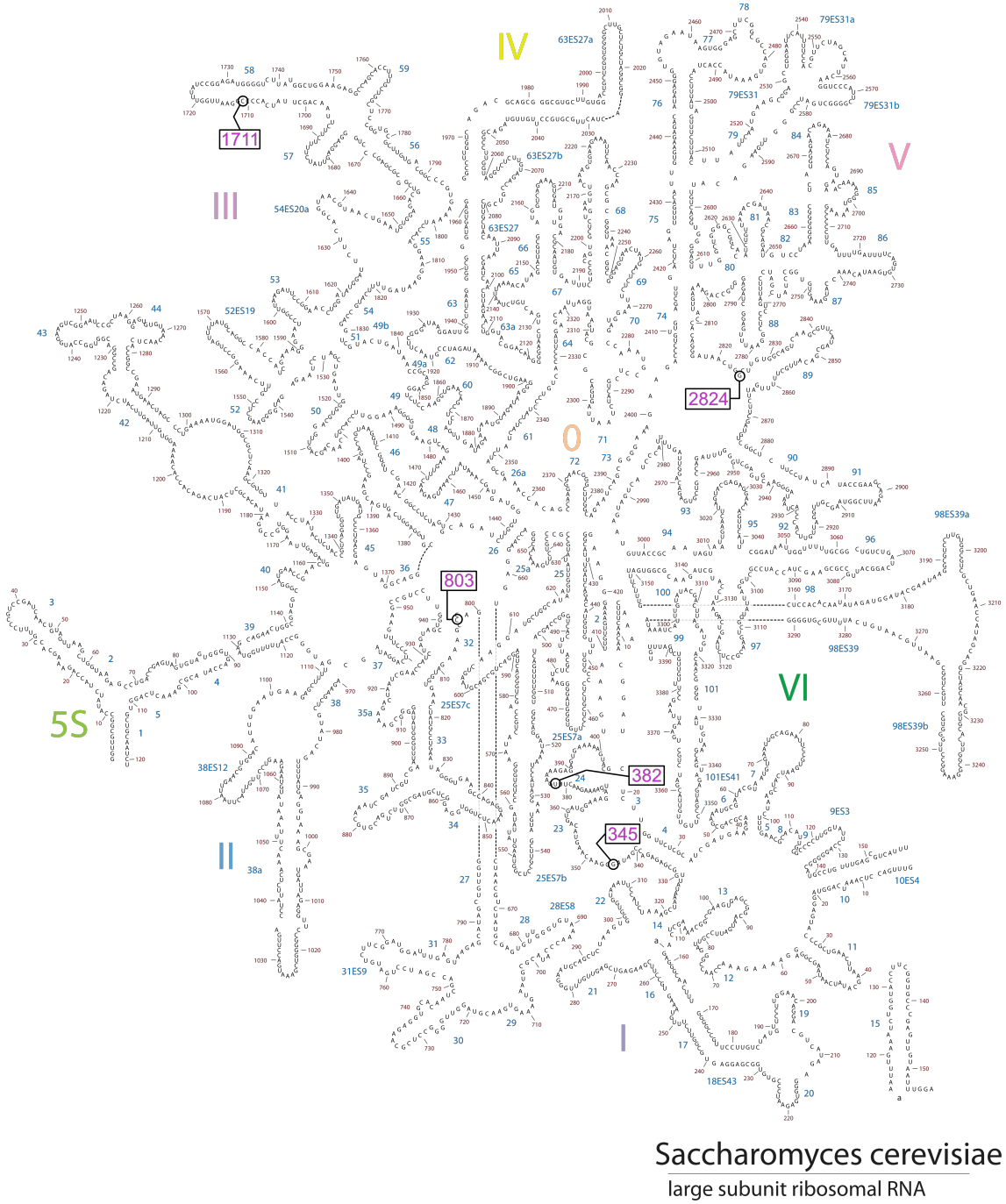


Figure 3.10. Conserved strong cisplatin-derived adducts present at 100 and 200 μM cisplatin mapped on a secondary structure of the large ribosomal subunit (Petrov et al., 2014), shown in pink. All strong stops except C1711 are located in the highly conserved core of the large ribosomal subunit.

Our strongest cisplatin-derived stops are located primarily in the highly conserved core of the ribosome (Figure 3.10). C1711 and U382 are located on the solvent-exposed surface of the 25S large ribosomal subunit (Figure 11A). G345 is also on the surface of the 25S but hidden behind the 5.8S ribosomal subunit (Figure 3.11B), close to the nascent peptide exit tunnel, the interior cavity from which newly synthesized peptides exit the ribosome. Adjusting the transparency of the 25S reveals the internal location of C803 and G2824 (Figure 3.11C). C803 is in a solvent-exposed cavity just off the exit tunnel, and is part of helix H32, which supports the PTC wall (Yassin and Mankin). G2824 is particularly interesting, as it is located in the core of the ribosome, near the peptidyl transferase center, where peptide bond formation occurs. Numerous eukaryotic-specific antibiotics bind in this region, just two nucleotides away (Anisomycin shown, Figure 3.12) and mutations to this region results in nonviable ribosomes, suggesting this region is particularly sensitive to modification. Cisplatin has previously been crystallized to regions near ribosome-targeting antibiotics (Melnikov et al., 2016). The authors suggest cisplatin may inhibit ribosomes similar to certain antibiotics. Our work suggests this is a subject worth investigation.

Our strongest azaplatin stops occur on more diverse regions. Two were on expansion segments ES7 and ES27 (C596 and A1972), which have previously shown to be important for ribosome function and biogenesis (Sweeney et al., 1994; Jeeninga et al., 1997). Three stops were located on H32 (G805, A807, and A808), near the cisplatin stop previously discussed. It is suggested this purine-rich internal loop plays an important structural role, as it supports the PTC wall, and single-nucleotide mutations to this region are lethal (Yassin and Mankin., 2007). We see reproducible stops across biological

replicates in this region (Figure B.6), as well as a tight clustering effect to nucleotides from 804-808, lending more evidence this region may be heavily modified by platinum reagents (Table B.1). Finally, G2249 is at the base of an important loop (H69), and a cisplatin lesion has previously been mapped to this location in *in vitro* modeled *E. coli* H69 (Dedduwa-Mudalige and Chow, 2015).

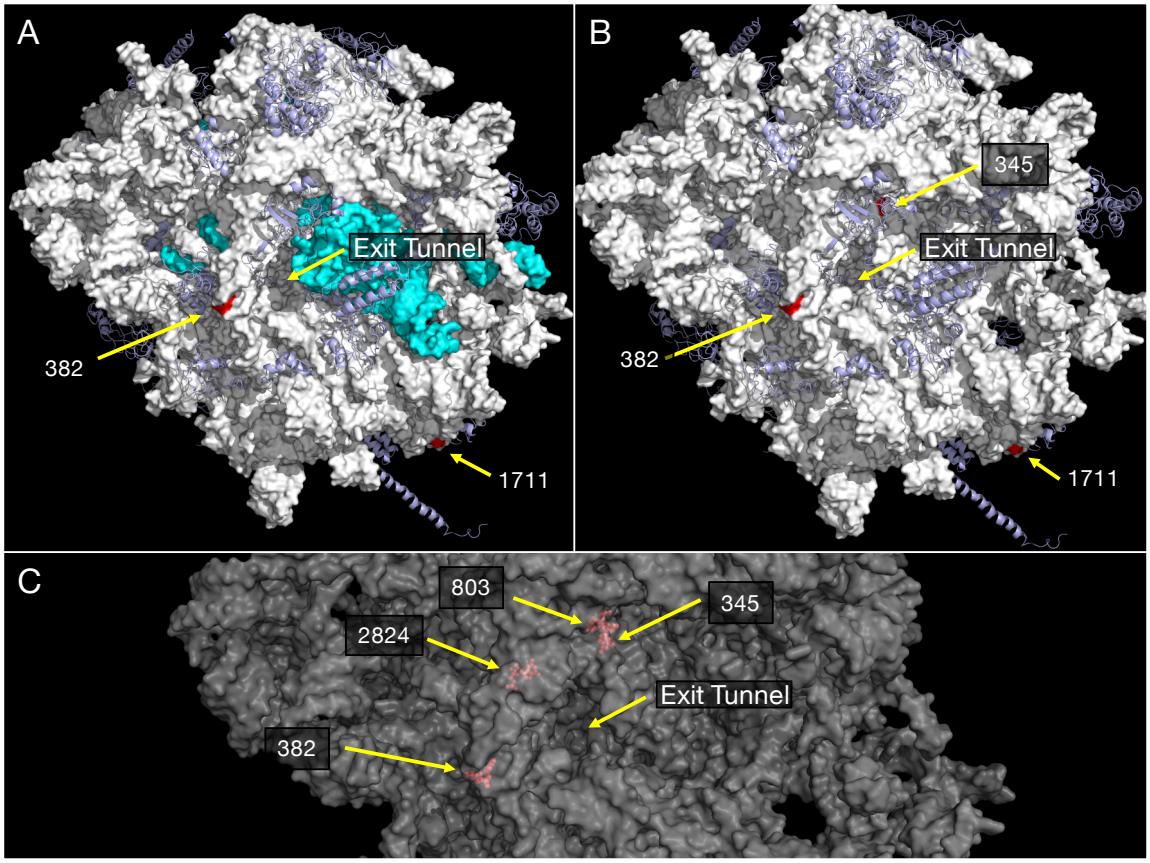


Figure 3.11. Three-dimensional views of the yeast ribosome from the solvent-exposed side of the large ribosomal subunit with major cisplatin adduct sites noted in red. 25S rRNA is shown in white surface, 5S and 5.8S rRNA are shown as cyan surface, ribosomal proteins as purple cartoon. A) Looking down the exit tunnel with 5S and 5.8S rRNA present. B) removing the 5S and 5.8S ribosomal subunits reveals adduct 345, just above the exit tunnel in this view. C) Partially transparent surface render reveals adducts 803 and 2824 and demonstrates how these strong stops are each located near the nascent peptide exit tunnel. 2824 is actually located at the core of the ribosome, in the peptidyl transferase center, and 803 is in a solvent-exposed cleft just off the exit tunnel. Figure created using PyMol, based on crystal structure PDB 4V7R

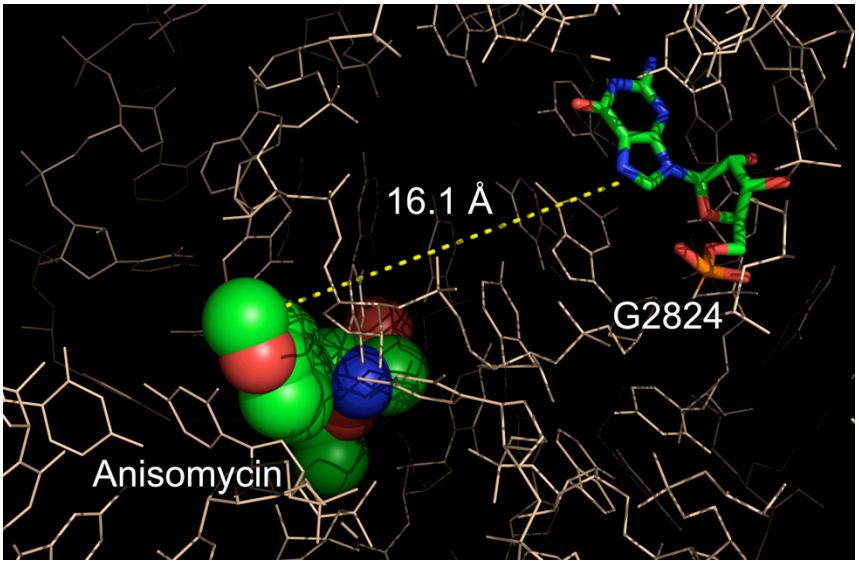


Figure 3.12. Pt-seq predicts a cisplatin-derived adduct at G2824, a residue near the PTC. Anisomycin is one of several eukaryotic-specific antibiotics which bind in a nearby pocket. (Garreau de Loubresse et al., 2014)

Notably, not all strong PT-SEQ stop sites are represented in the gel-based RT assays and vice-versa. We used Circligase to create single-stranded circular vector-like products post-RT for PCR amplification as described by Talkish et al. (2014). Circularization avoids a second potentially low-yielding ligation step and Circligase itself is very efficient: the manufacturer reports near 100% conversion from linear to circular products of control sequences. Circligase does have a 3' substrate bias toward primarily A and T residues in the 3' position for circularization (Epicentre). Transforming our Pt-seq results from adduct site to stop site (+1 nt) demonstrates a strong preference for A and T residues in the 3' position at both treatment concentrations (Figure B.3). This bias is present for both concentrations of cisplatin and also observed in control treatment (not shown), and suggests that results are biased towards those adjacent to a preferred Circligase substrate. This factor may explain the failure to recapitulate some strong stops

seen by others by traditional RT analysis in the yeast peptidyl transferase center (Hostetter et al., 2012; Osborn et al., 2014).

Conclusions

Our work using high-throughput sequencing demonstrates cisplatin binds to a broad number of ribosomal sites and numerous cellular RNAs (Tables B.1-B.4). The presence of stops significantly stronger than background suggests there are some highly preferred sites, as previously observed (Osborn et al., 2014; Melnikov et al., 2016). Work here has discovered interesting potential sites modified by cisplatin, including G2824 in the peptidyl transferase center and in a region near ribosome inhibitor binding sites (Garreau de Loubresse et al., 2014). This suggests a previously unexplored mechanism for cisplatin inhibition of the ribosome which warrants further investigation. Furthermore, our approach has been demonstrated for another platinum-based compound (azaplatin), demonstrating broader applicability for determination of ribosome modification sites by small molecule platinum reagents.

Summary of Chapter III and bridge to Chapter IV

In this chapter we presented a method which couples reverse transcriptase stalling at Pt(II) lesions with high-throughput sequencing to map all sites of platinum accumulation in cellular RNA, which we have termed "Platinum-seq." Using *Saccharomyces cerevisiae* as a model system, we have demonstrated Platinum-seq's utility by mapping Pt adducts following treatment with cisplatin and a derivative, azaplatin, on the large ribosomal subunit at two concentrations. We find dose-dependent

responses with both compounds, and map numerous binding sites to single-nucleotide resolution on the yeast large ribosomal subunit. High-throughput sequencing data gathered in this work is further expanded upon in Chapter III, in which we perform preliminary differential expression analysis and compare cellular response to treatment with the two different drugs.

CHAPTER IV

DIFFERENTIAL GENE EXPRESSION OF *S. CEREVISIAE* CELLS TREATED WITH CISPLATIN AND AZAPLATIN, A NOVEL CLICK REACTIVE PT(II)-BASED SMALL MOLECULE

This chapter contains contributions from Emily E. Reister and Victoria DeRose. Vickie provided guidance and direction throughout the project, and edited the final manuscript. Emily provided advice and assistance with differential expression, specifically performing analysis in R. I performed and analyzed the experiments and wrote the text.

Introduction

Cisplatin and related platinum-based anticancer compounds are present in 40-50% of all chemotherapy treatments (Dyson and Sava, 2006; Harper et al., 2010). Despite being used for approximately 40 years, mechanisms of cisplatin action are not fully understood (Siddik, 2003). Cisplatin is capable of forming exchange-inert complexes with DNA, RNA, proteins, membrane components, and small molecules (Harper et al., 2010). To better understand this important compound, we have developed a suite of cisplatin-derived small molecules capable of participating in click chemistry (White et al., 2016). Click chemistry describes a class of reactions which are fast, high-yielding, and bio-orthogonal and are well suited for adding unique functionality to small molecules (Kolb et al., 2001). We have used our click-functionalized platinum(II) compounds to append fluorescent molecules for monitoring platinum accumulation on synthetic DNA,

ex vivo RNA, extracted RNA, and begun cellular colocalization studies by click labeling platinum compounds in cells post-fixation (reviewed in White et al., 2016). We have recently developed a high-throughput sequencing approach to identify to single-nucleotide resolution cisplatin and azaplatin (Figure 4.1) binding sites on cellular RNA ("Pt-seq," Plakos and DeRose, 2016). Here, using data gathered from that experiment, we present preliminary differential expression analysis on non-housekeeping genes in *S. cerevisiae* cells treated with cisplatin and azaplatin (Figure 4.1).

Cisplatin treatment elicits a variety of cellular responses, including DNA damage response pathways and P53 activation (Jung and Lippard 2007; Basu and Krishnamurthy, 2010). RNA expression analysis has suggested that cisplatin treatment results in differential expression of a number of oncogenes (Jung and Lippard, 2007). Furthermore, a recent microarray study demonstrated that cisplatin and one of its two approved derivatives, carboplatin, exhibit unique differential gene expression profiles (Bicaku et al., 2012), likely stemming from their inherent differences in reactivity. Given this, it is important to understand how our modified Pt(II) compounds influence gene expression in comparison with the parent compound cisplatin, and to further develop a contextual understanding of their various effects on cellular function.

Results and Discussion

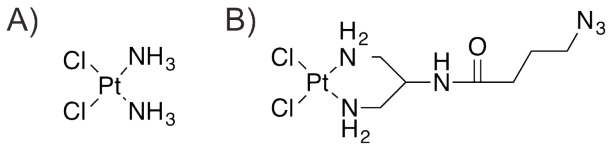


Figure 4.1. Platinum compounds used in this work. A) Cisplatin. B) Azaplatin (Wirth et al., 2015).

In our Pt-Seq experiment (Plakos and DeRose, 2016), *Saccharomyces cerevisiae* were treated with 100 and 200 μM of either cisplatin or azaplatin (Figure 4.1) for six hours. Total RNA was isolated and libraries prepared as described (Plakos and DeRose, 2016). From these samples, 35.8 million reads were obtained across six samples (two biological replicates each of control, 100 μM, and 200 μM azaplatin, Table 4.1). Data were aligned to the yeast genome using Bowtie (v1.1.2, Langmead et al., 2009), allowing 3 mismatches. Despite using data from an experiment designed for detecting RNA modifications (Plakos and DeRose, 2016), we found sufficient non-ribosomal transcriptome coverage to perform preliminary differential expression analysis of non-housekeeping genes. To do so, we masked ribosomal RNA, mitochondrial ribosomal RNA, and tRNA genes, and then counted all aligned reads across remaining genome features using HTSeq-count (Anders et al., 2005). Comparison of read counts across non-housekeeping shows that biological replicates correlate strongly ($r^2 > 0.95$ (data not shown), Figure C.1). We then performed differential expression analysis using the R package DESeq2 (Love et al., 2014).

	Total sequencing reads	Reads aligned to yeast genome	Percent aligned to yeast genome	Reads on non-housekeeping RNAs*
Control-1	4,113,364	3,826,004	93.0	213,627
Control-2	7,218,695	6,897,208	95.5	455,296
100 µM CisPt-1	4,851,790	4,608,522	95.0	294,912
100 µM CisPt-2	7,294,189	6,979,260	95.7	435,655
200 µM CisPt-1	5,925,462	5,548,737	93.6	294,181
200 µM CisPt-2	3,843,239	3,560,316	92.6	180,119
100 µM AzPt-1	5,152,589	4,925,655	95.6	365,552
100 µM AzPt-2	7,164,539	6,801,095	94.9	523,765
200 µM AzPt-1	5,549,454	5,305,073	95.6	379,955
200 µM AzPt-2	6,598,879	6,290,981	95.3	458,769
total	57,712,200			

* sum of reads aligned to regions other than rRNA and tRNA genes

Table 4.1. Sequencing read counts. Data represent two biological replicates each of control, 100 µM cisplatin (CisPt) treatment, 200 µM cisplatin treatment, 100 µM azaplatin treatment (AzPt), and 200 µM azaplatin treatment. Total obtained reads are presented, along with the count of reads that aligned to the yeast genome, percent alignment, and counts of reads aligned excluding rRNA and tRNA genes. Reads from the final column are used for differential expression analysis.

Data were regularized-log transformed using blind dispersion estimation for principal component analysis. PCA results show that samples cluster by treatment (Figure 4.2). The two cisplatin treatments are more closely related to control than azaplatin treatment, together constituting 12% of the sample variance. By contrast, 58% of sample variance is attributed to azaplatin treatment (Figure 4.2). Interestingly, while the two cisplatin concentrations deviate in the PCA, the two concentrations of azaplatin cluster extremely closely. We have not yet established an LD_{50} for azaplatin in yeast, but this compound has poor solubility in water and it is possible that the effective concentration is not actually doubled with 200 µM treatment.

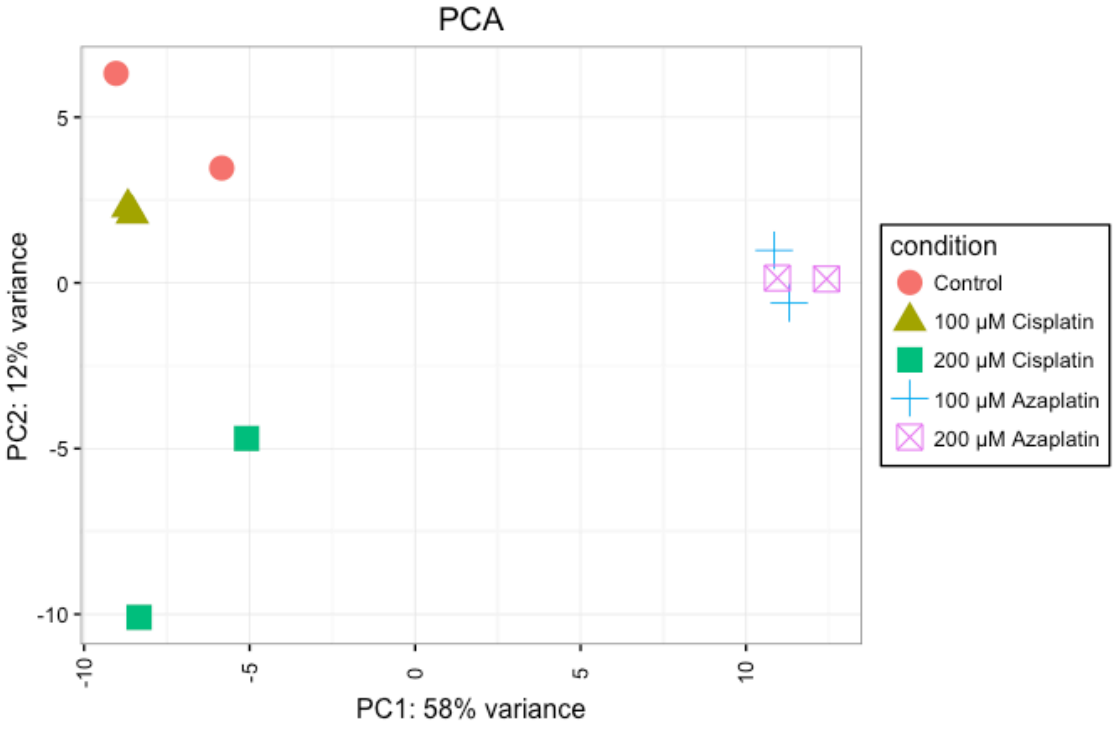


Figure 4.2. Principal component analysis of differential expression data for yeast cells treated with 100 µM cisplatin (yellow triangles), 200 µM cisplatin (green squares), 100 µM azaplatin (blue plus), 200 µM azaplatin (pink crossed boxes), and untreated control (orange circles). Most variance is attributed to treatment with azaplatin. Controls cluster most closely with cisplatin treatments, and azaplatin treatments cluster together.

Differential expression analysis reveals 12 total significantly differentially expressed genes under 100 µM cisplatin treatment, 55 under 200 µM cisplatin treatment, 213 under 100 µM azaplatin treatment, and 233 under 200 µM azaplatin treatment (Table 4.2). MA plots show that significantly expressed genes ($p_{adj} < 0.1$) are biased toward higher gene expression (Figure C.2, red points). Significantly differentially expressed genes are generally upregulated. Of note, the results from these exeriments are subject to bias due to low coverage, as upregulated genes are more likely to be sequenced since they represent a greater proportion of the RNA population (Table 4.2, Figure C.2). The 10 most upregulated and 10 most downregulated genes of each pairwise comparison are

shown in Tables 4.3, 4.4, and 4.5. Human homologs, when available, were determined using SGD's YeastMine and are also included (Cherry et al., 2012).

100 µM cisplatin	count	percent
Upregulated	4	0.07%
Downregulated	8	0.13%

100 µM azaplatin	count	percent
Upregulated	160	2.60%
Downregulated	53	0.88%

200 uM cisplatin	count	percent
Upregulated	38	0.63%
Downregulated	17	0.28%

200 uM azaplatin	count	percent
Upregulated	174	2.90%
Downregulated	59	0.98%

Out of 6049 genes with nonzero read count
padj < 0.1

Table 4.2. Counts of genes upregulated and downregulated by treatment with cisplatin and azaplatin, each at 100 µM and 200 µM. Gene expression is considered significantly different if adjusted p-value is less than 0.1.

Among the most highly expressed genes in both cisplatin treatments are RNR2 and RNR4 (Table 4.3). RNR2 and 4 are genes encoding the small subunit of ribonucleotide reductase, an essential enzyme responsible for the rate-limiting step of dNTP synthesis (Yao et al., 2003; Tsaponina et al., 2011). RNR2 and 4 are known to distribute to the cytoplasm upon DNA replication stress, such as induced by the alkylating agent methane methylsulfonate (MMS), which methylates the N7 position of guanine and the N3 position of adenine and is commonly used to induce DNA damage response pathways (Hurd et al., 1987; Lundin et al., 2005; Tkach et al., 2012). MMS's similarity to cisplatin is striking: cisplatin also modifies the N7 position of guanine, and adenine N3 is a preferred adenine platination site in gas-phase studies but is blocked from *in*-vivo DNA adduct formation by its location within the interior groove of the DNA double-helix (Chiavarino et al., 2015). RNR2 and RNR4 are upregulated upon cisplatin but not azaplatin treatment, and follow a dose-dependent response (RNR4 log2-fold response from 1.7 to 2.3 under 100 and 200 µM cisplatin treatment, respectively; RNR2

log2-fold response from 1.6 to 2.2 under 100 and 200 μM cisplatin treatment, respectively (Table 4.3)).

Several other cisplatin-induced genes are shared between the two cisplatin drug concentrations. HPF1 is upregulated in a dose-dependent response. Aside from its wine-clarifying attributes, no biological function has yet been ascribed to this gene (Brown et al., 2007). PDR5 (Pleiotropic Drug Resistance 5), a multidrug resistance protein involved in cation resistance (Miyahara et al., 1996; Bauer et al., 1999), is likewise upregulated in both treatments. Downregulated genes observed in both cisplatin treatments include DSE1, which is observed localizing to the cytoplasm upon DNA damage (Tkach et al., 2012), as well as INO1, CTS1, SCW11, PRY3, DSE2, and EGT2. Downregulated genes consistently show a dose-dependent response, with a greater log2-fold change observed upon greater cisplatin treatment (Table 4.3).

TSL1, TPS1, TPS2, and TPS3 are all upregulated upon azaplatin treatment. All three of these are slightly more upregulated under 100 μM azaplatin than 200 μM azaplatin (Tables 4.4 and 4.5). These genes are part of the trehalose synthesis pathway and are upregulated upon heat, nutrient, and osmotic stress (Winderickx et al., 1996). PHM7 is equally upregulated under both treatment conditions (2.9 log2-fold increased for both 100 and 200 μM azaplatin) and is known to localize to the cytoplasm upon DNA damage (Tkach et al., 2012). Numerous heat shock proteins are upregulated upon both azaplatin treatments, including MGA1, SSE2, HSP104, HSP78, and SSA4. No significant treatment effect was observed (Tables 4.4 and 4.5). Continuing the theme of DNA damage response, TMA10, a protein of unknown function which accumulates upon DNA damage (Fleischer et al., 2006), is upregulated equally in both azaplatin treatment

conditions (2.5 log2-fold). YNL194C is an unknown protein which, like the RNR genes seen to increase with cisplatin treatment, increases upon MMS treatment (Lee et al., 2007). YNL194C is increased upon both azaplatin treatments (Tables 4.4 and 4.5).

Azaplatin treatment downregulates the mitochondrial decarboxylase complex (GCV1, GCV2, GCV3) (Tables 4.4 and 4.5). PHO11 and PHO12 are paralogs, both seen downregulated under azaplatin treatment. PHO11 is significantly downregulated in 100 μM azaplatin (-2.2), PHO12 is downregulated under both treatments, in a slight dose-dependent manor (-2.2 at 100 μM azaplatin, -2.5 at 200 μM azaplatin). URA3, part of the pyrimidine synthesis pathway, is downregulated by both treatment conditions. PHO84 is downregulated in a slight dose-dependent manner by azaplatin (-3.1 Log2-fold increase under 100 μM azaplatin and -3.8 log2-fold increase under 200 μM azaplatin) (Tables 4.4 and 4.5). This gene is known as an inorganic phosphate transporter but has also been implicated in metal ion accumulation (Bun-Ya et al., 1991; Jensen et al., 2003). PHO84Δ yeast strains have reduced metal ion uptake and greater resistance to the toxic effects of manganese accumulation (Jensen et al., 2003).

100 µM cisplatin

SGD ID	Common Name	log2(FC)	padj	Human Homolog	Brief Description
YGR180C	RNR4	1.7	1.03E-05	RRM2B RRM2	Ribonucleotide-diphosphate reductase (RNR) small subunit
YJL026W	RNR2	1.6	1.03E-05	RRM2B RRM2	Ribonucleotide-diphosphate reductase (RNR), small subunit
YOL155C	HPF1	0.9	1.77E-02	NA	Haze-protective mannoprotein
YOR153W	PDR5	0.9	2.10E-02	ABCG4 ABCG1	Plasma membrane ATP-binding cassette (ABC) transporter
YNL327W	EGT2	-1.0	4.98E-03	NA	Glycosylphosphatidylinositol (GPI)-anchored cell wall endoglucanase
YHR143W	DSE2	-1.0	2.50E-02	NA	Daughter cell-specific secreted protein with similarity to glucanases
YJL078C	PRY3	-1.1	9.95E-02	GLIPR2	Cell wall-associated protein involved in export of acetylated sterols
YGL028C	SCW11	-1.2	2.06E-02	NA	Cell wall protein with similarity to glucanases
snR67	SNR67	-1.2	2.66E-02	NA	C/D box small nucleolar RNA (snoRNA)
YLR286C	CTS1	-1.2	1.87E-03	NA	Endochitinase
YJL153C	INO1	-1.2	6.50E-02	ISYNA1	Inositol-3-phosphate synthase
YER124C	DSE1	-1.4	9.97E-03	NA	Daughter cell-specific protein

200 µM Cisplatin

SGD ID	Common Name	log2(FC)	padj	Human Homolog	Brief Description
YOR382W	FIT2	2.5	1.09E-06	NA	Mannoprotein that is incorporated into the cell wall
YGR180C	RNR4	2.3	2.14E-11	RRM2B RRM2	Ribonucleotide-diphosphate reductase (RNR) small subunit
YJL026W	RNR2	2.2	3.88E-12	RRM2B RRM2	Ribonucleotide-diphosphate reductase (RNR), small subunit
YOR383C	FIT3	2.1	1.50E-03	NA	Mannoprotein that is incorporated into the cell wall
YOL155C	HPF1	1.6	4.66E-09	NA	Haze-protective mannoprotein
YBR072W	HSP26	1.6	7.38E-03	HSPB6 CRYAA CRYAB HSPB8 HSPB7 HSPB1 HSPB2 HSPB3 HSPB9	Small heat shock protein (sHSP) with chaperone activity
YDR001C	NTH1	1.5	2.78E-03	TREH	Neutral trehalase, degrades trehalose
YNL160W	YGP1	1.5	8.12E-07	ASPG	Cell wall-related secretory glycoprotein
YMR011W	HXT2	1.5	7.26E-03	SLC2A6 SLC2A8	High-affinity glucose transporter of the major facilitator superfamily
YLR034C	SMF3	1.3	9.24E-02	SLC11A2 SLC11A1	Putative divalent metal ion transporter involved in iron homeostasis
YKL164C	PIR1	-1.3	1.74E-02	NA	O-glycosylated protein required for cell wall stability
YGL028C	SCW11	-1.3	2.75E-03	NA	Cell wall protein with similarity to glucanases
YJL078C	PRY3	-1.5	3.22E-03	GLIPR2	Cell wall-associated protein involved in export of acetylated sterols
YBR158W	AMN1	-1.6	1.32E-02	NA	Protein required for daughter cell separation
YNL327W	EGT2	-1.6	1.25E-08	NA	Glycosylphosphatidylinositol (GPI)-anchored cell wall endoglucanase
YDL128W	VCX1	-1.6	1.41E-02	NA	Vacuolar membrane antiporter with Ca2+/H+ and K+/H+ exchange activity
YER124C	DSE1	-1.6	1.67E-03	NA	Daughter cell-specific protein
YJL153C	INO1	-1.8	9.44E-04	ISYNA1	Inositol-3-phosphate synthase
YLR286C	CTS1	-1.9	2.69E-09	NA	Endochitinase
YHR143W	DSE2	-2.1	5.66E-09	NA	Daughter cell-specific secreted protein with similarity to glucanases

Table 4.3. 10 most upregulated and 10 most downregulated genes for 100 µM and 200 µM cisplatin treatment vs. control. Human homologs are given, when available. Expression level given as log2(fold change). All genes exhibit adjusted p-values less than 0.1, as shown. Data from 100 µM cisplatin treatment only exhibited twelve genes with p_{adj} less than 0.1. Human homologs and gene descriptions found using SGD's YeastMine (Cherry et al., 2012).

100 μM Azaplatin

SGD ID	Common Name	log2(FC)	padj	Human Homolog	Brief Description
YML100W	TSL1	4.0	1.78E-32	NA	Large subunit of trehalose 6-phosphate synthase/phosphatase complex
YOL084W	PHM7	2.9	3.46E-07	NA	Protein of unknown function
YGR249W	MGA1	2.8	3.07E-06	HSFX1 HSF5 HSFY2 HSFY1	Protein similar to heat shock transcription factor
YDR074W	TPS2	2.8	3.50E-11	NA	Phosphatase subunit of the trehalose-6-P synthase/phosphatase complex
YBR169C	SSE2	2.7	5.25E-12	HYOU1 HSPH1 HSPA4L HSPA4	Member of Hsp110 subclass of the heat shock protein 70 (HSP70) family
YLL026W	HSP104	2.7	2.58E-16	CLPB	Disaggregase
YNL194C	NA	2.6	4.45E-05	NA	Integral membrane protein
YDR258C	HSP78	2.5	4.43E-11	CLPB	Oligomeric mitochondrial matrix chaperone
YLR327C	TMA10	2.5	3.20E-04	HABP4 SERBP1	Protein of unknown function that associates with ribosomes
YER103W	SSA4	2.5	1.08E-10	HSPA1A HSPA1B HSPA1L HSPA2 HSPA6 HSPA7 HSPA8	Heat shock protein that is highly induced upon stress
YNL250W	RAD50	-1.9	1.90E-02	RAD50	Subunit of MRX complex with Mre11p and Xrs2p
YHR215W	PHO12	-2.2	2.64E-03	NA	One of three repressible acid phosphatases
YAR071W	PHO11	-2.2	4.64E-03	NA	One of three repressible acid phosphatases
YHR216W	IMD2	-2.3	1.23E-04	GMPR IMPDH1 IMPDH2 GMPR2	Inosine monophosphate dehydrogenase
YBR296C	PHO89	-2.4	5.84E-04	SLC20A1 SLC20A2	Plasma membrane Na+/Pi cotransporter
YDR019C	GCV1	-2.5	2.38E-06	AMT	T subunit of the mitochondrial glycine decarboxylase complex
YMR189W	GCV2	-2.6	4.09E-09	GLDC	P subunit of the mitochondrial glycine decarboxylase complex
YAL044C	GCV3	-2.6	4.24E-05	GCSH	H subunit of the mitochondrial glycine decarboxylase complex
YML123C	PHO84	-3.1	9.63E-24	NA	High-affinity inorganic phosphate (Pi) transporter
YEL021W	URA3	-3.4	8.23E-32	UMPS	Orotidine-5'-phosphate (OMP) decarboxylase

Table 4.4. The 10 most upregulated and 10 most downregulated genes for 100 azaplatin treatment vs. control. Human homologs are given, when available. Expression level given as log2(fold change). All genes exhibit adjusted p-values less than 0.1, as shown. Human homologs and gene descriptions found using SGD's YeastMine (Cherry et al., 2012).

200 μM Azaplatin

SGD ID	Common Name	log2(FC)	padj	Human Homolog	Brief Description
YNR056C	BIO5	4.6	2.11E-13	SLC7A9 SLC7A13 SLC7A8 SLC7A11 SLC7A10 SLC7A5 SLC7A7 SLC7A6	Putative transmembrane protein involved in the biotin biosynthesis
YGR065C	VHT1	4.0	3.19E-23	NA	High-affinity plasma membrane H+-biotin (vitamin H) symporter
YNR058W	BIO3	3.7	3.01E-09	NA	7,8-diamino-pelargonic acid aminotransferase (DAPA)
YML100W	TSL1	3.3	6.25E-22	NA	Large subunit of trehalose 6-phosphate synthase/phosphatase complex
YOL084W	PHM7	2.9	3.90E-07	NA	Protein of unknown function
YNL194C	NA	2.8	1.15E-05	NA	Integral membrane protein
YLR177W	NA	2.7	1.20E-05	NA	Putative protein of unknown function
YLL026W	HSP104	2.5	1.13E-14	CLPB	Disaggregase
YER103W	SSA4	2.5	6.90E-11	HSPA1A HSPA1B HSPA1L HSPA2 HSPA6 HSPA7 HSPA8	Heat shock protein that is highly induced upon stress
YLR327C	TMA10	2.5	3.59E-04	HABP4 SERBP1	Protein of unknown function that associates with ribosomes
YMR189W	GCV2	-1.8	8.24E-05	GLDC	P subunit of the mitochondrial glycine decarboxylase complex
YHR216W	IMD2	-1.8	5.15E-03	GMPR IMPDH1 IMPDH2 GMPR2	Inosine monophosphate dehydrogenase
YAL044C	GCV3	-1.8	6.80E-03	GCSH	H subunit of the mitochondrial glycine decarboxylase complex
YDL227C	HO	-2.0	6.92E-03	NA	Site-specific endonuclease
YDR019C	GCV1	-2.0	2.33E-04	AMT	T subunit of the mitochondrial glycine decarboxylase complex
YBR296C	PHO89	-2.1	5.54E-03	SLC20A1 SLC20A2	Plasma membrane Na+/Pi cotransporter
YPL019C	VTC3	-2.1	1.20E-05	NA	Regulatory subunit of the vacuolar transporter chaperone (VTC) complex
YHR215W	PHO12	-2.5	3.59E-04	NA	One of three repressible acid phosphatases
YEL021W	URA3	-3.3	1.00E-29	UMPS	Orotidine-5'-phosphate (OMP) decarboxylase
YML123C	PHO84	-3.8	1.09E-30	NA	High-affinity inorganic phosphate (Pi) transporter

Table 4.5. The 10 most upregulated and 10 most downregulated genes for 200 μM azaplatin treatment vs. control. Human homologs are given, when available. Expression level given as log2(fold change). All genes exhibit adjusted p-values less than 0.1, as shown. Human homologs and gene descriptions found using SGD's YeastMine (Cherry et al., 2012).

We were next interested in which cellular processes were affected by cisplatin and azaplatin treatment. We took our list of differentially expressed genes under each condition with adjusted p-values less than 0.1 and performed KEGG pathway analysis and gene ontology (GO-term) analysis using the gage R package (Luo et al., 2009; Kanehisa et al., 2016). KEGG pathway analysis showed no significantly enriched pathways for either cisplatin treatment (p-value <0.1), which may be due to low coverage. Data from 100 μM azaplatin treatment shows significantly increased expression from "starch and sucrose metabolism" (p-value .009), and significantly decreased "purine metabolism" (p-value .03) and "metabolic pathways" (p-value .08), which is a broad category covering most metabolism in the cell. Samples from 200 μM azaplatin treatment show increased expression from "starch and sucrose metabolism" pathways (p-value .01) and no significantly decreased pathways.

A broader perspective of differential expression can be gained from exploring gene ontology of significantly expressed genes. Gene ontology (GO) classifies genes into categories describing their molecular function, biological process, and cellular component. We used SGD's GO Slim mapper to classify our differentially expressed genes into GO terms relating to function and process. GO Slim specifically uses curated lists of GO terms and is suitable for a broad overview without the fine-grained distinctions offered by the full GO database. We submitted significant gene lists from each treatment to SGD's GO Slim Process and Function categories, isolated the top 10 GO terms for each treatment by category (Tables 4.6 through 4.9), combined results

across treatments, and calculated the percent each GO term is represented within each pool of genes (Figure 4.3).

100 µM cisplatin

Biological Process

GO ID	GO term	Frequency	Gene(s)
55086	nucleobase-containing small molecule metabolic process	2	RNR4,RNR2
42221	response to chemical	2	DSE1,PDR5
6873	cellular ion homeostasis	1	PDR5
71554	cell wall organization or biogenesis	1	HPF1
7124	pseudohyphal growth	1	DSE2
23052	signaling	1	DSE1
6869	lipid transport	1	PRY3
6364	rRNA processing	1	SNR67
746	conjugation	1	DSE1
5975	carbohydrate metabolic process	1	INO1

Cellular Function

GO ID	GO term	Frequency	Gene(s)
16787	hydrolase activity	6	SCW11,DSE2,CTS1,EGT2,HPF1,PDR5
16798	hydrolase activity, acting on glycosyl bonds	5	SCW11,DSE2,CTS1,EGT2,HPF1
3674	molecular function unknown	2	DSE1,PRY3
16491	oxidoreductase activity	2	RNR4,RNR2
19843	rRNA binding	1	SNR67
30555	RNA modification guide activity	1	SNR67
16853	isomerase activity	1	INO1
22857	transmembrane transporter activity	1	PDR5
3723	RNA binding	1	SNR67
16887	ATPase activity	1	PDR5

Table 4.6. Gene ontology term analysis for 100 µM cisplatin treatment. Shown are Biological Process and Cellular Function GO terms for each treatment. Frequency is the number of significantly differentially expressed genes which fall into each category (p_{adj} <0.1).

200 μM cisplatin

Biological Process

GOID	GO term	Frequency	Gene(s)
71554	cell wall organization or biogenesis	6	SED1,GSC2,HSP150,PIR1,YGP1,HPF1
9451	RNA modification	6	SNR13,SNR161,SNR56,SNR63,SNR64,SNR75
42221	response to chemical	6	NTH1,DSE1,HSP104,AHP1,ASC1,PDR5
6364	rRNA processing	6	SNR13,SNR161,SNR56,SNR63,SNR64,SNR75
32196	transposition	6	YBL005W-B,YDR210W-B,YLR157C-B,YLR256W-A,YPL257W-B,YPR158C-D
5975	carbohydrate metabolic process	5	NTH1,GSC2,INO1,HSP104,GPH1
43934	sporulation	3	HHT1,PRB1,GSC2
55085	transmembrane transport	3	BAP2,VCX1,ITR1
55086	nucleobase-containing small molecule metabolic process	3	RNR1,RNR4,RNR2
6873	cellular ion homeostasis	3	VCX1,SMF3,PDR5

Cellular Function

GOID	GO term	Frequency	Gene(s)
16787	hydrolase activity	16	YBL005W-B,NTH1,YDR210W-B,PRB1,SCW11,DSE2,HSP150,HSP104,YLR157C-B,CTS1,SUN4,EGT2,HPF1,PDR5,YPL257W-B,YPR158C-D
3723	RNA binding	13	YBL005W-B,HSP26,YDR210W-B,YLR157C-B,YLR256W-A,YPL257W-B,YPR158C-D,SNR13,SNR161,SNR56,SNR63,SNR64,SNR75
16740	transferase activity	9	TLC1,YBL005W-B,YDR210W-B,MET6,GSC2,YLR157C-B,YPL257W-B,YPR158C-D,GPH1
3674	molecular function unknown	8	DSE1,SPI1,PRY3,TOH1,YGP1,FIT2,FIT3,YPR108W-A
16798	hydrolase activity, acting on glycosyl bonds	7	NTH1,SCW11,DSE2,CTS1,SUN4,EGT2,HPF1
22857	transmembrane transporter activity	7	BAP2,VCX1,ITR1,HXT4,SMF3,HXT2,PDR5
19843	rRNA binding	6	SNR13,SNR161,SNR56,SNR63,SNR64,SNR75
30555	RNA modification guide activity	6	SNR13,SNR161,SNR56,SNR63,SNR64,SNR75
16779	nucleotidyltransferase activity	6	TLC1,YBL005W-B,YDR210W-B,YLR157C-B,YPL257W-B,YPR158C-D
8233	peptidase activity	6	YBL005W-B,YDR210W-B,PRB1,YLR157C-B,YPL257W-B,YPR158C-D

Table 4.7. Gene ontology term analysis for 200 μM cisplatin treatment. Shown are Biological Process and Cellular Function GO terms for each treatment. Frequency is the number of significantly differentially expressed genes which fall into each category (p_{adj} <0.1).

100 µM azaplatin

Biological Process

GOID	GO term	Frequency	Gene(s)
5975	carbohydrate metabolic process	26	TPS1,GID7,GLK1,CIT2,NTH1,TPS2,GLC3,HXK1,TOS3,GSC2,VID28,KAR2,TDH1,INO1,UGP1,HSP104,GSY2,TSL1,PGM2,GID8,TPS3,GAC1,BUD7,TYE7,GPH1,GDB1
6364	rRNA processing	25	UTP20,NUG1,NOP7,MTR4,UTP10,RPS1A,RPS1B,RRP5,DBP2,PUS7,RRP12,SNR161,SNR17A,SNR17B,SNR30,SNR39,SNR41,SNR42,SNR43,SNR5,SNR60,SNR80,SNR81,SNR82,SNR83
8150	biological process unknown	22	AI3,AI5_BETA,YRO2,YBR287W,NRP1,JIP4,YER158C,YGR130C,RTS3,RIE1,PHO12,YLR177W,TMA10,STP3,YNL144C,PHM7,DUF1,YPL247C,YPR108W-A,YPR117W,SNR190,SNR4
55086	nucleobase-containing small molecule metabolic process	20	COX1,ATP8,ATP6,COB,OLI1,COX2,COX3,GLK1,URA3,RNR1,HXK1,PNC1,IMD2,TDH1,IMD3,APT1,AAH1,IRA2,VHS3,TYE7
42221	response to chemical	19	FLO1,UGA2,FES1,TPS1,HSP30,NTH1,ADR1,MTL1,RIM101,VHR1,KAR2,IML2,TMA19,HSP104,HAP1,GAD1,SIS1,IRA2,AZF1
55085	transmembrane transport	16	ATP8,ATP6,OLI1,SSA1,BAP2,AGP2,PHO89,AGP1,HXT7,HXT6,PIC2,SSA4,PMC1,KAR2,SSA2,PHO84
6091	generation of precursor metabolites and energy	16	COX1,COB,COX2,COX3,GLK1,GLC3,HXK1,TDH1,UGP1,HAP1,GSY2,PGM2,GAC1,TYE7,GPH1,GDB1
6457	protein folding	14	SSA1,HSP26,SSE2,AHA1,HSP78,SSA4,MDJ1,BTN2,SSA2,HSP104,SIS1,STI1,HSP82,CUR1
71554	cell wall organization or biogenesis	13	ZRG8,MTL1,GSC2,SNG1,RIM101,KAR2,MHP1,YPS1,LST8,YGP1,HPF1,AZF1,BUD7
9408	response to heat	13	FLO1,HSP26,TPS1,HSP30,TPS2,AHA1,HSP78,SSA4,MDJ1,HSP104,GAC1,LSP1,CUR1

Cellular Function

GO ID	GO term	Frequency	Gene(s)
16787	hydrolase activity	42	AI3,AI4,ATP8,ATP6,OLI1,SCEI,RPR1,SSA1,PHO11,HO,NTH1,DOA4,TPS2,HSP78,PRB1,NUG1,SSA4,DUG1,PMC1,PNC1, DSE2,PHO12,KAR2,MTR4,SMC3,CPS1,TIF1,SSA2,HSP104,YPS1,YEF3,NTE1,TSL1,TPS3,DBP2,AAH1,RAD50,HPF1,STE13, NEW1,HSP82,GDB1
3674	molecular function unknown	41	AI5_BETA,YRO2,AGP2,SDS24,YBR287W,GID7,HSP30,NRP1,PAM1,MTH1,JIP4,ZRG8,SPI1,YER158C,VTC2,MDS3,MTL1, YGR130C,RTS3,SNG1,VID28,IML2,TMA19,YLR177W,TMA10,MSC1,GID8,DIA1,YNL144C,YGP1,YNL194C,PHM7,RIM20, BUD7,VTC3,YPL247C,ATG41,SPO24,YPR108W-A,YPR117W,PIN3
3723	RNA binding	41	LSR1,SSA1,UTP20,HSP26,NUG1,MPT5,NOP7,RIE1,IMD2,MTR4,UTP10,JSN1,RPL14A,TIF1,SSA2,YEF3,IMD3,CLU1,RRP5, TIF11,STI1,RRP12,NEW1,SNR161,SNR17A,SNR17B,SNR19,SNR190,SNR30,SNR39,SNR4,SNR41,SNR42,SNR43,SNR5, SNR6,SNR60,SNR80,SNR81,SNR82,SNR83
16740	transferase activity	32	AI2,TLC1,TPS1,GLK1,CIT2,GIN4,GLC3,MET6,HXK1,TOS3,GSC2,VTC4,YAK1,HAL5,PTK2,UGP1,HSL1,PTK1,SHM2,GSY2, CAR2,APT1,TSL1,TPS3,RAD50,CMK2,HRK1,RPA190,PYK2,RPA135,GPH1,GDB1
22857	transmembrane transporter activity	19	COX1,ATP8,ATP6,COB,OLI1,COX2,COX3,BAP2,PHO89,AGP1,HXT7,HXT6,PIC2,PMC1,TPO2,PHO84,LYP1,TPO4,TPO3
16887	ATPase activity	18	ATP8,ATP6,OLI1,SSA1,HSP78,SSA4,PMC1,KAR2,MTR4,SMC3,TIF1,SSA2,HSP104,YEF3,DBP2,RAD50,NEW1,HSP82
19843	rRNA binding	16	RRP5,SNR161,SNR17A,SNR17B,SNR190,SNR30,SNR39,SNR41,SNR42,SNR43,SNR5,SNR60,SNR80,SNR81,SNR82,SNR83
3729	mRNA binding	15	UTP20,HSP26,NUG1,MPT5,NOP7,RIE1,IMD2,MTR4,JSN1,IMD3,CLU1,RRP5,STI1,RRP12,NEW1
30234	enzyme regulator activity	14	FES1,SSE2,AHA1,MDJ1,TFS1,TSL1,TPS3,LST8,APJ1,IRA2,STI1,VHS3,GAC1,CIP1
16491	oxidoreductase activity	13	COX1,COB,COX2,COX3,GCV3,UGA2,GCV1,RNR1,ERG1,IMD2,TDH1,IMD3,GCV2

Table 4.8. Gene ontology term analysis for 100 µM azaplatin treatment. Shown are Biological Process and Cellular Function GO terms for each treatment. Frequency is the number of significantly differentially expressed genes which fall into each category (p_{adj} <0.1).

200 µM azaplatin

Biological Process

GO ID	GO term	Frequency	Gene(s)
6364	rRNA processing	31	UTP20,SAS10,HCA4,MTR4,UTP10,MRT4,EBP2,NOP56,RPS1B,RRP5,NOP4,NIP7,RPS23B,SNR128,SNR161,SNR17A,SNR17B,SNR30,SNR39,SNR40,SNR41,SNR42,SNR47,SNR49,SNR5,SNR60,SNR8,SNR80,SNR81,SNR82,SNR83
5975	carbohydrate metabolic process	25	VID24,TPS1,GID7,GLK1,CIT2,NTH1,TPS2,GLC3,HXK1,TOS3,GSC2,ENO1,KAR2,TDH1,INO1,UGP1,HSP104,GSY2,TSL1, PGM2,GPD2,GAC1,TYE7,GPH1,GDB1
42221	response to chemical	21	FES1,TPS1,HSP30,RPN4,YDL124W,NTH1,ADR1,HSP12,HAC1,MTL1,YHB1,VHR1,KAR2,TMA19,HSP104,HAP1,ORM2,SIS1,BXI1,DDR2,IRA2
8150	biological process unknown	21	YRO2,STP4,YDR186C,JIP4,YER158C,RTS3,RIE1,PHO12,TOH1,MHO1,YLR177W,TMA10,STP3,PHM7,ZPS1,YPL247C, YPR108W-A,YPR117W,TDA6,SNR190,SNR4
55085	transmembrane transport	18	ATP6,OLI1,SSA1,BAP2,AGP2,PHO89,AGP1,CCC2,HXT7,HXT6,HXT3,SIT1,PIC2,SSA4,PMC1,KAR2,SSA2,PHO84
9451	RNA modification	16	NOP56,SNR128,SNR161,SNR39,SNR40,SNR41,SNR42,SNR47,SNR49,SNR5,SNR60,SNR8,SNR80,SNR81,SNR82,SNR83
71554	cell wall organization or biogenesis	16	TIP1,FMP45,PST1,SED1,MTL1,GSC2,KAR2,MHP1,HSP150,YPS1,LST8,YGP1,KRE1,HPF1,FLC1,NCE102
6091	generation of precursor metabolites and energy	16	COX1,COX3,GLK1,GLC3,HXK1,ENO1,TDH1,UGP1,PET10,HAP1,GSY2,PGM2,GAC1,TYE7,GPH1,GDB1
55086	nucleobase-containing small molecule metabolic process	15	COX1,ATP6,OLI1,COX3,GLK1,URA3,RNR1,HXK1,PNC1,ENO1,IMD2,TDH1,GPD2,IRA2,TYE7
6457	protein folding	14	SSA1,HSP26,SSE2,AHA1,HSP78,SSA4,BTN2,PHB2,SSA2,HSP104,SIS1,STI1,FLC1,HSP82

Cellular Function

GOID	GO term	Frequency	Gene(s)
3674	molecular function unknown	49	YRO2,VID24,AGP2,SDS24,GID7,HSP30,YCR061W,PRM7,SAS10,FMP45,PST1,YDR186C,MTH1,JIP4,SPI1,YER158C,VTC2,MDS3,MTL1,RTS3,PHB2,NCA3,TOH1,MHO1,TMA19,PET10,YLR177W,MSC3,TMA10,ORM2,MSC1,DIA1,YGP1,YNL194C,BXI1,DDR2,MAM3,PHM7,ZPS1,RIM20,VTC3,NIP7,YPL247C,ATG41,SPO24,YPR108W-A,YPR117W,NCE102,TDA6
3723	RNA binding	44	LSR1,SSA1,UTP20,HSP26,MSS116,PUF6,HYP2,VTC1,MPT5,RIE1,IMD2,MTR4,UTP10,JSN1,MRT4,EBP2,SSA2,NOP56,RRP5,STI1,NOP4,PUF2,SNR128,SNR161,SNR17A,SNR17B,SNR19,SNR190,SNR30,SNR39,SNR4,SNR40,SNR41,SNR42,SNR47,SNR49,SNR5,SNR6,SNR60,SNR8,SNR80,SNR81,SNR82,SNR83
16787	hydrolase activity	35	ATP6,BI2,OLI1,RPR1,SSA1,TIP1,DUR1,2,APE3,HO,NTH1,TPS2,MSS116,HSP78,CCC2,PRB1,SSA4,DUG1,PMC1,PNC1,ESP1,PHO12,HCA4,KAR2,MTR4,SMC3,HSP150,CPS1,SSA2,HSP104,YPS1,TSL1,HPF1,STE13,HSP82,GDB1
16740	transferase activity	31	TLC1,HMT1,TPS1,GLK1,CIT2,GLC3,MET6,HXK1,TOS3,GSC2,BIO2,VTC4,YAK1,HAL5,OPI3,UGP1,HSL1,PTK1,SHM2,SAM1,GSY2,CAR2,TSL1,BIO3,CMK2,HRK1,RPA190,PYK2,SPE3,GPH1,GDB1
19843	rRNA binding	21	MRT4,RRP5,SNR128,SNR161,SNR17A,SNR17B,SNR190,SNR30,SNR39,SNR40,SNR41,SNR42,SNR47,SNR49,SNR5,SNR60,SNR8,SNR80,SNR81,SNR82,SNR83
22857	transmembrane transporter activity	21	COX1,ATP6,OLI1,COX3,BAP2,PHO89,AGP1,CCC2,HXT7,HXT6,HXT3,SIT1,PIC2,PMC1,TPO2,PHO84,LYP1,BIO5,TPO4,FLC1,TPO3
3729	mRNA binding	16	UTP20,HSP26,MSS116,PUF6,VTC1,MPT5,RIE1,IMD2,MTR4,JSN1,EBP2,NOP56,RRP5,STI1,NOP4,PUF2
16887	ATPase activity	16	ATP6,OLI1,SSA1,MSS116,HSP78,CCC2,SSA4,PMC1,HCA4,KAR2,MTR4,SMC3,HSP150,SSA2,HSP104,HSP82
30555	RNA modification guide activity	16	SNR128,SNR161,SNR190,SNR39,SNR40,SNR41,SNR42,SNR47,SNR49,SNR5,SNR60,SNR8,SNR80,SNR81,SNR82,SNR83
16491	oxidoreductase activity	15	COX1,COX3,GCV3,YDL124W,GCV1,SUR2,RNR1,LPD1,YHB1,IMD2,TDH1,MAE1,HFD1,GCV2,GPD2

Table 4.9. Gene ontology term analysis for 200 µM azaplatin treatment. Shown are Biological Process and Cellular Function GO terms for each treatment. Frequency is the number of significantly differentially expressed genes which fall into each category (p_{adj} <0.1).

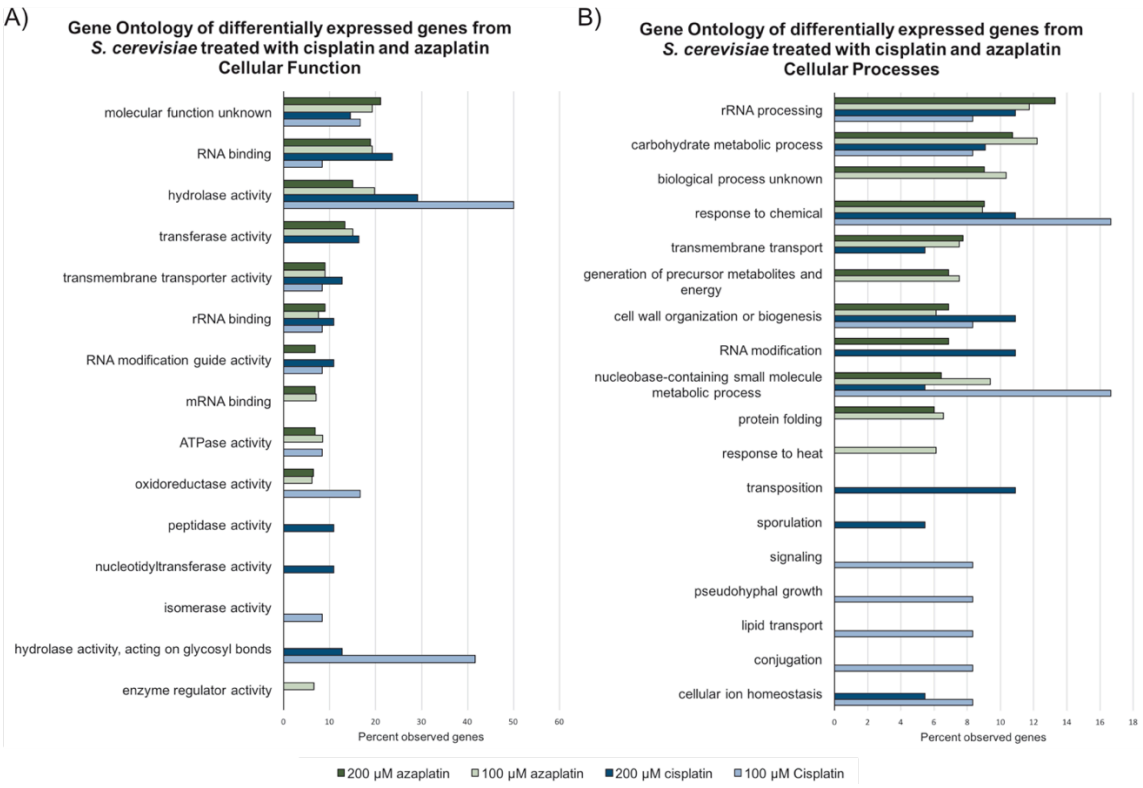


Figure 4.3. Gene ontology of differentially expressed genes from *Saccharomyces cerevisiae* treated with cisplatin (100 µM and 200 µM, blue and light blue) or azaplatin (100 µM and 200 µM, green and light green). Gene IDs from differential expression results (p_{adj}<0.1) were extracted and sorted by gene ontology using Go Slim Mapper by the Saccharomyces genome database (www.yeastgenome.org). Percent of each sample dataset within each GO term is shown as bars. A) Cellular function gene ontology of *S. cerevisiae* treated with 100 µM azaplatin. B) Cellular process gene ontology of *S. cerevisiae* treated with 200 µM azaplatin.

GO Slim analysis reveals similarities in each differentially expressed gene set (Figure 4.3 A and B). It is important to note that these categories are not mutually exclusive. For example, numerous snoRNAs were sequenced, and these appear under "RNA binding," "rRNA binding," and "RNA modification guide activity" categories, explaining their high enrichment (Tables 4.6 through 4.9). These RNA classes are interesting because cisplatin is observed to accumulate on cellular RNA (Hostetter et al., 2012) and the ribosome hypothesized to act as a platinum sink. We also observe

enrichment of "transmembrane transfer activity" genes (Figure 4.3, Tables 4.6 through 4.9). Cisplatin results in increased expression of genes related to metal ion transfer across membranes (Table C.2), in accordance with the observation that cisplatin is primarily shuttled into cells by metal ion transporters (Roos and Kaina, 2012).

Conclusions

This work is a pilot study, using data from an experiment designed to analyze cisplatin and azaplatin adducts on yeast ribosomal RNA to derive additional information on differential gene expression. Because of the experimental design, read counts outside ribosomal RNA and tRNA are quite low (Table 4.1) and we likely only received sufficient coverage to derive statistically significant expression information on highly expressed genes (Figure C.2, Table 4.2). Nevertheless, we found that biological replicates of each treatment cluster with each other via principal component analysis (Figure 4.2), and plots of raw gene counts between replicates show strong correlation (Figure C.1). We were able to find 12 significantly differentially expressed genes for 100 μM cisplatin treatment, 55 for 200 μM cisplatin, 213 for 100 μM azaplatin, and 233 for 200 μM azaplatin. GO Slim analysis shows that our results comprise primarily genes associated with cellular function GO terms including RNA binding, hydrolase activity, transferase activity, transmembrane transporter activity, rRNA binding and RNA modification guide activity. These RNA-driven terms are enriched due to the large number of snoRNA's sequenced (Table C.2). Similarly, rRNA processing is the most well-represented cellular process GO term, and RNA modification are well represented biological process GO terms. (Figure 4.3B).

Taken together, these results indicate that *S. cerevisiae* responds differently to treatment with these two Pt(II) compounds. Coverage was extremely low for 100 μM cisplatin so we are unable to make many conclusions from that data. Considering 200 μM cisplatin, 12 of the 20 most upregulated genes are unique to cisplatin treatment. These include genes such as RNR2 and 4 (shared with 100 μM cisplatin and discussed above), which are involved in DNA damage response, SED1, which assists with translating ribosomes SMF3, a putative divalent metal ion transporter which increases under DNA replication stress (Cherry et al., 2012). In addition to inducing DNA damage, Azaplatin appears to activate a myriad of metal ion transporters. PMC1, PIC2, and CCC2 are all involved in either Calcium transportation (PMC1) or Copper transportation (PIC2, and CCC2). Additionally, ATP synthase genes are significantly overexpressed in azaplatin treatment, hinting that the compound may interact with the mitochondria.

Cisplatin treatment follows a dose-dependent pattern much more than azaplatin treatment, which correlates with the relationships between treatments shown in Figure 4.2. Azaplatin is a new Pt(II) compound, recently synthesized by our lab (Wirth et al., 2015). Azaplatin has shown to be poorly soluble in water and we are working to establish its LD_{50} in yeast cells. In this study, principal component analysis reveals that 100 and 200 μM azaplatin treatments cluster extremely closely with each other (Figure 4.2), which is reflected in their inconsistent dose-dependent gene expression profiles (Tables 4.8 and 4.9). The similar responses seen by each treatment concentration indicate we may be working outside an effective treatment range for azaplatin. Future studies would benefit from titrating azaplatin's dosage and analyzing differential expression.

Ultimately, despite low coverage we were able to perform a pilot differential expression analysis. This information is very valuable to have in a high-throughput sequencing strategy such as Platinum-seq. This work has hinted at differences in the way *S. cerevisiae* respond to cisplatin versus our click-functionalized cisplatin derivative azaplatin. Understanding the expression-level response organisms have to azaplatin, especially as compared with cisplatin, is crucial, as so much of cellular response will dictate how the compound interacts. Future studies would benefit from repeating this work using ribosomal depletion, to reduce background noise.

Summary of Chapter IV and bridge to Chapter V

In this chapter we performed preliminary differential expression analysis of *S. cerevisiae* treated with 100 and 200 μM of either cisplatin or its click-functionalized derivative azaplatin. Using sequencing reads from Chapter III, we find azaplatin treatment resulted in more differentially expressed genes than cisplatin treatment. Cisplatin treatment follows a dose-dependent response more closely than azaplatin. Azaplatin upregulates numerous metal ion transporters not seen in cisplatin treatment. Together, these data suggest different responses of yeast cells to these two compounds. Chapter V is an exploration of a DNA hybridization-based method to enrich for platinated nucleic acids pre-sequencing.

CHAPTER V

DNA HYBRIDIZATION PULLDOWN OF CLICK-BASED AFFINITY TAGGED DNA

This chapter contains contributions from Anna Hickey and Victoria DeRose. Professor DeRose provided insight and direction throughout the work, and edited the final manuscript. Click reactions, pulldown experiments, and gels were performed by Anna Hickey. Methods were developed by myself and Anna Hickey. The writing is my own, and figures were assembled by myself with Anna's assistance.

Introduction

Cisplatin is the founding member of a class of platinum anticancer drugs that are used in approximately 40-50% of all chemotherapy treatments (Dyson and Sava, 2006; Harper et al., 2010) (Figure 5.1A). Cisplatin treatment is accompanied by numerous problematic side effects, including nephrotoxicity, ototoxicity, and hair loss, and tumor resistance limits long-term effectiveness of the drug (Florea and Büsselberg, 2011). Cisplatin's cytotoxic effects are attributed to its accumulation on cellular DNA, primarily forming covalent adducts with adjacent guanine residues (Fichtinger-Schepman et al., 1985; Jung et al., 2007). Cisplatin can induce apoptosis, but the mechanisms are not yet fully described, as evidence supports both nucleus-dependent (Basu and Krishnamurthy, 2010) and nucleus-independent pathways (Mandic et al., 2003). Despite the fact that cisplatin is a known DNA-binding drug, we have a surprisingly poor understanding of the genomic landscape of cisplatin-induced Pt adducts. It is, however, understood that

cisplatin can disrupt transcription through stalling RNA Pol II (Damsma et al., 2007), suggesting that cisplatin accumulation on certain genetic locations may play a role in cytotoxicity.

Recent high-throughput sequencing approaches seek to better understand the genomic targets of small molecules (Rodriguez and Miller, 2014). Methods such as CHIP-seq probe DNA binding proteins and histone modifications using immunoprecipitation (Park, 2009). Researchers have also recently used immunoprecipitation techniques to identify novel sites of RNA modification (Dominissini et al., 2016). Chem-seq, a new technology inspired by CHIP-seq, seeks to identify and sequence the binding sites of small molecules via high-throughput sequencing (Anders et al., 2014).

We have recently developed a suite of functionalized cisplatin-derived small molecules with the goal of applying similar methods to identify the cellular targets of Pt(II) compounds. Our compounds, one of which, azaplatin, is shown in Figure 5.1B, enable use of the prototypical azide-alkyne cycloaddition "click" reaction (Figure 5.1C, Kolb et al., 2001) to bioorthogonally label platinum compounds attached to biomolecules both *in vitro* and in cell culture (Wirth et al., 2015). Here we describe an *in vitro* approach to using these compounds in conjunction with a DNA hybridization-based streptavidin bead pulldown to selectively isolate a platinated DNA oligonucleotide. We are interested in methods of pre-enrichment of platinated nucleic acid fragments for high-throughput sequencing analysis of platinated DNA and RNA. DNA hybridization-based methods are well represented in the literature (Horn, 2012) and are valued for their ease of use, sensitivity, and specificity (Mamanova et al., 2010).

Results

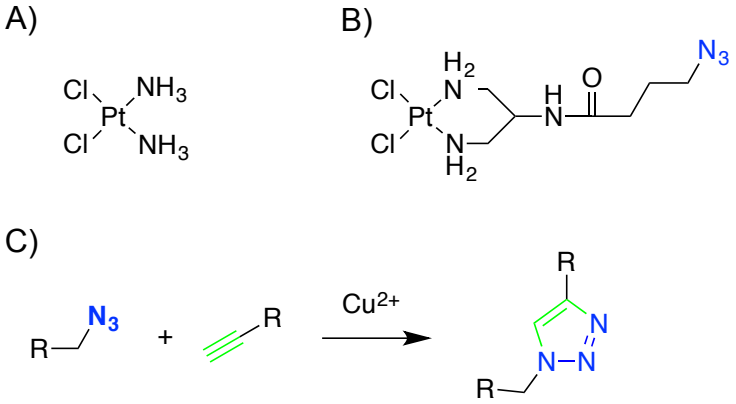


Figure 5.1. Azaplatin is a Pt(II)-based click-reactive small molecule (Wirth et al., 2015). A) Cisplatin. B) Azaplatin. C) Copper catalyzed azide-alkyne cycloaddition ("click" reaction).

We first prepared a control hybridization probe by clicking a fluorescent molecule, rhodamine-B azide (Rho), to a 5'-hexynyl functionalized synthetic DNA (HEX) (Figure 5.2, lane 2, bands B and C). This fluorescently labeled DNA (HEX-Rho) was used to optimize our bead binding and pulldown conditions. We found our most efficient approach was to pre-hybridize HEX-Rho with its 5'-biotinylated complement (BIO), then incubate overnight on MyOne C1 streptavidin magnetic beads (4 °C with rotation) that had been washed and pre-blocked with a 58-nt oligonucleotide (BLOCK) (Figure 5.2, band "BLOCK"). Pre-incubation with BLOCK served to block nonspecific binding sites on the streptavidin beads without interfering with the streptavidin-biotin interaction. Upon heating in Elution Buffer (see Methods), fluorescently labeled HEX was released (Figure 5.2, lane 7, bands HEX-Rho or BIO) along with a small amount of nonspecifically bound BLOCK (Figure 5.2, lane 7, band BLOCK).

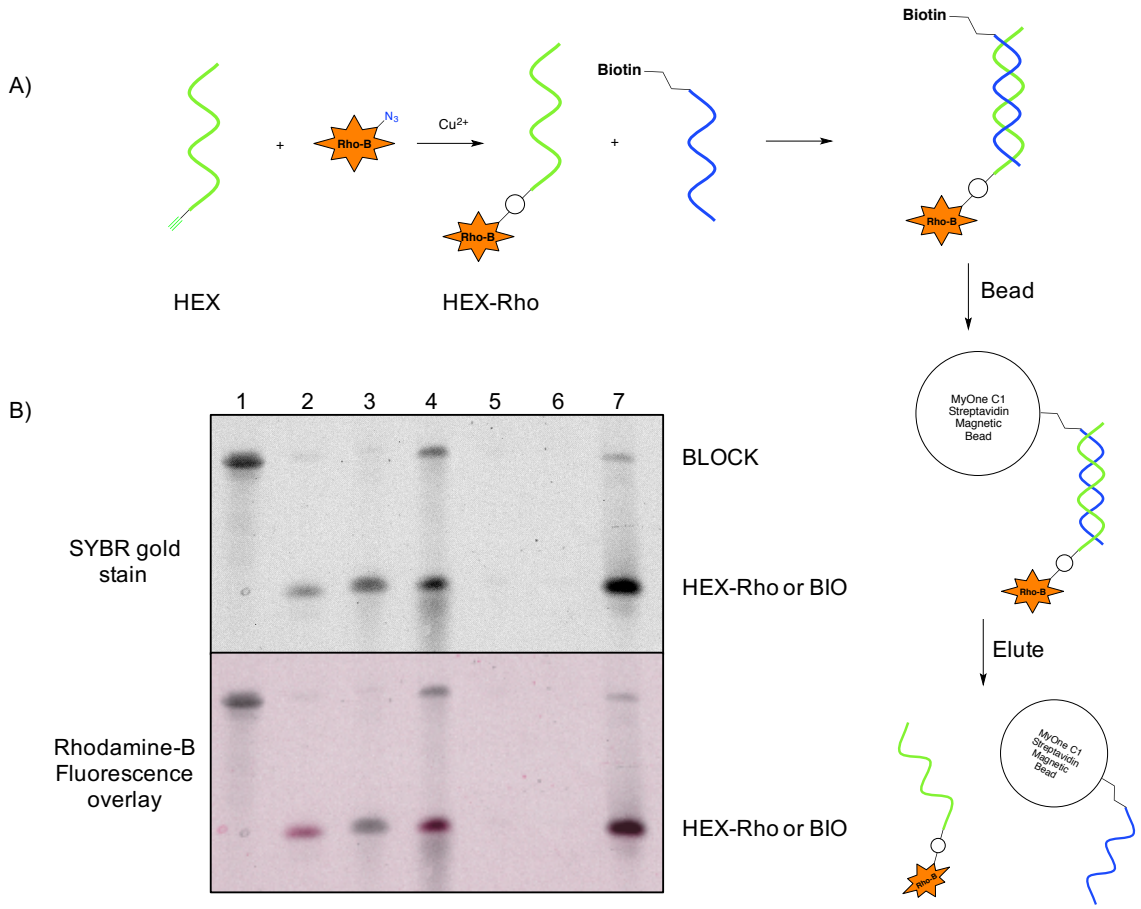


Figure 5.2. Pulldown of 5'-hexynyl DNA clicked to rhodamine-B azide. A) reaction scheme. B) PAGE gel demonstrating pulldown. Top image is SYBR gold stain, which reveals all nucleic acids in gel. Bottom image is rhodamine fluorescence overlay, artificially colored magenta. Colors inverted for clarity. Lane 1: blocking DNA only. Lane 2: Rho-clicked hex DNA only. Lane 3: 5'-biotin complementary strand only. Lane 4: pulldown supernatant. Lanes 5 and 6: washes 1 and 3, respectively. Lane 7: elution. Bands are labeled as follows: Blocking DNA (BLOCK). Rho-clicked hex DNA (HEX-Rho) (lane 2) or 5'-biotin complementary strand (BIO) (lane 3). In the bottom gel, we include rhodamine fluorescence overlay, pre-SYBR gold stain. Magenta colors over bands in lanes 2, 4, and 7 reveal eluted rhodamine-clicked HEX.

We next platinated a model cisplatin binding site, a synthetic DNA hairpin containing a single dGpG platinum(II) binding site (HP) using azaplatin (AzaPt, Wirth et al., 2015), a click-functionalized platinum(II) compound (Figure 5.1B). Platination was performed by incubating the DNA hairpin and azaplatin in a buffered solution overnight to allow aquation of the platinum species and coordination to the dGpG binding site

within HP. Unplatinated and platinated HP are separated by dPAGE wherein a lower band (Figure 5.3B, lane 1) migrates equivalent to an unplatinated HP control (not shown). Addition of azaplatin adds a +1 or +2 charge and a mass increase, resulting in a second band which migrates more slowly (Figure 5.3B, lane 1). Pt-HP clicks to a rhodamine-B azide in a Cu-dependent manner, and the clicked product further slows migration and results in an additional band shift (Figure 5.3B) present only in (+) $CuSO_4$ lanes. The successful click reaction is also demonstrated by rhodamine-B fluorescence visible in the gel pre-staining (artificially colored magenta in Figure 5.3B, Rhodamine-B fluorescence, lanes 3, 5, 7). We determined that under our reaction conditions, 75 μM $CuSO_4$ gives the greatest click conversion and minimal degradation of DNA (Figure 5.3B, lane 5).

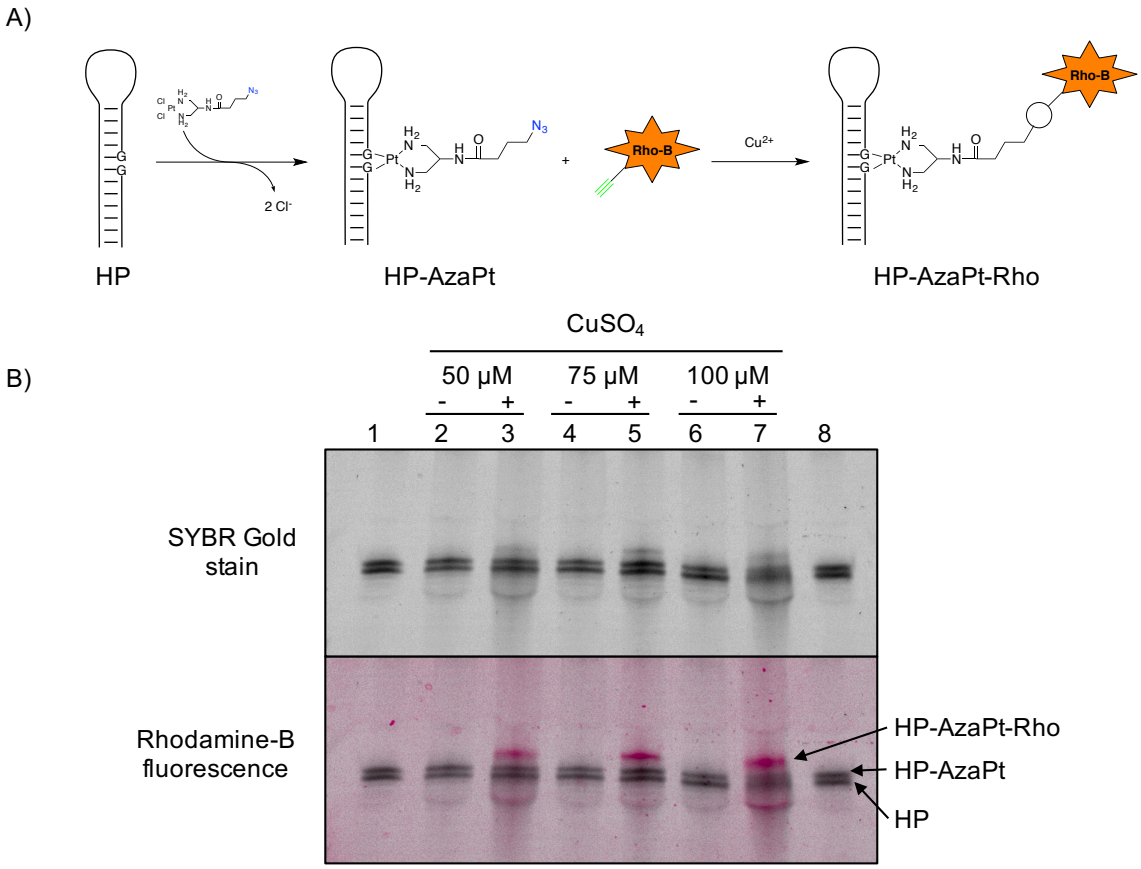


Figure 5.3. Rhoxamine-B alkyne clicked to azaplatin bound to DNA hairpin. A) reaction scheme. For simplicity rhodamine-B alkyne is shown as an orange star. B) PAGE gel demonstrating click reaction. Lanes 1 and 8: hairpin DNA treated with azaplatin (HP-AzaPt). Lanes 2-7: click reactions (-) and (+) $CuSO_4$ catalyst, as labeled. Click conditions are (final concentrations): 3 µM azaplatin-DNA, 13 uM rhodamine-B alkyne, 375 µM THPTA, 150 µM sodium ascorbate, 100 mM phosphate buffer, pH 7. Copper concentrations in each (+) $CuSO_4$ reaction: 50 µM (lane 3), 75 µM (lane 5), 100 µM (lane 7). Top gel is SYBYR gold stain only, bottom is SYBYR gold plus rhodamine fluorescence overlay (pre-SYBR gold stain). Bands are labeled as follows: platinated hairpin clicked to rhodamine (HP-AzaPt-Rho). Platinated hairpin only (HP-AzaPt). Unplatinated hairpin (HP).

Having optimized click conditions with rhodamine, we set out to click HP-AzaPt to HEX. We platinated HP with AzaPt as described, then reacted this complex with equimolar HEX for 2 hours at 37 °C. The resulting products were separated by dPAGE. The HEX-HP clicked DNA (Figure 5.4, band HP-AzaPt-HEX) migrates slower than

either unclicked complex (Figure 5.4, lane 5, compare A to B, C, and D) and is much more prevalent in (+) $CuSO_4$ conditions (lane 5, as compared with lane 4).

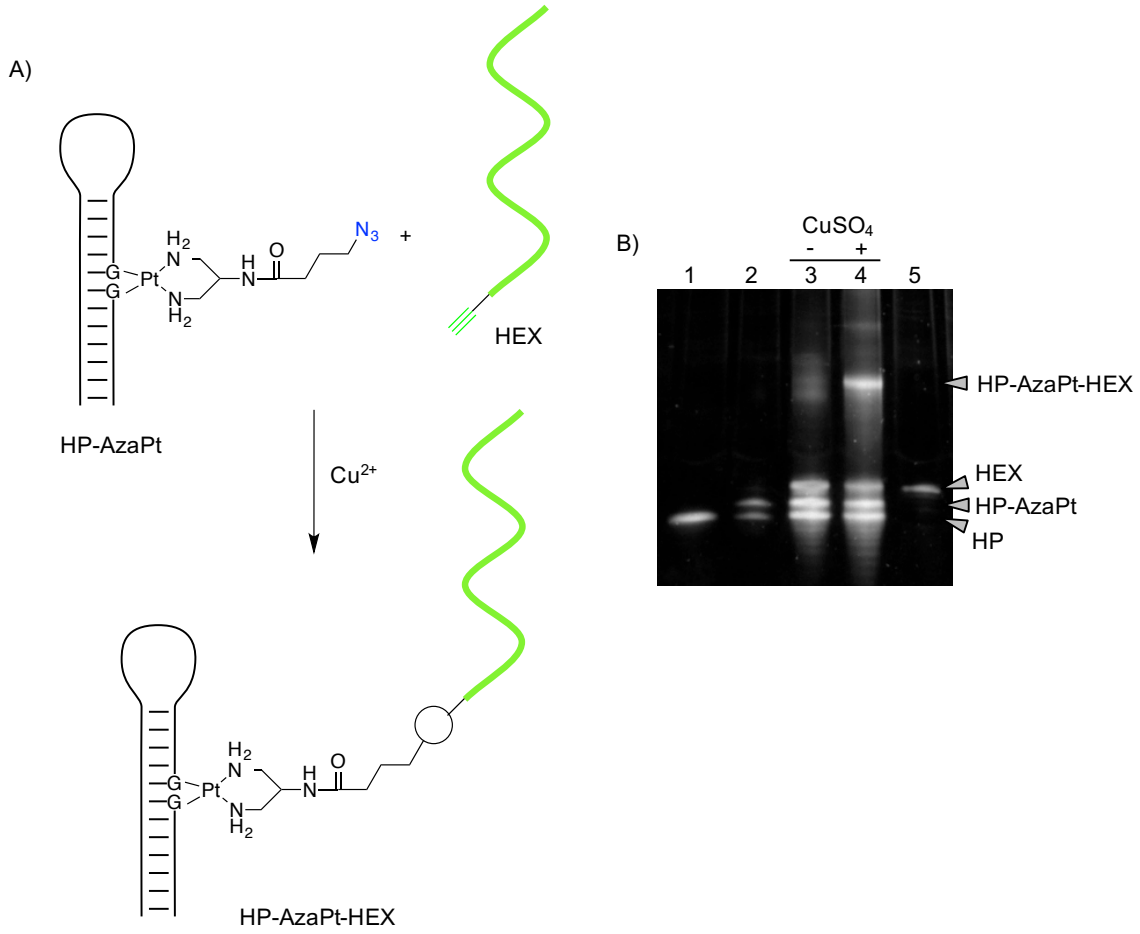


Figure 5.4. Platinated Hairpin clicks to hexynyl-functionalized oligonucleotide HEX. A) reaction scheme. B) PAGE gel demonstrating successful click reaction. Lane 1: HP only. Lane 2: HP-AzPt. Lane 3: HP-AzPt + HEX click reaction (-) $CuSO_4$. Lane 4: HP-AzPt + HEX click reaction (+) $CuSO_4$. Lane 5: HEX only. Bands are labeled as follows: Click product (HP-AzaPt-HEX), unclicked hexynyl oligonucleotide (HEX), unclicked HP-AzaPt (HP-AzaPt), unreacted HP (HP).

We prepared more HP-AzaPt-HEX (Figure 5.5B, lane 5), and then incubated the reaction mixture with the 5' biotinylated complement (BIO) (see Methods). This mixture was then incubated with washed and blocked streptavidin beads overnight at 4 °C with rotation. After washing (Figure 5.5B, lanes 7-9), samples were eluted from the beads by

heating to 95 °C for 5 minutes in Elution Buffer (see Methods), then placed on a magnet for 2 minutes to separate beads from eluant (Figure 5.5B, lane 10). Once again, reacting HP-AzaPt with HEX in the presence of 75 μM $CuSO_4$ results in a visible click product band (Figure 5.5B, "HP-AzaPt-HEX"). This band is not visible in supernatant or any of the bead washes, suggesting successful capture of the complex hybridized to BIO by streptavidin. Upon eluting from the beads, the click product band is visible (Figure 5.5B, lane 10, "HP-AzaPt-HEX") along with released BIO and HEX, suggesting these conditions have also captured and released unclicked HEX-BIO pairs which were enriched by the streptavidin beads.

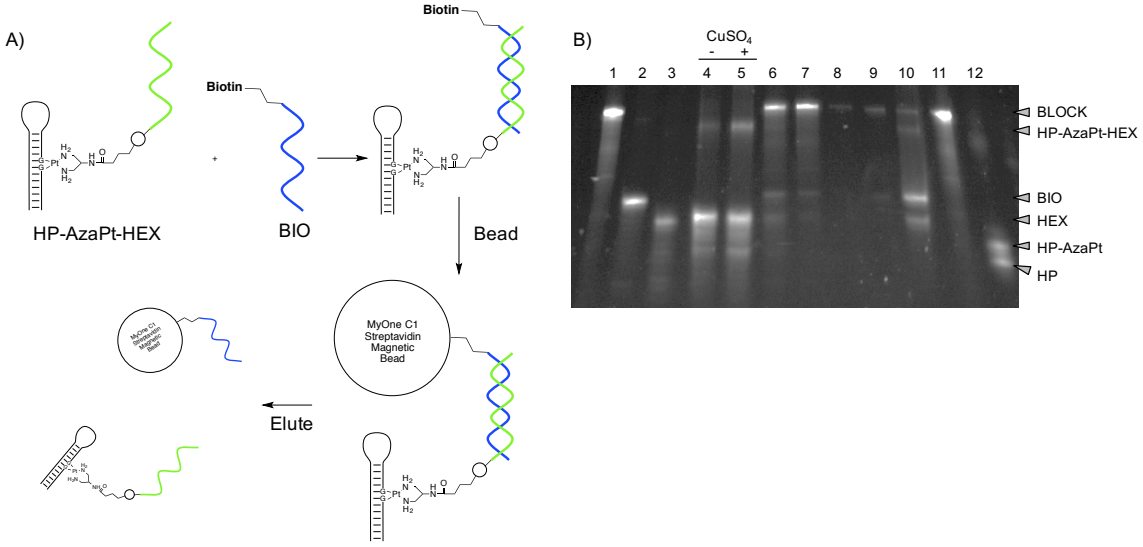


Figure 5.5. DNA hybridization-based pulldown of platinated and clicked hairpin DNA. A) reaction scheme. Blocking treatment to magnetic beads not shown. B) PAGE gel showing pulldown of clicked HP-AzaPt-HEX. Lane 1: BLOCK only. Lane 2: BIO only. Lane 3: HEX only. Lane 4: HP-AzPt + HEX click reaction (-) $CuSO_4$. Lane 5: HP-AzPt + HEX click reaction (+) $CuSO_4$. Lane 6: Supernatant remaining after overnight incubation of 5 with BLOCK-washed streptavidin beads. Lanes 7-9: washes 1, 3, 5 respectively. Lane 10: Bead elution. Lane 11: BLOCK only. Lane 12: HP-AzaPt. Bands are labeled as follows: Blocking DNA (BLOCK), click product (HP-AzaPt-HEX), 5'-biotin oligonucleotide (BIO), 5'hexynyl functionalized complementary oligonucleotide (HEX), unclicked platainated hairpin (HP-AzaPt), unplatinated hairpin (HP).

Conclusions

Toward the goal of developing a DNA hybridization-based pulldown technique, we have successfully platinated a model DNA hairpin (HP) with a click-functionalized Pt(II) compound (Azaplatin), and clicked this compound to a 5'-hexynyl functionalized DNA oligonucleotide and captured it by biotin-streptavidin pulldown. This hybridization-based pulldown technique shows great promise for isolating platinated oligonucleotides from complex solution while preserving the platinum-DNA bonds. DNA hybridization-based methods are well-used for high-throughput sequencing target enrichment (Mamanova et al., 2010; Horn, 2012). Hybridization-based target capture is easy to use, highly specific, modular, and affordable, especially for large numbers of samples (Mamanova et al., 2010). Efforts are underway in our lab to platinate, fragment, and isolate genomic DNA, and we foresee this technique as a way of isolating those platinum sites intact for downstream sequencing approaches.

CONCLUDING REMARKS

This work represents a significant step toward high-throughput mapping of Pt(II) adducts in cells treated with Pt(II) reagents. We have successfully mapped thousands of potential platination sites on hundreds of unique *S. cerevisiae* RNAs. Furthermore, we have demonstrated differential expression analysis from Platinum-seq datasets, a key step toward comprehensively understanding RNAs to which platinum reagents bind and the genetic response mounted against platinum reagent binding. This also allows us to assess relative abundance of different RNAs, improving models for platinum binding. Our work with the HDV ribozyme demonstrates great potential for using Pt(II) reagents for mapping important metal ion binding centers, which play key roles in coordinating RNA structure. Our pulldown methodology shows promise for enriching sample pools for platinated RNA pre-sequencing, decreasing experimental background noise and potentially allowing the discovery of low-copy number platinum bound nucleic acids. Future directions include using Platinum-seq to explore RNA binding in ribosome depleted samples, in order to sequence less abundant RNAs. We also plan to expand this approach toward looking at human cancer lines. Principles developed here for sequencing cisplatin binding sites are applicable to studying platinum lesions in genomic DNA, an important target that has not been thoroughly explored.

APPENDIX A

Materials and Methods

WT HDV ribozyme enzyme strand (ES) was produced by the Golden Lab (Purdue University, Lafayette, IN) using in-vitro transcription. Substrate strands (SS) were synthesized commercially and purchased from Dharmacon (GE Healthcare). DNA reverse transcriptase primers and controls were purchased from IDT. All RNA and DNA were 5' end-labeled using T4 Polynucleotide Kinase (Thermo Scientific), following typical methods. Before labeling, the 5' end phosphate on the HDV enzyme strand was removed using FastAP (Thermo Scientific), following manufacturer's protocols. Sequences for each strand used can be found in the appendix A.

13-mer substrate strand

Crosslinking

Crosslinked species were formed by mixing 40 μM HDV enzyme strand (ES), 40 μM substrate strand (SS, 13-mer,PS,dU1), and trace 32P ES or SS (as indicated) in 10 mM pH 7 phosphate buffer (PB), and 100 mM $NaNO_3$, then heating to 90 °C for 2 minutes and slowly cooling to 20 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes, then $Mg(NO_3)_2$ was added to a final concentration of 30 mM, and cisplatin to 120 μM. Nitrate salts are used instead of halides in order to promote aquation of cisplatin. Samples were crosslinked for 0-6 at 37 °C, as indicated, then quenched in 3x volume formamide loading buffer and resolved on 8 M urea, 1x TBE, 20% 29:1 bis:mono acrylamide gel.

Crosslinking kinetics

20 μM ES and trace 32P SS (13-mer,PS,dU1) were mixed in 10 mM pH 7 PB, and 100 mM $NaNO_3$, then folded by heating to 90 °C for 5 minutes and slowly cooling to 25 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes, then $Mg(NO_3)_2$ and cisplatin were added to final concentrations of 30 mM and 60 μM, respectively, to initiate crosslinking. The reaction was allowed to proceed at 37 °C and sampled at times ranging from 0-1400 minutes, and crosslinking was quenched by addition of 3x volume of formamide loading buffer. Samples were resolved on 8 M urea, 1x TBE, 20% 29:1 bis:mono acrylamide gel.

Reverse transcriptase primer extension analysis

40 μM ES and 40 μM SS (13-mer,PS,dU1) were mixed in 10 mM pH 7 PB, then folded by heating to XX for XX minutes and slowly cooling to 25 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes, then $Mg(NO_3)_2$ and cisplatin were added to final concentrations of 30 mM and 120 μM, respectively, to initiate crosslinking. The reaction was allowed to proceed at 37 °C for 6 hours then quenched via desalting (Sephadex G-25 column). Reverse transcription of crosslinked RNA was performed as described previously (Hostetter et al., 2012), using a 32P 5'-end labeled RT primer. RT products are resolved on an 8M urea, 1x TBE, 8% 19:1 mono:bis acrylamide sequencing gel.

11-mer Substrate Strand

Crosslinking

40 μM 32P 5' end-labeled substrate strand ((11-mer,PS,dU1) or (11-mer,PO,dU1), where noted) and 40 μM HDV enzyme strand were folded in the presence of 10 mM PB (pH 7), and 100 mM $NaNO_3$ by heating to XX for XX minutes and slowly cooling to 25 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes, and crosslinking was initiated by adding $Mg(NO_3)_2$ to a final concentration of 1 mM and cisplatin to final concentrations ranging from 0-120 μM and allowed to react for 6 hours at 37 °C. The reaction was quenched by the addition of 3x volume formamide loading buffer. Samples were resolved on 8 M urea, 1x TBE, 20% 29:1 bis:mono acrylamide gel.

Crosslinking kinetics

4 μM ES, 4 uM SS (11-mer,PS,dU1), and trace 32P SS(11-mer,PS,dU1) were mixed in 10 mM PB (pH 7), and 100 mM $NaNO_3$,then folded by heating to XX for XX minutes and slowly cooling to 25 °C over 20 minutes. After folding, crosslinking was initiated by adding $Mg(NO_3)_2$ and cisplatin to a final concentration of 10 mM and 4 μM, respectively. Crosslinking proceeded at 37°C and the reaction was sampled and quenched in formamide loading buffer at times from 0-300 minutes. Samples were resolved on 8M urea, 1x TBE, 20% 29:1 bis:mono acrylamide gels.

Cleavage kinetics

40 μM ES and Trace 32P SS (11-mer,PO) were mixed in 10 mM PB (pH 7) then folded by heating to XX for XX minutes and slowly cooling to 25 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes and $Mg(NO_3)_2$ was added to a concentration of 30 mM to initiate cleavage. Samples were taken from 0-600 minutes and quenched in 3x volume formamide loading buffer, then resolved on an 8M urea, 1x TBE, 20% 29:1 bis:mono acrylamide gel. Kinetics are determined using SigmaPlot (Chapman and DeRose 2012).

Alkali hydrolysis mapping

Trace 32P 5' end-labeled ES and 40 μM SS(11-mer,PS,dU1) were mixed in 10 mM PB (pH 7) and 100 mM $NaNO_3$, then folded by heating to XX for XX minutes and slowly cooling to 25 °C over 20 minutes. After folding, samples were heated to 37 °C for 10 minutes and crosslinking was initiated by adding $Mg(NO_3)_2$ and either cisplatin or 2-ADAP Pt to final concentrations of 10 mM and 40 μM, respectively. The reaction was allowed to proceed for 6 hours at 37 °C, then stopped by the addition of 3x volume of formamide loading buffer and the entire sample was run on a 15% 29:1 mono:bis acrylamide gel (8M urea, 1x TBE). After sufficient separation was achieved, products were resolved using autoradiography and crosslinked bands were excised and purified by crushing and soaking in 1 mL 100 mM acetic acid (pH adjusted to 3.4 using TEMED) at 20 °C while rotating overnight. Eluted RNA was concentrated and purified using ethanol precipitation, followed by further desalting (Sephadex G-25 column) then thoroughly dried via speed-vac. Purified crosslinked complex is resuspended in water and

hydrolyzed by treatment with equal volume hydrolysis buffer (50 mM carbonate buffer, pH 9.5, 1 mM EDTA) at 90 °C for times ranging from 0-10 minutes. Hydrolysis was quenched using equal volume of 50 mM citrate buffer, pH 3.5 and hydrolysis fragments were resolved on a 15% 19:1 mono:bis acrylamide, 1xTBE, 8M urea sequencing gel. Reference ladders were generated using RNAse T1, following the manufacturer's protocol (Thermo Scientific), to partially cleave 32P 5' end-labeled ES and a thorough hydrolysis ladder was generated by hydrolyzing a small amount of 32P 5' end-labeled ES completely for 18 minutes in hydrolysis buffer.

Supplementary Figures

HDV Enzyme Strand (ES)
5′-GGUCCGCAGCCUCCUCGCGGCGCAAGCUGGGCAACAUUCCGAAAGGUAAUGGCGAAUGCGGACC-3′

Substrate Strands (SS)

13-mer Substrates

Cleavable (13-mer,PO)
5′-CUUUUGGCUUGCA -3′
Phosphorothioate-substituted (13-mer,PS,dU1)
5′-CUUUdU(PS)GGCUUGCA -3′
Inhibited (13-mer,PO,dU1)
5′-CUUUdUGGCUUGCA -3′

11-mer Substrates

Cleavable (11-mer,PO)
5′-CCUGGCUUGCA -3′
Phosphorothioate-substituted (11-mer,PS,dU1)
5′-CCdU(PS)GGCUUGCA -3′
Inhibited (11-mer,PO,dU1)
5′-CCdUGGCUUGCA -3′

Figure A.1. Sequences used. HDV enzyme strand was synthesized and provided by the Golden Lab (Perdue University, Lafayette IN). Synthetic RNA was purchased from Dharmacon, synthetic DNA was purchased from IDT.

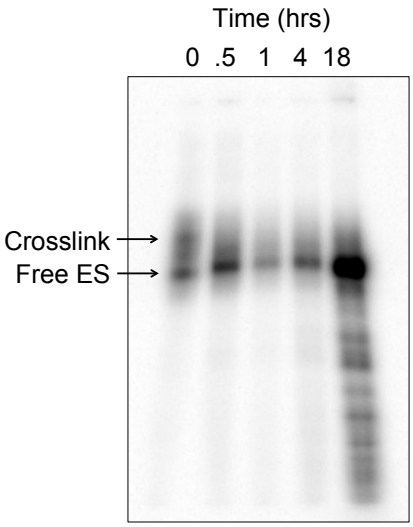


Figure A.2. Thiourea reversal of crosslinked ES-SS(11-mer,PS,dU1) complex. 32P 5'-end labeled ES was crosslinked to SS as described in Methods, then treated with an equal volume of saturated thiourea at 37°C for times ranging from 0-18 hours.

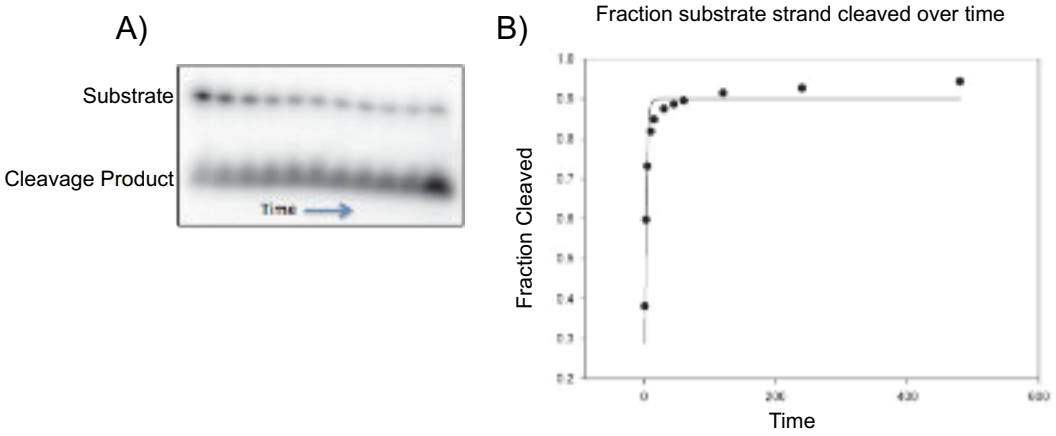


Figure A.3. 11-mer substrate strand cleavage kinetics. A) Autoradiograph demonstrating cleavage of a 32P 5'-end-labeled SS (11-mer,PO). Catalytic activity of the HDV ribozyme against a cleavable substrate strand indicates proper folding of the enzyme-substrate complex. A) Quantification of the image in A) demonstrates a rate of 0.39 ± 0.04 min-1. This compares favorably to the rate of the native complex (0.34 ± 0.02 min-1(Ananvoranich and Perreault, 1998)), suggesting this complex folds well.

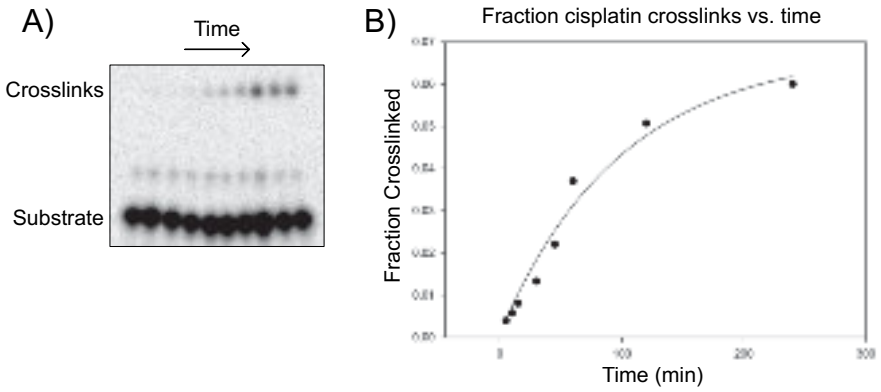


Figure A.4. 11-mer substrate strand crosslink kinetics. A) Autoradiograph demonstrating crosslinking between folded HDV ribozyme and 32P 5'-end-labeled phosphorothioate SS (11-mer,PS,dU1). B) Quantification of A). Cisplatin crosslinks appear with an observed rate constant of kobs = .01 +/- 0.003 min-1, and a pseudo second-order rate constant calculated to be k2, calc = 42 +/- 13 M-1 s-1. k2, calc was estimated by dividing kobs by the concentration of cisplatin used (Chapman and DeRose, 2012).

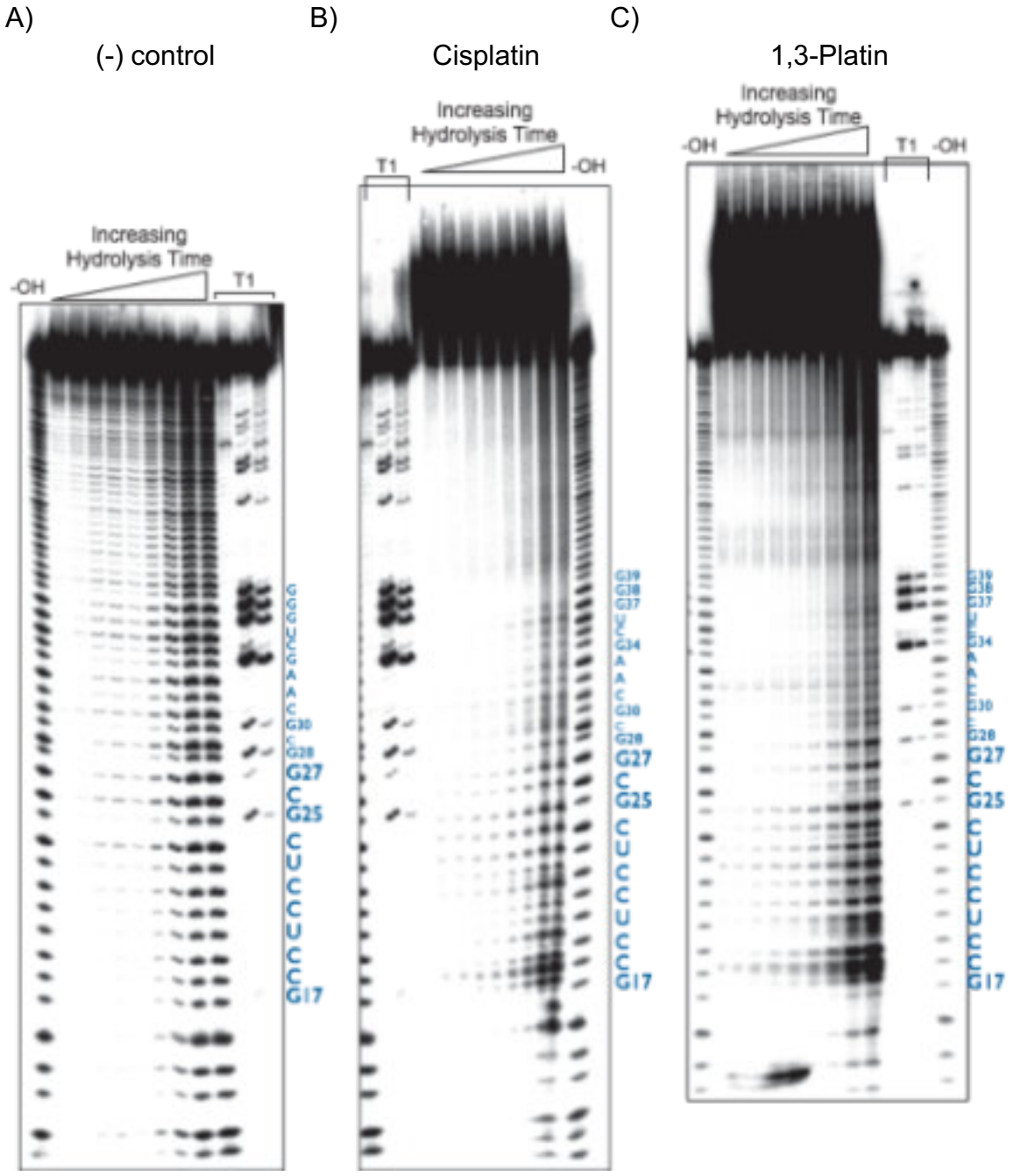


Figure A.5. Alkali hydrolysis mapping full gels. A) (-) control is ES-PS11 mer folded and mock crosslinked with H2O. B) Cisplatin and C) 2-ADAP Pt are platinated with 40 uM final concentration of their respective platinum(II) compounds. All three samples were separated, purified, and crosslinked, and hydrolyzed identically.

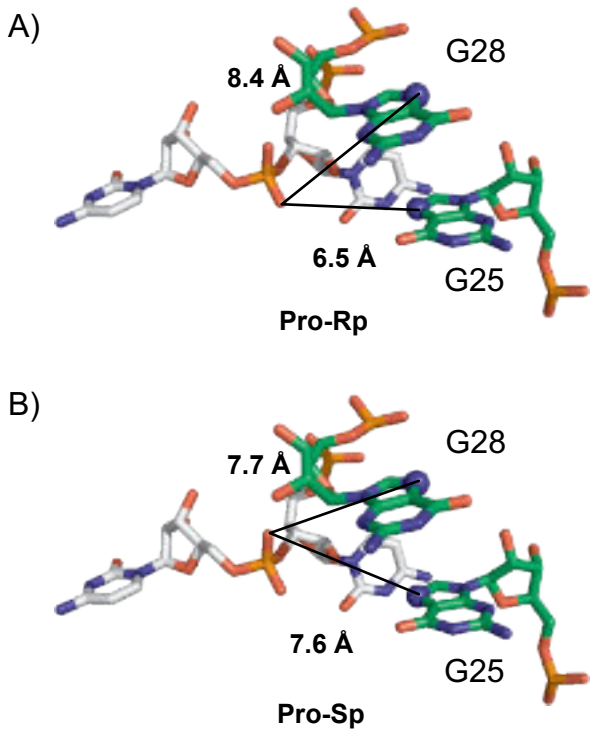


Figure A.6. Distance from Pt(II) sites to phosphorothioate-substituted scissile phosphate stereoisomers. PS substrate strands are a racemic mixture of the Pro-Rp and Pro-Sp substituted phosphorothioate. Measurements from N7 of G25 (expected) and N7 of G28 (observed) are estimated to the Pro-Rp A) and Pro-Sp B) position. These structures are obtained by modeling the nucleotides immediately surrounding the scissile phosphate in the hammerhead ribozyme (2OEU, Martick et al., 2008) into the crystal structure of the HDV ribozyme (3NKB), as described by Chen et al (2011).

APPENDIX B

Materials and Methods

The following synthetic DNA oligonucleotides were purchased from IDT:

Mod-seq RT primer:

AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC/iSp1

8/CACTCA/iSp18/TTCAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG

Mod-seq PCR FWD primer:

AATGATACGGCGACCACCGAGATCTACAC.

Mod-seq Indexed PCR Rev Primer:

CAAGCAGAAGACGGCATACGAGATxxxxxxGTGACTGGAGTTCAGACGTGTGC

TCTTCCG

(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, then diluted to OD 0.2 in SD broth supplemented with 1x amino acids and either mock, cisplatin, or azaplatin at final concentrations of 100 and 200 μM, in a final volume of 10 mL. Each sample was prepared in duplicate. Samples were grown at 37 °C for six hours while shaking, then spun at 4 °C, 4K RPM for 10 minutes to pellet.

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 sent for 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. 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) 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.

Reverse transcription primer extension analysis

Yeast was grown as described and treated with either mock or 200 μM cisplatin for six hours. RNA was extracted as described and reverse transcription was performed using Superscript II (Thermo Fisher) following manufacturer's instructions, using ^{32}P-labeled primers. The following primers were used for extension: CTCTTCCAGCCATAAGACCCC, which anneals 22 nt downstream of C1711 and CACCGGAGCCAGCAAAGG, which anneals 50 nt downstream of C1711. Dideoxy sequencing ladders were generated by adding 1 μL of 10 mM ddA, ddT, ddC, or ddG in their respective reactions. Samples were run on an 8% 19:1 mono:bis acrylamide gel.

Data Processing

Sequencing adapters were trimmed using CutAdapt (Martin et al., 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

Yeast was grown, treated with 0 or 200 μM cisplatin, and RNA was extracted as described above. RT primers corresponding with regions 1733-1753 and 1761-1779 of the yeast 25S rRNA were 5'end labeled with ^{32}P using PNK (Thermo Fisher). 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). dideoxy sequencing ladders were generated by reverse transcribing 3 μg of control RNA in the presence of 1 μL 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.

Supplementary Figures

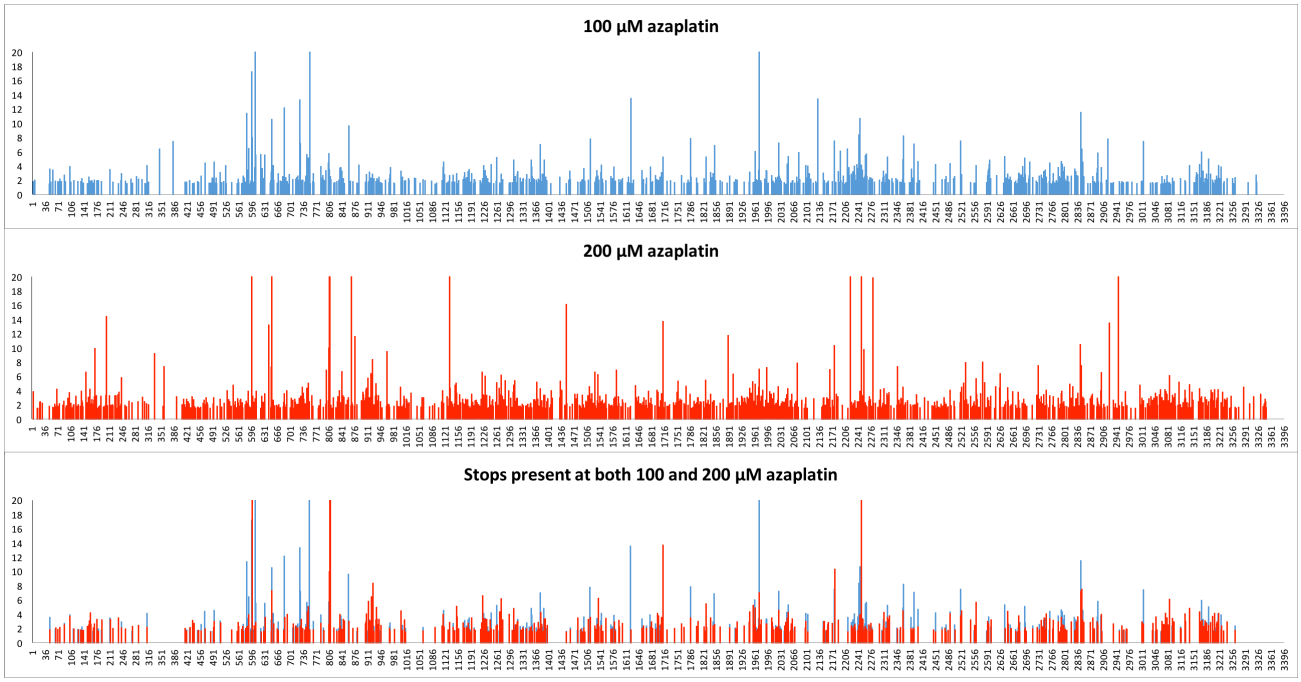


Figure B.1. Azaplatin stop sites on the yeast 25S rRNA. 100 µM, 200 µM and all matching stops.

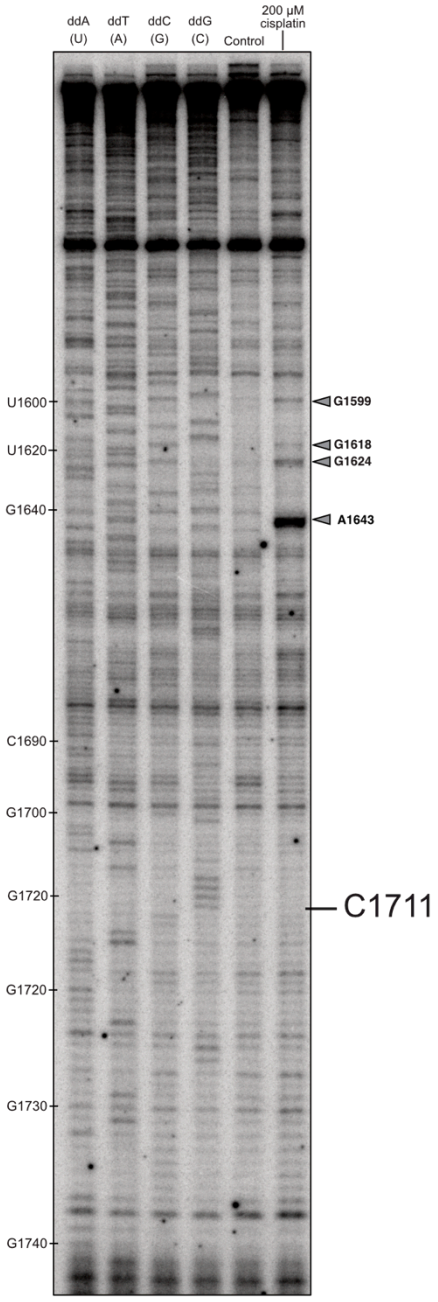


Figure B.2. Complete sequencing gel from Figure 3.9A

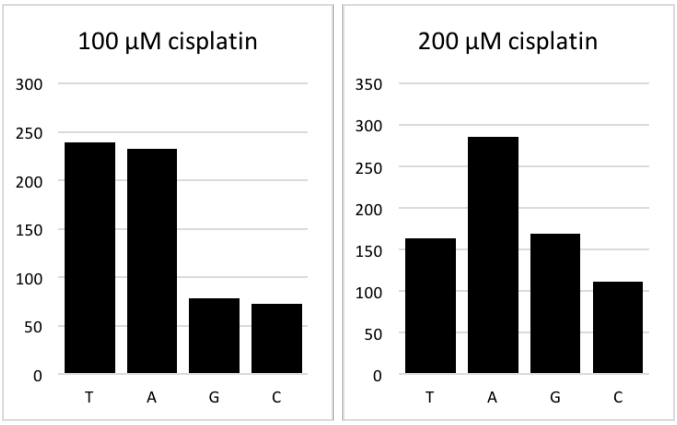


Figure B.3. Circligase preferentially ligates A and T residues in the 3' position. Use of Circligase to generate final templates for sequencing library preparation induces a strong bias toward A and T residues on the 3' end. This effect is less apparent at 200 µM cisplatin treatment.

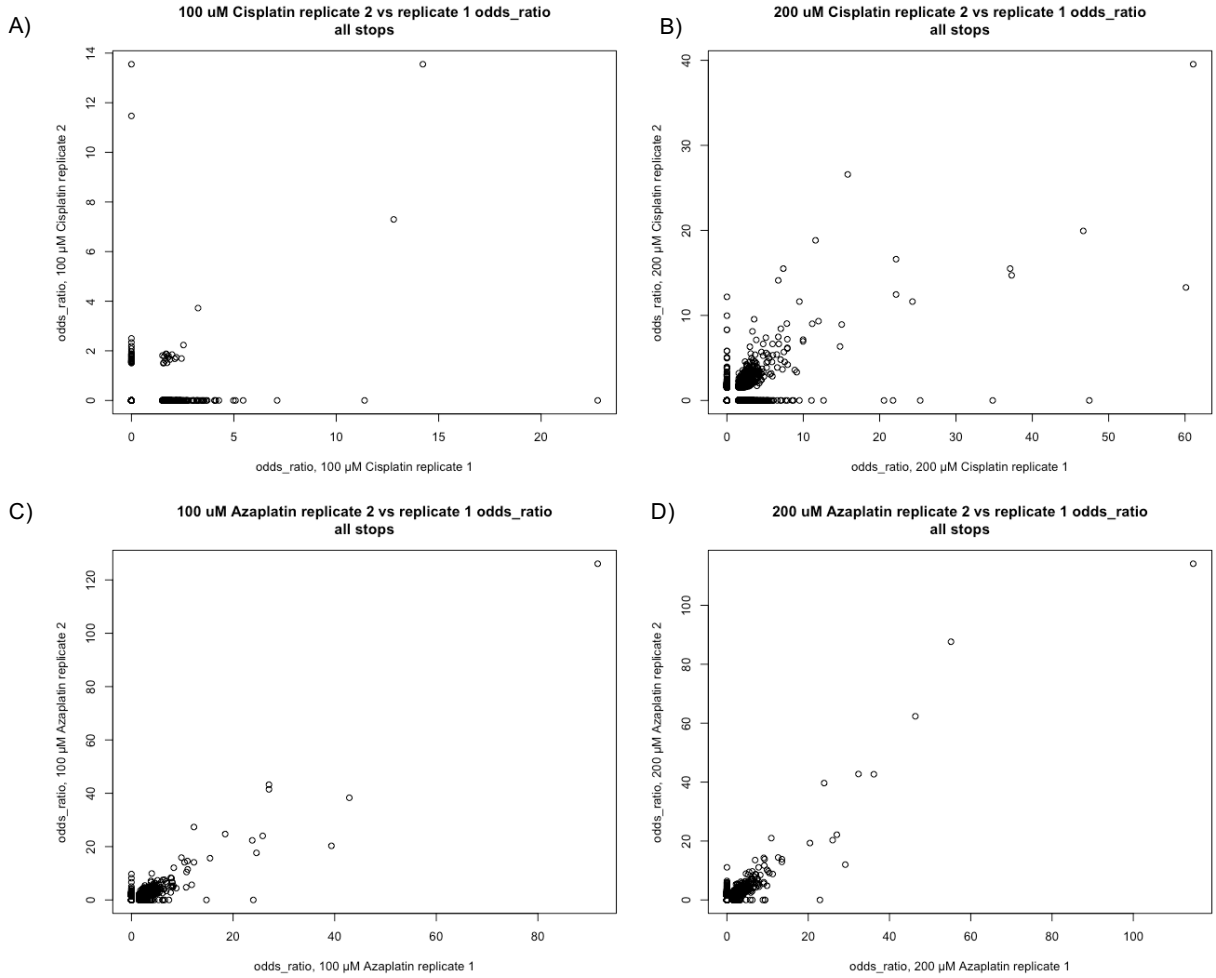


Figure B.4. Plot of fold-enrichment for each biological replicate by position compared against the same control. All stops greater than 1.5 fold-enriched over control are considered. Data was fit to a linear model to find R^2 values in order to assess correlation. A) 100 µM cisplatin (R^2 =0 .07). B) 200 µM cisplatin (R^2 =0 .42). C) 100 µM azaplatin (R^2 =0 .80). D) 200 µM azaplatin (R^2 =0 .85). Data points at 0 on the X and Y axes are stops present in one replicate but not the other. Gaps exhibited between these data points and the main data cluster are due to our chosen 1.5-fold minimum enrichment threshold.

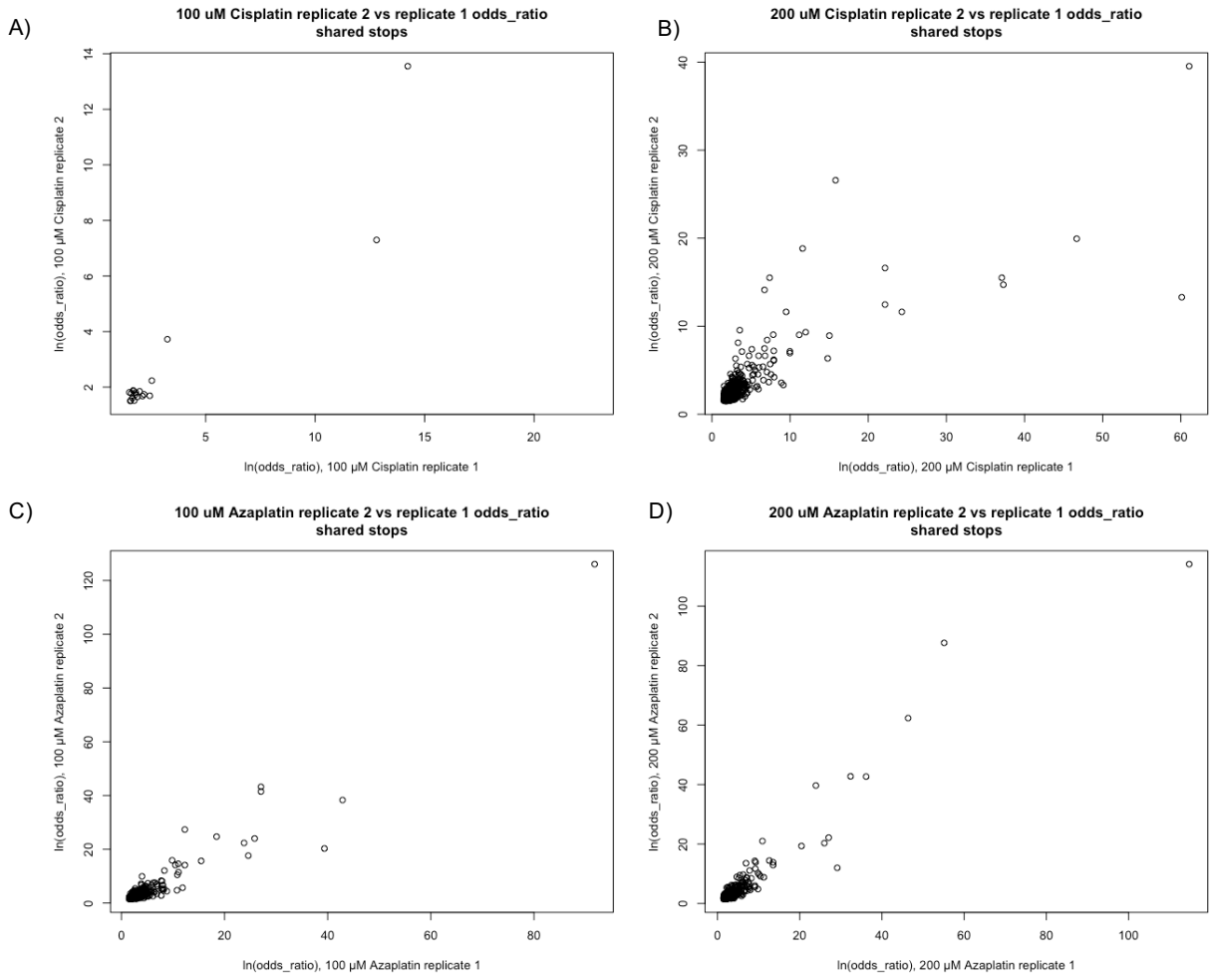


Figure B.5. Plot of fold-enrichment for each biological replicate by position compared against the same control. Only shared stops greater than 1.5 fold-enriched over control considered. A) 100 µM cisplatin (R^2 =0 .90). B) 200 µM cisplatin (R^2 =0 .66). C) 100 µM azaplatin (R^2 =0 .89). D) 200 µM azaplatin (R^2 =0 .92).

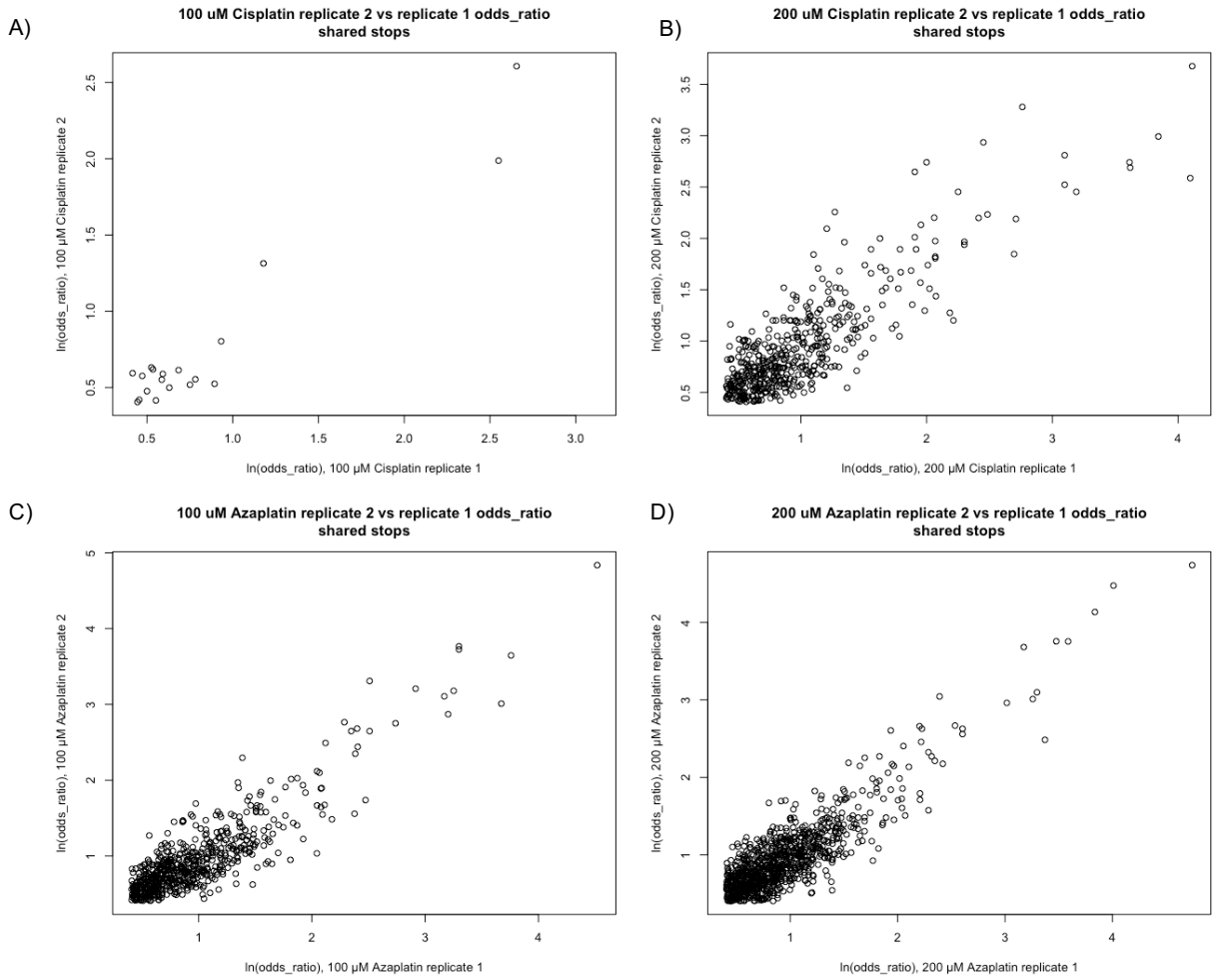


Figure B.6 – Plot of ln-transformed fold-enrichment for each biological replicate by position compared against the same control (data from figure B.5) demonstrate reproducible correlation between stop instensity and position. Only shared stops greater than 1.5-fold enriched over control considered. A) 100 µM cisplatin. B) 200 µM cisplatin. C) 100 µM azaplatin. D) 200 µM azaplatin.

Table B.1. Pt-seq stops for positions 1710-1720 on the 25S rRNA by treatment and replicate. Values presented are fold-enrichment over conrol (control replicate 2 for each experiment).

position	100 μM cisplatin		200 μM cisplatin		100 μM azaplatin		200 μM azaplatin	
on gene	**rep 1**	**rep 2**	**rep 1**	**rep 2**	**rep 1**	**rep 2**	**rep 1**	**rep 2**
1700	-	-	-	-	-	-	-	-
1701	-	-	-	-	-	-	-	-
1702	-	-	-	-	-	-	-	-
1703	-	-	-	-	-	-	1.8	1.6
1704	-	-	-	-	-	-	-	-
1705	1.7	-	-	-	1.8	1.7	2.5	2.6
1706	-	-	-	-	1.5	1.8	2.1	-
1707	2.4	-	-	-	2.0	2.4	4.0	3.1
1708	1.9	-	-	-	-	1.5	2.1	-
1709	1.5	-	-	-	-	-	-	-
1710	-	-	-	-	-	-	-	-
1711	14.2	13.6	15.8	26.6	12.3	14.1	13.5	13.8
1712	-	-	2.6	-	-	-	-	-
1713	-	-	-	-	-	-	-	1.8
1714	1.7	-	-	-	-	-	-	-
1715	1.6	-	2.5	2.8	-	-	1.6	-
1716	1.7	-	1.9	-	-	-	-	-
1717	-	-	-	-	-	-	2.1	2.6
1718	-	-	-	-	1.5	-	1.5	-
1719	-	-	-	-	-	-	-	-
1720	-	-	-	-	-	-	-	-

Table B.2. Pt-seq stops for positions 780-839 on the 25S rRNA by treatment and replicate. Values presented are fold-enrichment over conrol (control replicate 2 for each experiment).

position	100 μM cisplatin		200 μM cisplatin		100 μM azaplatin		200 μM azaplatin	
on gene	rep 1	rep 2	rep 1	rep 2	rep 1	rep 2	rep 1	rep 2
780	-	-	-	-	-	-	-	-
781	1.6	-	-	-	-	-	-	-
782	1.8	-	-	-	2.8	1.5	3.1	2.3
783	-	-	-	-	-	-	-	-
784	1.8	-	-	-	-	-	1.7	-
785	-	-	-	-	-	-	-	-
786	-	-	-	-	-	-	-	-
787	1.9	-	-	-	-	-	-	-
788	-	-	-	-	2.3	-	1.6	-
789	-	-	-	-	-	-	-	-
790	-	-	-	-	-	1.8	1.9	1.5
791	-	-	-	-	-	-	-	-
792	-	-	-	-	-	-	-	-
793	-	-	-	-	-	-	-	-
794	-	-	-	-	-	-	-	-
795	-	-	-	-	-	-	-	-
796	-	-	-	-	-	-	-	-
797	-	-	2.9	-	-	-	2.2	-
798	-	-	-	-	-	-	9.4	-
799	-	-	-	-	-	-	-	-
800	-	-	-	-	-	-	-	-
801	1.8	-	-	-	-	-	-	-
802	2.0	-	-	-	-	-	-	-
803	4.1	-	-	-	-	-	-	-
804	-	-	-	-	5.1	3.3	3.3	-
805	-	-	5.9	4.5	11.0	14.6	11.3	8.8
806	3.5	-	2.4	2.3	3.6	4.7	3.8	3.7
807	3.7	-	61.1	39.5	91.8	126.1	114.7	114.1
808	-	-	5.3	-	23.8	22.4	27.0	22.1
809	-	-	-	-	-	-	-	-
810	-	-	-	-	-	-	-	-
811	-	-	2.7	2.3	1.9	2.4	2.0	1.9
812	-	-	-	-	-	-	-	-
813	-	-	-	-	-	-	-	-
814	-	-	2.0	-	-	-	-	1.6
815	-	-	-	-	-	-	-	-
816	-	-	-	-	-	-	-	-
817	-	-	2.0	-	1.9	1.8	3.0	3.0
818	-	-	-	1.7	2.1	2.0	2.2	2.0
819	-	-	-	1.5	-	-	1.5	-
820	-	-	-	-	-	-	1.6	-
821	-	-	-	-	-	-	1.5	-
822	-	-	-	-	-	-	-	-
823	-	-	-	-	-	-	-	-
824	-	-	1.6	-	-	-	-	-
825	-	-	1.6	-	-	1.6	-	1.7
826	-	-	-	-	-	-	1.6	1.7
827	-	-	-	-	-	-	-	-
828	1.6	-	1.8	2.2	1.5	-	1.7	-
829	-	-	2.1	2.1	2.6	2.0	2.9	2.0
830	-	-	3.4	-	2.8	-	3.4	2.9
831	-	-	-	-	-	-	1.8	-
832	-	-	-	-	-	-	-	-
833	-	-	-	-	-	-	-	-
834	-	-	-	-	2.8	2.2	3.9	4.1
835	-	-	-	-	-	-	1.6	-
836	-	-	-	-	2.0	2.1	2.2	2.3
837	-	-	-	-	1.5	1.6	1.7	2.1
838	-	-	-	-	-	-	1.6	-
839	-	-	-	-	1.6	1.6	2.3	-

Table B.3. Counts of stops observed on all *S. cerevisiae* RNA at 100 µM cisplatin

100 µM Cisplatin

GeneName	Stops
15S_rRNA	23
21S_rRNA	16
RDN18-1	74
RDN18-2	74
RDN25-1	141
RDN25-2	141
RDN37-1	214
RDN37-2	214
RDN5-1	3
RDN5-2	2
RDN5-3	2
RDN5-4	2
RDN5-5	2
RDN5-6	3
RDN58-1	1
RDN58-2	1
SCR1	13
snR17a	1
snR56	1
snR9	1
tD(GUC)B	1
tD(GUC)D	1
tD(GUC)G1	1
tD(GUC)G2	1
tD(GUC)I1	1
tD(GUC)I2	1
tD(GUC)J1	1

GeneName	Stops
tD(GUC)J2	1
tD(GUC)J3	1
tD(GUC)J4	1
tD(GUC)K	1
tD(GUC)L1	1
tD(GUC)L2	1
tD(GUC)M	1
tD(GUC)N	1
tD(GUC)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	2
tE(UUC)P	1
tG(GCC)B	1
tG(GCC)C	1

GeneName	Stops
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
tG(GCC)P1	1
tG(GCC)P2	1
tL(UAA)B1	1
tL(UAA)B2	1
tL(UAA)D	1
tL(UAA)J	1
tL(UAA)K	1
tL(UAA)L	1
tL(UAA)N	1
YAR009C	1
YBL005W-B	1
YBR012W-B	1
YBR118W	1
YDR098C-B	1
YDR210C-D	1

GeneName	Stops
YDR261C-D	1
YDR316W-B	1
YDR365W-B	1
YER138C	1
YER160C	1
YGR038C-B	1
YGR161C-D	1
YHR214C-B	1
YJR027W	1
YJR029W	1
YLR035C-A	1
YLR154C-G	4
YLR157C-B	1
YML039W	1
YML045W	1
YMR045C	1
YMR050C	1
YNL054W-B	1
YNL284C-B	1
YOL103W-B	1
YOR142W-B	1
YPL257W-B	1
YPR080W	1
YPR137C-B	1
YPR158W-B	1

Table B.4. 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

Table B.5. Counts of stops observed on all *S. cerevisiae* RNA at 100 µM azaplatin

100 µM Azaplatin

GeneName	Stops
15S_rRNA	104
21S_rRNA	89
LSR1	31
NME1	1
Q0055	2
Q0250	1
RDN18-1	465
RDN18-2	465
RDN25-1	835
RDN25-2	835
RDN37-1	2155
RDN37-2	2155
RDN5-1	24
RDN5-2	25
RDN5-3	25
RDN5-4	25
RDN5-5	25
RDN5-6	24
RDN58-1	43
RDN58-2	43
RPR1	2
SCR1	69
snR10	1
snR11	1
snR128	6
snR13	1
snR14	4
snR17a	10
snR17b	2
snR189	1
snR19	8
snR190	3
snR24	1
snR3	5
snR30	4
snR31	1
snR32	2
snR34	1
snR37	10
snR40	1
snR43	3
snR44	3
snR47	2
snR48	1

GeneName	Stops
snR5	1
snR56	2
snR6	6
snR7-L	3
snR7-S	5
snR70	1
snR72	1
snR73	1
snR77	1
snR8	1
snR81	1
snR9	9
tA(UGC)A	2
tA(UGC)E	2
tA(UGC)G	2
tA(UGC)L	2
tA(UGC)O	2
tD(GUC)B	5
tD(GUC)D	5
tD(GUC)G1	5
tD(GUC)G2	5
tD(GUC)I1	5
tD(GUC)I2	5
tD(GUC)J1	5
tD(GUC)J2	5
tD(GUC)J3	5
tD(GUC)J4	5
tD(GUC)K	5
tD(GUC)L1	5
tD(GUC)L2	5
tD(GUC)M	5
tD(GUC)N	5
tD(GUC)O	5
tE(CUC)D	5
tE(CUC)I	5
tE(UUC)B	6
tE(UUC)C	6
tE(UUC)E1	6
tE(UUC)E2	6
tE(UUC)E3	6
tE(UUC)G1	6
tE(UUC)G2	6
tE(UUC)G3	6
tE(UUC)I	6

GeneName	Stops
tE(UUC)J	6
tE(UUC)K	6
tE(UUC)L	6
tE(UUC)M	6
tE(UUC)P	6
tG(CCC)D	2
tG(UCC)G	1
tG(UCC)N	1
tG(UCC)O	1
tR(CCG)L	1
tS(UGA)E	1
tS(UGA)I	1
tS(UGA)P	1
YAR009C	5
YAR010C	3
YBL005W-B	3
YBR012W-A	2
YBR012W-B	5
YDR098C-A	2
YDR098C-B	11
YDR210C-C	2
YDR210C-D	6
YDR261C-C	2
YDR261C-D	12
YDR278C	6
YDR316W-A	2
YDR316W-B	10
YDR365W-A	2
YDR365W-B	13
YEL021W	1
YER137C-A	2
YER138C	13
YER159C-A	2
YER160C	7
YGL008C	3
YGL088W	1
YGR027W-A	3
YGR027W-B	12
YGR038C-A	2
YGR038C-B	14
YGR161C-C	2
YGR161C-D	7
YGR192C	2
YHR214C-B	7

GeneName	Stops
YHR214C-C	2
YJR026W	2
YJR027W	12
YJR028W	2
YJR029W	10
YLR035C-A	3
YLR044C	1
YLR110C	5
YLR154C-G	28
YLR157C-A	2
YLR157C-B	13
YLR227W-A	2
YLR227W-B	7
YLR256W-A	2
YML039W	13
YML040W	2
YML045W	9
YML045W-A	2
YMR045C	5
YMR050C	8
YMR051C	2
YNL054W-A	2
YNL054W-B	8
YNL284C-A	1
YNL284C-B	5
YOL103W-A	2
YOL103W-B	12
YOR142W-A	2
YOR142W-B	9
YPL257W-A	2
YPL257W-B	15
YPR137C-A	3
YPR137C-B	12
YPR158C-C	2
YPR158C-D	7
YPR158W-A	2
YPR158W-B	9

Table B.6. Counts of stops observed on all *S. cerevisiae* RNA at 200 µM azaplatin

200 µM Azaplatin

GeneName	Stops	GeneName	Stops	GeneName	Stops	GeneName	Stops
15S_rRNA	171	snR70	1	tG(GCC)O1	4	YER137C-A	4
21S_rRNA	213	snR8	1	tG(GCC)O2	4	YER138C	15
LSR1	47	snR86	1	tG(GCC)P1	4	YER159C-A	2
Q0045	1	snR9	8	tG(GCC)P2	4	YER160C	7
Q0050	6	tA(UGC)A	2	tG(UCC)G	2	YGL008C	2
Q0055	5	tA(UGC)E	2	tG(UCC)N	2	YGL088W	3
Q0070	1	tA(UGC)G	2	tG(UCC)O	2	YGR027W-A	3
Q0110	2	tA(UGC)L	2	tG(UCC)Q	1	YGR027W-B	9
Q0115	1	tA(UGC)O	2	tH(GUG)E1	1	YGR038C-A	2
Q0130	7	tD(GUC)B	12	tH(GUG)E2	1	YGR038C-B	9
Q0275	1	tD(GUC)D	12	tH(GUG)G1	1	YGR161C-C	2
RDN18-1	579	tD(GUC)G1	12	tH(GUG)G2	1	YGR161C-D	7
RDN18-2	579	tD(GUC)G2	12	tH(GUG)H	1	YGR192C	5
RDN25-1	1188	tD(GUC)I1	12	tH(GUG)K	1	YHR174W	4
RDN25-2	1188	tD(GUC)I2	12	tH(GUG)M	1	YHR214C-B	6
RDN37-1	1839	tD(GUC)J1	12	tI(UAU)D	3	YHR214C-C	1
RDN37-2	1839	tD(GUC)J2	12	tI(UAU)L	3	YJR026W	4
RDN5-1	20	tD(GUC)J3	12	tL(CAA)A	1	YJR027W	14
RDN5-2	16	tD(GUC)J4	12	tL(CAA)C	2	YJR028W	2
RDN5-3	16	tD(GUC)K	12	tL(CAA)D	1	YJR029W	13
RDN5-4	16	tD(GUC)L1	12	tL(CAA)G1	2	YKL060C	4
RDN5-5	16	tD(GUC)L2	12	tL(CAA)G2	2	YLR035C-A	6
RDN5-6	20	tD(GUC)M	12	tL(CAA)G3	2	YLR044C	1
RDN58-1	51	tD(GUC)N	10	tL(CAA)K	2	YLR110C	9
RDN58-2	51	tD(GUC)O	12	tL(CAA)L	1	YLR157C-A	4
SCR1	70	tD(GUC)Q	1	tL(CAA)M	2	YLR157C-B	14
snR10	3	tE(CUC)D	4	tL(CAA)N	2	YLR227W-A	4
snR11	2	tE(CUC)I	4	tL(GAG)G	1	YLR227W-B	9
snR128	2	tE(UUC)B	11	tQ(UUG)B	2	YLR256W-A	4
snR13	1	tE(UUC)C	11	tV(CAC)D	2	YML039W	15
snR14	3	tE(UUC)E1	11	tV(CAC)H	2	YML040W	4
snR17a	9	tE(UUC)E2	11	tW(CCA)G1	2	YML045W	11
snR17b	2	tE(UUC)E3	11	tW(CCA)G2	2	YML045W-A	1
snR19	7	tE(UUC)G1	11	tW(CCA)J	2	YML123C	6
snR190	5	tE(UUC)G2	11	tW(CCA)K	2	YMR045C	4
snR191	1	tE(UUC)G3	11	tW(CCA)M	2	YMR050C	12
snR24	1	tE(UUC)I	11	tW(CCA)P	2	YMR051C	2
snR3	8	tE(UUC)J	11	YAL038W	3	YNL054W-A	2
snR30	22	tE(UUC)K	11	YAR009C	7	YNL054W-B	7
snR32	2	tE(UUC)L	11	YAR010C	2	YNL284C-A	2
snR34	3	tE(UUC)M	8	YBL005W-B	5	YNL284C-B	8
snR35	1	tE(UUC)P	11	YBR012W-A	2	YOL013W-A	1
snR37	12	tE(UUC)Q	3	YBR012W-B	7	YOL103W-A	3
snR39	1	tG(CCC)D	4	YBR118W	3	YOL103W-B	15
snR40	1	tG(CCC)O	1	YCR012W	1	YOL154W	1
snR42	1	tG(GCC)B	4	YDR098C-A	2	YOR142W-A	2
snR43	2	tG(GCC)C	4	YDR098C-B	9	YOR142W-B	10
snR44	1	tG(GCC)D1	4	YDR210C-C	2	YPL257W-A	4
snR45	1	tG(GCC)D2	4	YDR210C-D	9	YPL257W-B	14
snR47	4	tG(GCC)E	4	YDR261C-C	3	YPR080W	4
snR48	1	tG(GCC)F1	4	YDR261C-D	11	YPR137C-A	2
snR49	1	tG(GCC)F2	4	YDR278C	8	YPR137C-B	11
snR5	2	tG(GCC)G1	4	YDR316W-A	2	YPR158C-C	4
snR56	1	tG(GCC)G2	4	YDR316W-B	9	YPR158C-D	11
snR6	9	tG(GCC)J1	4	YDR365W-A	4	YPR158W-A	2
snR7-L	8	tG(GCC)J2	4	YDR365W-B	11	YPR158W-B	10
snR7-S	6	tG(GCC)M	4	YEL021W	8		

APPENDIX C

Methods

Yeast BY4741 cells were treated with either 0, 100, or 200 μM cisplatin or azaplatin and sequencing libraries prepared as in Plakos and DeRose (2016). Reads were aligned to the yeast genome with up to one allowed mismatch using Bowtie **(**bowtie -v 3 -a --best --strata --chunkmbs 500 -p 16 -t –S) (v1.1.2, Langmead et al., 2009), then alignments within genome features counted using HTSeq-count (Anders et al., 2005). This data were imported into R and differential expression analysis performed with the DESEQ2 R package (Love et al., 2014). Genes with a p_{adj} of >0.1 were considered for further analysis. Differential expression results were plotted with R and Excel. Gene names were translated from SGD Gene ID to human homologs, where available, using SGD's YeastMine (Cherry et al., 2012). KEGG pathway analysis and GO-term analysis were performed using the gage R package (Luo et al., 2009). KEGG and GO data was sourced as in Supplementary Table 1.

Supplementary Information

Table C.1. KEGG and GO analysis annotation and reference information. Data printed from gage (Luo et al., 2009)

OrgDb object:
DBSCHEMAVERSION: 2.1
Db type: OrgDb
Supporting package: AnnotationDbi
DBSCHEMA: YEAST_DB
ORGANISM: Saccharomyces cerevisiae
SPECIES: Yeast
YGSOURCENAME: Yeast Genome
YGSOURCEURL: http://downloads.yeastgenome.org/
YGSOURCEDATE: 12-Mar-2016
CENTRALID: ORF
TAXID: 559292
KEGGSOURCENAME: KEGG GENOME
KEGGSOURCEURL: ftp://ftp.genome.jp/pub/kegg/genomes
KEGGSOURCEDATE: 2011-Mar15
GOSOURCENAME: Gene Ontology
OSOURCEURL: ftp://ftp.geneontology.org/pub/go/godatabase/archive/latest-lite/
GOSOURCEDATE: 20160305
EGSOURCEDATE: 2016-Mar14
EGSOURCENAME: Entrez Gene
EGSOURCEURL: ftp://ftp.ncbi.nlm.nih.gov/gene/DATA
ENSOURCEDATE: 2016-Mar9
ENSOURCENAME: Ensembl
ENSOURCEURL: ftp://ftp.ensembl.org/pub/current_fasta
UPSOURCENAME: Uniprot
UPSOURCEURL: http://www.UniProt.org/
UPSOURCEDATE: Wed Mar 23 16:02:02 2016

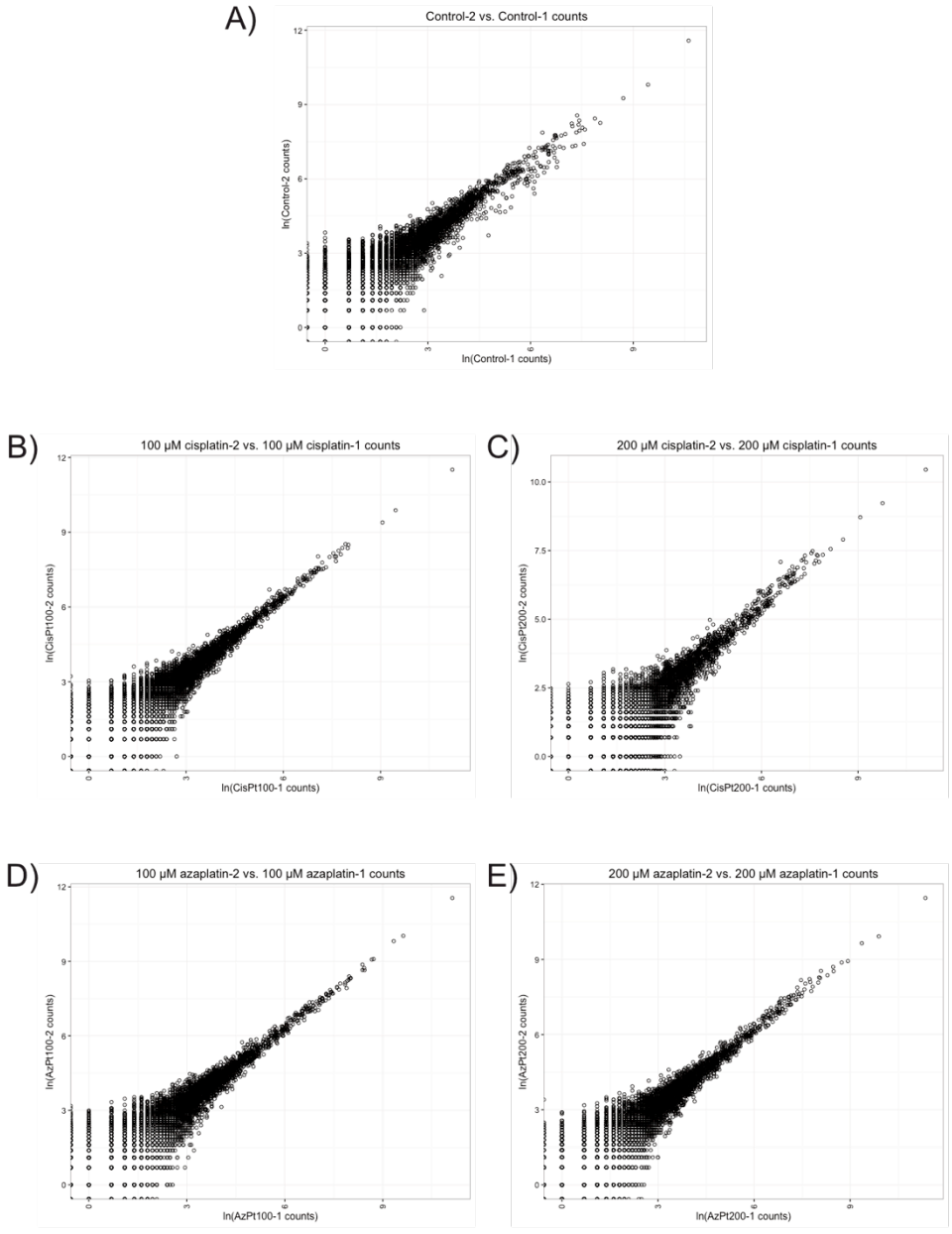


Figure C.1. Plots of natural-log transformed raw counts of alignments on all genes except rRNA and tRNA genes for A) Controls. B) 100 µM cisplatin treatment. C) 200 µM cisplatin treatment. D) 100 µM azaplatin treatment. E) 200 µM azaplatin treatment. Data is strongly linear for all samples (r-squared >0.9, not shown).

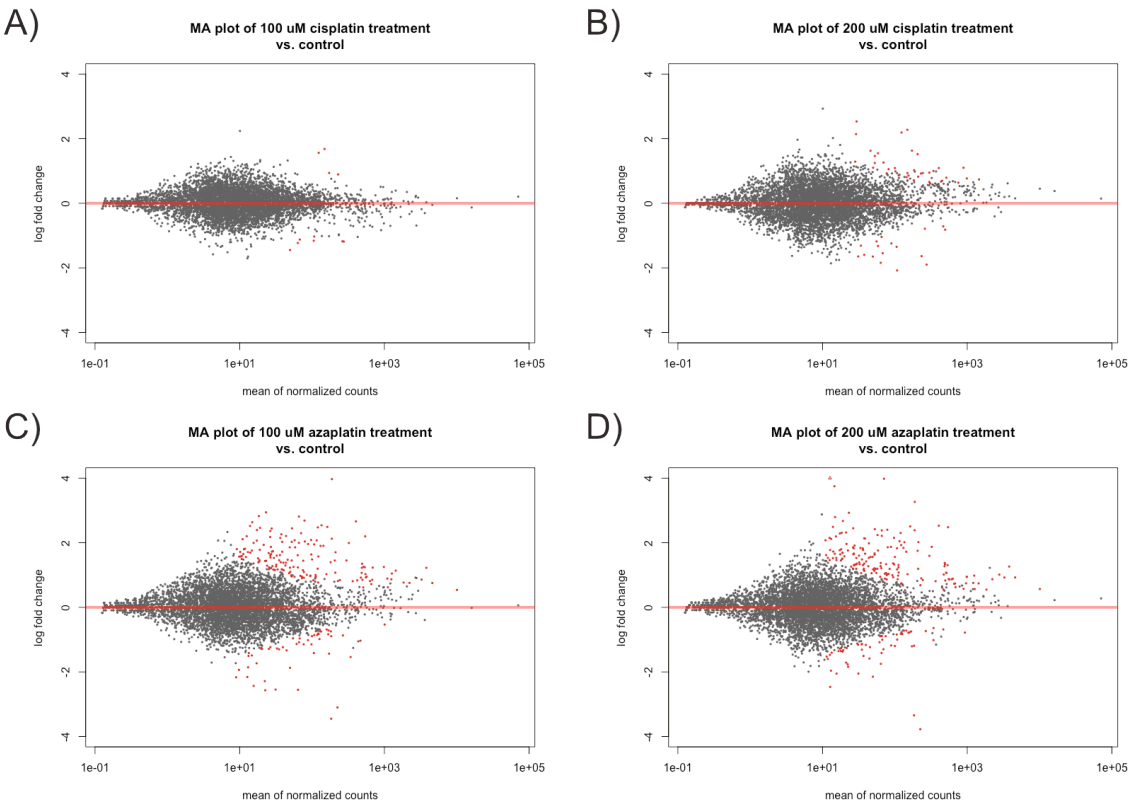


Figure C.2. MA plots of A) 100 µM cisplatin vs. control. B) 200 µM cisplatin vs. control. C) 100 µM azaplatin vs control. D) 200 µM azaplatin vs control. Points plotted are representative of genes with nonzero read coverage. Gray points exhibit adjusted p-values greater than 0.1, red points less than 0.1.

Table C.2. List of genes significantly differentially expressed under each treatment condition. SGD ID, log2fold enrichment, and p_{adj} are given for each.

100 µM cisplatin

SGD ID	log2(FC)	padj
YGR180C	1.7	1.03E-05
YJL026W	1.6	1.03E-05
YNL327W	-1.0	4.98E-03
YHR143W	-1.0	2.50E-02
YJL078C	-1.1	9.95E-02
YGL028C	-1.2	2.06E-02
snR67	-1.2	2.66E-02
YLR286C	-1.2	1.87E-03
YJL153C	-1.2	6.50E-02
YER124C	-1.4	9.97E-03

200 µM cisplatin

SGD ID	log2(FC)	padj
YOR382W	2.5	1.09E-06
YGR180C	2.3	2.14E-11
YJL026W	2.2	3.88E-12
YOR383C	2.1	1.50E-03
YOL155C	1.6	4.66E-09
YBR072W	1.6	7.38E-03
YDR001C	1.5	2.78E-03
YNL160W	1.5	8.12E-07
YMR011W	1.5	7.26E-03
YLR034C	1.3	9.24E-02
YHR092C	1.3	2.78E-02
YER150W	1.2	8.20E-02
YPR160W	1.1	8.20E-02
YDR077W	1.1	8.48E-07
YLL026W	1.1	1.74E-02
YEL060C	1.1	3.44E-02
YJL171C	1.1	2.92E-02
YER070W	1.0	4.58E-03
YPR108W-A	1.0	3.97E-02
YGR060W	1.0	2.95E-02
YJR109C	-1.0	9.98E-02
YBR010W	-1.2	4.71E-02
YNL066W	-1.2	1.23E-02
YBL003C	-1.3	9.24E-02
YKL164C	-1.3	1.74E-02
YGL028C	-1.3	2.75E-03
YJL078C	-1.5	3.22E-03
YBR158W	-1.6	1.32E-02
YNL327W	-1.6	1.25E-08
YDL128W	-1.6	1.41E-02
YER124C	-1.6	1.67E-03
YJL153C	-1.8	9.44E-04
YLR286C	-1.9	2.69E-09
YHR143W	-2.1	5.66E-09

100 µM azaplatin

SGD ID	log2(FC)	padj
YML100W	4.0	1.78E-32
YOL084W	2.9	3.46E-07
YGR249W	2.8	3.07E-06
YDR074W	2.8	3.50E-11
YBR169C	2.7	5.25E-12
YLL026W	2.7	2.58E-16
YNL194C	2.6	4.45E-05
YDR258C	2.5	4.43E-11
YLR327C	2.5	3.20E-04
YER103W	2.5	1.08E-10
YKR093W	2.5	1.78E-14
YLR177W	2.5	8.42E-05
YER150W	2.4	9.14E-08
YJL141C	2.4	1.93E-05
YEL011W	2.4	1.15E-04
YNL077W	2.4	4.51E-05
YGR065C	2.3	2.61E-07
snR60	2.3	3.48E-06
YJL165C	2.2	8.75E-07
YMR316W	2.2	2.06E-03
YOL081W	2.2	1.32E-17
YOL016C	2.2	1.22E-04
YBR101C	2.1	1.04E-05
YCR021C	2.1	1.77E-03
YGL179C	2.1	4.64E-03
YCL040W	2.1	5.82E-09
YPL014W	2.1	1.85E-06
YDR171W	2.1	1.23E-03
YGR161C	2.1	2.71E-03
YGL006W	2.0	3.49E-08
YDR216W	2.0	7.58E-06
YBR126C	2.0	5.40E-05
YCR005C	2.0	3.90E-07
YER053C	2.0	4.01E-03
YPL240C	2.0	2.12E-09
YOR027W	2.0	1.12E-06
YGR142W	1.9	4.19E-03
YBR214W	1.9	1.61E-03
YGR250C	1.9	2.61E-08
YJL042W	1.9	1.18E-07
YER158C	1.9	1.43E-03
YFR053C	1.9	3.42E-04
YDR001C	1.9	1.43E-05
YOR054C	1.8	1.87E-02
YPR158W	1.8	3.09E-02
YDR069C	1.8	2.55E-02
YPL247C	1.8	8.10E-03

100 µM azaplatin

SGD ID	log2(FC)	padj
YCL039W	1.8	2.17E-02
YPR154W	1.7	3.21E-03
YOR219C	1.7	2.12E-03
YPR184W	1.7	6.86E-04
YER033C	1.7	4.38E-02
YOR113W	1.7	2.87E-02
YGR138C	1.7	2.64E-09
YOR347C	1.7	3.68E-02
YGR130C	1.7	7.96E-03
YOR275C	1.7	5.77E-02
YNL006W	1.6	1.67E-02
YGR211W	1.6	4.36E-04
YML128C	1.6	6.02E-03
YNL007C	1.6	2.31E-04
YOR299W	1.6	2.09E-02
YGL037C	1.6	3.07E-03
YBR006W	1.6	5.82E-02
YBR054W	1.6	4.65E-05
YGR197C	1.6	7.14E-02
YIL017C	1.6	7.41E-02
YNL144C	1.6	6.00E-02
YIL056W	1.5	1.71E-02
YDR214W	1.5	1.09E-02
YOR178C	1.5	6.35E-02
YMR135C	1.5	9.66E-02
YEL060C	1.5	1.45E-04
YBR072W	1.5	1.01E-02
YMR105C	1.5	6.14E-02
YOR298C-A	1.5	1.67E-02
YGR032W	1.5	1.65E-08
YGL197W	1.5	2.12E-03
YNL160W	1.4	1.12E-06
YOR273C	1.4	1.93E-03
YJR059W	1.4	5.82E-02
YKL198C	1.4	3.40E-02
YBR132C	1.4	5.38E-02
YPL250C	1.4	2.26E-02
Q0075	1.4	2.13E-06
YLR178C	1.4	7.58E-02
YPR036W-A	1.4	7.18E-03
YGL178W	1.4	1.83E-02
YDR277C	1.4	9.85E-02
snR80	1.4	1.62E-06
YLR258W	1.3	3.98E-02
YFR044C	1.3	2.22E-02
YDR251W	1.3	8.40E-02
YBL101C	1.3	5.48E-02

Table C.2. (continued)

100 µM azaplatin		
SGD ID	log2(FC)	padj
YLR438W	1.3	7.68E-02
snR5	1.3	1.56E-07
YKL035W	1.3	1.23E-02
Q0045	1.2	5.82E-09
YAL005C	1.2	4.21E-07
YHL027W	1.2	8.06E-02
YJL153C	1.2	4.10E-03
Q0085	1.2	2.33E-06
YPL004C	1.2	1.87E-02
snR82	1.2	1.27E-04
snR19	1.2	1.25E-08
Q0080	1.2	9.55E-03
YOL087C	1.2	9.66E-02
YFL016C	1.2	6.36E-02
YGR023W	1.2	7.10E-02
YLR375W	1.2	9.66E-02
YMR261C	1.2	3.37E-02
YDR475C	1.2	9.37E-02
YJL082W	1.2	9.43E-02
YPR117W	1.1	1.19E-02
YMR250W	1.1	4.83E-02
Q0130	1.1	1.25E-08
YPR160W	1.1	4.06E-02
YJL005W	1.1	9.55E-03
TLC1	1.1	7.35E-04
snR17b	1.1	5.00E-04
YBR287W	1.1	5.82E-02
Q0275	1.0	6.22E-07
Q0105	1.0	5.87E-05
Q0060	1.0	3.84E-03
YDL070W	1.0	1.90E-02
YPR156C	1.0	8.67E-03
YML059C	1.0	7.68E-02
YNL268W	1.0	1.63E-02
YOR317W	1.0	5.10E-02
YDR343C	1.0	3.55E-02
YOR344C	1.0	8.93E-02
YGR175C	-1.0	8.59E-02
YPL012W	-1.0	4.06E-02
YLR058C	-1.0	7.66E-03
YER110C	-1.0	1.23E-02
YER091C	-1.0	1.04E-05
YKL006W	-1.1	1.21E-02
YKL101W	-1.1	2.14E-02
YPR010C	-1.1	3.90E-03
YLR432W	-1.1	1.13E-02
YER006W	-1.2	9.30E-02

100 µM azaplatin		
SGD ID	log2(FC)	padj
YJL050W	-1.2	5.01E-02
YDL167C	-1.3	9.83E-02
YGR103W	-1.3	7.95E-02
YDR507C	-1.3	9.74E-02
YNL141W	-1.3	6.28E-02
YMR260C	-1.3	9.43E-02
YFL004W	-1.3	5.13E-02
YJL012C	-1.3	1.23E-02
YJL109C	-1.4	2.84E-05
YJL074C	-1.4	9.83E-02
YOL155C	-1.4	3.98E-06
YOR243C	-1.5	6.00E-02
YAR050W	-1.5	9.85E-02
YPR108W-A	-1.5	4.17E-05
YML022W	-1.7	1.90E-02
YDL227C	-1.9	1.37E-02
YPL019C	-1.9	1.53E-04
YNL250W	-1.9	1.90E-02
YHR215W	-2.2	2.64E-03
YAR071W	-2.2	4.64E-03
YHR216W	-2.3	1.23E-04
YBR296C	-2.4	5.84E-04
YDR019C	-2.5	2.38E-06
YMR189W	-2.6	4.09E-09
YAL044C	-2.6	4.24E-05
YML123C	-3.1	9.63E-24
YEL021W	-3.4	8.23E-32

200 µM azaplatin		
SGD ID	log2(FC)	padj
YNR056C	4.6	2.11E-13
YGR065C	4.0	3.19E-23
YNR058W	3.7	3.01E-09
YML100W	3.3	6.25E-22
YOL084W	2.9	3.90E-07
YNL194C	2.8	1.15E-05
YLR177W	2.7	1.20E-05
YLL026W	2.5	1.13E-14
YER103W	2.5	6.90E-11
YLR327C	2.5	3.59E-04
YOL081W	2.5	7.30E-23
YNL077W	2.4	1.95E-05
YER150W	2.4	1.37E-07
YBR169C	2.4	1.23E-09
YCR021C	2.4	2.29E-04
YCR005C	2.4	3.07E-10
YGR286C	2.3	1.62E-04
YGR249W	2.3	3.53E-04
YKR093W	2.3	1.72E-12
YEL011W	2.3	2.90E-04
YPL014W	2.2	1.70E-07
YDR074W	2.2	5.97E-07
YPR157W	2.2	6.62E-04
YCL040W	2.2	1.23E-09
YMR316W	2.1	2.55E-03
snR60	2.1	1.77E-05
YBR126C	2.1	1.75E-05
YDL222C	2.1	4.48E-03
YGL179C	2.1	5.15E-03
YDR171W	2.0	1.30E-03
YGR138C	2.0	2.39E-13
YDR258C	2.0	5.30E-07
YJL165C	2.0	1.87E-05
YBR214W	2.0	1.17E-03
YBR054W	1.9	1.70E-07
YJL141C	1.9	2.06E-03
YPL240C	1.9	1.93E-08
YOR027W	1.9	4.67E-06
YGL006W	1.9	8.28E-07
YOL016C	1.9	1.98E-03
YJL042W	1.8	3.57E-07
YGR023W	1.8	2.94E-04
YGR142W	1.8	7.98E-03
YER158C	1.8	1.85E-03
YML128C	1.8	9.70E-04
YLR178C	1.8	6.46E-03
YOL052C-A	1.8	1.98E-02

Table C.2. (continued)

200 µM azaplatin		
SGD ID	log2(FC)	padj
YBR101C	1.8	6.65E-04
YGR161C	1.8	1.58E-02
YKL198C	1.8	2.99E-03
YDR001C	1.8	5.61E-05
YNL160W	1.8	9.65E-10
YOR347C	1.7	2.59E-02
YGR250C	1.7	8.28E-07
YMR105C	1.7	1.57E-02
YPL247C	1.7	1.30E-02
YGL197W	1.7	1.83E-04
YNL006W	1.7	1.30E-02
YPR184W	1.7	1.17E-03
YCL039W	1.7	3.17E-02
YIL056W	1.6	7.15E-03
YLR350W	1.6	7.01E-03
YKR046C	1.6	2.25E-02
YDR216W	1.6	8.19E-04
YOR275C	1.6	5.60E-02
YLR438W	1.6	8.29E-03
YGR032W	1.6	2.78E-10
YLL023C	1.6	2.76E-02
YFR053C	1.6	4.75E-03
YBR072W	1.6	4.53E-03
YEL060C	1.6	5.35E-05
YPR036W-A	1.5	1.55E-03
YDR277C	1.5	3.75E-02
YMR110C	1.5	7.42E-02
YLR375W	1.5	8.89E-03
YJL116C	1.5	5.78E-02
YBR132C	1.5	2.79E-02
YNL305C	1.5	3.29E-02
YOR298C-A	1.5	1.24E-02
YER053C	1.5	6.96E-02
YJR008W	1.5	7.16E-02
YOR178C	1.5	7.43E-02
YGL178W	1.4	1.20E-02
YLR258W	1.4	2.15E-02
snR82	1.4	1.01E-05
YJL153C	1.4	8.19E-04
YDR343C	1.4	3.78E-04
YFL014W	1.4	2.15E-02
YDR214W	1.4	3.12E-02
YOR273C	1.3	4.94E-03
YGL037C	1.3	2.23E-02
snR80	1.3	2.56E-06
YGR211W	1.3	7.91E-03
YDR475C	1.3	3.17E-02

200 µM azaplatin		
SGD ID	log2(FC)	padj
YOR219C	1.3	3.55E-02
YDR342C	1.3	9.44E-04
YDR270W	1.3	4.99E-02
YOL060C	1.3	1.98E-02
YDL020C	1.3	6.56E-02
YAL005C	1.3	8.41E-08
YNL007C	1.3	7.77E-03
YFR044C	1.3	2.32E-02
snR17b	1.3	1.20E-05
snR19	1.3	2.59E-09
YBR208C	1.3	9.94E-02
YPR156C	1.2	3.46E-04
YER088C	1.2	3.85E-03
YOR344C	1.2	1.12E-02
YLR219W	1.2	7.16E-02
YJR091C	1.2	8.16E-03
YPR042C	1.2	4.45E-02
TLC1	1.2	9.59E-05
YBR105C	1.2	7.61E-02
YDL070W	1.2	2.27E-03
YDL124W	1.2	2.63E-02
YPR160W	1.2	1.96E-02
YBR067C	1.2	1.25E-03
YPL250C	1.2	8.21E-02
YDR297W	1.1	3.22E-02
YCR061W	1.1	6.24E-02
YBR068C	1.1	2.27E-03
YJL052W	1.1	2.79E-06
YKL035W	1.1	4.51E-02
snR5	1.1	1.77E-05
YPR117W	1.1	2.26E-02
YOL059W	1.1	9.25E-02
YKL029C	1.1	4.87E-02
YLR120C	1.0	2.13E-02
YJL005W	1.0	1.48E-02
YJL171C	1.0	1.31E-02
YDL039C	1.0	2.35E-03
snR81	1.0	6.65E-05
snR128	1.0	9.98E-02
YDR055W	1.0	4.39E-02
YPR149W	1.0	1.11E-02
YDL048C	1.0	1.87E-02
RPR1	1.0	7.46E-05
YDR186C	1.0	9.75E-02
Q0110	-1.0	4.02E-02
YEL065W	-1.0	4.73E-02
YPR108W-A	-1.0	1.87E-02

200 µM azaplatin		
SGD ID	log2(FC)	padj
YKL101W	-1.1	2.93E-02
YPR069C	-1.1	7.89E-02
YFL004W	-1.1	9.98E-02
YFL018C	-1.1	8.49E-02
YDR496C	-1.2	7.89E-02
YJL050W	-1.2	6.30E-02
YGR234W	-1.2	6.92E-03
YER070W	-1.2	3.73E-04
YPL043W	-1.2	9.92E-02
YMR229C	-1.2	3.69E-03
YOL155C	-1.2	1.94E-04
YHR141C	-1.2	1.76E-02
YER072W	-1.3	8.49E-02
YGR098C	-1.3	9.81E-02
YGR231C	-1.3	9.92E-02
YDL133C-A	-1.3	1.83E-03
YBR034C	-1.3	2.59E-02
YKL172W	-1.3	9.32E-02
YHR052W	-1.4	5.88E-02
YOL049W	-1.4	8.23E-02
YJL033W	-1.4	6.30E-02
YJL074C	-1.4	9.25E-02
YDL184C	-1.4	1.71E-03
YPL118W	-1.5	7.41E-02
YDR194C	-1.5	6.69E-03
YPL211W	-1.5	7.02E-02
YJL012C	-1.6	1.51E-03
YDL153C	-1.6	4.85E-02
YKL009W	-1.7	1.37E-02
YMR189W	-1.8	8.24E-05
YHR216W	-1.8	5.15E-03
YAL044C	-1.8	6.80E-03
YDL227C	-2.0	6.92E-03
YDR019C	-2.0	2.33E-04
YBR296C	-2.1	5.54E-03
YPL019C	-2.1	1.20E-05
YHR215W	-2.5	3.59E-04
YEL021W	-3.3	1.00E-29
YML123C	-3.8	1.09E-30

APPENDIX D

Materials and Methods

Oligonucleotides used were purchased from IDT and are as follows:

5' hexynyl functionalized oligonucleotide (HEX):

5'-/5Hexynyl/TTTTTTTCTGTAGGCACCATCAAT-3'

5' Biotinylated hexynyl complement (BIO):

5'-/5Biosg/TTTTTTTTTATTGATGGTGCCTACAG-3'

Nonspecific blocking oligonucleotide (BLOCK):

5'-

CAAGCAGAAGACGGCATACGAGATGTCATGTACTGGAGTTCAGACGTGTGCT

CTTCCG-3'

DNA Hairpin containing a preferred Pt(II) binding site (HP):

5'-TATGGTATTTTTATACCATA -3'

Oligonucleotides were suspended in nuclease free H_2O to a final concentration of 1 mM, aliquoted and stored at -30 °C, then diluted to working concentrations immediately before use.

Rhodamine-B Azide and alkyne

Rhodamine-B azide and rhodamine-B alkyne were prepared in house (Wirth et al., 2015).

Preparation of MyOne C1 magnetic beads

10 µL MyOne strepatividin C1 beads were washed and prepared according to manufacturer's instructions (Thermo Fisher). Prior to biotin binding, beads were treated with 50 µM BLOCK for 20 minutes, rotating at room temperature. Beads were drawn from solution using a magnet and rinsed three times with 40 µL Nitrate Bind and Wash buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M $NaNO_3$) by aspirating the beads using a micropipette and returning the tube to the magnet for two minutes to remove beads from solution.

Binding to streptavidin beads

Simultaneously, 100 pmol of BIO was annealed to 100 pmol of 5' HEX clicked to Rho-B-azide in 40 µL Nitrate Bind and Wash buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M $NaNO_3$) by heating to 90 °C for five minutes, then cooled to room temperature for 20 minutes. This hybridized solution was applied to the washed and blocked beads and incubated at 4 °C overnight with rotation.

Elution of captured oligonucleotides

After incubation, DNA-bound beads were washed six times with 40 µL of decreasing concentrations of $NaNO_3$ in 5 mM tris-HCl, 0.5 mM EDTA (from 1 M $NaNO_3$ to 100 µM $NaNO_3$ in six increments). Each wash step was left on the magnet for two minutes to remove beads from solution, then supernatant was removed and beads re-suspended in the next wash buffer. Wash supernatants were saved for later analysis.

After completing the wash steps, DNA was re-suspended in 20 µL Elution buffer (0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) and eluted from beads by heating to 95

°C for 5 minutes, then placed on a magnet to extract supernatant. Samples were ethanol precipitated and visualized on a 15% 19:1 mono:bis acrylamide gel using UV transillumination.

Platination of hairpin DNA

HP-AzaPt was prepared as follows: 40 μM *in vitro* synthetic hairpin DNA was added to 10 mM phosphate buffer, 100 mM $NaNO_3$ (final concentration), then denatured by heating to 90 °C for 5 minutes and slow cooled to 4 °C. Samples were warmed to room temperature then $Mg(NO_3)_2$ and azaplatin (Figure 1) were added to final concentrations of 1 mM and 120 μM, respectively, then incubated overnight at room temperature with rotation to allow azaplatin to react with hairpin DNA (Wirth et al., 2015).

Click reactions

Bulk solution of HEX-Rho was prepared for pulldown optimization as follows: HEX was added to a solution containing 100 mM phosphate buffer, pH 7 and premixed 150 μM $CuSO_4$, 1.5 mM THPTA. Rhodamine-B azide was added to a final concentration of 200 μM, and sodium ascorbate was added to a final concentration of 150 μM to initiate the click reaction. Reaction was allowed to proceed for two hours at room temperature, with rotation, away from light. Reaction was quenched by running sample through a sephadex G-25 column.

Experimental click reactions were performed as follows: HP-AzaPt was prepared as described, then added to a final concentration of 3 μM to a solution containing (final

concentrations): Alkyne constituent (either HEX or rhodamine-B alkyne) 13 μM; 100 μM phosphate buffer, pH 7; 75 μM CuSO4 premixed with 375 μM THPTA. Sodium ascorbate was added to a final concentration of 150 μM to initiate the click reaction. The reaction was allowed to proceed for 2 hours at 37 °C with rotation, then quenched by running samples through sephadex G-25 columns. Bead pulldown was performed as described for all clicked samples. Samples were visualized on a 20% 19:1 mono:bis acrylamide gel and stained with SYBR gold (Thermo Fisher).

REFERENCES CITED

Akaboshi, M.; Kawai, K.; Maki, H.; Akuta, K.; Ujeno, Y.; Miyahara, T. The number of platinum atoms binding to DNA, RNA and protein molecules of HeLa cells treated with cisplatin at its mean lethal concentration. *Jpn. J. Cancer Res*. **1992**, *83*, 522–526.

Alderden, R. A.; Hall, M. D.; Hambley, T. W. Products of Chemistry The Discovery and Development of Cisplatin. *J. Chem. Educ.* **2006***, 83,* 728–734.

Ananvoranich, S. Substrate Specificity of delta Ribozyme Cleavage. *J. Biol. Chem.* **1998**, *273,* 13182–13188.

Anders, L.; Guenther, M. G.; Qi, J.; Fan, Z. P.; Marineau, J. J.; Rahl, P. B.; Lovén, J.; Sigova, A. A.; Smith, W. B.; Lee, T. I.; Bradner, J. E.; Young, R. A. Genome-wide localization of small molecules. *Nat. Biotechnol.* **2014,** *32*, 92–96.

Anders, S.; Pyl, P. T.; Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics.* **2015,** *31*, 166–169.

Astakhova, I. K.; Wengel, J. Interfacing click chemistry with automated oligonucleotide synthesis for the preparation of fluorescent DNA probes containing internal xanthene and cyanine dyes. *Chem. Eur. J.* **2013,** *19*, 1112–1122.

Barski, A.; Cuddapah, S.; Cui, K.; Roh, T. Y.; Schones, D. E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell.* **2007,** *129,* 823–837.

Basu, A.; Krishnamurthy, S. Cellular Responses to Cisplatin-Induced DNA Damage. *J. Nucleic Acids.* **2010.** *2010*, 1–16.

Bauer, B. E.; Wolfger, H.; Kuchler, K. Inventory and function of yeast ABC proteins: About sex, stress, pleiotropic drug and heavy metal resistance. *Biochim. Biophys. Acta.* **1999,** *1461*, 217–236.

Beaudoin, J. D.; Jodoin, R.; Perreault, J. P. In-line probing of RNA G-quadruplexes. *Methods.* **2013,** *64*, 79–87.

Benedetti, B. T.; Peterson, E. J.; Kabolizadeh, P.; Martínez, A.; Kipping, R.; Farrell, N. P. Effects of noncovalent platinum drug-protein interactions on drug efficacy: Use of fluorescent conjugates as probes for drug metabolism. *Mol. Pharm.* **2011,** *8*, 940–948.

Bicaku, E.; Xiong, Y.; Marchion, D. C.; Chon, H. S.; Stickles, X. B.; Chen, N.; Judson, P. L.; Hakam, A.; Gonzalez-Bosquet, J.; Wenham, R. M.; Apte, S. M.; Fulp, W.; Cubitt, C. L.; Chen, D. T.; Lancaster, J. M. In vitro analysis of ovarian cancer response to cisplatin, carboplatin, and paclitaxel identifies common pathways that are also associated with overall patient survival. *Br. J. Cancer* **2012,** *106,* 1967–1975.

Bifano, A. L.; Atassi, T.; Ferrara, T.; Driscoll, D. M. Identification of nucleotides and amino acids that mediate the interaction between ribosomal protein L30 and the SECIS element. *BMC Mol. Biol.* **2013,** *14,* 1-12.

Brown, S. L.; Stockdale, V. J.; Pettolino, F.; Pocock, K. F.; de Barros Lopes, M.; Williams, P. J; Bacic, A.; Fincher, G. B.; Høj P. B.; Waters, E. J. Reducing haziness in white wine by overexpression of Saccharomyces cerevisiae genes YOL155c and YDR055w. *Appl. Microbiol. Biotechnol.* **2007,** *73*, 1363–1376.

Bun-ya, M.; Nishimura, M.; Harashima, S.; Oshima, Y. The PH084 Gene of Saccharomyces cerevisiae Encodes an Inorganic Phosphate Transporter. *Mol. Cell Biol.* **1991,** *11*, 3229–3238.

Burger, A. M.; Double, J. A.; Newell, D. R. Inhibition of telomerase activity by cisplatin in human testicular cancer cells. *Eur. J. Cancer.* **1997,** *33*, 638–644.

Butcher, S. E.; Pyle, A. M. The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks Interactions That Stabilize RNA Tertiary Structure. *Acc. Chem. Res.* **1302,** *44,* 1302–1311.

Chapman, E. G.; DeRose, V. J. Enzymatic processing of platinated RNAs. *J. Am. Chem. Soc.* **2010,** *132*, 1946–1952.

Chapman, E. G.; DeRose, V. J. Site-specific platinum(II) cross-linking in a ribozyme active site. J. *Am. Chem. Soc.* **2012,** *134*, 256–262.

Chapman, E. G.; Hostetter, A. A.; Osborn, M. F.; Miller, A. L.; DeRose, V. J. Binding of kinetically inert metal ions to RNA: the case of platinum(II). *Met. Ions Life Sci.* **2011,** *9*, 347–377.

Chapman, E. G.; Moon, S. L.; Wilusz, J.; Kieft, J. S. RNA structures that resist degradation by Xrn1 produce a pathogenic dengue virus RNA. *Elife.* **2014,** *2014*, 1–25.

Chen, J. H.; Yajima, R.; Chadalavada, D. M; Chase, E.; Bevilacqua, P. C.; Golden, B. L. A 1.9 Å Crystal Structure of the HDV Ribozyme Precleavage Suggests both Lewis Acid and General Acid Mechanisms Contribute to Phosphodiester Cleavage. *Biochemistry.* **2010,** *49*, 6508–6518.

Cherry, J. M.; Hong, E. L.; Amundsen, C.; Balakrishnan, R.; Binkley, G.; Chan, E. T.; Christie, K. R.; Costanzo, M. C.; Dwight, S. S.; Engel, S. R.; Fisk, D. G.; Hirschman, J. E.; Hitz, B. C.; Karra, K.; Krieger, C. J.; Miyasato, S. R.; Nash, R. S.; Park, J.; Skrzypek, M. S.; Simison, M.; Weng, S.; Wong, E. D. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **2012,** *40*, D700-705.

Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Scuderi, D.; Salpin, J. Y. Interaction of cisplatin with 5′-dgmp: A combined irmpd and theoretical study. *Inorg. Chem.* **2015,** *54*, 3513–3522.

Dalbiès, R.; Payet, D.; Leng, M. DNA double helix promotes a linkage isomerization reaction in trans-diamminedichloroplatinum(II)-modified DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91*, 8147–8151.

Damsma, G. E.; Alt, A.; Brueckner, F.; Carell, T.; Cramer, P. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat. Struct. Mol. Biol.* **2007,** *14*, 1127–1133.

Dedduwa-Mudalige, G. N. P.; Chow, C. S. Cisplatin targeting of bacterial ribosomal RNA hairpins. *Int. J. Mol. Sci.* **2015,** *16*, 21392–21409.

Deigan, K. E.; Li, T. W.; Mathews, D. H.; Weeks, K. M. Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 97–102.

Ding, S.; Qiao, X.; Suryadi, J,; Marrs, G. S.; Kucera, G. L.; Bierbach, U. Using fluorescent post-labeling to probe the subcellular localization of DNA-targeted platinum anticancer agents. *Angew. Chemie Int. Ed.* **2013,** *52,* 3350–3354.

Dominissini, D,; Nachtergaele, S.; Moshitch-Moshkovitz, S,; Peer, E.; Kol, N.; Ben-Haim, M. S.; Dai, Q.; Di Segni, A.; Salmon-Divon, M.; Clark, W. C.; Zheng, G.; Pan, T.; Solomon, O.; Eyal, E.; Hershkovitz, V.; Han, D.; Doré, L. C.; Amariglio, N.; Rechavi, G.; He, C. The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature.* **2016,** *530*, 441–446.

Dyson, P. J.; Sava, G. Metal-based antitumour drugs in the post genomic era. *Dalton Trans.* **2006,** *16*. 1929–1933.

Eickbush, D. G.; Burke, W. D.; Eickbush, T. H. Evolution of the R2 retrotransposon ribozyme and its self-cleavage site. *PLoS One* [Online] **2013,** *8,* e66441.

Elmroth, S. K. C.; Lippard, S. J. Platinum binding to d(GpG) target sequences and phosphorothioate linkages in DNA occurs more rapidly with increasing oligonucleotide length. *J. Am. Chem. Soc.* **1994,** *116*, 3633–3634.

Engel, S. R.; Dietrich, F. S.; Fisk, D. G.; Binkley, G.; Balakrishnan, R.; Costanzo, M. C.; Dwight, S. S.; Hitz, B. C.; Karra, K.; Nash, R. S.; Weng, S.; Wong, E. D.; Lloyd, P.; Skrzypek, M. S.; Miyasato, S. R.; Simison, M.; Cherry, J. M. The Reference Genome Sequence of Saccharomyces cerevisiae: Then and Now. *G3 (Bethesda)* **2013,** *4*. 389-398.

Fauzi, H.; Kawakami, J.; Nishikawa, F.; Nishikawa, S. Analysis of the cleavage reaction of a trans-acting human hepatitis delta virus ribozyme. *Nucleic Acids Res.* **1997,** *25*, 3124–3130.

Favre, A.; Saintomé, C.; Fourrey, J. Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *J. Photochem. Photobiol. B Biol.* **1998,** *42*, 109–124.

Fichtinger-Schepman, M.; van der Veer, J. L.; den Hartog, J. H.; Lohman, P. H.; Reedijk, J. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry* **1985,** *24*, 707–713.

Fleischer, T. C.; Weaver, C. M.; McAfee, K. J.; Jennings, J. L.; Link, A. J. Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev.* **2006,** *20*, 1294–1307.

Florea, A.-M.; Büsselberg, D. Cisplatin as an anti-tumor drug: cellular mechanisms of activity, drug resistance and induced side effects. *Cancers* **2011,** *(Basel). 3*, 1351–1371.

Garreau de Loubresse, N.; Prokhorova, I.; Holtkamp, W.; Rodnina, M. V.; Yusupova, G.; Yusupov, M. Structural basis for the inhibition of the eukaryotic ribosome. *Nature* **2014,** *513*, 517–522.

Gelasco, A.; Lippard, S. J. NMR solution structure of a DNA dodecamer duplex containing a cis- diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochem.* **1998,** *37*, 9230–9239.

Golden, B. L.; Kundrot, C. E. RNA crystallization. *J. Struct. Biol.* **2003,** *142*, 98–107.

Golden, B. Two distinct catalytic strategies in the hepatitis delta virus ribozyme cleavage reaction. *Biochem.* **2011,** *50*, 9424–9433.

Graifer, D.; Karpova, G. Photoactivatable RNA derivatives as tools for studying the structural and functional organization of complex cellular ribonucleoprotein machineries. *RSC Adv.* **2013,** *3*, 2858-2872.

Halic, M.; Becker, T.; Frank, J.; Spahn, C. M.; Beckmann, R. Localization and dynamic behavior of ribosomal protein L30e. *Nat. Struct. Mol. Biol.* **2005,** *12*, 467–468.

Halvorsen, M.; Martin, J. S.; Broadaway, S.; Laederach, A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet.* **2010,** *6*, 1–11.

Hampel, K. J.; Burke, J. M. Time-resolved hydroxyl-radical footprinting of RNA using Fe(II)-EDTA. *Methods.* **2001,** *23*, 233–239.

Harper, B. W,; Krause-Heuer, A. M.; Grant, M. P.; Manohar, M.; Garbutcheon-Singh, K. B.; Aldrich-Wright, J. R. Advances in platinum chemotherapeutics. *Chemistry.* **2010,** *16*, 7064–7077.

Harris, M. E.; Christian, E. L. RNA crosslinking methods. *Methods in Enzymol.* **2009,** *468*, 127–46.

Holbrook, S. R.; Kim, S. H. RNA crystallography. *Biopolym*. **1997,** *44,* 3–21.

Horn, S. Target enrichment via DNA hybridization capture. *Meth. Mol. Biol.* **2012,** *840*, 177–188

Hostetter, A. A.; Chapman, E. G.; DeRose, V. J. Rapid cross-linking of an RNA internal loop by the anticancer drug cisplatin. *J. Am. Chem. Soc.* **2009,** *131*, 9250–9257.

Hostetter, A. A.; Osborn, M. F.; Derose, V. J. RNA-Pt adducts following cisplatin treatment of saccharomyces cerevisiae. *ACS Chem. Biol.* **2012,** *7*, 218–225.

Hurd, H. K.; Roberts, C. W.; Roberts, J. W. Identification of the gene for the yeast ribonucleotide reductase small subunit and its inducibility by methyl methanesulfonate. *Mol. Cell. Biol.* **1987,** *7*, 3673–3677.

Incarnato, D.; Neri, F.; Anselmi, F.; Oliviero, S. Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biol.* **2014,** *15*, 491.

Ingolia, N. T.; Brar, G. A.; Rouskin, S.; McGeachy, A. M.; Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* **2012,** *7*, 1534–1550.

Jagodinsky, J. C.; Sulima, A.; Cao, Y.; Poprawski, J. E.; Blackman, B. N.; Lloyd, J. R.; Swenson R. E.; Gottesman, M. M.; Hall, M. D. Evaluation of fluorophore-tethered platinum complexes to monitor the fate of cisplatin analogs. *J. Biol. Inorg. Chem.* **2015,** *20*, 1081–1095.

Jamieson, E. R.; Lippard, S. J. Structure, recognition, and processing of cisplatin-DNA adducts. *Chem. Rev.* **1999,** *99*, 2467–2498.

Jeeninga, R. E. et al. Variable regions V13 and V3 of Saccharomyces cerevisiae contain structural features essential for normal biogenesis and stability of 5.8S and 25S rRNA. *RNA* **1997,** *3*, 476–488.

Jeeninga, R. E.; Van Delft, Y.; de Graaff-Vincent, M.; Dirks-Mulder, A.; Venema, J.; Raué, H. A. Crystal structure of the 80S yeast ribosome. *Curr. Opin. Struct. Biol.* **2012,** *22*, 759–767.

Jensen, L. T.; Ajua-Alemanji, M.; Culotta, V. C. The Saccharomyces cerevisiae High Affinity Phosphate Transporter Encoded by PHO84 Also Functions in Manganese Homeostasis. *J. Biol. Chem.* **2003,** *278*, 42036–42040.

Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. *Science* **2007,** *80*. 316, 1497–1502.

Jung, Y.; Lippard, S. J. Direct cellular responses to platinum-induced DNA damage. *Chem. Rev.* **2007.** *107*, 1387–1407.

Juzumiene, D.; Shapkina, T., Kirillov, S.; Wollenzien, P. Short-range RNA-RNA crosslinking methods to determine rRNA structure and interactions. *Methods.* **2001,** *25,* 333–343.

Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.;Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016** *44*, D457–D462.

Kapoor, S.; Waldmann, H.; Ziegler, S. Novel Approaches to Map Small Molecule-Target Interactions. *Bioorg. Med. Chem.* **2016,** *24,* 3232–3245.

Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed. Engl. 40*, **2001,** 2004–2021.

Lai, M. M. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* **1995,** *64,* 259–286.

Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature.* **2001,** *409,* 860–921.

Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009,** *10*, 25.1-25.10.

Lawley, P. D.; Brookes, P. Further Studies on the Alkylation of Nucleic Acids and Their Constituent Nucleotides. *Biochem. J.* **1963,** *89*, 127–138.

Lee, M. Y.; Kim, M. A.; Kim, H. J.; Bae, Y. S.; Park, J. I.; Kwak, J. Y.; Chung, J. H.; and Yun, J. Alkylating agent methyl methanesulfonate (MMS) induces a wave of global protein hyperacetylation: Implications in cancer cell death. *Biochem. Biophys. Res. Commun.* **2007,** *360*, 483–489.

Li, J.; Kim, T.; Nutiu, R.; Ray, D.; Hughes, T. R.; Zhang, Z. Identifying mRNA sequence elements for target recognition by human Argonaute proteins. Genome Res. **2014,** *5*, 775–785.

Lindell, M.; Romby, P.; Wagner, E.G.H. Lead(II) as a probe for investigating RNA structure in vivo. *RNA.* **2002,** *8*, 534–451.

Love, M. I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014,** *15,* pp. 550-571.

Lundin, C.; North, M.; Erixon, K.; Walters, K.; Jenssen, D.; Goldman, A. S.; Helleday, T. Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res.* **2005,** *33*, 3799–3811.

Luo, W.; Friedman, M. S.; Shedden, K.; Hankenson, K. D.; Woolf, P. J. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*. **2009,** *10.161.*

Luo, W.; Friedman, M.; Shedden, K.; Hankenson, K.; Woolf, P. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics,* **2009,** *10*, pp. 161-178.

Makley, L. N.; Gestwicki, J. E. Expanding the number of 'druggable' targets: non-enzymes and protein-protein interactions. *Chem. Biol. Drug Des.* **2013,** *81*, 22-32.

Mamanova, L.; Coffey, A. J.; Scott, C. E.; Kozarewa, I.; Turner, E. H.; Kumar, A.; Howard, E.; Shendure, J.; Turner, D. J. Target-enrichment strategies for next-generation sequencing. *Nat. Methods* **2010,** *7*, 111–118.

Mandic, A.; Hansson, J.; Linder, S.; Shoshan, M. C. Cisplatin induces endoplasmic reticulum stress and nucleus-independent apoptotic signaling. *J. Biol. Chem.* **2003,** *278*, 9100–9106.

Marchand, V.; Blanloeil-Oillo, F.; Helm, M.; Motorin, Y. Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res.* **2016,** *44.*

Martick, M.; Lee, T.; York, D. M.; Scott, W. G. Solvent Structure and Hammerhead Ribozyme Catalysis. Chem. Biol. **2008,** *15,* 332-342.

Martin, M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnet J.* **2011,** *17*, 10–12.

Mattick, J. S. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Rep.* **2001,** *2*, 986–991.

Mattick, J. S. RNA regulation: a new genetics? *Nat. Rev. Genet.* **2004,** *5*, 316-323.

Mattick, J. S.; Makunin, I. V. Non-coding RNA. *Hum. Mol. Genet.* **2006,** *15,* 17–29.

McGhee, J. D.; von Hippel, P. H. Formaldehyde as a probe of DNA structure. I. Reaction with exocyclic amino groups of DNA bases. *Biochemistry.* **1975,** *14*, 1281–1296.

Melnikov, S. V.; Dieter, S.; Steitz, T. A.; Polikanov, Y. S. Insights into RNA binding by the anticancer drug cisplatin from the crystal structure of cisplatin-modified ribosome. *Nucleic Acids Res.* **2016,** *44*, 4978–4987.

Miyahara, K.; Mizunuma, M.; Hirata, D.; Tsuchiya, E.; Miyakawa, T. The involvement of the Saccharomyces cerevisiae multidrug resistance transporters Pdr5p and Snq2p in cation resistance. *FEBS Lett.* **1996,** *399*, 317–320.

Moghaddam, A. D.; White, J. D.; Cunningham, R. M.; Loes, A. N.; Haley, M. M.; DeRose, V. J. Convenient detection of metal–DNA, metal–RNA, and metal–protein adducts with a click-modified Pt(II) complex. *Dalton Trans.* **2015,** *44,* 3536– 3539,

Molenaar, C.; Teuben, J. M.; Heetebrij, R. J.; Tanke, H. J.; Reedijk, J. New insights in the cellular processing of platinum antitumor compounds, using fluorophore-labeled

platinum complexes and digital fluorescence microscopy. *J Biol Inorg Chem.* **2000,** *5,* 655-665.

Moore, P. B.; Steitz, T. A. After the ribosome structures: how does peptidyl transferase work? *RNA,* **2003,** *9,* 155–159.

Morris, K. V.; Mattick, J. S. The rise of regulatory RNA. *Nat. Rev. Genet.* **2014,** *15,* 423–437.

Nakano, S. General Acid-Base Catalysis in the Mechanism of a Hepatitis Delta Virus Ribozyme. *Science.* **2000,** *287,* 1493–1497.

Ohno, S. So much 'junk DNA' in our genome. *Brookhaven Symp. Biol.* **1972,** *23,* 366-370.

Osborn, M. F.; White, J. D.; Haley, M. M.; DeRose, V. J. Platinum-RNA modifications following drug treatment in S. cerevisiae Identified by click chemistry and enzymatic mapping. *ACS Chem. Biol.* **2014,** *9,* 2404–2411.

Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **2006,** *5,* 993–996.

Park, P. J. ChIP-Seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **2009,** *10,* 669–680.

Pascoe, J. M.; Roberts, J. J.; Interactions between mammalian cell DNA and inorganic platinum compounds. II. Interstrand cross-linking of isolated and cellular DNA by platinum(IV) compounds. *Biochem. Pharmacol.* **1974,** *23,* 1345-1357.

Petrov, A. N.; Meskauskas, A.; Roshwalb, S. C.; Dinman, J. D. Yeast ribosomal protein L10 helps coordinate tRNA movement through the large subunit. *Nucleic Acids Res.* **2008,** *36,* 6187–6198.

Petrov, A. S.; Bernier, C. R.; Gulen, B.; Waterbury, C. C.; Hershkovits, E.; Hsiao, C.; Harvey, S. C.; Hud, N. V.; Fox, G. E.; Wartell, R. M.; Williams, L. D. Secondary Structures of rRNAs from All Three Domains of Life. *PLoS One* **2014,** *9,* 1–6.

Piekna-Przybylska, D.; Decatur, W. A.; Fournier, M. J. New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA,* **2007,** *13*, 305–312.

Rakauskaite, R.; Dinman, J. D. Mutations of highly conserved bases in the peptidyltransferase center induce compensatory rearrangements in yeast ribosomes. *RNA.* **2011** *17,* 855–864.

Ramesh, M.; Woolford, J. L. Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. *RNA*. **2016,** *8*.

Reuter, J. A.; Spacek, D. V.; Snyder, M. P. High-Throughput Sequencing Technologies. *Mol. Cell.* **2015,** *58,* 586–597.

Rhodes, D.; Piper, P. W.; Clark, B. F. C. Location of a platinum binding site in the structure of yeast phenylalanine transfer RNA. *J. Mol. Biol.* **1974,** *89,* 469-475.

Rijal, K.; Chow, C. S. A new role for cisplatin: probing ribosomal RNA structure. *Chem. Commun. (Camb).* **2009,** *1,* 107–109.

Rodriguez, R.; Miller, K. M. Unravelling the genomic targets of small molecules using high-throughput sequencing. *Nat. Rev. Genet.* **2014,** *15,* 783–796.

Roos, W. P.; Kaina, B. DNA damage-induced cell death: From specific DNA lesions to the DNA damage response and apoptosis. *Cancer Lett.* **2013,** *332*, 237–248.

Rosenberg, B. Anticancer activity of cis-dichlorodiammineplatinum(II) and some relevant chemistry. *Cancer Treat. Rep.* **1979,** *63*, 1433-1438.

Rosenberg, B.; Vancamp, L.; Krigas, T. Inhibition of Cell Division in Escherichia Coli By Electrolysis Products From a Platinum Electrode. *Nature.* **1965,** *205*, 698–699.

Rosenberg, B.; vanCamp, L.; Trosko, J. E.; Mansour, V. H. Platinum compounds: a new class of potent antitumour agents. *Nature.* **1969,** *222,* 385–386.

Rosenberg, J.; Sato, P. Messenger RNA loses the ability to direct in vitro peptide synthesis following incubation with cisplatin. *Mol. Pharmacol.* **1988,** *33,* 611–616.

Safaei, R.; Howell, S. B. Copper transporters regulate the cellular pharmacology and sensitivity to Pt drugs. *Crit. Rev. Oncol. Hematol.* **2005,** *53,* 13–23.

Salehi-Ashtiani, K.; Lupták, A.; Litovchick, A.; Szostak, J. W. A Genomewide Search for Ribozymes Reveals an HDV-Like Sequence in the Human CPEB3 Gene. *Science.* **2006,** *313,* 1788–1792.

Sharp, P. A. The centrality of RNA. *Cell.* **2009,** *136*, 577–580.

Sherman, S. E.; Lippard, S. J. Structural aspects of platinum anticancer drug interactions with DNA. *Chem. Rev.* **1987,** *87,* 1153–1181.

Siddik, Z. H. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene.* **2003,** *22,* 7265–7279.

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

Speers, A. E.; Cravatt, B. F.; Jolla, L. Profiling Enzyme Activities In Vivo Using Click Chemistry *Methods.* **2004,** *11,* 535–546.

Spitale, R. C.; Flynn, R. A.; Zhang, Q. C.; Crisalli, P.; Lee, B.; Jung, J. W.; Kuchelmeister, H. Y.; Batista, P. J.; Torre, E. A.; Kool, E. T.; Chang, H. Y. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature.* **2015,** *519,* 486–490.

Stern, S.; Moazed, D.; Noller, H. F. Structural Analysis of RNA Using Chemical and Enzymatic Probing Monitored by Primer Extension. *Methods Enzymol.* **1988,** *164,* 481–489.

Sundaralingam, M.; Rubin, J. R.; Rao, S. T. X-ray studies on the interaction of the anticancer agent cis-[Pt(NH3)2cl2] to tRNAphe. A mechanism for the formation of the intrastrand cross-link to adjacent guanines in DNA. *Prog. Clin. Biol. Res.* **1985,** *172B*, 175-184.

Sweeney, R.; Chen, L.; Yao, M. C. An rRNA variable region has an evolutionarily conserved essential role despite sequence divergence. *Mol. Cell. Biol.* **1994,** *14,* 4203–4215.

Talkish, J.; May, G.; Lin, Y.; Woolford, J. L.; Mcmanus, C. J. Mod-seq : high-throughput sequencing for chemical probing of RNA structure. *RNA.* **2014,** *20,* 713–720.

Tang, Y.; Bouvier, E.; Kwok, C. K.; Ding, Y.; Nekrutenko, A.; Bevilacqua, P. C.; Assmann, S. M. StructureFold: Genome-wide RNA secondary structure mapping and reconstruction in vivo. *Bioinformatics.* **2015,** *31,* 2668–2675.

Thomas, J. R.; Hergenrother, P. J. Targeting RNA with Small Molecules. *Chem. Rev.* **2008,** *108,* 1171–1224.

Tijerina, P.; Mohr, S.; Russell, R. DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.* **2007,** *2,* 2608–2623.

Tkach, J. M.; Yimit, A.; Lee, A. Y.; Riffle, M.; Costanzo, M.; Jaschob, D.; Hendry, J. A.; Ou, J.; Moffat, J.; Boone, C.; Davis, T. N.; Nislow, C.; Brown, G. W. Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat. Cell. Biol.* **2012,** *14*, 966–976.

Tsaponina, O.; Barsoum, E.; Åström, S. U.; Chabes, A. Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools. *PLoS Genet.* **2011,** *7*.

Tumer, N. E.; Li, X. P. Interaction of Ricin and Shiga Toxins with Ribosomes. Curr. Top. *Microbiol. Immunol.* **2012,** *357*, 1–18.

Ulitsky, I.; Bartel, D. P. LincRNAs: Genomics, evolution, and mechanisms. *Cell.* **2013,** *154*, 26–46.

Valouev, A.; Johnson, D. S.; Sundquist, A.; Medina, C.; Anton, E.; Batzoglou, S.; Myers, R. M.; Sidow, A. Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat. Methods* **2008,** *5,* 829–834.

Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic acid catalysis: metals, nucleobases, and other cofactors. *Chem. Rev.* **2014,** *114,* 4318–4342.

Warner, J. R. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **1999,** *24,* 437–440.

Watters, K. E.; Yu, A. M.; Strobel, E. J.; Settle, A. H.; Lucks, J. B. Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods.* **2015,** *103,* 34–48.

White, J. D.; Haley, M. M.; DeRose, V. J. Multifunctional Pt(II) Reagents: Covalent Modifications of Pt Complexes Enable Diverse Structural Variation and In-Cell Detection. *Acc. Chem. Res.* **2016,** *49,* 56–66.

Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* **2014,** *12,* 35–48.

Winderickx, J.; de Winde, J. H.; Crauwels, M.; Hino, A.; Hohmann, S.; Van Dijck, P.; Thevelein, J. M. Regulation of genes encoding subunits of the trehalose synthase complex in Saccharomycescerevisiae: Novel variations of STRE-mediated transcription control? *Mol. Gen. Genet.* **1996,** *252,* 470-482.

Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011,** *67*, 235–242.

Wirth, R.; White, J. D.; Moghaddam, A. D.; Ginzburg, A. L.; Zakharov, L. N.; Haley, M. M.; DeRose, V. J. Azide vs Alkyne Functionalization in Pt(II) Complexes for Post-treatment Click Modification: Solid-State Structure, Fluorescent Labeling, and Cellular Fate. *J. Am. Chem. Soc.* **2015,** *137,* 15169–15175.

Wrzesinski, J.; Wichłacz, A.; Nijakowska, D.; Rebowska, B. Nawrotb, B.; Ciesiołka, J. Phosphate residues of antigenomic HDV ribozyme important for catalysis that are revealed by phosphorothioate modification. *New J. Chem.* **2010,** *34,* 1018–1026.

Wu, J.; Delneri, D.; O' Keefe, R. T. Non-coding RNAs in Saccharomyces cerevisiae : What is the function? *Biochem. Soc. Trans.* **2012,** *40,* 907–911.

Yao, R.; Zhang, Z.; An, X.; Bucci, B.; Perlstein, D. L.; Stubbe, J.; Huang, M. Subcellular localization of yeast ribonucleotide reductase regulated by the DNA replication and damage checkpoint pathways. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100,* 6628–6633.

Yassin, A.; Mankin, A. S. Potential new antibiotic sites in the ribosome revealed by deleterious mutations in RNA of the large ribosomal subunit. *J. Biol. Chem.* **2007,** *282,* 24329–24342.

Zhirnov, O. V.; Wollenzien, P. Action spectra for UV-light induced RNA-RNA crosslinking in 16S ribosomal RNA in the ribosome. *Photochem. Photobiol. Sci.* **2003,** *2,* 688–693.