Investigating a Biological Role for Paranemic Crossover (PX) DNA in *S. cerevisiae*

Jillian Armenia
Mirkin Lab
Tufts University Department of Biology
Senior Honors Thesis 2020

Table of Contents

Acknowledgements……..3

Abstract……4

Introduction……..5

Materials and Methods……8

Results……13

Discussion……..22

Conclusions……32

Literature Cited……..33

Acknowledgements

I would like to first acknowledge and thank Dr. Sergei Mirkin and Dr. Sasha Khristich for their constant help and encouragement over the past three years. They have both shown me how amazing science can be, and I never would have gotten here without them. Many thanks to Dr. Mitch McVey for his guidance as a member of my thesis committee and as a wonderful molecular biology professor. I would also like to thank my academic advisor, Dr. Kate Mirkin for her constant kindness and for the encouragement to pursue a senior thesis.

Additionally, I would like to acknowledge the family of Russell L. Carpenter for the generous award I was lucky enough to receive last summer, *The Russell L. Carpenter Summer Internship*. My project would not have been possible without the resources this internship gave to me.

A huge thank you to Chiara Masnovo and Gary Zhang for their help with the experimental side of this project. I would also like to thank every single member of the Mirkin Lab for everything they have done for me. I could not have asked for a better group of people to surround myself with. Thank you, Elle, Katia, Julia, Nastya, Grace, Tara, Yanchen, and Emily for being such a big part of the best experience of my college career.

I would finally like to thank my friends and family for their love and encouragement. Special thanks to my sister, Erin, for proofreading this manuscript. Mom, Dad, Andrew, Erin, and Chris, thank you for always supporting me and being proud of my accomplishments big and small. I couldn't have done it without you.

Abstract

Paranemic Crossover DNA (PX-DNA) is a four stranded DNA structure which can form at regions of double stranded DNA homology. This structure is a multiple Holliday junction, which begs the question: could PX-DNA be involved in the pairing of homologous chromosomes or homologous recombination? Here, we present promising preliminary data that points to the existence of a biological role for this unusual DNA structure. We show that PX-DNA forming sequences (PX motifs) undergo two mechanistically distinct types of genetic rearrangements in an *S. cerevisiae* model system. The first type of rearrangement yields junctions formed at microhomologies flanking the PX motifs on either side. We found that the presence of PX motifs is required for this type of rearrangement and we propose that they take place during break-induced replication. The second type of rearrangement forms junctions within the PX motifs and likely happens through microhomology-mediated end joining. Further investigation is required to fully elucidate the mechanisms and precise role of PX-DNA in these genetic rearrangements.

Introduction

For over one-hundred years, scientists have attempted to understand the pairing of homologous chromosomes in living cells. In 1909, F.A. Janssens published his chiasmatype theory, which postulated that physical interaction between homologous chromosomes corresponded to the location of crossover events during meiosis (reviewed by Koszul et al., 2012). Since then, our understanding of meiotic crossing over has been refined into what is today widely accepted as chromosomal exchange via homologous recombination.

In the mainstream accepted model, recombination between homologous chromosomes is initiated by the formation of a double strand break (DSB) by Spo11 (Keeney et al., 1997). Next, the ends of the DSB are resected and capped by Rad51. Rad51 searches for a region of homology within the homologous chromosome to utilize as a template for repair (Shinohara et al., 1992). After the formation and resolution of a Holliday junction by structure-specific endonucleases, (reviewed by Liu and West, 2004), recombinant chromosomes with interchanged genetic material may form. This process increases the genetic diversity of subsequent generations.

DSBs in DNA can be highly mutagenic, especially if a homologous template is not readily available and cells undergo error-prone repair pathways (reviewed by Rodgers and McVey, 2015). This might lead to detrimental consequences such as gross chromosomal rearrangements (Varga and Aplan, 2005). Therefore, many scientists have speculated that meiotic chromosomes pair up before strand exchange in order to ensure correct repair of Spo11-induced DSBs. This hypothetical process is termed DSB-independent pairing.

There has been evidence that Z-form DNA could be involved in homologous pairing (Kmiec and Holloman, 1986) and that chromosomes can pair before the initiation of meiosis via interstitial interactions in regions of homology (Kleckner and Weiner, 1994). In both cases, these interactions seem to correspond to crossing-over events, but neither of these theories have gained much traction over the years. Here, we investigate the potential role of a more recently discovered non-B-DNA form, PX-DNA, in the pairing of homologous regions of DNA.

In 2004, the Seeman lab at New York University published a paper describing a novel DNA structure called Paranemic Crossover DNA, or PX-DNA (Shen et al., 2004). PX-DNA is a four stranded DNA molecule that forms between two DNA duplexes that share a partial or full homology. It involves both interhelical and intrahelical pairing and consists of multiple Holliday junctions. In other words, each strand pairs with its original partner for several bases, then with the complementary strand from the other duplex. This pattern repeats for the length of the structure (Figure 1). Currently, PX-DNA is being studied as a potential tool in DNA nanotechnology and bioinformatics (reviewed by Wang et al., 2019), but little is known about a potential biological role for the structure.

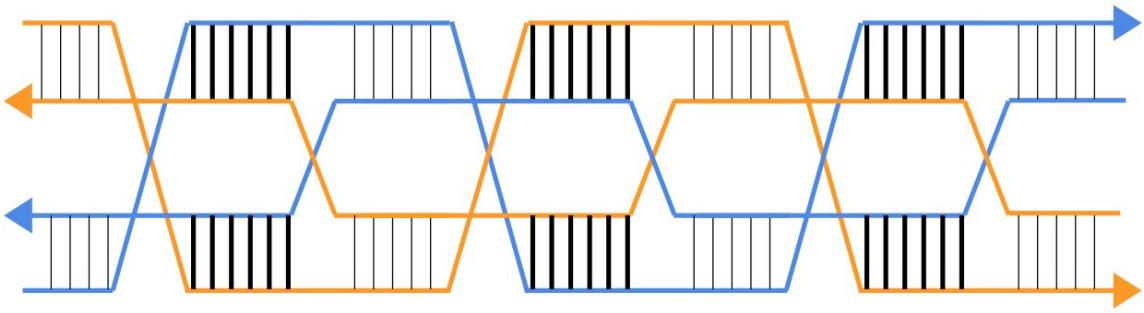

Figure 1. Structure of PX-DNA

Orange and blue strands represent two original DNA double helices. Thin black lines represent base pairs between strands from the same helix (intra-helix base pairing). Thick black lines represent base pairing between strands from opposite helices (inter-helix base pairing). Arrowheads represent the 3' ends of DNA strands. (Shen et al., 2004).

In 2010, the Seeman lab published an additional study, in which they confirmed that PX-DNA forms at regions of double-stranded DNA homology and further characterized the sequence requirements for PX structure formation. They concluded that a stable PX structure can be formed by two duplexes that share 6 to 9 bp of homology, followed by 5 non-homologous base pairs. The homologous and non-homologous regions likely correspond to the major and minor grooves of the structure, respectively. Importantly, they showed that PX-DNA can form *in vivo* (Wang et al, 2010). This was expected given that the formation of PX-DNA is topologically equivalent to unwound DNA and therefore it is energetically favorable to form the PX structure in highly superhelical DNA. Data regarding the *in vivo* role of PX-DNA is scarce, however, one protein, *E. coli* DNA polymerase I, was found to specifically bind to the PX-DNA structure (Gao et al., 2019). This observation supports the speculation that PX-DNA might have a biological function.

In this study, we used an experimental yeast system to explore a potential biological role for PX-DNA. We have found that this structure facilitates at least one type of genomic rearrangement, and we propose two models to explain our observations. In addition, we propose several follow-up experiments to confirm the validity of these models and better understand what kind of role PX DNA might serve *in vivo*.

Materials and Methods

Cassette Design

After designing our experimental and control cassettes in SnapGene, a synthesized oligonucleotide was ordered from GenScript as an insert in the pUC57 vector, which contains the gene for ampicillin resistance to be used for selection in bacteria.

Polymerase Chain Reaction (PCR)

For the amplification of sequences to be used for transformation of cells or DNA sequencing, Phusion high-fidelity DNA polymerase (Thermofisher Scientific) was used with 5x high fidelity buffer. For all other reactions, Taq polymerase was used with Sibenzyme green buffer (Sibenzyme). DNA was extracted from yeast by dissolving a small amount of a patch or colony in 2 microliters of lyticase, which was then incubated for ten minutes at 37˚C. Fifty microliters of water were then added to the mixture and the resulting suspension was boiled for 5 minutes in a PCR block. Tubes were then centrifuged to pellet the cells, and 1 microliter of DNA taken from the supernatant was added to a tube of PCR master mix made according to the polymerase manufacturer's protocol.

PCR Product Purification

Purification of DNA from PCR was performed using a GeneJET PCR Purification Kit, following the manufacturer’s protocol (GeneJet).

Yeast Transformation

For cassette integration, our parent strain (SKY1) was grown overnight in 10 mL of liquid YPDUA (yeast-extract peptone dextrose, with added uracil and adenine) media at 30˚C. The culture was then diluted in 40 milliliters of liquid YPDUA (Yeast Peptone Dextrose with added Uracil and Adenine) and allowed to grow for an additional three hours. The resulting culture was centrifuged at 2500 RPM for 2 minutes and the supernatant was discarded. The cells were washed in 25mL of water, centrifuged again, and the supernatant was discarded. The cells were then resuspended in 1mL of 0.1M lithium acetate and transferred into a 1.5mL tube and centrifuged at 3000 RPM for 2 minutes. The supernatant was discarded and 400 microliters of 0.1M lithium acetate were added. A 50 microliter aliquot was taken from this mixture, centrifuged and the supernatant was discarded. To the remaining cells the following was added; 240 microliters of PEG, 36 microliters of 1.0M lithium acetate, and 10 microliters of carrier DNA which had been boiled for five minutes and then cooled on ice. Then, approximately 1000ng of the fragment containing our cassette was added. Finally, water was added to reach a total of 360 microliters. The mixture was incubated at room temperature for 25 minutes and then placed into a 42˚C water bath for exactly 23 minutes. After centrifugation at 3,000 RPM for 1 minute, the supernatant was removed and the cells were gently resuspended in 100 microliters of water and plated on media lacking tryptophan. The plates were incubated for two days at 30˚C. The above protocol was also used to knock out genes of interest from strains containing our experimental cassette.

Fluctuation Test

For determining URA+ rate, a strain was first patched from culture frozen at -80°C onto either solid YPDUA or YPGalUA (Yeast Peptone Galactose with added Uracil and Adenine) media and grown at 30°C. The following day, a small amount of yeast from the patch was suspended in 1000ul of sterile water, and diluted 10,000 times. 100ul of the final dilution was then spread onto the same media type as the original patch. After four days at 30°C, single colonies were inoculated in 50ml of liquid media (either YPDUA or YPGalUA) and grown at 30°C with shaking for 24 hours. Cultures were then transferred to 50ml conical tubes and centrifuged at 2500 RPM for 2 minutes to pellet out the cells. The supernatant was removed and replaced with 5ml of water. Cells were then vortexed to resuspend, and centrifuged again. This wash step was repeated, and the remaining water was removed with a pipette to ensure that the cells were totally dry. This pellet was resuspended in 300ul of sterile water. Ten fold serial dilutions were made up to 10,000,000-fold. 100ul of the 10-fold dilution was plated on selective media, and 100ul of the 10,000,000-fold dilution was plated on complete media. Cells on complete media were allowed to grow for 2 days at 30°C and cells on selective media were allowed to grow for 4 days at 30°C before colonies were counted. The Ura^+ rate was then calculated using FluCalc (Radchenko et al., 2018).

Sanger Sequencing

Sequencing of DNA was performed via Sanger sequencing (after column purification) by GeneWiz.

Strains and Primers Used

Table 1. Yeast Strains Used in This Study

Name	**Genotype**
SKY1	*MATa, leu2-Δ1, trp1-Δ63, ura3-52, his3-200, ade2Δ::KanMX4, HAP1-wt, ura3Δ*
JAYPX2	SKY1*, ChrIII(75594- 75641)::PGal1-UR-PX1-TRP1-PX2-A3*
JAYPX16	JAYPX2, *yku70::HphMX4*
JAYPX17	JAYPX2, *yku70::HphMX4*
JAYPX20	JAYPX2, *rad52::HphMX4*
JAYPX21	JAYPX2, *rad52::HphMX4*
JAYFH1	*SKY1, ChrIII(75594- 75641)::PGal1-UR-PX1-TRP1-PX1-A3*
JAYFH7	JAYFH1, *rad52::HphMX4*
JAYFH8	JAYFH1, *rad52::HphMX4*

Table 2. Primers Used in This Study

Name	Sequence
Misc_PX_F	CTCTCACTTCATATTATTTCACTTATATTTTCCTCTCCACTTCCATCAC GCAATCGTTAGTGCTTTTTATTTCGGATTAGAAGCCGCCGAGC
Misc_PX_R	CTTGCCTTTTTTAAAAATCTACGTTAAGAATTTTTAATTTTGTTAAGGT GTAATTAGCATTAGAATGACTTAGTTTTGCTGGCCGCATCTTCTC
A36b-F	ACGTGTACAGTTCTCTTTACATCATC
Alex-URA3 -RT- Spl-F (75)	ATCCTAGTCCTGTTGCTGCCAA
5URA3rev	TAGAATTGGGACCGTGCAATTCTTC
TRP1_F	TGGAGATGAGTCGTGGCAAG
TRP1_R	TCTCTTGCCTTCCAACCCAG
URA3_1In_ R	CTTCCCTTTGCAAATAGTCCTCTTCC
A36a-R	AGGGTCGTTGCCTTCTGGT
JK183_hygRl eft_rev	ACAGTCACATCATGCCCCTG
Ku70_int_F	GCATGAAGATATCAGACAAGAAGC
Ku70_int_R	TGCTCGATGAACGGAACC
RAD52_chk_ in_F	GCCAAGAAATCTGCCGTTAC
RAD52_chk_ in_R	TGAGCTTTCGCTGATTTCATCC

Results

PX Motifs Promote Genetic Rearrangements

In order to test whether PX-DNA formation could promote genetic rearrangements, we took advantage of the fact that yeast cannot splice out introns greater than 1 kilobase (kb) in length. We placed two PX-forming DNA sequences (PX1 and PX2 motifs), on either side of the *TRP1* gene under the control of its wildtype (WT) promoter inside of an artificial intron in the *URA3* gene. We refer to this cassette as "PX-cassette" (Figure 2A). The PX-forming sequences contain 6 base pairs (bp) of homology, followed by 5 bp of DNA unique to each motif. This 11 bp pattern repeats eight times. The whole intron is ~2kb long, meaning that yeast are unable to splice it out, making the phenotype of strains with this cassette Ura^-, Trp^+. We reasoned that selection for the Ura^+ phenotype would yield colonies with a deletion of the *TRP1* gene, potentially promoted by the PX motifs.

As a positive control, we generated a cassette wherein we replaced the PX2 motif with a copy of the PX1 motif, resulting in two sequences of 83 base pairs of full homology (FH) (Figure 2B). We anticipated a high rate of recombination in this "FH-cassette". As a negative control, we replaced the PX2 motif with a scrambled sequence (PX2S) by randomly shuffling the bases from the PX2 motif. This eliminated all micro-homologies between the two motifs (Figure 2C). In this "Scrambled cassette" we did not expect to see genetic rearrangements of any kind.

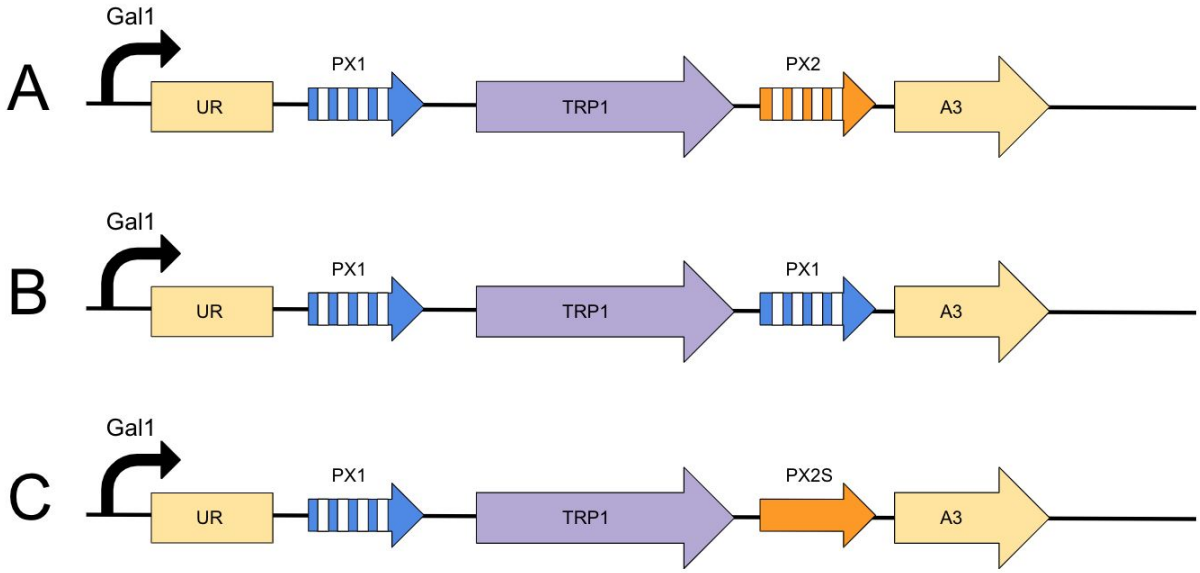

Figure 2. Experimental Cassettes Inserted into S. cerevisiae Chromosome III

Schematics of the cassettes which were inserted into a parent strain (SKY1) lacking the endogenous *URA3* gene. All three cassettes use the *URA3* gene under the Gal1 promoter with an artificial intron containing the functional *TRP1* gene under its wildtype promoter flanked by the PX motifs or their alternatives. **A)** "PX" cassette contains two regions of PX homology, with blue and orange arrows representing 83 base pair PX motifs (PX1 and PX2) containing 6:5 bases of homology:non-homology. White lines represent bases shared between the two motifs.
B) "FH" cassette contains 83 base pairs fully homologous to one another (PX1 and PX1).
C) "Scrambled" cassette contains the PX1 motif on one side of the intron, and a scrambled PX2 motif (PX2S) containing the same bases as the original PX2 motif in a random order.

After inserting our experimental cassettes into the *S. cerevisiae* genome, we performed fluctuation tests to determine the rate of genetic rearrangements in the three cassettes described above. As expected, the Ura^+ rate of the FH cassette was high (~9 x 10^{-6}) and all tested Ura^+ colonies were a product of recombination within the FH sequences. This was evidenced by the PCR pattern obtained from amplifying the DNA from Ura^+ colonies with primers URA3-RT-Spl-F and URA3_1In_R. In the Scrambled cassette, we did not recover any Ura^+ colonies, meaning that the rate of formation of Ura^+ colonies in this cassette is smaller than our detection threshold (2.5 x 10^{-10}) (Figure 3).

In the PX-cassette, the Ura$^+$ rate was rather small, but still detectable (~9x10^{-9})(Figure 3). Therefore, PX-motifs likely promote genetic rearrangements, although not as efficiently as the FH sequences.

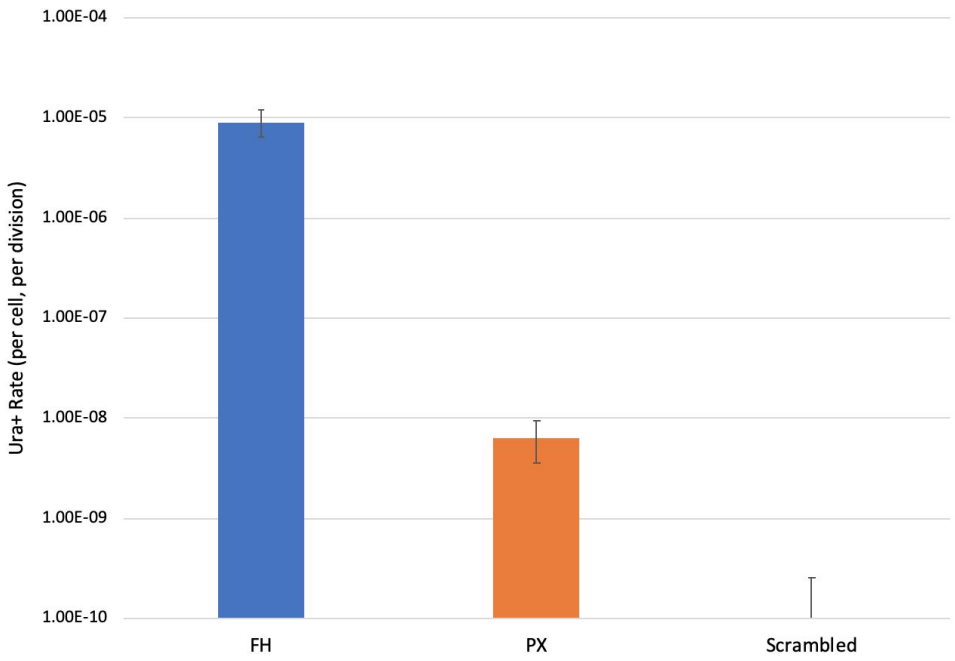

Figure 3. Rate of Events Leading to Ura$^+$ Colonies in Three Experimental Cassettes

URA$^+$ colony formation rate in the Full Homology, Scrambled, and PX cassettes (see Figure 1). Error bars denote 95% confidence intervals. Rates were considered significantly different if 95% confidence intervals did not overlap. For the scrambled cassette, no growth was detected on selective media, so the level of detection is plotted.

PX-promoted rearrangements do not depend on the level of transcription or transcription-replication collisions

We wondered whether transcription through our cassette or transcription-replication collisions might affect the rate of PX-promoted rearrangements. Transcription generates supercoiling in DNA (Liu and Wang, 1987), which might promote formation of the PX-DNA structure. Therefore, we decided to test whether higher transcription through the PX cassette

would increase the Ura$^+$ rate. Our cassette is under the control of the Gal1 promoter, so we were able to induce transcription by culturing cells in media containing galactose. Surprisingly, we saw no significant difference in Ura$^+$ rate between glucose repression and galactose induction conditions (Figure 4A).

Replication-transcription collisions can induce DSBs or promote recombination events through transcription-associated recombination (TAR) (reviewed by Aguilera, 2002), and head-to-head collisions are considered to be more detrimental than head-to-tail collisions (French, 1992). We initially integrated our cassettes near a well-characterized, efficient origin of replication, ARS305, such that replication-transcription collisions occur in a head-to-tail manner. We wondered whether forcing head-to-head collisions would increase the rate of rearrangements. To do this, we inverted the PX cassette with respect to the origin of replication. Unexpectedly, we saw no significant difference in the Ura$^+$ rate when we inverted the cassette so that replication and transcription were proceeding in opposite directions (Figure 4B). We concluded that the amount of transcription does not play a role in PX-promoted rearrangements, nor do transcription-replication collisions.

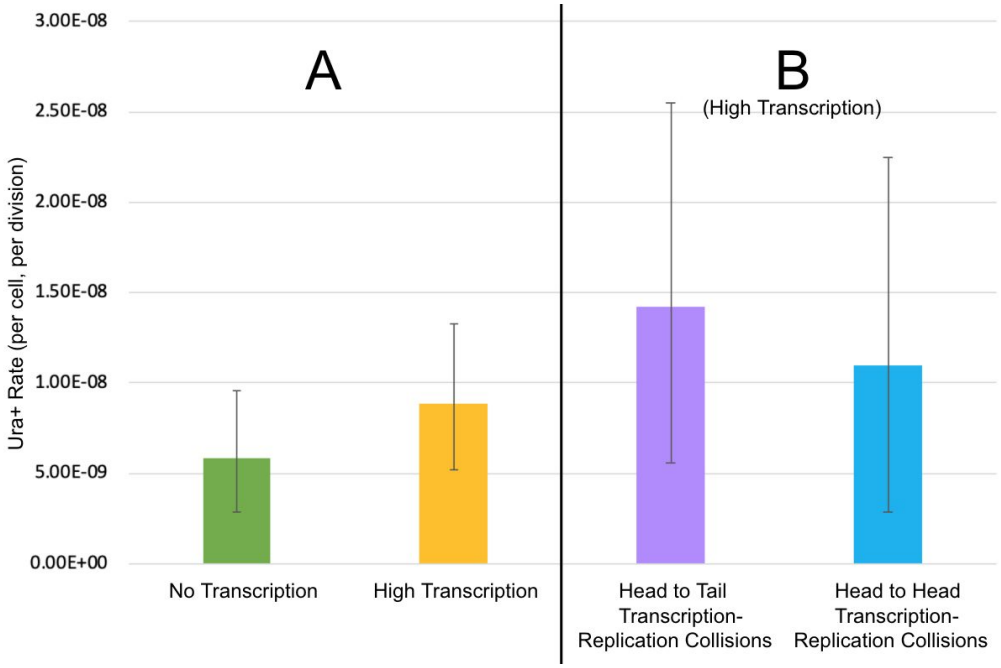

Figure 4. The Effect of Transcription and Replication on URA+ Colony Formation

A) Role of transcription in the Ura$^+$ rate of the PX cassette. Error bars denote 95% confidence intervals. Rates were considered significantly different if 95% confidence intervals did not overlap. **B)** Role of transcription-replication collisions orientation in the Ura$^+$ rate of PX cassette in the high transcription state. Error bars denote 95% confidence intervals. Rates were considered significantly different if 95% confidence intervals did not overlap.

PX Motifs Promote Different Events than Fully Homologous Sequences

In order to better understand the nature of events that led to the Ura$^+$ phenotype in the PX cassette, we amplified the entire intron of our cassette using primers annealing inside the *URA3* gene. We anticipated that these Ura$^+$ colonies would represent a deletion of the *TRP1* gene using microhomologies within the PX motifs, which would result in a hybrid PX1/PX2 motif. In this case, amplification of this region using primers URA3-RT-Spl-F and URA3_1In_R would produce a PCR product of 1173 base pairs (Figure 5A), similar to what we observed for Ura$^+$ colonies originated from strains with the FH cassette. Unexpectedly, we saw that the length of the PCR product in the majority of Ura$^+$ colonies recovered from strains with the PX cassette was

shorter than the predicted size. This meant that the PX-cassette was undergoing different events than the FH-cassette (Figure 5B).

Indeed, following Sanger sequencing of our amplified DNA fragments, we saw that most of the PX colonies were recombining at an AATT microhomology 53 bp ustream and 7 bp downstream of the PX1 and PX2 motifs, respectively (Figure 5A). After screening over 300 colonies, we calculated that approximately 91% of Ura^+ PX colonies recombined at the AATT microhomology, while the other 9% recombined inside of the PX motif. Herein, we refer to the events that involve the AATT microhomology as "AATT-mediated deletions", and to the events that produce hybrid PX motifs as "within-PX rearrangements". Since no AATT-mediated deletions were detected in the Scrambled cassette, we concluded that these events depend on the presence of the PX motifs.

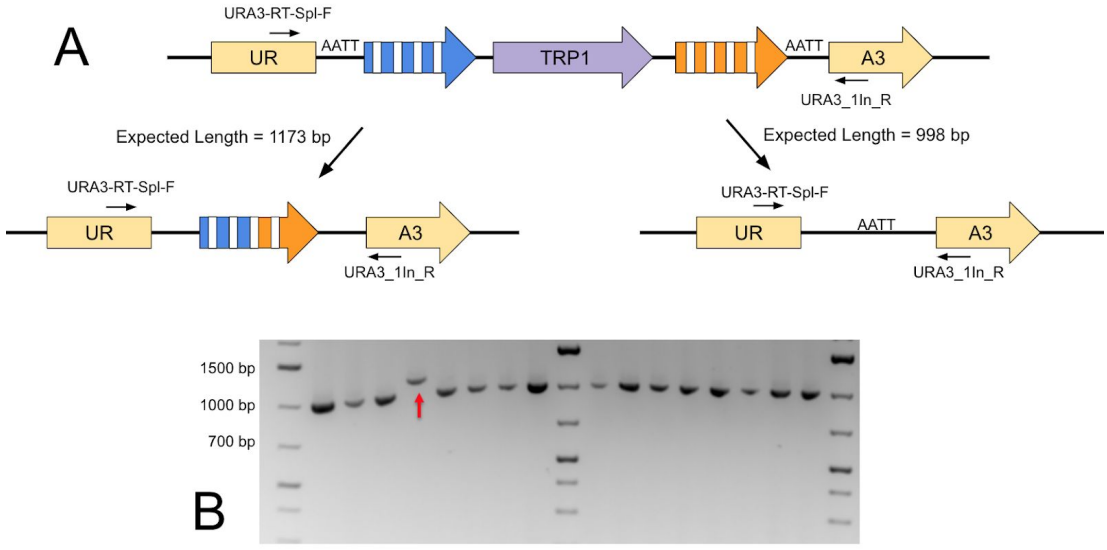

Figure 5. Analysis of Events Leading to URA+ Colonies

A) Two events can lead to Ura^+ colonies, yielding to two different sized bands when introns are amplified with primers URA3-RT-Spl-F and URA3_1In_R **B)** A section of a gel with PCR products amplified from the Ura^+ colonies containing the PX cassette. Most URA^+ colonies from the PX cassette recombine at an AATT microhomology resulting in a band of 994bp, but occasionally recombination occurs within the PX motif resulting in a band of 1173 bp, indicated by a red arrow.

Rad52 promotes AATT-mediated deletions

After our surprising discovery that two distinct types of events happen in the PX cassette, we wanted to establish the genetic control for both of those events. To begin, we decided to test two genes, *RAD52* and *YKU70*. These genes play an essential role in homologous recombination (HR) and non-homologous end joining (NHEJ,) respectively. Surprisingly, we observed no significant change in the Ura$^+$ rate after knocking out either of these genes (Figure 6).

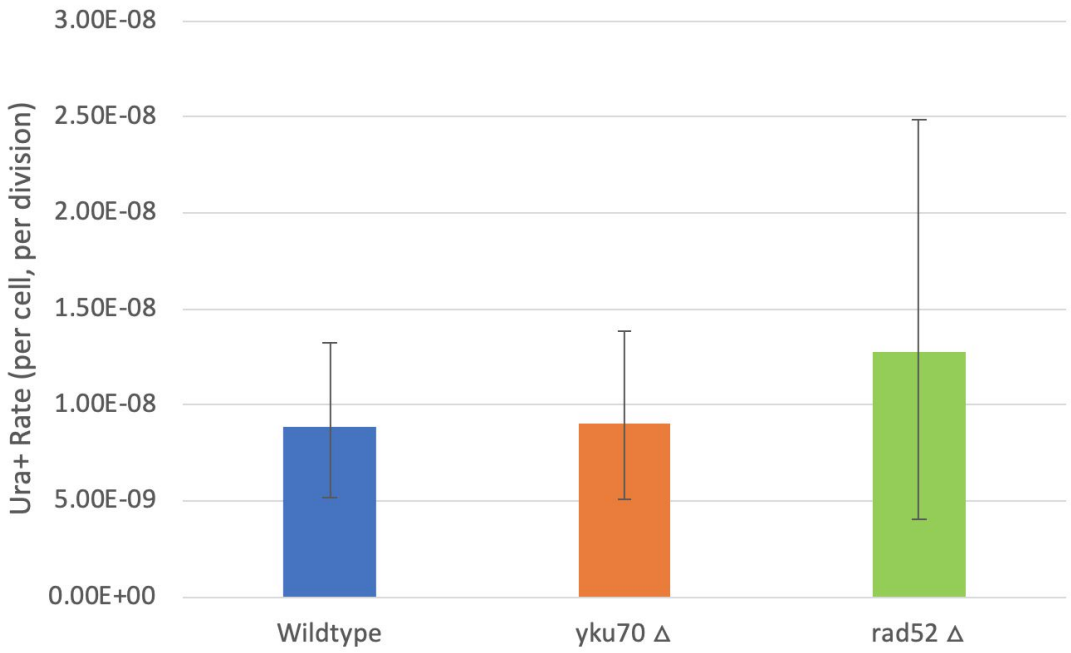

Figure 6. Ura+ Rate Following Deletion of Yku70 or Rad52

Role of Yku70 and Rad52 in the Ura$^+$ rate of the PX cassette. Error bars represent 95% confidence intervals. Rates were considered significantly different if 95% confidence intervals did not overlap.

We then amplified the introns of recovered Ura$^+$ colonies to see what type of events were happening in the *yku70*Δ and *rad52*Δ mutants. Remarkably, we found that none of the Ura$^+$ colonies (N=15) in the *rad52Δ* strain had undergone AATT-mediated deletions (Figure 7). This implies that AATT-mediated deletions are fully dependent on one of the HR-subpathways. On

the other hand, the rate of within-PX rearrangements was slightly increased in the *rad52Δ* mutants compared to the WT. Therefore, within-PX rearrangements appear to be somewhat counteracted by Rad52. Interestingly, the deletion of *YKU70* affected neither the rate of within-PX rearrangements nor AATT-mediated deletions. This means that NHEJ does not play a role in either of these processes.

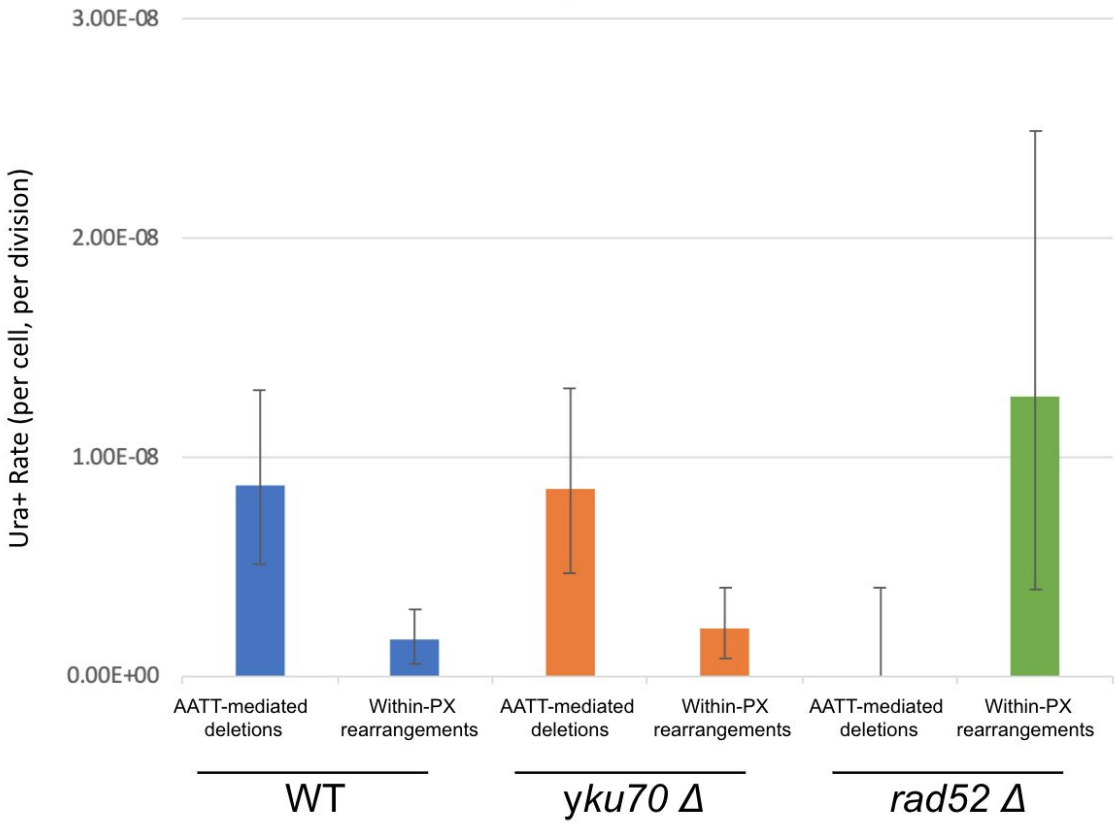

Figure 7. Rate of Different Events Leading to Ura+ Colony Formation
Role of Yku70 and Rad52 in the formation of various events in the PX cassette. Error bars represent 95% confidence intervals. In the *rad52Δ* mutant, no AATT-mediated deletions were observed, so the level of detection is plotted.

We then wondered what the genetic mechanism was for the within-PX rearrangements in the FH cassette. We reasoned that in this cassette all the Ura$^+$ colonies would be a product of HR between the two identical motifs of 83 bp. As expected, we observed a marked decrease in the

Ura$^+$ rate compared to the wild type FH cassette. However, the Ura$^+$ rate in the *rad52Δ* FH cassette was still an order of magnitude higher than in the wild type PX cassette (Figure 8).

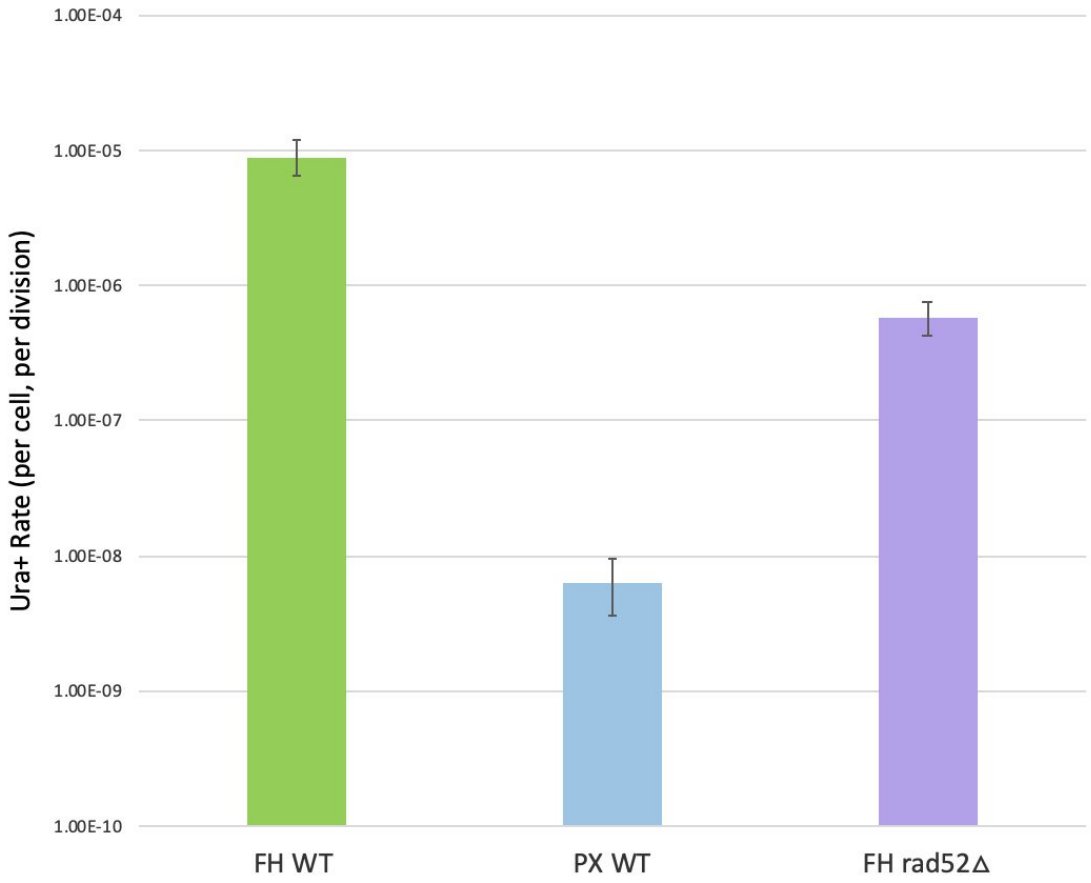

Figure 8. Effect of Rad52 Deletion in Fully Homologous Cassette
Role of Rad52 in the Ura$^+$ rate of the FH cassette. Error bars represent 95% confidence intervals. Rates were considered significantly different if 95% confidence intervals did not overlap.

Taking all of our experimental evidence together, we have found that (i) PX-promoted rearrangements do not happen via NHEJ, (ii) AATT-mediated deletions require Rad52, and (iii) within-PX rearrangements can happen in the absence of Rad52 in both the PX and FH cassettes.

Discussion

PX-DNA Forming Sequences Promote Two Types of Rearrangements

PX-DNA is an unusual DNA structure which forms both *in vitro* and *in vivo* at regions of double stranded DNA homology. Although PX-DNA has applications in DNA nanotechnology, little is known about a biological role for the structure. Because PX-DNA consists of Holliday junctions, we hypothesize that this structure might play a role in genetic rearrangements. Indeed, we have identified that, in the presence of PX-DNA forming sequences, two types of genetic rearrangements can happen: forming junctions both within (within-PX rearrangements) and outside the PX motifs (AATT-mediated deletions). Moreover, we have shown that at least one of these events, the AATT-mediated deletion, is promoted by PX motifs.

There remains the question of what mechanisms lead to these two types of rearrangements. One possibility which we considered was that the formation of PX-DNA was leading to double strand breaks that could undergo incorrect repair, resulting in intron shortening and the Ura+ phenotype that we observe. Based on what we know about the mutagenic repair of double strand breaks (reviewed by Rodgers and McVey, 2015), the most likely mechanisms were non-homologous end joining (NHEJ) which is Yku70-dependent, single strand annealing (SSA) which is Rad52-dependent, and microhomology mediated end joining (MMEJ) which depends on neither Rad52 nor Yku70. Another possibility was that DSBs were being repaired through break induced replication (BIR), which depends on Rad52 and Pol32. We hypothesize that a template switch during BIR could lead to AATT mediated deletions. This process would be dependent on Rad5.

We were quickly able to rule out NHEJ because yeast with a y*ku70Δ* genotype showed no significant difference in the rate of either event. Yku70 is an end capping protein and is an absolute requirement for NHEJ in *S. cerevisiae* (Milne et al., 1996). Since PX-promoted rearrangements were completely Yku70 independent, we began to suspect that MMEJ could be responsible for our observations. MMEJ is an error prone DNA repair pathway which is similar to Rad52-dependent SSA. It involves a double strand break in DNA, followed by resection to a region of homology and reannealing of the newly generated sticky DNA. MMEJ, however, differs from SSA in the fact that it is completely Rad52-independent (McVey and Lee, 2008).

Within-PX Rearrangements are Consistent with MMEJ

We found that within-PX rearrangements are independent of Yku70 and might be slightly counteracted by Rad52. This is consistent with the MMEJ process as there is experimental evidence that a deletion of *RAD52* can increase the rate of MMEJ events using 6 bp of homology by approximately 5 fold (Villarreal et al., 2012). In our system, we observe an increase of about 7 fold (Figure 7). Our model for MMEJ-mediated within-PX rearrangements is presented in Figure 9.

The role of MMEJ in within-PX rearrangements is also supported by our data in the FH cassette (Figure 8). When we deleted *RAD52* in the FH cassette, we saw a decreased, but still substantial rate of Ura^+ colonies, all of which likely underwent within-PX rearrangements. We hypothesize that within-PX rearrangements in the *rad52*Δ mutant in the FH cassette were also a product of MMEJ. In the FH cassette, the Ura^+ rate was over an order of magnitude higher than in the PX cassette. We can explain this by the availability of a much larger region of homology for MMEJ in the FH cassette (83 bp) when compared to the PX cassette (6 bp).

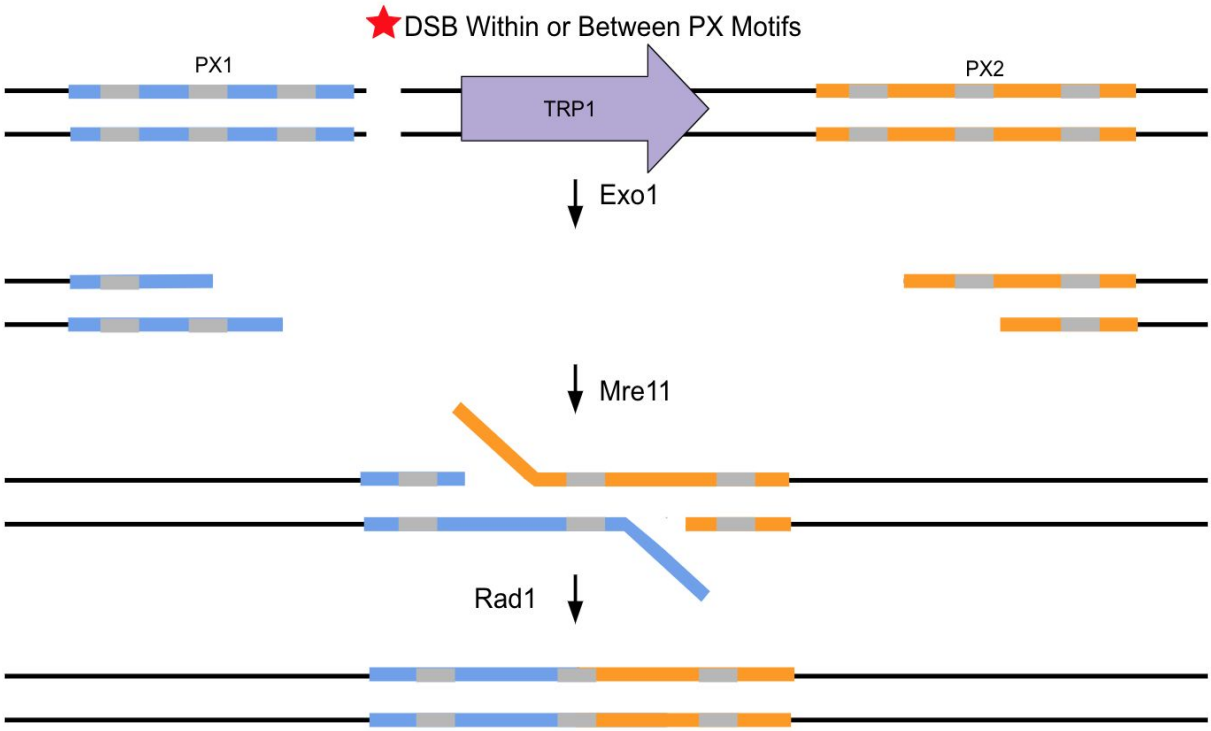

Figure 9. Proposed Mechanism for Recombination Events Occurring Inside PX Motifs
Blue and orange regions represent PX motifs. Grey bands represent regions of 6 base pair homology. In actuality there are 8 such regions, but for simplicity only three are drawn here.

To confirm that within-PX rearrangements happen via MMEJ, we plan to measure the rate of these events in *mre11*Δ and *rad1*Δ strains, as these two genes are essential for MMEJ in *S. cerevisiae* (Ma et al., 2003). MMEJ might be triggered by a random DSB in between the two PX motifs. In this case, over 1kb of DNA resection is required. This amount of resection is likely not possible with only the exonuclease activity of Mre11, so we imagine this pathway might also be reliant on Exo1, which can create larger deletions (Lam et al., 2008).

We have not detected any within-PX rearrangements in the scrambled cassette, which could be either because there are no microhomologies present, or because there is no potential for the formation of PX-DNA. Therefore, it is not yet clear whether the PX-DNA structure promotes these events. One possibility for the involvement of PX-DNA in this process might be that this structure facilitates formation of DSBs in between the two PX motifs. It is possible that these DSBs are a result of cleavage of the PX-DNA structure by a Holliday junction resolvase, such as Mus81 or Yen1 (Gaskell et al., 2007, Ip et al., 2008). This cleavage would create a DSB flanked by the two PX motifs.

In order to unambiguously determine whether it is the presence of the microhomologies or the formation of PX-DNA that allows for the within-PX rearrangements, we plan to use a sequence with the same microhomologies as in the PX-motifs but which is unable to form the PX-DNA structure. After consulting with Dr. Nadrian Seeman, we decided to use a PX motif with 6:10 base pairs of homology:non-homology. These PX-motifs will form the PX structure very poorly, if at all, but will contain the same length of microhomology as our original cassette. If MMEJ does not depend on the formation of PX-DNA, we expect to still observe rearrangements within the PX motifs. However, if the PX structure itself plays a role in the formation of the DSB that triggers within-PX rearrangements, we expect to detect a decrease in the rate of within-PX rearrangements.

If the experiment described above shows that PX-DNA plays a role in within-PX rearrangements, we will test deletions of Holliday junction resolvases, Mus81 and Yen1, in our system. If knockouts of these genes show an effect on the frequency of within-PX

rearrangements, it would mean that the cleavage of PX-DNA promotes a DSB which leads to rearrangements.

AATT-Mediated Deletions May Occur by Template Switching During BIR

Unlike within-PX rearrangements, AATT-mediated deletions were completely absent in the *rad52*Δ mutant. This indicates that within-PX rearrangements and AATT-mediated deletions are two fundamentally different processes. A genetic rearrangement with a 4 bp homology is unlikely to be a result of conventional HR pathways, such as synthesis-dependent strand annealing or SSA. We propose, instead, that AATT-mediated deletions occur via a template switch during break-induced replication (BIR) triggered by a random DSB upstream or downstream of the cassette (Figure 10). We believe that the formation of PX-DNA is essential for this model as no AATT-mediated deletions were detected in the scrambled cassette.

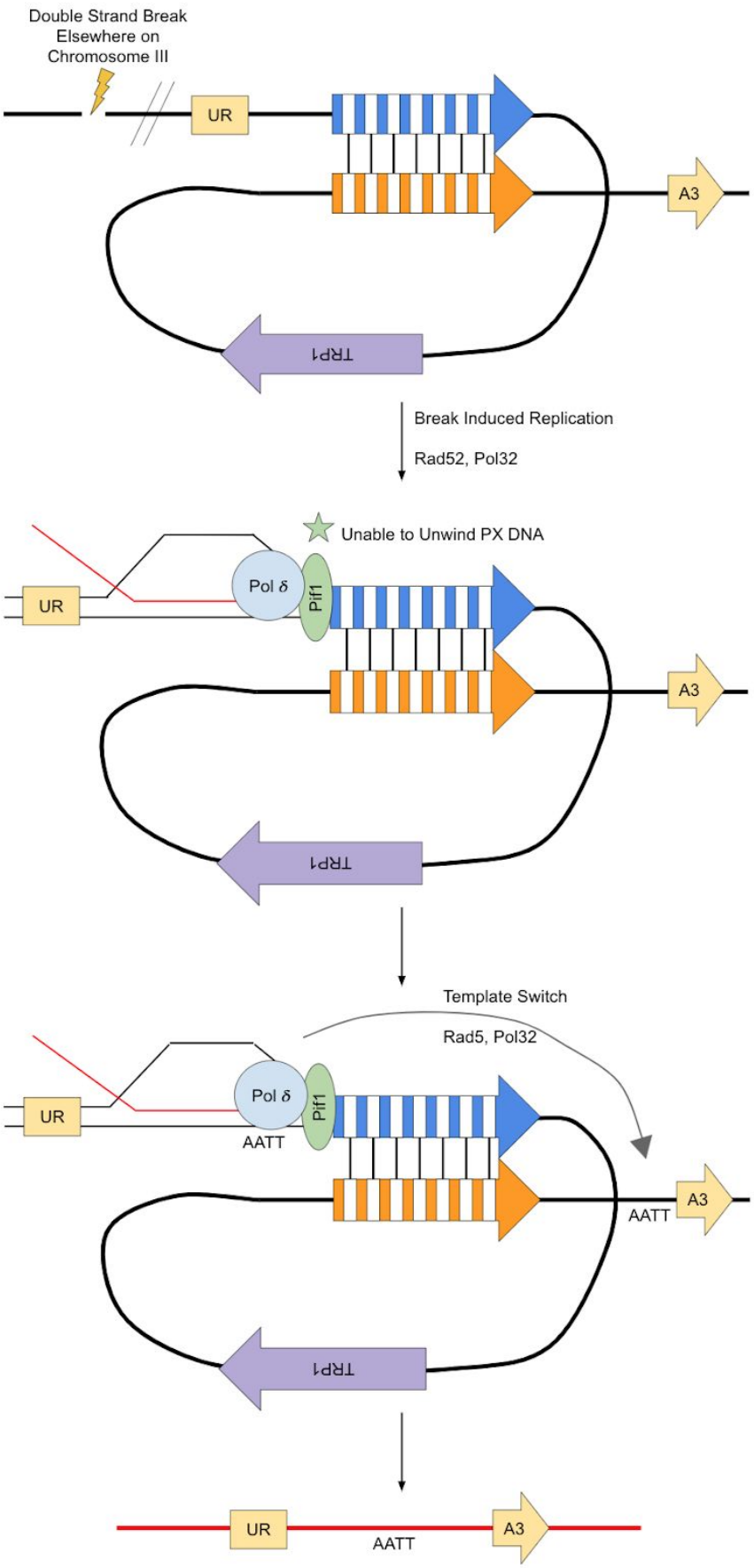

Figure 10. Proposed Mechanism for Recombination Events at AATT Microhomology

If a random DSB occurs outside of the PX motifs (blue and orange arrows, white lines represent regions of 6 base pair microhomology) on chromosome III, BIR could initiate. The resulting migrating bubble stalls on the PX-DNA structure as Pif1 is unable to unwind it. This could lead to a template switch using the AATT microhomology on the opposite side of the structure, resulting in the deletion of both PX motifs as well as the TRP1 gene in between them in the nascent DNA molecule (red strands).

BIR involves the asynchronous conservative synthesis of DNA strands via a migrating bubble (D-Loop) which uses different replication machinery than normal DNA replication (Saini et al., 2013). We imagine that the helicase associated with BIR, Pif1, could be unable to unwind PX DNA. Since the BIR machinery is prone to template switching (Smith et al., 2007), it might use the AATT microhomology as a template to bypass the PX-DNA structure entirely. This BIR model accounts for the absolute Rad52 dependence of AATT-mediated deletions, as BIR cannot take place in the absence of Rad52 (Malkova et al., 1996). It also explains why the level of transcription through the PX motifs does not have an effect on the rate, nor does the orientation of the cassette with respect to the origin of replication, as the formation of a DSB in this scenario happens outside the cassette.

Similar to the proposed mechanism for within-PX rearrangements, it is theoretically possible that the DSB which initiates BIR is formed by the cleavage of PX-DNA by structure-specific nucleases such as Mus81 or Yen1. In this scenario, the break is flanked by the AATT microhomology, which could be used for repair through BIR. If this is the case, the rate of AATT-mediated deletions would depend on the presence of the relevant nucleases. However, this is not a likely scenario because it is unclear why a DSB would not be repaired by MMEJ or NHEJ. This model also assumes that the first step of BIR would use only 4 bp of homology for Rad52/Rad51-dependent strand invasion, when Rad51 generally requires at least 8 bp of homology to complete this step successfully (Qi et al., 2015).

In either case, we propose that AATT-mediated deletions arise via BIR. This hypothesis is conditional on several genes which we plan to test. Pol32 is an absolute requirement for BIR (Lydeard et al., 2007), so we predict that we would no longer observe AATT-mediated deletions

in a *pol32*Δ mutant, resulting in a phenotype similar to *rad52*Δ. This knockout will allow us to unambiguously establish the role of BIR in AATT-mediated deletion. Next, we plan to test a knockout of Rad5, as this gene is required for template switching (Blastyák et al., 2007). This experiment will help us determine whether it is a template switch during BIR or AATT-mediated strand invasion that leads to AATT-mediated deletions.

Additionally, we plan to invert the first PX motif inverted relative to the rest of the cassette such that the two PX motifs would run in a head to head orientation (Figure 11). We would then measure the Ura^+ rate in this inverted cassette. We predict that head to head PX motifs should allow PX-DNA to form with higher efficiency than in our original cassette (Wang et al., 2010). However, the within-PX microhomologies would no longer be in the correct orientation to anneal during MMEJ. Thus, the only events we would anticipate observing are AATT-mediated deletions.

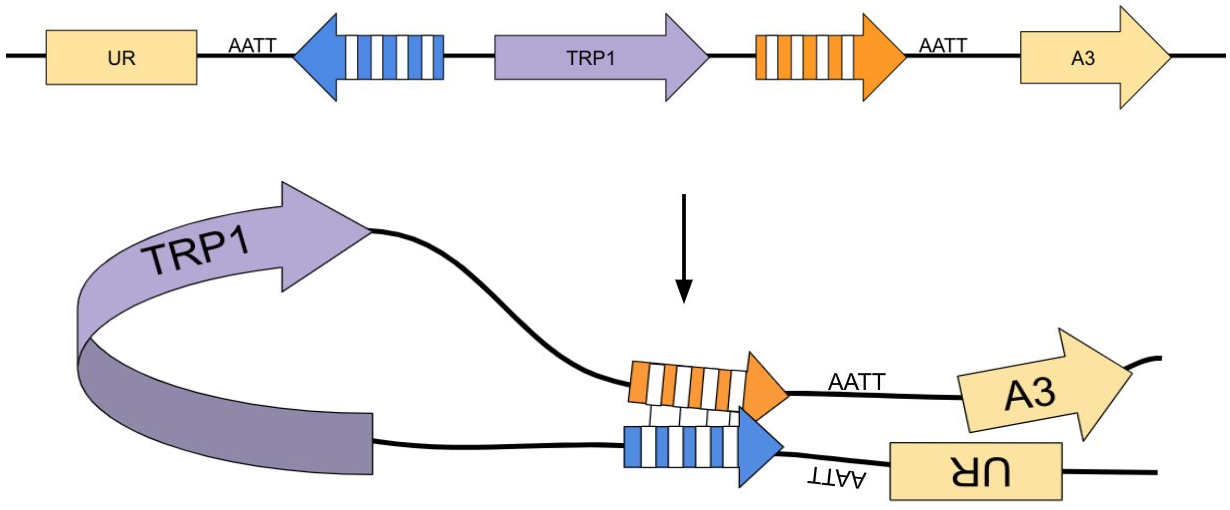

Figure 11. Schematic of Head to Head Orientation of PX Motifs and Resultant Structure
By inverting the PX1 motif (blue) and putting the PX motifs in the head to head orientation, PX DNA formation might result in the structure shown.

Future directions

The inspiration for this project, and the ultimate goal, is to determine whether PX DNA has a role in the pairing of chromosomes. Therefore, as we take this project further into the future, it will be important to investigate the consequences of PX-DNA forming between two chromosomes. As a first step, we have devised an experiment based on previous data obtained in our lab which would allow us to quantify physical interaction between PX motifs on separate chromosomes. We know from other projects in our lab that the *URA3* gene which we inserted on chromosome III can recombine with the *ura3-52* allele at the endogenous location on chromosome V in strains with this genotype (McGinty et al., 2017). This might mean that these two chromosomal regions on chromosomes III and V come into close proximity with one another, which we intend to use to our advantage in our studies of PX-DNA. In this experiment, we would place only the first half of the *URA3* gene and the artificial intron containing the PX1 motif on chromosome III. We would then place the PX2 motif and the other half of the artificial intron and *URA3* gene on chromosome V (Figure 12).

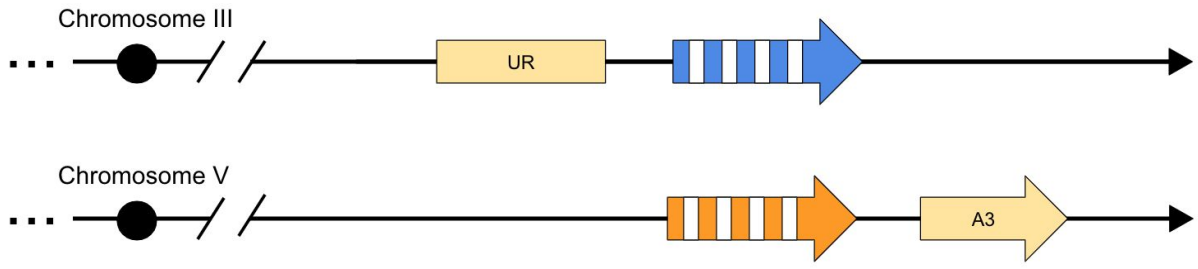

Figure 12. Proposed System for Studying PX-DNA Formation Between Chromosomes
PX motifs, represented by blue and orange arrows (with white bars representing 6 bp regions of homology) and halves of the *URA3* gene with an artificial intron to be placed on chromosomes III and V. Black circles represent centromeres. Arrowheads represent telomeres.

A genetic rearrangement between these motifs by some mechanism could result in restoration of the functional *URA3* gene. If we fail to detect the rate of these events, we will facilitate them by inducing a double strand break using CRISPR/Cas9. Additionally, we plan to try and find direct physical evidence of PX-DNA formation using chromosome conformation capture, otherwise known as 3C (Dekker et al., 2002). By cross-linking base paired DNA and generating a recombinant molecule through restriction digest and ligation, we would be able to quantify the interaction between PX sequences using qPCR. This would allow us to directly measure the level of the physical interaction between PX motifs on two separate chromosomes.

Conclusions

Here we demonstrate promising preliminary results about a potential biological role for the PX-DNA structure. We have shown that the presence of PX-DNA forming sequences leads to genetic rearrangements which occur in two distinct fashions. The majority of events we observe are AATT-mediated deletions, which occur outside of the PX motifs at an AATT microhomology. We propose that these events are taking place during BIR. PX-DNA appears to play a direct role either in the formation of a BIR-initiating DSB or in promoting template switching between the two AATT microhomologies adjacent to the PX-motifs. In addition, we have observed genetic rearrangements within the PX motifs. We believe that these events are occurring via MMEJ, which might also be initiated by a DSB caused by PX-DNA formation.

In future experiments, we plan to investigate the involvement of the PX-DNA structure in these processes by using sequences with different potentials for PX-DNA formation. We also plan to test our mechanistic hypothesis through a series of gene knockouts. Additionally, we will examine the genetic and physical interaction between two PX motifs by placing them on two separate chromosomes and conducting genetic analysis and 3C experiments.

Additional experiments will be required to understand the precise involvement of PX-DNA in *in vivo* processes, but our preliminary data suggest that PX-DNA plays some important biological role, potentially in the pairing of homologous chromosomes.

Literature Cited

Aguilera A. (2002). The connection between transcription and genomic instability. *The EMBO Journal*, *21*(3), 195–201. doi: 10.1093/emboj/21.3.195

Blastyák, A., Pintér, L., Unk, I., Prakash, L., Prakash, S., & Haracska, L. (2007). Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. *Molecular cell*, *28*(1), 167–175. doi: 10.1016/j.molcel.2007.07.030

Dekker, J. (2002). Capturing Chromosome Conformation. *Science*, *295*(5558), 1306–1311. doi: 10.1126/science.1067799

French, S. (1992). Consequences of replication fork movement through transcription units in vivo. *Science*, *258*(5086), 1362–1365. doi: 10.1126/science.1455232

Gao, X. P., Gethers, M. A., Han, S. P., Goddard, W. C., Cunningham, R. undefined, & Seeman, N. undefined. (2019). The PX Motif of DNA Binds Specifically to Escherichia coli DNA Polymerase I. *Biochemistry*, *58*(6), 575–581. doi: 10.1021/acs.biochem.8b01148

Gaskell, L. J., Osman, F., Gilbert, R. J. C., & Whitby, M. C. (2007). Mus81 cleavage of Holliday junctions: a failsafe for processing meiotic recombination intermediates? *The EMBO Journal*, *26*(7), 1891–1901. doi: 10.1038/sj.emboj.7601645

Ip, S. C. Y., Rass, U., Blanco, M. G., Flynn, H. R., Skehel, J. M., & West, S. C. (2008). Identification of Holliday junction resolvases from humans and yeast. *Nature*, *456*(7220), 357–361. doi: 10.1038/nature07470

Keeney, S., Giroux, C. N., & Kleckner, N. (1997). Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. *Cell*, *88*(3), 375–384. doi: 10.1016/s0092-8674(00)81876-0

Kmiec, E. B., & Holloman, W. K. (1986). Homologous pairing of DNA molecules by Ustilago rec1 protein is promoted by sequences of Z-DNA. *Cell*, *44*(4), 545–554. doi: 10.1016/0092-8674(86)90264-3

Koszul, R., Meselson, M., Doninck, K. V., Vandenhaute, J., & Zickler, D. (2012). The Centenary of Janssens's Chiasmatype Theory. *Genetics*, *191*(2), 309–317. doi: 10.1534/genetics.112.139733

Lam, A. F., Krogh, B. O., & Symington, L. S. (2008). Unique and overlapping functions of the Exo1, Mre11 and Pso2 nucleases in DNA repair. *DNA Repair*, *7*(4), 655–662. doi: 10.1016/j.dnarep.2007.12.014

Lydeard, J., Jain, S., Yamaguchi, M. *et al.* Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature* 448, 820–823 (2007). doi: 10.1038/nature06047

Liu, LF, Wang, JC. Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA*. 1987;84(20):7024–7027. doi:10.1073/pnas.84.20.7024

Liu, Y., West, S. Happy Hollidays: 40th anniversary of the Holliday junction. *Nat Rev Mol Cell Biol* 5, 937–944 (2004). doi: 10.1038/nrm1502

Malkova, A., Ivanov, E. L., & Haber, J. E. (1996). Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proceedings of the National Academy of Sciences*, *93*(14), 7131–7136. doi: 10.1073/pnas.93.14.7131

Ma, J-L., Kim, E. M., Haber, J. E., & Lee, S. E. (2003). Yeast Mre11 and Rad1 Proteins Define a Ku-Independent Mechanism To Repair Double-Strand Breaks Lacking Overlapping End Sequences. *Molecular and Cellular Biology*, *23*(23), 8820–8828. doi: 10.1128/mcb.23.23.8820-8828.2003

Mazur, A. K., Nguyen, T.-S., & Gladyshev, E. (2020). Direct Homologous dsDNA–dsDNA Pairing: How, Where, and Why? *Journal of Molecular Biology*, *432*(3), 737–744. doi: 10.1016/j.jmb.2019.11.005

Mcginty, R. J., Rubinstein, R. G., Neil, A. J., Dominska, M., Kiktev, D., Petes, T. D., & Mirkin, S. M. (2017). Nanopore sequencing of complex genomic rearrangements in yeast reveals mechanisms of repeat-mediated double-strand break repair. *Genome Research*, *27*(12), 2072–2082. doi: 10.1101/gr.228148.117

McVey, M, Lee, SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet*. 2008;24(11):529–538. doi:10.1016/j.tig.2008.08.007
Milne GT, et al. (1996) Mutations in two Ku homologs define a DNA end-joining repair pathway in Saccharomyces cerevisiae. *Mol Cell Biol* 16(8):4189-98

Qi, Z., Redding, S., Lee, J. Y., Gibb, B., Kwon, Y., Niu, H., … Greene, E. C. (2015). DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination. *Cell*, *160*(5), 856–869. doi: 10.1016/j.cell.2015.01.029

Radchenko, E. A, et al. (2018). Quantitative Analysis of the Rates for Repeat-Mediated Genome Instability in a Yeast Experimental System. *Methods in Molecular Biology, 1672,* 421-438.

Rodgers, K., & Mcvey, M. (2015). Error-Prone Repair of DNA Double-Strand Breaks. *Journal of Cellular Physiology*, *231*(1), 15–24. doi: 10.1002/jcp.25053

Saini, N., Ramakrishnan, S., Elango, R., Ayyar, S., Zhang, Y., Deem, A., … Malkova, A. (2013). Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature*, *502*(7471), 389–392. doi: 10.1038/nature12584

Shen, Z., Yan, H., Wang, T., & Seeman, N. C. (2004). Paranemic Crossover DNA: A Generalized Holliday Structure with Applications in Nanotechnology. *Journal of the American Chemical Society*, *126*(6), 1666–1674. doi: 10.1021/ja038381e

Shinohara, A., Ogawa, H., & Ogawa, T. (2002). Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell 69(3),* 457-470. doi:10.1016/0092-8674(92)90447-K

Smith, C., Llorente, B. & Symington, L. (2007) Template switching during break-induced replication. *Nature. 447*, 102–105 . doi: 10.1038/nature05723

Varga, T., & Aplan, P. D. (2005). Chromosomal aberrations induced by double strand DNA breaks. *DNA repair*, *4*(9), 1038–1046. https://doi.org/10.1016/j.dnarep.2005.05.004

Villarreal, D. D., Lee, K., Deem, A., Shim, E. Y., Malkova, A., & Lee, S. E. (2012). Microhomology Directs Diverse DNA Break Repair Pathways and Chromosomal Translocations. *PLoS Genetics*, *8*(11). doi: 10.1371/journal.pgen.1003026

Weiner, B., & Kleckner, N. (1994). Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell*, *77*(7), 977–991. doi: 10.1016/0092-8674(94)90438-3

Wang, X., Chandrasekaran, A.R., Shen Z., Ohayon, Y.P., Wang, T., Kizer, M.E., Sha R., Mao, C., Yan, H., Zhang, X., Liao, S., Ding, B., Chakraborty, B., Jonoska, N., Niu, D., Gu, H., Chao, J., Gao, X., Li, Y., Ciengshin, T., and Seeman, N.C. (2019). Paranemic Crossover DNA: There and Back Again. *Chemical Reviews*, *119* (10), 6273-6289 doi: 10.1021/acs.chemrev.8b00207

Wang, X., Zhang, X., Mao, C., & Seeman, N. C. (2010). Double-stranded DNA homology produces a physical signature. *Proceedings of the National Academy of Sciences*, *107*(28), 12547–12552. doi: 10.1073/pnas.1000105107