Published in final edited form as:

J Mol Biol. 2013 January 23; 425(2): 232–243. doi:10.1016/j.jmb.2012.11.016.

Mechanism of Replicative DNA Polymerase Delta Pausing and a Potential Role for DNA Polymerase Kappa in Common Fragile Site Replication

Erin Walsh^{1,2}, Xiaoxiao Wang³, Marietta Y. Lee³, and Kristin A. Eckert^{1,2,*}

¹Cellular and Molecular Biology Graduate Program, Penn State University College of Medicine. 500 University Drive, Hershey, PA 17033.

²Department of Pathology, Jake Gittlen Cancer Research Foundation, Penn State University College of Medicine, 500 University Drive Mailcode H059, Hershey, PA 17033.

³Department of Biochemistry and Molecular Biology, New York Medical College, 40 Sunshine Cottage Road, Valhalla, NY 10595.

Abstract

Common fragile sites (CFS) are hotspots of chromosomal breakage, and CFS breakage models involve perturbations of DNA replication. Here, we analyzed the contribution of specific repetitive DNA sequence elements within CFS to the inhibition of DNA synthesis by replicative and specialized DNA polymerases. The efficiency of *in vitro* DNA synthesis was quantitated using templates corresponding to regions within FRA16D and FRA3B harboring AT-rich microsatellite and quasi-palindrome (QP) sequences. QPs were predicted to form stems of ~75-100% self homology, separated by 3-9 bases of intervening sequences. Analysis of DNA synthesis progression by human DNA polymerase δ (Pol δ) demonstrated significant synthesis perturbation both at $[A]_n$ and $[TA]_n$ repeats in a length-dependent manner, and at short (<40 basepairs) QP sequences. DNA synthesis by the Y-family polymerase κ was significantly more efficient than Pol δ through both types of repetitive elements. Using DNA trap experiments, we show that Pol δ pauses within CFS sequences are sites of enzyme dissociation, and dissociation was observed in the presence of RFC-loaded PCNA. We propose that enrichment of microsatellite and QP elements at CFS regions contributes to fragility by perturbing replication through multiple mechanisms, including replicative DNA polymerase pausing and dissociation. Our finding that Pol δ dissociates at specific CFS sequences is significant, since dissociation of the replication machinery and inability to efficiently recover the replication fork can lead to fork collapse and/or formation of double-strand breaks *in vivo*. Our biochemical studies also extend the potential involvement of Y-family polymerases in CFS maintenance to include polymerase κ .

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding author: Phone: 717-531-4065 Fax: 717-531-5634 kae4@psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCESSION NUMBERS

FRA16D GenBank Accession Number: [AF217490](#), FRA3B Genbank Accession Number: [183583557](#).

INTRODUCTION

Genomic instability is a driving force of tumor evolution and numerous biochemical mechanisms are required for accurate genome maintenance and duplication in normal cells. Chromosomal fragile sites are genomic loci that are prone to instability in the form of gaps and breaks under specific cellular conditions¹. Common fragile sites (CFS) exist in all individuals as normal chromosomal components, but undergo breakage at increased frequencies upon treatment of cells with specific chemicals, such as aphidicolin^{1, 2}. Of clinical significance, breakpoints of chromosomal translocations and deletions in several tumor types have been mapped to CFS^{1, 3, 4}. A large number of studies have shown that loss of specific replication, cell cycle checkpoint and DNA repair proteins leads to CFS instability^{1, 5, 6, 7}. These findings suggest that maintenance of CFS stability in normal cells requires DNA replication checkpoint and repair responses. Several non-mutually exclusive models have been proposed to explain the propensity for specific chromosomal breakage within CFS regions. These models include, but are not limited to, late timing of CFS replication⁸, paucity of replication initiation events within CFS⁹, failure to activate additional replication origins under stress¹⁰, altered epigenetic patterns¹¹, and collision of the transcription and replication machinery within long genes¹².

Ectopic integration of ~140-150 kb FRA3B sequences led to increased aphidicolin-induced gaps and breaks at the new chromosomal location¹³. In addition to studies showing altered DNA replication initiation within CFS, non-B DNA secondary structures, such as hairpins, cruciforms, and large AT-rich flexibility regions, have been hypothesized to hinder elongation by the replication machinery, thus contributing to slowed replication at CFS^{1, 7, 14, 15}. Clusters of AT-rich flexibility peaks may act as sinks for the superhelical density generated ahead of the replication fork, thereby hindering efficient topoisomerase activity and inhibiting replication fork progression¹⁴. Experimentally, an [AT/TA]₃₄ microsatellite derived from FRA16D and predicted to form a large hairpin structure was shown to induce replication fork stalling and chromosomal fragility in an *S. cerevisiae* model¹⁶. In human cells, replication blockage was also shown to occur preferentially near AT-rich regions within FRA16C, even in the absence of aphidicolin, and the rate of replication fork progression was found to be decreased within FRA16C, compared to the whole genome¹⁰. While these experiments demonstrate that AT-rich regions within CFS impede the replication fork, the degree to which DNA synthesis by the replicative polymerases is affected has not been extensively investigated. We previously showed that a mononucleotide [A]₂₈ repeat derived from a region within FRA16D contributed to the pausing of DNA synthesis *in vitro* by the replicative polymerases α and δ (Pol δ)¹⁷. In a recent, genome-wide statistical modeling study of genomic features distinguishing CFS from non-fragile regions, we found that mononucleotide A/T microsatellite coverage (associated with Alu repeats), low complexity AT coverage and DNA flexibility are significant predictors of CFS regions in alternative models¹⁸.

Specialized DNA polymerases may be required for synthesis of naturally occurring alternative DNA structures that replicate late in S-phase, such as those found within CFS. Depletion of DNA polymerase η (Pol η) leads to increased chromosomal breakage at FRA7H, a slight delay in S phase, and a decrease in the number of late S phase replication foci¹⁹. Depletion of either Pol η or polymerase κ (Pol κ) results in increased double stranded breaks upon ectopic integration of constructs containing non-B DNA elements²⁰. We recently showed that Pol κ displays significantly less synthesis termination within mononucleotide [T]₁₁²¹ and dinucleotide [GT]₁₀ microsatellites, compared to replicative polymerases²². These findings led us to hypothesize that Pol κ also may be important for maintaining efficient replication through CFS sequences.

The goals of this study were to understand 1) the mechanisms limiting replicative Pol δ DNA synthesis through specific sequence elements contained within CFS regions, and 2) the potential contribution of the specialized Pol κ to efficient CFS DNA synthesis. We demonstrate that Pol δ undergoes significant pausing at three types of repetitive sequence elements, AT-rich microsatellites (A/T_n and AT/TA_n) and quasi-palindromes (QP), in a length-dependent manner. Mechanistically, we show that Pol δ pause sites within CFS sequences represent sites of enzyme dissociation from the DNA template. Finally, we show that DNA synthesis by Pol κ is significantly more efficient through these types of repetitive elements found within CFS templates. Based on our *in vitro* findings, we propose that the specialized Pol κ might function in replication of repetitive DNA sequences in order to maintain CFS stability.

RESULTS

Polymerase Pausing Assay

We analyzed the efficiency of polymerase DNA synthesis using templates corresponding to CFS sequences, and an *in vitro* primer extension assay that models DNA synthesis on the lagging strand of the replication fork. Single-stranded DNA templates were constructed that contain ~120-170 base pairs derived from the reference sequences for FRA16D or FRA3B (Table 1 and Figure 1). These templates harbor repetitive sequence elements identified using the Non-B DNA Database (nonb.abcc.nci.fcr.gov/apps/Query-GFF/Features) or Repeat Masker (www.repeatmasker.org), along with surrounding reference genome sequence. We focused our attention on three types of sequence elements: 1) mononucleotide A/T repeats; 2) dinucleotide AT/TA repeats; and 3) quasi-palindromes. The AT1 template¹⁷ contains an $[A/T]_{28}$ mononucleotide repeat and an interrupted $[AT/TA]_{24}$ dinucleotide repeat. AT2, previously identified as a region of high DNA flexibility²³, contains an $[A/T]_9$ repeat and an $[AT/TA]_8$. Based on previous findings that interrupted AT/TA repeats contribute to only minor pausing by Pol δ ¹⁷, a template containing a pure $[AT/TA]_{25}$ repeat (as well as an $[A/T]_{22}$) was also generated (AT 3 template). We previously observed strong pausing by Pals δ and α at a quasi-palindrome sequence predicted to form a hairpin (QP 1¹⁷). Here, we generated additional templates containing quasi-palindrome sequences of varying lengths (37 bases, QP 2; 29 bases, QP 3) to test the generality of Pol δ pausing at quasi-palindromes. $[AT/TA]_n$. All QP repeats were predicted to form stable hairpin structures by Mfold analysis (<http://mfold.rna.albany.edu/?q=mfold/>). Characteristics of these potential hairpin structures are displayed in Table 2. The control template is composed of an AT-rich sequence present within FRA16D that does not contain any repeat elements or potential hairpin structures.

Pol δ Pausing at CFS AT-rich Sequence Elements

Pol δ synthesis efficiency through non-B DNA elements was assessed by primer extension reactions, performed at nuclear physiological pH²⁴, with single-stranded CFS templates and excess Pol δ in the presence of PCNA. Pausing was identified as sites of accumulated reaction products. CFS region-specific pausing was quantified, relative to the control template (devoid of repetitive sequences and non-B DNA structure potential, Fig. 2A). The extent of Pol δ pausing within the three AT-rich templates was sequence-dependent. Significant pausing was observed throughout the $[A]_{28}$ repeat of the AT1 template, as observed previously¹⁷ using a different reaction buffer (Fig. 2B). Normalized percent transit past this CFS region increased from 5% (2 minutes) to only ~29% (15 minutes), and was significantly decreased as compared to the control template at all time points ($p=0.0009$, <0.0001 , 0.0006 , respectively, Student t-test). Some pausing was also observed within template AT2 at the shorter $[A]_9$ repeat; however, the percent transit rapidly increased from 11% (2 minutes) to 68% (15 minutes), demonstrating that the degree of pausing at mononucleotide A repeats is repeat length-dependent. Pausing was not observed for Pol δ at

the [T]₂₂ repeat of the AT3 template, suggesting that Pol δ may be biased towards less efficient synthesis of mononucleotide A repeats. Pol δ pausing on reverse complement templates also exhibited this strand bias pattern. A pause site was observed at the [A]₂₂ repeat (AT3b template), while no pause site was detected at the [T]₉ repeat (AT2b template) (Supplemental Fig. 1A). Quantitatively, normalized percent transit was greater for the AT3b template ([A]₂₂ repeat) than the AT1 template ([A]₂₈ repeat) (Fig. 2B and Supplemental Fig. 1A), further demonstrating the length-dependence of Pol δ pausing at mononucleotide A repeats.

Pol δ pausing at dinucleotide AT/TA repeats was also dependent on tract length. Qualitatively, we observed no pausing at the [TA]₈ dinucleotide repeat within the AT2 template, or the interrupted [TA]_{24i} within the AT1 template (Fig. 1B). Based on previous findings that replication stalling is dependent on AT/TA repeat length¹⁶, Pol δ pausing was assessed on the AT3 template, which contains a pure [TA]₂₅ repeat, predicted by Mfold to form a stable hairpin structure. Interestingly, statistically significant pausing was observed at the base of the predicted hairpin for the 2 and 5 minute reaction time points ($p = 0.0009$, 0.0037). This corresponded to normalized percent transit of 7.9% and 24%, respectively, after which synthesis was able to continue past the predicted hairpin. Surprisingly, Pol δ did not exhibit pausing at the [AT]₂₅ repeat of the reverse complement AT3b template (Supplemental Fig. 1A). This could be explained by the effect of sequences flanking the AT/TA repeat (which differ due to base complementarity) on hairpin stability; however, we cannot rule out other possibilities at this time.

Pol δ Pausing at CFS Quasi-palindrome Sequence Elements

A strong Pol δ pause site is present at the base of a quasi-palindrome hairpin within the QP1 (QP₃₆) template (Figure 2C). Normalized percent transit changed from ~14% (2 and 5 minutes), to only 31% (15 minutes), a statistically significant decrease from the control template at each time point ($p = 0.0016$, 0.0001 , 0.0039 , respectively). We constructed DNA templates containing other quasi-palindrome sequences (QP₃₇ and QP₂₉), to test the generality of QP-induced pausing and the effect of altered stem homology (Table 2). As shown in Figure 2C, Pol δ displays strong pausing at the base of all predicted quasi-palindrome hairpins. Because Pol δ displays differences in pausing throughout each CFS region due to DNA sequence elements other than QPs (Figure 2), we calculated the termination probability at the base of each predicted QP to correlate the extent of pausing with QP homology. Termination probabilities were determined for the time points when significant pausing was initially observed at the quasi-palindrome. The quasi-palindrome within QP3 exhibited the greatest termination probability (0.72, 2 minute time point), compared to QP1 (0.51, 15 minutes) and QP2 (0.38, 15 minutes), consistent with a potentially greater stability of the QP3 predicted stem structure, as it possesses 100% self-homology (Table 2). Pol δ pausing was also assessed on reverse complement templates (Supplemental Fig. 1B). Pausing on templates QP2b and QP3b was similar to that on the complementary strand, with distinct pause sites located at the base of the quasi-palindrome hairpins. These data demonstrate that relatively short, quasi-palindromes located within CFS can impede synthesis by Pol δ and contribute to decreased synthesis rates.

Mechanism of Pol δ Pausing

Pol δ pausing at repetitive sequences decreases quantitatively with reaction time (Fig. 2), suggesting that Pol δ does not undergo an absolute inhibition of DNA synthesis. Two, non-mutually exclusive mechanisms could account for these observations: 1) CFS repetitive sequences contribute to a decreased rate of processive DNA synthesis, and 2) CFS repetitive DNA sequences cause dissociation of Pol δ from the DNA template, and reassociation events are required to complete synthesis past the site. To determine whether Pol δ

undergoes sequence-specific dissociation, reactions were carried out with Pol δ and CFS templates in the presence of excess primed, unlabeled single-stranded DNA trap. This DNA was effective as a trap for Pol δ , out-competing labeled template DNA, as no synthesis was detected in control reactions when trap was added before starting reactions (Fig. 3, + trap control lanes). In reactions using the AT1 template with no trap DNA added, Pol δ displayed a strong pause within the $[A]_{28}$ repeat at the 2 minute time point, but pausing disappeared as synthesis continued past the repeat by the 15 minute time point (Fig. 2B; Fig. 3A, - trap lanes). In contrast, upon addition of the DNA trap at the 2 minute time point, the pause site persisted over the course of the experiment (Fig. 3A, + trap experiment lanes), demonstrating that Pol δ dissociated from the DNA template at the pause site. To examine whether catalytic loading of PCNA might stabilize the Pol δ -DNA interaction and prevent dissociation at CFS repetitive sequences, DNA trap experiments were carried out in the presence of RFC and PCNA. Upon addition of RFC to reactions with PCNA and Pol δ using the AT1 template, an increased number of extended DNA products were visibly detectable above the primer region, relative to reactions containing only PCNA and Pol δ , or RFC and Pol δ (data not shown). Using reaction conditions established for optimal RFC activity, a broad distribution of pause sites was observed within the $[A]_{28}$ and $[AT]_{24i}$ repeats (Fig. 3C, - trap control lanes). However, upon addition of the DNA trap at 2 minutes, no products of Pol δ synthesis were observed past the major $[A]_{28}$ and $[AT]_{24i}$ pause sites. Therefore, loading of PCNA onto CFS templates does not prevent Pol δ dissociation at these non-B DNA sequence elements. As a further control, trap reactions were also carried out on the control CFS template that is devoid of repetitive DNA sequences. After addition of the DNA trap at the 2 minute time point, very long DNA products were observed to accumulate by the 5 and 15 minute time points (asterisk in Fig. 3B), demonstrating that some Pol δ molecules remained associated to the control template and continued synthesis past the pause site. The absence of such long products for the AT1 template (Fig. 3C) strongly suggests that specific DNA sequence elements within the CFS region present a strong block to processive synthesis by the DNA Pol δ holoenzyme.

Pol κ Pausing at CFS Sequence Elements

DNA Pol κ has been implicated specifically in DNA synthesis of repetitive DNA sequence elements^{20, 22}. Therefore, we investigated the efficiency of Pol κ synthesis through CFS repetitive elements and compared this to our Pol δ results. Optimal reaction conditions, including reaction buffer and Pol quantities, were established in order to yield approximately equal percent transit between Pools δ and κ on the control template. Normalized percent transit for Pol κ was higher than Pol δ for the AT-rich templates, differences that are statistically significant at the 5 and 15 minute time points for the AT1 template ($p = 0.0106$ and 0.0028) and 5 minute time point for the AT3 template ($p = 0.0103$; Fig. 4B). Pol κ pausing was observed at the mononucleotide $[A]_{28}$ repeat of the AT1 template and the dinucleotide $[TA]_{25}$ repeat of the AT3 template, but was more transient than Pol δ . Similar to Pol δ , no pausing was observed at mononucleotide T repeats and shorter dinucleotide AT/TA repeats (Fig. 4B and Supplemental Fig. 2A). Pol κ also displayed increased normalized percent transit relative to Pol δ on CFS templates containing quasi-palindrome sequences (Fig. 4B). This was most pronounced for the QP2 template, in which normalized percent transit by Pol κ was significantly increased relative to Pol δ at all time points ($p = 0.0311$, 0.0022 , <0.0001 , respectively). Although Pol κ paused at the base of predicted quasi-palindrome hairpins, this was much more transient than Pol δ pausing. Indeed, normalized percent transit by Pol κ reached $>80\%$ at the 15 minute time point for each template, while Pol δ normalized percent transit at 15 minutes was $\sim 50\%$ on the QP1 and QP2 templates. Similar results were observed for Pol κ on reverse complement templates (Supplemental Fig. 2B). These findings suggest that Pol κ synthesizes DNA more efficiently than does Pol δ at repetitive sequences with CFS.

DISCUSSION

Genome instability is a hallmark of tumorigenesis, and chromosomal rearrangements in tumor cells frequently occur at CFS. Determining the genomic features that contribute to CFS instability and the replication proteins that maintain stability in normal cells will expand our understanding of how genome instability arises during tumorigenesis. Here, we investigated the DNA sequence elements and DNA polymerases that affect the efficiency of DNA synthesis, in order to gain insight into whether the elongation stage of DNA replication is altered within CFS regions. We demonstrate that the replicative human Pol δ exhibits significant pausing at mononucleotide $[A]_n$ and dinucleotide $[TA]_n$ microsatellites, and quasi-palindrome sequences of sufficient length within model CFS templates. Importantly, we demonstrate that these pause sites correspond to sequences where Pol δ , in the presence of the RF-C/PCNA complex, dissociates from the DNA template. We also discovered that synthesis of CFS repetitive sequences by the specialized Pol κ is significantly more efficient than Pol δ . Our biochemical studies extend the potential involvement of Y-family Pols in CFS maintenance to include Pol κ , in addition to Pol η . In a recent genome-wide analysis of predictors of CFS fragility, we identified several genomic features distinguishing CFS and non-fragile regions¹⁷. Positive predictors of fragility included, but were not limited to, Twistflex value and *Alu* repeat coverage, which could be replaced with low-complexity AT-rich regions and mononucleotide A/T repeats, respectively, in alternative models. Combined with the findings presented here, we propose that enrichment of $[A/T]_n$, $[AT/TA]_n$, and QP repeat elements contribute to perturbed DNA replication fork progression through CFS regions by inhibiting or slowing synthesis by the lagging strand DNA polymerase, Pol δ . Importantly, previous evidence supports a role for the paucity of replication initiation events within CFS, the failure to activate additional replication origins under stress, and altered epigenetic patterns in fragility^{5; 6; 7}. Therefore, multiple, non-mutually exclusive mechanisms are likely involved in the propensity for instability at CFS.

Cis-acting CFS Sequence Elements that Contribute to Replicative Polymerase Pausing

Many lines of evidence support the hypothesis that AT-rich regions within CFS contribute to instability^{1; 7; 25}. Early characterization of several CFS identified regions of increased DNA torsional flexibility at regions of high AT content²⁵, which were hypothesized to generate torsional restraint to the topoisomerase complex during replication¹¹. An $[AT]_{34}$ repeat within the FRA16D Flex1 sequence, predicted to form a stable cruciform structure, was previously shown to cause replication inhibition *in vivo*¹⁶. Additionally, sites of replication blockages were located in relative proximity to AT-rich regions of FRA16C¹⁰. Our analyses of DNA synthesis progression through AT-rich CFS sequence elements extends these studies by demonstrating that these sequences also may contribute to replication perturbation by directly impeding the lagging strand DNA polymerase. AT-rich flexibility peaks harbor $[AT/TA]_n$ repeats, which can form hairpin and cruciform DNA secondary structures. Here, we demonstrate that an AT/TA repeat of sufficient length (25 units) impedes Pol δ synthesis, whereas shorter AT/TA repeats do not contribute to significant pausing (Figure 2b). A length dependence of hairpin formation, replication inhibition and chromosomal instability at TA/AT tracts has been previously demonstrated *in vivo*^{16; 26}. Interestingly, AT/TA dinucleotide repeats at flexibility peaks appear to be short (~2-10 units in length) and exist as clusters of interrupted, rather than pure, tracts^{14; 25}. Our results show that only longer repeats (> 25 units) are capable of directly impeding the lagging strand Pol δ . Possibly, longer stretches of interrupted AT/TA repeats, such as those found within flexibility peaks, are also capable of impeding Pol δ .

We demonstrate here that mononucleotide $[A]_n$ tracts specifically contribute to strong pausing by Pol δ , and that the intensity of pausing is directly correlated with repeat length

(Fig. 2). We previously proposed that such Pol δ pausing was due to the formation of bent DNA^{17; 21}. This alteration of the B-form DNA helix could disrupt the polymerase-DNA interaction during synthesis and perturb elongation. It is also possible that slipped strands within the repeat tract could distort the DNA helix and impede the lagging strand polymerase²⁷. Our observation of Pol δ pausing only within template dA tracts, and not template dT tracts, might reflect the greater stability of slipped strands within the template purine strand. Alternatively, poly(dA/dT) tracts within CFS sequences might favor formation of triplex DNA. Homopurine-homopyrimidine sequences have been shown to form triplex structures and contribute to polymerase pausing during *in vitro* DNA synthesis^{28; 29} and replication stalling *in vivo*³⁰. Importantly, the effect of mononucleotide A/T repeats on replication stalling within CFS has not yet been examined *in vivo*. Although replication blockage was shown to occur in relative proximity to AT-rich sequences within FRA16C¹⁰, the precise location of replication stalling, which could potentially occur within other repetitive sequence elements, was not examined at the nucleotide level. Future directions will focus on understanding the contribution of mononucleotide A/T repeats to CFS replication perturbation in cells.

We also identified QPs within CFS sequences as elements that significantly contribute to significant Pol δ pausing. Quasi-palindromes are hot spots of double-strand breaks, template switching, and rearrangements that contribute to genomic instability^{31; 32; 33}. Fragile sites in yeast that undergo chromosomal aberrations similar to human fragile sites have been identified at sites of large (several kb in length) palindromes³⁴. Palindromes formed by *Alu* elements of hundreds of base pairs in length within the human genome have been shown to cause replication stalling *in vivo*³⁵. Here, we demonstrate that much shorter, quasi-palindrome repeats (from 29-37 nucleotides in length) can directly impede the lagging strand polymerase *in vitro*. Based on predictions by the Non-B DNA database and Mfold, we propose that the formation of hairpins possessing ~75-100% self homology within stem structures, and intervening loop sequences of 3-9 bases, might be responsible for Pol δ pausing. Future computational modeling is required to determine whether CFS are enriched for such quasi-palindrome repeat elements, as compared to non-fragile regions.

Mechanism of Pol δ Pausing within CFS

What mechanism accounts for Pol δ pausing at CFS repetitive sequence elements? Our findings demonstrate that Pol δ pause sites in CFS sequences represent sites of enzyme dissociation, rather than an absolute inhibition of DNA synthesis. Pol δ dissociation at mononucleotide [A]_n and dinucleotide [TA]_n sequence elements was observed in the presence of RFC-loaded PCNA (Fig. 3C), suggesting that the processivity factor does not prevent dissociation by stabilizing Pol δ -DNA interactions. Differences in the efficiency of Pol δ re-association with the primer-template might also account for the decreased transit we observed within CFS templates containing repetitive sequence elements. In order for productive Pol δ binding and extension to occur after dissociation, the primer terminus must either realign downstream of the inhibitory element, or potential DNA structures within the repetitive element must be resolved. In contrast, pause sites that occur independently of repetitive sequences (such as within the control template) might occur as a consequence of limited Pol δ processivity, and Pol δ likely could reassociate efficiently with the primer terminus following dissociation. Importantly, dissociation of the replication machinery and the inability to efficiently recover the replication fork at CFS repetitive sequences might contribute to fork collapse and/or formation of double strand breaks *in vivo* (as reviewed in³⁶).

Pol κ Pausing at CFS Repetitive Sequence Elements

Recently, the specialized, Y-family Pol η was implicated in maintenance of CFS stability¹⁹. Combined with our previous findings that the Y-family Pol κ can synthesize specific microsatellite elements more efficiently than replicative Pols^{21; 22}, we were interested in examining the potential role of Pol κ in CFS stability. We demonstrate in this study that Pol κ is significantly more efficient at *in vitro* synthesis of CFS repetitive sequence elements relative to Pol δ . However, some pausing was observed by Pol κ , particularly within AT-rich mononucleotide and dinucleotide repeats (Fig. 5B and Supplemental Fig. 4). Betous et al. (2009) proposed that Pol κ might function directly in replication of non-B DNA structures, through a mechanism similar to translesion synthesis; alternatively, it was suggested that the polymerase could function in repair or recombination events downstream of double-stranded breaks initiated at these sites²⁰. Our findings provide further support for the involvement of Pol κ in replication of repetitive and potential non-B DNA elements. We have previously demonstrated that Pol κ can freely exchange with Pol δ *in vitro* to relieve pausing at a [GT]₁₀ microsatellite²². Perhaps, Pol κ could undergo switching with the replicative polymerase upon Pol δ pausing at repetitive sequences, similar to translesion synthesis³⁷. Alternatively, Pol κ could function in a replication complex with Pol δ to continue synthesis past the non-B DNA sequence, as Pols κ and δ have been shown to function within the same repair complex during nucleotide excision repair³⁸. Relief of Pol δ pausing may also require recruitment of the WRN protein, as WRN is required to maintain CFS stability³⁹, and has been shown to functionally interact with Pols κ and δ ^{40; 41}. Additionally, we demonstrated that WRN stimulates more efficient Pol δ synthesis through CFS repetitive sequences¹⁷. Future *ex vivo* experiments are required to define the role of Pol κ in CFS maintenance. Nevertheless, the expression of the specialized Pols κ and η can be decreased in colorectal, stomach, and lung tumors^{20; 42; 43}. CFS are known hotspots of chromosomal breakage that occur during tumorigenesis⁴, and deletions of up to hundreds of kilobases at FRA3B have been identified in lung, digestive, kidney and breast tumors^{44; 45; 46}. Large deletions and rearrangements frequently occur at FRA16D in breast and prostate cancers, and multiple myeloma^{47; 48; 49}. Understanding the genetic consequences of lowered specialized DNA polymerase expression is crucial for elucidating mechanisms of instability at CFS regions during tumorigenesis.

MATERIALS & METHODS

Replication proteins

Recombinant human four subunit Pol δ and PCNA were purified using baculovirus expression systems, as described previously⁵⁰. Human Pol κ and additional PCNA were purchased from Enzymax (Lexington, KY). Yeast RFC was a generous gift from Dr. Gregory D. Bowman.

Vector Constructs

Oligonucleotides corresponding to FRA16D or FRA3B regions (derived from GenBank accessions AF217490 and 183583557, respectively) were purchased from IDT (Coralville, IA). Double stranded DNA for cloning was generated for each oligonucleotide using T7 polymerase (USB, Cleveland, Ohio) and primers complementary to underlined sequences below (IDT). Fragile site sequences were cloned into the pGem3zf(-) vector as previously described¹⁷ and confirmed by DNA sequence analysis. Due to the sequence complexity of the templates, some minor sequence alterations occurred during construction; clones with sequences closest to original oligonucleotides were chosen. The final sequences are provided in supplementary information (methods).

***In vitro* polymerase pausing assay**

Constructs in both ligation orientations were used to generate single stranded DNA (ssDNA) templates for *in vitro* analysis, as previously described¹⁷. Primer extension reactions were carried out with CFS-containing templates and replication proteins of interest, as described¹⁷, with the following modification. Titration experiments with each polymerase were carried out on the control template to determine the enzyme:DNA ratio required for complete synthesis, defined as 75-80% product extension completely past the CFS region by the 15 minute time point. Preliminary experiments were carried out to develop a suitable reaction buffer for both Pols δ and κ , that reflected the nuclear cellular pH range, as previously determined²⁴. A KPO₄ buffer was used because Pol κ activity was significantly reduced in Tris buffer. Standard reactions contained 25 mM KPO₄ pH 7.6, 5 mM MgCl₂, 2.5 mM DTT, 0.2 mg/ml non-acetylated BSA, 250 μ M dNTPs, 100 fmol ssDNA-primer hybrid template and 0.5-1.0 pmol Pol δ or 0.5 pmol Pol κ , unless indicated otherwise. DNA template, buffer and replication accessory proteins (when included) were pre-incubated at 37°C for 3 minutes and reactions initiated by addition of polymerase. Reactions were terminated by addition of an equal volume of stop dye at indicated time points, reaction products were separated by polyacrylamide electrophoresis, and quantified as previously described¹⁷.

DNA Trap Experiments

Primer extension reactions for trap experiments in the absence of RFC were carried out exactly as described for the polymerase pausing assay, with the addition of a DNA trap. The DNA trap consisted of unlabeled, primed QP1 ssDNA template. Reactions were initiated by addition of Pol δ , and trap DNA (8-11 pmol) was added at the 2 minute time point. Reactions were stopped at 5, 15 and 30 minute time points and separated by denaturing gel electrophoresis. Control reactions in the absence of trap were carried out simultaneously to monitor Pol δ pausing at 2, 5, 15, and 30 minute time points. To monitor trap effectiveness, an additional control reaction was carried out in which unlabelled trap and template DNA were combined prior to initiation of the reaction with Pol δ . Reactions to test the effect of RFC-loaded PCNA contained 100 fmol ³²P end-labeled primed, single stranded CFS templates, 300 fmol human four-subunit Pol δ , 400 fmol human PCNA, 50 fmol yeast RFC, 5 mM ATP, 50 mM Tris-HCl at pH 7.0, 50 mM MgCl₂, 2 mM DTT, 0.2 mg/ml BSA, 2% glycerol, 50 mM NaCl, and 250 μ M dNTPs.

Quantification and Statistical Analysis

Reaction products were grouped into four regions: unextended primer, 5' to the CFS sequence, CFS sequence region, and 3' to the CFS sequence. The number of DNA molecules in each region was determined using the ImageQuant 5.2 software and corrected for background pixels and loading by normalizing to the no polymerase control. Percent transit for each reaction is defined as [number of molecules in the 3' region] $\div$ [number of molecules in the CFS region + 3' region combined] $\times$ 100. For unbiased comparisons between polymerases, normalized percent transit was calculated as [percent transit on the CFS template] $\div$ [percent transit on the control template], for each time point. Statistical significance (defined as $p < 0.05$) was determined by t-test analysis of three independent reactions. Termination probabilities for quasi-palindromes was calculated as [number of molecules within the quasi-palindrome region] $\div$ [number of molecules within the quasi-palindrome region + past the quasi-palindrome region].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Suzanne Hile for her technical assistance in the initial stages of this project, and Dr. Erin Gestl for his critical reading of this manuscript. We also thank Arakachai Fungtammasan (Penn State University) for his critical discussions and genome-wide computational analyses of specific DNA sequence element enrichment within CFS regions.

FUNDING

This work was supported by the National Institutes of Health [grant number CA100060 to K.A.E.] and [grant number GM313973 to M.Y.L.]; and the Gittlen Cancer Research Foundation (Penn State Hershey Medical Center). This project was also funded, in part, under a grant from the Pennsylvania Department of Health using Tobacco CURE Funds [SAP# 4100042746 to K.A.E.]. The Department of Health specifically disclaims responsibility for any analyses, interpretations, or conclusions.

Abbreviations used

CFS	common fragile site(s)
Pol	DNA polymerase
QP	quasi-palindrome
ssDNA	single-stranded DNA
Ub-PCNA	monoubiquitinated PCNA

GLOSSARY

Common fragile site	Genomic loci that undergo a high frequency of chromosomal breakage under specific cellular conditions of replication stress
Microsatellite	DNA sequences of 1-6 base pairs repeated in tandem
Quasi-palindromes	Repetitive DNA sequence elements composed of an imperfect palindrome (inverted repeat sequence) separated symmetrically by an intervening sequence. DNA hairpins predicted to form within quasi-palindromes: stem structures are formed by intramolecular base pairing within the palindromic sequence, and loop structures are formed by the intervening sequence
Replicative DNA polymerases	DNA polymerases of the B family, including polymerases δ , α -primase, and ϵ in eukaryotes, which are responsible for replication of the bulk genome
Lagging strand DNA polymerase	The DNA polymerase (polymerase δ in eukaryotes) shown to be responsible for replication on the lagging strand of the replication fork
Y-family Specialized DNA polymerases	DNA polymerases of the Y family, including κ and η in humans, that function in DNA metabolic processes including repair, translesion synthesis, recombination, and resolution of stalled replication forks
Polymerase pausing	Accumulation of <i>in vitro</i> DNA synthesis products at a specific sequence, representing sites at which the DNA polymerase exhibited either a slowed synthesis rate or an elevated frequency of dissociation from the DNA template

PCNA	The polymerase sliding clamp replication protein in eukaryotes that tethers polymerases to DNA and allows for processive DNA synthesis
RFC	The eukaryotic clamp-loading protein, responsible for loading the PCNA clamp protein onto DNA

REFERENCES

1. Durkin SG, Glover TW. Chromosome Fragile Sites. *Annual Review of Genetics*. 2007; 41:169–192.
2. Dillon LW, Burrow AA, Wang YH. DNA instability at chromosomal fragile sites in cancer. *Curr Genomics*. 2010; 11:326–37. [PubMed: 21286310]
3. Cox C, Bignell G, Greenman C, Stabenau A, Warren W, Stephens P, Davies H, Watt S, Teague J, Edkins S, Birney E, Easton DF, Wooster R, Futreal PA, Stratton MR. A survey of homozygous deletions in human cancer genomes. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:4542–4547. [PubMed: 15761058]
4. Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM, Buck G, Chen L, Beare D, Latimer C, Widaa S, Hinton J, Fahey C, Fu B, Swamy S, Dalgliesh GL, Teh BT, Deloukas P, Yang F, Campbell PJ, Futreal PA, Stratton MR. Signatures of mutation and selection in the cancer genome. *Nature*. 2010; 463:893–898. [PubMed: 20164919]
5. Franchitto A, Pichierri P. Understanding the molecular basis of common fragile sites instability: Role of the proteins involved in the recovery of stalled replication forks. *Cell Cycle*. 2011; 10:4039–4046. [PubMed: 22101264]
6. Debatisse M, Le Tallec B. t. Letessier A, Dutrillaux B, Brison O. Common fragile sites: mechanisms of instability revisited. *Trends in Genetics*. 2012; 28:22–32. [PubMed: 22094264]
7. Ozeri-Galai E, Bester AC, Kerem B. The complex basis underlying common fragile site instability in cancer. *Trends in Genetics*. 2012; 28:295–302. [PubMed: 22465609]
8. Le Beau MM, Rassool FV, Neilly ME, Espinosa R, Glover TW, Smith DI, McKeithan TW. Replication of a Common Fragile Site, FRA3B, Occurs Late in S Phase and is Delayed Further Upon Induction: Implications for the Mechanism of Fragile Site Induction. *Human Molecular Genetics*. 1998; 7:755–761. [PubMed: 9499431]
9. Letessier A, Millot GA, Koundrioukoff S, Lachages A-M, Vogt N, Hansen RS, Malfoy B, Brison O, Debatisse M. Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*. 2011; 470:120–123. [PubMed: 21258320]
10. Ozeri-Galai E, Lebofsky R, Rahat A, Bester, Assaf C, Bensimon A, Kerem B. Failure of Origin Activation in Response to Fork Stalling Leads to Chromosomal Instability at Fragile Sites. *Molecular Cell*. 2011; 43:122–131. [PubMed: 21726815]
11. Jiang Y, Lucas I, Young DJ, Davis EM, Karrison T, Rest JS, Le Beau MM. Common fragile sites are characterized by histone hypoacetylation. *Human Molecular Genetics*. 2009; 18:4501–4512. [PubMed: 19717471]
12. Helmrich A, Ballarino M, Tora L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Molecular Cell*. 2011; 44:966–977. [PubMed: 22195969]
13. Ragland RL, Glynn MW, Arlt MF, Glover TW. Stably transfected common fragile site sequences exhibit instability at ectopic sites. *Genes, Chromosomes and Cancer*. 2008; 47:860–872. [PubMed: 18615677]
14. Zlotorynski E, Rahat A, Skaug J, Ben-Porat N, Ozeri E, Hersberg R, Levi A, Scherer SW, Margalit H, Kerem B. Molecular Basis for Expression of Common and Rare Fragile Sites. *Mol. Cell. Biol*. 2003; 23:7143–7151. [PubMed: 14517285]
15. Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfrond Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui L-C, Rosenthal A, Kerem B. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the

- cloning of a simian virus 40 integration site. *Proceedings of the National Academy of Sciences*. 1998; 95:8141–8146.
16. Zhang H, Freudenreich CH. An AT-Rich Sequence in Human Common Fragile Site FRA16D Causes Fork Stalling and Chromosome Breakage in *S. cerevisiae*. *Molecular Cell*. 2007; 27:367–379. [PubMed: 17679088]
 17. Shah SN, Opresko PL, Meng X, Lee MYWT, Eckert KA. DNA structure and the Werner protein modulate human DNA polymerase delta-dependent replication dynamics within the common fragile site FRA16D. *Nucleic Acids Research*. 2010; 38:1149–1162. [PubMed: 19969545]
 18. Fungtammasan A, Walsh E, Chiaromonte F, Eckert KA, Makova KD. A genome-wide analysis of common fragile sites: What features determine chromosomal instability in the human genome? *Genome Research*. 2012; 22:993–1005. [PubMed: 22456607]
 19. Rey L, Sidorova JM, Puget N, Boudsocq F, Biard DSF, Monnat RJ Jr, Cazaux C, Hoffmann J-S. Human DNA Polymerase {eta} Is Required for Common Fragile Site Stability during Unperturbed DNA Replication. *Mol. Cell. Biol*. 2009; 29:3344–3354. [PubMed: 19380493]
 20. Bétous R, Rey L, Wang G, Pillaire M-J, Puget N, Selves J, Biard DSF, Shin-ya K, Vasquez KM, Cazaux C, Hoffmann J-S. Role of TLS DNA polymerases eta and kappa in processing naturally occurring structured DNA in human cells. *Molecular Carcinogenesis*. 2009; 48:369–378. [PubMed: 19117014]
 21. Hile SE, Eckert KA. DNA polymerase kappa produces interrupted mutations and displays polar pausing within mononucleotide microsatellite sequences. *Nucleic Acids Research*. 2008; 36:688–696. [PubMed: 18079151]
 22. Hile SE, Wang X, Lee MYWT, Eckert KA. Beyond translesion synthesis: polymerase ϵ fidelity as a potential determinant of microsatellite stability. *Nucleic Acids Research*. 2012; 40:1636–1647. [PubMed: 22021378]
 23. Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S, Nancarrow JK, Woollatt E, Kremmidiotis G, Gardner A, Venter D, Baker E, Richards RI. Common chromosomal fragile site FRA16D sequence: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells. *Human Molecular Genetics*. 2000; 9:1651–1663. [PubMed: 10861292]
 24. Seksek O, Bolard J. Nuclear pH gradient in mammalian cells revealed by laser microspectrofluorimetry. *J Cell Sci*. 1996; 109:257–262. [PubMed: 8834810]
 25. Lukusa T, Fryns JP. Human chromosome fragility. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2008; 1779:3–16.
 26. Dayn A, Malkhosyan S, Duzhy D, Lyamichev V, Panchenko Y, Mirkin S. Formation of (dA-dT)_n cruciforms in *Escherichia coli* cells under different environmental conditions. *Journal of Bacteriology*. 1991; 173:2658–2664. [PubMed: 1849512]
 27. Garcia-Diaz M, Kunkel TA. Mechanism of a genetic glissando*: structural biology of indel mutations. *Trends in Biochemical Sciences*. 2006; 31:206–214. [PubMed: 16545956]
 28. Hile SE, Eckert KA. Positive Correlation Between DNA Polymerase [alpha]-Primase Pausing and Mutagenesis within Polypyrimidine/Polypurine Microsatellite Sequences. *Journal of Molecular Biology*. 2004; 335:745–759. [PubMed: 14687571]
 29. Baran N, Lapidot A, Manor H. Formation of DNA triplexes accounts for arrests of DNA synthesis at d(TC)_n and d(GA)_n tracts. *Proceedings of the National Academy of Sciences*. 1991; 88:507–511.
 30. Mirkin EV, Mirkin SM. Replication Fork Stalling at Natural Impediments. *Microbiology and Molecular Biology Reviews*. 2007; 71:13–35. [PubMed: 17347517]
 31. Dutra BE, Lovett ST. Cis and Trans-acting Effects on a Mutational Hotspot Involving a Replication Template Switch. *Journal of Molecular Biology*. 2006; 356:300–311. [PubMed: 16376936]
 32. Bissler JJ. DNA inverted repeats and human disease. *Frontiers in bioscience : a journal and virtual library*. 1998; 3:d408–18. [PubMed: 9516381]
 33. Seier T, Padgett DR, Zilberberg G, Sutera VA, Toha N, Lovett ST. Insights Into Mutagenesis Using *Escherichia coli* Chromosomal lacZ Strains That Enable Detection of a Wide Spectrum of Mutational Events. *Genetics*. 2011; 188:247–262. [PubMed: 21441210]

34. Casper AM, Greenwell PW, Tang W, Petes TD. Chromosome Aberrations Resulting From Double-Strand DNA Breaks at a Naturally Occurring Yeast Fragile Site Composed of Inverted Ty Elements Are Independent of Mre11p and Sae2p. *Genetics*. 2009; 183:423–439. [PubMed: 19635935]
35. Voineagu I, Narayanan V, Lobachev KS, Mirkin SM. Replication stalling at unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins. *Proceedings of the National Academy of Sciences*. 2008; 105:9936–9941.
36. Paulsen RD, Cimprich KA. The ATR pathway: Fine-tuning the fork. *DNA Repair*. 2007; 6:953–966. [PubMed: 17531546]
37. Chen J, Bozza W, Zhuang Z. Ubiquitination of PCNA and Its Essential Role in Eukaryotic Translesion Synthesis. *Cell Biochemistry and Biophysics*. 2011:1–14. [PubMed: 20740326]
38. Ogi T, Limsirichaikul S, Overmeer RM, Volker M, Takenaka K, Cloney R, Nakazawa Y, Niimi A, Miki Y, Jaspers NG, Mullenders LHF, Yamashita S, Fousteri MI, Lehmann AR. Three DNA Polymerases, Recruited by Different Mechanisms, Carry Out NER Repair Synthesis in Human Cells. *Molecular Cell*. 2010; 37:714–727. [PubMed: 20227374]
39. Pirzio LM, Pichierri P, Bignami M, Franchitto A. Werner syndrome helicase activity is essential in maintaining fragile site stability. *The Journal of Cell Biology*. 2008; 180:305–314. [PubMed: 18209099]
40. Kamath-Loeb AS, Lan L, Nakajima S, Yasui A, Loeb LA. Werner syndrome protein interacts functionally with translesion DNA polymerases. *Proceedings of the National Academy of Sciences*. 2007; 104:10394–10399.
41. Kamath-Loeb AS, Shen J-C, Schmitt MW, Loeb LA. The Werner Syndrome Exonuclease Facilitates DNA Degradation and High Fidelity DNA Polymerization by Human DNA Polymerase δ . *Journal of Biological Chemistry*. 2012; 287:12480–12490. [PubMed: 22351772]
42. Pan Q, Fang Y, Xu Y, Zhang K, Hu X. Down-regulation of DNA polymerases [kappa], [eta], [iota], and [zeta] in human lung, stomach, and colorectal cancers. *Cancer Letters*. 2005; 217:139–147. [PubMed: 15617831]
43. Pillaire MJ, Selves J, Gordien K, Gouraud PA, Gentil C, Danjoux M, Do C, Negre V, Bieth A, Guimbaud R, Trouche D, Pasero P, Mechali M, Hoffmann JS, Cazaux C. A “DNA replication” signature of progression and negative outcome in colorectal cancer. *Oncogene*. 2009; 29:876–887. [PubMed: 19901968]
44. Druck T, Hadaczek P, Fu T-B, Ohta M, Siprashvili Z, Baffa R, Negrini M, Kastury K, Veronese ML, Rosen D, Rothstein J, McCue P, Cotticelli MG, Inoue H, Croce CM, Huebner K. Structure and Expression of the Human FHIT Gene in Normal and Tumor Cells. *Cancer Research*. 1997; 57:504–512. [PubMed: 9012482]
45. Michael D, Beer DG, Wilke CW, Miller DE, Glover TW. Frequent deletions of FHIT and FRA3B in Barrett's metaplasia and esophageal adenocarcinomas. *Oncogene*. 1997; 15:1653–9. [PubMed: 9349498]
46. Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, Huebner K, Croce CM. Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. *Proceedings of the National Academy of Sciences*. 1999; 96:7456–7461.
47. Paige AJW, Taylor KJ, Taylor C, Hillier SG, Farrington S, Scott D, Porteous DJ, Smyth JF, Gabra H, Watson JEV. WWOX: A candidate tumor suppressor gene involved in multiple tumor types. *Proceedings of the National Academy of Sciences*. 2001; 98:11417–11422.
48. Mangelsdorf M, Ried K, Woollatt E, Dayan S, Eyre H, Finnis M, Hobson L, Nancarrow J, Venter D, Baker E, Richards RI. Chromosomal Fragile Site FRA16D and DNA Instability in Cancer. *Cancer Research*. 2000; 60:1683–1689. [PubMed: 10749140]
49. Paige AJW, Taylor KJ, Stewart A, Sgouros JG, Gabra H, Sellar GC, Smyth JF, Porteous DJ, Watson JEV. A 700-kb Physical Map of a Region of 16q23.2 Homozygously Deleted in Multiple Cancers and Spanning the Common Fragile Site FRA16D. *Cancer Research*. 2000; 60:1690–1697. [PubMed: 10749141]
50. Xie B, Mazloun N, Liu L, Rahmeh A, Li H, Lee MYWT. Reconstitution and Characterization of the Human DNA Polymerase Delta Four-Subunit Holoenzyme. *Biochemistry*. 2002; 41:13133–13142. [PubMed: 12403614]

51. Becker NA, Thorland EC, Denison SR, Phillips LA, Smith DI. Evidence that instability within the FRA3B region extends four megabases. *Oncogene*. 2002; 57:8713–8722. [PubMed: 12483524]
52. Krummel KA, Roberts LR, Kawakami M, Glover TW, Smith DI. The Characterization of the Common Fragile Site FRA16D and Its Involvement in Multiple Myeloma Translocations. *Genomics*. 2000; 69:37–46. [PubMed: 11013073]

\$watermark-text

\$watermark-text

\$watermark-text

Highlights

- Mechanisms by which CFS repeat sequences affect DNA polymerases are unknown.
- Polymerase delta pauses at microsatellite and palindromic sequences within CFS.
- Polymerase delta pausing is due to sequence-specific polymerase dissociation.
- Polymerase kappa displays more efficient synthesis of CFS repeat sequences.
- Polymerase kappa may be required for replication of CFS repeat sequences in cells.

Figure 1. Distribution of Template Sequences within Common Fragile Site Regions

Repetitive sequence elements were identified within FRA16D or FRA3B using Repeatmasker or the non-B DNA database. Repetitive sequences and surrounding sequence were cloned into pGem vectors to create ssDNA templates. (A.) Scheme of FRA3B (gray box) with approximate location of the template investigated here (short black line), based on previously defined aphidicolon-induced break points⁵¹. Approximate location of the FHIT gene and the FRA3B GenBank accession used to generate the templates are represented by black lines above the gray line. (B.) Schematic of FRA16D end points (gray box) as previously identified^{48; 52} showing the approximate distribution of templates investigated in this study (short black lines). The ~270 Kb GenBank accession used to generate templates, and approximate location of WWOX gene are represented by black lines.

Figure 2. Pol δ Pausing at CFS Repetitive Sequence Elements

Primer extension reactions were carried out using ^{32}P end-labeled primed single-stranded CFS templates. Reactions contained 100 fmol primer/template DNA, 500 fmol-1 pmol Pol δ and 500 fmol PCNA in standard reaction conditions. Products were separated by 8% denaturing polyacrylamide gel electrophoresis and quantified by calculating percent transit as: [amount of product in the 3' region of the template] ÷ [product in the CFS region + product in the 3' region]. (A.) Top, representative gel showing Pol δ synthesis on the control CFS template. Black triangle represents increasing time from 2, 5, and 15 minutes. -, No polymerase control. +, Hybridization control. TACG, sequencing ladder. Bottom, quantification of percent transit for three independent reactions on the control template. (B.) Top, representative gels showing synthesis of Pol δ on AT-repeat containing CFS templates. Repetitive sequences within CFS template are outlined in brackets. Bottom, percent transit, mean normalized to percent transit on the control template, for three independent Pol δ reactions on each template. White, gray and black bars represent 2, 5 and 15 minute reaction time points, respectively. Asterisks indicate statistically significant decreased % transit of Pol δ through the template of interest, relative to the control template (Student t-test). (C.) Top, representative gels for quasi-palindrome containing templates. Bottom, mean normalized percent transit for three independent reactions on each template.

Figure 3. DNA Trap Experiment to Determine the Mechanism of Pol δ Pausing

(A.) Dissociation of Pol δ within the AT1 template. Primer extension reactions were carried out using 100 fmol ^{32}P end-labeled primer/template DNA, 1 pmol Pol δ , 500 fmol PCNA, with or without 8-11 pmol unlabeled, primed ssDNA trap. *Left to right*: – Trap control reaction (black triangle represents 2, 5, 15 and 30 minute time points); + Trap control (cont.), in which AT1 template and trap DNA were incubated prior to reaction initiation with Pol δ (black triangle represents 5, 15, and 30 minute time points); + Trap experimental (exp.), in which reactions containing Pol δ and AT1 template DNA were carried out for 2 minutes, after which time DNA trap was added (black triangle represents 5, 15, and 30 minute time points). (B.) DNA trap experiment to determine dissociation of Pol δ within the control template. Reactions were carried out as in (A). Asterisk indicates accumulation of longer reaction products at the 5 and 15 minute time points. (C.) DNA trap experiment in the presence of RFC-loaded PCNA. Primer extension reactions were carried out using 100 fmol ^{32}P end-labeled primer/template DNA, 300 fmol Pol δ , 400 fmol PCNA, 50 fmol RFC, 5 mM ATP, 50 mM Tris-HCl at pH 7.0, 50 mM MgCl_2 , 2 mM DTT, 0.2 mg/ml BSA, 2% glycerol, 50 mM NaCl, and 250 μM dNTPs, with or without 8-11 pmol DNA trap. Trap experiments were carried out as in (A.). P, unextended primer.

Figure 4. Pol κ Pausing at CFS Repetitive Sequence Elements

Primer extension reactions were performed with 100 fmol ^{32}P -end labeled primed-single stranded DNA hybrid and 500 fmol Pol κ under standard reaction conditions. Reaction products were analyzed by gel electrophoresis and quantified as in Figure 2. (A.) Top, representative gel for a Pol κ reaction on the control template. Bottom, quantification of percent transit for five independent reactions. Black triangle, 2, 5 and 15 minute time points. -, No polymerase control. +, hybridization control. TACG, sequencing ladder. (B.) Top, representative gels for Pol κ reactions on AT repeat-containing templates. Bottom, quantification of percent transit, normalized to percent transit on the control template, for three independent Pol κ reactions. White, gray and black bars represent 2, 5 and 15 minute time points, respectively. Asterisks indicate statistically significant decreased % transit of Pol κ through the template of interest, compared to the control template (Student t-test). Hatch marks (#) represent significantly greater normalized % transit of Pol κ compared to Pol δ . (C.) Top, representative gels for Pol κ reactions on quasi-palindrome containing templates. Bottom, quantification of mean normalized percent transit for three independent reactions.

Table 1

Characteristics of CFS Templates.

Region	CFS ^a	% AT	Repeat Elements ^b
Control	16D (199256-199375)	76	None
AT 1 ^c	16D (191565-191712)	74	[A/T] ₂₈ , [AT/TA] _{24i}
AT 2	16D (91285-91458)	72	[A/T] ₉ , [AT/TA] ₈
AT 3	3B (837821-837929; 837963-838030)	80	[A/T] ₂₂ , [AT/TA] ₂₅
QP 1 ^c	16D (191713-191860)	53	[A/T] ₁₉ , QP ₃₆
QP 2	16D (51616-51796)	51	QP ₃₇
QP 3	16D (176735-176887)	64	QP ₂₉

^aFRA16D sequences derived from NCBI GenBank Accession AF217490; FRA3B sequence derived from NCBI GenBank Accession 183583557; parentheses indicate location of sequences within GenBank Accession.

^bIdentified with RepeatMasker or non-B DNA database (<http://nonb.abcc.ncifcrf.gov/apps/Query-GFF/Features>). QP, quasi-palindrome. i, interrupted/impure repeat.

^cTemplates previously investigated by¹⁷.

Table 2

Characteristics of Mfold-predicted hairpin structures within CFS templates.

Template	Repetitive Sequence Predicted to form Hairpin (length, base pairs)	% Hairpin Stem Structure Self-Homology ^a	Length of Hairpin Stem (base pairs) ^b	Length of Hairpin Loop (bases) ^c
AT 1	AT/TA _(24i)	100%	18	4
AT 2	AT/TA ₍₈₎	95%	11 ^d	4
AT 3	AT/TA ₍₂₅₎	100%	24 ^d	4
QP 1	QP ₍₃₆₎	76%	29	7
QP 2	QP ₍₃₇₎	82%	34	3
QP 3	QP ₍₂₉₎	100%	20	9

^aPercentage of stem predicted to be involved in base pairing.

^bLength of predicted stem structure, including unmatched bases

^cLength of intervening sequence between the two stem arms.

^dSequences outside the repetitive element were also predicted to form part of the hairpin structure