HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2017 June 19.

Published in final edited form as:

J Mol Biol. 2016 June 19; 428(12): 2542–2556. doi:10.1016/j.jmb.2015.10.021.

Base flipping by MTERF1 can accommodate multiple conformations and occurs in a stepwise fashion

James Byrnes[a,*], **Kevin Hauser**[b,*], **Leah Norona**[a], **Edison Mejia**[a], **Carlos Simmerling**[b,c], and **Miguel Garcia-Diaz**[a,**]

[a]Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY 11794, USA

[b]Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, USA

[c]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY 11794, USA

Abstract

Human mitochondrial transcription termination occurs within the leu-tRNA gene and is mediated by the DNA binding protein MTERF1. The crystal structure of MTERF1 bound to the canonical termination sequence reveals a rare base flipping event that involves the eversion of three nucleotides. These nucleotides are stabilized by stacking interactions with three MTERF1 residues, which are not only essential for base flipping but also for termination activity. To further understand the mechanism of base flipping we examined each of the individual stacking interactions in structural, energetic and functional detail. Individual substitutions of Arg162, Tyr288 and Phe243 have revealed unequal contributions to overall termination activity. Furthermore, our work identifies an important role for Phe322 in the base flipping mechanism and we demonstrate how Phe322 and Phe243 are important for coupling base flipping between the heavy and light strand DNA chains. We propose a step-wise model for the base flipping process that recapitulates our observations. Finally, we show that MTERF1 has the ability to accommodate alternate active conformations. The adaptability of base flipping has implications for MTERF1 function and for the putative function of MTERF1 at alternative binding sites in human mitochondria.

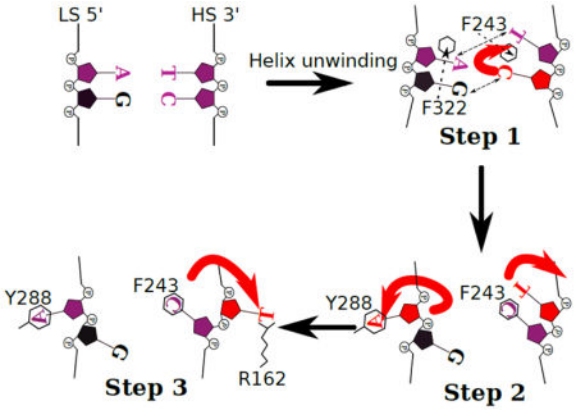
Graphical Abstract

**Corresponding author: ; Email: Miguel.garcia-diaz@stonybrook.edu. Tel: +1 631 444 3054; Fax: +1 631 444 9749

*These authors contributed equally to this work.

Accession codes

The atomic coordinates and structure factors have been deposited into the Protein Data Bank (PDB) with the following codes **5CKY, 5CO0, 5CRJ, 5CRK**, for the R162A, Y288A, F322A and F243A MTERF1 substitutions respectively.



Keywords

Mitochondria; transcription termination; X-ray crystallography; molecular dynamics; base flipping

INTRODUCTION

The human mitochondrial genome is a circular, closed, 16.5kb DNA molecule that encodes 22 tRNAs, 2 rRNAs and 13 mRNAs located on the heavy strand (HS) and light strand (LS) [1–5]. The expression of these genes produce protein subunits that make up the mitochondrial contribution of the oxidative phosphorylation system (OXPHOS), which is ultimately responsible for the generation of ATP [6,7]. Alterations to OXPHOS can result in decreased energy production and are associated with human pathology such as cancer and age-related diseases including aging [8]. OXPHOS is therefore dependent on proper expression of the mitochondrial genome. Expression is tightly regulated and depends on the import of nuclear proteins [9]. The first step in gene expression is transcription of the mitochondrial genome [10]. The initiation of transcription involves three proteins: a mitochondrial RNA Polymerase (POLRMT) and two transcription factors (TFAM and TFB2M) that assemble at the regulatory D-loop, an ~1kb non-coding region of the mitochondrial genome [11–13]. Here, bi-directional transcription initiates at promoters for the light strand (LS) and heavy strand (HS).

Transcription products of both strands can be terminated within the leu-tRNA gene and this termination event is dependent on the mitochondrial termination factor, MTERF1 [14]. The reason for and the precise location of termination at the MTERF1 binding site is not clearly understood. Because the termination site is located downstream of the heavy strand promoter (HSP) and the 2 rRNA genes, it was suggested that, in concert with appropriate selection of alternative HSP promoters through a secondary MTERF1 binding site, this site is important for regulating rRNA production [3,15–17]. Alternatively, genetic and biochemical evidence suggests that this termination site is important as an obligate site for LSP transcriptional termination [18]. Regardless of the reason for termination here, this event seems to be evolutionarily conserved, and this site is also the most strongly protected from DNase digestion in the mtDNA molecule, indicative of tight protein binding [19]. Furthermore, pathogenic mtDNA mutations are found in the termination site. These mutations include

A3243G, which is the mtDNA mutation most frequently found in the population. It has been demonstrated that some of these mutations can negatively affect termination activity [20–22]. In addition to mediating transcription termination, a role has been suggested for MTERF1-mediated pausing at alternative sites in the mtDNA genome during DNA replication [23].

The MTERF1 protein belongs to a family of factors (MTERFs) found in metazoans and plants [24,25]. MTERF proteins are exclusively localized to mitochondria and chloroplasts, where they often play essential roles [25,26]. Their functions have been shown to involve regulation of mitochondrial DNA replication, transcription and translation [15,27–29]. Crystal structures of MTERF proteins have shown that these proteins belong to a novel class of all helical tandem-repeat DNA binding proteins [22,30–32]. The crystal structure of MTERF1 bound to its termination sequence demonstrates a unique DNA binding mode, where sequence specificity is mostly maintained through interactions with the DNA major groove established by 5 conserved arginine residues [22]. Most strikingly, MTERF1 mediates a rare base flipping conformation that involves unwinding of the double stranded DNA and the eversion of three nucleotide bases. The everted nucleotides are extra-helically stabilized by stacking interactions with three residues, Arg162, Phe243 and Tyr288, which are critical for MTERF1 termination activity [22,33]. Our previous work established the importance of base flipping for termination activity through an R162A, F243A, Y288A (RFY) triple substitution that destabilizes base flipping. Although the RFY substitution and WT MTERF1 maintain a similar binding mode and are both capable of specifically recognizing the termination sequence, RFY is deficient in its ability to terminate transcription [22]. This suggests that the binding event alone is not sufficient for termination and underscores the role of base flipping for termination. However, the mechanism through which this unique mode of base flipping occurs remains obscure [34]. Since there are three bases involved, the mechanism of eversion is likely a multi-step, complex process. Therefore, we sought to understand the mechanical steps involved in nucleotide eversion. In this study, to break down the process of base flipping, we have created individual substitutions that disrupt the stacking interactions one at a time. We have evaluated the structural, energetic and functional consequences of these substitutions and show that individual stacking interactions do not contribute equally to function. We have identified Phe322 as an additional residue important for base flipping and propose a model for the base flipping mechanism. Our data also demonstrate the existence of alternate base flipped conformations, some of which are still compatible with termination activity. This could have implications for the ability of MTERF1 to adapt to functional roles at other sites in the mitochondrial genome.

RESULTS

Functional differences among Y288A, R162A and F243A substitutions

The X-ray crystal structure of human MTERF1 at 2.2Å bound to the termination sequence reveals an all-helical protein with a right-handed superhelical fold that wraps around the major groove of DNA (Supplementary Figure 1). Furthermore, the protein induces DNA bending and helix unwinding, which results in the eversion of three nucleotides (C3242 and T3243 in the heavy strand and A3243 in the light strand) stabilized by stacking interactions as shown schematically in Figure 1A and structurally in Figure 1B. Eliminating the three

stacking interactions by an RFY to AAA triple substitution results in a severe functional loss [22]. However RFY does not inform on the specific role of base flipping or the mechanism of eversion. To further investigate base flipping in MTERF1 we created, expressed and purified R162A, Y288A and F243A individual substitutions, and performed termination assays to assess their function on a model substrate as previously reported [22]. The resulting UREA–PAGE gel shows two distinct bands corresponding to a Full Length (FL) runoff RNA transcript and a smaller termination (T) transcript that represents MTERF1 mediated termination (Figure 1C). A control experiment (C) containing only the initiation machinery yields a single band corresponding to the FL transcript. The percentage of termination for each substitution is shown graphically in Figure 1D. Interestingly, the termination defect for each mutant is not equal. Whereas both R162A and F243A have a severe functional defect that cannot be rescued by increasing the protein concentration, the Y288A substitution results in a protein that is deficient at low concentrations, but closer to wild type at higher concentrations. Interestingly, R162A displays the strongest effect, with only residual termination activity. Thus, unexpectedly, the termination assays demonstrate that the functional contribution of each stacking residue is clearly not equal with respect to function.

Energy decomposition explains the importance of Arg162, suggests a role for Phe243

The observed functional differences when replacing each one of the stacking residues in WT MTERF1 with alanine suggest that each of these residues could contribute unequally to the stability of the functional complex. Thus, we tested the hypothesis that Arg, Phe and Tyr residues contribute unequally to the stability of the flipped bases. If true, this would imply that either the energetic contribution of each stacking interaction is different or, besides stacking, additional interactions established by Tyr288, Phe243 and Arg162 to the DNA result in a differential energetic contribution. Structural analysis of the MTERF1 crystal structure does not immediately explain these differences. Thus, in order to understand how the interactions between the stacking residues and the DNA (Figure 1B) compare with each other, we used MD simulations and energy decomposition to evaluate the energies of WT MTERF1-DNA interactions (Figure 2). We reasoned that the functional differences could in part be explained by the presence of additional interactions between Arg162 and the DNA backbone. As a control, energy decomposition identified interactions expected to be favorable - between Phe243, Tyr288 and Arg162 with the everted nucleotides (Figure 2A, B, C). Next, we analyzed all the interactions between these residues and the DNA to better understand how Tyr288, Phe243 and Arg162 support the everted bases and stabilize the altered backbone conformation observed in the crystal structure.

This energy analysis reveals the importance of the nonspecific interactions established between Arg162 and the DNA backbone (Figure 2C). In addition to stacking with T3243, Arg162 forms a nonspecific interaction between its positively charged guanidinium group and the negatively charged phosphate group of T3241 (2.6Å; Figure 1B). Thus, in addition to stacking with one of the everted residues, like Phe243 or Tyr288, Arg162 also interacts with the DNA backbone through nonspecific electrostatic interactions. Moreover, Arg162 might be key to stabilizing the otherwise unfavorable dipole-phosphate interaction between helix 2 in the fourth mterf motif and the phosphate group of T3241 (Figure 2E) [22].

Furthermore, this interaction might stabilize the severe kink in the HS backbone that takes place in the base flipping region (arrow in Figure 1B). This kink is essential to facilitate the conformation observed in the crystal structure. An *in silico* R162A mutation reveals the dramatic change in electrostatic potential between the dipole of helix 2 and T3241, C3242, and T3243 (Figure 2E, F). Hence, Arg162-DNA nonspecific interactions provide a potential explanation for the severity of the phenotype in the R162A substitution.

Energy decomposition also revealed that Phe243 establishes a weak average interaction with T3243 (Figure 2A), −0.7 kcal/mol. We decided to inspect each of the four independent MD simulations to structurally understand this unexpected interaction. In the simulations, T3243 spontaneously flipped back into the duplex by forming a new double t-stacking interaction with C3242 and Phe243 (Supplementary Figure 2A, B). The average F243-T3243 interaction energy in these simulations was −1.2 kcal/mol. These data suggest that T3243 flipping and C3242 flipping may be coupled via Phe243.

X-ray crystal structures of R162A, Y288A and F243A substitutions

The MD analysis indicated that Arg162 makes the strongest energetic contribution to the stability of the functional termination complex, and therefore helps explain the importance of this side chain for termination. The simulations also suggest a possible role of Phe243 in stabilizing the flipping of T3243, since we observed the base to spontaneously form a double t-stacking interaction with C3242 and Phe243 after partial reversion into the duplex (Supplementary Figure 2). This suggests that the functional termination difference might result from contributions of these residues to the base flipping process, which involves intermediates not observed in the WT crystal structure. Indeed, base flipping of three residues, combined with duplex unwinding, is likely to be a complex, multi-step mechanical process, and the importance of specific residues cannot be recapitulated by potential energy decomposition analysis of a single conformation.

We thus decided to crystallize each individual substitution in order to assess any conformational changes that would provide insight into the base flipping mechanism. Each structure was solved using the wild type (PDB: 3MVA) structure as a search model for molecular replacement. The structures were solved to a resolution of 2.6Å, 2.6Å and 2.5Å for R162A, Y288A and F243A, respectively. Crystallographic and refinement statistics are shown in Table 1.

The overlay of R162A (yellow) on wild type (gray) shows an altered conformation of the HS (Figure 3A). As expected (since Arg162 stacks with T3243 in the wild-type crystal structure) the conformation of T3243 is altered and located within the two DNA strands (Figure 3A). T3243 now interacts with Arg195 through a hydrogen bond and with Phe243 through a t-stacking interaction. (Supplementary Figure 3A). Interestingly, C3242 fully swings out of the unwound DNA helix and occupies the space where the Arg162 side chain was in the WT structure (gradient arrows, Figure 3A). In addition, the conformation of C3242 is identical to that of the triple RFY mutant, both of which are devoid of termination activity [22]. In contrast, the conformation of the LS remains largely unaltered:[22] A3243 is stacked with Tyr288 in the same fashion as observed in the wild type structure (Figure 3A).

By contrast, the loss of the stacking interaction with A3243 in the Y288A structure results in an altered conformation of the A3243 on the LS (Figure 3B). The overlay of Y288A (cyan) and wild type (gray) shows that the A3243 is only partially flipped out from the unwound DNA helix. Moreover, the conformation of the HS in the Y288A structure is also altered, but not as drastically compared with what was observed in R162A. Whereas the conformations of Arg162 and C3242 are similar to those observed in the wild type structure, T3243 is located within the helix as in the R162A substitution (Figure 3B, gradient arrow). The conformation of T3243 is very similar to that observed in the R162A substitution, forming a t-stacking interaction with Phe243 (Supplementary Figure 3A). However, the hydrogen bond with Arg195 is not observed, but instead T3243 interacts with its neighboring nucleotide through a hydrogen bond between O2 of T3243 and N4 of C3244 (3.4Å; Supplementary Figure 3B). Interestingly, the perturbation of the HS conformation resulting from a substitution that only directly affects the LS implies a coupling between the conformations of the LS and the HS.

Consistent with the severe termination defect, the structure of F243A reveals multiple alterations of both the LS and HS (Figure 3C). Phe243 forms a stacking interaction with C3242 in the wild-type structure and importantly, C3243 is the only everted base that is not fully extrahelical. As expected, in the F243A structure the conformation of C3242 is severely altered. The absence of the hydrophobic side chain seems to make the WT conformation of C3242 unstable and the Cytosine is now in a position where it is still forming Watson-Crick-like hydrogen bonds with its Guanine pair (Figure 3D). Interestingly, the conformation of T3243 is very close to that in the wild type structure, stabilized by stacking with Arg162. However, further stressing the coupling between HS and LS conformations, A3243 is located in a helical position.

Overall, these structures reveal that the conformation of C3242 closely correlates with termination activity, suggesting that the conformation of this base is critical to stabilize the transcriptionally active conformation of the protein. The structures also suggest that some of the aberrant conformations adopted by the MTERF1 substitutions can be stabilized by compensatory interactions, which presumably stabilize the final state sufficiently to enable close to wild-type termination activity in the case of Y288A.

One of the most striking observations from the structures is the conformation of C3242 in the F243A structure. As stated above, C3242 is still essentially base-paired with G3242, although the base pair is buckled (Figure 3D). This structure might reveal a conformation that represents an intermediate in the base flipping mechanism, and suggests a pathway for the movement of C3242 from its G-C base pair to the everted conformation observed in the wild type structure bound to DNA. Phe-mediated buckling is a conserved mechanism of base flipping in bacterial and human glycosylases [37]. An overlay between the F243A and wild-type structures suggests that this transition is mediated by a conformational change of the side chain of Asp283 (Figure 3E). This residue interacts with the N4 of C3242 and appears to facilitate breaking the G/C pair and follow the movement of C3242 from its intrahelical position to its final position where it stacks with Phe243 and hydrogen-bonds to Glu280 (Figure 3E).

Phe322 and Phe243 are important for the base flipping mechanism

As mentioned above, the F243A structure might represent an intermediate in the base flipping mechanism where A3243 is still located in an intrahelical conformation. Importantly, A3243 in this conformation is stabilized by a t-stacking interaction with Phe322 (Figure 3F). Phe322 is not involved in any obvious interactions with the DNA in the final state captured in the wild-type crystal structure. Moreover, the energy decomposition finds no Phe322-DNA interaction stronger than 1 kcal/mol (Figure 2D). This indicates its role in stabilizing the functional complex is less important than Arg162 and Tyr288. Since the F243A mutant led to diminished termination activity and an altered DNA structure (Figures 1C, D & Figure 3F), the energy decomposition deductively suggests that its role in function might be to stabilize recognition intermediates and base flipping. Both observations are consistent with the idea that Phe322 might be important to stabilize intermediate steps during base flipping.

Interestingly, in the R162A and Y288A structures, T3243 is located within the helix, adopting conformations that are likely intermediate in the base flipping process. Observing the environment of T3243 in those structures reveals a striking parallelism with the Phe322-A3243 interaction seen in the F322A structure: Phe243 forms an analogous t-stacking interaction with T3243 (Supplementary Figure 3A). Similarly to what is observed with Phe322, no interactions are observed between Phe243 and T3243 in the wild-type crystal structure. The energy decomposition analysis, however, did indicate that Phe243 and T3243 interact (1 kcal/mol) in the WT structure (Figure 2A and Supplementary Figure 2A).

Since the results of the MD simulations were based on the final WT functional state, the interactions observed in our substitution crystal structures provide new insight into the flipping mechanism. Our data suggests that both Phe243 and Phe322 play an important role in the base flipping process, helping to stabilize intermediates before the bases are flipped-out of the helix. Interestingly, an overlay between the Y288A and F243A structures reveals a hypothetical intermediate during base flipping where both T3243 and A3243 are still between both DNA strands (Figure 4A). This suggests a model where Phe243 and Phe322 intercalate between adjacent bases in the HS and LS, respectively, and stabilize their intrahelical conformations. Sequence conservation information analysis reveals the deep conservation of Phe322, supporting the functional importance of Phe322 as a wedge (Supplementary Figure 4), similar to that seen in glycosylases [37]. Importantly, given the position of both phenylalanines, Phe322 and Phe243 appear to be important to help buckle and break the A3243/T3243 base pair (dashed arrow, Figure 4A).

To test this hypothesis, we constructed an F322A substitution and tested its biochemical effect on MTERF1 function using a termination assay as above. As can be seen (Figure 4B), F322A results in a large functional defect, almost of a similar magnitude to that observed with an F243A substitution. Given the analogous role that could be played by both residues (wedge), we then constructed a double F243A/F322A substitution to test whether MTERF1 utilizes a double-wedge mechanism of base flipping. If this hypothesis was true, the double mutation should have an additive defect in termination activity. The resulting gels show that both substitutions indeed have additive effects, as the double mutant results in a protein severely impaired in its ability to mediate termination activity (Figure 4B).

Phe322 and Phe243 play an important role in coordinating the base flipping mechanism

To further investigate the importance of Phe322 for the base flipping process, we decided to determine the structure of the F322A substitution. The structure was solved to 2.6Å resolution (Table I). Interestingly, despite the fact that Phe322 stabilizes LS intermediate conformations, the F322A structure (Figure 4C) reveals that the substitution results in a perturbed final state in which the LS is in a wild-type conformation, but where C3242 is located in an aberrant conformation outside the helix. Since Phe322 does not play a role in stabilizing the final conformation, this result further supports that the function of Phe322 is important during the base flipping process. Moreover, the fact that the F322A substitution results in alterations in the HS conformation, even though the residue would mainly appear to interact with residues in the LS, implies that intermediate LS conformations affect the base flipping process in the HS, substantiating that the mechanisms of base flipping in the HS and LS are coupled.

DISCUSSION

In this study we have biochemically, computationally and structurally characterized the importance of several key residues directly involved in the mechanism of base flipping by MTERF1. Consistent with the observation that sequence recognition is mostly mediated by direct readout [38], base flipping is not strictly required for DNA binding by MTERF1. However, base flipping is important for MTERF1 transcriptional termination activity. Although the precise mechanism of termination is still unknown, base flipping presumably allows the protein to kinetically anchor itself to the termination sequence, allowing MTERF1 to block the transcription machinery. This mechanism involves the eversion of three nucleotides, implying that the base flipping mechanism is likely complex. We have discovered that residues Tyr288, Arg162 and Phe243 that directly interact with each of the everted nucleotides are all important to stabilize the final state of base flipping. Eliminating these side chains results in altered conformations of the base-flipped state, highlighting the importance of π-stacking interactions in driving and stabilizing an intricate mechanism to flip out three bases. Our results also show that the three side chains are not equally critical for function. In the absence of key interactions observed in the wild-type conformation, MTERF1 adopts alternate binding modes that can maintain an appreciable amount of activity. The structural adaptation suggests that the MTERF1 binding mode is sufficiently flexible to support binding and at least partial termination activity in different sequence contexts. This might have implications for the alternate MTERF1 binding sites that have been characterized in the mitochondrial genome [23], where the full consensus sequence observed in the wild-type crystal structure is not present.

MD simulations and energy decomposition of the WT MTERF1-DNA complex illuminated the importance of Arg162 in establishing multiple nonspecific interactions with the DNA backbone. These interactions presumably stabilize the kinked backbone of DNA between T3241 - C3242 and C3242 - T3243. The positive charge carried by the guanidinium group of Arg162 not only stabilizes the precarious geometry of 3 phosphate groups, but also mediates a destabilizing electrostatic force imparted onto the phosphate groups by the dipole of helix 2 of mterf motif 4. The energy analysis also led to a peculiar result that Phe243

established a weak, ~1 kcal/mol interaction with T3243. Upon inspection of the MD trajectories, we found that T3243 spontaneously reverted into a pseudo-helical position in which the pyrimidine ring t-stacked with C3242 and Phe243. The simulations suggested that the key residues hypothesized to be key for base flipping may serve to not only stabilize the final flipped state but also to, perhaps, stabilize the flipping process itself (Supplementary Figure 2).

To test these specific hypotheses, we solved the crystal structures of substitutions for all four residues (Arg162, Phe243, Tyr288, Phe322). Our crystal structures reveal altered conformations that might be representative of intermediate base flipping states, and thus provide information on the MTERF1 eversion mechanism. The structures reveal interactions not present in the final state of base flipping, and have allowed us to identify additional residues important for this process. Furthermore, the structures provide evidence that a high degree of coordination exists between base flipping events in the heavy and light strands.

A key mechanism that appears to be at the core of this coordination is related to the role of two phenylalanine side chains, Phe243 and Phe322. In addition to the role of Phe243 stabilizing C3243 in the final state, these residues help stabilize intermediate conformations during base flipping. As we have shown, alterations to this phenylalanine-mediated mechanism lead to perturbations in the conformation of the final state, and reveal adjustments in the coupling between HS and LS conformations. Furthermore, the structures provide details on the nature of this mechanism. For instance, Phe243 appears to stabilize the intrahelical conformation of T3243. Once T3243 is flipped out of the DNA helix, Phe243 interacts exclusively with C3242. However, Phe243 is able to simultaneously interact with both residues, as shown in the Y288A structure (Supplementary Figure 3B) and MD simulations (Supplementary Figure 2), implying that base flipping in the HS follows a stepwise mechanism. Thus, C3242 is first delivered to its final position through a network of interactions involving Asp283 (Figure 3D), while Phe243 stabilizes the intrahelical conformation of T3243. Similarly, Phe322 stabilizes the intrahelical conformation of A3243, preceding base flipping. The importance of Phe322 for base flipping in the HS suggests that this intermediate conformation must be at least transiently maintained in order to properly coordinate HS base flipping. Although the exact mechanism(s) by which HS-LS coordinate base flipping is still unclear, our data illuminate the importance of a network of hydrogen bonds (Asp283 and Arg195) and sequential t-stacking and π-stacking interactions between Phe243, Phe322, T3243 and C3242. We demonstrate that the absence of Phe322 leads to perturbed base flipping, suggesting that this residue is important for coordinating the steps preceding base flipping and in conjunction with Phe243, seem to be important in buckling the A3243/T3243 base pair.

Based on our observations we are able to propose an ordered step-wise model for the base flipping mechanism. Helped by the helix unwinding that occurs upon MTERF1 binding to its target sequence, the first crucial step is the intercalation of Phe322 and Phe243, which stabilize the intrahelical conformations of A3243 and T3243 and buckle the base pair, similar to the phenylalanine wedge (F-wedge) in glycosylases [37]. The formation of a hydrogen bond between Asp283 and C3242 destabilizes the Watson-Crick hydrogen bonds between C3242/G3242, extrudes C3242 from the helix, and subsequently delivers the C3242

to its final conformation. Since the final conformation of C3242 is incompatible with the intrahelical conformation of A3243, it is likely that the motion of C3242 towards Phe243 is coupled with A3243 flipping. Lastly, the interaction between Phe243 and C3242 provide a hydrophobic pocket along the trajectory by which T3243 would flip out of the duplex (Figure 5).

Our structural observations recapitulate our step-wise model for the mechanism of MTERF1 base flipping. For example, the coupling of the C3243 motion to A3243 flipping is consistent with the conformations observed in the F243A structure, where in the absence of the motion of C3242 towards Phe243, A3243 is located within the duplex and stabilized by Phe322. Furthermore, any alteration to this mechanism results in an altered final conformation. For instance, decoupling C3242 from the flipping of A3243 in the F322A structure, or premature base flipping of C3242 in the R162A structure, results in an aberrant conformation of C3242 that prevents T3243 from flipping out of the duplex. Moreover, incomplete base flipping of A3243 due to the absence of the Tyr288 side chain is coupled to T3243 base flipping through a parallel mechanism.

In summary, we have shown that the eversion of three nucleotides by MTERF1 is a complex multi-step process that is important for transcriptional termination. However, we have demonstrated that this process is relatively flexible, and that some conformations are consistent with a certain degree of termination activity. This might explain the residual termination activity observed in pathogenic mtDNA mutations like A3243T [22] [20], suspected to affect base flipping. Importantly, base flipping involves not only residues that stabilize the final base flipped state but also residues responsible for establishing intermediate interactions that are important in coordinating this tightly coupled process. Our work has taken an important step towards elucidating the complete mechanism underlying the process of base flipping that takes place in MTERF1. Since the base flipping process might be more tolerant of sequence alterations than previously thought, structural adaptation also has implications for the ability of MTERF1 to perform putative additional roles at alternative binding sites as has been previously suggested. It is of interest to note that wedge residues with bulky hydrophobic side chains – phenylalanine in particular – are conserved in bacterial (Fpg) and human (hoGG1) glycosylases to flip nucleobases from within a duplex for DNA damage repair. The biochemical nature of phenylalanine may therefore be an evolutionary solution to the problem of base flipping and our work underscores the role of phenylalanine in the rare 3 nucleotide eversion process.

MATERIALS AND METHODS

Mutagenesis, Protein Expression & Purification

Wild type MTERF1 (residues 57–399) was expressed as a fusion with maltose binding protein (MBP) and a 6X His tag [22]. MTERF1 substitution constructs were created using the QuikChange Site-Directed Mutagenesis protocol (Stratagene, La Jolla, CA). PCR products were cleaned using the MinElute PCR purification cleanup kit (Qiagen) and successful mutagenesis was confirmed by DNA sequencing. Expression and purification of all MTERF1 substitutions were carried out as previously described [22].

Transcription termination

Transcription assays were adapted from Asin-Cayuela [39] and described in Yakubovskaya, et al [22]. Briefly, a circular template DNA plasmid (pet 22b, Novagen) containing the heavy strand promoter (HSP) and the MTERF1 binding sequence 100bp downstream in reverse orientation with respect to HSP was linearized with HindIII. The template DNA was then incubated with MTERF1 (0.8pmole, 0.4pmoles or 0.2pmoles), rNTPs (0.4mM ATP, 0.15mM CTP and GTP, 0.01mM UTP), DNA template (30ng, 8.3nmoles), 0.5ul (5uCi) of α-^{32}P labeled UTP and the initiation machinery TFAM (0.4fmoles), TFB2 (0.2fmol) and POLRMT (0.2fmol) and transcription buffer (150mM KCl, 20mM Hepes pH8.0, 5mM DTT, 1mM EDTA and 10mM $MgCl_2$) in a total reaction volume of 20ul. The reaction was carried out for 30 minutes at 32°C and stopped using 100ul of 1%SDS, 20mM EDTA, 300mM Sodium Acetate and 20ug Calf thymus DNA. The samples were then ethanol precipitated and run on a 5% UREA-PAGE gel. The gel was dried and exposed to a phosphorimager (Amersham Biosciences) and scanned using the Typhoon FLA 9000 scanner and software package (V1.2, GE). Each experiment was repeated at least three times. Densitometry analysis was performed using the ImageQuant TL (GE) software and for each gel a control lane (no MTERF1) was used to account for background and subtracted from the termination band. In addition a correction factor was applied to account for the increased UTP incorporation observed in the runoff band and the percentage termination was calculated. Graphs were made using Prism 5 for Mac OS X version 5.0c.

X-ray Crystallography

The wild type 22mer DNA oligo was added to each of the MTERF1 substitution proteins for a final ratio of 2.5:1 and diluted with buffer containing 200mM KCl, 20mM Hepes pH8.0, 1mM DTT and 1mM EDTA. Crystals were grown using the hanging drop vapor diffusion method at room temperature in 2ul drops containing a 1:1 ratio of reservoir solution (0.2M Sodium acetate, 0.1M Tris HCl pH 8.0, 15.5% Peg 4000) for R162A, Y288A, F243A and F322A to the protein/DNA mixture. Crystals were cryoprotected using 25–30% ethylene glycol and flash frozen in liquid nitrogen. Diffraction data was collected using the x9, x25 and x29 beamlines (Table 1) at the National Synchrotron Light Source (NSLS, Upton, NY). Data were then processed using XDS [40] and Scala/AIMLESS, carried out within the autoPROC toolbox [41]. Molecular replacement (MR) and refinement was then performed using MOLREP [42] and REFMAC5 [43,44] in the ccp4 software suite [45] and manual model building was done using COOT [46]. Figures were made using pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC).

MD simulations

System Preparation—We prepared chemical topologies and coordinates for the WT MTERF1-DNA complex based on the previously published crystal structure, PDB ID 3MVA [22]. The unassigned N-terminal glutamate and aspartate residue side chains of the crystal structure were built using Amber libraries [36]. MolProbity [47] was used to analyze all rotamers. We selected the A conformation of Ser292. Hydrogen atoms were added using standard procedures in Amber [36]. We used a truncated octahedron solvent with a minimum 10 Å buffer from any solute atom to edge of the periodic box. This fully solvated system was

61042 atoms and its dimensions were used to determine that 83 K+/Cl− ions be added to mimic the in vitro 0.2 M concentration; 32 neutralizing K+ ions were also added. Ion placements were randomized. The in silico R162A mutation followed the exact procedure as above, except the sidechain atoms of R162 were replaced by a methyl.

Equilibration and production MD—We equilibrated the system using restraints to only atoms resolved in the crystal structure, except crystallographic water. In stage 1 we minimized the model with complete force evaluations, periodic boundary conditions, and a non-bonded interaction cutoff of 8.0 Å, for 10,000 steepest descent steps; atomic coordinates from the crystallographic model were positionally restrained with a force constant of 100 kcal/mol Å^2 to their initial coordinates. In stage 2 we linearly ramped the temperature from 100 K to 300 K over 200 ps, with SHAKE [48] constraints on all bonds involving hydrogen; we used the Berendsen thermostat [49] with temperature and pressure bath coupling constants of 0.1 ps^{-1}, with a force constant of 100 kcal/mol Å^2 to the minimized structure, integrated over 1 fs steps, and velocities initialized. In stage 3, we resumed the second stage velocities and thermostat at 300 K with the same coupling constants but with pressure held constant, a 100 kcal/mol Å^2 positional restraint force constant to the final structure from stage 2, was used for 100 ps. In stage 4, we resumed conditions from stage 3 but with 0.5 ps^{-1} pressure and temperature bath coupling constants and weaker positional restraint force constant of 10 kcal/mol Å^2 to the final structure from stage 3 for 250 ps. In stage 5, we changed the restrained atoms to only MTERF1 backbone (CA, N, C) or DNA backbone (O5′, C5′, C4′, C3′, O3′, O2P, O1P, P, O4′, C2′, C1′) atoms with a force constant of 10 kcal/mol Å^2 to the final structure from stage 4, for 10,000 steps. In stage 6 we use identical conditions as stage 4 but with a 10 kcal/mol Å^2 restraint force constant to the final structure from stage 5 for 100 ps. In stage 7, we resumed velocities and conditions from stage 6 with a weaker 1 kcal/mol Å^2 restraint force constant to the final structure from stage 6 for 100 ps. In stage 8, we resumed from stage 7 and restrained to its final structure with a force constant of 0.1 kcal/mol Å^2 for 100 ps. In the final stage of equilibration, free of restraints, we resumed the conditions and velocities from stage 8 for 250 ps. Production dynamics resumed velocities and conditions from the final equilibrated structure, with the integration step increased to 2 fs, sampling for 1 ns. Equilibration and production calculations utilized the single precision-double precision PMEMD implementation of SANDER for graphical processing units (GPUs), release 14 by Le Grand et al.[50]; the minimization calculations utilized the CPU implementation of PMEMD. Four independent trajectories were generated by utilizing different random seeds for velocity initialization in stage 2 of the equilibration.

Energy decomposition—For each of the four independent 1 ns production trajectories of the WT MTERF1-DNA complex, pair-wise per residue potential energy decomposition was performed using the following standard procedures [37]. We used AMBER energy decomposition [51] to calculate the pairwise per-residue interaction energies, $\bar{G} = \bar{E}_{MM} + \bar{G}_{GBSA} - TS_{MM}$,[52] where $\bar{G}_{GBSA} = G_{ELE} + \gamma A$, with G_{ELE} calculated using the Generalized Born model of Onufriev, Bashford and Case [53], A the solvent accessible surface area calculated using icosahedral recursion [54], and the scaling factor γ set to 0.005 kcal/molÅ^2 [55]. TS_{MM} was assumed to contribute negligibly to the energy [52] and E_{MM} only included the electrostatic and van der Waals energies [51]. For each of the four residues

R162, F243, Y288, and F322, we calculated the interaction energy, $\bar{G}$, of all the atoms within the residue to all the atoms in each of the nucleotides, leading to 44 residue-nucleotide energies for each of the 2000 MD snapshots (1 ns). Each residue-nucleotide interaction was averaged and standard deviations calculated to estimate data precision. Poisson-Boltzmann electrostatic surfaces were calculated using standard procedures in Amber. The dipole of helix 2 was calculated by including only the backbone atoms of the helix. Figures were generated using VMD version 1.9.1[56].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank the NSLS Protein Crystallography group for beamline support and the Laufer Center for computational support. NSLS beamlines ×9, ×25 and ×29 are mainly supported by the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and the National Center for Research Resources of the National Institutes of Health. This work was supported by NIH R00-ES015421, R01-GM100021, a UMDF grant to M.G.-D., NIH F31-GM101946 to K.H. and the Stony Brook Chemical Biology Training grant (NIH T32-GM092714) to J.B. and K.H.

ABBREVIATIONS

HSP	Heavy Strand Promoter
LSP	Light Strand Promoter
MTERF	Mitochondrial Transcription Termination Factor
RFY	R162A F243A Y288A triple substitution

References

1. Shutt TE, Shadel GS. A compendium of human mitochondrial gene expression machinery with links to disease. Environ Mol Mutagen. 2010; 51:360–79. [PubMed: 20544879]
2. Anderson S, Bankier AT, Barrell BG, De Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature. 1981; 290:457–65. [PubMed: 7219534]
3. Lodeiro MF, Uchida A, Bestwick M, Moustafa IM, Arnold JJ, Shadel GS, et al. Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro. Proc Natl Acad Sci USa. 2012; 109:6513–8. [PubMed: 22493245]
4. Asin-Cayuela J, Gustafsson CM. Mitochondrial transcription and its regulation in mammalian cells. Trends Biochem Sci. 2007; 32:111–7. [PubMed: 17291767]
5. Gaspari M, Larsson N-G, Gustafsson CM. The transcription machinery in mammalian mitochondria. Acta-Bioenerg. 2004; 1659:148–52.
6. Scheffler, IE. Mitochondria. John Wiley & Sons; 2011.
7. Hatefi Y. The mitochondrial electron transport and oxidative phosphorylation system. Annu Rev Biochem. 1985
8. Wallace DC. Mitochondrial DNA mutations in disease and aging. Environ Mol Mutagen. 2010; 51:440–50. [PubMed: 20544884]
9. Scarpulla RC. Transcriptional activators and coactivators in the nuclear control of mitochondrial function in mammalian cells. Gene. 2002; 286:81–9. [PubMed: 11943463]
10. Scarpulla RC. Nuclear control of respiratory chain expression in mammalian cells. J Bioenerg Biomembr. 1997; 29:109–19. [PubMed: 9239537]

11. Sologub M, Litonin D, Anikin M, Mustaev A, Temiakov D. TFB2 Is a Transient Component of the Catalytic Site of the Human Mitochondrial RNA Polymerase. Cell. 2009; 139:934–44. [PubMed: 19945377]
12. Arnold JJ, Smidansky ED, Moustafa IM, Cameron CE. Human mitochondrial RNA polymerase: structure-function, mechanism and inhibition. Biochim Biophys Acta. 2012; 1819:948–60. [PubMed: 22551784]
13. Schwinghammer K, Cheung ACM, Morozov YI, Agaronyan K, Temiakov D, Cramer P. Structure of human mitochondrial RNA polymerase elongation complex. Nat Struct Mol Biol. 2013; 20:1298–303. [PubMed: 24096365]
14. KRUSE B, NARASIMHAN N, ATTARDI G. Termination of Transcription in Human Mitochondria - Identification and Purification of a Dna-Binding Protein Factor That Promotes Termination. Cell. 1989; 58:391–7. [PubMed: 2752429]
15. Roberti M, Polosa PL, Bruni F, Manzari C, Deceglie S, Gadaleta MN, et al. The MTERF family proteins: mitochondrial transcription regulators and beyond. Biochim Biophys Acta. 2009; 1787:303–11. [PubMed: 19366610]
16. Zollo O, Tiranti V, Sondheimer N. Transcriptional requirements of the distal heavy-strand promoter of mtDNA. Proc Natl Acad Sci USa. 2012; 109:6508–12. [PubMed: 22454497]
17. Martin M, Cho J, Cesare AJ, Griffith JD, Attardi G. Termination Factor-Mediated DNA Loop between Termination and Initiation Sites Drives Mitochondrial rRNA Synthesis. Cell. 2005; 123:1227–40. [PubMed: 16377564]
18. Terzioglu M, Ruzzenente B, Harmel J, Mourier A, Jemt E, López MD, et al. MTERF1 Binds mtDNA to Prevent Transcriptional Interference at the Light-Strand Promoter but Is Dispensable for rRNA Gene Transcription Regulation. Cell Metab. 2013; 17:618–26. [PubMed: 23562081]
19. Rebelo AP, Williams SL, Moraes CT. In vivo methylation of mtDNA reveals the dynamics of protein-mtDNA interactions. Nucleic Acids Res. 2009; 37:6701–15. [PubMed: 19740762]
20. Helm MM, Florentz CC, Chomyn AA, Attardi GG. Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human tRNALys and tRNALeu(UUR). Nucleic Acids Res. 1999; 27:756–63. [PubMed: 9889270]
21. Chomyn A, MARTINUZZI A, YONEDA M, DAGA A, HURKO O, JOHNS D, et al. Melas Mutation in Mtdna Binding-Site for Transcription Termination Factor Causes Defects in Protein-Synthesis and in Respiration but No Change in Levels of Upstream and Downstream Mature Transcripts. Proc Natl Acad Sci USa. 1992; 89:4221–5. [PubMed: 1584755]
22. Yakubovskaya EE, Mejia EE, Byrnes JJ, Hambardjieva EE, Garcia-Diaz MM. Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. Cell. 2010; 141:982–93. [PubMed: 20550934]
23. Hyvärinen AK, Pohjoismäki JLO, Reyes A, Wanrooij S, Yasukawa T, Karhunen PJ, et al. The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. Nucleic Acids Res. 2007; 35:6458–74. [PubMed: 17884915]
24. Babiychuk E, Vandepoele K, Wissing J, Garcia-Diaz M, De Rycke R, Akbari H, et al. Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. Proc Natl Acad Sci USa. 2011; 108:6674–9. [PubMed: 21464319]
25. Linder T, Park CB, Asin-Cayuela J, Pellegrini M, Larsson N-G, Falkenberg M, et al. A family of putative transcription termination factors shared amongst metazoans and plants. Curr Genet. 2005; 48:265–9. [PubMed: 16193327]
26. Peralta S, Wang X, Moraes CT. Mitochondrial transcription: Lessons from mouse models. Biochimica Et Biophysica Acta (BBA)-Gene …. 2012
27. Park CB, Asin-Cayuela J, Cámara Y, Shi Y, Pellegrini M, Gaspari M, et al. MTERF3 Is a Negative Regulator of Mammalian mtDNA Transcription. Cell. 2007; 130:273–85. [PubMed: 17662942]
28. Pellegrini M, Asin-Cayuela J, Erdjument-Bromage H, Tempst P, Larsson N-G, Gustafsson CM. MTERF2 is a nucleoid component in mammalian mitochondria. Acta-Bioenerg. 2009; 1787:296–302.
29. Cámara Y, Asin-Cayuela J, Park CB, Metodiev MD, Shi Y, Ruzzenente B, et al. MTERF4 Regulates Translation by Targeting the Methyltransferase NSUN4 to the Mammalian Mitochondrial Ribosome. Cell Metab. 2011; 13:527–39. [PubMed: 21531335]

30. Jiménez-Menéndez N, Fernández-Millán P, Rubio-Cosials A, Arnan C, Montoya J, Jacobs HT, et al. Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat. Nat Struct Mol Biol. 2010; 17:891–3. [PubMed: 20543826]
31. Rubinson EH, Eichman BF. Nucleic acid recognition by tandem helical repeats. Curr Opin Struct Biol. 2012; 22:101–9. [PubMed: 22154606]
32. Spåhr H, Samuelsson T, Hallberg BM, Gustafsson CM. Structure of mitochondrial transcription termination factor 3 reveals a novel nucleic acid-binding domain. Biochem Biophys Res Commun. 2010; 397:386–90. [PubMed: 20430012]
33. Guja KE, Garcia-Diaz M. Hitting the brakes: termination of mitochondrial transcription. Biochim Biophys Acta. 2012; 1819:939–47. [PubMed: 22137970]
34. Byrnes J, Garcia-Diaz M. Mitochondrial transcription: How does it end? Transcription. 2011; 2
35. Ye X, Wang J, Wang J, Luo R, Luo R. A Revised Density Function for Molecular Surface Calculation in Continuum Solvent Models. J Chem Theory Comput. 2010; 6:1157–69. [PubMed: 24723844]
36. Case, DA.; VB; Berryman, JT.; Betz, RM.; Cai, Q.; Cerutti, DS.; Cheatham, TE., III; Darden, TA.; Duke, RE.; Gohlke, H.; Goetz, AW.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, TS.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, KM.; Paesani, F.; Roe, DR.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, CL.; Smith, W.; Swails, J.; Walker, RC.; Wang, J.; Wolf, RM.; Wu, X.; Kollman, PA. Amber. Vol. 14. University of California; San Francisco: 2014.
37. Kuznetsov NA, Kuznetsov NA, Bergonzo C, Bergonzo C, Campbell AJ, Campbell AJ, et al. Active destabilization of base pairs by a DNA glycosylase wedge initiates damage recognition. Nucleic Acids Res. 2015; 43:272–81. [PubMed: 25520195]
38. Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of Specificity in Protein-DNA Recognition. Annu Rev Biochem. 2010; 79:233–69. [PubMed: 20334529]
39. Asin-Cayuela J, Schwend T, Farge G, Gustafsson CM. The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form. J Biol Chem. 2005; 280:25499–505. [PubMed: 15899902]
40. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–32. [PubMed: 20124692]
41. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr. 2011; 67:293–302. [PubMed: 21460447]
42. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr. 2010; 66:22–5. [PubMed: 20057045]
43. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011; 67:355–67. [PubMed: 21460454]
44. Nicholls RA, Long F, Murshudov GN. Low-resolution refinement tools in REFMAC5. Acta Crystallogr D Biol Crystallogr. 2012; 68:404–17. [PubMed: 22505260]
45. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011; 67:235–42. [PubMed: 21460441]
46. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]
47. Chen VB, Arendall WB3, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]
48. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys. 1977; 23:327–41.
49. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. Molecular-Dynamics with Coupling to an External Bath. J Chem Phys. 1984; 81:3684–90.
50. Le Grand S, Götz AW, Walker RC. SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. Computer Physics Communications. 2013

51. Gohlke H, Kiel C, Case DA. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. Journal of Molecular Biology. 2003; 330:891–913. [PubMed: 12850155]

52. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res. 2000; 33:889–97. [PubMed: 11123888]

53. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004; 55:383–94. [PubMed: 15048829]

54. AMBER12 2012.

55. Sitkoff D, Sharp KA, Honig B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. Journal of Physical Chemistry. 1993; 98:1978–88.

56. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996; 14 33-8-27-8.

Highlights

- MTERF1 mediates a complex 3-base flipping process stabilized by stacking interactions
- Phe322 and Phe243 act as wedge residues preparing the DNA for flipping
- The mechanism of flipping 3 nucleotides occurs in a stepwise manner
- Perturbations to the stepwise mechanism alter MTERF1 termination ability

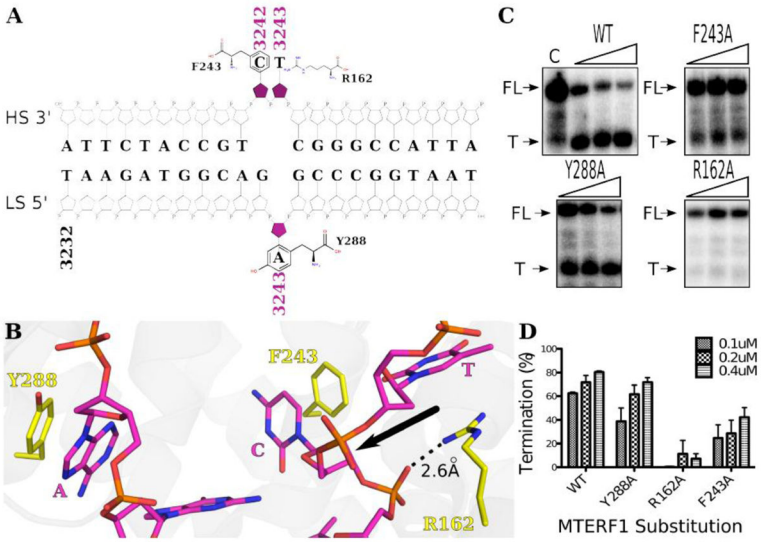


Figure 1.

(A) Schematic representation of the 22bp binding sequence of MTERF1 located within the leu-tRNA gene. The three everted nucleotides are highlighted in magenta and the stacking interactions are shown. (B) Wild type (PDB: 3MVA) structural conformation of the base-flipped region (see Supplementary Figure 1 for overall view of MTERF1 structure). The crystal structure reveals an interaction between Arg162 and the DNA backbone (2.6 Å) in addition to the interaction with T3243 (T) that seems to stabilize a critical kink in the DNA backbone at the site of base flipping (black arrow). (C) Transcription termination assays were run on a 5% UREA-PAGE gel. Full length (FL) and termination (T) bands are shown for each of the individual stacking substitutions. (D) The percentage termination by each MTERF substitution was calculated for each concentration and the values correspond to the mean +/− SEM from three independent experiments.

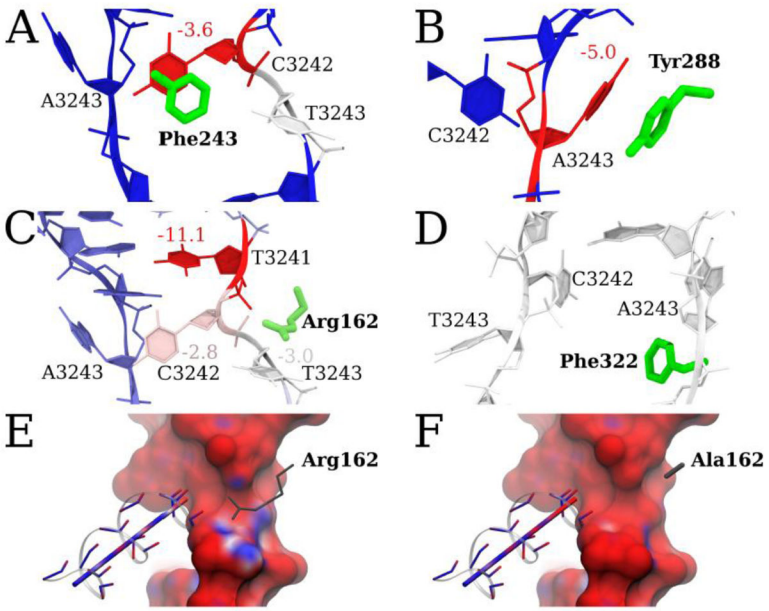


Figure 2.

Pairwise per-residue potential energy decomposition analysis showing the interaction energies between Phe243, Tyr288, Arg162 and Phe322 and the DNA. (A) Phe243 π-stacks with C3242 (−3.6 kcal/mol) as would be expected from the WT crystal structure. Unexpectedly, T3243 interacted with Phe243 (−0.7 kcal/mol averaged over all 4 simulations; −1.2 kcal/mol for one simulation in which T3243 flipped-in). (B) Y288 π-stacks with A3243 (−5.0 kcal/mol). (C) Arg162 and the HS. Arg162 interacts most strongly with T3241 (−11.1 kcal/mol) but also interacts with C3242 (−2.8 kcal/mol) and T3243 (−3.0 kcal/mol). (E) The Poisson-Boltzmann electrostatic potential energy surface of the DNA after the second stage of equilibration, using default parameters in Amber14 [35,36] and (F) the identical calculation except with Arg162 mutated to A. Panels E and F are colored from −6 kcal/e (red) to +6 kcal/e (blue). Comparison of E and F shows that Arg162 adds a positive electrostatic force on the T3243 nucleobase and the backbone of C3242. In the absence of Arg162 (R162A), the electrostatic potential of helix 2 becomes more negative near the negative DNA backbone, an unfavorable configuration that should not be stable in an R162A mutant.

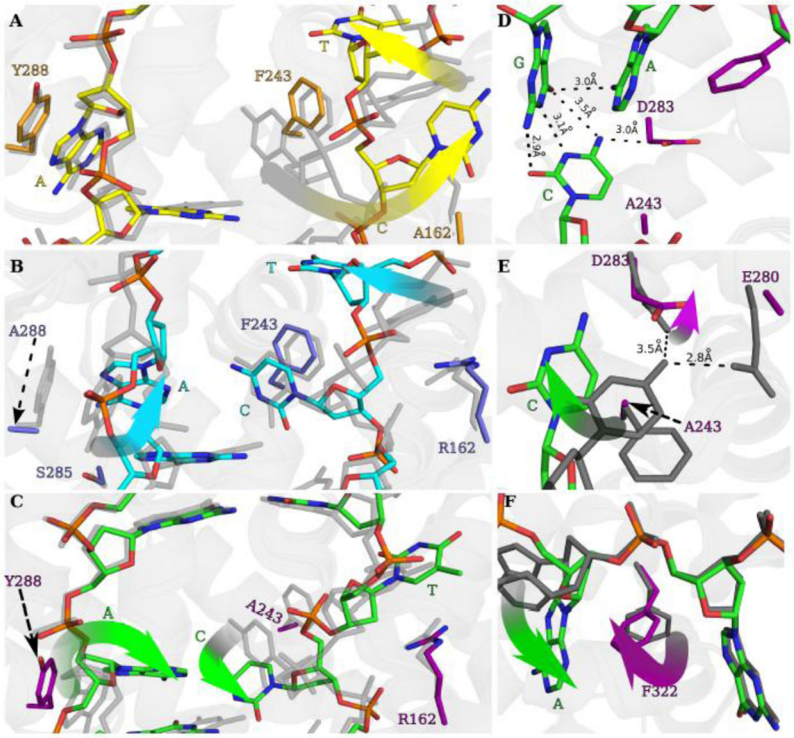


Figure 3.

X ray crystal structures of the R162A, Y288A and F243A substitutions. Panels A–C represent identical views of the base flipping region within the R162A, Y288A and F243A structures, respectively. Panels D–F represent regions within the F243A crystal structure. An overlay of wild type (gray) with (A) R162A (yellow and orange), (B) Y288A (cyan and blue) and (C) F243A (green and magenta). Alterations of the base-flipped nucleotides from their wild type conformations are highlighted with gradient arrows. (D) The F243A structure demonstrates the inability of C3243 to form a stacking interaction with Phe243 and reveals C3242 in a perturbed conformation yet, still associating with its G base pair as denoted by the hydrogen bond distances. In addition, a new contact with Asp283 and C3242 forms as the G/C base pair is being broken. (E) An overlay of wild type (gray) with the F243A structure (green and magenta) demonstrates a pathway for the movement of C3242 from its G/C base pair to its wild type conformation stacked with Phe243 and hydrogen bonded to Glu280. Note the alteration of the Asp283 side chain in the F243A structure (magenta, gradient arrow) that follows the movement of C3242, indicating its role to help break its G/C base pair. (F) An overlay of wild type (gray) with the F243A structure that demonstrates the intrahelical conformation of A3243 (green gradient arrow) that is stabilized by an altered, intrahelical conformation of Phe322 (magenta gradient arrow).

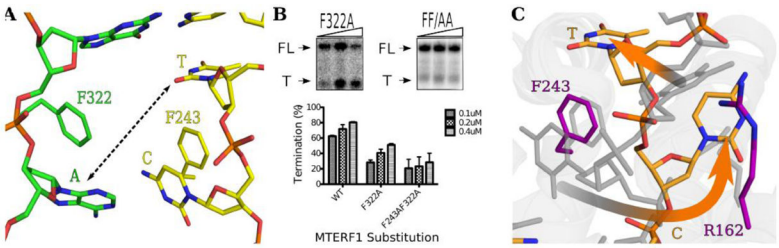


Figure 4.

(A) A composite structure of the F243A light strand DNA (green) with the Y288A heavy strand DNA (yellow) demonstrates that both Phe322 and Phe243 play a role in A/T base pair separation. (B) Termination assays for the F322A substitution and the double F243A/F322A substitutions. The bar graph values correspond the mean +/− SEM percentage termination from three independent experiments. (C) X-ray crystal structure of the F322A substitution (orange and magenta) with wild type (gray). The perturbed base flipping observed in this structure (gradient arrows) suggests Phe322 plays a role in coordinating proper flipping of the heavy strand nucleotides.

Figure 5.

Proposed model for the stepwise order of base flipping by MTERF1. Upon MTERF1 binding, helix unwinding and the intercalation of Phe243 and Phe322 into the two DNA strands, the A/T base pair is broken. C3242 (red, Step1) breaks away from its G base pair and moves towards the Phe243. Secondly, as C3242 moves, the motion is coupled to the flipping of A3243 (red, Step2) towards Tyr288. Concomitantly as C3242 moves over Phe243 and forms a π-stacking interaction, T3243 begins to break its t-stacking interaction with Phe243 and starts to move out of the helix (red arrow, Step 2). Lastly, as C3242 becomes stably oriented over Phe243, T3243 is able to fully break its interaction with Phe243 and move out of the helix and interact with Arg162 (red, Step3).

Table 1

Data-collection and refinement statistics. Values in parentheses are for highest resolution shell.

	R162A	Y288A	F243A	F322A
Data Collection				
PDB Code	5CKY	5CO0	5CRK	5CRJ
R_{merge}	0.066 (0.807)	0.054 (0.786)	0.037 (0.708)	0.042 (0.825)
R_{meas}	0.073 (0.873)	0.060 (0.840)	0.041 (0.773)	0.047 (0.89)
R_{pim}	0.037 (0.444)	0.031 (0.437)	0.021 (0.388)	0.024 (0.452)
No. of total observations	143906 (1384)	137707 (1338)	166530 (1592)	149721 (1562)
I/σI	25.4 (2.8)	27.3 (2.6)	34.1 (2.9)	29.1 (2.5)
$CC_{1/2}$	0.999 (0.823)	1.0 (0.793)	0.999 (0.851)	1.0 (0.787)
Completeness %	100 (100)	99.8 (100)	99.3 (100)	100 (100)
Multiplicity	7.3 (7.4)	7.2 (7.2)	7.3 (7.5)	7.2 (7.4)
Cell Dimensions (a,b,c)	88.22, 91.6, 159.10	87.53, 90.43, 169.50	89.25, 90.14, 161.43	89.15, 90.19, 160.85
α,β,γ	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Space Group	C2221	C2221	C2221	C2221
Resolution Å	2.62	2.59	2.48	2.59
Refinement				
R_{work}/R_{free}	0.21/0.26	0.21/0.27	0.24/0.27	0.21/0.28
R.m.s.d Bond Angles Å	1.698	1.588	1.692	1.580
R.m.s.d. Bond Lengths Å	0.012	0.012	0.013	0.011
Average B- factor	53.18	60.866	69.086	72.032