DNA Repair 8 (2009) 1321–1327

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

Brief report

# The roles of SbcCD and RNaseE in the transcription of GAA·TTC repeats in *Escherichia coli*

Xuefeng Pan[a,b,*], Yunfeng Ding[a,1], Lingfeng Shi[a,1]

[a] School of Life Science, Beijing Institute of Technology, Beijing 100081, China
[b] Institute of Cell and Molecular Biology, Edinburgh University, Edinburgh EH9 3JR, Scotland, UK

ARTICLE INFO

*Article history:*
Received 17 February 2009
Received in revised form 24 July 2009
Accepted 9 August 2009
Available online 4 September 2009

*Keywords:*
Frederich's ataxia
GAA·TTC repeats
α-Complementation SbcCD
RNaseE

ABSTRACT

Expansion of GAA·TTC repeats in *FXN* gene is associated with decreased frataxin production in Frederich's ataxia patients. To study this effect, we have engineered a set of GAA·TTC repeats in the *Eco*RI site of *lac*Z gene of plasmid pUC18 as part of the transcription template of the *lac*Z gene, while keeping its ORF unchanged. The effects of the GAA·TTC repeats on the *lac*Z expression were investigated in *Escherichia coli* JM83 and its mutants deficiency in RNA processing, homologous recombination and DNA repair. We found that transcriptions of the GAA strand with different sizes and organizations displayed normal α-complementation when *RNase* E was functional. By contrast, transcriptions of TTC repeats containing more than 13 triplets failed to support α-complementation, showing RNase-independent but length-dependent effects of TTC repeats on *lac*Z expression. In addition, we also found that functions of SbcCD, a DNA structure specific nuclease, were needed in the RNase E-dependent *lac*Z expression of the GAA repeats. These suggested that processing of DNA and RNA is essential to the transcription of the repeats-carrying gene *in vivo*.



## 1. Introduction

Friedreich's ataxia (FRDA) is one of the most common recessive genetic diseases with the symptoms of muscle weakness, skeletal abnormalities, limb ataxia, progressive gait, vision impairment, heart disease and diabetes. It can be caused by a decreased expression of *FXN* gene (named as gene X25 previously) due to the expansion of GAA·TTC repeats in the first intron of the *FXN* gene [1,2]. *FXN* encodes an essential protein-frataxin responsible for iron metabolism in mitochondria [3–8]. Normal frataxin alleles have 5–34 units whereas the FRDA alleles may have 66–1700 repeats units. It is well documented that the expanded GAA·TTC repeating sequences tend to form non-B DNA structures, e.g. triplexes [9–12], sticky DNA [10,13], hairpins [14,15] and parallel DNA [16], during DNA replication, transcription, recombination, and repair [17–21]. These non-B secondary structures can also influence the DNA metabolic processes, such as stalling of DNA replication forks *in vivo* and *in vitro* [11,12,22–27].

Long $(GAA)_n$ template has shown to reduce the transcription of *FXN* gene more vigorously *in vitro*, in comparison to the long $(TTC)_n$ [26]. The $(GAA)_n$ repeats corresponding to the native orientation of the GAA sequence in *FXN* gene can disrupt the RNA transcription more profoundly [4,25–27]. One example is the pausing of T7 RNA polymerase in RNA transcription by a longer $(GAA)_n$ template *in vitro*, during which large amounts of full-length transcripts rather than the prematurely terminated transcripts were reduced [27]. Recently, impairment on RNA polymerase progression by GAA repeats as the sense strand in transcription, and truncation of RNA species by TTC repeats as the sense strand, were also reported *in vivo* [28,29]. Interestingly, the impeded RNA transcripts by the $(TTC)_n$ template were found to form RNA·DNA hybrid, while the truncated RNA transcripts by $(GAA)_n$ template degraded via the cleavage of $(UUC)_n$ repeats by an *Escherichia coli* degradosome [28,29].

However, mechanism of transcriptional repression of *FXN* gene in FDRA patients remains a controversy. Inhibition of transcription by forming triplex structures by GAA·TTC repeats is not unique. Similar inhibitory effects have also been observed in a transient transfection of the repeats into COS-7 *in vivo* [25], nuclear protein(s) binding upon the repeats (as noted in HeLa nuclear extracts) [10–12,25], altered RNA splicing and processing [27,30]; and $(GAA)_n$ induced heterochromatin formation in the peripheral region around *FXN* gene that silences the frataxin gene in a reversible fashion [31–34].

To help understanding of how GAA·TTC repeats affect the expression of *FXN* gene in FRDA, we carried out a systematic evaluation on the effects of different GAA·TTC repeats in pUC18 plasmid on its α-peptide gene transcription and α-complementation in

* Corresponding author at: School of Life Science, Beijing Institute of Technology, Beijing 100081, China. Tel.: +86 10 68914495x805.
*E-mail address:* xuefengpancam@yahoo.com.cn (X. Pan).
[1] YFD and LFS, contributed to this work equally.

1568-7864/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2009.08.001

*E. coli* JM83, and its mutants lacking of the conserved functions of RNA processing, DNA homologous recombination and mismatch repair. Toward this, we have engineered a set of GAA·TTC repeats in the *Eco*RI site of *lac*Z gene of plasmid pUC18, where the open reading frames of the α-peptide gene (the N-terminal of *lacZ* gene) retained. The effects of differently organized GAA·TTC repeats on the α-peptide gene expression, including the transcription and translation, were monitored by α-complementation analysis, and molecular cloning, RT-PCR and gene knockout etc. We found that the processing of UUC repeats in mRNA transcripts by RNaseE, and the avoidance of non-B secondary structure formation by SbcCD, an equivalent of Mre11-Rad50 in eukaryotes, were both critical for the transcription of GAA repeats *in vivo*.

## 2. Materials and methods

### 2.1. Bacteria and medium

*E. coli* strains used in this work were JM83 and its derivatives, including DL324 (JM83 wild-type), DL887*recA::Cm*$^R$, DL941 *mutS::miniTn10*, DL733*sbcCD::Kan*$^R$ (a gift from D Leach in Edinburgh University, UK); strain HAT103 carrying an *rne*$^{ts}$ mutation (a gift from Dr. Wachi in Tokyo Institute of Technology, Japan). The following strains: XP1100, JM83*rne*$^{ts}$ XP1101, JM83 *sbcCD::Kan*$^R$ *rne*$^{ts}$, XP 1184, *mutS::miniTn10sbcCD::Kan*$^R$ were constructed in this work by using P1 transduction; LB broth and LB agar were used as media for bacteria cultivation [36].

### 2.2. Plasmid construction and DNA sequencing

Constructions of differently organized GAA·TTC trinucleotide repeats in plasmid pUC18 were performed as follows: different oligonucleotides of AATTC$(GAA)_{24}$, AATTC $(GAA)_{36}$, AATTC $(TTC)_{24}$, and AATTC $(TTC)_{36}$ all flanked with an *Eco*RI restriction site were synthesized by Oswell company (UK). For repeats constructions, equal molar of AATTC $(GAA)_{24}$/AATTC $(TTC)_{24}$ and AATTC $(GAA)_{36}$/AATTC $(TTC)_{36}$ oligonucleotides was mixed in an annealing buffer containing 200 mM Tris-Cl pH 7.9, 500 mM NaCl, 250 mM $MgCl_2$. The mixture was initially incubated at 90 °C for 5 min and cooled down to the room temperature by 60 min. The annealed oligonucleotides were phosphorylated by using polynucleotide kinase in the presence of ATP molecules before used to build up the tandem linked repeats. Plasmids pGAA24, pTTC24, pGAA37, and pTTC37 were constructed by making the GAA or the TTC repeats as the lagging strand template of replication fork. These plasmid constructs were confirmed by DNA sequencing, and the DNA sequencing was carried out by using a thermal cycling kit or by using the sequencing services of Huada Gene Co. Ltd. (Beijing). Plasmid p$(GAA)_{50}$ was obtained through propagation of a p$GAA_{37}$ plasmid in a *sbc*CD mutant; plasmid p$TTC_{50}$ was constructed by reversing the GAA strand of the p$(GAA)_{50}$ into the opposite orientation, during which a *mut*S mutant was used as the host.

### 2.3. Plasmid DNA transformation and recovery

The above-mentioned plasmids were transformed the indicator strains, either the wild-type or mutants carrying different gene mutations by a $CaCl_2$ method [36]. Propagations of plasmid – carrying indicator strains were carried out by using LB broth containing 100 μg/mL ampicillin. Plasmid DNA recovered after a period of propagation by Qagene-plasmid miniprep kit and was further analyzed on the repeats' instabilities.

### 2.4. LacZ complementation assay

LacZ complementation assay was performed by using a method described elsewhere in the Ref. [36]. LB plates containing 40 mM IPTG and 20 mM X-gal were utilized in the α-complementation analysis [36].

### 2.5. β-Galactosidase activity assay

β-Galactosidase activity of strains carrying different plasmid constructs when cultivated in liquid LB cultures was measured by using a method developed by Miller [37].

### 2.6. RT-PCR analysis on TTC repeats transcripts

RT-PCR was conducted by using a Qiagen One-step RT-PCR kit. In brief, the white colonies formed by DL324 (JM83) harboring plasmid p$GAA_{37}$ and the blue colonies with pUC 18 plasmids (control) were collected, respectively, and total mRNA was isolated by using a RNA extraction kit (purchased from Promega UK). Purity of the total RNA was confirmed by using a pilot PCR amplification with DNA primers as follows: lac-F: 5′AACGTCGTGACTGGGAAAAC, lac-R: 5′GTGAAATACCGCACAGATGC; Amp-F: 5′GCGGCCAACTTACTTCTGAC; Amp-R: 5′AACTTTATCCGCCTCCATCC; PUC-F: 5′AGTCGTGTCTTACCGGGTTG and PUC-R: 5′GACGAGCATCACAAAAATCG (synthesized by MWG Oligo Synthesis Service). Molecular weights of the RT-PCR products were expected to be 186 bp with lac-F and lac-R primers, 249 bp with Amp-F and Amp-R, and 313 bp with PUC-F and PUC-R primers, respectively.

### 2.7. Analysis of repeat expansion and contraction instability

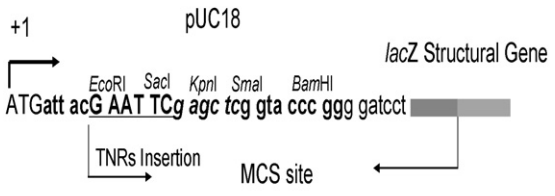
Repeat expansion and contraction instabilities associated with the propagation of GAA·TTC repeats were analyzed either by directly sequencing the repeats DNA or by comparing the repeats variations of the *Eco*RI DNA fragments on agarose gel or on PAGE after being isolated from a small population of cells as described in our previous work [18].

## 3. Results

### 3.1. Experimental rationale

α-Complementation assay is a useful technique in gene cloning, based upon the complementary capacities of the two fragmented β-galactosidase genes in the *E. coli* genome and in the plasmid pUC18. When the two-fragmented gene products were associated into a functional β-galactosidase *in vivo*, blue colonies form in the presence of X-gal, otherwise white colonies form [35]. In this work, we have utilized this capacity to investigate the effects of transcriptions of differently organized GAA·TTC repeats on α-peptide gene expression in various JM83 *E. coli* strains. 13 links of GAA·TTC repeat, differing in sizes and organizations, were cloned in the *EcoRI* site of pUC18 plasmid, respectively, while retaining the open reading frame of the α-peptide gene (Fig. 1A). Amongst them, p$GAA_{24}$, p$GAA_{37}$, p$TTC_{24}$, p$TTC_{37}$ were constructed as either GAA or TTC strand to serve as the template for RNA transcription (Fig. 1A). Plasmids p$(GAA)_{24}$TTC$(GAA)_{24,}$ p$(GAA)_{48}$, p$(TTC)_{48}$, p$(TTC)_{24}$GAA$(TTC)_{24,}$ and (p3$(GAA)_{24}$) were constructed by cloning the tandemly linked $GAA_{24}$ or $TTC_{24}$ oligonucleotides at the *Eco*RI site of pUC18. Plasmid p$GAA_{50}$, containing 50 triplets of GAA repeating units, was obtained by propagation of a p$GAA_{37}$ plasmid in a JM83 *sbc*CD mutant (DL733, From Leach), and p$TTC_{50}$ plasmid was constructed by reversing the $GAA_{50}$ repeats of the p$GAA_{50}$ into its opposite orientation. By transforming the plasmid constructs into *E. coli* JM83 cells, we conducted a systematic analysis on the

(A)




(B)

| Trinucleotide Repeats | LacZ complementation | Source |
|---|---|---|
| $(TTC)_6(TTT)(TTC)_5$ | + | Deletion of $(TTC)_{36}$ |
| $(GAA)_{24}$ | + | |
| $(TTC)_{24}$ | - | |
| $(GAA)_{36}$ | + | |
| $(TTC)_{36}$ | - | |
| $(TTC)_{14}$-(TTT)-$(TTC)_{21}$ | - | From 2X$(TTC)_{24}$ |
| $(TTC)_{24}$-(GAA)-$(TTC)_{24}$ | - | 2X$(TTC)_{24}$ |
| $(GAA)_{24}$-(TTC)-$(GAA)_{24}$ | + | 2X$(GAA)_{24}$ |
| $(GAA)_{24}$-(TTC)-$(GAA)_{24}$-(TTC)-$(GAA)_{24}$ | + | 3X$(GAA)_{24}$ |
| $(GAA)_{21}$(AAA)$(GAA)_{29}$ | + | Expansion of $(GAA)_{36}$ |
| $(TTC)_{21}$-(TTT)-$(TTC)_{29}$ | - | Expansion of $(GAA)_{36}$ |
| $(CAG)_{28}$ | + | |
| $(CTG)_{28}$ | + | |

**Fig. 1.** α-Complementation assay with GAA·TTC trinucleotide repeats containing plasmid in JM83 strain. (A) arrangements of GAA·TTC repeats in the plasmid pUC18; and (B). the effects of differently organized repeats on α-peptide expression as tested on LB plates containing IPTG and X-Gal. Formation of blue colored colonies by the plasmids containing (GAA)n, $(TTC)_{13}$ and $(CTG)_{28}$, $(CAG)_{28}$ repeats were denoted "+", and white colored colony-forming plasmids by $(TCC)_{>13}$ repeats were denoted "−".

effects of differently constructed repeats in plasmid pUC18 on their α-complementation capacities in wild-type JM83 and its derivative stains carrying mutations on RNA processing, recombination and repair (Fig. 1). Formation of blue or white colonies through α-complementation analysis on LB plates containing IPTG and X-gal was recognized as functions of effects of repeats array on α-peptide expression *in vivo*. Alternatively, β-galactosidase activity was measured when they were cultivated in liquid LB medium. The results of these analyses were listed in Fig. 1B.

### 3.2. Effects of transcriptions of GAA·TTC repeats on LacZ α-complementation

Several investigations have reported that insertion of certain DNA sequences in the multiple cloning site (MCS) of α-peptide gene of pUC18 can retain the α-complementation capability of *lac*Z gene when the open reading frame of the *lac*Z gene fragment was also retained in *E. coli* JM83 etc [38]. Indeed, a similar phenomenon was noted in this work when the GAA repeats but not the TTC repeats were inserted in the *Eco*RI site in pUC18, of which the open reading frame of the α-peptide gene was unaffected, revealing an orientation dependent α-complementation property of GAA and TTC repeats in JM83 grown in LB plates containing IPTG and X-gal (Fig. 1). Transcription of GAA repeats in *cis* to the (-peptide gene template, which was supposed to yield UUC repeats-containing LacZ mRNA, showed normal α-complementation by $pGAA_{24}$,$pGAA_{37}$, $pGAA_{51}$, $p(GAA)_{48}$, $p(GAA)_{72}$ in wild-type JM83 strain while transcription of plasmid $pTTC_{24}$, $pTTC_{37}$, $pTTT_{37}$, $pTTC_{48}$, and $pTTC_{51}$ forming white colonies (Fig. 1B). Showing orientation dependent effects of transcriptions of differently cloned GAA·TTC repeat on the LacZ α-complementation in JM83.

In addition to the formation of the white colonies, a small portion of blue colonies could also be yielded by the TTC repeats transcription in JM83 wild-type. DNA plasmids carried by these blue colonies were further characterized by DNA sequencing, and were shown only 8–13 triplets of GAA·TTC repeats left in the *Eco*RI site (data not shown), indicating that TTC repeats containing less than 13 triplets could yield positive LacZ α-complementation in *E. coli* JM83 (Fig. 2).

### 3.3. RNA transcription of TTC repeats containing more than 13 triplets

The unable to show α-complementation by the TTC repeats longer than 13 triplets may imply as possibilities of either the longer TTC repeats blocking RNA transcription, or interfering with the subsequent translation of the α-peptide mRNA [26,27]. These possibilities were tested by using a $TTC_{37}$ plasmid as a representative for RT-PCR examinations for the length of the RNA molecules transcribed *in vivo* from the white colonies (Data not shown). To do these, we synthesized and utilized a set of DNA primers for the RT-PCR tests (see Section 2, and Fig. 3). Results of the RT-PCR amplifications were presented in Fig. 3. It was seen in Fig. 3 that RNA transcriptions using pUC18 and $pTTC_{37}$ *in vivo* did not make any significant difference in PCR products, suggesting that failure of showing α-complementation by TTC repeats *in vivo* was not due to the impediments of the repeat array on its transcription, instead the effects of TTC transcripts on RNA translation were suggested.

### 3.4. Functional RNaseE is required by the normal α-complementation by the UUC repeats-containing mRNA

It can be seen in Fig. 1B, transcription of different links of GAA repeats in wild-type JM83 did not affect *lacZ* α-complementation, nor did the transcription of TTC repeats containing less than 13 triplets in the same situation (Fig. 2). To understand the reasons underlying these observations, we have further analyzed the repeats' processing of the mRNA molecules transcribed in an *rne*$^{ts}$ mutant (a temperature sensitive *rne* mutant). It was known that *rne* gene of *E. coli* encodes RNaseE, a core component of RNA processing complex that cleaves RNA molecules at the AU-rich region. The *rne*$^{ts}$ mutant we were using can grow below a temperature of 30 °C, but it cannot grow well when the temperature was beyond 42 °C, under which the *rne*$^{ts}$ gene was supposed to be turned off by the high temperature. To this end, plasmids pUC18 and $pGAA_{48}$ were transformed into JM83 *rne*$^{ts}$ mutant, respectively, and the strains carrying each of the plasmids were grown initially at 30 °C to form colonies of visible size, and then, they were incubated at 30 °C and 42 °C, separately. We found that JM83*rne*$^{ts}$ mutant strains carrying plasmid pUC18 formed blue colonies either at 30 °C or at 42 °C. In contrast, plasmid $pGAA_{48}$ carried in JM83*rne*$^{ts}$ mutant formed blue colonies at 30 °C, but will form white colored colonies when at 42 °C (data not shown). These clearly show that UUC-repeats in *lacZ* mRNA molecules were responsible for the failure in the LacZ α-complementation, and, which have to be cleaved by RNase E.

### 3.5. SbcCD is partially needed by making positive α-complementation in GAA repeats transcription in the presence of RNaseE

In addition to see the roles of RNaseE in α-complementation in GAA repeats transcription, we also observed that ~23% of colonies could appear to be white colonies when $pGAA_{48}$ plasmid was in *sbc*CD alone mutant in the presence of IPTG and X-gal, indicating that SbcCD was also required in the transcription of UUC repeats containing *lacZ* mRNA, regardless of the function of RNaseE.

The roles of SbcCD on LacZ α-complementation had been further investigated by comparing the α-complementation of plasmid $pGAA_{48}$ in JM83*mutS*, JM83*recA* and JM83*mutS, sbc*CD double mutants. It was found that white colonies could only be formed by *sbc*CD alone and *mutSsbc*CD double mutants, but could not be formed either by *mut*S alone, or the *rec*A alone mutant. Indicating

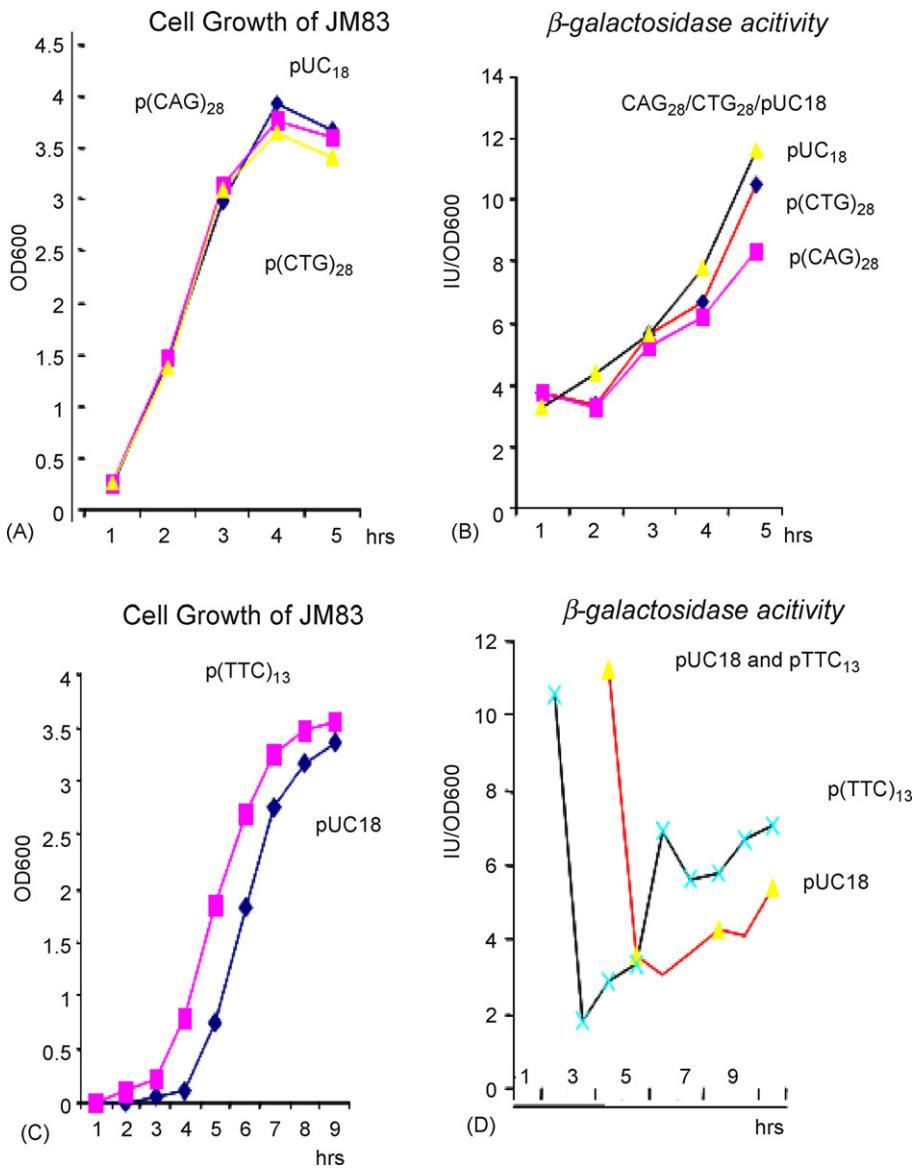



**Fig. 2.** β-Galactosidase activity associated with the expression of plasmids pUC18 containing $(CTG)_{28}$, $(CAG)_{28}$, $(TTC)_{13}$ repeats and pUC18. (A) Growths of JM83 cells carrying $p(CTG)_{28}$, $p(CAG)_{28}$, and pUC18 plasmids, respectively, which correspond to (B) the **β**-galactosidase activities of JM83 carrying $p(CTG)_{28}$, $p(CAG)_{28}$ and pUC18; (C) growths of JM83 cells carrying $p(TTC)_{13}$ and pUC18 plasmids, which correspond to (D) the **β**-galactosidase activities of JM83 carrying $p(TTC)_{13}$, and pUC18, respectively. An increased expression of **β**-galactosidase was observed in JM83 carrying $p(TTC)_{13}$.

that the role of SbcCD in GAA48 repeats transcription was neither the gene mutation prevention as associated with *mut*S mutant, nor the homologous recombination as with the *rec*A gene mutant (Data not shown). SbcCD was characterized as a structural specific DNA nuclease that recognizes DNA hairpin structure *in vitro* and *in vivo* [18,39,40]. The requirements of SbcCD in $GAA_{48}$ repeats transcription may therefore hint at the possibility of the processing of non-B DNA repeat structure, which in this case may interfere with RNA transcription *in vivo* [28].

### 3.6. *SbcCD functions before RNaseE in the transcription of GAA repeats*

To understand the possible relationships between the roles of SbcCD and RNaseE in α-peptide mRNA transcription and/or in α-complementation, we have constructed a JM83*sbcCDrne*$^{ts}$ double mutant, and conducted the α-complementation analysis by using this double mutant. We found that colonies of JM83*sbcCDrne*$^{ts}$ double mutants carrying $pGAA_{48}$ were all white colored when grown at 30 °C and incubated at 42 °C, and a substantial number of white colonies turned into blue colored once the *rne*$^{ts}$ gene was reactivated by incubation at 30 °C over a period of time (Data not shown). Among them, a small number of white colonies were still remained, making distinctions to the situation of forming all blue colonies as seen in *rne*$^{ts}$ alone mutant (data not shown). These suggested that the SbcCD epistatically functioned before the functions of the RNaseE to cleave the DNA structure formed in the transcription of the GAA repeats (Fig. 4). These observations made consistence to the observation that *sbc*CD alone mutant can yield white colonies by ~23% in the presence of functional RNaseE.

### 3.7. *Disassociation of formation of non-B secondary structure in transcription from the instabilities of repeat expansion and contraction*

To see if the formation of white colonies in *sbc*CD mutants could lead to the expansion and contraction instabilities associ-

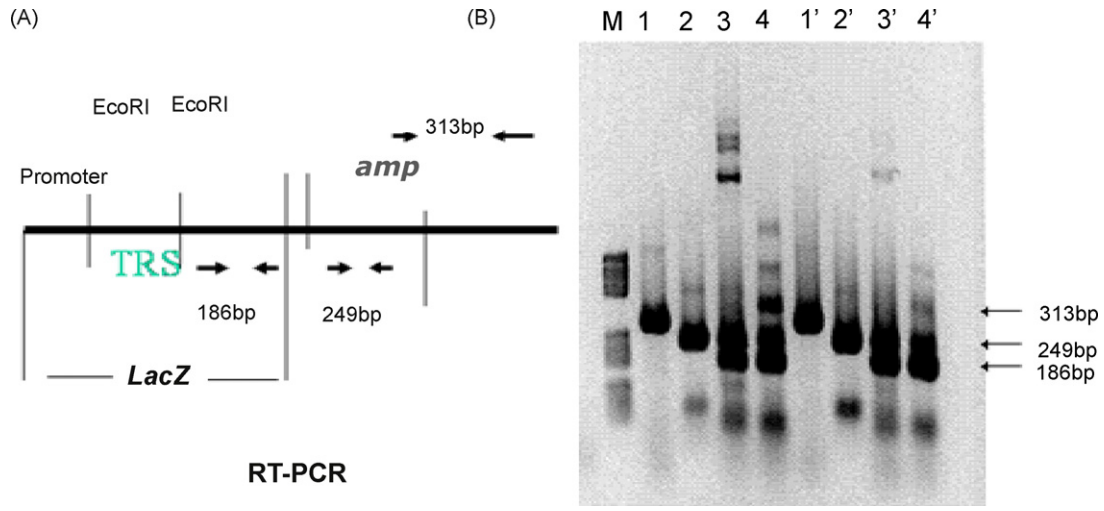



**Fig. 3.** RT-PCR examination of the RNA transcripts of $(TCC)_{37}$ repeats in JM83. (A) Schematic illustration of the RT-PCR examination; (B) RT-PCR products obtained with JM83 strains carrying $p(TTC)_{37}$ when using primers of PUC-F and PUC-R (lanes 1 and 1′), Amp-F/Amp-R(Lanes 2 and 2′), lac-F/lac-R/Amp-F/Amp-R (lanes 3 and 3′), and lac-F/lac-R (lanes 4 and 4′). Lane M was the DNA molecular weight marker. Total RNA isolated from white clones of JM83 strains carrying $p(GAA)_{37}$ and blue clones of JM83 strains carrying pUC18 plasmids have been utilized for RT-PCR, and was tested for any plasmid DNA contamination by using PUC-F/PUC-R primers in a PCR amplification with the total RNA samples.

ated with the GAA repeats, the links of the GAA repeats in the white colonies were examined by sequencing the repeat arrays in plasmid DNA or comparing alterations of the repeat sizes on an agarose or a PAGE gel [18]. In this regard, 36 individual plasmid DNA preparations were examined, and no any significant variations either in base or in repeats' sizes, except a small portion of contracting and inserting instabilities of 1 triplet gain or loss were noted, which can be explained simply by a mechanism of slippage replication of the repeats DNA (data not shown). Therefore, we argue a disassociation of the non-B secondary structure formed in the transcription of GAA repeats from the instability of the repeat expansion.

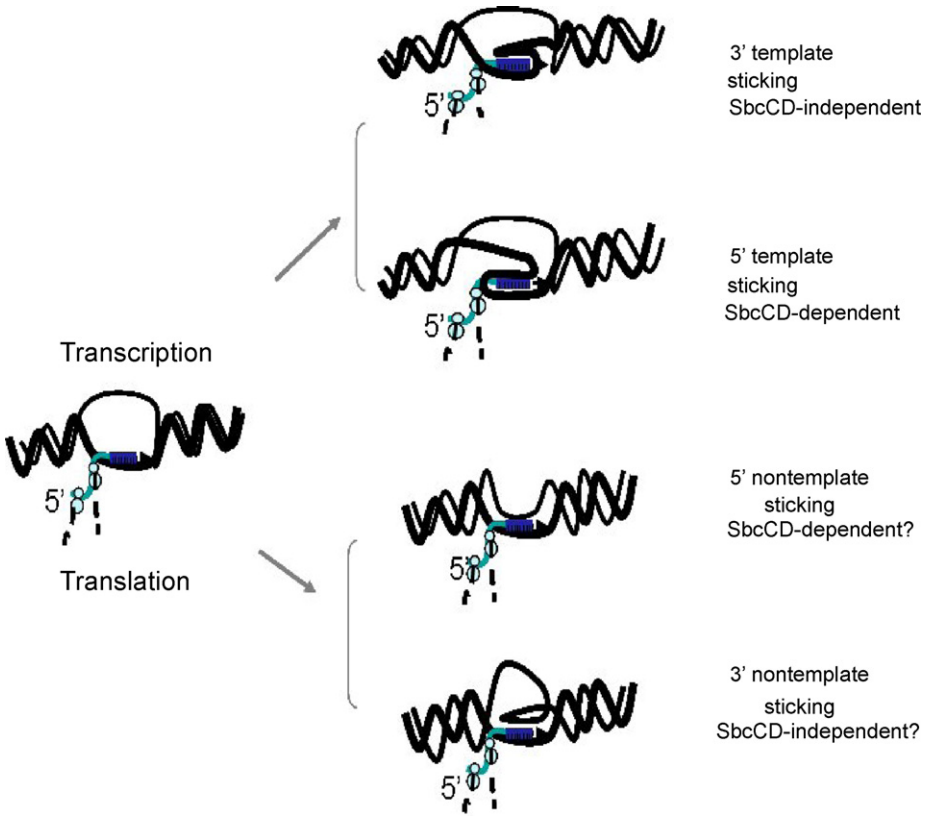



**Fig. 4.** Model of the formation of non-B repeats secondary structures in (GAA)n repeats transcription *in vivo*. In the system, RNA transcription through GAA repeats produced UUC repeats-containing mRNA, which in turn forms RNA·RNA·DNA or RNA·DNA·DNA triplex structures with the repeats DNA template by using either the 5′ side or the 3′ side of the mRNA molecules. Our results argue that a type of RNA·DNA·DNA triplex structure formed by the UUC repeats at 5′ terminus of the LacZ mRNA, and the repeats DNA template was suggested to fold back. This type of structure in RNA transcription does not stall the ongoing RNA polymerase, but may affect the proceeding α-peptide translation sequentially.

## 4. Discussion

In this work, we demonstrated that, using different lengths of the GAA·TTC repeats, the effects of transcriptions of GAA template on LacZ α-complementation depended fully on the functions of RNaseE and partially on the functions of SbcCD. RNaseE is a core component of RNA processing degradosome in *E. coli* that cleaves RNA molecules in AU-rich region, while SbcCD is a DNA structure-specific nuclease that recognizes DNA hairpin structures and double stranded DNA ends [18,39,40]. Our findings suggested that transcription of GAA repeats may lead to the formation of some non-B DNA secondary structure, consisting of the repeat DNA template and the RNA transcripts (Fig. 4) [10,13]. As it can be seen in Fig. 4, RNA transcription through GAA repeats could potentially produce either RNA·RNA·DNA or RNA·DNA·DNA triplex structure composed of the DNA repeat template and the UUC repeats of the transcribed mRNA molecule at its 5′ side. The involvement of SbcCD in the processing of the non-B secondary structure may suggest the formation of RNA·DNA·DNA triplex. This type of three-stranded DNA secondary structure can affect α-complementation, unless both SbcCD and RNaseE are functional.

However, transcription of TTC strands of the GAA·TTC repeats affected the *lac*Z gene α-complementation in a length-dependent but an RNaseE-independent manner. White colonies formed with the expression of TTC repeats containing more than 13 triplets, while blue colonies formed when the length of TTC repeats is less than 13 triplets. This length-dependent effect on α-complementation by TTC repeats was found to be irrelevant to the transcription of TTC strand. As analyzed by RT-PCR, full length of *lac*Z mRNA was transcribed. These put together reveal a significant difference of the transcriptions of two strands of the GAA·TTC repeats on affecting the *lac*Z gene expression *in vivo*.

The formation of hairpin structures with G·A mismatched base pairs by GAA strand, and parallel duplex by GAA and TTC repeats was reported *in vitro* [14–16]. However, this is different from our study *in vivo*. As shown in Fig. 2, the $CAG_{28}$ and $CTG_{28}$ repeats formed only hairpin structures *in vivo*, but their transcriptions did not affect the α-complementation significantly in our system.

It was suggested that formation of RNA·DNA·DNA triplex structure in GAA repeats transcription may be responsible for the instabilities of GAA repeats. This was examined here by using a *sbc*CD mutant. The transcription of the GAA repeats did not show dramatic repeats instabilities in the white colonies, implicating a bipartite influence of non-B secondary structure formation in GAA repeats transcription and expansion [7,11,12,19,20,41].

## 5. Conclusion

The effects of the transcriptions of a set of GAA·TTC repeats in plasmid pUC18 on *lac*Z gene expression were tested in *E. coli* JM83. We found that transcriptions of GAA strand in *lac*Z gene did not affect α-complementation in the presence of RNaseE and/or SbcCD, while transcription of the TTC strand affected α-complementation in a length-dependent manner. Our results suggested the formation of triplex structure comprising of UUC repeats in mRNA and GAA repeats template in the transcription of GAA strand of the GAA·TTC repeats, during which RNA processing by RNaseE and DNA processing by SbcCD are essential. However, the formation of triplex structure did not cause drastic expansion instabilities of the GAA repeats.

## Conflict of interest statement

The authors declare that there are no conflicts of interest.

## Acknowledgements

This work was supported by grants from Beijing Institute of Technology Natural Science Foundation of China to XP, and was once supported by a grant from MRC (UK). We are very grateful to Professor B. Michel at Institut National de la Recherche Agronomique, France for strains and Professor Masaaki Wachi at Tokyo Institute of Technology Japan for *rne*$^{ts}$ mutant. Thanks must also go to Professor D leach at University of Edinburgh for both his strains and his constructive comments; and also Dr Zhilian Yue at Institute of Bioengineering and Nanotechnology, Singapore for her thorough reading of the manuscript.

## References

[1] V.L. Campuzano, M.D. Montermini, L. Molto, M. Pianese, F. Cossee, E. Cavalcanti, F. Monros, F.R. Duclos, Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion, Science (1996) 1423–1427, 7195254.
[2] M. Pandolfo, Molecular genetics and pathogenesis of Friedreich ataxia, Neuromuscul. Disord. 8 (1998) 409–415.
[3] J. Kaplan, Friedreich's ataxia is a mitochondrial disorder, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 10948–10949.
[4] L. Montermini, E. Andermann, A. Richter, M. Pandolfo, F. Cavalcanti, L. Pianese, L. Iodice, G. Farina, A. Monticelli, M. Turano, A. Filla, G. De. Michele, S. Cocozza, The Friedreich ataxia GAA triplet repeat: premutation and normal alleles, Hum. Mol. Genet. 6 (1997) 1261–1266.
[5] M. Cossee, M. Schmitt, V. Campuzano, L. Reutenauer, C. Moutou, Evolution of the Friedreich's ataxia trinucleotide repeats expansion: Founder effect and permutations, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 7452–7457.
[6] P.I. Patel, G. Isaya, Friedreich ataxia: from GAA triplet-repeat expansion to frataxin deficiency, Am. J. Hum. Genet. 69 (2001) 15–24.
[7] S.M. Mirkin, Expandable DNA repeats and human disease, Nature 447 (2007) 932–940.
[8] A.M. Gacy, G.M. Goellner, C. Spiro, X. Chen, G. Gupta, E.M. Bradbury, R.B. Dyer, M.J. Mikesell, J.Z. Yao, A.J. Johnson, A. Richter, S.B. Melancon, C.T. McMurray, GAA instability in Friedreich's ataxia shares a common, DNA directed and intraallelic mechanism with other trinucleotide diseases, Mol. Cell 1 (1998) 583–593.
[9] S.V. Mariappan, P. Catasti, L.A. Silks 3rd, E.M. Bradbury, G. Gupta, The high-resolution structure of the triplex formed by the GAA·TTC triplet repeat associated with Friedreich's ataxia, J. Mol. Biol. 285 (1999) 2035–2052.
[10] N. Sakamoto, P.D. Chastain, P. Parniewski, K. Ohshima, M. Pandolof, J.D. Griffith, R.D. Wells, Sticky DNA:Self-association properties of long GAA·TTC repeats in R.R.Y triplex structures from Friedreich's ataxia, Mol. Cell 3 (1999) 465–475.
[11] R.D. Wells, Non-B conformations, mutagenesis, and diseases, Trends Biochem. Sci. 32 (2007) 271–278.
[12] R.D. Wells, DNA triplexes and Friedreich ataxia, FASEB J. 22 (2008) 1625–1634.
[13] L.S. Son, A. Bacolla, R.D. Wells, Sticky DNA: *in vivo* formation in *E. coli* and *in vitro* association of long GAA·TTC tracts to generate two independent supercoiled domains, J. Mol. Biol. 360 (2006) 267–284.
[14] I.S. Suen, J.N. Rhodes, M. Christy, B. McEwen, D.M. Gray, M. Mitas, Structural properties of Friedreich's ataxia d(GAA) repeats, Biochim. Biophys. Acta. 1444 (1999) 14–24.
[15] B.L. Heidenfelder, A.M. Makhov, M.D. Topal, Hairpin formation in Friedreich's ataxia triplet repeat expansion, J. Biol. Chem. 278 (2003) 2425–2431.
[16] E.M. LeProust, C.E. Pearson, R.R. Sinden, X. Gao, Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia, J. Mol. Biol. 302 (2000) 1063–1080.
[17] H. Moore, P.W. Greenwell, C.-P. Liu, N. Arnheim, T. Petes, Triplet repeats form secondary structures that escape DNA repair in yeast, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 1504–1509.
[18] X. Pan, D.R.F. Leach, The roles of *mutS, sbcCD* and *recA* in the propagation TGG repeats in *E. coli*, Nucleic Acids Res. 28 (2000) 3178–3184.
[19] X. Pan, Self-maintenance of Gene and Occurrence of Genetic Diseases, Science Press, Beijing, 2004, pp.1–403.
[20] X. Pan, Mechanism of trinucleotide repeats instabilities: the necessities of repeat non-b secondary structure formation and the roles of cellular transacting factors, J. Genet. Genomics 33 (2006) 1–11.
[21] J.F. Qiu, X. Pan, The linkage of DNA replication, recombination and repair, Prog. Biochem. Biophys. 35 (2008) 751–756.
[22] G.M. Samadashwily, G. Raca, S.M. Mirkin, Trinucleotide repeats affect DNA replication *in vivo*, Nat. Genet. 17 (1997) 298–304.
[23] M.M. Krasilnikova, S.M. Mirkin, Replication stalling at Friedreich's ataxia (GAA)n repeats *in vivo*, Mol. Cell Biol. 24 (2004) 2286–2295.
[24] S.M. Mirkin, DNA structures, repeat expansions and human hereditary disorders, Curr. Opin. Struct. Biol. 3 (2006) 351–358.
[25] K. Ohshima, L. Montermini, R.D. Wells, M. Pandolfo, Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication *in vivo*, J. Biol. Chem. 273 (1998) 14588–14595.

[26] S.I. Bidichandani, T. Ashizawa, P.I. Patel, The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure, Am. J. Hum. Genet. 62 (1998) 111–121.
[27] E. Grabczyk, K. Usdin, The GAA·TTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner, Nucleic Acids Res. 28 (2000) 2815–2822.
[28] M.M. Krasilnikova, M.L. Kireeva, V. Petrovic, N. Knijnikova, M. Kashlev, S.M. Mirkin, Effects of Friedreich's ataxia (GAA)n·(TTC)n repeats on RNA synthesis and stability, Nucleic Acids Res. 35 (2007) 1075–1084.
[29] E. Grabczyk, M. Mancuso, M.C. Sammarco, A persistent RNA·DNA hybrid formed by transcription of the Friedreich's ataxia repeat in live bacteria, and by T7 RNAP *in vitro*, Nucleic Acids Res. 35 (2007) 5351–5359.
[30] M. Baralle, T.E. Pastor, B.F. Pagani, Influence of Friedreich ataxia GAA noncoding repeat expansions on pre-mRNA processing, Am. J. Hum. Genet. 83 (2008) 77–88.
[31] D. Herman, K. Jenssen, R. Burnett, E. Soragni, S. Perlman, J.M. Gottesfeld, Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia, Nat. Chem. Biol. 2 (2006) 551–558.
[32] E. Greene, L. Mahishi, A. Entezam, D. Kumari, K. Usdin, Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia, Nucleic Acids Res. 35 (2007) 3383–3390.
[33] E. Soragni, D. Herman, S.Y. Dent, J.M. Gottesfeld, R.D. Wells, M. Napierala, Long intronic GAA·TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia, Nucleic Acids Res. 36 (2008) 6056–6065.
[34] S. Al-Mahdawi, R.M. Pinto, O. Ismail, D. Varshney, S. Lymperi, C. Sandi, D. Trabzuni, M. Pook, The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues, Hum. Mol. Genet. 17 (2008) 735–746.
[35] C. Yanisch-Perron, J. Vieira, J. Messing, Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, Gene 33 (1985) 103–119.
[36] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
[37] J.H. Miller, Experiments in Molecular Genetics, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1972, pp.352–355.
[38] T.J. Close, J.L. Christmann, R.L. Rodriguez, M13 bacteriophage and pUC plasmids containing DNA inserts but still capable of beta-galactosidase alpha-complementation, Gene 23 (1983) 131–136.
[39] J. Connelly, D.R.F. Leach, The *sbc*C and *sbc*D genes of *Escherichia coli* encode a nuclease involved in palindromic inviability and genetic recombination, Genes Cell 1 (1996) 285–291.
[40] J. Connelly, E.S. delau, D.R.F. Leach, DNA cleavge and degradation by the SbcCD complex from *Escherichia coli*, Nucleic Acid Res. 27 (1999) 1039–1046.
[41] C.T. McMurray, DNA secondary structure: A common and causative factor for expansion in human disease, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 1823–1825.