Journal of
Proteomics & Bioinformatics
ISSN: 0974-276X

Davoudi-Dehaghani et al., J Proteomics Bioinform 2012, 5:9
DOI: 10.4172/jpb.1000236

Research Article Open Access

Amplicon Secondary Structure Formation and Elongation during the Process of Sequencing

Elham Davoudi-Dehaghani[1,2], Sirous Zeinali[2,3]*, Nejat Mahdieh[3,4], Atefeh Shirkavand[5] and Hamideh Bagherian[3]

[1]Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[2]Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Pasteur St., Tehran, Iran
[3]Medical Genetics Lab, Kawsar Human Genetics Research Center, No. 41 Majlesi St., Vali Asr St., Tehran, Iran
[4]Growth and Development Research Center, Tehran University of Medical Sciences, Iran
[5]Department of Biology, Faculty of Sciences, Razi University, Kermanshah, Iran

Abstract

Understanding the exact mechanism involved in different molecular processes such as replication, transcription and recombination, as well as improving different applications of single stranded DNA, requires increasing of our knowledge regarding different properties of single stranded DNA secondary structure. Despite some efforts in this field, our understanding of the exact mechanism is still very limited.

During our investigation on hearing loss, we encountered an unusual sequencing result for exon 5 of TMC1 gene in a patient suspected of having mutation in this gene. BLASTN was used to find out the origin of extra segment seen in the sequencing result. The Mfold web server was used to investigate if secondary structure formation of the amplicon during sequencing can be the cause of this elongation.

Our investigation on the DNA sequence revealed that the extra sequence was an inverted repeat of the first 270 nucleotides inserted just after the sequence of the reverse primer. To investigate the origin of this repeat, we run the original PCR product on an agarose gel and no size increase was observed. This suggested that the inverted repeat has been formed during sequencing as a result of secondary structure formation and subsequent extension during sequencing.

Despite of the above the Mfold web server did not predict this hypothetical structure. However, when we submitted RNA equivalent for this sequence, the above server suggested formation of several secondary structures in which one of them was very similar to the one which had been predicted. Inability of the Mfold to predict DNA secondary structure for our sequence may lie in its thermodynamic parameters.

This finding, during DNA sequencing, suggests that single stranded DNA elongation in the presence of dNTPs and DNA polymerase can be used as a new way for studying the secondary structure in the single stranded DNA.

Keywords: Single stranded DNA; Secondary structure; DNA sequencing; DNA elongation; Secondary structure prediction; Mfold web server; Single stranded DNA folding

Introduction

DNA usually exists as double stranded macromolecule in living organisms but it can form, transiently, a single-stranded structure during vital cell processes such as replication, transcription and DNA repair [1,2]. Single stranded DNA (ssDNA) is much more flexible than double stranded DNA and can be folded easily [3,4]. It is proposed that, remarkable characteristics of the ssDNA such as enzymatic activities are as a result of its secondary structure [5-7]. This unique property of the ssDNA has made it a suitable tool for different applications in molecular genetic studies [8]. In recent years its application in medical diagnosis, treatment of disorders and nanotechnology have attracted attention and is expected to increase [9-11]. Despite the importance of knowledge regarding the ssDNA secondary structure, in clarifying precise mechanism of different cellular processes and increasing the efficiency of it in various applications, our understanding of thermodynamic properties of ssDNA secondary structure is very limited [12,10].

It is suggested that elongation of short oligodeoxynucleotides to a very long nucleic acids, in the presence of only dNTPs and DNA polymerase, is one of the results of secondary structure formation in the single stranded end of these oligodeoxynucleotides [13,14]. Liang et al. have proposed a terminal hairpin formation and self-priming extension (THF-SPE) model for efficient DNA elongation in this condition [14]. According to this model, a terminal hairpin structure at the 3′-end of DNA can act as a self-priming complex and lead to elongation of the molecule in the absence of specific primers. It is suggested that defining the precise mechanism of this event can be helpful to increase our knowledge of molecular evolution in the past [12].

During our investigation on hearing loss, trying to link the disease to a particular gene in different families, we found that TMC1 gene may be defected in a family. While sequencing patient's DNA, we encountered an unusual sequencing result for exon 5 of this gene. This unusual sequencing result suggested amplicon secondary structure construction and elongation during the sequencing process. This is the first report of its kind in that ssDNA can fold and elongate during cycle sequencing.

***Corresponding author:** Sirous Zeinali, Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Pasteur St., Tehran, Iran, Postal Code: 1595645513, Tel: 989121372040; Fax: 98 21 88939140; E-mail: zeinali@kawsar.ir

Received July 07, 2012; **Accepted** August 23, 2012; **Published** August 25, 2012

Citation: Davoudi-Dehaghani E, Zeinali S, Mahdieh N, Shirkavand A, Bagherian H (2012) Amplicon Secondary Structure Formation and Elongation during the Process of Sequencing. J Proteomics Bioinform 5: 204-207. doi:10.4172/jpb.1000236

Materials and Methods

Several families with hearing loss were admitted into our study. Informed consent forms were filled for each family. From each person 5-10 ml peripheral blood was taken in EDTA as anticoagulant. DNA was extracted using salting out method [15]. Sequencing was done by KBC (KBC, Tehran, Iran) using Big dye terminator kit™ (Life Technologies, USA) according to manufacturer's recommendation and the reaction product was run on a 3130XL Genetic Analyzer (Life Technologies, USA). The sequence of forward and reverse primers used to perform PCR and sequencing TMC1 exon 5 were: TMC1 E5F 5'-GTACTCCGTCCCTTAAATGGC-3' and TMC1 E5R 5'-GTCTTTCTGAGTGAAAGCGAGG.

BLASTN was used to find the origin of the 270-bp extra segment seen in the sequencing result [16]. Secondary structure prediction was performed by Mfold web server (version 3.5, online: http://www.bioinfo.rpi.edu/applications/mfold/old/dna/) [17].

Results

During our hearing loss study, linkage analysis using short tandem repeats was performed for several investigated families. We found linkage between the hearing loss and TMC1 gene in a family. PCR reaction was used to amplify all 24 exons of TMC1 gene in DNA sample of a patient suspected of having mutation in this gene followed by sequencing. During sequence analysis we encountered an unusual sequencing result for exon 5 of the gene. Analysis of the sequencing result showed that a 270 bp extra segment is formed. We checked this phenomenon by repeating PCR and running the PCR product on agarose gel to see if it has been present in the patient's DNA or it may have occurred during PCR or sequencing. No size increase was observed by gel electrophoresis (Figure 2). This process was repeated few times and no change was seen. This suggested that the elongation has happened during the sequencing reaction.

Further investigation and alignment of the DNA sequence showed that the 270 bp extra segment had been initiated just after the last nucleotide of reverse primer (Figure 1). BLASTN analysis showed that this extra sequence is an inverted repeat of the first part of this amplicon. The position of forward and reverse primers and extra segment on sequenced segment are shown in Figure 1.

We re-sequenced the same exon ten times, using the forward primer, to confirm the initial results. We expected to see a 509 bp length sequencing result if no elongation had occurred. However, we observed that the length of three out of ten sequences obtained were longer than expected. The length for the other seven were as expected (i.e. 509 bp). Interestingly no size increase was observed using the reverse primer.

Comparing the DNA sequences of the extra segment and its origin borders led us to find two homologous regions (HR1, 2). One of them was located 279 bp after the beginning of the forward primer with a length of 11 bp. The second one was at the 11 bp end of the 509 bp fragment (Figure 1). To find out if formation of secondary structure can explain the unusual sequencing result, DNA sequence of this region was submitted to Mfold web server. None of the suggested secondary structures was what we had expected. However, when RNA equivalent of the sequenced region was submitted, a secondary structure was predicted which could explain such an elongation during sequencing amplification. In this structure, 11 consecutive nucleotides at the 3' end of the nucleic acid (HR2) have been paired with another 11 bp homologous nucleotides just after the replicated segment (HR1) (Figure 3).

Based on our findings, we are proposing presence of a potential mechanism which explains the unusual sequencing observed in our study. Figure 4 shows schematic diagram of our hypothetical amplicon secondary structure and potential mechanism for the formation of the additional segment. As shown in Figure 4, terminal hairpin formation can be followed by elongation using the first part of the amplicon as a template.

Discussion

In this study we described an unusual sequencing result of exon 5 of the TMC1 gene in a patient with autosomal recessive non-syndromic hearing loss. This sequencing result showed that the sequenced segment is longer than was expected. The results of the BLASTN showed that the extra sequence is an inverted repeat of the first part of the amplicon.

At first, this result indicated an inverted duplication which may be present in the patient's DNA. This hypothesis was rejected since

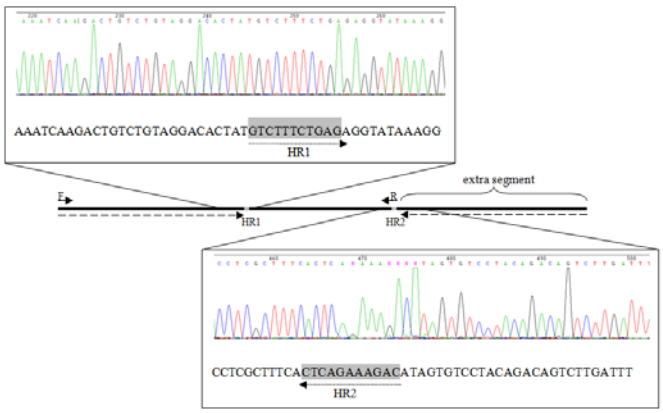

Figure 1: The position of forward and reverse primers, extra segment and homologous regions on the sequenced segment. Forward and reverse primers are shown by arrows. The sequence of homology regions 1, 2 (HR1, 2) and their borders are shown in boxes. Homologous regions are highlighted. Forward and reverse primers are shown by arrows. As shown in this figure, homologous region 2 is a part of the reverse primer sequence. Dashed arrows indicate the direction of sequences.

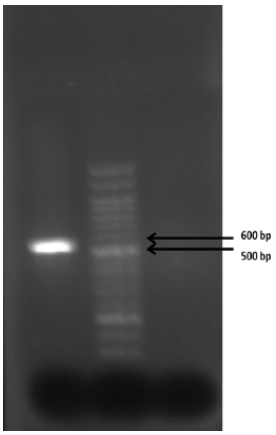

Figure 2: Gel electrophoresis result. The result of gel electrophoresis shows that the size of the amplicon is between 500-600 bp. As regards the sequenced segment is about 780 bp, it can be suggested that elongation of amplicon has happened during the sequencing process.

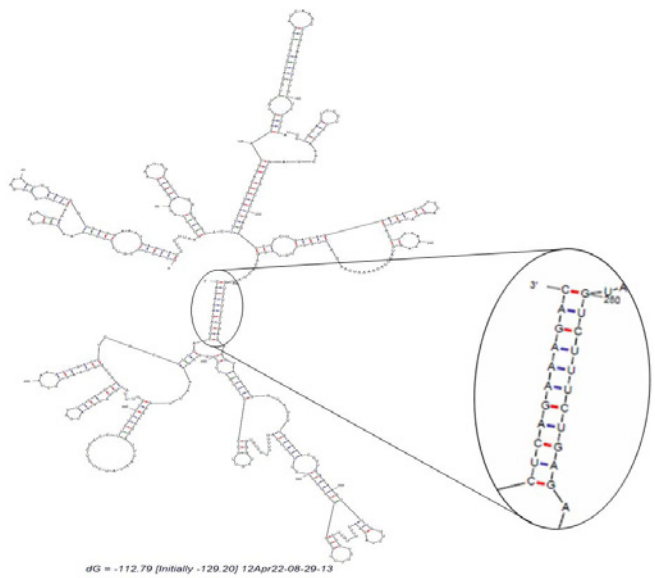

Figure 3: One of the predicted structures for the RNA equivalent of amplicon. In this structure, 11 nucleotides in the 3' end of the nucleic acid have been paired with nucleotides just after the replicated segment. Pairing of these regions is magnified in the right-hand image.

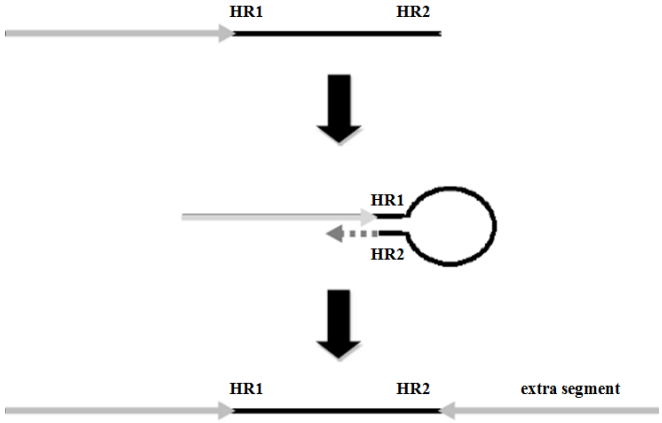

Figure 4: Schematic diagram of our hypothetical amplicon secondary structure and mechanism of its elongation during sequencing. As shown in this figure, terminal hairpin formation can be followed by elongation using the first part of the amplicon as a template.

gel electrophoresis finding didn't support size increase during PCR. Observation of this unusual result, in normal DNA samples, showed that this event cannot be related to the hearing loss. These results, suggested that the extra sequence could be the result of a secondary structure formation in the single stranded amplicon and elongation. This could have occurred through terminal hairpin formation and self-priming extension (THF-SPE) model suggested by Liang et al. [14]. One explanation for observation of this result in the sequencing but not in the PCR result may be that during PCR both forward and reverse primers are present and the normal size can participate in exponential amplification but the elongated types will amplify linearly which will not be detectable by agarose gel. However, absence of reverse primer and presence of only forward primer during the sequencing can explain formation of this molecular event during sequencing. One possible explanation is that only non-exponential amplification is generated during cycle sequencing.

Repeating sequencing with similar condition, with forward primer, showed that only three out of ten sequences were longer than we expected. It shows that structural stability of this hypothetical construction is not enough to form most of the time. However this structure can be one of the predicted constructions for the single stranded amplicon.

Inability of Mfold web server to predict this structure can be explained by indefinite thermodynamic parameters that are used by this program [12,10]. It seems that much more knowledge of the ssDNA properties is needed to predict all likely structures. Evidence for ssDNA folding in special conditions obtained from designed experiments or by sheer chance, like what was obtained in this study, can be used to investigate different parameters needed to identify accurate conformation of the ssDNA in different situations.

Although none of the predicted secondary structures for this ssDNA by Mfold web server was similar to our hypothetical structure, our hypothetical construction was predicted for the RNA equivalent. Similarity between properties of RNA and ssDNA can explain this result. Therefore the parameters used in the Mfold for DNA cannot be verified.

Continuing this study by performing experimental and computational investigations can prove formation of our hypothetical secondary structure and help us to determine actual conformation of this amplicon as a clue for increasing our knowledge about the ssDNA secondary structure formation.

From some years ago it has been shown that the nucleic acid secondary structure is an important interfering factor of different nucleic acid hybridization assays [18,19]. Lane et al. showed that amplicon secondary structure can interfere with hybridization of targets to their respective microarray oligonucleotide probes [20]. Measuring of hybridization kinetics for some designed probe-target pairs by Gao et al. also showed that probe and target conformations can affect the DNA hybridization kinetics in different situations [21].

Previous studies also have shown that secondary structure formation in the sequencing template can affect DNA polymerase activity and lead to no signal or weak signal in the electropherogram [22]. Here, we presented another result of ssDNA folding during DNA sequencing. It seems that specific forms of folding in the template can cause elongation during sequencing. More information about the detailed mechanism of DNA polymerase is needed to understand in what situations the secondary structure will result in halting the DNA polymerase and when it lead to elongation of template.

Until now several studies have been performed to predict folding of nucleic acids and some computational and experimental methods are suggested for this purpose but none of these methods can certainty predict ssDNA secondary structure [17, 23-26]. Dong et al. demonstrated that using of both experimental method by using Taq Exo 5'-nuclease and Mfold web server can help to increase the probability of correct folding prediction [10]. Studies of Liang et al. also have shown that the investigation of catenanes formed from two ssDNA molecules can provide worthwhile information about the ssDNA folding [12]. Our finding suggests a new way for studying the ssDNA folding. Investigation of the ssDNA elongation in the presence of dNTPs and DNA polymerase can help to detect terminal hairpin loop formation.

Elongation studies by using DNA segments with different sequences and various conditions can increase our knowledge about factors affecting ssDNA secondary structure formation.

Acknowledgments

This work was supported by the Kawsar Human Genetics Research Center. We gratefully acknowledge all staffs in Kawsar Center and Medical Genetic laboratory of Dr. Zeinali that contribute to this study.

References

1. Watson JD, Crick FHC (1953) A Structure for Deoxyribose Nucleic Acid. Nature 171: 737-738.
2. Kodadek T (1998) Mechanistic parallels between DNA replication, recombination and transcription. Trends Biochem Sci 23: 79-83.
3. Smith SB, Cui Y, Bustamante C (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. Science 271: 795-799.
4. Zhang Y, Zhou H, Ou-Yang ZC (2001) Stretching single-stranded DNA: interplay of electrostatic, base-pairing, and base-pair stacking interactions. Biophys J 81: 1133-1143.
5. Breaker RR (1997) DNA aptamers and DNA enzymes. Curr Opin Chem Biol 1: 26-31.
6. Sreedhara A, Li Y, Breaker RR (2004) Ligating DNA with DNA. J Am Chem Soc 126: 3454-3460.
7. Li Y, Liu Y, Breaker RR (2000) Capping DNA with DNA. Biochemistry 21: 3106-3114.
8. Gyllensten UB, Erlich HA (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc Natl Acad Sci U S A 85: 7652-7656.
9. Tan W, Wang H, Chen Y, Zhang X, Zhu H, et al. (2011) Molecular aptamers for drug delivery. Trends Biotechnol 29: 634-640.
10. Dong F, Allawi HT, Anderson T, Neri BP, Lyamichev VI (2001) Secondary structure prediction and structure-specific sequence analysis of single-stranded DNA. Nucleic Acids Res 29: 3248-3257.
11. Saccà B, Niemeyer CM (2012) DNA origami: the art of folding DNA. Angew Chem Int Ed Engl 51: 58-66.
12. Liang X, Kuhn H, Frank-Kamenetskii MD (2006) Monitoring single-stranded DNA secondary structure formation by determining the topological state of DNA catenanes. Biophys J 90: 2877-2889.
13. Kornberg A, Bertsch LL, Jackson JF, Khorana HG (1964) Enzymatic synthesis of deoxyribonucleic acid, XVI. Oligonucleotides as templates and the mechanism of their replication. Proc Natl Acad Sci U S A 51: 315-323.
14. Liang X, Kato T, Asanuma H (2008) Mechanism of DNA elongation during de novo DNA synthesis. Nucleic Acids Symp Ser 52: 411-412.
15. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.
16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.
17. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415.
18. Gamper HB, Cimino GD, Hearst JE (1987) Solution hybridization of crosslinkable DNA oligonucleotides to bacteriophage M13 DNA. Effect of secondary structure on hybridization kinetics and equilibria. J Mol Biol 197: 349-362.
19. Lima WF, Monia BP, Ecker DJ, Freier SM (1992) Implication of RNA structure on antisense oligonucleotide hybridization kinetics. Biochemistry 31: 12055-12061.
20. Lane S, Evermann J, Loge F, Call DR (2004) Amplicon secondary structure prevents target hybridization to oligonucleotide microarrays. Biosens Bioelectron 20: 728-735.
21. Gao Y, Wolf LK, Georgiadis RM (2006) Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. Nucleic Acids Res 34: 3370-3377.
22. McMurray AA, Sulston JE, Quail MA (1998) Short-insert libraries as a method of problem solving in genome sequencing. Genome Res 8: 562-566.
23. Ehresmann C, Baudin F, Mougel M, Romby P, Ebel JP, et al. (1987) Probing the structure of RNAs in solution. Nucleic Acids Res 15: 9109-9128.
24. Knapp G (1989) Enzymatic approaches to probing of RNA secondary and tertiary structure. Methods Enzymol 180: 192-212.
25. Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J Mol Biol 216: 585-610.
26. Gaspin C, Westhof E (1995) An interactive framework for RNA secondary structure prediction with a dynamical treatment of constraints. J Mol Biol 254: 163-174.