Acta Scientiarum

http://www.uem.br/acta
ISSN printed: 1679-9283
ISSN on-line: 1807-863X
Doi: 10.4025/actascibiolsci.v39i4.34600

Prediction of mutations on structure primase of the archaeon *Sulfolobus solfataricus*

Eden Silva e Souza and Michely Correia Diniz*

*Colegiado de Ciências Biológicas, Grupo de Estudos Integrados do Semiárido, Universidade Federal do Vale do São Francisco, Rod. BR-407, Km 12, Lote 543, s/n, 56300-990, Campus Ciências Agrárias, Projeto de Irrigação - Nilo Coelho, Petrolina, Pernambuco, Brazil. *Author for correspondence. E-mail: michely.diniz@univasf.edu.br; souza.eden@outlook.com*

ABSTRACT. All living organisms need a DNA replication mechanism and it has been conserved in the three domains of life throughout evolutionary process. Primase is the enzyme responsible for synthesizing *de novo* RNA primers in DNA replication. Archaeo-Eukaryotic Primase (AEP) is the superfamily that typically forms a heterodimeric complex containing both a small catalytic subunit (PriS) and a large accessory noncatalytic subunit (PriL). *Sulfolobus solfataricus* is a model organism for research on the Genetics field. The aim of this work was to evaluate, via Bioinformatics tools, three mutations in the large subunit (PriL) of the archaeon *Sulfolobus solfataricus*. The aspartic acid residue in the positions (Asp) 62, (Asp) 235, (Asp) 241 have been substituted by glutamic acid (Glu). The highest positive free energy variation of the three substitutions analyzed occurred with the mutation at the (Asp) 241 site. The *in silico* analysis suggested that these mutations in PriL may destabilize its tridimensional structure interfering with replication mechanisms of *Sulfolobus solfataricus*. Moreover, it may also alter interactions with other molecules, making salt bridges, for instance.

Keywords: *in silico*, PriL, 3D structure, DNA replication, primer, mutations.

Predição de mutações na estrutura da primase da arquea *Sulfolobus solfataricus*

RESUMO. Todos os organismos vivos necessitam de um eficiente mecanismo de replicação de DNA. Ao longo da evolução biológica foi observado que esse mecanismo é conservado nos três domínios da vida. Uma enzima importante que participa desse mecanismo é a RNA primase, a qual é responsável pela síntese *de novo* de iniciadores de RNA na replicação do DNA. Em Arquea-Eucariota, RNA Primase (AEP) tipicamente forma um complexo heterodimérico, que contém uma pequena subunidade catalítica (PriS) e uma subunidade maior não catalítica acessória (PriL). *Sulfolobus solfataricus* é um organismo modelo de Arquea para a pesquisa no campo da genética. O objetivo deste trabalho foi avaliar, por meio de ferramentas de bioinformática, três mutações pontuais na subunidade maior (PriL) de *Sulfolobus solfataricus*. Nas sequências mutantes, os resíduos de ácido aspártico nas posições (Asp) 62, (Asp) 235, (Asp) 241 foram substituídos por ácido glutâmico (Glu). A maior variação de energia livre positiva das três mutações analisadas ocorreu no sítio (Asp) 241. A análise *in silico* sugeriu que essas mutações em PriL podem desestabilizar sua estrutura tridimensional, interferindo com os mecanismos de replicação de *Sulfolobus solfataricus*. Além disso, podem alterar interações com outras moléculas, formando pontes salinas.

Palavras-chave: *in silico*, PriL, estrutura 3D, Replicação de DNA, iniciador, mutações.

Introduction

Archaea is one of the three domains of life which encompass organisms that live in extreme environments, such as strict anaerobic conditions, high temperature, and high salinity. This domain can be taxonomically divided into five Phyla. Crenarchaeota and Euryarchaeota are best characterized, and this taxonomic division is strongly supported by comparative genomics (Ishino, Kelman, Kelman, & Ishino, 2013; Sarmiento, Long, Cann, & Whitman, 2014). These five Phyla have 163 genomes that have already been revealed with 50 genomes from Crenarchaeota and 113 from Euryarchaeota (Jozwiakowski, Gholami, & Doherty, 2015). The genome of the crenarchaeon *Sulfolobus solfataricus* contains around 2,992,245 bp on a single chromosome and encodes 2,977 proteins and many RNAs (She et al., 2001).

Sulfolobus solfataricus (Sso) is an aerobic crenarcheon which grows at 80ºC. It is the most studied species of the crenarchaeal branch, what makes this species a model organism for research on the genetics field (Farkas, Picking, & Santangelo, 2013). Also, the simplicity of the replication

machinery and similarity to eukaryotes replication makes this organism valuable in the study of conserved DNA mechanisms in eukaryotes (Zuo, Xu, & Hao, 2015).

All living organisms need a DNA replication mechanism and it has been conserved in the three domains of life throughout evolutionary process (Schwob, 2004). However, bioinformatic, biochemical, structural, and genetic studies have established that this mechanism and its proteins involved in archaeal DNA replication are more similar to those in eukaryotic DNA replication than in bacterial DNA replication. In addition, that mechanism has some archaeal-specific particularities, instead of being a simpler version of the eukaryotic replication machinery (Kelman & Kelman, 2014). The study of archaeal DNA replication was initiated shortly after the recognition of Archaea as the third domain of life (Ishino et al., 2013).

Primase is the enzyme responsible to synthesize *de novo* RNA primers. It works inserting the first 7 to 10 nucleotides during the replication process (Kuchta & Stengel, 2010). Although primase has been conserved through time, two families of primase evolved: one common to bacteria and phages, and the other present in eukaryotic organisms and archea. The latter belongs to Archaeo-Eukaryotic Primase (AEP) superfamily that typically forms a heterodimeric complex containing both a small catalytic subunit (PriS/Prim1) and a large accessory noncatalytic subunit (PriL) (Guilliam, Keen, Brissett, & Doherty, 2015). Only in eukaryotic these two primase subunits are associated with DNA polymerase α and B subunit, which all together form the pol α/primase tetrameric complex (Frick & Richardson, 2001).

Although archea primase is similar to the eukaryotic primase, it is a heterodimer which lacks the two subunits of DNA polymerase (Rowen & Kornberg, 1978). Primase structure was elucidated in the following species: *Pyrococcus furiosus* (Pfu), *Pyrococcus horikoshii* (Pho), *Pyrococcus abyssi* (Pab), *Sulfolobus solfataricus* (Sso) (Liu et al., 2001; Lao-Sirieix, Nookala, Roversi, Bell, & Pellegrini, 2005; Cohen et al., 2003; Rowen & Kornberg, 1978). These structures revel that eukaryotic and archaea primase have unrelated polymerases fold, showing, however, a high conservation level and a three dimension arrangement of aspartate residues (Guilliam et al., 2015).

The substitution of aspartic acid residue (or aspartate), a polar amino acid negatively charged, by glutamic acid (or glutamate), also negatively charged (and very similar) may happen spontaneously when a mutation in the nucleotide sequence occurs (Betts & Russel, 2003). The aim of this work was to evaluate, via bioinformatics tools, three mutations in the large subunit of the archaeon *Sulfolobus solfataricus* (Sso; access code Q97Z83) primase. The aspartic acid residues in the positions (Asp) 62, (Asp) 235, (Asp) 241 have been substituted by a glutamic acid (Glu). We analyzed the probable effects of these mutations.

Material and methods

The sequence of the large subunit of the Sso primase was retrieved from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), under the access code Q97Z83 (330 amino acids) and the 3D structure from the Protein Data Bank (www.rcsb.org/pdb/home/home.do) under the code 1ZT2. To determine the positions which would be mutated, reverse transcription of the protein sequence into DNA was performed, and after transcription, from BioEdit v.7.2.6 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

Structural effects of a point mutation in the protein sequence (Q97Z83) was made using HOPE (http://www.cmbi.ru.nl/hope/home), which is an automatic server that analyses and gives insights in the structural features of native proteins and its mutant models. HOPE provides the 3D structural visualization of mutated proteins. The amino acid substitution is performed on the protein sequence. HOPE server predicts the output in the form of structural variation between mutant and wild type residues (Venselaar, te Beek, Kuipers, Hekkelman, & Vriend, 2010).

Energy of force field between residues was performed in Swiss-PdbViewer or DeepView v4.1 (http://spdbv.vital-it.ch/) (Guex & Peitsch, 1997). DeepView is a server that integrates protein visualization, sequence manipulation, creating mutations and changing the angle. DeepView allows the user to download PDB files, create sequence alignments, and structural superposition (Schwede, Kopp, Guex, & Peitsch, 2003).

Results and discussion

Table 1 summarizes the residues that were mutated and their respective positions in the mRNA. The aspartic acid (Asp) is encoded by the GAU and GAC, while glutamic acid (Glu) is encoded by the codons GAA and GAG. The substitution of uracil by adenine in the codon GAU or the substitution of cytosine by guanine in the codon GAC results in a nonsynonymous mutation (Graur & Li, 2000).

Table 1. Residues mutated and their respective positions in the mRNA.

Residues positions	Codon positions mRNA
Aspartic acid (asp 62)	184 -186
Aspartic acid (asp 235)	703-705
Aspartic acid (asp 241)	721-703

Figure 1 shows the Sso large subunit (PriL). The core primase has lower molecular mass (35.7 kDa) compared to the small subunit (37.6 kDa), and it is composed of only α-helices. This subunit is composed of a multi-domain residues (PriL-α), which includes the amino acids ranging from 1 to 94 and from 183 to 289. The function of PriL is still unclear, albeit it is known that this subunit has high affinity for single-stranded DNA and it might play a role in the process of template-DNA binding. It was suggested that PriL is a key molecule in the regulation of PriS activity (Lao-Sirieix et al., 2005).

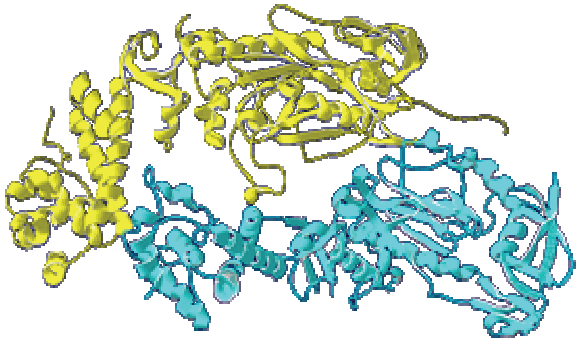

Figure 1. Structure of Sso primase core before mutation through DeepView. Large subunit colored in yellow; Small subunit colored in blue.

The residue (Asp) 62 is located on the surface of the PriL protein. The mutant residue is larger than the wild type. Mutations like this can interfere on how the primase interacts with other molecules or other parts of the protein. The residue (Asp) 62 forms a salt bridge with the Arginine 61 (Venselaar et al., 2010). This interaction is probably not changed, since both amino acids, wild and the mutant, have negatively charged side chains (Figure 2).

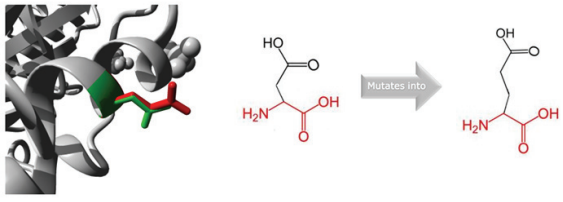

Figure 2. Molecular view of the amino acid substitution (mutation) (Asp) 62 (Glu). The protein is colored grey, the side chains of both the wild-type and the mutant residue are shown colored green and red respectively.

The literature uses different nomenclature for free energy (ΔΔG), which predicts the stability of a molecule (Cheng, Randall, & Baldi, 2006). ΔΔG is defined: ΔΔG = ΔG (mutant) – ΔG (wild/native). When Energy < 0, it is energetically more costly to have the native type structure than the mutant one, therefore, the mutation is more favorable to the structure stability. If Energy > 0, the mutant structure ΔG is higher than the native type one thus the mutation is less favorable to the structure stability.

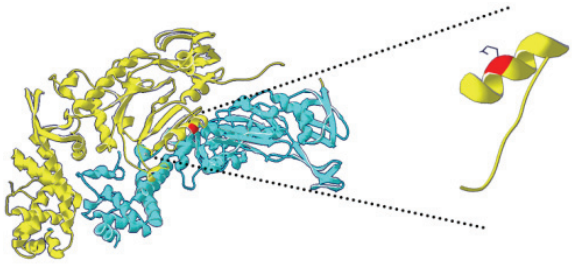

Figure 3. Location of the residue aspartic acid 62, labeled in red. Extension of the (Asp) 62 backbone of the mutant residue (Glu) 62.

The energy of force field to the wild primase (Asp) 62 was -29,880.172. However, after the mutation (Asp) 62 (Glu), this energy reduced to -29,873.285, ΔΔG = 6,89. The energy variation was positive and, therefore, the mutation may foment a destabilization of the structure.

The residue (Asp) 235 is located on the primase active site and it is part of the interpro domain IPR014052. This residue interacts, through salt bridges, with arginine 176 and arginine 244 (Venselaar et al., 2010). Once again this mutation may not change these interactions. Mutations in the active site usually result in loss of function. However, the amino acids Asp and Glu share some common properties, like negatively charged side chains. Therfore, this mutation may occur in some rare cases (Figure 4). Although, the mutation is likely to damage the primase structure and function. Again, as the residue (Asp) 235 is located on the protein surface, consequently this mutation may result in changes on how the primase interacts with other parts of the primase or other molecules.

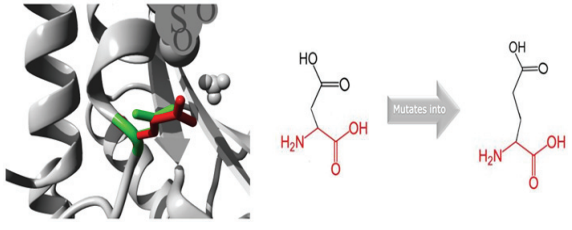

Figure 4. Molecular view of the amino acid substitution (Asp) 235 (Glu). The protein is colored grey, the side chains of both the wild-type and the mutant residue are shown colored green and red respectively. Also, it is shown in the figure the interaction of the residues, mutant and wild, with the salt bridges.

DeepView calculates the energy of force field by analysing the angles, bonds, torsion, electrostatic, constraint, and whether the residue is improper (Figure 3, 5 and 7). Force field energy analysis shows the initial energy of the primase with its native residues was -29,880.172. The force field energy calculated after the substitution (Asp) 235 (Glu) was -29,871.424, $\Delta\Delta G$ = 8,748. This result suggests that this positive mutation changes may foment a destabilization of the structure.

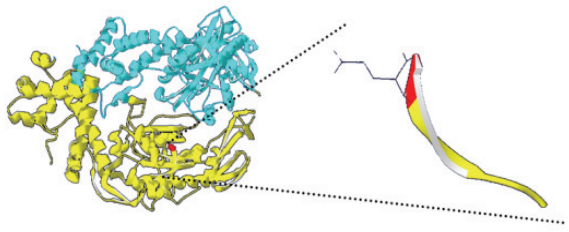

Figure 5. Location of the residue aspartic acid 235, labeled in red. Extension of the Asp 235 backbone of the mutant residue (Glu) 235.

A substitution of the aspartic acid 241, located on the chain C, by a glutamic acid was performed (Figure 6). The native amino acid was the only of its type found at this position. Mutations of conserved residues often result in damage to the protein. However, the mutant residue shares some similar properties with the wild residue, which indicates this mutation may occur in some rare cases. Aspartic acid 241 forms a salt bridge with the arginine 245, this interaction is not likely to change since these two residues have negatively side chains (Venselaar et al., 2010). The mutant residue is larger than the wild. This residue is located on the protein surface and, thus, this mutation may interfere on the interactions of the primase.

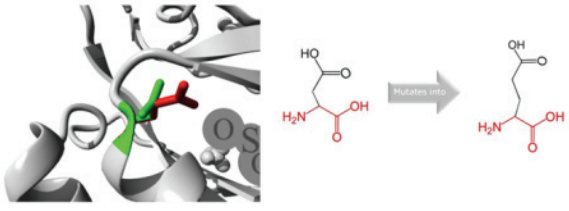

Figure 6. Molecular view of the amino acid substitution (Asp) 241 (Glu). The protein is colored grey, the side chains of both the wild-type and the mutant residue are shown colored green and red respectively. Also, it is shown in the figure the interaction of the residues, mutant and wild, with the salt bridges.

The energy of force field to the wild primase (Asp) 241 was -29,880.172, however after the mutation, this energy reduced to -29,791.828, $\Delta\Delta G$ = 88,342. This result suggests that this positive mutation changes may foment a destabilization of the structure.

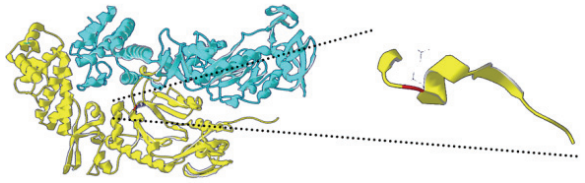

Figure 7. Location of the residue aspartic acid 241, labeled in red. Extension of the Asp 241 backbone of the mutant residue (Glu) 241.

The highest positive free energy variation of the three substitutions analyzed occurred with the mutation at the (Asp) 241 site. With these tools we can infer that this mutation could cause greater destabilization of the primase larger subunit (PriL) of the archaeon *Sulfolobus solfataricus*.

Lao-Sirieix et al. (2005) reinforces through a combination of DNA/RNA modeling and structure-based mutagenesis that there are evidences to support a two-fold role of the large subunit. The PriL would provide additional points of contact with the DNA template and would increase the stability of the primase–DNA complex helping to keep the primase in the correct orientation for synthesis initiation. The PriL would also be poised to take part in the mechanism of RNA counting and further primer processing probably due to its position within the heterodimeric primase relative to the nascent RNA primer.

A recent study showed a novel archea primase noncatalytic subunit, denoted PriX, from *Sulfolobus solfataricus*. Genetic analysis shows that PriX and PriL are essential for the growth of *S. solfataricus* and that PriX, PriL and PriS form a stable heterotrimer for efficient primer synthesis. Highly conserved PriX homologues are present in many members of the Phylum Crenarchaeota (Liu et al., 2015).

Conclusion

The mutation of the residue aspartic acid by a glutamic acid may occur naturally due to a misrepair on the DNA replication and by a substitution of the third nucleotide of the codons GAU and GAC to GAA to GAG, corresponding to aspartic acid and glutamic acid, respectively. The *in silico* analysis suggested that these mutations in PriL may cause destabilization on its structure interfering with replication mechanisms of *Sulfolobus solfataricus*. In addition, the mutation may alter the interactions with other molecules, such as salt bridges.

Acknowledgements

The authors thank Dr. Eduardo Miranda Dantas, from UNIVASF, for the contributions with orthographic and grammatical revision.

References

Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. In M. R. Barnes (Ed.), *Bioinformatics for geneticists: A Bioinformatics primer for the analysis of genetic data* (2nd Ed., p. 311-342).

Cheng, J., Randall, A., & Baldi, P. (2006). Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins: Structure, Function, and Bioinformatics*, *62*(4), 1125-1132.

Cohen, G. N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., & Thierry, J. C. (2003). An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Molecular microbiology*, *47*(6), 1495-1512.

Farkas, J. A., Picking, J. W., & Santangelo, T. J. (2013). Genetic techniques for the archaea. *Annual Review of Genetics*, *47*(1), 539-561.

Frick, D. N., & Richardson, C. C. (2001). DNA primases. *Annual Review of Biochemistry*, *70*(1), 39-80.

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis*, *18*(15), 2714-2723.

Guilliam, T. A., Keen, B. A., Brissett, N. C., & Doherty, A. J. (2015). Primase-polymerases are a functionally diverse superfamily of replication and repair enzymes. *Nucleic Acids Research*, *43*(14), 6651-6664.

Graur, D., & Li, W. H. (2000). *Fundamentals of molecular evolution* (2nd ed.). Sunderland, MA: Sinauer Associates.

Ishino, S., Kelman, L. M., Kelman, Z., & Ishino, Y. (2013). The archaeal DNA replication machinery: past, present and future. *Genes and Genetic Systems*, *88*(6), 315-319.

Jozwiakowski, S. K., Gholami, F. B., & Doherty, A. J. (2015). Archaeal replicative primases can perform translesion DNA synthesis. *Proceedings of the National Academy of Sciences*, *112*(7), 633-638.

Kelman, Z., & Kelman, Z. (2014). Archaeal DNA replication and repair. *Current opinion in microbiology*, *8*(6), 669-676.

Kuchta, R. D., & Stengel, G. (2010). Mechanism and evolution of DNA primases. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1804*(5), 1180-1189.

Lao-Sirieix, S. H., Nookala, R. K., Roversi, P., Bell, S. D., & Pellegrini, L. (2005). Structure of the heterodimeric core primase. *Nature Structural and Molecular Biology*, *12*(12), 1137-1144.

Liu, B., Ouyang, S., Makarova, K. S., Xia, Q., Zhu, Y., Li, Z., ... Huang, L. (2015). A primase subunit essential for efficient primer synthesis by an archaeal eukaryotic-type primase. *Nature Communications*, *6*, Article number: 7300. doi: 10.1038/ncomms8300.

Liu, L., Komori, K., Ishino, S., Bocquier, A. A., Cann, I. K., Kohda, D., & Ishino, Y. (2001). The Archaeal DNA Primase biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *Journal of Biological Chemistry*, *276*(48), 45484-45490.

Rowen, L., & Kornberg, A. (1978). A ribo-deoxyribonucleotide primer synthesized by primase. *Journal of Biological Chemistry*, *253*(3), 770-774.

Sarmiento, F., Long, F., Cann, I., & Whitman, W. B. (2014). Diversity of the DNA replication system in the archaea domain. *Archaea*, 2014. Article ID 675946. doi: 10.1155/2014/675946

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, *31*(13), 3381-3385.

Schwob, E. (2004). Flexibility and governance in eukaryotic DNA replication. *Current Opinion in Microbiology*, 7(6), 680-690.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., & Erauso, G. (2001). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proceedings of the National Academy of Sciences*, *98*(14), 7835-7840.

Venselaar, H., te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*, *11*, 548. doi: 10.1186/1471-2105-11-548

Zuo, G., Xu, Z., & Hao, B. (2015). Phylogeny and taxonomy of Archaea: a comparison of the whole-genome-based CVTree approach with 16S rRNA sequence analysis. *Life*, *5*(1), 949-968.

Received on December 19, 2016.
Accepted on September 5, 2017.