MOLECULAR MODELING OF *Leptospira biflexa* OxyR REGULATOR

L. M. Bomediano.*, A. S. K. Braz**, R. M. A. Costa.** and A. C. Q. Simoes.*

*Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas, UFABC, Santo André, Brazil
**Centro de Ciências Naturais e Humanas, UFABC, Santo André, Brazil

e-mail: livia.bomediano@ufabc.edu.br

Abstract: Leptospirosis is an important zoonosis and a neglected tropical disease responsible for serious public health problems that result in elevated costs to the economy. The disease is caused by pathogenic bacteria of the genus *Leptospira*, which has both pathogenic (the *interrogans* complex) and saprophytic (the *biflexa* complex) species, belonging to the order *Spirochaetales*. As any infectious agent, leptospires must overcome the reactive oxygen and nitrogen produced by macrophages of the host immune system. The genome sequencing of *Leptospira spp*. revealed the absence of global regulators against oxidative stress. However, our *in silico* analysis identified a potential OxyR orthologous in *Leptospira biflexa*. This work aimed the characterization of the putative OxyR regulator in *Leptospira biflexa* through molecular modeling. The structural modeling was investigated using the software Modeller and confirms the presence of the domains that are essential for OxyR function.

Keywords: oxidative stress response, molecular modeling, *Leptospira biflexa*.

Introduction

Leptospirosis or Weil's disease is a neglected tropical disease caused by pathogenic bacteria from the genus *Leptospira*, which embraces both pathogenic and saprophytic species, belonging to the order *Spirochaetales*. Leptospires are spiral-shaped bacteria that are 6-20 μm long and 0.1 μm in diameter with a wavelength of about 0.5 μm. Because they are so thin, live leptospires are best observed by dark field microscopy. Leptospira can be classified in a simplest way, in two serological groups: pathogenic species in the 'interrogans' complex and saprophytic species in the 'biflexa' complex [1].

As any infectious agent, leptospires must overcome a great challenge posed by the host immune system: the reactive oxygen and nitrogen produced by macrophages. These reactive radicals, like hydrogen peroxide (H_2O_2), radical superoxide, and nitric oxide (NO), interact with biomolecules such as DNA, RNA, proteins and lipids modifying their structure and function. The living organisms have a defense mechanism that maintains these radicals in low level within the cells. In bacteria, the adaptive response against these radicals is well studied in the model organism *Escherichia coli*. In this work, the model organism is *Escherichia coli str. K-12 substr. W3110*, RefSeq NC_007779.1.

The model organism responds very quickly against oxidative stress mainly through the activation of antioxidant genes. There are two important regulons in this process: OxyR and SoxRS. These two regulons activate a transcription of more than 16 genes related to the response to oxidative stress [2]. Products of OxyR regulon detoxify the cell from hydrogen peroxide [3] and protect the DNA from the oxidative attack [4]. The OxyR protein is highly sensitive to the oxidation promoted by the H_2O_2 molecule and only its oxidized form is capable of bonding to the promoter region and activates the transcription of antioxidant genes.

The most important genes activated by the oxidized form of OxyR are *katG* (peroxidase), *dps* (binding protein), *gorA* (GSH reductase), *grxA* (glutaredoxine), *aphCF* (alquil hidroperoxide-NADPH oxido-reductase) e *fur* (iron transport repressor). However, in *Leptospira* this response mechanism against oxidative stress is not well known. The genome sequencing of *Leptospira spp.* revealed the absence of global regulators against oxidative stress.

Our *in silico* analysis identified a potential OxyR orthologous in *Leptospira biflexa serovar Patoc str. 'Patoc 1 (Paris)'*, RefSeq NC_010602.1. This work aimed the characterization of the putative OxyR regulator in *Leptospira biflexa* through molecular modeling analysis. The structural modeling was investigated using the Modeller software that confirmed the presence of the domains that are essential for OxyR regulation function. Therefore, our results indicate the existence of a functional OxyR orthologous in *Leptospira biflexa*.

Materials and methods

Identification of the genomic and protein sequences – The open reading frames (ORFs) were identified by bioinformatics tools available at the National Center for Biotechnology Information (NCBI) website. BLASTP and BLASTN algorithms [5] were used to select sequences with high similarity to the OxyR sequences between the model organism *Escherichia coli str. K-12 substr. W3110*, RefSeq

NC_007779.1 and *Leptospira biflexa serovar Patoc str. 'Patoc 1 (Paris)', RefSeq NC_010602.1*.

Analysis of the structure of the sequences - The identified protein sequence from OxyR in *Leptospira biflexa* (geneID: 6221978, RefSeq YP_001840353.1 NCBI) and the known OxyR protein sequence of the model organism (geneID: 12934463, RefSeq YP_001840353.1, NCBI) were analyzed by the Pfam tool [6] to compare and confirm the presence of protein domains well known in the protein of the model organism *Escherichia coli* and the protein of interest.

Structural modeling of the OxyR protein - The most similar secondary structures of the protein OxyR of *Leptospira biflexa* were identified using the Hhpred tool [7]. The Modeller software [8] was used to build the initial structure of the OxyR protein. Modeller gives a prediction of a structure based on a protein sequence and its homology with the related sequences in the study. The quality of the protein model was studied with the Ramachandran plot [9].

Results

The protein sequence of the OxyR regulator in the model organism and the one identified in *Leptospira biflexa* were analysed with the Pfam tool and the main domains found are shown in Figure 1 and Figure 2, respectively.

Figure 1 - Domains found in OxyR regulator of *Escherichia coli*.

Figure 2 - Domains found in probably OxyR regulator of *Leptospira biflexa*.

The HHpred tool was used to identify the most similar secondary structures to the protein sequence of the identified OxyR regulator in *Leptospira biflexa*. This algorithm aligns the sequence given with all the sequences deposited in its data bank, showing the top sequences with the higher structure homology between each other. The Table 1 shows the most similar sequences for the OxyR in *Leptospira biflexa*.

Table 1 – PDB structures with higher homology with OxyR protein sequence of *Leptospira biflexa*.

PDB ID	Protein	Organism
3HO7	OxyR	*P. gingivalis*
3JV9	OxyR	*Neisseria meningitidis*
1I6A	OxyR	*Escherichia coli*

The sequences of the structures chosen as templates were aligned with the query sequenced to be modeled. After that, spatial features such as hydrogen bounds, alpha carbons distance and angles of the main chain and side chain were calculated and transferred from the template sequences to the query. The model obtained has to satisfy as much restrains as it is possible to be validated. Figure 3 shows the initial structure obtained.

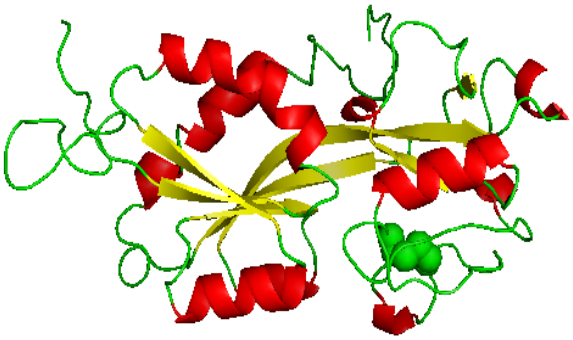

Figure 3 – Initial structure predicted for the probably OxyR protein of *Leptospira biflexa*.

The structure found was validated through the analysis of the Ramachandran plot built by the Modeller software. Figure 4 and Figure 5 show the Ramachandran plots obtained for the structure found. The Ramachandran plot reveals the backbone dihedral angles ψ (psi) against φ (phi) of amino acid residues in protein structure. The plot shows the empirical distribution of information observed in a structure where theoretically regions are favored to appear. From the kind of conformation these angles can assume it is possible to correlate it to specific secondary structures. Alpha-helices and beta-sheets adopt a limited set of phi-psi angle interactions. Some specific amino acids, such as glycine and proline, have unique Ramachandran plots because they differ from other amino acids in its structure. According to the literature [9], if the percentage of amino acids in favored regions is greater than 88% to 90% the model is good enough for *in silico* studies.

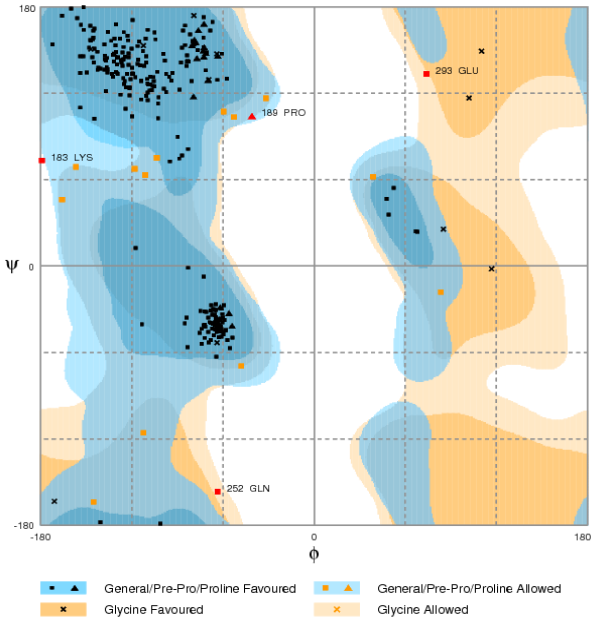

Figure 4 – Ramachandran plot of OxyR model obtained for *Leptospira biflexa*.

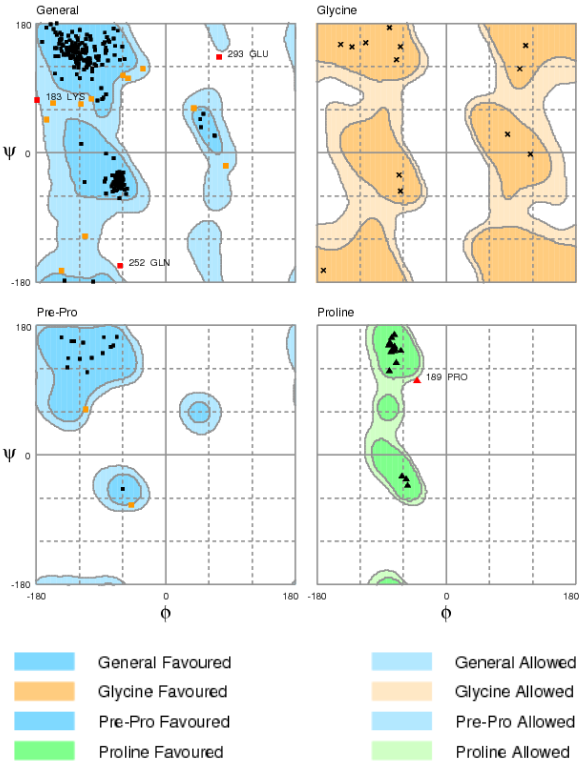

Figure 5 – Specific Ramachandran plot for the favoured regions of particular amino acids.

In our analisys the number of residues in favoured region represents 94.4% of the total residues of the sequence.

Discussion

The domains found in both *Leptospira biflexa* and *Escherichia coli* organisms are conserved and this is an indicative of the same function. The main domain, LysR substrate binding domain, is part of a conservative family of regulators with varied regulatory function of procariont genes such as the ones related to metabolism, sensibility, mobility and virulence [10]. The other domain is the HTH1 helix-turn-helix domain, important in DNA interaction proteins that are related to genomic expression regulation [11].

The three structures most similar found through Hhpred search tool were chosen as templates because they share an important alignment factor: the disulfide bonds that appear in the sequence of the OxyR protein in the model organism. The disulfide bond is directly correlated with the activity of this protein and it is characterized by the presence of conservative cysteines close to each other in the sequence. Figure 6 shows the aligned cysteines positions for the template organism chosen, *Escherichia coli*, compared to the query of *Leptospira biflexa*.

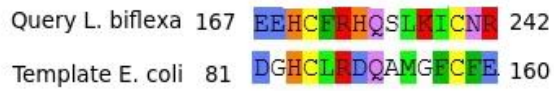

Figure 6 – Cysteines positions comparison.

The complete alignment of the sequences in the Hhpred tool is available as a complementary document and it also shows the predicted secondary structure for the sequences. It shows that the cysteines positions have the same secondary structure predicted. With the Modeller software it was possible to calculate and to obtain the predicted tridimensional structure with the disulfide bridge as two points at the right side of the molecule, as it was expected, since the disulfide bond is a very important characteristic of the OxyR protein function.

The Ramachandran plot ensured that the structure obtained is good enough to start a more robust analysis of the protein studied, for example, a molecular dynamic analysis. With more than 90% of the amino acids residues in favored regions it is clear that the identified structure corresponds to that one in the model organism and has the same function as an OxyR regulatory protein.

Conclusion

The analysis showed that the annotated LysR in *Leptospira biflexa* is actually an OxyR protein, homologous to the OxyR of the *Escherichia coli* organism. The structure found is very similar to its template and it maintains the same domains that are required for the protein to have the regulatory OxyR function. The evaluation of this initial prediction indicates the structure as a reliable one for further molecular analysis.

References

[1] Bharti A.R., Nally J.R., Ricaldi J.N., Mathias M.A., Diaz M.M., Levvet P.N., Gilman R.H., Willig M.R., Gotuzzo E., Vinetz J.M., Peru-United States Leptospirosis Consortium. Leptospirosis: a zoonotic disease of global importance. Lancet Infect Dis. 2003; 757-771.

[2] Pomposiello P.J., Demple B. Redox-operated genetic switches: the SoxR and OxyR transcription factors. Trends in Biotechnology 2001; 19: 109-114.

[3] Zheng M., Aslund F., Storz G. Activation of the OxyR transcription factor by reversible disulfide bond formation. Science 1998; 279: 1718-1721.

[4] Touati D. Sensing and protecting against superoxide stress in *Escherichia coli* – how many ways are there to trigger SoxRS response? Redox 2000; Rep 5: 287-293.

[5] Altschul D.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. J. Mol Biol. 1990; 215: 403-410.

[6] Punta P.C., Coggill R.Y., Eberhardt J., Mistry J., Tate C., Boursnell N., Pang K., Forslund G., Geric J., Clements A., Heger L., Holm E.L.L., Sonnhammer S.R., Eddy A., Bateman R.D. The Pfam portein families database. Nucleic Acids Research 2012; 40: 290-301.

[7] Soding J., Biegert A., Lupas A.N. The Hhpred interactive server for protein homology detection and structure prediction. Nucleic Acids Research 2005; 33: 244-248.

[8] Sali A., Potterton L., Yuan F., van Vlikmen H., Karplus M. Evaluation of comparative protein modelling by Modeller. Proteins 1995; 23: 318-326.

[9]Ramachandran G.N., Ramakrishnan C., Sasisekharan V. Stereochemestry of polypeptide chain configurations. Journal of Molecular Biology 1963; 7: 95-9.

[10]Maddocks S.E., Oyston P.C.F. Structure and function of the LysR-tyoe transcriptional regulator family proteins. Microbiology 2008; 154: 12369-3629.

[11] Brennan R.G., Mathews B.W. The helix-turn-helix DNA binding motif. J. Biol. Chem. 1989; 264:1903-6.