Identification of evolutionary conserved DNA sequence and corresponding S21 ribosomal protein region for diagnostic purposes of all *Borrelia* spirochetes

Roman Kotłowski and Lucyna Holec-Gąsior

Department of Molecular Biotechnology and Microbiology, Gdańsk University of Technology, Gdańsk, Poland

It is still under investigation whether all *Borrelia* sp. causing Lyme borreliosis and other diseases are already identified and properly classified as human pathogens. For this reason, it is of great importance to develop a diagnostic ELISA test that detects all *Borrelia* sp. The aim of this study was to identify conserved DNA and protein regions present in all currently known *Borrelia* sp. In experimental studies 31 available *Borrelia* sp. genomes were aligned and screened for the presence of evolutionary conserved regions. As a result of bioinformatics analysis, one evolutionally conserved DNA region encoding a core fragment of the S21 ribosomal protein was identified. Both a couple of genus-specific PCR primers and the S21 protein B-cell epitope were designed for prospective diagnostic purposes.

**Keywords:** *Borrelia* sp., Lyme disease, diagnostic

**Abbreviations:** ELISA, Enzyme-Linked Immunosorbent Assay; IgG, Immunoglobulin G; PCR, Polymerase Chain Reaction; PDB, Protein Data Bank

INTRODUCTION

It is still under investigation if all the *Borrelia* species causing Lyme borreliosis and other diseases (Barstad et al., 2018; Cutler et al., 2016; Jahfari et al., 2017; Stanek & Reiter, 2011), mainly around the N hemisphere, are fully identified and properly classified as human pathogens (Schneider et al., 2008; Rudenko et al., 2008). In the summer, people visiting the woodlands are exposed to the attack of ticks that can be vectors of virulent *Borrelia* sp. From a diagnostic point of view, the high genetic diversity of *Borrelia* sp. makes it difficult to detect specifically those species that are responsible for the development of the disease. Therefore, it is not easy to make progress in the field of accurate medical treatment of infected people and epidemiological studies. In the present study, conserved DNA sequences and the corresponding S21 ribosomal protein region were identified and found to be common to all *Borrelia* sp. deposited at the Gene Banks in NCBI (National Center for Biotechnology Information) and DDBJ (DNA Databank of Japan). Identification of all *Borrelia* sp. at a time can help in the detection of a rare pathogenic strain that cannot be identified using currently available tests specific to *Borrelia* sp. sensu lato and *Borrelia* sp. sensu stricto (Jahfari et al., 2017). In the 16SrDNA gene it is possible to identify conserved regions shared by all *Borrelia* sp. (Parola et al., 2011; Ni et al., 2014). However, greater variability within the flaB gene requires optimization of DNA amplification conditions if genus-specific primers are used (Wodecka et al., 2010). The conserved region of S21 gene discovered in our study can therefore be used as an alternative target in the diagnostics of *Borrelia* sp. by PCR, but mainly, as an additional test to exclude false positive results obtained for the 16SrDNA or flaB genes, in the case of contamination of the samples by previous amplicons. However, we recommend the use of PCR-based tests to detect *Borrelia* sp. directly from ticks, rather than from clinical samples.

The immunological assays available commercially and assays developed in many laboratories worldwide have a widely divergent sensitivity and specificity (Kodym et al., 2018). They are based on native whole-cell antigens, purified antigens (flagellar components), or whole-cell antigens combined with recombinant antigens. Furthermore, the vast majority of borrelial proteins are cross-reactive (Reed, 2002). The complexity of the antigenic composition among *Borrelia* genospecies and differential expression of proteins in the host and vector (temporal and spatial antigenic variability) has posed challenges for the serodiagnosis of borreliosis. This variability impedes the development of a single immunodiagnostic assay for all major *Borrelia* genospecies. The optimal combination of the most specific epitopes in the form of recombinant chimeric proteins could increase the discrimination abilities of serodiagnostic tests. In the present study we also designed a chimeric antigen for the specific detection of *Borrelia* sp.

MATERIALS AND METHODS

In this study, 31 complete *Borrelia* sp. genomes were aligned using MAFFT v. 7.271 software (Katoh, 2002), and a DNA fragment encoding the most conserved protein within the entire *Borrelia* sp. proteome, S21, was identified. The conserved amino-acid sequence of the S21 protein was then compared with all protein sequences deposited at the Proteins Data Bank of Japan. Next, we used RaptorX Structure Prediction server (http://raptorx.uchicago.edu/) to predict the 3D structure of the whole S21 ribosomal protein and its conserved fragment (Figs. 3A and 3B, respectively). The 5MMJ PDB protein structure from the small subunit of the chloroplast *Spinacia oleracea* ribosome was selected by the RaptorX online software as a template (Bieti et al., 2017) after comparison with all available 3D protein structures from Protein Data Bank (PDB) (Källberg et al., 2012). B-cell epitopes present within the whole S21...
Figure 1. Multiple sequence alignment of the S21 ribosomal protein gene fragment for 31 genomic sequences deposited at the National Centre for Biotechnology Information (NCBI).

Bor_spF and Bor_spR are the degenerated primer sequences selected for PCR detection of all Borrelia sp.

Figure 2. Multiple sequence alignment of the S21 protein fragment of 29 Borrelia/Borreliella sp. amino acid sequences deposited in the Protein Bank of Japan.

Highlighted in grey is the conserved region characteristic for ribosomal S21 protein of Borrelia/Borreliella sp. The predicted B-cell epitope is distinguished on a black background.
Diagnostic of all *Borrelia* sp.


**RESULTS AND DISCUSSION**

Multiple alignments of *Borrelia* sp. genomes allowed for the selection of the most conserved DNA protein coding region (Fig. 1). The identified DNA region was used to design slightly degenerate PCR primers: Bor_spF and Bor_spR specific to all *Borrelia* sp. Translation of the selected DNA region revealed, irrespective of point mutations in the DNA sequences from different genomes, an identical amino-acid chain among all of the analyzed *Borrelia* sp, suggesting that this core part of the ribosomal S21 protein is evolutionary conserved. All amino-acid sequences with 100% similarity identified using BLAST server at DDBJ belonged to the genus *Borrelia/Borreliella* (Fig. 2). The conserved part of S21 ribosomal protein may potentially serve as an intracellular antigen, indicating active human infection caused by all of the *Borrelia* spirochetes on the basis of ribosomal metabolism. The structure prediction analysis confirmed a similar 3D-conformation for both the entire S21 protein and for the conserved region only (Figs. 3A and 3B, respectively). Therefore, based on bioinformatics analysis, we can speculate that it will be possible to immunize animals for the production of specific antibodies using the S21 protein conserved region. Because the selected S21 protein fragment has a molecular weight below 5kDa, it may act as a hapten that is capable of inducing an immune response and recognizing specific antibodies only when attached to a large carrier such as a protein or assembled into one chimeric protein with other fragments of various antigens. This is why we suggest combining two genus-specific protein fragments, the S21 conserved peptide (Fig. 2) and, for example, a part of the FlaB protein: QIRGLSQASRNTSKAINFIQTTEGNL. Successful application of chimeric proteins, assembled using molecular biology techniques, in immunological assays for the diagnosis of viral, bacterial and parasitic diseases has already been demonstrated (Alcaro et al., 2003; Holec-Gąsior et al., 2012a, Holec-Gąsior et al., 2012b; Drapała et al., 2014; Ferra et al., 2015). To date, only a few studies have shown the reactivity of this kind of proteins with specific IgG antibodies from human sera of individuals with Lyme borreliosis (Gomes-Solecki et al., 2000; Schreterova et al., 2017). However, chimeric proteins are a new generation of recombinant products which have the potential to replace the native antigen (e.i. crude fractions of sonicated cells of microorganisms). Furthermore, the construction of recombinant chimeras containing genes from several genospecies can allow generating one protein that confers antigenicity to multiple strains. In addition, the use of pure chimeric proteins as diagnostic antigens provides greater flexibility in adapting the test to different assay formats. Thus, the S21-FlaB fusion peptide proposed in this study appears to be a very promising antigen. The core region of the S21 protein contains the B-cell epitope (Figs. 3C and 3D), whereas FlaB, the major endoflagellar filament protein (Motaleb et al.,...
acts as an antigen targeting the initial antibody response of the host (Aguero-Rosenfeld et al., 1993). The advantage of the chimeric antigen designed in this study is 100% specificity to all Borrelia sp. and the presence of two different antigens.

Acknowledgements

This work was not supported by any grant or other form of funding.

LITERATURE