Journal of Theoretical Biology 400 (2016) 11–18

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Identification of CD4+ T-cell epitope and investigation of HLA distribution for the immunogenic proteins of *Burkholderia pseudomallei* using *in silico* approaches – A key vaccine development strategy for melioidosis

Rayapadi G. Swetha, Madangopal Sandhya, Sudha Ramaiah, Anand Anbarasu *

Medical & Biological Computing Laboratory, School of BioSciences and Technology, VIT University, Vellore 632014, India

HIGHLIGHTS

- Identification of CD4+ T-cell epitope in immunogenic proteins of *B. pseudomallei.*
- Epitope set of bipB, fliC, groEL and ompA with maximum HLA-DR alleles was predicted.
- Population coverage analysis reveals the identified epitopes have > 90% occurrence rate.
- Identified epitopes serve as good candidates for design of vaccine for melioidosis.

ARTICLE INFO

Article history:
Received 7 December 2015
Received in revised form
18 March 2016
Accepted 8 April 2016
Available online 13 April 2016

Keywords:
Melioidosis
Burkholderia pseudomallei
Epitope
Vaccine
Population coverage

ABSTRACT

Melioidosis is a serious infectious diseases affecting multi-organ system in humans with high mortality rate. The disease is caused by the bacterium, *Burkholderia pseudomallei* and it is intrinsically resistant to many antibiotics. Thus, there is an urgent need for protective vaccine against *B. pseudomallei*; which may reduce morbidity and mortality in endemic areas. The identification of peptides that bind to major histocompatibility complex II class helps in understanding the nature of immune response and identifying T-cell epitopes for the design of new vaccines. Previous studies indicate that, ompA, bipB, fliC and groEL proteins of *B. pseudomallei* stimulate CD4+ T-cell immune response and act as protective immunogens. However, the data for CD4+ T-cell epitopes of these immunogenic proteins are very limited. Hence, in this present study we attempted to identify CD4+ T-cell epitopes in *B. pseudomallei* immunogenic proteins using *in silico* approaches. We did population coverage analysis for these identified epitopic core sequences to identify individuals in endemic areas expected to respond to a given set of these epitopes on the basis of HLA genotype frequencies. We observed that eight epitopic core sequences, two from each immunogenic protein, were associated with the maximum number of HLA-DR binding alleles. These eight peptides are found to be immunogenic in more than 90% of population in endemic areas considered. Thus, these eight peptides containing epitopic core sequences may act as probable vaccine candidates and they may be considered for the development of epitope-based vaccines for melioidosis.

1. Introduction

Melioidosis, also known as "Whitemore disease", is a fatal bacterial disease caused by the Gram-negative bacterium *B. pseudomallei* (Leelarasamee and Bovornkitti, 1989). The organism is classified as category B priority pathogen by the Centers for Disease Control and Prevention (Bondi and Goldberg, 2008; Moran, 2002). Melioidosis has been called the "Great Imitator" as the disease does not exhibit any particular clinical features but has symptoms of many other infectious diseases (Jakribettu et al., 2014; Walsh and Wuthiekanun, 1996). Melioidosis is endemic in tropical regions such as south-east Asia and northern Australia

Abbreviations: bipB, Cell invasion protein; fliC, Flagellin; groEL, Molecular chaperone; HLA, Human leukocyte antigen; IEDB, Immuno epitope database; MHC, Major histocompatibility complex; ompA, outer membrane protein A

* Corresponding author.
E-mail address: aanand@vit.ac.in (A. Anbarasu).

http://dx.doi.org/10.1016/j.jtbi.2016.04.009
0022-5193/

with high fatality rates due to lack of timely diagnosis, prompt treatment and licensed vaccine (Currie, 2015; Nithichanon et al., 2015). *B. pseudomallei* exhibits resistance to a wide variety of antibiotics including β-lactams, macrolides, aminoglycosides and polymyxins. The drug of choice for the treatment of severe melioidosis is Ceftazidime but the mortality rate is more than 40% even in treated patients (Anutrakunchai et al., 2015; Dance, 2002). Reports from many parts of the world reveal that *B. pseudomallei* strains have developed resistance to Ceftazidime and the treatment is unsuccessful as the disease is often associated with relapse (Anutrakunchai et al., 2015; Dance et al., 1989). Researchers have attempted to discover vaccines against *B. pseudomallei* but most of them failed to have an effective impact (Benoit et al., 2015). Many techniques have been applied to study the immunogenic properties of *B. pseudomallei* but a good vaccine has not yet been designed to prevent melioidosis. Naturally attenuated vaccines did not provide complete protection over chronic melioidosis. Killed whole cell vaccines were developed for melioidosis but further experiments gave conflicting results (Patel et al., 2011). Several studies on mice have reported that, live attenuated mutants of *B. pseudomallei* were able to induce protective immunity against virulent *B. pseudomallei.* However, it is doubtful that such mutants would be useful as a vaccine candidate for human use (Atkins et al., 2002; Breitbach et al., 2008; Cuccui et al., 2007; Rodrigues et al., 2006; Srilunchang et al., 2009; Stevens et al., 2004). The vaccine development for *B. pseudomallei* has posed challenges such as identifying broadly protective antigens, designing efficient vaccine delivery and adjuvant systems and better understanding of both acute and chronic infection (Silva et al., 2013). An effective and ever safer multicomponent vaccine should be designed by understanding the cellular and humoral responses involved in the protection against *B. pseudomallei.* In this regard, the use of an epitope (a small part of an antigen) rather than considering full-length protein represents an alternative choice for the development of an effective vaccine candidate. Epitopes restrict the cross-reactivity and induce protective antibody response. The vaccine developed based on this approach might have increased potency and will be safer for human use (Armstrong, 2007; Sarkar-Tyson and Titball, 2010; Sette and Fikes, 2003). The selection and identification of peptides that have high immune response and memory is an important step in vaccine discovery. The prime requisite for an efficient humoral immune response is to activate CD4+ T-cells and thereby, it generates lytic IgG and memory B cells. The antigen peptides associated with major histocompatibility complex (MHC) class II on antigen-presenting cells are recognized by activated CD4+ T-cells (Patronov and Doytchinova, 2013). Thus, the objective of our study is to develop epitope-based vaccine candidates based on proteins which evoke CD4+ T-cell response using computational approaches. The results from the CD4+ T-cell response are more accurate when compared to whole genome analysis.

The advancements in field of immunoinformatics have revolutionised the way of dealing with fatal pathogens and have also enhanced the way to understand the mechanism of immune response when the body encounters antigens (De Groot et al., 2002). Several studies predicted T-cell epitopes by using HLA class II peptide-binding assays (Musson et al., 2014) and by using structure and sequence based prediction tools (Nithichanon et al., 2015). Differentiated T cells produce immune response to pathogens and react with epitopes. The epitope identification is an important prerequisite to develop a safe and potent epitope-based vaccine (Lauemoller et al., 2000; Sette et al., 2002). Advancements in computational immunology methods have greatly minimized the time and effort for screening potential epitopes (De Groot et al., 2001; De Groot et al., 2002). The prediction of immunogenic epitopes for *B. pseudomallei* remains crucial and challenging task. The main challenges in epitope prediction for melioidosis include selection and identification of peptides in *B. pseudomallei* that have high immune response and memory. When the MHC encounters an antigen, it presents the antigen to the T-cells which then activate the CD4+ cells which in turn trigger immune response. Hence, antigen recognition and presentation is also a crucial step that decides the antigenicity of the epitope. T-cell epitope identification currently depends on predicting the most restrictive step of antigen presentation: peptide binding to MHCs. However, this is an essential, but not sufficient, condition for epitope recognition.

Previous reports suggest that outer membrane protein A (ompA), cell invasion protein (bipB), flagellin (fliC) and molecular chaperone (groEL) proteins from *B. pseudomallei* are highly effective in eliciting CD4+ T-cell immune response and they also act as protective immunogens (Haque et al., 2006; Suwannasaen et al., 2011; Tippayawat et al., 2011; Ye et al., 2008). Hence, we have selected these four proteins of *B. pseudomallei* for the present study. Our study involves the identification and HLA distribution analysis of CD4+ T-cell epitopes in these immunogenic proteins of *B. Pseudomallei.* The predicted putative epitopes may serve as good vaccine candidates in the development of epitope-based vaccines for melioidosis.

As demonstrated by a series of recent publications (Jia et al., 2016a; Liu et al., 2016a, 2016b) in compliance with Chou's 5-step rule (Chou, 2011), to establish a really useful sequence-based statistical method for a biological system, the five guidelines should be followed: (1) construct or select a valid benchmark dataset to train and test the predictor (2) formulate the biological sequence samples with an effective mathematical expression that can truly reflect their intrinsic correlation with the target to be predicted (3) introduce or develop a powerful algorithm to operate the prediction (4) properly perform cross-validation tests to objectively evaluate the anticipated accuracy of the predictor and 5) establish a user-friendly web-server for the predictor that is accessible to the public. We described how to deal with these steps one-by-one.

2. Materials and methods

2.1. Selection and Retrieval of protein sequences

The protein sequences of ompA (Locus tag: BPSL2522, Accession No: YP_109118.1), bipB (Locus tag: BPSS1532, Accession No: YP_111538.1), fliC (Locus tag: BPSL3319, Accession No: YP_109915.1) and groEL (Locus tag: BPSL2697, Accession No: YP_109293.1) were retrieved from GenBank database (Benson et al., 2013).

2.2. Prediction of T-cell epitope

The most commonly used tool to predict the T-cell epitope is NetMHCIIpan (Version 3). NetMHCIIpan is the best suited tool for studying the pan-specific binding affinity of peptides with entire human leukocyte antigen (HLA) class II molecules including HLA-DR, HLA-DP and HLA-DQ from a known protein sequence. The tool can also identify the possible epitopes that trigger immune response. The pan-specific T-cell epitope prediction tools available for MHC class II molecules are very limited. The classical MHC class II predictor, TEPITOPE (Sturniolo et al., 1999) and MHC2MIL (Xu et al., 2014) tools have certain limitations and hence we decided to employ NetMHCIIpan.

NetMHCIIpan is one of the most accurate prediction servers that uses artificial neural networks algorithm to identify the binding affinities of peptides and has been trained on more than 50,000 quantitative peptide-binding measurements (Karosiene