Molecular Immunology 97 (2018) 56–62

Contents lists available at ScienceDirect

Molecular Immunology

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

In silico design of *Mycobacterium tuberculosis* epitope ensemble vaccines

Preksha Shah[a], Jaymisha Mistry[a], Pedro A. Reche[b], Derek Gatherer[c], Darren R. Flower[a,*]

[a] School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK
[b] Immunomedicine Group, Facultad de Medicina, Departamento de Inmunologia & O2, Universidad Complutense de Madrid, Madrid, Spain
[c] Division of Biomedical & Life Sciences, Faculty of Health & Medicine, Lancaster University, Lancaster LA1 4YW, UK

ARTICLE INFO

Keywords:
Tuberculosis
Vaccine
Immunoinformatics

ABSTRACT

Effective control of *Mycobacterium tuberculosis* is a global necessity. In 2015, tuberculosis (TB) caused more deaths than HIV. Considering the increasing prevalence of multi-drug resistant forms of *M. tuberculosis*, the need for effective TB vaccines becomes imperative. Currently, the only licensed TB vaccine is *Bacillus Calmette-Guérin* (BCG). Yet, BCG has many drawbacks limiting its efficacy and applicability. We applied advanced computational procedures to derive a universal TB vaccine and one targeting East Africa. Our approach selects an optimal set of highly conserved, experimentally validated epitopes, with high projected population coverage (PPC). Through rigorous data analysis, five different potential vaccine combinations were selected each with PPC above 80% for East Africa and above 90% for the World. Two potential vaccines only contained CD8+ epitopes, while the others included both CD4+ and CD8+ epitopes. Our prime vaccine candidate was a putative seven-epitope ensemble comprising: SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL, QSSFYSDW and KMRCGAPRY, with a 97.4% global PPC and a 92.7% East African PPC.

1. Introduction

Mycobacterium tuberculosis infects populations worldwide. Due in part to troubling rates of new and relapsing tuberculosis (TB), the estimated 2015 death toll from TB was 1.8 million, with 10.5 million new cases recorded (WHO, 2017). Mortality rates are disproportionately high in Africa (Migliori et al., 2010): for example, in 2015, Kenya was reported to have TB affecting over 81,000 people and causing the deaths of at least 16,000 people. TB usually presents as a pulmonary disease transmitted by droplet inhalation resulting in symptoms such as persistent cough, fever, and night sweats (Heemskerk et al., 2015). Individuals with healthy immune systems can often suppress the disease, typically being asymptomatic but with a latent infection. Problems arise for immunocompromised patients who cannot mount an immune response sufficient for the suppression of symptoms.

The *M. tuberculosis* genome comprises over 4 million base pairs and approximately 4000 genes (Cole et al., 1998). The immune response mounted against *M. tuberculosis* mainly involves cellular immunity, including CD4+ and CD8+ T cells (Mendez-Samperio, 2016a; Mendez-Samperio, 2016b). Once activated, both CD4+ and CD8+ T cells secrete cytokines inducing an immune response. The CD8+ cells also mediate cytotoxicity and lysis of infected cells. Effective T cell responses are essential for *M. tuberculosis* elimination

The slow growth rate of *M. tuberculosis* (792–932 min doubling time), its complex pathogenesis, and its capacity to remain dormant, are major challenges to the development of effective treatments against TB (Zenteno-Cuevas, 2017). This is compounded by widespread resistance to antibiotics, such as isoniazid and rifampicin, with 480,000 cases of multi-drug resistant TB emerging annually (Raviglione and Sulis, 2016). Thus, vaccination against TB remains a priority; especially in the developing world, particularly many African countries, India, and Indonesia, where the disease is widespread (WHO, 2017).

The *Bacillus Calmette-Guérin* (BCG) vaccine – the only currently available TB vaccine offering prophylaxis – is an attenuated form of *Mycobacterium bovis*, used globally since 1923. Estimates suggest over 3 billion people have been vaccinated with BCG (Franco-Paredes et al., 2006). BCG has subsequently come under much scrutiny. Each BCG vaccine dose, whilst containing a preparation of attenuated *M. bovis*, has different biological effects, as the amount of viable versus dead organisms in each dose varies (Bali et al., 2015). Depending on the strain used, immunogenicity, reactogenicity and viability varies by manufacturer. BCG has efficacy ranging from 0 to 80% against adult pulmonary TB (Bali et al., 2015; Romano and Huygen, 2012), providing protection for 10–20 years from immunisation. As a partially-effective vaccine, BCG only protects paediatric patients from severe TB. BCG, as a live-attenuated vaccine, is seen as having a low safety profile due to the risks associated with its use in immunocompromised people and the possibility of the bacterium reverting to its virulent form (Detmer and

* Corresponding author.
E-mail address: d.r.flower@aston.ac.uk (D.R. Flower).

https://doi.org/10.1016/j.molimm.2018.03.007
Received 12 December 2017; Received in revised form 8 March 2018; Accepted 10 March 2018
Available online 19 March 2018

Glenting, 2006). BCG's varying efficacy is a particular problem in developing countries, such as East Africa. The need for effective, safe TB vaccines is clear.

Sixteen TB vaccine candidates are currently in phase I, II, and III (Evans et al., 2016). Almost half of these candidate vaccines contain a live attenuated form of *M. tuberculosis*. Such an approach comes with many safety concerns. Indeed, BCG is contraindicated in patients with human immunodeficiency virus (HIV). Immunocompromised people are at increased risk of disseminated BCG disease (WHO, 2017). Disseminated BCG disease is a rare, life-threatening complication presenting with symptoms such as persistent fever, infection, weight loss and night sweats (Eccles and Mehta, 2011).

Viral vector-based vaccines include recombinant Vaccinia strain Ankara virus (MVA), synthesised to express antigen 85A of *M. tuberculosis* and an adenovirus expressing the mycobacterial antigens 85A, 85B and TB10.4 known as the Crucell-Ad35/AERAS-402 vaccine (Wilkie and McShane, 2015). However, previous exposure to the vector can reduce vaccine efficacy. Prophylactic live vaccines are also being developed, such as recombinant BCG VPM1002 and MTBVAC (Mendez-Samperio, 2016a; Mendez-Samperio, 2016b). Live vaccines can revert to a pathogenic form. This is particularly dangerous for immunocompromised individuals. Inoculating patients with subunit-based vaccines eliminates risks of reversion to virulence. Subunit vaccines include the H1 vaccine which combines 85B and ESAT-6 antigens, H4 which combines antigens 85B and TB10.4 and M72 combining antigens 39A and 32A. However, subunit vaccines lack intrinsic immunogenicity, and are seldom able to induce long-term immunity against diseases, necessitating multiple vaccinations and the inclusion of adjuvants.

The Phase I trial of AERAS-422, a recombinant BCG vaccine, studies were discontinued when two participants developed the Varicella-Zoster virus (Hoft et al., 2016). M72/AS01$_E$ is a candidate subunit vaccine initially deemed clinical safe in both healthy and TB-infected adults. However, during Stage II trials, many volunteers suffered local reactions at injection sites ending the study prematurely (Gillard et al., 2016). Another subunit vaccine, MVA85A, initially looked promising. Proposed as an adjunct to conventional BCG vaccination, it showed effective protection in animal models. Unfortunately, MVA85A did not show similar efficacy in healthy infants and adults (Ndiaye et al., 2015).

Given the success of peptide subunit vaccine candidates such as H4/IC31, peptide vaccines should be considered strong potential TB vaccines. H4/IC31 had clinically safe profiles in Phase I trials inducing a positive immune response in healthy adults and infants already vaccinated with BCG (Kagina et al., 2014). Peptide vaccines can be freeze-dried maintaining stability without a cold-chain (Li et al., 2014). Apart from low manufacturing costs (Slingluff, 2011), peptide vaccines typically also have higher safety profiles, due to the use of epitopes without reactogenic responses (Li et al., 2014).

In this work we sought to design an anti-TB epitope ensemble vaccine applying an emerging approach in computational vaccinology. An ideal vaccine would concentrate on highly conserved immunogenic epitopes with a wide population coverage. We have recently begun to exemplify the approach by identifying epitope ensemble vaccine against Hepatitis C (Molero-Abraham et al., 2013), influenza (Sheikh et al., 2016), and malaria (Damfo et al., 2017). Likewise, here we selected *M. tuberculosis* epitopes of proven immunogenicity that can be combined to form an effective, widely-applicable epitope ensemble vaccine, defining both a global vaccine and one focussing on East Africa.

2. Methods

2.1. Collection of Mycobacterium tuberculosis-specific epitopes

The Immune Epitope Database and Analysis Resource (IEDB; URL: http://www.iedb.org/) (Peters et al., 2005a; Peters et al., 2005b) was used to collect *M. tuberculosis*-specific epitopes. Epitope search terms: any disease, antigen ID: Mycobacterium tuberculosis 1773, human host, and positive T cell assays. MHC Class I and II data were collected separately. The number of MHC I and MHC II epitopes were narrowed by selecting antigens with four or more epitopes (CD8 + , Class I) or fifteen or more epitopes (CD4 + , MHC Class II).

2.2. Protein sequences and multiple sequence alignments

Protein sequences of selected antigens were retrieved from the NCBI (National Centre for Biotechnology Information; http://www.ncbi.nlm.nih.gov/). For replaced or removed sequences, the most recently updated sequence was used. The retrieved antigen sequences were run against the NCBI Reference Proteins (RefSeq) database using automated BLASTP (Altschul et al., 1997), with maximum hit sequences limited to 100 and restricting the search to *Mycobacterium tuberculosis*. Multiple Sequence Alignments (MSAs) were generated using COBALT (Papadopoulos and Agarwala, 2007).

2.3. Analysis of epitope sequence variability

Conserved epitopes were identified by determining variable versus conserved sequence positions by analysing each separate MSA using the Protein Variability Server, PVS hereafter (Garcia-Boronat et al., 2008) (URL: http://imed.med.ucm.es/PVS/). Sequence variability was masked and only fragments with a length greater or equal to 9 were selected. The Shannon entropy threshold was set to 0.5.

2.4. Prediction of HLA binding profile for conserved epitopes

IEDB was used to calculate the binding profiles of MHC Class I (http://tools.iedb.org/mhci/) and Class II (http://tools.iedb.org/mhcii/) highly conserved epitopes. HLA I reference set was used for MHC I epitopes (Weiskopf et al., 2013) and an HLA II reference set was used for MHC II epitopes (Greenbaum et al., 2011). For MHC Class I, epitopes were chosen with a percentile rank less than or equal to one. For MHC Class II, epitopes with a percentile rank greater than or equal to five were omitted.

2.5. Calculation of projected population coverage (PPC) and optimal epitope identification

The PPC of highly conserved epitopes were calculated using IEDB (http://tools.iedb.org/tools/population/iedb_input). MHC I and MHC II epitopes were then ranked by PPC. Epitopes were combined within each class to calculate PPC's for the World and East Africa. MHC I and II epitopes were also combined and the PPC calculated. Vaccine combinations were ranked by PPC's in descending order with criteria of 90% or higher for the world and 80% or higher for East Africa.

3. Results

3.1. MHC class I epitope selection and combination

Using IEDB, a total of 400 epitopes were obtained from 122 different TB antigens. Following Damfo, Reche (Damfo et al., 2017), we focussed on epitope-rich antigens containing four or more epitopes: 18 different antigens were identified containing a total of 259 epitopes (Fig. 1A). Of these, 151 were conserved, as determined by PVS (Garcia-Boronat et al., 2008). The binding profile and percentage population coverage of each epitope was calculated using IEDB. Epitopes were selected on the allele diversity of their binding profiles and a PPC of at least 20% for the World or East Africa (Table 1).

By choosing the epitope with the highest global PPC (MLLAVTVSL) and then selecting epitopes that bound different HLA I alleles, PPC values for epitope combinations were calculated. Combining MHC I