Infection, Genetics and Evolution 21 (2014) 315–319

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Common antigens prediction in bacterial bioweapons: A perspective for vaccine design

Varun Jaiswal, Rajinder S. Chauhan, Chittaranjan Rout *

Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh 173234, India

ARTICLE INFO

Article history:
Received 2 September 2013
Received in revised form 10 November 2013
Accepted 11 November 2013
Available online 1 December 2013

Keywords:
Protein vaccine candidates (PVCs)
Computational vaccine design
Bioweapons (BWs)
Epitope
BW broad-specific vaccine design

ABSTRACT

Bioweapons (BWs) are a serious threat to mankind and the lack of efficient vaccines against bacterial bioweapons (BBWs) further worsens the situation in face of BW attack. Experts believe that difficulties in detection and ease in dissemination of deadly pathogens make BW a better option for attack compared to nuclear weapons. Molecular biology techniques facilitate the use of genetically modified BBWs thus creating uncertainty on which bacteria will be used for BW attack. In the present work, available resources such as proteomic sequences of BBWs, protective antigenic proteins (PAPs) reported in Protegen database and VaxiJen dataset, and immunogenic epitopes in immune epitope database (IEDB) were used to predict potential broad-specific vaccine candidates against BBWs. Comparison of proteomes sequences of BBWs and their analyses using in-house PERL scripts identified 44 conserved proteins and many of them were known to be immunogenic. Comparison of conserved proteins against PAPs identified six either as PAPs or their homologues with a potential of providing protection against multiple pathogens. Similarly, mapping of conserved proteins against experimentally known IEDB epitopes identified six epitopes which had exact epitope match in four proteins including three from earlier predicted six PAPs. These epitopes were also reported to provide protection against several pathogens. In the backdrop of conserved heat shock GroEL protein from *Salmonella enterica* providing protection against five diverse bacterial pathogens involved in different diseases, and synthetic proteins produced by combination of epitopes from *Mycobacterium tuberculosis* and 4 viruses providing protection against both bacterium and viruses, the identified putative immunogenic conserved proteins and immune-protective epitopes can further be explored for their potential as broad-specific vaccine candidates against BBWs.

1. Introduction

Organisms with potential to be used as bioweapons (BWs) have unique importance despite their low or rare involvement during pandemic in normal situations. By and large, the anthrax infection is a rare event in humans but it was the major concern after 2001 terrorist attack in USA. Intelligence have estimated that BW threat is greater than the nuclear weapons because of their ease in dissemination and difficulties in detection of deadly pathogens (D'Agostino and Martin, 2009). Lack of appropriate and efficient licensed vaccines against the bacterial bioweapons (BBWs) further worsens the situation in face of BW attack. Although vaccines are available for some BBWs, they still have limitations. For example, anthrax vaccines have serious side effects and they require yearly boosters (Weiss et al., 2007). Vaccine against *Francisella tularensis* is not fully licensed and the data about efficacy of plague vaccine is not available (Jefferson et al., 1998).

Despite low risk of infection by BBWs, uncertainty remains on which bacteria should be used in BW attack. The use of genetically modified BBWs in attacks adds uncertainty on the nature of BBWs (D'Agostino and Martin, 2009). Therefore, there is a need for developing broad-specific vaccines which can provide immunization against most of the BBWs. This can be achieved through producing a vaccine formulation containing conserved protective antigen(s) or immunogenic epitope(s) from several BW organisms. Such vaccines will have immense potential to provide protection against bacterial BW organisms. Recently, it has been observed that a single antigen (conserved heat shock protein (HSP)) from *Salmonella enterica* can provide protective immunity against multiple bacteria such as *Shigella flexneri*, *Shigella boydii*, enteropathogenic *Escherichia coli* (*EPEC*), *Klebsiella pneumonia* and *Pseudomonas aeruginosa* that even cause different diseases (Chitradevi et al., 2013). Experimental findings have shown that synthetic proteins produced by combination of epitopes from *Mycobacterium tuberculosis* and 4 viruses (vesicular *Stomatitis virus*, *Sendai virus*, respiratory *Syncytial virus* and lymphocytic *Choriomeningitis virus*) can provide protective immune response against infection caused by all these organisms (An and Whitton, 1997).

* Corresponding author. Tel.: +91 1792239306; fax: +91 1792245362.
E-mail addresses: computationalvarun@gmail.com (V. Jaiswal), rajinder.chauhan@juit.ac.in (R.S. Chauhan), chittaranjan.rout@juit.ac.in (C. Rout).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.11.011

According to Centre for Disease Control and Prevention (CDC) the potential BBWs and their associated diseases are given as follows: *Bacillus anthracis* (anthrax); *Brucella abortus*, *Brucella canis*, *Brucella melitensis* and *Brucella suis* (brucellosis); *Vibrio cholerae* (cholera); *Burkholderia mallei* (glanders); *Burkholderia pseudomallei* (melioidosis); *Clostridium botulinum* (botulism); *E. coli* (food poisoning); *S. enterica* (salmonellosis); *F. tularensis* (tularemia); *Coxiella burnetii* (Q fever); and *Yersinia pestis* (plague). Taking into consideration the facts that conserved protective antigens from an organism or immunogenic epitopes or combination of both from one or more organisms may provide broad-specific protection against diverse group of diseases, the current work intended to exploit the available resources such as genome sequences of BBWs, protective antigenic proteins (PAPs) reported in Protegen database (Yang et al., 2011) and VaxiJen dataset (Doytchinova and Flower, 2007), and immunogenic epitopes accumulated in immune epitope database (IEDB) (Vita et al., 2010) to predict potential conserved vaccine candidates across BBWs which can be formulated as broad-specific vaccines against BBWs (Fig. 1).

2. Materials and methods

2.1. Data collection

With an objective to predict common immunogenic proteins and epitopes across all BBWs, the proteome sequence files of fourteen BBWs (*B. anthracis* str. Ames, *B. abortus* bv. 1 str. 9-941, *B. canis* ATCC 23365, *B. melitensis* bv. 1 str. 16 M, *B. suis* 1330, *B. mallei* ATCC 23344, *B. pseudomallei* K96243, *C. botulinum* A str. ATCC 3502, *E. coli* O157:H7 str. EDL933, *S. enterica* subsp. enterica serovar Choleraesuis str. SC-B67, *V. cholerae* O395, *F. tularensis* subsp. tularensis strain FSC 198, *C. burnetii* strain RSA 331 and *Y. pestis* strain A1122) were downloaded from NCBI ftp site (Fig. 1). Standalone NCBI–BLAST (ftp://ftp.ncbi.nlm.nih.gov/BLAST/executables/BLAST+/LATEST/) was used to find out common proteins among all 14 selected BBWs proteomes as well as for other sequences comparisons required at different stages of analyses.

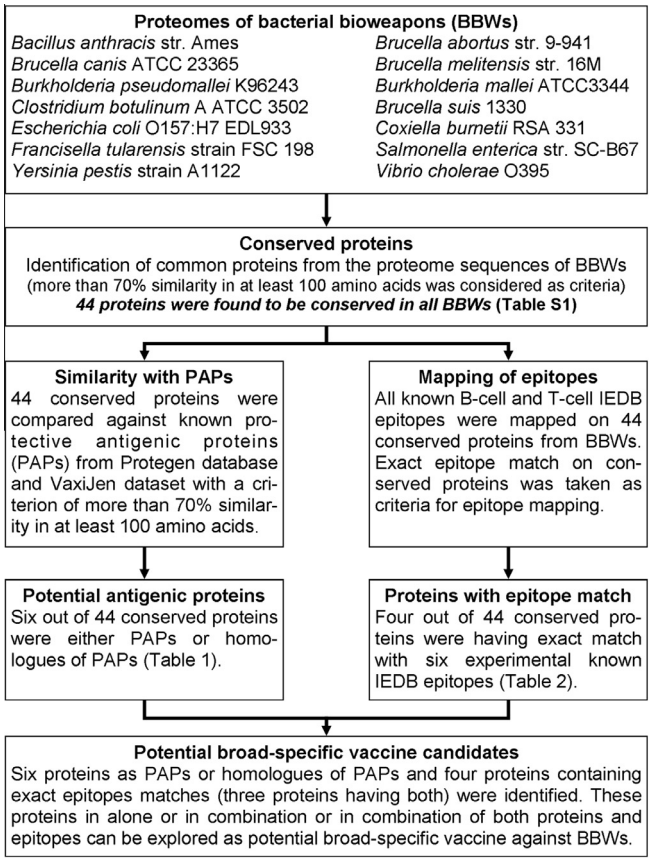

Fig. 1. Flow diagram for identification of broad-specific vaccine candidates (proteins and epitopes) from bacterial bioweapons (BBWs).

2.2. Identification of conserved proteins across BBWs

For identification of conserved proteins in all 14 selected bacterial BWs, the pathogenic strain sequences of 13 bacterial proteomes (mentioned in "Data collection" section) were taken as database and the *V. cholerae* O395 proteome was used as query sequences for BLAST search. The output of BLAST was then analyzed through in-house program to determine conservation of sequences in all 14 selected BBWs. The query proteins (proteome of *V. cholerae* O395) with >70% similarity in at least 100 amino acids in all 13 bacterial proteomes (taken as database) were considered as conserved or common proteins. If any of the proteome had more than one protein matching with the query sequence (according to the above criteria) then the protein with maximum significant match (minimum BLAST e-value) was considered as common protein. The functional similarity among common proteins was also manually verified.

2.3. Identification of putative conserved immunogenic proteins across BBWs

For the identification of conserved immunogenic proteins across BBWs, known protective antigenic protein (PAP) sequences reported in Protegen database (Yang et al., 2011) and positive dataset used for VaxiJen server development (Doytchinova and Flower, 2007) were retrieved. From the protein vaccine candidates (PVCs) reported for bacteria, viruses and eukaryotes in Protegen database, sequences of only bacterial PAPs were collected. Similarly, bacterial PVCs used in VaxiJen server development were also considered. 257 and 100 PVCs sequences from Protegen database and VaxiJen dataset (Jaiswal et al., 2013), respectively were compared against 44 conserved proteins across BBWs using BLAST. The outputs of BLAST were analyzed through in-house PERL script, and only those common proteins with $\geqslant$70% similarity in at least 100 amino acids length against sequences of bacterial protective antigens were considered as putative vaccine candidates which were further manually cross checked to confirm their functional similarity with the known protective antigens.

2.4. Mapping of IEDB epitopes to predict putative immunogenic regions in conserved proteins from BBWs

Experimentally known immunogenic epitope sequences of all T-cell epitopes (TCEs) and B-cell epitopes (BCEs) data, available as assay files at IEDB (http://www.iedb.org/), were downloaded (Vita et al., 2010). Peptide epitope sequences having quantitative measurement positive (positive, positive-low, positive-intermediate or positive-high) with literature reference, epitope ID, GI of source protein, and source and host organisms' information were extracted from these TCEs and BCEs assays. In case of TCEs, MHCs allele names were also extracted. All epitope sequences (BCEs and TCEs) were stored in 'fasta' format for comparison against 44 common proteins from BBWs using BLAST to predict their immunogenic regions. Exact match of epitope was considered as cut-off for epitopes mapping. In-house PERL script was used to perform BLAST search for all epitopes as well as analysis of output. Output of PERL script was manually checked to confirm the exact epitope