Vaccine 30 (2012) 7547–7560

Contents lists available at SciVerse ScienceDirect

Vaccine

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

Conservation of HIV-1 T cell epitopes across time and clades: Validation of immunogenic HLA-A2 epitopes selected for the GAIA HIV vaccine

Lauren Levitz[a], Ousmane A. Koita[d], Kotou Sangare[d], Matthew T. Ardito[a], Christine M. Boyle[a], John Rozehnal[b], Karamoko Tounkara[c], Sounkalo M. Dao[e], Youssouf Koné[c], Zoumana Koty[c], Soren Buus[f], Leonard Moise[a,b], William D. Martin[a], Anne S. De Groot[a,b,c,*]

[a] *EpiVax, Inc., Providence, Rhode Island, United States*
[b] *Institute for Immunology and Informatics, University of Rhode Island, United States*
[c] *GAIA Vaccine Foundation, Providence, Rhode Island and Bamako, Mali*
[d] *Laboratory of Applied Molecular Biology and Faculty of Medicine, University of Bamako, Mali*
[e] *Infectious Disease and Tropical Medicine Unit, Point G, Bamako, Mali*
[f] *University of Copenhagen, Denmark*

ARTICLE INFO

Article history:
Received 7 October 2011
Received in revised form
10 September 2012
Accepted 11 October 2012
Available online 24 October 2012

Keywords:
T cell epitopes
HLA-A2 epitope
HIV-1 vaccine
Immunoinformatics
GAIA vaccine

ABSTRACT

HIV genomic sequence variability has complicated efforts to generate an effective globally relevant vaccine. Regions of the viral genome conserved in sequence and across time may represent the "Achilles' heel" of HIV. In this study, highly conserved T-cell epitopes were selected using immunoinformatics tools combining HLA-A2 supertype binding predictions with relative global conservation. Analysis performed in 2002 on 10,803 HIV-1 sequences, and again in 2009, on 43,822 sequences, yielded 38 HLA-A2 epitopes. These epitopes were experimentally validated for HLA binding and immunogenicity with PBMCs from HIV-infected patients in Providence, Rhode Island, and/or Bamako, Mali. Thirty-five (92%) stimulated an IFNγ response in PBMCs from at least one subject. Eleven of fourteen peptides (79%) were confirmed as HLA-A2 epitopes in both locations. Validation of these HLA-A2 epitopes conserved across time, clades, and geography supports the hypothesis that such epitopes could provide effective coverage of virus diversity and would be appropriate for inclusion in a globally relevant HIV vaccine.

1. Introduction

The development of a safe and efficacious HIV vaccine is believed to be essential for stopping the AIDS pandemic [1–3]. Two major factors confounding vaccine design have been the extensive viral diversity of HIV-1 worldwide and the ongoing evolution and adaptation of virus sequences to HLA class I molecules driven by CD8+ cytotoxic T-cell (CTL)-mediated immune pressure [4,5]. In addition, the insufficient understanding of the complex roles of innate and adaptive immune responses in natural infection, as well as the immune correlates of protection, has made developing a vaccine capable of responding to these changes difficult. Indeed, the variability of HIV-1 may in part help explain the failure of recent HIV-1 candidate vaccines to elicit immune responses that recognize contemporaneous circulating virus stains. Neither the AIDSVAX vaccine [6–8], designed to generate antibody responses, nor the Merck AD5 [9,10], designed to raise T-cell responses, was able to prevent infection or alter disease among high-risk HIV-negative individuals. It has been suggested that these failures may be due to the inability of these vaccines to elicit cross-reactive broadly neutralizing antibodies and sufficient breadth and magnitude of T-cell responses at mucosal portals of entry [11–13]. The RV144 vaccine trial demonstrated modest success, leading to a 31% lowered rate of HIV-1 infection in a specific subset of vaccinees versus placebo groups [14]. While the correlates of immunity of that trial remain to be understood, viral diversity is likely to be at least partially responsible for the limited coverage.

HIV-1-specific CD4+ T helper cells and CD8+ cytotoxic T cells have been shown to play a central role in control of the virus following infection [15–21]. CD4+ T helper cells are essential for the generation of both humoral and cellular responses against the virus [22,23], while cytotoxic T cells play an important role in the resolution of acute viremia and in control of persistent HIV-1 viral replication [17,24]. Recent longitudinal studies following first CD8+ CTL responses to founder virus in early infection have defined a narrow window of opportunity for the CTL response to control infection and revealed multiple evolutionary pathways utilized by the virus during acute infection to retain replicative fitness [25–28]. Moreover, roles for both cytolytic function of CD8+ T cells during

* Corresponding author at: EpiVax, Inc., 146 Clifford Street, Providence, RI 02903, United States. Tel.: +1 401 272 2123; fax: +1 401 272 7562.
E-mail address: AnnieD@EpiVax.com (A.S. De Groot).

0264-410X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.vaccine.2012.10.042

nonproductive infection and noncytolytic functions (e.g., MIP-1β, MIP-1α, IFNγ, TNFα, and IL-1) in resolution of peak viremia have been identified [29,30]. Therefore, vaccines that stimulate virus-specific T-cell responses may be able to boost humoral immune responses and may also delay the progression of HIV-1 to AIDS in infected individuals. A robust T-cell response will be a necessary component of any successful HIV vaccine; however, the ability of a vaccine to account for the extraordinary viral diversity of HIV-1 continues to be a challenge. This diversity extends not only to T-cell epitope differences across clades, but also to isolates from a number of diverse clades that occupy a single geographic area [31].

One approach to address the problem of HIV-1 diversity is to develop multiple vaccines. These vaccines could be developed on a clade-by-clade basis, whereby a single vaccine represents isolates from a single clade, or on a geographically specific basis, whereby vaccines are derived from isolates commonly circulating in a particular country or region. However, this multiple vaccine approach raises the question of how many vaccines would be needed to protect against each of the many clades of HIV. In a time of increasing global connectedness and mobility, the notion of controlling a particular viral population and keeping it geographically sequestered is unlikely to bear fruit.

In contrast to region-specific vaccine efforts, our approach is to develop a globally effective vaccine. This vaccine would be comprised of epitopes targeting specific regions that are conserved across clades and regional variations, which are considered to be the most stable elements of the rapidly changing HIV-1 genome [32,33]. These regions may represent the "Achilles' heel" of the virus, as their persistence across time and space suggests they lie in regions of the HIV genome that may be resistant to selective immunologic pressure because they ensure viral fitness [34,35]. Other universal vaccine design strategies, such as the Mosaic Vaccine Constructs and Conserved Elements concepts currently undergoing preclinical studies, proffer global coverage based upon consensus plus most common variants and Center-Of-Tree derivation [36–39].

"Protective" HLA class I alleles are associated with CTL responses that target conserved regions of the viral genome located in functional or structural domains that, when mutated, impart a substantial fitness cost on the virus [40,41]. Population-based studies have shown that the number and rate of reverting mutations were highest in conserved residues in GAG, POL, and NEF (at equal frequency), while escape without reversion occurred in more variable regions [42]. Another study found that the highest fitness cost, based upon identification of reverting mutations across the entire HIV-1 subtype C proteome, occurred in target genes in the rank order VPR > GAG > REV > POL > NEF > VIF >TAT > ENV > VPU [42]. CD8+ CTL responses broadly targeting GAG have proven to be important in virus control as well as elite suppression in some individuals possessing "protective" HLA-B*57, HLA-B*5808, and HLA-B*27 alleles [43]. It could be argued that only epitopes that can undergo escape reversion mutations will elicit effective antiviral responses [44,45].

The biggest challenge for the rational design of an effective CD8+ T cell vaccine is the identification of HLA-class I-restricted immunodominant epitopes in HIV-1 that are under similar structural and functional constraint. Therefore, our strategy for HIV-1 vaccine design is to select epitopes that can induce broad and dominant HLA-restricted immune responses targeted to the regions of the viral genome least capable of mutation due to the high cost to fitness and low selective advantage to the virus. Both DeLisi and Sette have shown that epitope-based vaccines containing epitopes restricted by the six supertype HLA can provide the broadest possible coverage of the human population [46,47]. Thus epitopes that are restricted by common HLA alleles and conserved over time in the HIV genome are good targets for an epitope-based vaccine. Previously, we described the identification of 45 such HIV-1 epitopes for HLA-B7 [32], sixteen for HLA-A3 [48], and immunogenic consensus sequence epitopes representing highly immunogenic class II epitopes [49]. In this study, we focus on the identification and selection of highly conserved and immunogenic HLA-A2 HIV-1 epitopes. The goal is to provide valuable information and strategies that would contribute to the development of the GAIA vaccine or any other multi-epitope, pan-HLA-reactive, globally relevant HIV vaccine.

The HLA-A2 supertype allele is highly prevalent in much of the world, especially in those geographic areas under severe threat of HIV-1. It is common among Caucasian North Americans, but slightly less common in African American (20%) and Hispanic populations (34%) [50]. In China, where an HIV epidemic is beginning to emerge, HLA-A2 prevalence is 53.3% [51]. Among the African population, HLA-A2 frequency ranges from 36% to 63% with Mali, in particular, at 43% [52]. In this study, we present data using advanced immunoinformatics tools to identify highly conserved putative HLA-A2 epitopes for HIV-1. This analysis was conducted and epitopes were selected at two time points: first in 2002, and again in 2009. These two data sets allowed us to assess the persistence and conservation of the selected epitopes, as the number of available HIV sequences expanded four-fold over this time period. The immunogenicity of the 2002 and 2009 selected epitopes were confirmed with in vitro assays using blood from HIV-positive subjects in Providence, Rhode Island, and Bamako, Mali.

2. Materials and methods

2.1. Selecting a highly conserved HIV-1 sequence data set

2.1.1. 2002 sequence set

The sequences of all HIV-1 strains published on GenBank between January 1st, 1990, and June 2002 were obtained. Sequences posted to GenBank prior to December 31st, 1989, were excluded based on our observation that early sequences were more likely to be derived from HIV clade B. Sequences shorter than 80% and longer than 105% of a given protein's nominal length were also excluded. Short sequences were excluded because inclusion of these fragments skews the selection of conserved epitopes in favor of regions of particular interest to researchers, such as the CD4 binding domain or the V3 loop of HIV (unpublished observation). Longer sequences were excluded because these sequences tend to cross protein boundaries, confusing the categorization process. A second dataset was downloaded from the Los Alamos HIV Database using the same criteria, and the two datasets were merged. The combined 2002 dataset contained 10,803 unique entries selected for the next phase of analysis.

2.1.2. 2009 sequence set

In June–July 2009, the informatics component was repeated to assess the extent to which the predicted epitopes had been maintained in the expanding and evolving set of available viral sequences. In addition, the EpiMatrix algorithm had undergone revision which enabled it to be better at eliminating false positives (see Section 2.1.4 below); this updated EpiMatrix was employed to analyze the expanded sequence database. The same steps described above were repeated with the sequences posted between January 1st, 1990, and June 30th, 2009. All other inclusion criteria were unchanged. Due to the expansion of available HIV sequences, the combined dataset grew from 10,803 to 43,822 sequences. At this time we also performed a retrospective analysis of HIV sequences by year (Fig. 1) and selected additional epitopes (below).