ELSEVIER

Vaccine 21 (2003) 605–610

Vaccine

www.elsevier.com/locate/vaccine

Two years into reverse vaccinology

Jeannette Adu-Bobie, Barbara Capecchi, Davide Serruto, Rino Rappuoli*, Mariagrazia Pizza

IRIS, Chiron SpA, Via Fiorentina 1, 53100 Siena, Italy

Abstract

During the last century, several approaches have been used for the development of vaccines, going from the immunization with live-attenuated bacteria up to the formulation of the safer subunit vaccines. This conventional approach to vaccine development requires cultivation of the pathogen and its dissection using biochemical, immunological and microbiological methods. Although successful in several cases, this method is time-consuming and failed to provide a solution for many human pathogens. Now genomic approaches allow for the design of vaccines starting from the prediction of all antigens *in silico*, independently of their abundance and without the need to grow the microorganism in vitro. A new strategy, termed "Reverse Vaccinology", which has been successfully applied in the last few years, has revolutionized the approach to vaccine research. The *Neisseria meningitidis* serogroup B project, the first example of Reverse Vaccinology, as well as the application of this strategy to develop novel vaccines against other human pathogens are discussed.

Keywords: Vaccines; Reverse Vaccinology; *Neisseria meningitidis* B

1. Introduction

The approaches to vaccine development have made remarkable progress since the 18th and 19th centuries, when Edward Jenner and Louis Pasteur pioneered the use of attenuated and inactivated vaccines. Recently, there have been two major revolutions in vaccine design. The first revolution was the use of modern recombinant DNA technology to produce subunit vaccines based on specific antigens. In this approach, the pathogen is first studied to identify the factors important in the pathogenesis and immunity, and the identified factors that are produced in large scale by recombinant DNA. This approach has generated two very efficacious recombinant vaccines. The first was the hepatitis B vaccine based on a highly purified capsid protein [1], while the second was the acellular vaccine against *Bordetella pertussis* containing three highly pure proteins [2]. The latter vaccine also pioneered the use of structure–function studies to produce a genetically altered pertussis toxin that lacked toxicity but maintained an unaltered antigenic conformation [3]. Many of the vaccines developed during the last few decades are based on the subunit approach and contain one or more protective antigens. This conventional approach to vaccine development requires the pathogen to be grown in laboratory conditions, individual components to be identified and produced in a pure form, either directly from the bacterium or through recombinant DNA technology and then tested for their ability to induce immunity (Fig. 1). The approach is time-consuming and allows for the identification only of those antigens which can be purified in quantities suitable for vaccine testing. Since the most abundant proteins are most often not suitable vaccine candidates, and the genetic tools required to identify the less available components may be inadequate or not available, using this approach it could take decades to develop a vaccine. The conventional approach also means that vaccine development is not possible when the pathogen cannot be grown in laboratory conditions.

The second revolution in vaccine development occurred at the end of the 20th century as a result of the use of the genomic technology. The possibility of determining the complete genome sequence of a bacterium in a few months at low cost allowed the sequencing of the genome of most bacterial pathogens in a short period of time. Today, databases contain the complete genomic sequence of more than 80 bacteria, including most bacterial pathogens. More than hundred additional bacterial genomes are in progress. Large genomes of parasites such as malaria have been sequenced [4]. Powerful technologies such as genome sequencing, *in silico* analysis, proteomics (two-dimensional (2D) gel electrophoresis and mass spectrometry), DNA microarrays, in vivo expression technology (IVET) and signature tagged mutagenesis (STM) have revolutionized the way of studying bacterial pathogenesis and vaccine design. The availability of complete genome

* Corresponding author. Tel.: +39-577-243414; fax: +39-577-243564.
E-mail address: rino_rappuoli@chiron.it (R. Rappuoli).

0264-410X/02/$ – see front matter
PII: S0264-410X(02)00566-2

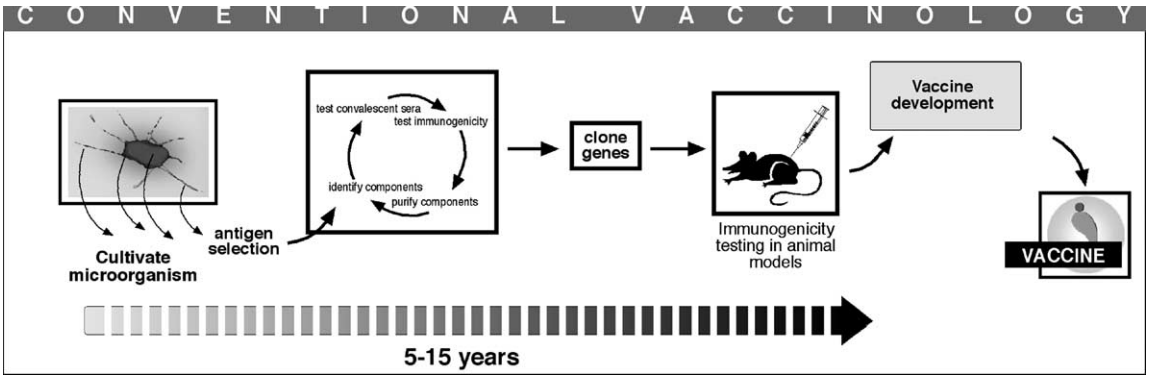

Fig. 1. Schematic representation of the conventional approach to vaccine development.

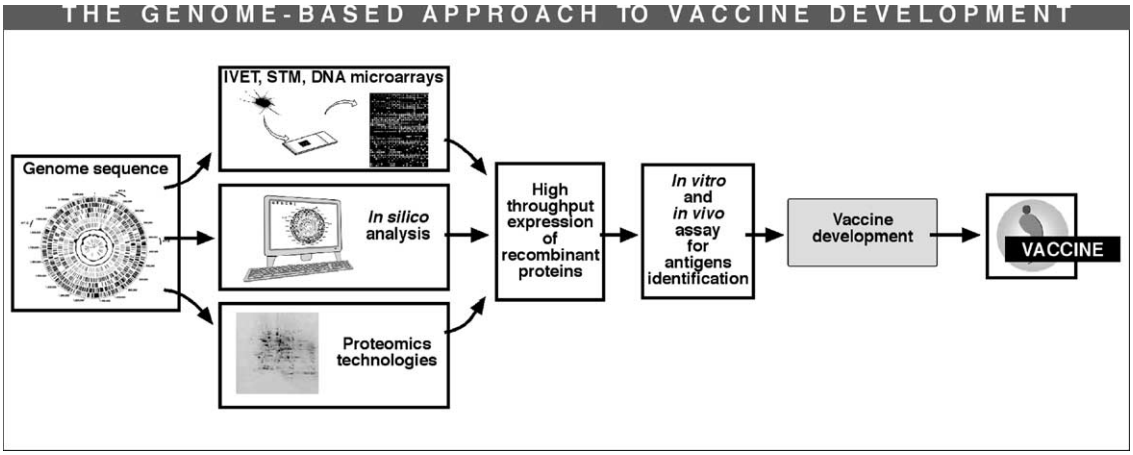

Fig. 2. The genome-based approach to vaccine development.

sequences and recent advances in molecular biology technology mean that every single antigen of a pathogen can be tested for its ability to induce a protective immune response. In addition, the use of microarray technology and proteomics provide us with more information on these potential antigens (Fig. 2).

Here we discuss the application of an approach coined "reverse vaccinology" on developing novel and effective vaccines against various human pathogens.

2. Meningococcus B: the pioneer for reverse vaccinology

Neisseria meningitidis is a major cause of meningitis and sepsis in children and young adults. While polysaccharide-based vaccines are available for A, C, Y and W135 serogroups, conventional vaccinology has failed for serogroup B. Group B meningococcus (MenB) represents the first example to which reverse vaccinology has been applied. MenB complete genome from strain MC58 was obtained by the random shotgun strategy [5]. While the sequencing was still in progress, the MenB genome was screened, using several softwares, in order to select putative ORF coding for surface-exposed or secreted proteins. Among the 2158 putative open reading frames (ORFs) annotated, 600 ORFs were selected on the basis of these criteria. These putative antigens included different classes of proteins, according to their predicted localization on the bacterial surface: outer membrane or secreted proteins, lipoproteins, inner membrane proteins, periplasmic proteins and also proteins with homologies to bacterial factors involved in virulence and pathogenesis [6].

The selected 600 ORFs were amplified from meningococcus by PCR, and cloned into *Escherichia coli* in order to express each gene as His-tag or GST fusion protein. Out of these 600 putative ORFs, 350 were successfully expressed, purified and used to immunize mice. Screening of immune sera was performed by Western blot on meningococcus total cell lysates and outer membrane vesicles to verify whether the protein was really expressed in meningococcus and to determine its subcellular localization. The surface-exposure of each antigen was then confirmed by fluorensce-activated cell sorter (FACS) analysis and ELISA on whole cell bacteria. Finally, sera were tested in bactericidal assay, an assay which is known to correlate with the protection in humans. Ninety-one proteins were found to be surface-exposed, 29 of them were able to induce bactericidal antibodies.

Most of the antigens identified by the conventional approach showed strain variability or were expressed only in some strains, resulting protective in humans against the homologous but not against the heterologous strains. Therefore, the candidate antigens selected by genome-analysis, were evaluated for gene presence and sequence conservation in a panel of 31 MenB strains isolated in different parts of the world, and representative of the major serogroups. Sequence alignment revealed that the majority of the antigens analyzed are well conserved while only a few of them present highly variable regions and carry multiple epitopes conserved in most strains. Finally, most of these antigens, tested in bactericidal assays, were able to induce cross-protection against heterologous strains. Surprisingly, the antigens identified by the genomic approach are quite different from those identified using conventional vaccinology. In fact, in addition to classical outer membrane proteins with variable loops, many of the new selected antigens were lipoproteins or surface-exposed proteins with a globular structure and without membrane crossing domains. Furthermore, some of these antigens are not abundant on the bacterial surface.

In conclusion, in only a few years reverse vaccinology has resulted in the identification of more vaccine candidates as compared to those discovered during the previous 40 years. Even if is too early to speculate when a vaccine against meningococcus will be available, now 15 very promising antigens are under investigation and are likely to enter into the development step (Fig. 3).

The use of genomic technologies allows to discover previously unknown and undescribed proteins. It is crucial to further characterize these novel molecules as vaccine candidates but also understand their role and function. Many of the novel outer membrane or surface-exposed proteins identified in MenB, share interesting homologies to known virulence factors [5,6]. Among these the newly-identified antigens GNA33, NadA and GNA992 have been further characterized from the biochemical and functional point of view.

GNA33 (genome derived *Neisseria* antigen) is a lipoprotein highly conserved among meningococcus serogroup B strains, other meningococcal serogroups and gonococcus. GNA33 shows 33% identity to a membrane bound lytic transglycolase (MltA) from *E. coli*. Biochemical analysis confirmed that the molecule is a murein hydrolase of the lytic transglycosylase class as it is capable of degrading both insoluble murein sacculi and unsubstituted glycan strands [7]. It has been shown that the recombinant GNA33 elicits antisera that are bactericidal and confer passive protection against bacteraemia in infant rats by mimicking a surface-exposed epitope on loop 4 of Porin A in strains with serosubtype P1.2 [8]. Epitope mapping of a bactericidal anti-GNA33 monoclonal antibody, identified a short motif (QTP) present in GNA33, which is essential for recognition. The QTP motif is also present in the loop 4 of PorA and was also found to be essential but not sufficient for the binding of the Mab to PorA [8].

NadA (*Neisseria* adhesin A, NMB1994) induces strong bactericidal antibodies against both homologous and heterologous strains, suggesting that this protein could be a good candidate for a vaccine [9]. Sequence analysis reveals that NadA is homologous to YadA, a non-pilus associated adhesin of enteropathogenic *Yersinia*, and to UspA2, an ubiquitous protein involved in serum resistance of *Moraxella catarrhalis*. In spite of the low level of the amino acid sequence homology, the three proteins show a well conserved secondary structure. These proteins have, in fact, a carboxyl terminal membrane anchor domain and an internal region with high coiled-coil probability. Interestingly, NadA forms very stable high molecular weight oligomers, like YadA and UspA2, and such oligomers are anchored to the outer membrane of meningococcus. The gene is present in three out of four hypervirulent lineages and its sequence is highly conserved among different strains. Furthermore, NadA is able to bind to human cells in vitro, suggesting that the structural homology with YadA is conserved also at the functional level.

GNA992 appears to be strictly related to two adhesins encoded by *Haemophilus influenzae*, Hsf and its allelic variant Hia, both involved in the formation of type b fibrils [10]. The elevated amino acid sequence similarity of GNA992 with Hsf and Hia (57 and 51% identity, respectively) and

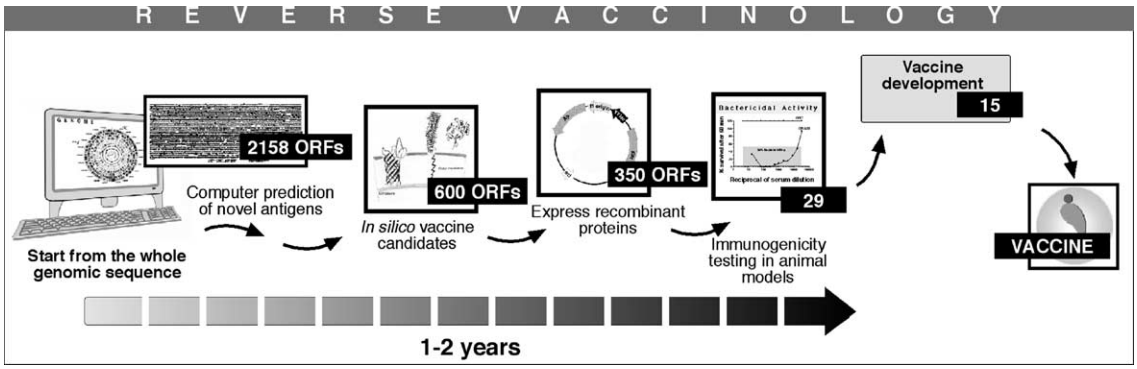

Fig. 3. The application of reverse vaccinology to identify vaccine candidates against *Neisseria meningitidis* serogroup B.

the similar topology of the three proteins suggest that they could share a common role in the mechanism of adherence.

The three adhesins have a modular structure and are composed of a different number of repeats. The core of these repetitive units has been identified and the conserved motif has been shown to be present in other adhesive molecules of *H. influenzae* (HMW1) as well as in human proteins belonging to the family of cell adhesin molecules (CAMs), such as NB-2 [11], which also share a conserved folding. On the basis of these common features, GNA992 has been postulated to promote adherence of meningococcus to host cells by mimicking the cell–cell recognition phenomena that occur at the neural level [12].

By FACS analysis and Western blot on outer membrane vesicles, GNA992 has been shown to be surface-exposed. Antibodies elicited by GNA992 are bactericidal against a subgroup of MenB strains and therefore this antigen is being regarded as a possible component of a multi component protein-based vaccine.

3. Other human pathogens: following in the footsteps of MenB

Progress in DNA sequence technology shows that it is now possible to obtain the genome map of any bacterium in a very short time. So far, the complete genome sequence of over 80 bacteria has been determined. The availability of these genomes indicates that the reverse vaccinology approach can be used on a wide scale. Following the success of the MenB paradigm other groups, in the last two years, have utilized this approach for the identification of vaccine candidates against main human pathogens.

3.1. Streptococcus pneumoniae

Streptococcus pneumoniae is a major cause of septicemia, pneumonia, meningitis and otitis media in young children in the US. In the US, efficacious conjugate vaccines against this pathogen are available. However, they are based on the polysaccharide capsule of only seven of the more than 70 serotypes known to cause disease in man. Koening and co-workers have exploited the whole genome sequence of *S. pneumoniae* in order to identify potential vaccine candidates [13]. All the 2687 open reading frames of the genome were evaluated to determine whether the gene products contained sequence motifs predictive of their localization on the surface of the bacterium. This led to the identification of 130 ORFs. Mice were immunized with 108 of these proteins and six conferred protection against disseminated *S. pneumoniae* infection. FACS analysis confirmed the surface localization of several of these antigens. The six protective targets were broadly distributed among pneumococcus strains and showed immunogenicity during human infection. Consequently, these new antigens may represent components of an improved vaccine against *S. pneumoniae*.

3.2. Porphyromonas gingivalis

The reverse vaccinology strategy was applied for the periodontal gram-negative pathogen *Porphyromonas gingivalis*, which is implicated in the etiology of chronic adult periodontitis. Using a series of bioinformatics methods Barr and co-workers have identified 120 genes from the genome of *P. gingivalis* [14]. The selected genes were cloned for expression in *E. coli* and screened with *P. gingivalis* antisera before purification. A set of 40 proteins were reactive by Western blot. They were purified and used to immunize mice, which were subsequently challenged with live bacteria in an abscess model. Interestingly, two antigens (showing homology to *Pseudomonas* sp. OprF protein) demonstrated significant protection in this mouse model of infection and therefore could represent potential vaccine candidates.

3.3. Chlamydia pneumoniae

Chlamydia pneumoniae is an obligate intracellular human pathogen that causes pneumonia and is also associated with atherosclerotic and cardiovascular disease. The bacterium has two distinct developmental phases, a spore-like infectious form, the elementary bodies (EBs), and an intracellular replicative form, the reticulate bodies. Because of the intrinsic difficulty in working with *C. pneumoniae* and the lack of adequate methods for its genetic manipulation, not much is known regarding protein components of the EB cell surface.

Grandi and co-workers have combined the reverse vaccinology strategy and proteome technologies in order to identify chlamydial surface antigens [15]. As a result of *in silico* analysis of *C. pneumoniae* genome, the authors identified 157 putative surface-exposed proteins. They used recombinant forms of these proteins expressed in *E. coli* to raise antisera, which were then utilized to assess surface location on the EBs by FACS analysis. Finally, 2D gel electrophoresis and mass spectroscopy were used to confirm the expression of the FACS-positive antigens in the elementary body phase of development. The result of this systematic genome–proteome combined approach represents the first successful attempt to define surface protein organization of *C. pneumoniae* and opens the way to the selection of suitable components of a novel vaccine.

3.4. Staphylococcus aureus

The work by Meinke and colleagues is another example of an approach which makes use of whole genome sequence information [16]. The availability of the genomic sequence of *Staphylococcus aureus* has allowed the development of a comprehensive approach for the identification of immunogenic proteins in this human pathogen. They made use of a system based on genomic peptide libraries. *S. aureus* peptides were displayed on the surface of *E. coli* via fusion to one or two outer membrane proteins (LamB and FhuA) and

Table 1
Comparison of conventional and genomic approaches to vaccine development

Conventional vaccinology	Reverse vaccinology
Most abundant antigens during disease	All antigens immunogenic during disease
Antigens expressed in vitro	Antigens expressed in vitro and in vivo
Cultivable microorganism	Antigens even in non-cultivable microorganisms
Animal models essential	Animal models essential
Correlates of protection useful	Correlates of protection essential
Structural components of microorganism	Non-structural components, including early proteins of viruses Correct folding in recombinant expression important High throughput expression/analysis important
Polysaccharides may be used as antigens	Non-proteic antigens cannot be used

probed with sera selected for high antibodies titer and opsonic activity. The whole screening of these libraries by magnetic cell sorting determines the profile of antigens, which are expressed in vivo and elicit an immune response in humans. They have identified a total of 60 antigenic proteins (most of which are located or predicted to be located on the surface of the bacterium) that could represent promising vaccine candidates for further evaluation.

3.5. Viruses

Many viral pathogens have quite small genomes whose sequence has been available for several years. However, the concept of reverse vaccinology is novel for viruses. The approach to vaccines against disease caused by viruses has been conventional, only structural antigens (envelope and core) have been usually considered until now.

Most attempts at designing a vaccine against HIV have been at the envelope glycoprotein (gp120, gp140 and gp160), and the core protein gag, using traditional approaches. However, these proteins are also highly variable due to immune system pressure. Only very recently promising data from experiments with Tat, Nef, Rev and Pol [17–19] strongly support the idea that the genome could provide information about potential antigens, which are either not part of the final viral particle or are present in such low quantities that they could not be purified and used as antigens by conventional strategies. Perhaps, if a reverse vaccinology approach had been taken to research into an HIV vaccine, a vaccine may already be available.

4. Conclusions

In the last two years, several positive results (obtained for meningococcus and other microorganisms) showed that the reverse vaccinology could likely be the solution to approach numerous human pathogens. The list of the pathogens where the traditional strategies to vaccine development have failed or provided only partial solutions is extensive. Conventional approaches in this field are time-consuming, identify only abundant antigens that may or may not provide immunity, and failed when the pathogen cannot be cultivated under laboratory conditions. Using reverse vaccinology, if a suitable assay to predict vaccine efficacy is available, every protein synthesized by the pathogen at any time can be tested as a vaccine candidate without any prior selection based on the limited knowledge of the pathogenesis of the microorganism and of the immune response induced following infection (Table 1). In conclusion, it is clear that genomic-based strategies to vaccine development has established a new paradigm in approaches to bacterial and viral pathogenesis. In this new context, it is expected that the number of candidate antigens for new vaccines will increase considerably in the near future.

References

[1] Andre FE. Overview of a 5-year clinical experience with a yeast-derived hepatitis B vaccine. Vaccine 1990;8(Suppl):S74–78, discussion S79–80.
[2] Greco D, Salmaso S, Mastrantonio P, et al. A controlled trial of two acellular vaccines and one whole-cell vaccine against pertussis. Progetto Pertosse Working Group. N Engl J Med 1996;334(6):341–8.
[3] Pizza M, Covacci A, Bartoloni A, et al. Mutants of pertussis toxin suitable for vaccine development. Science 1989;246(4929):497–500.
[4] Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium talciparum*. Nature 2002;419(6906):498–511.
[5] Tettelin H, Saunders NJ, Heidelberg J, et al. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science 2000;287(5459):1809–15.
[6] Pizza M, Scarlato V, Masignani V, et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 2000;287(5459):1816–20 [see comments].
[7] Jennings GT, Savino S, Marchetti E, et al. GNA33 from *Neisseria meningitidis* serogroup B encodes a membrane-bound lytic transglycosylase (MltA). Eur J Biochem 2002;269(15):3722–31.
[8] Granoff DM, Moe GR, Giuliani MM, et al. A novel mimetic antigen eliciting protective antibody to *Neisseria meningitidis*. J Immunol 2001;167(11):6487–96.
[9] Comanducci M, Bambini S, Brunelli B, et al. NadA, a novel vaccine candidate of *Neisseria meningitidis*. J Exp Med 2002;195(11):1445–54.
[10] St. Geme JW, Cutter D, Barenkamp SJ. Characterization of the genetic locus encoding *Haemophilus influenzae* type b surface fibrils. J Bacteriol 1996;178(21):6281–7.

[11] Ogawa J, Keneko H, Masuda T, Nagata S, Hosoya H, Watanabe K. Novel adhesion molecules in the Contactin/F3 subgroup of the immunoglubulin superfamily: isolation and characterization of cDNAs from rat brain. Neurosci Lett 1996;218(3):173–6.
[12] Scarselli M, Rappuoli R, Scarlato V. A common conserved amino acid motif module shared by bacterial and intercellular adhesions: bacterial adherence mimicking cell–cell recognition? Microbiology 2001;147:250–2.
[13] Wizemann TM, Heinrichs JH, Adamou JE, et al. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. Infect Immun 2001;69(3):1593–8.
[14] Ross BC, Czajkowski L, Hocking D, et al. Identification of vaccine candidate antigens from a genomic analysis of *Porphyromonas gingivalis*. Vaccine 2001;19(30):4135–42.
[15] Montigiani S, Falugi F, Scarselli M, et al. Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. Infect Immun 2002;70(1):368–79.
[16] Etz H, Minh DB, Henics T, et al. Identification of in vivo expressed vaccine candidate antigens from *Staphylococcus aureus*. Proc Natl Acad Sci USA 2002;99(10):6573–8.
[17] Cafaro A, Caputo A, Maggiorella MT, et al. SHIV89.6P pathogenicity in cynomolgus monkeys and control of viral replication and disease onset by human immunodeficiency virus type 1 Tat vaccine. J Med Primatol 2000;29(3/4):193–208.
[18] Osterhaus AD, van Baalen CA, Gruters RA, et al. Vaccination with Rev and Tat against AIDS. Vaccine 1999;17(20/21):2713–4.
[19] Pauza CD, Trivedi P, Wallace M, et al. Vaccination with Tat toxoid attenuates disease in simian/HIV-challenged macaques. Proc Natl Acad Sci USA 2000;97(7):3515.