Pakistan
Journal of Biological Sciences

Computational Comparison of T-cell Epitopes of gp120 of Iranian HIV-1 with Different Subtypes of the Virus

Hassan Mohabatkar and Reza Mohammadzadegan
Department of Biology, College of Sciences, Shiraz University, Shiraz 71454, Iran

Abstract: In the present study, T-cell epitopes of gp120 of an Iranian isolate have been compared to different subtypes of HIV-1. At first, the amino acid sequences of gp120 were fetched from data banks. Then T-cell epitopes, disulfide bonding states, protein kinase C phosphorylation sites, cAMP-dependent protein kinase phosphorylation sites, casein kinase II phosphorylation sites, N-myristoylation sites and amidation sites were predicted using different soft wares. According to this computational analysis 6 good disulfide binding states in Iranian gp120 were predicted. From the viewpoint of cAMP-dependent protein kinase phosphorylation site (1 site) Iranian isolate was similar to clades B and F. Like subtype C 1 amidation site was predicted in the Iranian subtype. In the Iranian isolate 7 sites protein kinase C phosphorylation sites and 4 N-myristoylation sites were predicted. Since the number of individuals infected with HIV-1 in Iran, like many other countries is increasing, study of similarities and differences between the Iranian samples and different clades of HIV-1 can help us in identification of the origin and understanding the changes in the virus.

Key words: T-cell epitopes, phosphorylation, amidation, N-myristoylation, disulfide bonds

INTRODUCTION

Different subtypes of (Human Immunodeficiency Virus type 1) HIV-1 predominate in different parts of the globe. HIV-1 subtype A, D and G are predominate in most of sub-Saharan Africa, subtype B in North America and Europe, subtype C in India and south Africa, subtype E in Thailand, subtype F in Brasilia and subtype H in some European countries (De Groot et al., 2002).

HIV-1 is a retrovirus that causes Acquired Immune Deficiency Syndrome (AIDS), a condition in which the immune system begins to fail, leading to life-threatening opportunistic infections (Podlasin et al., 2006). The envelope glycoprotein of HIV-1 consists of a complex of gp120 and gp 41 proteins which are good candidates for use in detection procedures and vaccine development. gp120 determines viral tropism by binding to target-cell receptors, while gp41 fusions between viral membrane and cellular membrane (Chan et al., 1997).

T-cells are vital components of the machinery of protective immunity, both directly by recognizing and eliminating the self-altered cells and indirectly by controlling the production of antibodies by B cells (De Groot et al., 2002). Cytotoxic T lymphocyte (CTL) epitopes are potential candidates for subunit vaccine design for various diseases (Bhasin and Raghava, 2004). CTLs recognize proteolysed fragments of the

protein in combination with MHC class I molecules (Hammerling et al., 1999; Watts and Dowis, 1999). They identify short peptides of 8-10 amino acids. Hence, identification of CTL epitopes is crucial in understanding the rules of T-cell activation and designing of synthetic vaccines (Brunak and Buus, 2000).

In Iran, HIV-1 infection was first documented in 1986 among five hemophiliacs who received infected blood factors (Iranian Disease Management Center, 2005); and subtypes A and B are the dominant HIV-1 subtypes circulating at present (Sarrami-Forooshani et al., 2006). In the last few decades, highly predictive patterns have occurred suggesting that local primary amino acid sequence can predict CTL epitopic sites. Here, in the present study, the CTL epitopic sites and some other feature of gp120 of Iranian and different subtypes of HIV-1 have been studied and compared; taking into consideration that predictions cannot be taken as ultimate conclusions but they give information to design new experiments (Chen et al., 2006; Mohabatkar and Kar, 2004). Considering two facts that accurate T-cell epitope prediction is necessary for developing computational peptide vaccines and immunodiagnostic methods and HIV-1 gp120 undergoes a lot of variation, the main objective of this study is prediction of T-cell epitopes of gp120 of Iranian HIV-1 and comparison of these epitopes with epitopes of different subtypes of the virus.

Corresponding Author: Hassan Mohabatkar, Department of Biology, College of Sciences, Shiraz University, Shiraz 71454, Iran
Tel: +9871l-6137373 Fax: +98711-2280916

MATERIALS AND METHODS

Amino acid sequences: Nine different HIV-1 gp120 sequences (from subtypes A-H and Iranian subtype) were fetched from the Protein Data Bank (Berman *et al.*, 2000). The amino acid sequence of gp120 of the Iranian virus (accession number AAU10463) was compared to amino acid sequence of different subtypes of the virus (Sarrami-Forooshani *et al.*, 2006).

Theoretical analysis: All prediction calculations were based on propensity scales for each of the 20 amino acids. The method integrates prediction of T-cell epitopes. Prediction was performed using artificial neural networks. In addition, other scales were used to predict the N-myristoylation sites, protein kinase C phosphorylation sites, Casein kinase II phosphorylation sites (Bairoch *et al.*, 2000), disulfide bonds (Vullo and Frasconi, 2004), number of predicted exposed residues, amidation sites, cyclic AMP sites and accessible residues percentages (Rost, 1996) of above-mentioned sequences. These analyses were conducted in Shiraz University, from October 2006 to March 2007.

RESULTS

Results of the computer-assisted comparison of the number of CTL epitopic sites of HIV-1 gp120 are shown in Fig. 1. It seems that there are some parts which are to some extant similar in most of the patterns, but the Iranian isolate is most similar to subtype G. On the other hand, it seems that Iranian and D subtypes are very different.

According to this computational analysis, the exposed residues of Iranian subtype were more than the exposed residues of other subtypes. Subtypes D and G had the least and the most predicted ratio of accessible residues, respectively (Table 1). From cAMP-dependent protein kinase phosphorylation site aspect of view, Iranian isolate was similar to clades B and F. In the Iranian isolate more than any s ubtype protein kinase C phosphorylation sites and less than any subtype N-myristoylation sites were predicted. It was interesting that Iranian glycoprotein was the only gp120 in which both cAMP-dependent protein kinase phosphorylation site and amidation site were predicted.

Other important differences between Iranian subtype and others were the number of cAMP and amidation sites. As in Iranian subtype on the contrary to other isolates, presence of both of the sites were predicted.
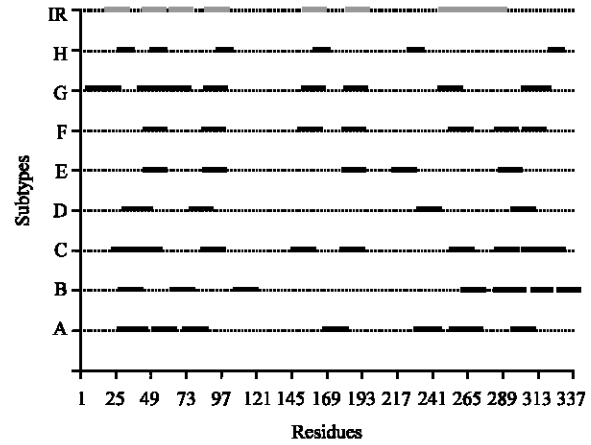

Fig. 1: Comparison of CTL epitopes of Iranian HIV-1 with subtype A-H of the virus

Table 1: Comparison of HIV-1 gp120 of Iranian isolate with gp120 of different subtypes of the virus

Subtype	PKC sites	CK2 sites	Myr sites	Dis sites	Expose Res.	Acces. Res. (%)	cAMP sites	Amidation sites
A	4	5	7	7	179	46.74	0	0
B	2	4	6	6	191	48.11	1	0
C	4	4	6	6	187	47.83	0	1
D	4	4	10	6	176	45.71	0	0
E	3	2	7	6	181	46.77	0	0
F	6	6	7	6	185	47.44	1	0
G	6	1	7	7	191	48.97	0	0
H	4	6	6	6	186	46.27	0	0
Iranian	7	3	4	6	192	47.76	1	1

PKC sites: The number of Protein kinase C phosphorylation sites, Ck2 sites: The number of Casein kinase II phosphorylation sites, Myr sites: The number of N-myristoylation sites, Dis sites: The number of disulfide bonds, Expose sites: The number of predicted exposed sites, Acces Res %: The accessible residues ratio, cAMPsites: The number of cyclic AMP phosphorylation sites, Amid sites: The number of amidation sites

DISCUSSION

The vast majority of the new HIV-1 infections that occur every day are acquired by individuals who live in developing countries, where the isolates of HIV-1 that are transmitted are significantly different from the isolates selected for most of the HIV-1 vaccines currently under development (Chan *et al.*, 1997).

CTL epitopes are potential candidates for subunit vaccine design for various diseases. Identification of CTL epitopes is crucial in understanding the rules of T-cell activation and designing of synthetic vaccines.

The purpose of this study was to use a specific bioinformatics approach to identify the Iranian HIV-1 T-cell epitopes from among a lot of epitope candidates. Almost certainly, this kind of studies would help the developing of the experimental methodologies by omitting non functional processes. As a result the present approach might pave the way to the concrete possibility of safe immunotherapy.

In view of the fact that co-translational and post-translational modifications of proteins can affect their functions, we use the computational analysis to compare the molecular properties of the Iranian subtype with other subtypes of the virus. In an investigation it was proved that in HIV a Nef mutant with alanine substitutions in myristoylation site elicited lower Nef-specific T lymphocyte responses than did the wild type Nef. A Nef mutant generated by simple alanine substitutions at myristoylation site had reduced immunogenicity (Liang *et al.*, 2002). It was reported also that, the amidation of the carboxy terminus of the natural determinant resulted in a peptide that exhibited both greater propensity to form a helix and the ability to stimulate the clone at concentrations approximately two orders of magnitude lower than the native sequence (Rothbard *et al.*, 1989). It has been shown that amino acid replacements which are associated with the emergence of alternative myristylation and casein kinase II (CKII) phosphorylation sites may affect the conformation and antigenicity of regulatory proteins in HIV (Oliveira *et al.*, 2004). In a research for developing vaccine it was concluded that the stability of MART-1 peptide was clearly rolonged by C-terminal amidation of the C-terminus. These modified peptides were recognized by CTL (Brinckerhoff *et al.*, 1999).

Identification of CTL epitopes has paved a way towards vaccine development and identification of many infectious diseases. Study of different sequences of HIV-1 is important, because an HIV-1 vaccine that includes specific T-cell epitopes is likely to be the most effective type of vaccine. CTL epitopes are potential candidates for subunit vaccine design for various diseases. Identification of CTL epitopes is crucial in understanding the rules of T-cell activation and designing of synthetic vaccines.

ACKNOWLEDGMENT

REFERENCES

Bairoch, A., P. Bucher and K. Hofmann, 2000. PROSITE. Nucleic Acids Res., 25: 217-221.

Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat and H. Weissig, 2000. The protein data bank. Nucleic Acids Res., 28: 235-242.

Bhasin, M. and G.P.S. Raghava, 2004. Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine, 22: 3195-3204.

Brinckerhoff, L.H., V.V. Kalashnikov, L.W. Thompson, G.V. Yamshchikov, R.A. Pierce and H.S. Galavotti, 1999. Terminal modifications inhibit proteolytic degradation of an immunogenic MART-1(27-35) peptide: Implications for peptide vaccines. Int. J. Cancer, 83: 326-334.

Brunak, S. and S. Buus, 2000. Identifying cytotoxic T-cell epitopes from genomic and proteomic information: The Human MHC Project. Rev. Immunogenet., 2: 477-491.

Chan, D.C., D. Fass, J.M. Berger and P.S. Kim, 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell, 89: 263-273.

Chen, C.C., J.K. Hwang and J.M. Yang, 2006. (PS) 2: Protein structure prediction server. Nucleic Acids Res., 34: 152-157.

De Groot, A.S., H. Sbai, C.S. Aubin, J. McMurry and W. Martin, 2002. Immuno-informatics: Mining genomes for vaccine components. Immunol. Cell Biol., 80: 255-269.

Hammerling, G.J., A.B. Vogt and H. Kropshofer, 1999. Antigen processing and presentation-towards the millennium. Immunol. Rev., 172: 5-9.

Iranian Disease Management Center, 2005. Quarterly Reports of Disease Management Center (DMC) (2003-2005). Health Ministry TGoIRoI.

Liang, X., T.M. Fu, H. Xie, E.A. Emini and J.W. Shiver, 2002. Development of HIV-1 Nef vaccine components: Immunogenicity study of Nef mutants lacking myristoylation and dileucine motif in mice. Vaccine, 20: 3413-3421.

Mohabatkar, H. and S.K. Kar, 2004. Prediction of exposed domains of envelope glycoprotein in Indian HIV-1 isolates and experimental confirmation of their immunogenicity in humans. Braz. J. Med. Res., 36: 675-681.

Oliveira, T., M. Salemi, M. Gordon, A.M. Vandamme, E.J. Van Rensburg and S. Engelbrecht, 2004. Mapping sites of positive selection and amino acid diversification in the HIV genome: An alternative approach to vaccine design? Genetics, 167: 1047-1058.

Podlasin, R.B., A. Wiercinska-Drapalo, A. Olczak, M. Beniowski, T. Smiatacz, E. Malolepsza, J. Juszczyk, M. Leszczyszyn-Pynka, T. Mach, M. Mian, B. Knysz and A. Horban, 2006. Opportunistic infections and other AIDS-defining illnesses in Poland in 2000-2002. Infection, 34: 196-200.

Rost, B., 1996. PHD: Predicting one-dimensional protein structure by profile-based neural networks. Meth. Enzymol., 266: 525-553.

Rothbard, J.B., R. Busch, R. Lechler, J. Trowsdale and J.R. Lamb, 1989. Recognition of the HLA class II-peptide complex by T-cell receptor: Reversal of major histocompatibility complex restriction of a T-cell clone by a point mutation in the peptide determinant. Philos. Trans. R. Soc. Lond. B. Biol. Sci., 323: 553-564.

Sarrami-Forooshani, R., S.R. Das, F. Sabahi, A. Adeli, R. Esmaeili and B. Wahren, 2006. Molecular analysis and phylogenetic characterization of HIV in Iran. J. Med. Virol., 78: 853-863.

Vullo, A. and P. Frasconi, 2004. Disulfide connectivity prediction using recursive neural networks and evolutionary information. Bioinformatics, 20: 653-659.

Watts, C. and S. Powis, 1999. Pathways of antigen processing and presentation. Rev. Immunogenet., 1: 60-74.