Comparison of antigenic sites of the envelope glycoprotein of the Iranian isolate of human T-cell leukemia virus type 1 with different subtypes of the virus

Hassan Mohabatkar, MSc, PhD, Sahar Sadeghi, MSc.

ABSTRACT

Objective: Human T-cell leukemia virus type 1 (HTLV-1) is an enveloped retrovirus, which is associated with a T-cell malignancy known as adult T-cell leukemia (ATL). Variation in the HTLV-1 envelope nucleotide sequence has been extensively documented and has been used to classify HTLV-1 isolates into different subtypes. The virus occurs in at least 3 subtypes, which have been named A, B, and C. We conducted this study to compare the antigenic proprieties of the Iranian isolate of HTLV-1 with the homologous region of different subtypes of the virus.

Methods: This study took place in the Department of Biology, College of Sciences, Shiraz University, Iran in 2005. The predicted antigenic sites and secondary structure of the envelope glycoprotein of HTLV-1, present in Iran, have been compared with the antigenic sites and secondary structure of the homologous domains in subtypes A, B, C of the virus. To predict the epitopes of glycoproteins, 21 different scales were used.

Results: The number of helices in the Iranian isolate was equal to the number of these regions in all 3 subtypes, but the number of ß-sheets was more than other viruses. One potential glycosylation site, on all these studied envelope glycoproteins, was predicted. Antigenic sites in the Iranian isolate were almost similar to subtype A of the virus and the Iranian isolate of HTLV-1 may be belongs to subtype A.

Conclusion: Our results indicate the similarities and differences between the Iranian and other subtypes of HTLV-1. Antigenic sites represent potential candidates for use in a peptide vaccine against HTLV-1 glycoproteins and since most of the properties of a particular protein depend on its structural properties, this type of study can help in better understanding of HTLV-1 isolates present in Iran.

Saudi Med J 2006; Vol. 27 (8): 1121-1124

Human T-cell leukemia virus type 1 (HTLV-1), the first human oncoretrovirus isolated in 1980[1] is the etiologic agent of adult T-cell leukemia (ALT),[2] and of a chronic progressive/HTLV-1-associated myelopathy (HAM) [tropical spastic paraparesis (TSP)/HAM).[3] Furthermore, this virus has been associated with cases of uveitis, polymyositis, and infective dermatitis in areas where HTLV-1 is endemic.[4] Among the 15-25 million HTLV-1-infected individuals living throughout the world, roughly 1-5% will develop ATL, or TSP/HAM, depending on as-yet-unknown cofactors. This could vary according to the geographical location.[4] The few nucleotide changes observed among strains were specific for the geographical origins of the patients. However, there are no consistent differences between strains from

From the Department of Biology, College of Sciences, Shiraz University, Shiraz, *Iran*.

Received 7th February 2006. Accepted for publication in final form 24th April 2006.

Address correspondence and reprint request to: Dr. Hassan Mohabatkar, Department of Biology, College of Sciences, Shiraz University, Shiraz 71454, *Iran*. Tel. +98 (711) 2280916. Fax. +98 (711) 2280916. E-mail: mohabatkar@susc.ac.ir

patients with ATL versus those from patients with TSP/HAM.[5,6] Based upon sequence or restriction fragment length polymorphism data from the polymerase and envelope genes or the long terminal repeat of more than 250 different strains, there are 3 major geographical subtypes (or genotypes),[7] which are strongly supported by high bootstrap values in phylogenetic analysis. Each of these genotypes (cosmopolitan, HTLV-1 subtype A; central African, HTLV-1 subtype B; and Melanesian, HTLV-1 subtype C) appears to have arisen from interspecies transmission between Simian T-cell leukemia virus type 1 (STLV-1)-infected monkey and humans followed by variable periods of evolution in the human host.[8] The presence of glycans seem to be necessary for the functional disposition of this glycoprotein and for protecting the polypeptide from denaturation.[9] In the last few decades, highly predictive patterns have occurred suggesting that local primary amino acid sequence can predict secondary structure of a protein.[10] It should be mentioned here that predictions can be taken as definitive conclusions but they give information helping to interpret results or to design new experiments.[11,12] There are several publications describing the presence of HTLV-1 in Mash-had, Iran.[13,14] Since most of the properties of a particular protein depend on its structure, the aim of this work was a computer-assisted study of differences and similarities between the secondary structure, N-glycosylation sites and antigenic sites of the envelope glycoprotein of the Iranian isolate of HTLV-1 and the homologous region of different subtypes of the virus. This is the first investigation of this type on the Iranian HTLV-1 using the approaches of bioinformatics.

Methods.

Envelop glycoprotein amino acid sequences. To study the secondary structure of the envelope glycoprotein of the Iranian isolate, the sequence of the protein from Swiss-Prot databank was fetched. Using Basic Local Alignment Search Tool (BLAST)[15] the amino acid sequence of this glycoprotein (accession No ACC58034) was compared with amino acid sequence of the protein of different subtypes of the virus.[16]

Theoretical analysis. Plots displaying regions of a protein with antigenic potential were generated with PREDITOP program written in Turbo-Pascal 5.5 for IBM computers.[17] All prediction calculations were based on propensity scales for each of the 20 amino acids. Nine scales of inverted hydrophobicity and 2 scales of hydrophilicity, which were mostly derived from the study of partition coefficient of amino acids in 2 non-interacting isotropic phases were taken into consideration. Four scales of accessibility, which were constructed by measuring the accessible surface of all the residences in a number of proteins, were considered. Three scales of secondary structure, which were based on the prediction of turns and troops obtained from statistical analysis of proteins of know structure, were considered for secondary structure prediction, also, 2 scales of flexibility. The 21 scales were grouped as follows:[17] 9 scales of inverted hydrophobicity (Doolittle, Heijne, Manavala, Prils, Rose, Sweet, Totls, Ges, Zimmermann); 2 scales of hydrophilicity (Hopp, Parker); 4 scales of accessibility (Janin, Chothia, Chothia,[8] Acrophil); one scale of antigenicity (Welling), and 2 scales of flexibility (Karplus, Ragone) and 3 scales of secondary structure (Chouf,[3] Garnier,[3] Levitt). The amino acid sequence of each protein was read as a moving window of 7 residues and the values corresponding to each of the 21 scales were taken into consideration and the mean was plotted against the fourth residue of the window. To compare the profiles obtained by different methods, various scales were normalized where the original values of each scale were set between +3 and -3. N-glycosylation sites are searched as Asn-X-Thr or Asn-X-Ser sequences, where X is any residue.[18,19]

Results.

The results of the computer-assisted comparison of N-glycosylation sites and secondary structure of HTLV-1 envelope glycoproteins are shown in **Table 1**. The results show that 2 regions predicted to be a-helix, and 8 regions predicted to be ß-sheet. According to this analysis 22.5%, 28.3% and 20.8% of the protein were in a-helix, ß-sheet and turn forms. The same analysis was performed to study the homologous regions of envelope glycoprotein of 3 different subtypes of HTLV-1. Like other subtypes, the Iranian isolate contains one potential glycosylation site only. The comparison of the antigenic determinants of the Iranian isolates with different subtypes (**Figure 1**) shows that all subtypes are similar to each other, but the Iranian isolate has the most similarity to subtype A of the virus. Clearly, some sites such as residues number: 17-25, 50-56, 60-66, 72-80, 85-92, 98-101, 125-135 and 170-172 are almost similar between the Iranian isolate and subtype A.

Discussion.

Several publications describe the presence of the HTLV-1 in Jewish individuals born in Mash-had, Iran.[20,21] Phylogenic analysis of the viral DNA sequence indicated that the HTLV-1 present in Mash-had belonged to the HTLV-1 subtype A clade.[22] The BLAST analysis[15] showed that there was almost 99.4% amino acid sequence homology between the Iranian and the other subtype A isolate (172 out of 173 amino acid). It is interesting that in all 4 studied sequences, the glycosylation site is

Table 1 - Prediction of secondary structure of the envelope glycoprotein of subtypes of human T-cell lymphotropic virus type 1.

Subtype	No. of helices	No. of sheets	Percentage of helices	Percentage of sheet	Percentage of turn	No. of N-glycosylation sites
A	2	7	24.9	26	21.4	1
B	2	6	26.6	22	21.4	1
C	2	6	24.9	21.4	24.3	1
Iranian A	2	8	22.5	28.3	20.8	1

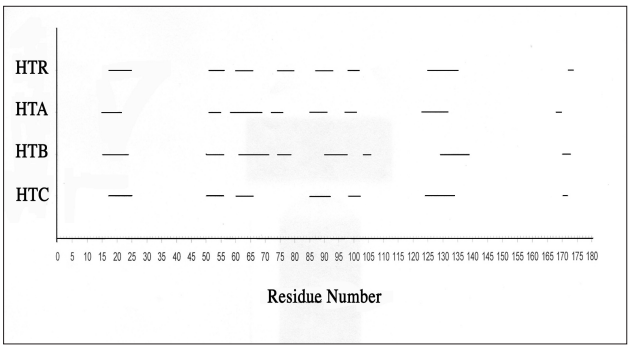

Figure 1 - Comparison of antigenic domains of Iranian human T-cell (HT) leukemia virus 1 with HT subtype A, HT subtype B, HT subtype C of the virus.

extremely conserved. One of the differences between the envelope glycoprotein of Iranian and non-Iranian subtype A of the virus is in the residue number 37. In this location, methionine has replaced isoleucine in the Iranian isolate. This replacement causes formation of one more ß-sheet conformation in the Iranian isolate. Like any other protein, prediction of secondary structure of HTLV-1 envelope glycoprotein can provide us important information regarding the interactions and functions of this protein.[23] The core of proteins usually contains a combination of helices and sheets, which are hydrophobic. In contrast, turns are situated on the surface of the protein in contact with solvent atoms. Thus, turns are accessible and hydrophilic.[17] In the Iranian isolate, the percentage of turns is less than other analyzed sequences (20.8%). The number of a-helices in the Iranian isolate is equal to the number of these structures in other subtypes, but the number of ß-sheets is more than other isolates (**Table 1**). Similarities and differences showed in **Table 1** could have a clear effect on the overall structure of the envelope glycoprotein of the Iranian isolate. It has already been shown that carbohydrates have a powerful effect on the antigenicity of proteins.[24] Like other subtypes, the Iranian isolate contains one potential glycosylation site. This similarity may play a significant role in the pathogenicity of the virus. The comparison of the antigenic determinants of the Iranian isolates and other scales (**Figure 1**) show that all subtypes are similar to each other, but the Iranian isolate has the most similarity to subtype A of the virus. It can confirm that the Iranian isolate belongs to subtype A of HTLV-1. Antigenic sites represent potential candidates for use in developing specific kits or peptide vaccines against viruses.[25,26] For non-homologous proteins, the reported methods exhibit a probability index, namely the percent of correctly predicted residues per predicted residues, of <70%. The limitation in the accuracy of predictions can be related essentially to the effect of long-range interactions specific for each protein family. The prediction of the tertiary structure of a protein is limited to the case of modeling a structure based on the known 3-dimensional structure of a homologous protein.[27] Since antigenic sites defined as continuous epitopes are strongly dependent on interactions with other parts of the overall tertiary fold, antigenicity predictions are plagued by the same problems encountered with other tertiary structure prediction algorithm.

Acknowledgment. The authors would like to thank Shiraz University for the support.

References

1. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Mina JD, Gallo RC. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. ***Proc Natl Acad Sci USA*** 1980; 77: 7415-7419.
2. Yoshida M, Miyoshi I, Hinuma Y. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. ***Proc Natl Acad Sci USA*** 1982; 79: 2031-2035.
3. Gout O, Baulac M, Gessain A, Semah F, Saal F, Peries J, et al. Rapid development of myelopathy after HTLV-I infection acquired by transfusion during cardiac transplantation. ***N Engl J Med*** 1990; 322: 383-388.
4. Muller N. The epidemiology of HTLV-1 infection. ***Cancer Causes Control*** 1991; 2: 37-52.
5. Yanagihara R. Geographic-specific genotypes or topotypes of human T-cell lymphotropic virus type I as markers for early and recent migrations of human populations. ***Adv Virus Res*** 1994; 43:147-186.

6. Vandamme AM, Liu HF, Goubau P, Desmyter J, Primate T. Lymphotropic virus type I LTR sequence variation and its phylogenetic analysis: compatibility with an African origin of PTLV-I. ***Virology*** 1994; 202: 212-223.
7. Saksena NK, Sherman MP, Yanagihara R, Dube DK, Poiesz BJ. LTR sequence and phylogenetic analyses of a newly discovered variant of HTLV-1 isolate from the Hagahai of Papua New Guinea. ***Virology*** 1992; 189: 1-9.
8. Ibrahim F, The G. de, Gessain A. Isolation and characterization of a new simian T-cell leukemia virus type 1 from naturally infected Celebes macaques (Macaca tonkeana): complete nucleotide sequence and phylogenetic relationship with the Australo-Melanesian human T-cell leukemia virus type 1. ***J Virol*** 1995; 69: 6980-6993.
9. Lasky LA, Groopman JE, Fennie CW, Benz PM, Capon DJ, Dowbenko DJ, et al. Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. ***Science*** 1986; 233: 209-212.
10. Alix AJ. Predictive estimation of protein linear epitopes by using the program people. ***Vaccine*** 1999; 18: 311-314.
11. Carmenes RS, Freije JP, Molina MM, Martin JM. Predict7 a program for protein structure prediction. ***Biochem Biophys Res Commun*** 1989; 159: 687-693.
12. Rooman MJ, Wodak SJ. Identification of predictive sequence motifs limited by protein structure data base size. ***Nature*** 1988; 335: 45-49.
13. Rafatpanah H, Pravica V, Farid R, Abbaszadegan MR, Tabatabaei A, Goharjoo A, et al. Association of a novel single nucleotide polymorphism in the human perforin gene with the outcome of HTLV-I infection in patients from northeast Iran (Mash-had). ***Hum Immunol*** 2004; 65: 839-846.
14. Sabouri AH, Saito M, Usuku K, Bajestan SN, Mahmoudi M, Forughipour M, et al. Differences in viral and host genetic risk factors for development of human T-cell lymphotropic virus type 1 (HTLV-1)-associated myelopathy/tropical spastic paraparesis between Iranian and Japanese HTLV-1-infected individuals. ***J Gen Virol*** 2005; 86: 773-781.
15. Ganesan N, Bennett NF, Velauthapillai M, Pattabiraman N, Squier R, Kalyanasundaram B. Web-based interface facilitating sequence-to-structure analysis of BLAST alignment reports. ***Biotechniques*** 2005; 2: 186-188.
16. Ellerbrok H, Fleischer C, Vandamme AM, Kucherer C, Pauli G. Sequence analysis of two HTLV type I infections imported to Germany. ***AIDS Res Hum Retroviruses*** 1997; 13: 1255-1258.
17. Pellequer JL, Westhof E, Van Regenmortel MD. Predicting location of continuous epitopes in proteins from their primary structures. ***Methods Enzymol*** 1991; 203: 176-201.
18. Bause E. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. ***Biochem J*** 1983; 209: 331-336.
19. Hubbard SC, Ivatt RJ. Synthesis and processing of asparagine-linked oligosaccharides. ***Annu Rev Biochem*** 1981; 50: 555-583.
20. Yamashita M, Kitze B, Miura T, Weber T, Fujiyoshi T, Takehisa J, et al. The phylogenetic relationship of HTLV type I from non-Mashhadi Iranians to that from Mashhadi Jew. ***AIDS Res Hum Retroviruses*** 1995; 11: 1533-1535.
21. Nerurkar VR, Achiron A, Song KJ, Melland RR, Pinhas-Hamiel O, Melamed E, et al. Human T-cell lymphotropic virus type I in Iranian-born Mashhadi Jews: genetic and phylogenetic evidence for common source of infection. ***J Med Virol*** 1995; 4: 361-366.
22. Safai B, Jun-Ling H, Farid R, Jahangir R, Schutzer P, Ahkami R, et al. Prevalence of HTLV-1 infection in Iran: a serological and genetic study. ***AIDS Res Hum Retroviruses*** 1996; 12: 1185-1190.
23. Englbrecht CC, Facius A. Bioinformatics challenges in proteinics. ***Comb Chem High Throughput Screen*** 2005; 8: 705-715.
24. Huang X, Barchi JJ Jr, Lung FD, Roller PP, Nara PL, Muschik J, et al. Glycosylation affects both the three dimensional structure and antibody binding properties of the HIV-1IIIB GP120 peptide RP135. ***Biochemistry*** 1997; 36: 10846-10856.
25. Mohabatkar H, Kar SK. Prediction of exposed domains of envelope glycoprotein in Indian HIV-1 isolates and experimental confirmation of their immunogenicity in humans. ***Braz J Med Biol Res*** 2004; 36: 675-681.
26. Sundaram R, Lynch MP, Rawale SV, Sun Y, Kazanji M, Kaumaya PT. De novo design of peptide immunogens that mimic the coiled coil region of human T-cell leukemia virus type-1 glycoprotein 21 transmembrane subunit for induction of native protein reactive neutralizing antibodies. ***J Biol Chem*** 2004; 279: 24141-24151.
27. Garnier J, Levin JM. The protein structure code: what is its present status? ***Comput Appl Biosci*** 1991; 7: 133-142.