J Biomed. 2016 September; 1(3):e8212. doi: 10.17795/jmb-8212.

Published online 2016 September 28. **Research Article**

A Potential In Silico Designed Chimeric Antigen for Precise Screening of HTLV-I and HTLV-II

Roxana Faramarzi,[1] and Samaneh Dolatabadi[1,*]

[1]Department of Microbiology, Neyshabur Branch, Islamic Azad University, Neyshabur, Khorasan Razavi, IR Iran

[*]*Corresponding author*: Samaneh Dolatabadi, Department of Microbiology, Neyshabur Branch, Islamic Azad University, Neyshabur, Khorasan Razavi, IR Iran. Tel: +98-5144446070, E-mail: s_dolatabadi1436@yahoo.com

Received 2016 August 11; **Accepted** 2016 August 29.

Abstract

Background: Separate expression and purification of peptides used in the diagnosis of infection with human T-cell lymphotrophic virus-I/-II (HTLV-I/-II) infection is time-consuming and costly.
Objectives: The aim of this study was to develop a single chimeric protein (CA) to be used in screening for infection HTLV-I/-II.
Methods: HTLV-I/-II antigen fragments were connected to each other using an amino acid linker and created a chimeric protein. Secondary and tertiary structures of this protein were predicted and assessed by Z-score calculation. The DNA sequence encoding the CA was optimized according to *E. coli* codon usage.
Results: The Z-score analysis showed that the three dimensional model has high quality. Linear epitopes on proteins are separately exposed on the CA surface and hence there is a possibility of their interactions with antibodies. Based on the computational analysis, this antigen has the chance to express at high levels in *E. coli* cells.
Conclusions: Based on bioinformatics analysis of this protein, it could be a suitable candidate for screening for infection with HTLV-I/-II.

Keywords: HTLV-I, HTLV-II, In Silico, Screening, Chimeric Protein

1. Background

Human T-cell lymphotropic virus types I and II (HTLV-I/-II) are retroviruses which are known to be causative agents for lymphopatic and neurological diseases. Infection with these viruses is endemic in some parts of the world such as South America, Japan, Iran and the Northeastern United States. Peoples who are infected with the virus have no clear clinical symptoms. Therefore, individuals can pass the virus to other people through common interactions such as sexual encounters, breastfeeding, and the use of shared syringes among drug users. The identification and education of carriers is the most important strategy that has been suggested for the prevention of infection with these viruses (1, 2).

The presence of specific antibodies in the blood indicates an HTLV infection. Peptides derived from proteins such as p19, gp46-I, gp21, and gp46-II can be used for antibody detection and by extension, HTLV infection screening. These peptides interact with HTLV antibodies in a specific manner (3, 4).

In this study, we used computational tools to design a novel antigen for the screening of HTLV contamination. In short, a chimeric protein containing dominant epitopes of HTLV-I/-II proteins was designed. The structure and potential reactivity of the chimeric protein was studied using in silico approaches. The results indicate that this antigen may be used as a cost-effective reagent for screening for infection with HTLV-I /-II (5, 6).

2. Methods

2.1. Study Design

This study was conducted in two stages. In the first stage, the dominant epitopes of HTLV-I and HTLV-II antigens were identified according to previous studies. In the second stage, a chimeric protein was designed and evaluated through structure prediction, epitope mapping; expression capacity.

2.2. Sequence Analysis

All the sequences used in this study were extracted from the database UniProt (http://www.uniprot.org/). Multiple sequences alignment was performed using the Muscle program (http://www.ebi.ac.uk/Tools/msa/muscle/).

2.3. Designed Chimeric Protein

The amino acid linker (sequence: EAAAK) was used to connect the epitopes together. Physico-chemical parameters including theoretical molecular weight (MW), isoelectric point (pI), and half-Life were calculated using Expasy Protparam (http://us.expasy.org/tools/protparam.html). The folding capacity of SCA was analyzed using FoldIndex online tools available at (http://bip.weizmann.ac.il/fldbin/findex). I-TASSER software (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) was used to predict the tertiary structure. The three dimensional structure of the protein was visualized by PyMol. To find possible errors in the models, PDB files were submitted to the Pro-SA server (https://prosa.services.came.sbg.ac.at/prosa.php) that is used to verify the tertiary structure. Solubility of the protein in *E. coli* cells were predicted using an online tools available in http://www.biotech.ou.edu/.

2.4. Optimization of Expression

The amino acid sequence of the chimeric protein was translated to a nucleotide sequence using a reverse translator (http://www.bioinformatics.org/sms2/rev_trans.html).The nucleotide sequence using JCat software (http://www.jcat.de/) was optimized based on the codon use *E. coli*. BamHI and EcoRI restrictive sites were added to the ends of the sequence for cloning purposes in a prokaryotic system.

3. Results

3.1. Antigen Selection

Epitopes of gp21(400-370), gp46 (215-165), p19 (130-100), and gp46-II (160-220) were used to design chimeric proteins. Based on the results of the comparison with Clustal W2, the sequence of these epitopes is conserved.

3.2. CA Design

In order to separate the epitopes, a linker consisting of EAAAK repeats was designed. Results of foldIndex indicate that the designed CA has folding capacity. The tertiary structure of CA is shown in Figure 1. The model has a C-value of -0.16. The C-value typically ranges from -5 to 2. A TM-score of 0.8 ± 0.13 confirms the accuracy of predicted structure (TM-score higher than 0.5, indicating the accuracy of the model topology). The Z-score analysis (Figure 2) indicates that the model has good quality. The localization of epitopes in CA are shown in Figure 3. It is shown that the link can easily separate sequences. The schematic structure of structures of the construct is shown in Figure 4. The molecular weight of the chimeric protein was calculated to be 24 kDa. The pI of protein was calculated to be 5. The half-life of protein in *E. coli* cells was more than 10 hours, which indicates that this protein has good stability. The results of the solubility prediction indicatethat the CA has no chance for soluble expression in *E. co*li cells.

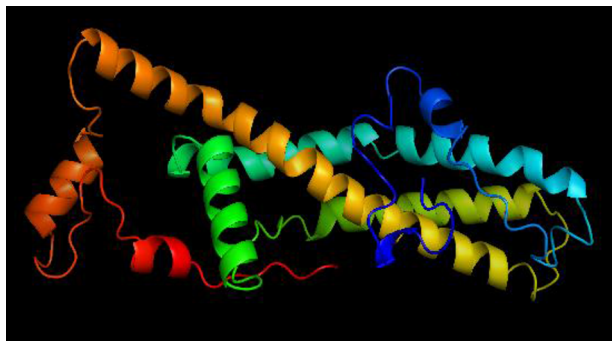

Figure 1. The tertiary structure of CA. The 3D structure predicted by I-TASSER software. The structure is shown in ribbon model.

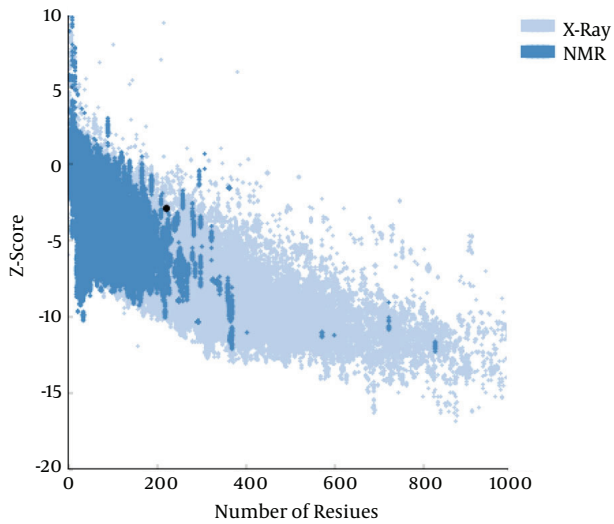

Figure 2. Z-score plot for tertiary structure of CA. The Z-score was used for model quality. The value of Z-score is highlighted as a black colored dot. The value is in the range of native conformations.

3.3. Optimization for Expression

The codon adaptation index (CAI) percent after optimization increased to 0.89, which indicates a high level expression chance. The guanine-cytosine content (GC) content changed to 55 percent. After optimization, almost 66% of the codons were the most frequently used codon.

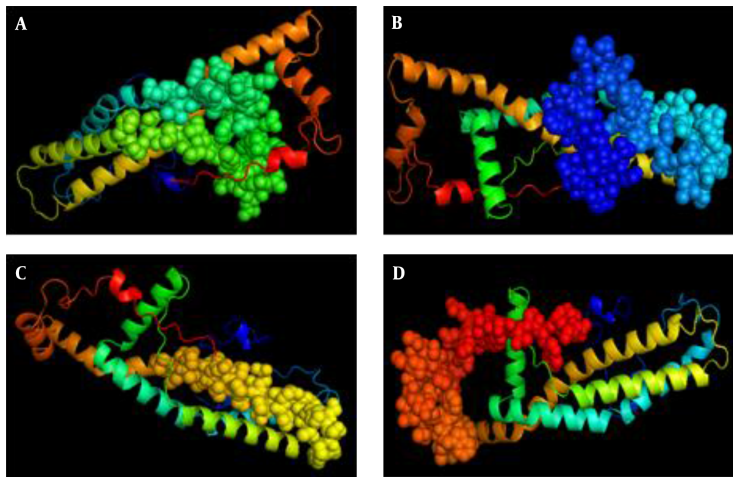

Figure 3. Localization of immunodominant peptides on the CA structure. A, gp46-I; B, gp21; C, p19; and D, gp46-II.

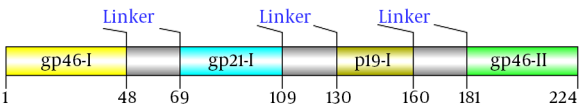

Figure 4. Graphical Display of CA. The Length of Fragment and Position of Linkers are Graphically Illustrated.

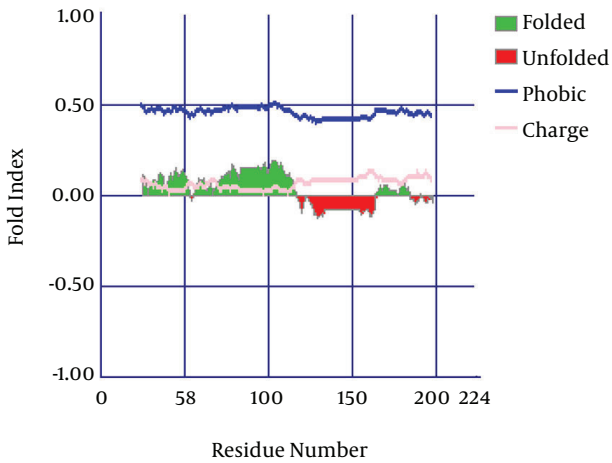

Figure 5. Disordered Region Prediction. The Threshold Value is Shown as Level Line.

There were no rare codons in the optimized sequence. The presence of rare codons (codons with lower than 30% usage) could hamper protein synthesis. These results indicate that CA could be overexpressed in *E. coli*.

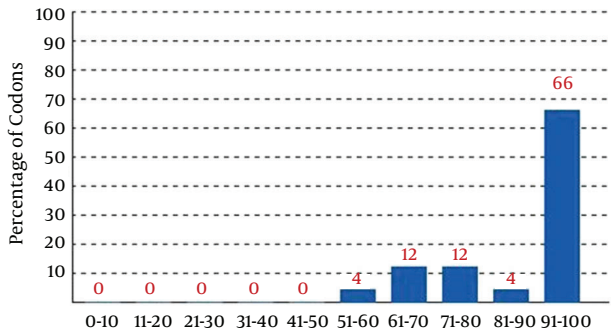

Figure 6. Rare codon analysis. There are no rare codons (codons with lower than 30% usage) in the optimized sequence.

4. Discussion

The use of peptides for detection of HTLV-I/-II antibodies has advantages, such as increasing its specificity, but the coating of peptides in the plate has problems. Peptides adsorbed to the plate by Van der Waals forces (7). These force are weak; therefore, more peptides are washed out during the washing steps. By increasing the length of the peptide, the adsorption of peptides to plates was increased. However, longer peptides also have high synthesis disadvantages (8, 9). Another way to produce peptides is the use of expression systems. However, the expression of small peptides using bacteria is extremely difficult. Usually fusion systems such as GST are used for peptide expression. For the detection of HTLV-I and HTLV-II infections, the application of at least three epitopes are necessary. Separate expression of these peptides is difficult and time-consuming. Also, antibodies against the GST tag are found

in the blood of some healthy people. The presence of such antibodies may cause false positive reaction in the screening. So, the ELISA systems usually benefit from horseradish peroxidase (HRP) conjugated chimeric protein instead of secondary antibodies to remove all false positive results (10, 11).

Several chimeric proteins were designed for the screening of HTLV-I/-II in previous studies. They use gp46, gp21, and p19 epitopes for the design of chimeric proteins. In this study, we used gp46-II epitopes in the construction of chimeric protein which allow for detection of HTLV-I and HTLV-II antibodies simultaneously.

An order of epitopes is one of the most important factors influencing chimeric protein structure and function. The primary sequence of protein affects its function by changing the tertiary structure. Linker also affects the quality of epitopes exposure. Four repeats of (EAAK) amino acid linker were used to connect the fragments (12). In previous studies, this linker is used to separate peptide fragments. This linker forms an elongated structure, which causes proper separation and exposure of epitopes (13).The evaluation of epitope exposure on protein surfaces creates the need for secondary and tertiary structure prediction. A TM-score of 0.8 ± 0.13 confirms the accuracy of the predicted structure, indicating the accuracy of the model topology. Also, the Z-score analysis confirms accuracy of the 3D model. The results of the evaluation show that all epitopes are located on the surface and antibodies can easily react with it (14).

One way to increase gene expression in bacteria is codon usage optimization. Rare codon frequency, GC, and CAI are criteria for analysis of expression capacity of sequence after optimization. Lack of rare codon, CAI above 0.8, and GC content of 55% indicate that this protein could be expressed in bacteria in high levels (15).

The results of solubility indicate that the antigen has a low solubility chance. There are various strategies for solubility enhancement of proteins in *E. coli*, such low temperature induction, low IPTG concentration, or solubility tags. With combination of these strategies, designed antigen could be expressed in a soluble form (16).

4.1. Conclusion

In this study, a chimeric antigen for detecting HTLV infection using antigen-antibody system was introduced. This antigen is insoluble and has multiple sites for binding to the antibody.

References

1. Takatsuki K. Discovery of adult T-cell leukemia. *Retrovirology.* 2005;**2**:16. doi: 10.1186/1742-4690-2-16. [PubMed: 15743528].
2. Lairmore MD, Montgomery A. Isolation and confirmation of human T-cell leukemia virus type 2 from peripheral blood mononuclear cells. *Methods Mol Biol.* 2005;**304**:113–23. doi: 10.1385/1-59259-907-9:113. [PubMed: 16061970].
3. Rafatpanah H, Hedayati-Moghaddam MR, Fathimoghadam F, Bidkhori HR, Shamsian SK, Ahmadi S, et al. High prevalence of HTLV-I infection in Mashhad, Northeast Iran: a population-based seroepidemiology survey. *J Clin Virol.* 2011;**52**(3):172–6. doi: 10.1016/j.jcv.2011.07.004. [PubMed: 21840754].
4. Kalavi K, Moradi A, Tabarraei A. Population-based Seroprevalence of HTLV-I Infection in Golestan Province, South East of Caspian Sea, Iran. *Iran J Basic Med Sci.* 2013;**16**(3):225–8. [PubMed: 24470867].
5. Proietti FA, Carneiro-Proietti AB, Catalan-Soares BC, Murphy EL. Global epidemiology of HTLV-I infection and associated diseases. *Oncogene.* 2005;**24**(39):6058–68. doi: 10.1038/sj.onc.1208968. [PubMed: 16155612].
6. Zarnagh HH, Ravanshad M, Pourfatollah A, Rasaee M. Expression and purification of a novel computationally designed antigen for simultaneously detection of htlv-1 and hbv antibodies. *Iranian J Allergy, Asthma Immunol.* 2015;**14**(2):168.
7. Davis GD, Elisee C, Newham DM, Harrison RG. New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnol Bioeng.* 1999;**65**(4):382–8. [PubMed: 10506413].
8. Thorstensson R, Albert J, Andersson S. Strategies for diagnosis of HTLV-I and -II. *Transfusion.* 2002;**42**(6):780–91. [PubMed: 12147033].
9. Zhang Y, Song S, Liu C, Wang Y, Xian X, He Y, et al. Generation of chimeric HBc proteins with epitopes in E.coli: formation of virus-like particles and a potent inducer of antigen-specific cytotoxic immune response and anti-tumor effect in vivo. *Cell Immunol.* 2007;**247**(1):18–27. doi: 10.1016/j.cellimm.2007.07.003. [PubMed: 17707782].
10. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins.* 2006;**65**(1):40–8. doi: 10.1002/prot.21078. [PubMed: 16894596].
11. Frangione-Beebe M, Albrecht B, Dakappagari N, Rose RT, Brooks CL, Schwendeman SP, et al. Enhanced immunogenicity of a conformational epitope of human T-lymphotropic virus type 1 using a novel chimeric peptide. *Vaccine.* 2000;**19**(9-10):1068–81. [PubMed: 11137241].
12. Dipti CA, Jain SK, Navin K. A novel recombinant multiepitope protein as a hepatitis C diagnostic intermediate of high sensitivity and specificity. *Protein Expr Purif.* 2006;**47**(1):319–28. doi: 10.1016/j.pep.2005.12.012. [PubMed: 16504539].
13. Carlson ED, Gan R, Hodgman CE, Jewett MC. Cell-free protein synthesis: applications come of age. *Biotechnol Adv.* 2012;**30**(5):1185–94. doi: 10.1016/j.biotechadv.2011.09.016. [PubMed: 22008973].
14. Weickert MJ, Doherty DH, Best EA, Olins PO. Optimization of heterologous protein production in Escherichia coli. *Curr Opin Biotechnol.* 1996;**7**(5):494–9. [PubMed: 8939627].
15. Baneyx F. Recombinant protein expression in Escherichia coli. *Curr Opin Biotechnol.* 1999;**10**(5):411–21. [PubMed: 10508629].
16. Vasina JA, Baneyx F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. *Protein Expr Purif.* 1997;**9**(2):211–8. doi: 10.1006/prep.1996.0678. [PubMed: 9056486].