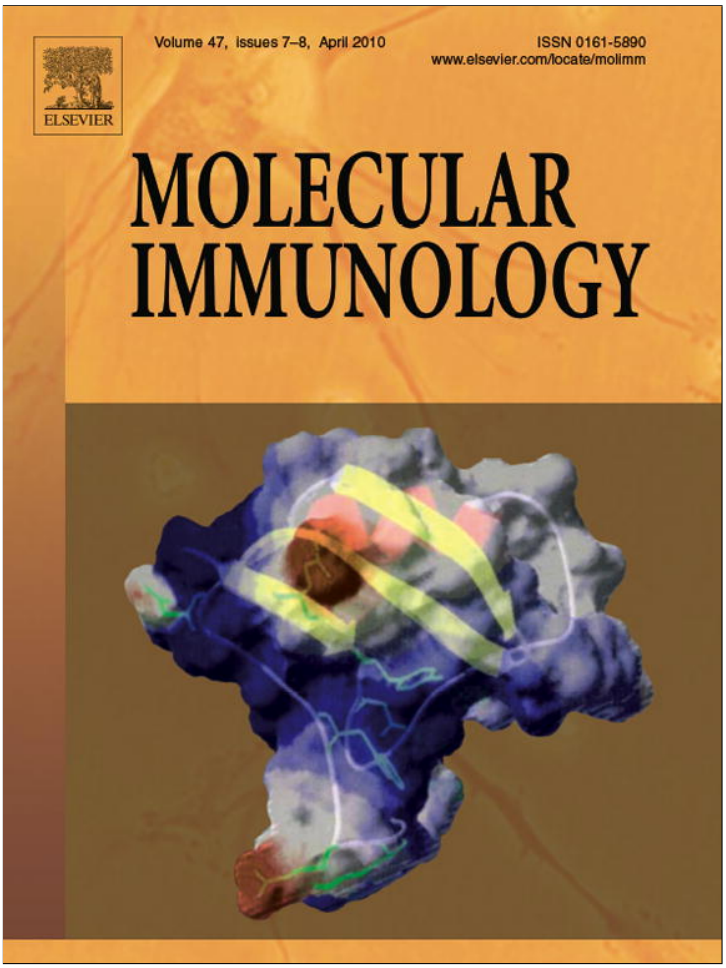
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Molecular Immunology 47 (2010) 1635–1639

Contents lists available at ScienceDirect

Molecular Immunology

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

Short communication

Mutational pressure makes HIV1 gp120 linear B-cell epitopes shorter and may lead to their disappearance

Vladislav Victorovich Khrustalev*

Department of General Chemistry, Belarussian State Medical University, 83 Dzerzinskogo Prospect, Minsk 220000, Belarus

ARTICLE INFO

Article history:
Received 6 November 2009
Received in revised form 7 January 2010
Accepted 10 January 2010
Available online 4 February 2010

Keywords:
HIV1
gp120
Linear B-cell epitopes
Mutational pressure
Acrophilicity
Acrophobicity
APOBEC1
APOBEC3

ABSTRACT

We showed that nucleotide sequences coding for linear B-cell epitopes of human immunodeficiency virus type 1 (HIV1) gp120 protein are enriched with codons containing cytosine and guanine in their first and second codon positions. Guanine and cytosine are the most mutable nucleotides in HIV1 genes (due to APOBEC3 and APOBEC1 editing of viral DNA and RNA, respectively, as well as due to reverse transcriptase preference to incorporate 8-oxo-G against C). We introduced all the possible G to A, C to U, C to A and G to U single nonsynonymous nucleotide mutations in gp120 coding region from the HIV1 reference strain. The BepiPred algorithm (www.cbs.dtu.dk/services/BepiPred) was used for the linear B-cell epitopes predictions. Results of this "in-silico directed mutagenesis" showed that: (i) single nonsynonymous G to A transitions will cause partial or complete destruction of linear epitopes in 18% of 229 possible cases; (ii) single nonsynonymous C to U transitions will cause partial or complete destruction of linear epitopes in 58% of 142 possible cases; (iii) single nonsynonymous C to A transversions will cause partial or complete destruction of linear epitopes in 28% of 184 possible cases; (iv) single nonsynonymous G to U transversions will cause partial or complete destruction of linear epitopes in 37% of 240 possible cases. Moreover, single transition of C to U direction leading to amino acid replacement inside an epitope will cause partial or complete destruction of this epitope at a probability of 98%.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The process of human immunodeficiency virus (HIV) immune escaping attracts a lot of attention among researchers in the field of virology, microbiology and immunology. In this study we decided to answer the question how exactly gp120 B-cell linear epitopes change their structure during HIV1 mutagenesis.

B-cell epitopes are parts of proteins or other molecules recognized and bound by antibodies produced by B-cells. There are two types of B-cell epitopes: discontinuous or 3D epitopes and linear or continuous epitopes. For approximately 10% of the 3D epitopes the corresponding antibodies are cross-reactive with a linear peptide fragment of the epitope (Larsen et al., 2006).

Certain types of nucleotide mutations occur more frequently than other types in the given genome (Sueoka, 1988). The most frequent nucleotide mutation in HIV genes is G to A transition (Berkhout and de Ronde, 2004). This mutation is thought to be caused by cytosine deamination in DNA minus strand of the viral genome (Pillai et al., 2008). The process of cytosine deamination in singlestranded DNA is catalyzed by several cellular enzymes from APOBEC3 family (Izumi et al., 2008).

The catalytic deamination of cytosine can also take place in HIV RNA. It has been shown in several experimental works that cellular enzymes from APOBEC1 family deaminate cytosine in HIV1 RNA causing C to U transitions (Petit et al., 2009). Although, this process is less frequent than "G to A hypermutagenesis" (Bishop et al., 2004).

During reverse transcription HIV1 reverse transcriptase incorporates an oxidized guanine (8-oxo-G) preferably in front of cytosine residues (Kamath-Loeb et al., 1997). The most of cellular polymerases, including RNA-polymerases, preferably incorporate adenine in front of 8-oxo-G (Kamath-Loeb et al., 1997). If 8-oxo-G incorporates in DNA minus strand, it will lead to C to A transversion in HIV RNA. Guanine residues can be oxidized directly in HIV DNA minus strands leading to the same result: C to A transversions (Gros et al., 2002).

So, we can conclude that there are three main directions of nucleotide mutations in HIV genes: G to A transitions, C to U transitions and C to A transversions. To study complete set of mutations forming AT-pressure we also simulated results of nonsynonymous G to U mutations. During "in-silico directed mutagenesis" (Khrustalev, 2009) we introduced all the possible single nonsynonymous nucleotide mutations (795 mutations) of GC to AT

* 7-24 Communisticheskaya Street, Minsk 220029, Belarus.
Tel.: +375 80172845957; fax: +375 80172845957.
E-mail address: vvkhrustalev@mail.ru.

0161-5890/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2010.01.006

Author's personal copy

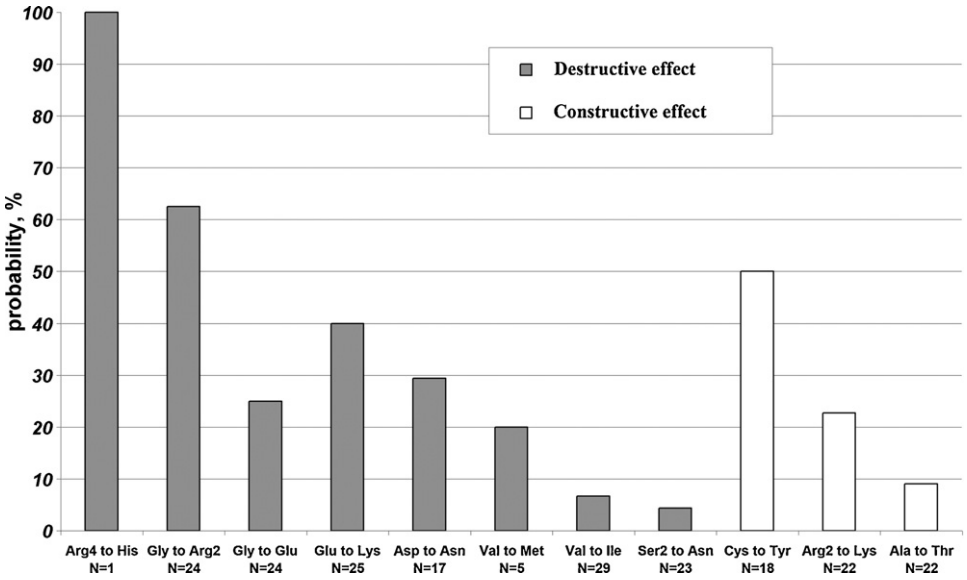


Fig. 1. Probability of destructive and constructive effects on the length of HIV1 gp120 linear B-cell epitopes for amino acid replacements caused by G to A transitions.

direction in the district coding for gp120 from HIV1 reference strain and evaluated their effects on the length of linear B-cell epitopes with the help of BepiPred algorithm (Larsen et al., 2006).

The BepiPred algorithm created by Larsen et al. (2006) includes a novel method of B-cell linear epitopes prediction based on hidden Markov model as well as on a new propensity scale combining antigenicity, hydrophilicity, inverted hydrophobicity, accessibility and secondary structure. One of the three data sets used for the creation of BepiPred algorithm itself consists of epitopes found in HIV proteins (Larsen et al., 2006).

We found out that C to U transition is the most destructive type of nonsynonymous mutation for the length and quantity of linear B-cell epitopes. About 63.3% of C to U nonsynonymous transitions in the region coding strictly for an epitope will make this epitope shorter, while about 34.7% of them will destroy this epitope completely. Probably, this "radical" way of immune escaping is the most beneficial one for the virus.

2. Material and methods

As the material we used nucleotide sequence of *env* gene coding for gp120 protein from the reference strain of HIV1 (GenBank accession number: NC_001802).

Nucleotide usage has been calculated with the help of our CodonChanges computer algorithm (www.barkovsky.hotmail.ru).

During the session of "in-silico directed mutagenesis" (Khrustalev, 2009) we analyzed results of all possible single nonsynonymous G to A transitions (229 mutations), C to U transitions (142 mutations), C to A transversions (184 mutations) and G to U transversions (240 mutations) by BepiPred algorithm (Larsen et al., 2006).

We distinguished between effects of amino acid substitutions inside the predicted epitopes (246 substitutions) and in epitope free regions flanking those epitopes (549 substitutions). Certain amino acid replacements in epitope flanking regions have a serious influence on the structure of epitopes situated nearby.

Except the estimation of probabilities at which the given type of nucleotide substitution will lead to epitope appearance, disappearance, elongation or to the reduction of its length, we counted average lengths of newly occurred epitopes, of the disappeared epitopes, of the increase and of the decrease of their length.

3. Results

3.1. Nucleotide content of the regions coding for predicted linear B-cell epitopes of HIV1 gp120

The average level of guanine (G) in parts of HIV1 *env* gene region coding for epitopes of gp120 is 24.6%, while the average level of G in parts of the same gene coding for epitope free regions is 19.8%. Interestingly, the level of G in third (neutral) codon positions is higher in parts of *env* gene coding for epitope free regions of gp120 than in parts of this gene coding for epitopes (16.3% versus 12.2%).

The average level of cytosine (C) is also much higher in parts of *env* gene coding for epitopes (22.0%), than in parts of this gene coding for epitope free regions (14.7%). The level of C in first codon positions is approximately the same (12.2% versus 12.8%), while the level of cytosine in second and third codon positions is much higher in parts of *env* gene coding for epitopes (2C = 31.3%; 3C = 22.6%), than in parts of this gene coding for epitope free regions (2C = 17.9%; 3C = 13.3%).

Indeed, predicted epitopes of gp120 are enriched with amino acids encoded by codons containing G and C in first and second codon positions. Levels of glycine (GGX) and proline (CCX) are 3.4 and 4.7 times higher, respectively, in epitopes relatively to epitope free regions of gp120. Epitopes are also enriched with serine (TCX and AGT/C), threonine (ACX), glutamic (GAT/C) and aspartic (GAA/G) acids, as well as with asparagine (AAT/C).

3.2. Consequences of G to A transitions

Results of nonsynonymous G to A mutations are shown in Fig. 1. About 62.5% of amino acid substitutions from Gly to Arg will lead to the decrease in length of the B-cell epitope. Two of these mutations will even lead to the complete disappearance of an epitope 5 amino acid residues in length. Amino acid substitutions from Gly to Glu, from Glu to Lys, from Asp to Asn will also lead to the decrease of the length of the epitope at the probabilities of 25.0, 40.0 and 29.4%, respectively (see Fig. 1). However, the "hidden" part of an epitope is always short (from 1 to 2 amino acid residues).

Cys to Tyr, Arg to Lys and Ala to Thr replacements are able to make epitopes a little bit longer. One of a 16 Cys to Tyr mutations in epitope free regions will lead to a new B-cell epitope formation.

Anyway, the probability of the "destructive" effect of amino acid replacements caused by G to A mutations on B-cell epitopes is

Author's personal copy

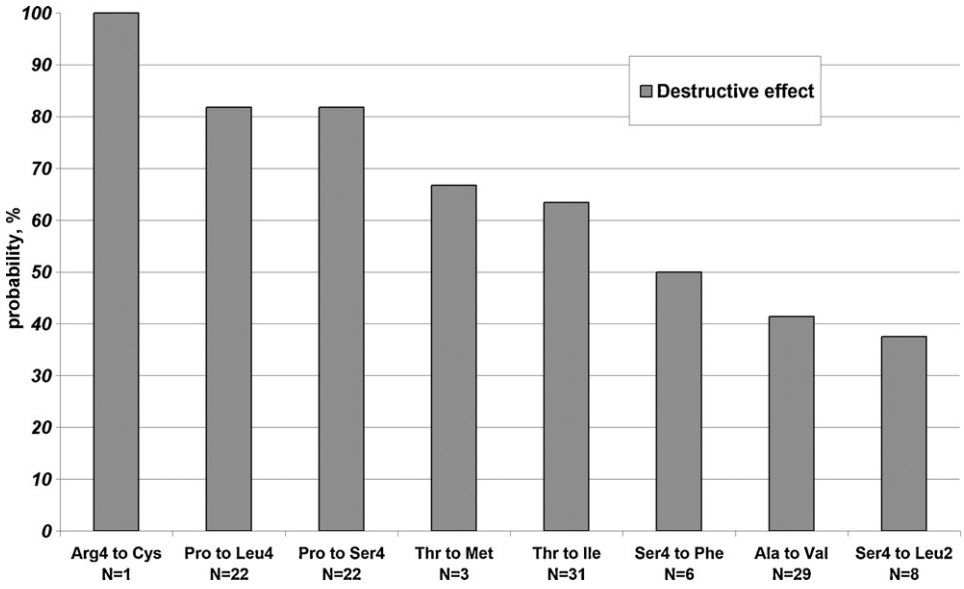


Fig. 2. Probability of destructive effect on the length of HIV1 gp120 linear B-cell epitopes for amino acid replacements caused by C to U transitions.

higher than the probability of their "constructive" effect. An average length of the disappeared part of an epitope is 1.22 ± 0.13 amino acid residues. An average length of the elongated part of an epitope is just 1.07 ± 0.13 amino acid residues.

3.3. Consequences of C to U transitions

Fig. 2 shows the consequences of single nonsynonymous C to U transitions. There can be only destructive consequences of this kind of mutation on the epitope length. All nonsynonymous C to U transitions leading to Pro to Leu4, Pro to Ser4, Ala to Val, Thr to Ile and Ser4 to Phe amino acid replacements inside an epitope will destroy this epitope partially or completely. Even relatively long linear B-cell epitopes (from 7 to 9 amino acid residues in length) can be completely destroyed by a single Pro to Leu substitution. A lot of nonsynonymous C to U transitions in parts of *env* gene coding for epitope free regions also make epitopes shorter and even make some of them disappear.

An average length of an epitope completely destroyed due to a single nonsynonymous C to U mutation is 5.00 ± 0.73 amino acid residues. An average length of the disappeared part of an epitope is 2.65 ± 0.46 amino acid residues.

3.4. Consequences of C to A transversions

As you can see in Fig. 3, nonsynonymous C to A mutations leading to the replacement of proline (Pro to Thr, Pro to His and Pro to Gln) have a strong destructive effect on the length of epitopes, as well as Ser4 to Tyr, Ser2 to Arg2 and Asn to Lys substitutions.

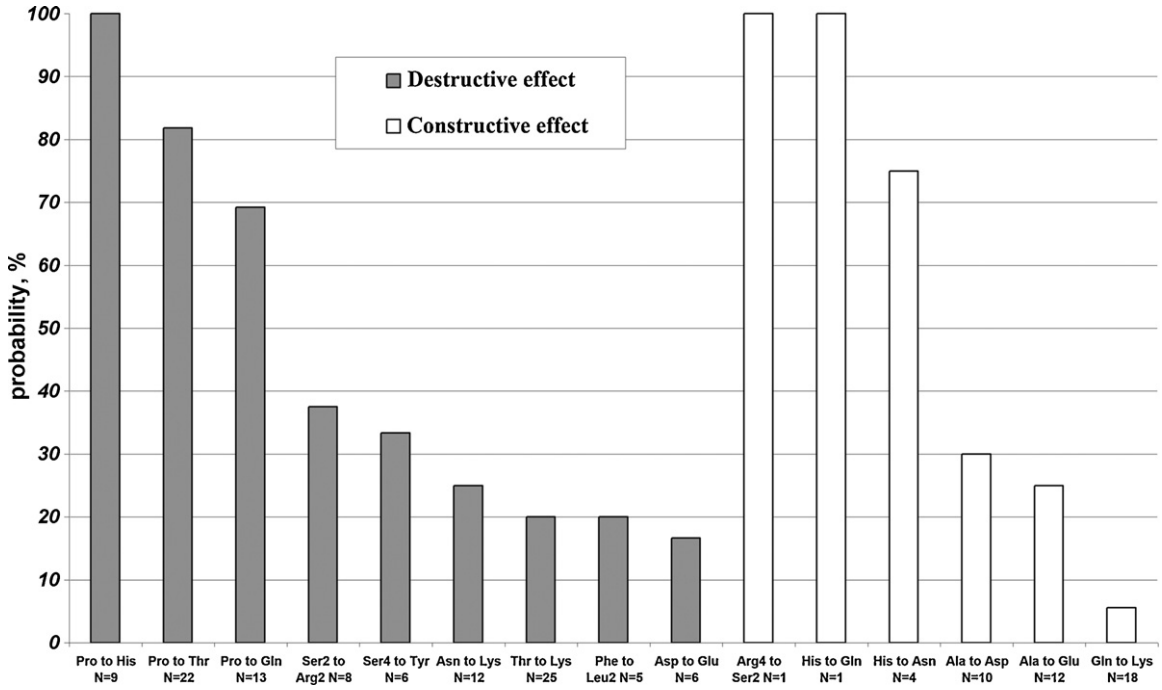


Fig. 3. Probability of destructive and constructive effects on the length of HIV1 gp120 linear B-cell epitopes for amino acid replacements caused by C to A transversions.

Author's personal copy

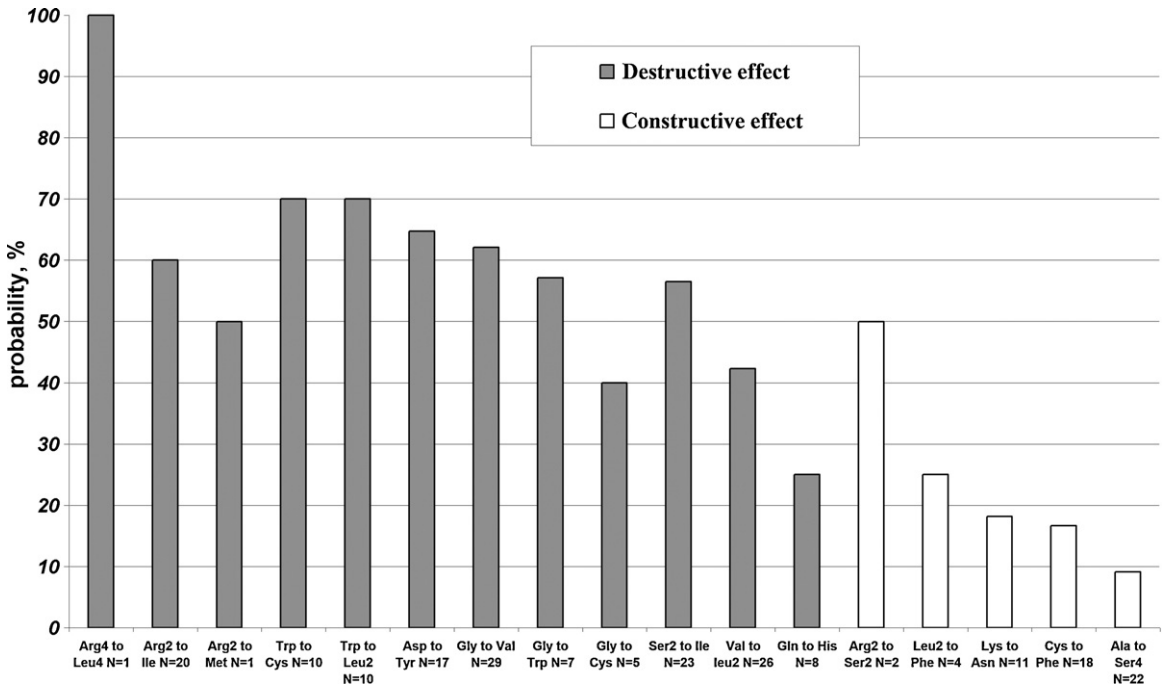


Fig. 4. Probability of destructive and constructive effects on the length of HIV1 gp120 linear B-cell epitopes for amino acid replacements caused by G to U transversions.

Table 1
Probabilities of the constructive ("+" effect) and destructive ("–" effect) consequences of nucleotide substitutions in the region of *env* gene coding for gp120 protein from the reference strain of HIV1 for the length of linear B-cell epitopes.

Type of nucleotide mutation	"+" effect, %	New epitope formation, %	Enlargement of existing epitope, %	"–" effect, %	Complete disappearance of epitope, %	Partial destruction of epitope, %
G to A	6.99	0.44	6.55	18.34	2.18	16.16
G to A inside a region coding for an epitope	7.35	–	7.35	36.76	4.41	32.35
C to U	0.00	0.00	0.00	58.45	19.72	38.73
C to U inside a region coding for an epitope	0.00	–	0.00	97.96	34.69	63.27
C to A	5.98	1.63	4.35	28.26	9.78	18.48
C to A inside a region coding for an epitope	1.67	–	1.67	60.00	20.00	40.00
G to U	3.75	0.42	3.33	37.08	5.42	31.67
G to U inside a region coding for an epitope	4.35	–	4.35	68.12	13.04	55.07

Two from 10 substitutions of alanine with aspartic acid will cause formation of a new B-cell epitope (1 and 3 amino acids in length). Substitutions of histidine with asparagine should also lead to the constructive effect at a high probability (75%). However, in general, C to A nonsynonymous transversions will lead to the destruction of linear B-cell epitopes.

An average length of an epitope completely destroyed due to a single nonsynonymous C to A mutation is 5.39 ± 0.86 amino acid residues. An average length of the disappeared part of an epitope is 2.68 ± 0.66 amino acid residues. The length of the elongated part of an epitope is strictly 1 amino acid residue.

3.5. Consequences of G to U transversions

Fig. 4 shows that appearance of acrophobic Val in place of acrophilic Gly will have a destructive effect on the epitope length at a probability of 62.1%. The same effect will take place in case of Ser2 to Ile, Arg2 to Ile, Val to Leu2, Gly to Cys substitutions at the probabilities of 56.5, 60.0, 42.3 and 40.0%, respectively.

However, the term "destructive effect" rarely means the complete disappearance of an epitope for G to U nonsynonymous mutations. An average length of an epitope completely destroyed due to a single nonsynonymous C to A mutation is 3.62 ± 1.96 amino acid residues. An average length of the disappeared part of an epitope is 1.96 ± 0.29 amino acid residues. The length of the elongated part of an epitope is always just 1 amino acid residue.

4. Discussion

Acrophilicity is the probability for amino acid to be situated on the surface of protein globule and so, to be included in linear B-cell epitopes. Glycine has the highest level of acrophilicity (3.0), the level of acrophilicity for proline (2.6) is some lower (Hopp and Woods, 1983). According to BepiPred predictions, replacement of proline inside the B-cell epitope usually leads to the decrease of the length of this epitope or to the complete disappearance of linear B-cell epitope from the surface of protein globule. Both C to U and C to A mutations may lead to the replacement of proline with amino acid of a lower acrophilicity (see Figs. 2 and 3). Replacements of glycine are not as destructive as replacements of proline (see Figs. 1 and 4).

Hydrophobic and acrophobic amino acids (leucine, isoleucine, valine, methionine and phenylalanine) are encoded by codons containing thymine (uracil) in their second codon positions; while

Author's personal copy

acrophilic proline, threonine and serine (the quartet of serine – Ser4) as well as fairly neutral alanine are encoded by codons containing cytosine in their second codon positions. Nonsynonymous nucleotide mutation of C to U direction often leads to the radical amino acid replacement which always has a great impact on a tertiary structure of a protein.

Indeed, the most destructive type of nucleotide mutation for gp120 linear epitopes is C to U transition (see Table 1). The second place is shared by C to A and G to U transversions. The common probability of the destructive effect is higher for G to U transversions, while the probability of the complete disappearance of an epitope is higher for C to A transversions. The most frequently occurring in HIV genes mutations (G to A transitions) have the lowest probability of the destructive effect on B-cell epitopes and the lowest strength of this effect.

A change of the amino acid sequence of linear B-cell epitope as well as a little decrease in its length caused by G to A nonsynonymous mutation surely may help the virus to escape the neutralizing antibody. But the probability that this mutated epitope will not be recognized by previously synthesized antibody should be relatively low. From this point of view nonsynonymous C to U mutation often leading to the complete disappearance of B-cell epitope from the surface of a protein should be much more beneficial for a virus in terms of immune escape. Due to APOBEC1 editing of viral RNA, HIV1 proteins (including gp120) tend to become less immunogenic during the infection.

Transversions of C to A direction should be more frequent than G to U transversions in HIV1 genes. Yet, the both of these types of transversions have stronger destructive effect on B-cell epitopes than G to A transitions.

This work should be useful for HIV1 vaccine designers. There is no reason to immunize a person to immunogenic but unstable B-cell epitope, which may simply disappear from the surface of gp120 protein after a single nonsynonymous C to U mutation (or after a few nonsynonymous C to U mutations) in the region of *env* gene coding for this epitope. Relatively stable epitopes should not contain or should contain a limited number of main targets for highly destructive C to U and C to A mutations, such as proline, threonine and serine.

References

Berkhout, B., de Ronde, A., 2004. APOBEC3G versus reverse transcriptase in the generation of HIV-1 drug-resistance mutations. AIDS 18 (13), 1861–1863.

Bishop, K.N., Holmes, R.K., Sheehy, A.M., Malim, M.H., 2004. APOBEC-mediated editing of viral RNA. Science 305, 645.

Gros, L., Saparbaev, M.K., Laval, J., 2002. Enzymology of the repair of free radicals-induced DNA damage. Oncogene 21, 8905–8925.

Hopp, T.P., Woods, K.R., 1983. A computer program for predicting protein antigenic determinants. Mol. Immunol. 20 (4), 483–489.

Izumi, T., Shirakawa, K., Takaori-Kondo, A., 2008. Cytidine deaminases as a weapon against retroviruses and a new target for antiviral therapy. Mini Rev. Med. Chem. 8, 231–238.

Kamath-Loeb, A.S., Hizi, A., Kasai, H., Loeb, L.A., 1997. Incorporation of the guanosine triphosphate analogs 8-oxo-dGTP and 8-NH_2-dGTP by reverse transcriptases and mammalian DNA polymerases. JBC 272 (9), 5892–5898.

Khrustalev, V.V., 2009. Can mutational GC-Pressure create new linear B-cell epitopes in Herpes Simplex virus type 1 glycoprotein B? Immunol. Invest. 38 (7), 613–623.

Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immun. Res. 2, 2.

Petit, V., Guétard, D., Renard, M., Keriel, A., Sitbon, M., Wain-Hobson, S., Vartanian, J.P., 2009. Murine APOBEC1 is a powerful mutator of retroviral and cellular RNA in vitro and in vivo. J. Mol. Biol. 385, 65–78.

Pillai, S.K., Wong, J.K., Barbour, J.D., 2008. Turning up the volume on mutational pressure: is more of a good thing always better? (a case study of HIV-1 *vif* and APOBEC3). Retrovirology 5, N26.

Sueoka, N., 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. U.S.A. 85, 2653–2657.