HA03 as an Iranian Candidate Concealed Antigen for Vaccination against *Hyalomma anatolicum anatolicum*: Comparative Structural and *In silico* Studies

Mohammadi1,3, A., Aghaiypour2, K., Keywanfar1, H., Paykari2, H., Tajbakhsh1, H., Jalali1, A.H., Safavieh2, S., Ghorbani2, R., Foroghi3, A.

1. Department of Microbiology, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran.
2. Department of Biotechnology, Razi Vaccine and Serum Research Institute, Karaj, Iran
3. Department of Human Viral Vaccine Razi Vaccine and Serum Research Institute, Karaj, Iran

Received 04 Feb 2012; accepted 21 May 2013

Abstract
In the last decades researchers had focused on developing a vaccine against tick based on protective antigen. Recombinant vaccines based on concealed antigen from *Boophilus microplus* have been developed in Australia and Cuba by the name of TICKGARD and GAVAC (De La Fuente and Kocan, 2006). Further studies on this antigen have shown some extent of protection against other species (De Vos et al., 2001). In Iran most important species is *Hyalomma anatolicum* and limited information about its control are available. This paper reports structural and polymorphic analysis of HA03 as an Iranian candidate concealed antigen of *H. a. anatolicum* deposited in Gen-Bank (Aghaeipour et al. GQ228820). The comparison between this antigen and other mid gut concealed antigen that their characteristics are available in GenBank showed there are high rate of similarity between them. The HA03 amino acid sequence had a homology of around 89%, 64%, 56% with HA98, BM86, BM95 respectively. Potential of MHC class I and II binding region indicated a considerable variation between BM86 antigen and its efficiency against Iranian *H. a. anatolicum*. In addition, predicted major of hydrophobisity and similarity in N-glycosylation besides large amount of cystein and seven EGF like regions presented in protein structure revealed that value of HA03 as a new protective antigen and the necessity of the development, BM86 homolog of H. a. anatolicum HA03 based recombinant vaccine.

Keywords: HA03, Concealed Antigen, BM86 Homologous, *Hyalomma*

INTRODUCTION
Ticks are important exoparasites in the field of veterinary and human hygiene. For many years, tick control extensively relies on the usage of chemical acaricides. However this application is often accompanied by serious problems, including the selection of acaricide resistant ticks, environmental contamination and contamination of milk and meat product with drug residues (GRISI 2002) (Nolan 1990) (Chema et al 1990, Pegram et al 1991) (Garcia-Garcia et al 2000). In the early 1990s, vaccines based on concealed antigen were developed could induce
immunological protection against tick infestation. Development of concealed antigens-based vaccine is attractive because the host and parasite will not have evolved immunological interaction. These antigens are not normally exposed to the host and therefore do not usually play a role in naturally acquired immunity. (de la Fuente et al 2007, de la Fuente et al 2006, de la Fuente et al 2006b, De La Fuente & Kocan 2006, Kocan 2003, Nuttall et al 2006, Willadsen 2006). In this paper we will report the results of molecular characterization and bioinformatics analysis of HA03, compared with the previous data about HA98, BM86 and BM95. These analyses are about the identity and MHC-I and II binding prediction sites and also comparing the N-glycosylation and hydrophobisity as well as the amount of cystein in epidermal growth factor like domain. Result can give us a general view to introduce a new potential target antigen for vaccine development against Iranian ticks in cattle.

MATERIALS AND METHODS

Ticks. Eggs, larvae, nymphs and adult of H. a. anatolicum used in this study were from reference laboratory of tick and tick born disease of Razi institute.

RNA extraction, cDNA synthesis and amplification of sequence. Total RNA of target eggs from a single female was isolated and purified. Eggs were frozen and thawed in liquid nitrogen. RNA was extracted using RNA- plus TM (Cinna Gen, Tehran, Iran) according to manufacture's instructions. cDNA was synthesized from total RNA using (Roch – Indianapolis, USA) RT-PCR kit, and used as template in PCR reaction with primers designed based on BM86, BM95 and HA98 conserved sequence (Aghaiypour.k unpublished data). Primer synthesis and RT-PCR product sequencing was ordered to MWG biotech Germany. The sequencing results analyzed with DNA MAN software 4.13. To determine the sequence of 5’ and 3’ ends of gene random amplification of coding ends (RACES) were used (invitrogen, USA) based on manufactures instruction (Aghaiypour, k. unpublished data).

PCR Product Cloning. The PCR products were cloned by T/A cloning kit (Fermentas, Lativa) following the instruction manual. The plasmid was purified by high pure plasmid purification kit (Roch – Indianapolis, USA).

Bioinformatics and in silico analysis. Sequencing analysis and polymorphic studies secondary structure, hydrophobicity and hydrophilicity were done by DNA MAN software version 4.13. N-glycosylation was analyzed using software based services available at (http://www. cbs.dtu.db/services/ NetNGlyc). Analysis the amino acid sequence pattern related to the epidermal growth factor (EGF) like domains (Xaa- Cys- Xaa-5-Cys-Xaa-3-Cys-Xaa-811-Cys- Xaa-13-Cys-Xaa-15-Cys-Xaa ,).(Rand et al 1989, Doolittle et al 1984, Appella et al 1998) sequence of cloned HA03 was identified. The comparative study of BM86, BM95 and HA98 were done by the use of online service (http: //www. expasi.ch/tools/sim-prot.html) by keeping gap open penalty as 12 and gap extension penalty as 4.

MHC Binding Prediction: The online service (http: //www. imtech.res.inr) was used to identity the potential MHC-I and II binding regions. A score greater than 4% for MHC-I and 3% for MHC-II was considered as predicted binder for selected MHC-alleles and was shown by different colors on each antigen (Singh & Raghava 2001).

RESULTS

3’ and 5’ RACE (random amplification of coding end) were used for characterization of the coding end, following by sequencing. An antigen was similar to BM86, BM95 and HA98 and finally a gene by 2110 bp was isolated which name HA03-1 and deposited in GenBank by Aghaiypour. K. et al. (GQ228820). The identity to BM86 (Australia M29323) and BM95 (Argentina AF 150891) and HA98 (Indiana AF 347079) measured as 62.01%, 56.17% and 89.46% (Fast alignment method (Higgins & Sharp 1989)) and all of the isolated items showed 78.45% identity with each other (Table 1) (Figure 3).
Table 1. Matrix identity between HA03, HA98, BM86, and BM95.

<table>
<thead>
<tr>
<th>Identity (per cent)</th>
<th>HA03</th>
<th>HA98</th>
<th>BM95</th>
<th>BM86</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA03</td>
<td>-</td>
<td>89.46</td>
<td>56.17</td>
<td>62.01</td>
</tr>
<tr>
<td>HA98</td>
<td>89.46</td>
<td>-</td>
<td>56.17</td>
<td>62.31</td>
</tr>
<tr>
<td>BM95</td>
<td>56.17</td>
<td>56.17</td>
<td>-</td>
<td>84.62</td>
</tr>
<tr>
<td>BM86</td>
<td>62.01</td>
<td>62.01</td>
<td>84.62</td>
<td>-</td>
</tr>
</tbody>
</table>

The prediction of secondary structure between four antigens and 3-dimentional structure revealed that the N-terminal (signal peptide) and C-terminal are buried into lipid bilayer cell membrane of mid-gut cells, but the main body of protein combines with glycosilated motifs are topologically budding out of membrane. Result from differential analysis of secondary structure showed that there are major differences between HA03 and HA98 in one side and BM95, BM86 from another side. It seems that except 71 position in HA03 the rest of amino acid sequence is completely the same as HA98. Results of secondary group (HA03 with BM86and BM95) were shown about 5 and also 200 different position between BM86 and BM95. In Figure 1 comparative result of hydrophobicity and hydrophilicity of all antigens is available. These results suggest some differences between HA03 and other antigens. The HA03 show regions of high immunogenic potential that were not observed in the other variants. The peaks in green show regions of high immunogenic potential in HA03 that were not observed in the BM86 and BM95 variants. However, HA03 and HA98 have a high potential of antigen covering. This result revealed that amino acid changes presented in the HA03 not observed in protein regions which normally exist to antibody recognition in BM86 and BM95. It can also predict that HA03 and HA98 have a high potential to cover antigen detecting of together meanwhile it seems to be an interference between these antigens and BM86 and BM95. It means that may be using BM95 based antibody against HA03 antigen could not recognize this antigen. Of course this is a hypothesis based on in silico data analysis. In general analysis of hydrophilicity that is likely to be potential sites of antigenicity showed that, all of four antigens have compatible structure with each other. Polymorphism of MHC-I and MHC-II were compared between four antigens. (Tables 2, 3, 4 and 5).

Table 2. No of peptide binder on MHC-I of HA03 and it comparison by BM95, BM86, HA98.

<table>
<thead>
<tr>
<th>Name of Antigen</th>
<th>HLA-A1</th>
<th>HLA-A2</th>
<th>HLA-A201</th>
<th>HLA-A203</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA03</td>
<td>13</td>
<td>16</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>HA98</td>
<td>18(42.8%)</td>
<td>10(62.5%)</td>
<td>10(52.6%)</td>
<td>16(48.2%)</td>
</tr>
<tr>
<td>BM95</td>
<td>2(5.8%)</td>
<td>1(8.3%)</td>
<td>3(16.6%)</td>
<td>0</td>
</tr>
<tr>
<td>BM86</td>
<td>0</td>
<td>1(6.25%)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. No of peptide binder on MHC-II of HA03 it comparison by BM95, BM86, HA98.

<table>
<thead>
<tr>
<th>Name of Antigen</th>
<th>DRB1.101</th>
<th>DRB1.102</th>
<th>DRB1.1301</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA98</td>
<td>5(33.3%)</td>
<td>5(29.4%)</td>
<td>5(37.5%)</td>
</tr>
<tr>
<td>BM95</td>
<td>1(12.5%)</td>
<td>2(28.5%)</td>
<td>4(44.4%)</td>
</tr>
<tr>
<td>BM86</td>
<td>2(16.6%)</td>
<td>1(7.14%)</td>
<td>2(22.2%)</td>
</tr>
</tbody>
</table>

Table 4. Similarity percent of MHC-I Locuses between HA03 and other antigens.

<table>
<thead>
<tr>
<th>MHC-I</th>
<th>HLA-A1</th>
<th>HLA-A2</th>
<th>HLA-A201</th>
<th>HLA-A203</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA98</td>
<td>18(42.8%)</td>
<td>10(62.5%)</td>
<td>10(52.6%)</td>
<td>16(48.2%)</td>
</tr>
<tr>
<td>BM95</td>
<td>2(5.8%)</td>
<td>1(8.3%)</td>
<td>3(16.6%)</td>
<td>0</td>
</tr>
<tr>
<td>BM86</td>
<td>0</td>
<td>1(6.25%)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5. Similarity percent of MHC-II Locuses between HA03 and other antigens.
These analysis was based on HLA-A1, HLA-A2, HLA-A201, HLA-A205 for MHC-I and DRBI (.101), DRBI (.102) and DRBI (.301) for MHC-II between HA03 and three other antigens (HA98, BM95 & BM86). As shown in table 2 the number of HA03 binding peptide on HLA-A1, HLA-A2, HLA-A201 & HLA-A205 were 49, 14, 19 and 27 that from these number 18 (42%), 10 (62.5%), 10 (52.6%) and 14 (48.2%) was common between HA03 and HA98. These result between HA03 with (BM95, BM86) in relation to all of the for MHC-I gene (HLA-A1, HLA-A2, HLA-A201 & HLA-A205) is very low percentage (5% - 16%) for BM95 and less than 6.25% for BM86.

Table 6. N-linked glycosylation sites on HA03 in comparison of BM86-BM95-HA98. All of antigens has five except HA98 that contains four N-glycosylation site.

<table>
<thead>
<tr>
<th>Name of Antigen and Accession No.</th>
<th>N-glycosylation site</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA03 (GQ228820)</td>
<td>148 (NKT); 187 (NST); 314 (NFT); 359 (NIS); 393 (NCT)</td>
</tr>
<tr>
<td>HA98 (AF3477079)</td>
<td>147 (NPT); 314 (NFT); 359 (NIS); 393 (NCT)</td>
</tr>
<tr>
<td>BM86 (M29323)</td>
<td>85 (NPS); 141 (NMT); 182 (NCS); 348 (NIS); 382 (NCT)</td>
</tr>
<tr>
<td>BM95 (AF150891)</td>
<td>85 (NPS); 148 (NKT); 182 (NCS); 348 (NIS); 382 (NCT)</td>
</tr>
</tbody>
</table>

Analysis of sequence of peptide binders shows that in locus of HLA-A2 (MHC-I) there are two peptide in region of 65-79 (contain NAAEQCEYK) and at region 161-169 (DLCEKNLL) that is common between BM95, HA98, HA03 and in locus of HLA-A201 on region of (440-450) there is 2 peptide binder NAAEKQCEYK, DLCEKNLL that is common between BM95, HA98, HA03. The most important result revealed from these analysis is that a common binder peptide in one region of 3 MHC-I locus (HLA-A2, HLA-A201 & HLA-A205) named (KLQACEHPI) and in all of four antigen (HA03, HA98, BM95, BM86) repeated. So theoretically we can use this peptide as a new candidate for designing common new antigen between Hyalomma and Boophilus. MHC-II sequence analysis showed that there is complete similarity on the first 485 base between BM86 and BM95 related to three locus on MHC-II (DRBI .301, DRBI .102, DRBI.101). Meanwhile between HA98 and BM86, just three peptide binders are similar as follow: DRBI.101 in region of (470-476), DRBI.102 (170-476), and DRBI.301 (461-470). The similarity between HA03 and BM86, BM95, HA98 has shown in table 4. There is high similarity between HA03 and HA98 (30%) related to all three loci DRBI.101 region (32-40) and (651- 664), DRBI.102 region (32-40) and 651-664 and DRBI.301 region (178-182) and 363-370. This antigen (HA03) has common peptide binder as follow with BM86 and BM95, DRBI.101 at region (5-10), DRBI.102 at region (5-10), DRBI.103 at (5-10) and (122-132). In comparison between MHC-I there isn't any common peptide binder on MHC-II between four antigen. Comparative analysis of N-glycosylation in HA03 and three other antigen (HA98 – BM86 – BM95) showed that all of antigens has five potential sites except HA98 that contain four sites. There isn't any similarity between HA03 and BM86 and BM95 meanwhile similarity with HA98 N-glycosylation are present at amino acid residue of , 314-316 (NFT), 359-361 (NIS) and 393-395 (NET). These characteristics for BM86 - BM95 and HA98 are compared and presented in table 6. Result of EGF like domain showed seven regions that completely matched the pattern. Where Xaa is any amino acid except cystein and 4 cysteins are present in one EGF like domain. These results were exactly the same as the reported (Rand et al 1989) and observed pattern for BM86. According to the figure 3. HA03 and HA98 are completely the same, except in 71 positions. However, comparison of HA03 with BM86 and BM95, revealed, 5 and 200 different positions, respectively.

**DISCUSSION**
Progressive uses of recombinant DNA and protein synthesis technology made it possible that even at the level of epitope the protein can be synthesized. So it can be said that designing and introducing a recombinant antigen as a tick vaccine would be combination of art and current technology. An effective tick vaccine should activate: antigen presenting cell to produce antigen and cytokine, activate B and T cell to make memory cell and stimulate T helper, cytotoxic T cells as well as their epitope to overcome on the limitation of MHC polymorphism immune response. The above-mentioned problem besides different tick species causes to search about protective antigen as potent vaccine. The result of using BM86, suggests that specific BM86 homologues could cover more protection (Liao et al 2007). In this paper, we characterize HA03 antigen, isolated from *Hyalomma a. anatolicum* from Iran and evaluate its sequence, structural homology, MHC-I and MHC-II peptide binder, N-glycosylation site, EGF like domain compared with HA98, BM86 and BM95 antigens. Our analysis showed HA03 (concealed antigen from midgut of *Hyalomma anatolicom anatolicom*) contain 7 EGF like domains with high content of cysteine (67 residues). These features are common between HA03...
Figure 3. Comparative HA03 sequence alignment with HA98, BM95 & BM86. Except in 71 position HA03 and HA98 show same alignment. However between HA03 and BM95, BM86, there are 200 different positions.

HA03-1: MCPPPLFVAAVLLIVGCAGQVLYAQPTTSCDGFQKFCQ
HA98: -----------g-i-----------r--------------
BM95: l-------------------------q-r----n-aa-pads------------
BM86: l-------------------------q-r----n-aa-pads------------

HA03-1: SAECEVIPTEDDFVCKCPRDIYNAAEKGYRTCTK
HA98: n-----------r-----------r-------------q-----------
BM95: n-----------r-----------r-------------q-----------
BM86: n-----------r-----------r-------------q-----------

HA03-1: KVCYGFCTCQVGFPRTACGCGVDTLTLLCGIQEMFSNDC
HA98: -a-----------n-v-s--i-----------r-------------ya-e-
BM95: re--r-ven-skgs-v-eas-d--q-k-kndyat--
BM86: re--r-ven-skas-v-eas-d--q-k-kndyat--

HA03-1: GRKGTAVLTDGFLGARCCGWEAKRRKNTQDGCKVPTTC
HA98: -----------r-------------g-m------------g-pn-
BM95: rr-r--k-i-t--------gam-m-trn--------
BM86: rr-r--k-i-t--------gam-m-trn--------

HA03-1: IRPDLCXKLCEKMLLLGKDTRCQGWNSDGSVPQGETY
HA98: -----------p-----------p-d-----------
BM95: l-------------------------q-r----n-aa-pads------------
BM86: l-------------------------q-r----n-aa-pads------------

HA03-1: CSPGIXGEDGCKDACKTEALLLCQGGCIRCQPKQAY
BM95: n-----v--a---------r-n--f-----q----d----
BM86: n-----v--a---------r-n--f-----q----d----

HA03-1: RLIASEPLSKEHVKLKQLACENPVADFCMLYPMFLIJKGSA
HA98: -----------k-----------m-----------m------------m-
BM95: -----ek--------y------igew---m-k-l-n--------
BM86: -----ek--------y------igew---m-k-l-n--------

HA03-1: SLSNDCFCEWYEEARLVKATIRAIAGVFKEVINLQDIKA
HA98: -----------k-----------m-----------m------------m-
BM95: -----ek--------y------igew---m-k-l-n--------
BM86: -----ek--------y------igew---m-k-l-n--------

HA03-1: CRGIEEENCESLLKNQEKAYKGVNKCAKVGDLYWFQCANG
BM95: ---e-------d-------a---q--v--dn-f--------d--
BM86: ---e-------d-------a---q--v--dn-f--------d--

HA03-1: YRAVEDEVARLRLLSVCAGVCSSTDEQLECKANKQICVF
BM95: -----------k-----------m-----------m------------m-
BM86: -----------k-----------m-----------m------------m-

HA03-1: CVYRDQKAECKCPQGTVDAGQGCSGEPASVTCEKNIAAC
BM98: -----------q--e------y------g-ve-----es--e--
BM95: -----------q--e------y------g-ve-----es--e--
BM86: -----------q--e------y------g-ve-----es--e--

HA03-1: ENEKPNCQCPFQVFGQQAARCTPCKFRECEDEEEKKE
BM98: -----------d-----------d-----------d-----------d-
BM95: -----------d-----------d-----------d-----------d-
BM86: -----------d-----------d-----------d-----------d-

HA03-1: CVYRDQKAECKCPQGTVDAGQGCSGEPASVTCEKNIAAC
BM98: -----------q--e------y------g-ve-----es--e--
BM95: -----------q--e------y------g-ve-----es--e--
BM86: -----------q--e------y------g-ve-----es--e--

HA03-1: RSNGQCRACENHRPFVCKEETSDEVTAEMMTCTAKDAPDFG
BM95: -----------g-v-----------i-----------g-v-----------i---------
BM86: -----------g-v-----------i-----------g-v-----------i---------

HA03-1: KSQGVAVSATT.LLLAASVAAA.
BM95: -------------------
BM86: -------------------

BM95: -----------k-----------m-----------m------------m-
and BM86 and HA98. Cysteine is sulphydryl – amino acids which cause cross linking disulfide binds between different parts of proteins and also is considered as the receptor or secretory gland function for these proteins (Liao et al 2007). In addition Cystein is characteristic of extracellular membrane bound proteins (Wise 1984). These amounts of cysteine represent the frequency of fivefold greater than the average (Miseta & Csutora 2000), which is responsible in protein structure, (Noiva 1994, Raina & Missiakas 1997) and has the important role in protein folding (Azhahianambi et al 2009). This conserved structure EGF like domains not only confirms the status of BM86 homologue of H. a. anatolicum but supports the important role of EGF like domains toward the structure and function of these molecules and its potency for vaccination target (Figure 2). As shown in table 1, the pattern of identity between BM95 and BM86 isolate was about 56 and 71% while the nucleotide sequence similarity of HA03 and HA98 is 89.46%. These results revealed the closeness between HA03 and HA98. The variation greater than 2.8% in the amino acid sequence of protein would be sufficient to confirm vaccination inefficiencies (Garcia-Garcia et al 1999), when heterologous recombinant antigens are used. By this in silico analysis we have shown that for controlling the Ha03 may have better results compared with BM86 or BM95. Prediction of HLA antigen recognition sites and epitope modeling suggested that differential binding of different alleles and explain the observed effect on antibody titers. The link between variation in MHC-II alleles and it's correlation with antibody response in two bovine genera were studied (Sitte et al 2002). The analysis of polymorphism of peptide binder on MHC-I and II between four antigens showed there are high similarities between HA03 and HA98, while the homology among HA03, BM86 and BM95 was not so high. These results also confirm the using of HA03 as new candidate antigen against H. a. anatolicum. It is difficult to measure the size of effect from the result presented between HA03 and other antigen, but it is clear that it explains only a small piece of antigen to antigen variation in antibody response. The presence of a common single peptide binder between four antigen on MHC-I suggested that this peptide can be used as a conserve recombinant antigen for immunization against all of ticks in future, of course these results revealed based on in silico analysis and for final judgment it's better focus on planning on epitope map and field trial result. Our next step is study on cloning and expression of HA03 in Prokaryotic and Eukaryotic system and study on it.

Acknowledgment

This project was financially supported by a grant for Ph.D degree from Razi Institute in Iran. The authors wish to thank Biotechnology – Genomic Department members for their helpful contribution to this project.

References


