**Introduction:**

The tick *Rhipicephalus (Boophilus) microplus* is an obligate haematophagous ectoparasite that feeds on a variety of domestic animals and mainly on cattle. *B. microplus* is distributed worldwide principally in tropical and subtropical regions and feeds on cattle. The development of vaccine is very important in order to save the livestock. Three glycosylphosphatidylinositol (GPI)-anchored proteins (Contig2828, Contig7420, and CK181624) from the gut of *B. microplus* were selected by VaxiJen server as highly immunogenic. As the proteins will be tested further on BalbC mice (haplotype d), the T-cell epitope predictions were performed on mouse major histocompatibility complex (MHC) class II proteins IAd and IEd. The predictions were made by three sequence-based methods implemented in the servers MHCPred, NetMHCII and RANKPEP and by one structure-based method – molecular docking. Based on the consensus prediction by the sequence- and structure-based methods, the peptides 1-14, 31-47 and 49-67 from Contig2828, peptides 6-32 and 75-101 for Contig7420, peptides 1-26 and 45-67 for CK181624 were selected for further experimental studies.

**Abstract.** Ticks are blood-sucking parasites that feed on variety of domestic animals. *Rhipicephalus (Boophilus) microplus* is a hard tick found in tropical and subtropical regions and feeds on cattle. The development of vaccine is very important in order to save the livestock. Three glycosylphosphatidylinositol (GPI)-anchored proteins (Contig2828, Contig7420, and CK181624) from the gut of *B. microplus* were selected by VaxiJen server as highly immunogenic. As the proteins will be tested further on BalbC mice (haplotype d), the T-cell epitope predictions were performed on mouse major histocompatibility complex (MHC) class II proteins IAd and IEd. The predictions were made by three sequence-based methods implemented in the servers MHCPred, NetMHCII and RANKPEP and by one structure-based method – molecular docking. Based on the consensus prediction by the sequence- and structure-based methods, the peptides 1-14, 31-47 and 49-67 from Contig2828, peptides 6-32 and 75-101 for Contig7420, peptides 1-26 and 45-67 for CK181624 were selected for further experimental studies.

**Keywords:** molecular docking, *Rhipicephalus (Boophilus) microplus*, MHCII
The effectiveness of the vaccine varies at different regions in the world. Another protein, named Bm95, isolated from Argentinean tick strain A and differentiated from Bm86 in 21 amino acid residues, has been used for preparation of a vaccine [15]. An immune response after vaccination with Bm95-based vaccine has been achieved in some cases, but a variable effect has been found in others, where the infected animals were insensitive to protein Bm86 [16]. In 2002, a synthetic anti- R. microplus vaccine, named SBm7462, containing oligopeptides from Bm86, was developed by Patarroyo and co-workers [17]. It was shown that the synthesized epitopes used in SBm746 are highly conserved among the South American tick strains [18, 19]. Although in most cases the Bm86-based vaccines elicit protective immune responses against R. microplus, the effectiveness vary considerably due to genetic variability of tick and bovine populations [20-25]. Therefore, the discovery of new tick antigens covering the majority of R. microplus populations could improve efficacy and reduce the variation in protection level afforded by the Bm86-based vaccines.

T-cell recognition is a fundamental mechanism underlying the adaptive immune system through the action of which the host identifies and responds to foreign antigens [26]. The T cell is a specialized type of immune cell that mediates cellular immunity. The T-cell receptors (TCRs), found on the surface of T cells, bind major histocompatibility complex (MHC) proteins, which are presented on antigen-presenting cell (APCs) surfaces. T-cell immune responses are driven by recognition of peptide antigens (T-cell epitopes) bound to MHC molecules. MHC proteins are highly polymorphic glycoproteins, which bind small peptide fragments, or epitopes, derived from both pathogen and host protein and present them at the cell surface for interaction by T cells. The combination of two or more preferred peptide residues (anchors) is called a motif. The experimental determination of motifs for every allele is prohibitively expensive in terms of labor, time and resources. The only practical alternative is to make use of a bioinformatics approach.

In this study we applied sequence-based and structure-based approaches to identify probable T-cell epitopes bound to mouse MHC class II proteins IAd and IEd.

**Methods**

**Immunogenicity prediction**

VaxiJen is a server for prediction of protective antigens based on the physicochemical properties of proteins [27-29]. The alignment-free approach implemented in the server allows recognition between immunogenic and non-immunogenic proteins from different origin. The protein structure is presented by uniform vectors, generated by auto- and cross-covariance transformation of protein sequence described by three amino acid z-scores accounting for hydrophobicity, size and polarity [30]. The accuracy of prediction of the implemented models into VaxiJen server varies between 70 and 90% for the internal and external test sets [27].

**Sequence-based prediction**

The sequence-based prediction for binding to mouse MHC class II proteins IAd and IEd was performed by the servers MHCpred [31-34], NetMHCII [35] and RANKPEP [36-38]. In MHCpred server, the quantitative prediction of peptide binding affinity to MHC is based on the additive method. A combination of individual amino acid contribution at each position of the peptide and contributions from side-chain – side-chain interactions are taken into account. Several allele specific quantitative structure-activity relationship (QSAR) models, derived by partial least square (PLS), are implemented in the server. The quantitative prediction of peptide binding affinity to MHC in NetMHCII server uses NN-align, which is an artificial neural network-based alignment algorithm [35]. In NetMHCII, the binding core and binding affinity are estimated simultaneously. The method incorporates explicit encoding (composition and length) of peptide flanking residues and deals
with the data redundancy inherent in the peptide data due to multiple examples of identical binding cores. RANKPEP server uses position specific scoring matrices (PSSM) or profiles in prediction of peptide binding potential to MHC [36-38]. PSSM are derived from a set of aligned known binders to a given MHC molecule and contain observed sequence-weighted frequencies of all amino acids from the sequence alignments. Comparison between a peptide, binding to a given MHC and the corresponding PSSM for that MHC molecule gives the predicted binding potential of the peptide to the MHC molecule.

**Structure-based prediction. Molecular docking**

Structure-based methods derive predictions by analyzing the interactions of one single 3D structure of the complex between peptide and MHC protein. A widely used structure-based method, together with protein threading [39, 40] and homology modeling [41, 42], is molecular docking [43-45]. The molecular docking is a tool for prediction the binding mode(s) of a ligand within a protein with known 3D structure. Docking methodology consists of two steps, (1) generation the poses as sampling positional, conformational, and configurational space of the ligand within the binding site of the receptor and (2) selection the poses using scoring functions for ranking the ligand poses. GOLD v. 5.0.2 was used in the present study. GOLD is based on genetic algorithms for pose generation and several scoring functions are available for ranking the poses [46]. It has proven successful in virtual screening, lead optimization and identification of the correct binding mode of active molecules [47, 48]. GOLD takes into account the flexibility of the ligand as well as the flexibility of the residues within the binding site.

**Results and Discussion**

The selection of immunogenic proteins from *R. (B.) microplus* was performed by VaxiJen server. The proteins selected as highly immunogenic proteins were Contig2828 (VaxiJen score 1.2176), Contig7420 (VaxiJen score 1.0345) and CK181624 (VaxiJen score 0.9919). Each protein was presented as a set of overlapping nonamer peptides. Thus, for Contig2828 a set of 59 overlapping peptides was generated, for Contig7420 – a set of 93 peptides for CK181624 – a set of 61 peptides.

As the proteins are intended to be tested on BalbC mice (haplotype d), the T-cell epitope predictions were performed on mouse MHC class II proteins IAd and IEd. Prediction by MHCPred identified high binders (IC50< 50 nM) only to IAd, but not to IEd (Fig. 1). NetMHCII predicts binding peptides only to IAd, but no high binder was identified. RANKPEP, similarly to MHCPred, identified binders only to IAd, but not to IEd (Fig. 2).

![Fig. 1. Binding affinity prediction to IAd by MHCPred. Predicted binders with IC50< 50 nM are highlighted in green.](image-url)
Fig. 2. Binding affinity prediction to IAd by RANKPEP. Predicted binders with binding threshold: 7.10 are highlighted in purple.

For molecular docking simulations, the X-ray structures of peptide – MHC protein complexes OVA123-139/IAd (pdb code: 2iad) and HB/IEk (pdb code: 1iea) were selected as input structures. As MHC alleles are expressed by Balb/c strain, haplotype d was used as experimental mouse model. The structure of IEd was modeled from IEk applying single amino acid substitution method. Subsequently, a simulated annealing optimization of each peptide-protein complex was performed by AMBER force field [49]. The binding site was defined from the corresponding X-ray structure (pdb id: 2iad and 1iea) within 6 Å radius (Fig. 3). ChemScore scoring function was used to rank the solutions. Protein and peptide backbone were fixed. All peptides originating from one B. microplus proteins were ranked according to their docking-derived binding score and the top two clusters of peptides were selected for further experimental studies.

Fig. 3. Two docking solutions of peptides derived from CK181624 in complex with IEd.
The molecular docking was performed at the following settings:
- 6 Å radius of the binding site;
- fixed protein and peptide backbones;
- flexible side chains of the binding site residues and flexible peptide side chains;
- ChemScore function for ranking the poses.

The best binding peptides (ChemScore > 76.40) for IAd and IEd are given in Fig. 4 and 5, respectively.

**Fig. 4.** Binding affinity prediction to IAd by molecular docking. The top two best predicted clusters of binders are highlighted in magenta.

**Fig. 5.** Binding affinity prediction to IEd by molecular docking. The top two best predicted clusters of binders are highlighted in magenta.
Peptides selected for further experimental studies based on consensus prediction to IAd and IEd by molecular docking, MHCPred and RANKPEP. Predicted binders are highlighted in blue.

**Acknowledgments.** The present research was supported by the National Research Fund, Bulgaria (Grant 02-1/2009).

**Conclusion**

Based on the consensus prediction by molecular docking, MHCPred and RANKPEP, we propose for further experimental studies peptides 1-14, 31-47 and 49-67 from Contig2828, peptides 6-32 and 75-101 for Contig7420, and peptides 1-26 and 45-67 for CK181624 (Fig. 6).

**References**


24. Rodriguez M, Massard CL, da Fonseca AH, Ramos NF, Machado H, Labarta V, de la...


Corresponding author:
Mariyana Atanasova
Chemistry Department
Faculty of Pharmacy
Medical University of Sofia
2 Dunav st., Sofia 1000, Bulgaria
email: matanasova@ddg-pharmfac.net