Journal of Virological Methods 315 (2023) 114704

Contents lists available at ScienceDirect

Journal of Virological Methods

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

Protocols

A double-strain TM (gp45) polypeptide antigen and its application in the serodiagnosis of equine infectious anemia

Angela Ostuni [a,*], Valentina Iovane [b], Magnus Monné [a], Maria Antonietta Crudele [a], Maria Teresa Scicluna [c], Roberto Nardini [c], Paolo Raimondi [d], Raffaele Frontoso [d,e], Raffaele Boni [a], Alfonso Bavoso [a]

[a] Department of Sciences, University of Basilicata, viale Ateneo Lucano 10, 85100 Potenza, Italy
[b] Dipartimento di Agraria - Università degli Studi di Napoli Federico II -Via Università, 100 - 80055 Portici, NA, Italy
[c] Istituto Zooprofilattico Sperimentale del Lazio e della Toscana "M. Aleandri", Via Appia Nuova, 1411, 00178 Roma, Italy
[d] OneHEco APS, 84047 Capaccio Paestum, SA, Italy
[e] Istituto Zooprofilattico Sperimentale del Mezzogiorno Via Salute, 2 - 80055 Portici, Napoli, Italy

ARTICLE INFO

Keywords:
EIAV, Antigen design
Protein clustering
Modelling and structural analysis
Synthetic gene
ELISA test

ABSTRACT

Lentiviruses, including equine infectious anemia virus (EIAV), are considered viral quasispecies because of their intrinsic genetic, structural and phenotypic variability. Immunoenzymatic tests (ELISA) for EIAV reported in the literature were obtained mainly by using the capsid protein p26, which is derived almost exclusively from a single strain (Wyoming), and do not reflect the great potential epitopic variability of the EIAV quasispecies. In this investigation, the GenBank database was exploited in a systematic approach to design a set of representative protein antigens useful for EIAV serodiagnosis. The main bioinformatic tools used were clustering, molecular modelling, epitope predictions and aggregative/ solubility predictions. This approach led to the design of two antigenic proteins, i.e. a full sequence p26 capsid protein and a doublestrain polypeptide derived from the gp45 transmembrane protein fused to Maltose Binding Protein (MBP) that were expressed by recombinant DNA technology starting from synthetic genes, and analyzed by circular dichroism (CD) spectroscopy. Both proteins were used in an indirect ELISA test that can address some of the high variability of EIAV. The novel addition of the gp45 double-strain antigen contributed to enhance the diagnostic sensitivity and could be also useful for immunoblotting application.

1. Introduction

Equine infectious anemia virus (EIAV) is a lentivirus (family *Retroviridae*) that infects horses, mules and donkeys. It shares many characteristics with other lentiviruses like HIV and SRLV including its viral structure, genome and certain aspects of its life cycle and transmission (Leroux et al., 2004). EIAV establishes a persistent infection and is naturally transmitted between hosts by biting insects like *Tabanidae*; inappropriate iatrogenic practices with contaminated blood can also be responsible for the transmission (Issel and Foil, 2015). The presence of the virus has been documented worldwide including Africa (Fathi et al., 2014; Wegdan et al., 2016) and Australia (https://www.business.qld.gov.au/industries/farms-fishing-forestry/agriculture/livestock/animal-welfare/pests-diseases-disorders/equine-infectious-anaemia), with a very high prevalence in warm swampy areas. Phylogeographic studies suggest that the virus originated in Hungary that was one of the most important centers of diversification for the disease (Jara et al., 2020). Historical EIAV spread was predominantly due to long-distance dissemination following globalized animal trade. EIAV infection is a substantial concern for equine health worldwide equids and requires reporting to the World Organization for Animal Health (WOAH). Like other lentiviruses it is considered a viral quasispecies due its very high genetic, structural and antigenic variability (Wang et al., 2018). Globally circulating strains characterized so far, is related to six phylogenetic clades, or lineages (Cursino et al., 2018).

EIAV infection is characterized by an acute stage, that occurs sporadically with high fever, anemia and thrombocytopenia, a chronic stage with fluctuating viral loads and recurring febrile episodes and an asymptomatic phase with a decreased viral load and no apparent clinical manifestations. In many countries, preventing transmission is obtained

* Corresponding author.
E-mail address: angela.ostuni@unibas.it (A. Ostuni).

https://doi.org/10.1016/j.jviromet.2023.114704
Received 4 January 2023; Received in revised form 21 February 2023; Accepted 23 February 2023
Available online 24 February 2023

by isolation or euthanasia of animals that test positive for infection (Leroux et al., 2004), since there are no treatments and no prophylactic vaccines in current use against EIAV.

Disease testing is performed using serological tools mainly Agar Gel Immunodiffusion Assay (AGID), that is considered the "gold standard", ELISA and western blotting. Recently ELISA was suggested to be the first choice of test for screening followed by AGID and western blotting as confirmatory tools, owing to the greater diagnostic sensitivity of the ELISA when compared to the AGID (Issel et al., 2013). Three structural viral proteins are the main antigens used in the EIAV serodiagnosis, the capsidic protein p26 (CA), the transmembrane protein gp45 (TM) and the surface protein gp90 (SU). AGID is exclusively, and ELISA, is preferentially based on the p26 protein, whereas all three proteins can be used in combination in western blotting (Issel et al., 2013). At the best of our knowledge, a single ELISA has so far been based on the combination of p26 and gp45 derived antigenic peptides and validated according to WOAH criteria (Rosati et al., 2004, Scicluna et al., 2018). It was reported that this test was superior to other ELISAs based only on p26 with respect to the precocity in antibody detection (Nardini et al., 2017). Other indirect ELISA based only on gp45 peptides as antigens were described (Soutullo et al., 2001; Naves et al., 2019). Almost all the p26-based ELISA tests depend on the prototypic Wyoming strain because it was one of the first to be isolated and considered as reference strain. In view of the high structural and immunogenic variability due to the "quasispecies" nature of EIAV, it could be interesting to assess the usefulness of combinations of peptide/proteins derived from different strains in immunochemical tests. This approach was applied for other lentivirus such as SRLV (Grego et al., 2002; Ostuni et al., 2021) to improve the diagnostic performances of the ELISAs and is currently used in certain commercial ELISAs for SRLV.

A relatively large number of EIAV protein sequences derived from different viral strains are available for p26 and gp45 in GenBank; these sequences were used as a starting point for this investigation. In this study bioinformatic tools were used, such as protein clustering, molecular modelling, epitope predictions and aggregative/solubility predictions, to develop a more systematic approach in the design of a set of representative protein antigens, based on p26 and gp45, useful in the EIAV serodiagnosis. The designed antigenic proteins were expressed by recombinant DNA technology starting from synthetic genes. Circular dichroism spectroscopy was exploited to analysed the secondary structure of the polypeptides in solution. ELISA and western blotting were used for the immunochemical characterization of the antigens.

2. Materials and methods

2.1. Sequence clustering analysis

The proteins investigated were the p26 capsid protein and the transmembrane (TM) protein gp45. The exodomain of the TM protein was identified using the program TM pred (Hofmann and Stoffel,1993). Cd-Hit, a sequence-based clustering software (Huang et al., 2010), was used to analyze the protein sequences retrieved with BlastP from the Non-Redundant database.

A clustering procedure to organize protein sequences into homologous and functionally similar groups was performed as previously reported (Ostuni et al., 2021). In the present analysis we used a two-level clustering. The sequences were clustered at 95 % identity threshold, and then the representative sequences of the individual clusters were "re-clustered" at lower identity thresholds between 75 % and 90 %. Default input parameters were used in CD-Hit.

2.2. Homology modelling and structural analyses

The structural homology model of p26 of EIAV (residues 18–224 of ABE03841.1) was generated with SwissModel (Waterhouse et al., 2018) with the primary template EIAV CA-SP (PDB ID: 6T61) (Dick et al., 2020). The 3D-models of two gp45 sequences of EIAV, gp45_1 (residues 483–630 of AFW99173.1) and gp45_2 (residues 483–630 of QIC50013.1), were built by Modeller (Fiser and Sali, 2003) by using the ectodomains of gp45 of EIAV (residues 474–530 and 567–601 of PDB ID: 3WMI), and gp41 of SIV (residues 1–123 of PDB ID: 1QCE) and HIV (residues 531–581 and 629–681 of PDB ID: 2×7R) (Buzon et al., 2010) as templates. In the gp45 models, structural restraints were used for intramolecular disulfide bond formation between C68 and C76 in each protomer. The gp45_1_2 polypeptide contains the sequences of gp45_1 and gp45_2 joined by a flexible linker. Two homology models of the double-strain gp45_1_2 polypeptide were generated starting from the model of the trimeric gp45_1: in both models the gp45_1 domain is placed in the position of chain A of the trimer, whereas gp45_2 is placed either in the position of chain B (gp45_1 _2_AB) or chain C (gp45_1_2_AC). All top evaluated models were energy minimized through the YASARA energy minimization server (Krieger et al., 2009).

For the structural analyses several programs were used. Protein geometry was evaluated by MolProbity (Williams et al., 2018). Multiple sequence alignments of p26 and gp45 homologs were generated by ClustalW in Seaview (Gouy et al., 2010) and sequence conservation was displayed on the structural models by ConSurf (Ashkenazy et al., 2016). Linear B-cell epitopes of the p26, gp45_1 and gp45_2 sequences were predicted by ABCpred (Saha and Raghava, 2006), BcePred (Saha and Raghava, 2004), BepiPred 2.0 (Larsen et al., 2006), COBEpro (Sweredoski and Baldi, 2009) and ElliPro linear (Ponomarenko et al., 2008). Discontinuous epitopes were predicted in the structural models by using Epitopia (Rubinstein et al., 2009), DiscoTope 2.0 (Kringelum et al., 2012), ElliPro discontinuous (Ponomarenko et al., 2008), EPCES (Liang et al., 2009) and SEPPA 3.0 (Zhou et al., 2019). The epitope consensus prediction for each position was calculated based on the results from the ten above-mentioned programs.

2.3. Computational aggregation study

An in-silico method that can predict the aggregative properties of proteins starting from the modelled 3D structures was used. Calculations were done with the Aggrescan3D software available at: http://biocomp.chem.uw.edu.pl/A3D2/ web site (Zambrano et al., 2015). Structures were used with no further energy minimization and default input parameters were applied. Results were analyzed with the tools provided by the software.

2.4. Synthesis and purification of the polypeptides

All polypeptides were obtained by recombinant DNA technology in *Escherichia coli* strain BL21(DE3) (Stratagene, La Jolla, CA, USA). The p26 synthetic gene (Eurofins Genomics, Ebersberg, Germany) was subcloned into the *Hin*dIII- *Bam*HI sites of plasmid pQE30 to obtain a protein with a N-terminal histidine tag (MRGSHHHHHHGS). The gp45_1_2 synthetic gene was cloned (GenScript Company) into the expression vector pMAL-c4X (Addgene) thus obtaining the recombinant polypeptide fused with the maltose-binding protein at the N-terminal; a histidine tag was added to the C-terminal.

All the proteins recovered from lysed IPTG-induced cells were analyzed by 4 %−12 % SDS-PAGE. Polypeptides were purified from soluble protein fraction by affinity chromatography with a HisTrap HP (GE Healthcare) column and analyzed by western blot using a Monoclonal Anti-poly-Histidine-Peroxidase antibody (Sigma A7058). Polypeptide concentrations were determined spectrophotometrically at 280 nm using an extinction coefficient of 22585 and 140050 $M^{-1}cm^{-1}$ at 280 nm (https://web.expasy.org/protparam/) for the p26 and gp45 constructs, respectively. Purity analysis was made by SDS-PAGE and HPLC on a reverse phase C4 column (Jupiter 5 mm, 300 A, 250 × 4.60 mm, Phenomenex).

2.5. CD spectroscopy analysis

CD spectra were acquired at 25 °C with a Jasco J-815 CD Spectrometer with a light source constituted by a Xenon 250 W lamp and equipped with a thermoelectric temperature controller. The acquisition was carried out in a thermostated cylindrical quartz cell with an optical path of 0.1 cm in the spectral range 190–250 nm, a scanning speed of 50 nm/min, 1.5 nm bandwidth, a time-constant of 1 s, 20 mdeg of sensitivity and a total number of 16 accumulations for each spectrum. Then, the baseline spectra of the solvents were subtracted and spectra were smoothed using the Fourier transform. Data were expressed in terms of the molar ellipticity per residues in units of deg × cm^2 × $dmol^{-1}$. Analysis of CD spectra for the evaluation of secondary structure content was performed with DICHROWEB (Miles et al., 2021) using the CONTINLL algorithm (Provencher and Glockner, 1981; Sreerama et al., 2000).

2.6. Sample collection

A total of 210 serum samples were used in this investigation of which 107 samples were collected from officially EIAV negative animals in the Basilicata and Campania Regions while of the 103 positive samples used, 100 were obtained from IZSLT (National Reference Center for EIA and WOAH Reference Laboratory for EIA) and 3 samples were collected in Basilicata region from EIAV positive animals confirmed by the National Reference Center for EIA. All sampling procedures were conducted in strict accordance with European legislation, regarding the protection of animals used for scientific purposes (European Directive 2010/63). Veterinary officers of the Italian National Health System collected blood samples during the compulsory, official eradication and surveillance program on EIA, for which equids must be periodically sampled.

2.7. Indirect ELISA

Preliminarily checkerboard titrations were conducted to optimize the in-house indirect ELISA tests. Maxi Binding Immuno plates SPL (Life Sciences) were coated with the p26 protein or the gp45 double-strain polypeptide (range 8–0,5 μg/ml) alone and with a combination of p26/ gp45 polypeptides in the mass ratio 1:5 at different total concentrations in the coating buffer (0.5 M carbonate buffer, pH 9.6), and then incubated at 4 °C overnight. The optimal antigens coating concentrations were 0.5 μg/100 μl and 0.1 μg/100 μl for gp45 and p26, respectively. Non-specific binding was inhibited with blocking buffer 2 % ovalbumin (Sigma-Aldrich) in PBS. Plates were washed twice with phosphate-buffered saline (PBS pH 7.4). Sera were diluted with PBS Tween 0.05 % (PBS-T) at 1:200. Protein G Peroxidase (Sigma-Aldrich) diluted in (PBS-T) at 1:60000 was added as secondary antibody; then, tetramethylbenzidine (TMB, Sigma-Aldrich) was used as a substrate. The reaction was stopped with 1 N sulfuric acid and the optical density (OD) read at 450 nm. The percentage of reactivity of each serum was calculated by applying the following formula: (OD sample - OD Negative control)/ (OD Positive Control - OD Negative Control) x 100. The diagnostic accuracy and the optimal cut off -point value were evaluated by ROC analysis performed using the MedRoc software (https://stenstat.com/MedRoc/MedRoc.htm).

2.8. Western blotting

The western blot analysis was performed as per a previously reported method (Martinelli et al., 2018) with some modifications. Both polypeptides separately resolved with SDS-PAGE on a 4 %−12 % gel were electrotransferred onto a nitrocellulose membrane (AmershamTM ProtranTM Nitrocellulose Blotting Membrane, GE Healthcare). After a 2 h blocking phase at RT, with 2 % non-fat dried milk in PBS with 0.05 % Tween 20 (PBS-T), membranes were probed overnight at 4 °C with sera 1:200 in PBS-T. The membranes were washed with PBS-T and incubated with Protein G- peroxidase 1:6000 at room temperature for 1 h, and the signals were visualized by the ECL™ Western Blotting Detection Reagents (GE Healthcare, Chicago, IL, USA), using the Chemidoc TM XRS detection system equipped with Image Lab Software for image acquisition (BioRad, Hercules, CA, USA). Densitometric analysis was performed by using GelAnalyzer 2010 software (Debrecen, Hungary). Protein expression level was expressed as a percentage of an EIAV positive control.

3. Results

3.1. Clustering analysis of the variable protein sequences of the EIAV p26 and gp45 (TM) proteins

To obtain EIAV sequence variants for the clustering analysis, the representative non-Wyoming strain EIAV sequences ABE03841.1 of p26, which has been exploited previously in two serological tests with good results (Alvarez et al., 2007; Singha et al. 2013), and AFW99173.1 of the TM exodomain were used in Blastp searches. Blastp retrieved 187 and 142 clear EIAV homologue sequences for the p26 protein and for the 145-residue-polypeptide fragment of the TM (gp45) protein (sequence 488–632 envelope glycoprotein accession AFW99173.1) (Quinlivan et al., 2013), respectively. In the Blastp analysis, the lowest sequence identities of the retrieved homologues were 82 % with respect to the p26 query sequence (ABE03841.1) (Alvarez et al., 2007) and 70 % with respect to the TM peptide query sequence (AFW99173.1). The 145-residue polypeptide represents the exodomain of the transmembrane protein and was chosen with the aid of a TM-pred analysis. Sequences shorter than the full-length one were also retrieved in the case of p26. Twenty and 18 clusters were obtained at 95 % identity for the p26 and TM sequences, respectively. This preliminary calculation was done to obtain a single representative sequence from very similar ones, for example those related to a specific outbreak. The centroids thus generated were the input for a successive clustering (Re-clustering). “Re-clustering” results are summarized in Table 1. In the case of the p26 at 80 % of sequence identity a single cluster is obtained with the representative sequence related to the accession number ABE03841.1 (Alvarez et al., 2007). The clustering of the gp45 (TM) polypeptide at 75 % of sequence identity shows that two main clusters are present with 13 and 3 sequences, respectively (these sequences are the centroids of the preceding clustering at 95 % threshold). They represent 140 out of the original 142 retrieved TM sequences (98.6 %). The representative sequences of the clusters are related to the accession numbers AFW99173.1 (Quinlivan et al., 2013) and QIC50013.1 (Malossi et al., 2020).

In conclusion, this analysis reduced a large data set of sequences to a limited number of representative sequences that were used for the antigen design.

3.2. Structural bioinformatics analysis of the EIAV p26 and TM proteins

To gain a better insight into the immunogenic properties of these proteins that complements the literature data presently available, a structural analysis and an epitope prediction was conducted both for p26 and for the transmembrane protein on the polypeptides gp45_1, gp45_2 and a double-strain gp45_1_2 polypeptide. A rational approach was taken to design recombinant constructs of p26 and gp45_1_2 considering structural and antigenicity aspects by using molecular homology modelling and epitope prediction.

3.2.1. Homology models of p26, gp45_1, gp45_2 and a double-strain polypeptide gp45_1_2

The homology model of p26 was generated on the primary template Gag polypeptide from EIAV, which share 91 % identical sequence. As its template, the p26 model has a two-domain structure: a N-terminal domain with seven α-helices, which is connected by a flexible linker to

Table 1
Reclustering results for p26 and gp45 (TM) proteins retrieved with BLASTP.

(a)

I.T.	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Cluster8
80%	ABE03841.1 20 Argentina							
85%	ABE03841.1 18 Argentina	ADJ93850.1 1 Italy	ACY02859.1 1 Italy					
90%	ABE03841.1 10 Argentina	ADU02647.1 3 China	AHL29039.1 3 Germany	AFV61763.1 1 Japan	QHD59425.1 1 USA	ADJ93850.1 1 Italy	ACY02859.1 1 Italy	

(b)

I.T.	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Cluster8
75%	AFW99173.1 13 ireland	QIC50013.1 3 Brasil	AAA43017.1 1 UK	QBC73619.1 1 UK				
80%	AFW99173.1 2 Ireland	QBC73613.1 7 UK	QHD59426.1 1 USA	BAB12107.1 4 Japan	QHD59411.1 1 USA	QIC50013.1 1 Brasil	QBC73619.1 1 UK	AFV61762.1 1 Japan

I.T. Identity Threshold. In green are highlighted the accession numbers related to the biosynthesized sequences. Under the accession numbers, the number of sequences in each cluster is indicated and the country from which they were isolated.
(a) Re-clustering results for the 187 sequences of the p26 protein retrieved with BLASTP. The Query sequence was the p26 (ABE03841.1)
(b) Re-clustering results for the 142 sequences of the gp45(TM) polypeptide retrieved with BLASTP. The Query sequence was TM-1 (AFW99173.1)

the C-terminal domain, which consists of four α-helices and contains a disulphide bond (Supplementary Fig. 1). The p26 model did not deviate much from the protein geometry of its template (Supplementary Table 1). Typical for globular proteins the majority of the hydrophobic residues are buried in the interior of the protein, whereas the polar residues are on the surface (Supplementary Fig. 2). The structurally buried residues coincide with evolutionary conserved ones while the non-conserved residues are found in exposed positions (Supplementary Fig. 3). Taken together, the physics- and knowledge-based analyses of the generated 3D-structure of p26 of EIAV suggest that the model is sound.

Structural homology models of gp45_1 and gp45_2 were generated, by using as templates the structures of p45 of EIAV (48 % similar to gp45_1 and gp45_2), and gp41 of SIV and HIV (28 % and 21 % similar to gp45_1 and gp45_2). As their templates, the obtained homology models of gp45_1 and gp45_2 consist of trimeric helix-loop-helix structures with a pseudo threefold symmetry axis and they were evaluated in several ways. Not surprisingly, both models display similar overall structure (Supplementary Fig. 4) and protein geometry deviations (Supplementary Table 2), which are comparable with their templates. As expected, hydrophobic residues are mainly buried in the inter-helical interfaces, whereas the polar residues are solvent-accessible (Supplementary Fig. 5). The structurally buried residues are also evolutionary conserved while the non-conserved residues are found in exposed positions (Supplementary Fig. 6). Furthermore, a structural model was built of the double-strain construct with the linker-fused sequences of gp45_1 and gp45_2 in the two possible conformations within the pseudo-three-fold symmetry of the trimer (Supplementary Fig. 7). These two versions were energy minimized and their protein geometry did not display distortions (Supplementary Table 2). In conclusion, the analyses imply that over-all the generated models of gp45 of EIAV are plausible.

Continuous (linear) and discontinuous (conformational) B-cell epitopes were predicted in the p26, gp45_1 and gp45_2 sequences and homology models, respectively (Supplementary Figs. 8–9). The consensus epitope predictions were mapped onto the models of p26, gp45_1 and gp45_2. In p26, the majority of the discontinuous epitopes are found on one side of the N-terminal domain and in the linker connecting the two domains on the opposite side as well as on most of the exposed surface of the C-terminal domain (Supplementary Fig. 10). In both gp45_1 and gp45_2, the discontinuous hotspot epitopes are mainly located in solvent-accessible residues in the beginning and end of the α-helices as well as in the connecting loop region (Supplementary Fig. 11). The same result, but maybe even clearer, is obtained when considering the epitope predictions from all programs, also including the continuous epitope predictions (Fig. 1**a–b**). It is worth to note that the first approximately 70 residues at the N-terminal of p26 do not show

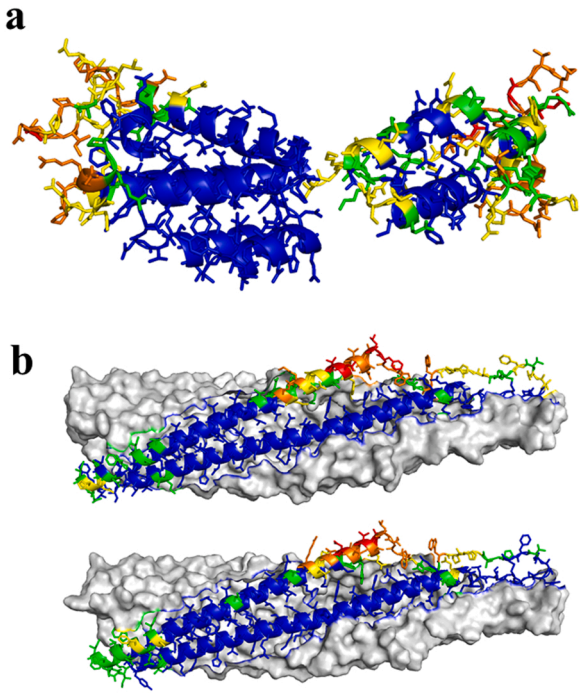

Fig. 1. (a) The consensus predictions for linear and discontinuous epitopes (A_CONS, Supplementary Fig. 8) are displayed on the homology model of p26 in blue with the consensus intervals ≥ 80 % (red), 70–79 % (orange), 60–69 % (yellow) and 50–59 % (green); **(b)** the consensus predictions for linear and discontinuous epitopes (A_CONS, Supplementary Fig. 9) are displayed on the homology models of gp45_1 (top) and gp45_2 (bottom) in blue with the consensus intervals ≥ 80 % (red), 70–79 % (orange), 60–69 % (yellow) and 50–59 % (green).

significant epitopes. The results on p26 are in a quite good agreement with the experimental results obtained by Soutullo et al. (2007) using epitope scanning techniques with overlapping synthetic peptides. It should be added that our analysis correctly identifies a major (immunodominant) C-terminus epitope in the region 193–223 of our sequence, that has already been reported (Soutullo et al. 2007; Hu et al. 2016). In gp45_1 and gp45_2, the consensus epitope predictions identify the major epitope hotspot in a region encompassing the loop (CIEKTHTFC and CVERSHTFC for gp45_1 and gp45_2, respectively) and the beginning of the second α-helix, which protrudes from the rest of the structure in both proteins (Fig. 1**b**). It is again noteworthy that the potentially epitopic region following the loop, found in our prediction, was not clearly identified by Naves et al. in a systematic investigation using epitope scanning techniques with synthetic peptides (Naves et al., 2019).

3.3. Design and production of the EIAV p26 and gp45 (TM) antigens

Antigen constructs of the EIAV p26 and gp45_1_2 double-strain polypeptides, which were designed according to considerations from the sequence clustering and structural modelling analyses described in the previous sections, were expressed in *E. coli* and purified.

Since a low solubility for a gp45 exodomain polypeptide was reported (Du et al., 2018), the aggregation propensity of structural models of gp45_1 (trimer), gp45_1 (monomer) and the double-strain gp45_1_2 was investigated by using the computational approach Aggrescan3D (Supplementary Figure 12). The gp45_1 trimer structure contains exposed aggregative residues in the loop region, which are also found in the other models, whereas additional aggregative surfaces are found on the monomer at the N terminal α-helix and the double-strain gp45_1_2 close to the N- and C-termini. This qualitative analysis predicts a risk of low solubility for both the monomers and the gp45_1_2 polypeptide. To avoid the expression of an insoluble form of the gp45_1_2 double-strain polypeptide, maltose-binding protein (MBP) (Austin et al., 2009) was fused to its N- terminus in an attempt to improve the solubility of the protein. It is worth to note that a p26 protein fused to MBP was already used in an EIAV test with no reported interference due to the presence of MBP (Fontes et al., 2018).

The qualitative analysis of the proteins produced was carried out by western blotting, SDS PAGE and HPLC chromatography (Supplementary Figure 13). Some of the sera were also assessed for reactivity to both polypeptides by western blotting (Supplementary Table 3). The immunoreactive bands obtained from the western blot analysis on the positive control serum were used to normalize the immunoreactive bands obtained from some of them, thus demonstrating that western blotting can be a confirmatory analysis at in-house made ELISA.

3.4. CD analysis of the EIAV p26 and gp45 antigens

To evaluate if the secondary structure implied by the structural modelling is consistent with that of the recombinantly obtained antigens, we performed a spectroscopic study of the polypeptides. CD spectra of p26 and gp45_1_2 are reported in Fig. 2. The predicted secondary structures deduced from the experimental CD spectra appear qualitatively coherent with those implied by the modelled 3D structures of p26 and gp45_1_2 (data not shown), thus suggesting an overall correct folding of the polypeptides.

3.5. ELISA to test the EIAV p26 and gp45 antigens

An indirect in-house ELISA was optimized from checkerboard titration performed for each of the antigens as well as combining p26 and gp45_1_2 polypeptide in a fixed mass ratio of 1:5.

All the positive samples were confirmed with a commercial indirect ELISA test (Eradikit -https://www.in3diagnostic.com/en/eradikit-eiav-elisa/) that recognizes antibodies to both p26 and gp45 antigens or with western blotting performed with the p26 and gp45_1_2 polypeptide

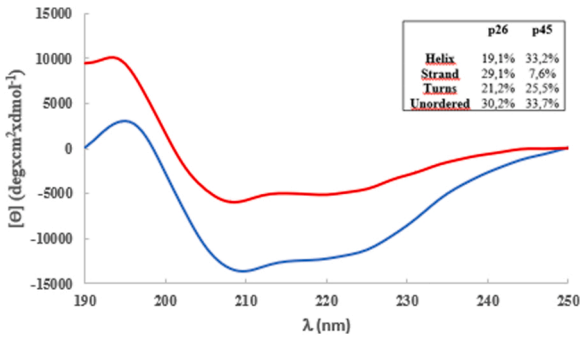

Fig. 2. CD spectra of p26 (red) and gp45_1_2 (blue) polypeptides in 20 mM phosphate buffer, pH 6.8. The inset shows the fractions of secondary structure estimated from the deconvoluted CD spectra by the Contin method.

antigens separately. The negative samples were confirmed with the commercial kit. The positive samples were retested with our in-house p26, gp45 and combined p26/ gp45_1_2 ELISA. Negative samples were only tested with the combined p26/gp45 ELISA. A preliminary Roc analysis relative to the combined p26/gp45 ELISA showed 96 % sensitivity and 100 % specificity for our test when a cut-off value for the reactivity of 50 % was applied using the 210 samples included in the analysis (103 positive and 107 negative). Fig. 3.

Considering only the 99 samples correctly identified as positive with the in-house test, 3 samples were reactive only to the gp45_1_2 antigen and one sample was reactive only to the p26 ELISA protein; the remaining samples were reactive to both antigens. Thus, the gp45_1_2 antigen in the combined ELISA test contributed to enhance its diagnostic sensitivity. It is important to note that we found 100 % agreement between our combined ELISA and the commercial Eradikit test in the analysis of the 210 samples.

4. Discussion and conclusions

Lentiviruses exhibit heterogeneous and complex populations with similar but non-identical genomes, since they evolve rapidly mainly due to the large population size, the high replication rate, the defective proofreading ability of their RNA-dependent DNA polymerase (Reverse

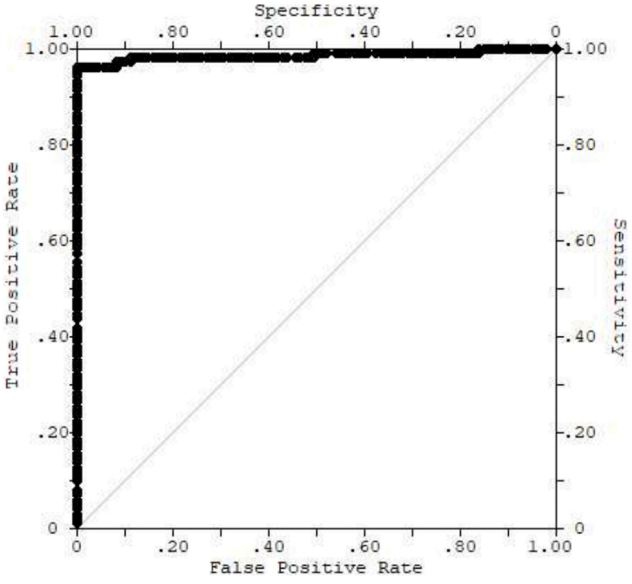

Fig. 3. ROC analysis- DATA Points: 210, Sensitivity: 96 %, Specificity: 100 %, Cut-off: 50 % reactivity. Area under the non-parametric ROC = 0.9851 Standard Error of Area = 0.0096.

Transcriptase); therefore, many authors have classified them as viral quasispecies (Holland et al., 1992).

Regarding EIAV, recent findings show an extended genetic variability between geographically distinct EIAV strains that were greatly underestimated in the past (Capomaccio et al., 2012a, 2012b; Dong et al., 2013; Quinlivan et al., 2013). As a result, the development of universally applicable direct detection molecular diagnostics were hindered (Cappelli et al., 2017). It is now quite clear that additional EIAV molecular characterization is required because the known sequences are numerically limited with respect to the circulating EIAV strains (Cappelli et al., 2017), even when compared to the number of known sequences of other lentiviruses such as HIV and SRLV. So far, the potential consequences of the genetic and structural variability of EIAV were scarcely considered when developing serological tests, in particular for ELISA. This could probably be ascribed to the almost exclusive use of the capsid p26 protein as target antigen, since this protein was considered quite conserved between the different strains. In addition, historical reasons have also led to the use of a unique strain (Wyoming) in most of the serological tests.

In the present investigation, a minimum of 82 % of identity was found in the sequences retrieved with respect to the query sequence in our Blastp analysis for p26. In the EIAV gp45 exodomain sequences a minimum 70 % of identity was found that is very similar to the result obtained for the corresponding TM protein of SRLV (68 %) (Ostuni et al., 2021). It is also possible that the differences in identity will be increased with the growing number of strains sequenced in the future. We reasoned that methodologies could be developed so as to take into account the intrinsic diversity of EIAV strains when designing protein/peptide antigens for serological diagnostic purposes with similar approaches as previously reported for SRLV (Ostuni et al., 2021). Starting from the known sequences, a simple algorithm reduced the data set to a reasonable number of representative protein sequences. Clustering with an acceptable threshold identity appeared quite appropriate for this scope. In the present investigation, the clustering approach has been slightly modified compared to the previously reported one (Ostuni et al., 2021). In fact, a “two steps” approach was used: an initial clustering at 95 % of identity was followed by a clustering of the centroids found. In this way, the preliminary groups contain very similar sequences, for example those likely derived from a single outbreak, that are hence represented by a single centroid in the successive analysis. This should result, in our opinion, in a more balanced and readable clustering information, without biasing the results of the analysis.

The clustering analysis performed provided a number of sequences of the same protein corresponding to different strains at various levels of identity. In our test, we used for p26 a single sequence of an isolate from Argentina obtained at 80 % identity. The expressed and used double-strain polypeptide antigen was obtained by the fusion of a multi-epitope region of the gp45 exodomain derived from two selected representative sequences. To identify the location of the potential epitopes in this domain, a preliminary consensus prediction based on linear and conformational (discontinuous) epitopes was performed. The knowledge of the protein 3D structure was necessary for the conformational epitope predictions and, since no experimentally determined structure was available for the specific sequences selected, homology modelling was used. In the case of the gp45_1_2 construct, an "artificial" fusion polypeptide was created through a linker and probably each antigen incorporated in the double-strain polypeptide very closely mimics its native state, against which most likely the antibodies in the infected animals are raised. Considering the results of the aggregative study, we decided to add the maltose binding protein in the gp45_1_2 fusion polypeptide, that likely enhanced the solubility profile of the construct and allowed us to obtain the polypeptide expressed in solution, thus avoiding a setback previously reported for a similar polypeptide (Du et al., 2018).

It is worth to note that the TM exodomain of gp45 contains an immunodominant epitope in the loop region. This epitope is common to many lentiviruses (Saman et al., 1999) and comprises a sequence of few residues between two conserved cysteines that are engaged in a disulphide bond. In the case of EIAV, as for other lentiviruses (Saman et al., 1999), the propensity of this region to be a good epitope much depends on the conformational restriction imposed by the disulphide bond. In fact, it has been reported that a cyclic peptide that mimics this immunodominant sequence of gp45 showed excellent reactivity in an EIAV ELISA test, thus suggesting that its functional activity depends significantly on its conformation, because very low reactivity was observed in the linear form of the peptide (Soutullo et al., 2001). Our goal was to have as many epitopes as possible in the TM antigen construct and a folded conformation of the immunodominant epitope described above. The results show that, very likely, the gp45_1_2 polypeptide has a folded structure in the loop regions due to the favourable intramolecular interactions between the four α-helices (two helix-loop-helix motifs) predicted in the model, thus retaining the immunogenic properties in the region between the two cysteines without any need of chemical oxidation of the polypeptide.

In our current ELISA, just one p26 antigen and one double-strain TM antigen are used, however, by further exploiting the clustering and epitope analyses, the set of useful p26 and gp45 antigens could easily be expanded to represent more accurately the structural and immunogenic variability of the deposited sequences. In any case, we believe that the present test can complement well the available ELISA tests that use a single protein from a single strain. In addition, the recombinant soluble gp45_1_2 double strain polypeptide was used in western blotting and can be potentially useful in an AGID test.

Immunoinformatics was applied to design suitable antigens for human and veterinary vaccine development over the last two decades (Tomar and De, 2014; De Groot et al., 2020). However, application of bioinformatic and biostructural tools to design antigens for veterinary immunoassays are quite a few. Our effort is focused on introducing an in-silico analysis coupled with recombinant protein expression starting from synthetic genes in the veterinary diagnostic sector, which hopefully could benefit from this approach.

Further validation is required to fully assess the use of the described test in the routine serological diagnosis of EIA.

CRediT authorship contribution statement

Angela Ostuni: Investigation, Visualization, Writing – review & editing; Valentina Iovane: Investigation;Magnus Monné: Investigation, Writing – original draft, Visualization; Maria Antonietta Crudele: Investigation, Data curation, Formal analysis; Maria Teresa Scicluna: Investigation, review; Roberto Nardini: Investigation, review; Paolo Raimondi: Resources; Raffaele Frontoso: Investigation; Raffaele Boni: Investigation; Alfonso Bavoso: Conceptualization, Project administration, Methodology, Writing – original draft, Writing - review & editing, Funding Acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jviromet.2023.114704.

References

Alvarez, I., Gutierrez, G., Vissani, A., Rodriguez, S., Barrandeguy, M., Trono, K., 2007. Standardization and validation of an agar gel immunodiffusion test for the diagnosis of equine infectious anemia using a recombinant p26 antigen, 3-4 Vet. Microbiol. 15 (121), 344–351. https://doi.org/10.1016/j.vetmic.2007.01.007.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., Ben-Tal, N., 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44, W344–W350.

Austin, B.P., Nallamsetty, S., Waugh, D.S., 2009. Hexahistidine-tagged maltose-binding protein as a fusion partner for the production of soluble recombinant proteins in Escherichia coli. Methods Mol. Biol. 498, 157–172. https://doi.org/10.1007/978-1-59745-196-3_11.

Buzon, V., Natrajan, G., Schibli, D., Campelo, F., Kozlov, M.M., Weissenhorn, W., 2010. Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. PLoS Pathog. 6 (5), e1000880 https://doi.org/10.1371/journal.ppat.1000880.

Capomaccio, S., Willand, Z.A., Cook, S.J., Issel, C.J., Santos, E.M., Reis, J.K., Cook, R.F., 2012a. Detection, molecular characterization and phylogenetic analysis of full-length equine infectious anemia (EIAV) gag genes isolated from Shackleford Banks wild horses. Vet. Microbiol. 157 (3–4), 320–332. https://doi.org/10.1016/j.vetmic.2012.01.015.

Capomaccio, S., Cappelli, K., Cook, R.F., Nardi, F., Gifford, R., Marenzoni, M.L., Passamonti, F., 2012b.). Geographic structuring of global EIAV isolates: a single origin for New World strains? Virus Res. 163 (2), 656–659.

Cappelli, K., Cook, R.F., Stefanetti, V., Passamonti, F., Autorino, G.L., Scicluna, M.T., Coletti, M., Verini Supplizi, A., Capomaccio, S., 2017. Deep sequencing and variant analysis of an Italian pathogenic field strain of equine infectious anaemia virus. Transbound. Emerg. Dis. 64 (6), 2104–2112. https://doi.org/10.1111/tbed.12631.

Cursino, A.E., Vilela, A.P.P., Franco-Luiz, A.P.M., de Oliveira, J.G., Nogueira, M.F., Júnior, J.P.A., de Aguiar, D.M., Kroon, E.G., 2018. Equine infectious anemia virus in naturally infected horses from the Brazilian Pantanal. Arch. Virol. 163 (9), 2385–2394. https://doi.org/10.1007/s00705-018-3877-8.

De Groot, A.S., Moise, L., Terry, F., Gutierrez, A.H., Hindocha, P., Richard, G., Hoft, D.F., Ross, T.M., Noe, A.R., Takahashi, Y., Kotraiah, V., Silk, S.E., Nielsen, C.M., Minassian, A.M., Ashfield, R., Ardito, M., Draper, S.J., Martin, W.D., 2020. Better epitope discovery, precision immune engineering, and accelerated vaccine design using immunoinformatics tools. Front. Immunol. 11, 442. https://doi.org/10.3389/fimmu.2020.00442.

Dick, R.A., Xu, C., Morado, D.R., Kravchuk, V., Ricana, C.L., Lyddon, T.D., Broad, A.M., Feathers, J.R., Johnson, M.C., Vogt, V.M., Perilla, J.R., Briggs, J., Schur, F., 2020. Structures of immature EIAV Gag lattices reveal a conserved role for IP6 in lentivirus assembly. PLoS Pathog. 16 (1), e1008277 https://doi.org/10.1371/journal.ppat.1008277.

Dong, J.-B., Zhu, W., Cook, F.R., Goto, Y., Horii, Y., Haga, T., 2013. Identification of a novel equine infectious anemia virus field strain isolated from feral horses in southern Japan. J. Gen. Virol. 94 (Pt 2), 360–365.

Du, C., Hu, Z., Hu, S.D., Lin, Y.Z., Wang, X., Li, Y.J., 2018. Development and Application of an Indirect ELISA for the Detection of gp45 Antibodies to Equine Infectious Anemia Virus. Journal of Equine Veterinary Science 62, 76–80. https://doi.org/10.1016/j.jevs.2017.10.018.

Fathi, A., Hala, A.S., Mona, M.S., 2014. The immunological and serological studies of Equine Infectious Anemia in Egyptian horses. Suez Canal Vet. Med. J. SCVMJ 19 (2), 149–154. https://doi.org/10.21608/scvmj.2014.65332.

Fiser, A., Sali, A., 2003. Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol. 374, 461–491. https://doi.org/10.1016/S0076-6879(03)74020-8.

Fontes, K.F.L.P., Silva-Júnior, L.C., Nascimento, S.A., Chaves, D.P., Pinheiro-Júnior, J.W., Freitas, A.C., Castro, R.S., Jesus, A.L.S., 2018. Enzyme-linked immunosorbent assay and agar gel immunodiffusion assay for diagnosis of equine infectious anemia employing p26 protein fused to the maltose-binding protein. Arch. Virol. 163 (10), 2871–2875. https://doi.org/10.1007/s00705-018-3923-6.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224.

Grego, E., Profiti, M., Giammarioli, M., Giannino, L., Rutili, D., Woodall, C., Rosati, S., 2002. Genetic heterogeneity of small ruminant lentiviruses involves immunodominant epitope of capsid antigen and affects sensitivity of single-strain-based immunoassay. Clin. Diagn. Lab. Immunol. 9 (4), 828–832. https://doi.org/10.1128/cdli.9.4.828-832.2002.

Hofmann, K., Stoffel, W., 1993. TMbase - a database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler 374, 166.

Holland, J.J., de La Torre, J.C., Steinhauer, D.A., 1992. RNA virus populations as quasispecies. Curr. Top. Microbiol Immunol. 176, 1–20.

Hu, Z., Chang, H., Chu, X., Li, S., Wang, M., Wang, X., 2016. Identification and characterization of a common B-cell epitope on EIAV capsid proteins. Appl. Microbiol Biotechnol. 100 (24), 10531–10542. https://doi.org/10.1007/s00253-016-7817-9.

Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics (Oxford, England) 26 (5), 680–682. https://doi.org/10.1093/bioinformatics/btq003.

Issel, C.J., Foil, L.D., 2015. Equine infectious anaemia and mechanical transmission: man and the wee beasties. Rev. Sci. Tech. 34 (2), 513–523. https://doi.org/10.20506/rst.34.2.2376.

Issel, C.J., Scicluna, M.T., Cook, S.J., Cook, R.F., Caprioli, A., Ricci, I., Rosone, F., Craigo, J.K., Montelaro, R.C., Autorino, G.L., 2013. Challenges and proposed solutions for more accurate serological diagnosis of equine infectious anaemia. Vet. Rec. 172 (8), 210. https://doi.org/10.1136/vr-2012-100735.

Jara, M., Frias-De-Diego, A., Machado, G., 2020. Phylogeography of equine infectious anemia virus. Front. Ecol. Evol. 8, 127.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., Karplus, K., 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins 77 (Suppl 9), 114–122.

Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Computational Biology 8 (12), e1002829. https://doi.org/10.1371/journal.pcbi.1002829.

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

Leroux, C., Cadoré, J.L., Montelaro, R.C., 2004. Equine Infectious Anemia Virus (EIAV): what has HIV's country cousin got to tell us? Vet. Res. 35 (4), 485–512. https://doi.org/10.1051/vetres:2004020.

Liang, S., Zheng, D., Zhang, C., Zacharias, M., 2009. Prediction of antigenic epitopes on protein surfaces by consensus scoring. BMC Bioinforma. 10, 302.

Malossi, C.D., Fioratti, E.G., Cardoso, J.F., Magro, A.J., Kroon, E.G., Aguiar, D.M., Borges, A.M.C.M., Nogueira, M.F., Ullmann, L.S., Araujo Jr., J.P., 2020. High genomic variability in equine infectious anemia virus obtained from naturally infected horses in Pantanal, Brazil: an endemic region case. Viruses 12 (2), 207. https://doi.org/10.3390/v12020207.

Martinelli, F., Cuviello, F., Pace, M.C., Armentano, M.F., Miglionico, R., Ostuni, A., Bisaccia, F., 2018. Extracellular ATP regulates CD73 and ABCC6 expression in HepG2 Cells. Front. Mol. Biosci. 5, 75. https://doi.org/10.3389/fmolb.2018.00075.

Miles, A.J., Ramalli, S.G., Wallace, B.A., 2021. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. Protein Sci. https://doi.org/10.1002/pro.4153.

Nardini, R., Autorino, G.L., Issel, C.J., Cook, R.F., Ricci, I., Frontoso, R., Rosone, F., Scicluna, M.T., 2017. Evaluation of six serological ELISA kits available in Italy as screening tests for equine infectious anaemia surveillance. BMC Vet. Res 13 (1), 105. https://doi.org/10.1186/s12917-017-1007-6.

Naves, J.H.F.F., Oliveira, F.G., Bicalho, J.M., Santos, P.S., Machado-de-Ávila, R.A., Chavez-Olortegui, C., Leite, R.C., Reis, J.K.P., 2019. Serological diagnosis of equine infectious anemia in horses, donkeys and mules using an ELISA with a gp45 synthetic peptide as antigen. J. Virol. Methods 266, 49–57. https://doi.org/10.1016/j.jviromet.2018.12.009.

Ostuni, A., Monné, M., Crudele, M.A., Cristinziano, P.L., Cecchini, S., Amati, M., De Vendel, J., Raimondi, P., Chassalevris, T., Dovas, C.I., Bavoso, A., 2021. Design and structural bioinformatic analysis of polypeptide antigens useful for the SRLV serodiagnosis. J. Virol. Methods 297, 114266. https://doi.org/10.1016/j.jviromet.2021.114266.

Ponomarenko, J., Bui, H.-H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinforma. 9, 514.

Provencher, S.W., Glockner, J., 1981. Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20, 33–37.

Quinlivan, M., Cook, F., Kenna, R., Callinan, J.J., Cullinane, A., 2013. Genetic characterization by composite sequence analysis of a new pathogenic field strain of equine infectious anemia virus from the 2006 outbreak in Ireland. J. Gen. Virol. 94 (Pt 3), 612–622. https://doi.org/10.1099/vir.0.047191-0.

Rosati, S., Profiti, M., Lorenzetti, R., Bandecchi, P., Mannelli, A., Ortoffi, M., Tolari, F., Ciabatti, I.M., 2004. Development of recombinant capsid antigen/transmembrane epitope fusion proteins for serological diagnosis of animal lentivirus infections. J. Virol. Methods 121 (1), 73–78. https://doi.org/10.1016/j.jviromet.2004.06.001.

Rubinstein, N.D., Mayrose, I., Martz, E., Pupko, T., 2009. Epitopia: a web-server for predicting B-cell epitopes. BMC Bioinformatics 10, 287. https://doi.org/10.1186/1471-2105-10-287.

Saha, S., Raghava, G.P.S., 2004. BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. In: Nicosia, G., Cutello, V., Bentley, P.J., Timmis, J. (Eds.), Artificial Immune Systems. Springer, Berlin, Heidelberg, pp. 197–204.

Saha, S., Raghava, G.P.S., 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65, 40–48.

Saman, E., Van Eynde, G., Lujan, L., Extramiana, B., Harkiss, G., Tolari, F., Gonzàlez, L., Amorena, B., Watt, N., Badiola, J., 1999. A new sensitive serological assay for detection of lentivirus infections in small ruminants. Clin. Diagn. Lab. Immunol. 6 (5), 734–740.

Scicluna, M.T., Autorino, G.L., Nogarol, C., Ricci, I., Frontoso, R., Rosone, F., Nardini, R., 2018. Validation of an indirect ELISA employing a chimeric recombinant gag and env peptide for the serological diagnosis of equine infectious anemia. J. Virol. Methods 251, 111–117. https://doi.org/10.1016/j.jviromet.2017.10.002.

Singha, H., Goyal, S.K., Malik, P., Khurana, S.K., Singh, R.K., 2013. Development, evaluation, and laboratory validation of immunoassays for the diagnosis of equine infectious anemia (EIA) using recombinant protein produced from a synthetic p26 gene of EIA virus. Indian J. Virol. 24 (3), 349–356. https://doi.org/10.1007/s13337-013-0149-9.

Soutullo, A., Verwimp, V., Riveros, M., Pauli, R., Tonarelli, G., 2001. Design and validation of an ELISA for equine infectious anemia (EIA) diagnosis using synthetic peptides. Vet. Microbiol 79 (2), 111–121. https://doi.org/10.1016/s0378-1135(00)00352-7.

Soutullo, A., Santi, M.N., Perin, J.C., Beltramini, L.M., Borel, I.M., Frank, R., Tonarelli, G. G., 2007. Systematic epitope analysis of the p26 EIAV core protein. J. Mol. Recognit. 20 (4), 227–237. https://doi.org/10.1002/jmr.825.

Sreerama, N., Venyaminov, S.Y., Woody, R.W., 2000. Estimation of protein secondary structure from CD spectra: Inclusion of denatured proteins with native protein in the analysis. Anal. Biochem. 287, 243–251.

Sweredoski, M.J., Baldi, P., 2009. COBEpro: a novel system for predicting continuous B-cell epitopes. Protein Eng. Des. Sel. 22, 113–120. https://doi.org/10.1007/978-1-4939-1115-8_3.

Tomar, N., De, R.K., 2014. Immunoinformatics: A Brief Review. In: De, R., Tomar, N. (Eds.), Immunoinformatics. Methods in Molecular Biology (Methods and Protocols), vol 1184. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-1115-8_3.

Wang, X.F., Liu, Q., Wang, Y.H., Wang, S., Chen, J., Lin, Y.Z., Ma, J., Zhou, J.H., Wang, X., 2018. Characterization of equine infectious anemia virus long terminal repeat quasispecies in vitro and in vivo. e02150-17 J. Virol. 92. https://doi.org/10.1128/JVI.02150-17.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46 (W1), W296–W303. https://doi.org/10.1093/nar/gky427.

Wegdan, H.A., Sahar, M.E., Ballal, A., Intisar, K.S., Shaza, M.M., Algezoli, O.A., Ihsan, H. A., Baraa, A.M., Taha, K.M., Nada, E.M., Manan, A.A., Ali, Y.H., Nouri, Y.M., 2016. Sero prevalence of equine infectious anemia (EIA) virus in selected regions in Sudan. Microbiol. Res. J. Int. 18 (1), 1–6. https://doi.org/10.9734/MRJI/2017/27451.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al., 2018. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 27, 293–315.

Zambrano, R., Jamroz, M., Szczasiuk, A., Pujols, J., Kmiecik, S., Ventura, S., 2015. AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures. Nucleic Acids Res. 43, W306–W313. https://doi.org/10.1093/nar/gkv359.

Zhou, C., Chen, Z., Zhang, L., Yan, D., Mao, T., Tang, K., Qiu, T., Cao, Z., 2019. SEPPA 3.0-enhanced spatial epitope prediction enabling glycoprotein antigens. Nucleic Acids Res 47, W388–W394.