



Identification of potential CD8+ epitopes in pp62 polyprotein of African swine fever virus using computational immunology

MARK LESTER C. GALICIA^{1*}, DALE JONATHAN M. MORALES¹, PRECIOUS GRACE B. POGADO¹,
ASHLEY L. QUEBRADO¹, LEANA RICH HERRERA-ONG²

¹Department of Physical Sciences, College of Science, Polytechnic University of the Philippines, Manila, Philippines

²Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines, Manila, Philippines

Received: 11 August 2022; revised: 30 January 2023; accepted: 3 April 2023

Abstract

The highly infectious African swine fever virus (ASFV) is currently the only known DNA arbovirus within the *Asfarviridae* family which primarily infects domestic pigs and wild boars. African swine fever (ASF) leads to a mortality rate of up to 100% which has caused massive socio-economic losses worldwide. Previous research indicates that ASFV's virulence can be attributed to polyprotein pp62, which plays a crucial role in viral assembly and core maturation. This particular study utilized *in silico* analysis to identify highly conserved cytotoxic T-cell epitopes in pp62 that can potentially serve as key components for future ASFV vaccines. To achieve this, the researchers retrieved, clustered, and aligned the peptide sequences of pp62. Subsequently, the aligned sequences were analyzed to identify epitopes that bind promiscuously to the swine major histocompatibility complex I (MHC I) alleles and exhibiting MHC IC₅₀ values <500 nM. Additionally, peptide sequences with positive proteasome and TAP scores were considered. Potential cross-reactivity was assessed by comparing the peptide sequences against available proteome sequences of *Sus scrofa domestica* in various databases. Furthermore, molecular docking was conducted to evaluate the binding of candidate epitopes with swine leukocyte antigen-1*0401 (SLA-1*0401). The dissociation constants, binding energies, root mean square deviation, and root mean square fluctuation values for the SLA-epitope complexes were compared with a positive reference. In the course of the study, 21 highly conserved CD8+ epitopes were identified, out of which four were further assessed for their potential immunogenicity. The results demonstrated that the highly conserved CD8+ epitopes discovered in this study are promising for integration into future ASFV vaccine formulations. As preliminary data, it is anticipated that these findings will subsequently undergo *in vitro* and *in vivo* studies in the future.

Key words: African swine fever virus, vaccine, epitopes, *in silico*, pp62 polyprotein

Introduction

African swine fever (ASF) continues to have a detrimental impact on the global pig industry, particularly in Asia—world's largest pig producer (Kedkovid et al., 2020). Despite efforts to control the disease through measures like culling infected pigs and implementing strict sanitary and monitoring policies, the lack of available vaccines for commercial use has hindered any significant progress. In addition, current management approaches and strategies have not effectively mitigated

the huge socio-economic losses caused by ASF (Blome et al., 2020).

According to the World Organization for Animal Health (WOAH) report in December 2022, ASF has affected 45 countries across five regions in the world since January 2020. Total animal losses have exceeded two million, wherein over 1 191 000 pigs and 37 000 wild boars were affected (WOAH, 2022).

During the initial waves of outbreaks, small-to-medium pig farms with weak biosecurity measures have

* Corresponding author: Department of Physical Sciences, College of Science, Polytechnic University of the Philippines, Anonas St., Sta. Mesa, Manila, Philippines; e-mail: marklestergalicia007@gmail.com; mcgalicia@up.edu.ph

been identified as a crucial contributing factor (Kedkovid et al., 2020). This is because small backyard farms, despite their size, can have collective supply strain which has more significant impact as compared to larger pig farms (Cooper et al., 2022). In countries like China and Vietnam, small farms constitute approximately 80% of all pig farms. The proximity of small and larger farms can potentially facilitate the rapid spread of ASFV in Asian farms (Kedkovid et al., 2020).

ASFV strains circulating in Asia and Europe, particularly the highly virulent genotype II strains, are known to cause acute and peracute diseases within a short span of 7–10 days, often resulting in a mortality rate of up to 100%. Common symptoms of ASF include loss of appetite, high fever, gastrointestinal and respiratory manifestations, ataxia, cyanosis, and rapid-onset death. Clinical signs may vary depending on the specific strain of the virus (Blome et al., 2020).

ASFV is a DNA arbovirus belonging to the *Asfarviridae* family and genus *Asfivirus*. The viral genome of ASFV encodes approximately 150–200 viral proteins, including 68 structural and 100 nonstructural proteins which further highlights its complexity (Wang et al., 2021). Its complication poses challenges in devising effective prevention and control measures for ASFV infections (Andrés et al., 2020).

The core shell of the ASFV viral particle is composed of processed products derived from two polyprotein precursors: pp220 and pp62. Between the two, polyprotein pp62 plays a crucial role in the proper assembly and maturation of the viral core. The polyprotein is consist of p15, p35, and p8 with a relative molecular weight of 60.5 kDa (Urbano et al., 2021; Jia et al., 2017), and is encoded by *CP503R* gene. Its sequence exhibits a high degree of conservation across different ASFV strains (Zhong et al., 2022). Its presence is necessary for proper assembly and maturation of viral particle's core. In its absence during viral morphogenesis, a homogeneous population of empty particles is produced. Given its vital role in ASFV pathogenesis, eliciting a successful immune response against pp62 can potentially prevent infection and serve as effective immunotherapeutic agent in combating ASF (Suárez, 2010).

The field of immunoinformatics has greatly enhanced the current understanding of diagnosis, pathogenesis, immune response, and computational vaccinology (Oli et al., 2020). With the recent advancements in immuno-

informatics tools and databases, the search for protein sequences and the identification of MHC-binding motifs for the epitopes have become more efficient and cost-effective. This study took the advantages that computational vaccinology approach can offer in strategically and efficiently mapping candidate CD8+ epitopes in ASFV. The main objective of the study was to identify highly conserved epitopes derived from the pp62 polyprotein of ASFV that can bind promiscuously to dominant swine leukocyte antigen (SLA). Immunoinformatics tools and databases were utilized to obtain protein sequences and identify epitopes which serve as promising candidates for the development of epitope-based vaccines. It is anticipated that the epitopes identified in this preliminary research will undergo *in vitro* and *in vivo* studies to determine their suitability for inclusion in future epitope-based vaccines against ASF.

Materials and methods

Retrieval and identification of highly conserved pp62 sequences

To generate information on the post-translationally modified sites of the pp62 polyprotein, related data on its sequence were obtained from the Universal Protein Resource (UniProt) server (<https://www.uniprot.org/>). The ASFV protein sequences were collected from the National Center for Biotechnology Information (NCBI) protein database (<https://www.ncbi.nlm.nih.gov/protein/>) on March 21, 2022, by filtering the sequences based on a length of 530–534 residues. To eliminate redundant protein sequences (Fu et al., 2012), sequences from NCBI were uploaded and analyzed in CD-HIT server (<https://sites.google.com/view/cd-hit?pli=1>) using 1.00 as identity cut-off value. Unique, representative sequences obtained from the CD-HIT server were subsequently aligned in Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

Conserved sequences were identified using the Protein Variability Server (PVS) tool (<http://imed.med.ucm.es/PVS/>). Consensus sequence was chosen as the representative sequence for further analysis (García-Boronat et al., 2008). Conserved fragments containing at least nine residues were included, while fragments with a Shannon variability scores ≥ 0.1 were masked and excluded. The resulting fragments were tabulated based on their alignment positions in the sequence.

CD8⁺ epitope mapping

The highly conserved sequences obtained were analyzed in the Immune Epitope Database (IEDB) using the Proteasomal cleavage/TAP transport/MHC class I combined predictor (<http://tools.iedb.org/processing/>) and the netMHCcons prediction method. The netMHCcons method is a consensus prediction method that combines netMHCpan, PickPocket, and netMHC (ANN) to provide more reliable results (Karosiene et al., 2012). The prediction method included all 45 available SLA alleles to generate nonameric CD8⁺ epitopes. The immunoproteasome type was used for proteasomal cleavage prediction, and the default parameters were used for TAP transport prediction. The tabulated results from these prediction methods include the proteasome score, TAP score, MHC score, processing score, and MHC-binding score for each peptide. Peptides with an MHC IC₅₀ scores ≤ 500 nM (Fleri et al., 2017; Jensen et al., 2018) and positive TAP and proteasome scores were selected for subsequent analysis.

To assess for potential cross-reactivity, mapped epitopes were compared against available protein sequences of domestic pigs (*Sus scrofa domesticus*) using BlastP (Protein Basic Local Alignment Search Tool) in NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). The database search excluded model, uncultured, and environmental sample sequences. Epitopes that did not exhibit significant sequence similarities with the sequences in the databases were selected as candidate ASFV pp62 CD8⁺ epitopes.

In silico validation of candidate epitopes

To further support the binding ability of mapped candidate epitopes with the dominant swine leukocyte antigen (SLA-1*0401), molecular docking analysis were conducted. Currently, it is the only available SLA I PDB crystal structure in all the protein structure databases. The crystal structure of SLA-1*0401 (PDB 3QQ3) was downloaded from the Protein Data Bank. The structure contains an influenza epitope (NSDTVGSWSW) bound to the peptide-binding groove of SLA-1*0401; thus, the influenza epitope was removed from the complex to enable docking of each candidate epitope in GalaxyPepDock server (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=PEPDOCK>). Other unrelated heterostructures were also eliminated as part of the complex preparation

for the docking procedures (Hernández-Santoyo et al., 2013).

Resulting complexes with the highest scores were retrieved and refined using the RefineComplex tool (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>). Refined complexes were visualized using Chimera X version 1.3. To further analyze the favorability and stability of binding, the binding energy and dissociation constant were estimated at 38.6 using the PRODIGY server (<https://wenmr.science.uu.nl/prodigy/>). These values were then compared to a positive reference peptide (NSDTVGSWSW).

To further assess the stability and flexibility of the SLA-epitope complex formation, molecular dynamics (MD) simulation was performed. The simulation was carried out using the OpenMM toolkit on the Google Colab framework. The root mean square deviation (RMSD) and root mean square fluctuation (RMSF) plots were obtained to analyze the system's behavior over time (Arantes et al., 2021). The AMBER MD setup was utilized for the MD simulation procedures. Solvation parameters were determined using the Amber ff19SB force field while the OPC water model was used in constructing the topology and state file (Tian et al., 2020). The simulation duration was set to 5 ns, and an energy minimization procedure consisting of 1000 steps was performed before the simulation. The system was neutralized with a 0.15 M NaCl concentration. A force constant of 500 kJ/mol was used to maintain a stable system at a temperature of 312.32 K and a pressure of 1 bar.

Results

Highly conserved sequences in ASFV pp62 polyprotein

Computational analysis of pp62 polyprotein sequence aided in the identification of 12 highly conserved fragments with a Shannon variability index ≤ 0.1. These fragments exhibit minimal variability among different strains of ASFV. The length of these highly conserved fragments varies from 11 to 86 residues. Table 1 provides the detailed information for the positions of these highly conserved fragments along the pp62 amino acid sequence.

Highly conserved and promiscuous CD8⁺ epitopes in ASFV pp62

Table 2 presents a total of 21 highly conserved CD8⁺ epitopes which were identified from the analysis of con-

Table 1. Highly conserved fragments of ASFV pp62 polyprotein

Start	End	Sequence
1	86	MPSNMKQFCKISVWLQQHDPDLLEIINNLCMLGNLSAAKYKHGVTFIYPKQAKIRDEIKKHAYSNDPSQA IKTLESLILPFYIPTP
88	105	EFTGEIGSYTGVKLEVEK
107	120	EANKVILKNGEAVL
122	164	PAADFKPFPDRRLAVWIMESGSMPLEGGPPYKRKKEGGNDPPV
166	186	KHISPYTPRTRIAIEVEKAFD
188	253	CMRQNWCSVNNPYLAKSVSLLSFLSLNHPTEFIKVLPLIDFDPLVTFYLLLEPYKTHGDDFLIPET
255	291	LFGPTWNGTDLYQSAMLEFKKFFFTQITRQTFMDIAD
293	312	ATKEVDVPICYSDPETVHSY
314	324	NHVRTEILHHN
326	464	VNKVTTPNLVVQAYNELETNTIRHYGPIFPESTINALRFWKKLWQDEQRFVIHGLHRTLMDQPTYETSE FAEIVRNLRFSRPGN
472	504	TSSLTDLVPTRLHSFLNHNLSKLNRAQQTV
506	530	NILSNDCLNQLKHYVKHTGKNEILK

served fragments of pp62. It includes information on the sequence locations of the epitopes, TAP scores, proteasome scores, and IC_{50} scores, along with their corresponding SLA binders. These epitopes were selected based on MHC IC_{50} scores and positive TAP and proteasome scores for effective MHC-binding ability.

Moreover, among the 21 CD8⁺ epitopes identified, five epitopes (FINSTDFLY, GTDLYQSAM, STDFLYTAI, TEFIKVLPL, and TLESLILPF) were classified as promiscuous because they can also bind to more than one MHC allele (Bhasin and Raghava, 2007).

Evaluation of binding energy and affinity to SLA I

The study focused on the molecular docking of candidate epitopes with SLA-1*0401, including the four promiscuously binding epitopes from pp62: FINSTDFLY, GTDLYQSAM, STDFLYTAI, and TLESLILPF. The positive reference used in the study was the influenza epitope (NSDTVGWSW) which is originally bound to SLA-1*0401. Figure 1 illustrates the 3D structures of the epitope-SLA complexes, showcasing the binding groove of the swine MHC I.

To visualize the interaction between the SLA-1*0401-epitope complexes, protein–ligand interactions were observed. Majority of interactions observed were hydrogen bonding between the candidate epitopes and SLA-1*0401, as H-bonds are known to facilitate binding between the protein and the ligand of interest (Chen et al.,

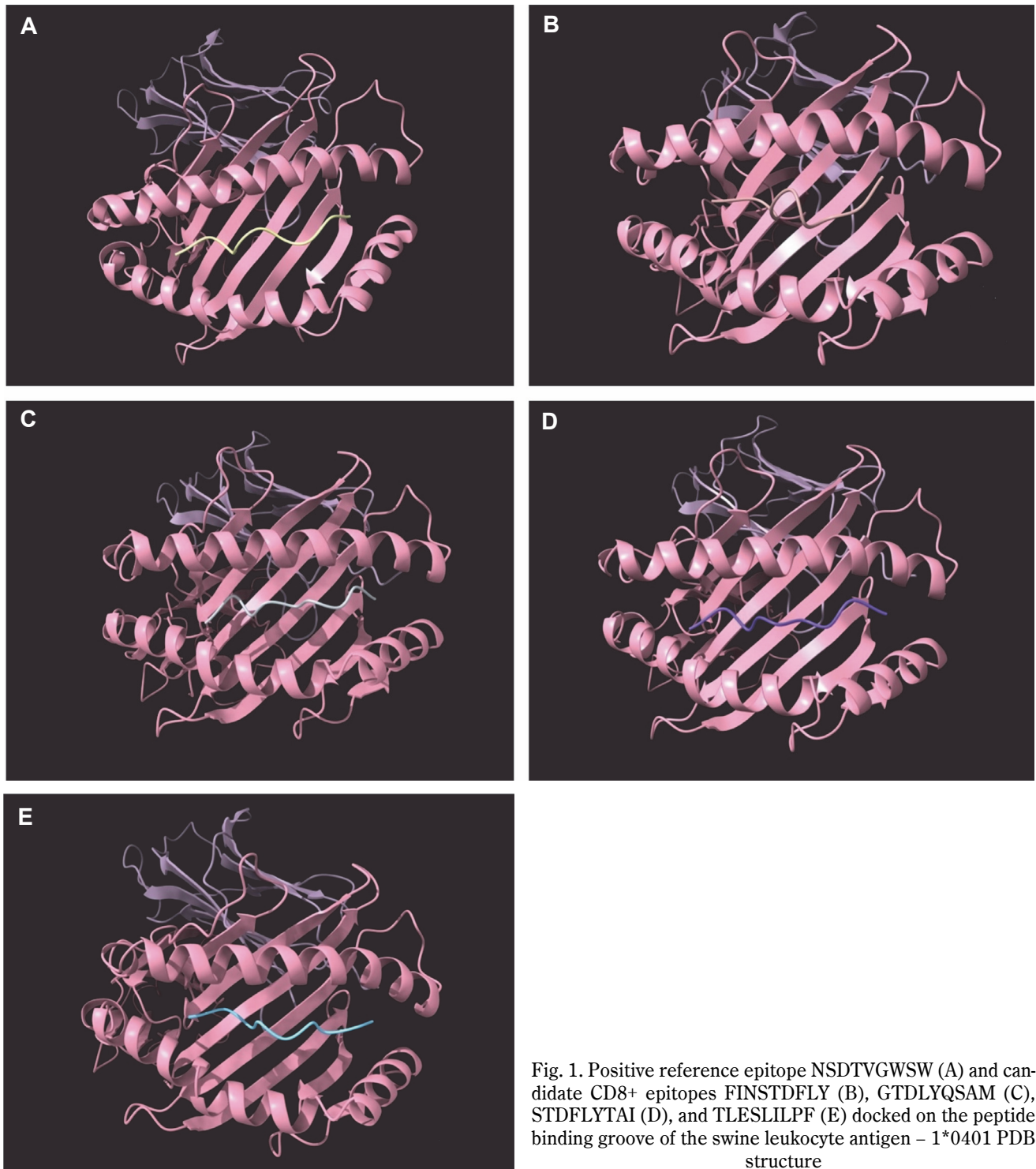
2016). Table 3 provides an overview of the protein–ligand interactions in the SLA-epitope complexes.

To assess favorable binding of the epitopes to SLA-1*0401 binding groove, in comparison to the positive reference, the dissociation constant (K_d) and binding energy (G_{bind}) of each SLA-epitope complex were estimated. Table 4 presents the calculated values for each SLA-epitope complex. Resulting values imply that the candidate epitopes exhibit comparable binding with the positive reference. Small K_d and G_{bind} values indicate favorable complex formation of each candidate epitope with SLA-1*0401.

MD simulation was performed to further assess stability and flexibility of epitope-SLA complex formed. The simulation resulted in average RMSD values 1.66, 2.03, 2.37 and 2.52 Å, for peptide-SLA complexes TLESLILPF, GTDLYQSAM, STDFLYTAI, and FINSTDFLY, respectively (Fig. 2). Moreover, average RMSF values for TLESLILPF, STDFLYTAI, FINSTDFLY, and GTDLYQSAM, are 1.88, 2.03, 2.27, and 2.69 Å, respectively (Fig. 3).

Discussion

The lack of commercially available vaccines as prophylactic and therapeutic agents against ASFV has motivated the researchers to identify potent peptides that can be incorporated in future vaccine formulations. Due to known high mutation rates of viral genomes, one of



the important strategies to consider in designing vaccines that are capable of inducing long-term immunity is the identification and epitope-mapping of highly conserved sequences. This approach can overcome the challenges associated with epitope immune evasion in rapidly mutating infectious diseases. With this essential

step in mind, this study chose to focus on CD8⁺ T-cell epitopes mapping using the highly conserved sequences of the ASFV pp62 polyprotein.

The sequence conservation analysis employed a Shannon variability threshold ≤ 0.1 . Sequences possessing a Shannon entropy of $H \leq 1.0$ is considered to be

Table 2. Highly conserved CD8+T-cell epitopes from ASFV polyprotein pp62

Epitope	Start	End	Allele	IC ₅₀ [nM]	Proteasome score	TAP score
AVLAPAADF	48	56	SLA-1*0401	456.7	1	1.16
DFDPLVTFY	17	25	SLA-1*0401	324.8	1.36	1.13
ESLILPFYI	5	13	SLA-2*0502	348.4	1	0.19
FINSTDFLY	25	33	SLA-1*0801	386.2	1.2	1.31
			SLA-2*0102	333.7	1.2	1.31
			SLA-2*1002	293.1	1.2	1.31
			SLA-1*1301	258.8	1.2	1.31
			SLA-1*0702	145.8	1.2	1.31
			SLA-1*0701	145.8	1.2	1.31
SLA-1*0401	73.4	1.2	1.31			
FPDRRLAVW	59	67	SLA-2*1101	382	1.42	0.07
FPESTINAL	5	13	SLA-2*1101	205.1	1.53	0.18
GEIGSYTGV	21	29	SLA-2*0601	471.8	1.06	0.03
GTDLYQSAM	53	61	SLA-1*0101	256	0.78	0.03
			SLA-1*0401	145.1	0.78	0.03
IADAATKEV	9	17	SLA-1*0101	464.2	0.99	0.03
LTDLVPTRL	55	63	SLA-1*0101	346.6	1.43	0.36
NELEQTNTI	60	68	SLA-2*0601	339.1	1.39	0.17
PVAKHISPY	23	31	SLA-1*0401	187.1	0.76	1.07
STDFLYTAI	28	36	SLA-1*0401	180.1	0.98	0.23
			SLA-2*0502	176.2	0.98	0.23
			SLA-1*0101	98.8	0.98	0.23
STINALRFW	8	16	SLA-1*0501	316.1	1.43	0.35
TEFIKVLPL	7	15	SLA-3*0501	369.8	1.2	0.33
			SLA-3*0502	369.8	1.2	0.33
			SLA-3*0503	369.8	1.2	0.33
			SLA-2*0601	217.6	1.2	0.33
TEILHHNAV	38	46	SLA-2*0601	439.7	1.04	0.15
TINALRFWK	9	17	SLA-2*0102	396.8	0.76	0.23
TLESLILPF	3	11	SLA-1*1301	312.7	1.11	1.08
			SLA-1*0401	202.9	1.11	1.08
TPAEFTGEI	15	23	SLA-2*1101	451.8	0.98	0.03
YETSEFAEI	41	49	SLA-2*0601	253.2	0.98	0.15
YLAKSVSLL	60	68	SLA-2*1002	469.2	1.32	0.36

highly conserved (Liu and Bahar, 2012). Findings of the study showed that mapped CD8+ epitopes occupy highly conserved positions in the pp62 polyprotein sequence (Crisuolo and Gribaldo, 2010) which imply that these

conserved epitopes can potentially induce long-term adaptive immunity in swine.

Cytotoxic T-cells play crucial roles in providing immunity against ASFV by eliminating virally infected cells.

Table 3. Protein–ligand interactions in SLA – epitope complexes

SLA – epitope complex	Number of hydrogen bonds	Donor–acceptor hydrogen distance range [Å]	Donor–acceptor distance range [Å]
SLA-1*0401-FINSTDFLY	6	1.929–2.036	2.805–2.984
SLA-1*0401-GTDLYQSAM	16	1.859–2.204	2.804–3.167
SLA-1*0401-STDFLYTAI	8	1.900–2.123	2.823–3.026
SLA-1*0401-TLESLILPF	9	1.856–2.066	2.805–3.431
SLA-1*0401-NSDTVGWSW (Reference)	None	N/A	2.618–3.213

Table 4. Binding energy and binding affinity of SLA 1*0401 – epitope complexes

Docked epitope	G_{bind} [kcal/mol]	K_d [mol/l]
FINSTDFLY	–8.5	1.1×10^{-6}
GTDLYQSAM	–10.8	2.8×10^{-8}
STDFLYTAI	–11.5	8.1×10^{-9}
TLESLILPF	–11.9	4.3×10^{-9}
NSDTVGWSW	–12.2	2.9×10^{-9}

CD8+ T-cells recognize antigenic peptides associated with MHC class I molecules on the surface of virus-infected target cells (Turgeon, 2014). This study successfully identified epitopes that can promiscuously bind to the most common SLA I alleles; thereby, providing broader coverage of the global swine population. MHC polymorphism can affect the interaction between epitopes and the MHC binding groove (Ebrahimi et al., 2019). Thus, the promiscuity of epitopes is another important factor to consider in designing epitope-based vaccines. The binding of these epitopes with the most dominant SLA I allele further supports wider swine population coverage.

A previous study had identified and analyzed highly conserved CD8+ T-cell epitopes of ASFV pp220, pp62, p72, p30, and CD2v. Similarly, researchers in this study docked epitopes that can bind to SLA-1*0401. However, the study was limited to using only the three most common SLAs (SLA-1*0101, SLA-1*0401, and SLA-1*0801) in determining candidate epitopes (Herrera and Bisa, 2021). In contrast, the current study included all SLAs available in the databases to comprehensively map highly conserved CD8+ pp62 epitopes. Moreover, this is the first study to identify and analyze a set of highly

conserved CD8+ T-cell epitopes in the ASFV polyprotein pp62 that are also promiscuous binders of 45 various SLA alleles. The researchers identified 21 candidate epitopes with MHC IC_{50} values ≤ 500 nM which indicate strong binding affinity between the candidate epitope and its corresponding SLA-binder (Fleri et al., 2017; Jensen et al., 2018).

Analysis of candidate epitopes generated other valuable data including proteasome scores which predict the efficiency of C-terminus generation, and the TAP scores which evaluate the ability of the peptide's N-terminal prolonged precursor. A more positive TAP score indicates high substrate affinity and MHC scores for the candidate epitopes (Abele and Tampe 1999; Fleri et al., 2017).

In vaccine design, safety profile is equally important as vaccine efficacy. To assess the safety profile of mapped CD8+ peptides, all mapped epitopes underwent filtering process in BLASTp against the collection of protein sequences from domestic pigs (*Sus scrofa domestica*). This selective step aimed to identify potential cross-reactive epitopes and prevent potential adverse reactions. The BLASTp analysis of the candidate epitopes resulted in E -values ranging from 2.2 to 75. Pepti-

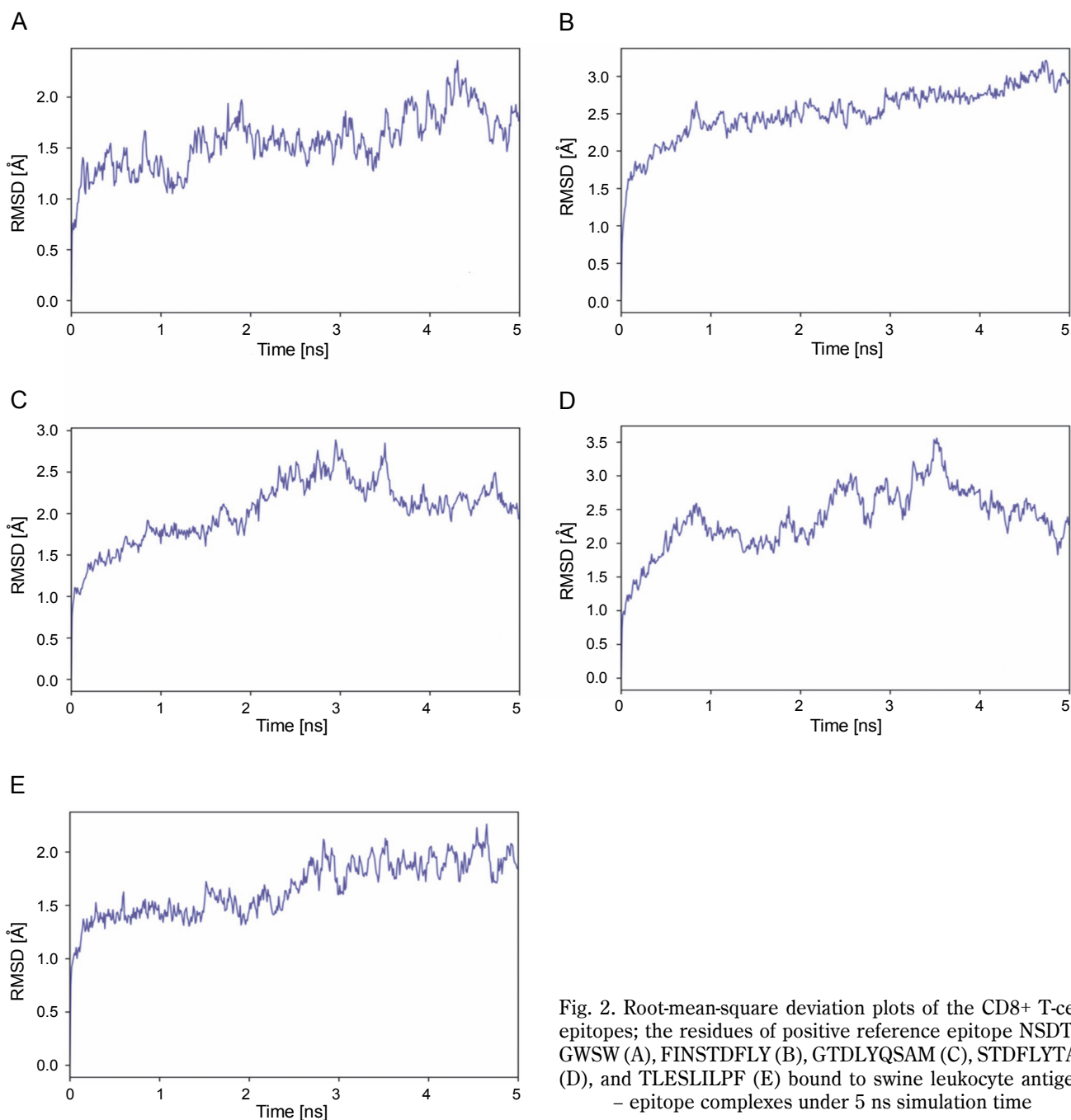


Fig. 2. Root-mean-square deviation plots of the CD8+ T-cell epitopes; the residues of positive reference epitope NSDTV GSW (A), FINSTDFLY (B), GTDLYQSAM (C), STDFLYTAI (D), and TLESLILPF (E) bound to swine leukocyte antigen – epitope complexes under 5 ns simulation time

des with E -values $< 1.0 \times 10^{-30}$ are considered cross-reactive (Hileman et al., 2002). Using this parameter, the results from this study suggest that all candidate ASFV CD8+ epitopes are less likely to exhibit cross-reactivity. This finding is a significant step as it helps mitigate the risk of autoimmune issues that could arise from administering the candidate epitopes in swine.

Furthermore, the analysis of peptide–SLA I interactions revealed that the presence of hydrogen bonds contribute favorably to stability of complex formation

(Pace et al., 2014). These hydrogen-bond pairings promote high-affinity receptor–ligand interactions by reducing competitive interference with water molecules (Chen et al., 2016). The strength of these bonds is characterized based on the donor–acceptor distances. Interactions with donor–acceptor distances ranging from 2.2 to 2.5 Å are considered “strong, mostly covalent”. Distance range of 2.5–3.2 Å is categorized as “moderate, mostly electrostatic”, while distance range 3.2–4.0 Å is regarded as “weak, electrostatic” interactions (Jeffrey, 1997).

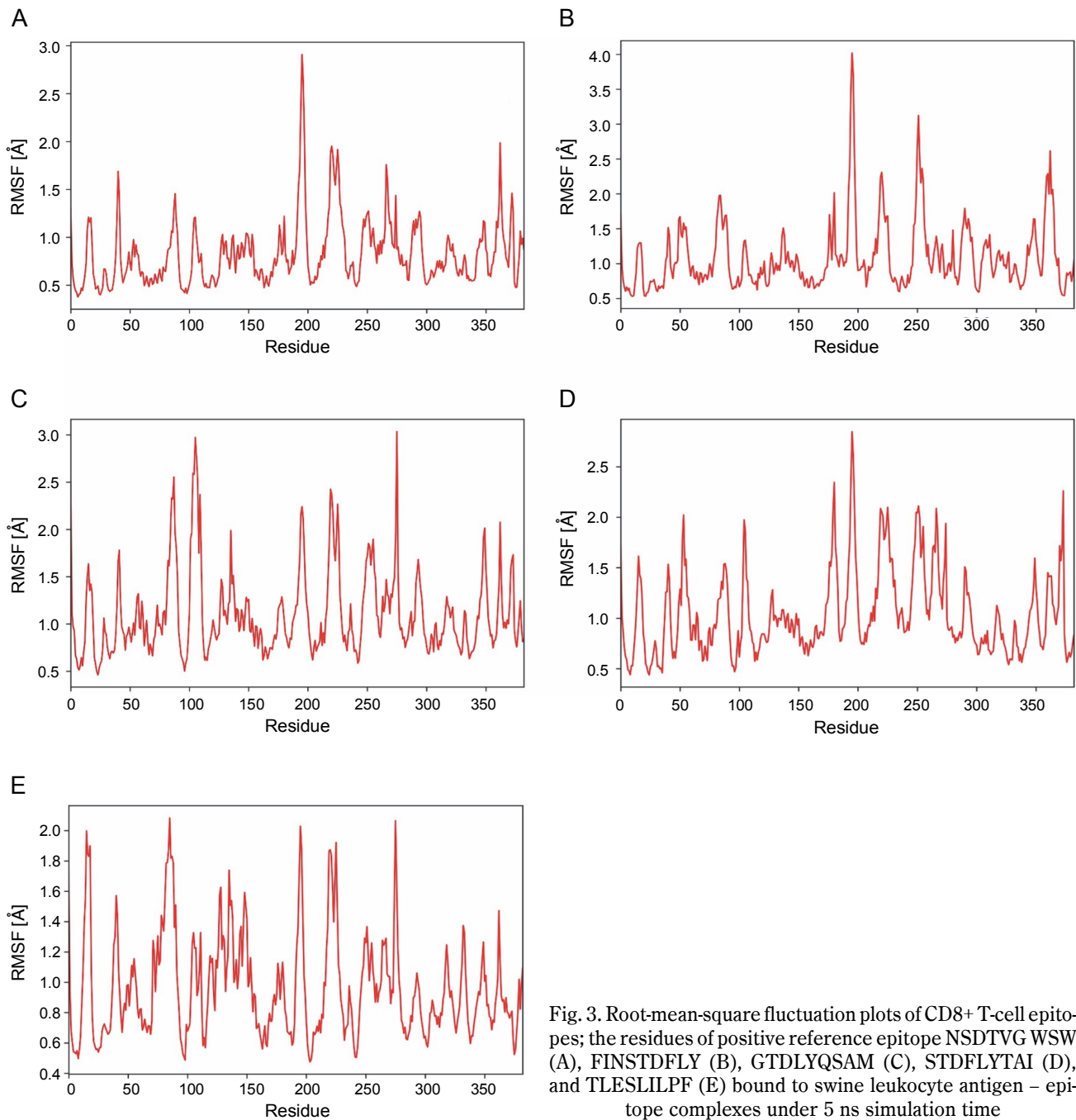


Fig. 3. Root-mean-square fluctuation plots of CD8⁺ T-cell epitopes; the residues of positive reference epitope NSDTVG WSW (A), FINSTDFLY (B), GTDLYQSAM (C), STDFLYTAI (D), and TLESLILPF (E) bound to swine leukocyte antigen – epitope complexes under 5 ns simulation time

The estimated dissociation constant values (K_d) for GTDLYQSAM, STDFLYTAI, and TLESLILPF are <100 nM which indicate favorable epitope binding. K_d values > 100 nM are classified as having larger dissociation rates (Motulsky and Neubig, 2010) where complex dissociation is more likely. Additionally, the binding energies (G_{bind}) for all the docked epitopes are negative which further underline the spontaneity of epitope-SLA I complex formation.

The stability and flexibility of docked immune complexes were assessed using RMSD and RMSF parameters. RMSD analysis was used to evaluate deviations during the simulation, with smaller deviations indicating higher stability of the protein complex (Aier et al., 2021). The RMSD value of the complex formation reflects the stability of the complex, as it is influenced by the binding of the ligands (Mawere and Kumar, 2014). Fluctuations within the range 1.0–3.0 Å are considered

acceptable and are expected to occur (Shawan et al., 2021). Results indicate that complex formation is stable based on the fluctuation values within a specific period. It is also important to note that the binding of ligands affects the RMSD of protein backbone (Mawere and Kumar, 2014). In terms of structural flexibility, the RMSF plots of the candidate epitopes indicate the fluctuations of the residues within the complex. An RMSF value with fluctuations above 4 Å in the amino acids of the binding site may suggest suboptimal binding, as the residues do not form strong interactions with the small molecule (de Vita et al., 2021). Upon evaluation of RMSD and RMSF values of each epitope–SLA complex analyzed in this study, results showed that the binding of candidate epitopes with their corresponding SLA are within the acceptable fluctuation range which suggest the formation of stable immune complexes. Taken together, estimated values of these parameters can further support the efficacy of identified candidate CD8⁺ epitopes in potentially inducing cytotoxic immune response against ASFV in swine.

Conclusion

This study successfully identified a set of highly conserved and promiscuous CD8⁺ epitopes in the ASFV polyprotein pp62 which can potentially induce immunogenicity upon incorporation in future ASF vaccine formulations. All candidate epitopes exhibited acceptable binding affinity and formed flexible yet stable epitope–SLA complexes, emphasizing their potential immunogenicity. More importantly, the low probability of cross-reactivity with the swine proteome highlights the safety profile of candidate peptides. Nevertheless, to establish the efficacy of these candidate epitopes, further evaluations through subsequent *in vitro* and *in vivo* studies are recommended.

Acknowledgments

This work was supported by the Department of Science and Technology-Science Education Institute (DOST-SEI).

References

- Abele R., Tampé R. (1999) *Function of the transport complex TAP in cellular immune recognition*. *Biochim. Biophys. Acta* 1461(2): 405–419. [https://doi.org/10.1016/S0005-2736\(99\)00171-6](https://doi.org/10.1016/S0005-2736(99)00171-6)
- Aier I., Varadwaj P.K., Raj U. (2016) *Structural insights into conformational stability of both wild-type and mutant EZH2 receptor*. *Sci. Rep.* 6: 34984. <https://doi.org/10.1038/srep34984>
- Andrés G., Charro D., Matamoros T., Dillard R.S., Abrescia N.G.A. (2020) *The cryo-EM structure of African swine fever virus unravels a unique architecture comprising two icosahedral protein capsids and two lipoprotein membranes*. *J. Biol. Chem.* 295(1): 1–12. <https://doi.org/10.1074/JBC.AC119.011196>
- Arantes P.R., Polêto M.D., Pedebos C., Ligabue-Braun R. (2021) *Making it rain: cloud-based molecular simulations for everyone*. *J. Chem. Inform. Model.* 61(10): 4852–4856. https://doi.org/10.1021/ACS.JCIM.1C00998/SUPPL_FILE/CI1C00998_SI_001.PDF
- Bhasin M., Raghava G. (2007) *A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes*. *J. Biosci.* 32: 31–42. <https://www.imtech.res.in/raghava/nhalpred/>
- Blome S., Franzke K., Beer M. (2020) *African swine fever – a review of current knowledge*. *Virus Res.* 287: 198099. <https://doi.org/10.1016/J.VIRUSRES.2020.198099>
- Chen D., Oezguen N., Urvil P., Ferguson C., Dann S.M., Savidge T.C. (2016) *Regulation of protein-ligand binding affinity by hydrogen bond pairing*. *Sci. Adv.* 2(3): e1501240. <https://doi.org/10.1126/sciadv.1501240>. PMID: 27051863; PMCID: PMC4820369
- Cooper T.L., Smith D., Gonzales M.J.C., Maghanay M.T., Sanderson S., Cornejo M.R.J.C., Pineda L.L., Sagun R.A.A., Salvacion O.P. (2022) *Beyond numbers: determining the socioeconomic and livelihood impacts of African swine fever and its control in the Philippines*. *Front. Vet. Sci.* 8: 734236. <https://doi.org/10.3389/fvets.2021.734236>
- Criscuolo A., Gribaldo S. (2010) *BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments*. *BMC Evol. Biol.* 10: 210. <https://doi.org/10.1186/1471-2148-10-210>
- Ebrahimi S., Mohabatkar H., Behbahani M. (2019) *Predicting promiscuous T cell epitopes for designing a vaccine against Streptococcus pyogenes*. *Appl. Biochem. Biotech.* 187(1): 90–100. <https://doi.org/10.1007/s12010-018-2804-5>
- Fleri W., Paul S., Dhanda S.K., Mahajan S., Xu X., Peters B., Sette A. (2017) *The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design*. *Front Immunol.* 8: 278. <https://doi.org/10.3389/FIMMU.2017.00278>
- Fu L., Niu B., Zhu Z., Wu S., Li W. (2012) *CD-HIT: accelerated for clustering the next-generation sequencing data*. *Bioinformatics* 28(23): 3150–3152. <https://doi.org/10.1093/BIOINFORMATICS/BTS565>
- Garcia-Boronat M., Diez-Rivero C.M., Reinherz E.L., Reche P.A. (2008) *PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery*. *Nucl. Acids Res.* 36: W35–41. <https://doi.org/10.1093/NAR/GKN211>
- Hernández-Santoyo A., Tenorio-Barajas A.Y., Victor Altuzar, Vivanco-Cid H., Mendoza-Barrera C. (2013) *Protein-protein and protein-ligand docking*. [in:] *Protein engineering*

- *technology and application*. Ed. T. Ogawa. InTech: 187. <https://doi.org/10.5772/56376>. [accessed 2023 Jan 22]. <https://www.intechopen.com/state.item.id>
- Herrera L.R.D.M., Bisa E.P. (2021) *In silico analysis of highly conserved cytotoxic T-cell epitopes in the structural proteins of African swine fever virus*. *Vet. World*. 14(10): 2625–2633. <https://doi.org/10.14202/vetworld.2021.2625-2633>
- Hileman R.E., Silvanovich A., Goodman R.E., Rice E.A., Holleschak G., Astwood J.D., Hefle S.L. (2002) *Bioinformatic methods for allergenicity assessment using a comprehensive allergen database*. *Int. Arch. Allergy Immunol.* 128(4): 280–291. <https://doi.org/10.1159/000063861>
- Jeffrey G.A. (1997) *An introduction to hydrogen bonding*. Oxford University Press.
- Jensen K.K., Andreatta M., Marcatili P., Buus S., Greenbaum J.A., Yan Z., Sette A., Peters B., Nielsen M. (2018) *Improved methods for predicting peptide binding affinity to MHC class II molecules*. *Immunology* 154(3): 394–406. <https://doi.org/10.1111/IMM.12889>
- Jia N., Ou Y., Pejsak Z., Zhang Y., Zhang J. (2017) *Roles of African swine fever virus structural proteins in viral infection*. *J. Vet. Res.* 61(2): 135–143. <https://doi.org/10.1515/jvetres-2017-0017>
- Karosiene E., Lundegaard C., Lund O., Nielsen M. (2012) *Net-MHCcons: a consensus method for the major histocompatibility complex class I predictions*. *Immunogenetics* 64(3): 177–186. <https://doi.org/10.1007/s00251-011-0579-8>
- Kedkovid R., Sirisereewan C., Thanawongnuwech R. (2020) *Major swine viral diseases: an Asian perspective after the African swine fever introduction*. *Porcine Health Manag.* 6(1): 20. <https://doi.org/10.1186/s40813-020-00159-x>
- Liu Y., Bahar I. (2012) *Sequence evolution correlates with structural dynamics*. *Mol. Biol. Evol.* 29(9): 2253–2263. <https://doi.org/10.1093/molbev/mss097>
- Mawere C., Kumar S. (2014) *Molecular dynamics simulation studies of HYD1*. *Inter. J. Sci. Res.* 3(8): 678–784. www.ijsr.net
- Motulsky H.J., Neubig R.R. (2010) *Analyzing binding data*. *Curr. Protoc. Neurosci.* chapt. 7: unit 7.5. <https://doi.org/10.1002/0471142301.ns0705s52>
- Oli A.N., Obialor W.O., Ifeanyichukwu M.O., Odimegwu D.C., Okoyeh J.N., Emechebe G.O., Adejumo S.A., Ibeanu G.C. (2020) *Immunoinformatics and vaccine development: an overview*. *ImmunoTarg. Ther.* 9: 13. <https://doi.org/10.2147/ITT.S241064>
- Pace C.N., Fu H., Lee Fryar K., Landua J., Trevino S.R., Schell D., Thurlkill R.L., Imura S., Scholtz J.M., Gajiwala K., Sevcik J., Urbanikova L., Myers J.K., Takano K., Hebert E.J., Shirley B.A., Grimsley G.R. (2014) *Contribution of hydrogen bonds to protein stability*. *Protein Sci.* (5): 652–661. <https://doi.org/10.1002/pro.2449>
- Shawan M.M.A.K., Halder S.K., Hasan M.dA. (2021) *Luteolin and abyssinone II as potential inhibitors of SARS-CoV-2: an in silico molecular modeling approach in battling the COVID-19 outbreak*. *Bull. Natl. Res. Cent.* 45(1): 27. <https://doi.org/10.1186/s42269-020-00479-6>
- Suárez C., Salas M.L., Rodríguez J.M. (2010) *African swine fever virus polyprotein pp62 is essential for viral core development*. *J. Virol.* 84(1): 176–187. <https://doi.org/10.1128/JVI.01858-09>
- Tian C., Kasavajhala K., Belfon K.A.A., Raguette L., Huang H., Miguez A.N., Bickel J., Wang Y., Pincay J., Wu Q., et al. (2020) *Ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution*. *J. Chem. Theory Comput.* 16(1): 528–552. <https://doi.org/10.1021/acs.jctc.9b00591>
- Turgeon M.L. (2014) *Immunology & serology in laboratory medicine*. <https://evolve.elsevier.com/Turgeon/immunology/>
- Urbano A., Forth J.H., Olesen A.S., Dixon L., Rasmussen T.B., Cackett G., Werner F., Karger A., Andrés G., Wang X. et al. (2021) *African swine fever virus: cellular and molecular aspects*. [in:] *Understanding and combatting African Swine Fever*. Wageningen Academic Publishers: 25–61.
- World Organisation for Animal Health (2022) *African swine fever (ASF) – situation report 24*. [accessed 2023 Jan 5]. <https://www.woah.org/app/uploads/2022/12/asf-report24.pdf>
- de Vita S., Chini M.G., Bifulco G., Lauro G. (2021) *Insights into the ligand binding to bromodomain containing protein 9 (BRD9): A guide to the selection of potential binders by computational methods*. *Molecules* 26(23): 7192. <https://doi.org/10.3390/molecules26237192>
- Zhong K., Zhu M., Yuan Q., Deng Z., Feng S., Liu D., Yuan X. (2022) *Development of an indirect ELISA to detect African swine fever virus pp62 protein-specific antibodies*. *Front. Vet. Sci.* 8: 798559. <https://doi.org/10.3389/fvets.2021.798559>