immuno

MDPI

Article

Integration of Physicochemical Profiling and HLA Class II Binding for the Identification of Conserved Epitopes in the Glycoprotein of Lyssaviruses from Phylogroups I and II

André Miller C. Lima [1], Taciana Fernandes S. B. Coelho [2] and Carlos Alberto M. Carvalho [1,*]

[1] Laboratory of Molecular Biology, Center for Biological and Health Sciences, University of Pará State, Belém 66095-662, PA, Brazil; andre.lima@aluno.uepa.br
[2] Section for Arbovirology and Hemorrhagic Fevers, Laboratory of Rabies, Evandro Chagas Institute, Ananindeua 67030-000, PA, Brazil; tacianabarbosa@iec.gov.br
* Correspondence: carlos.carvalho@uepa.br

Abstract

Lyssaviruses are neurotropic viruses that cause fatal encephalitis, with the rabies virus as the most prominent member. The viral glycoprotein (G) plays a key role in infection and is the main target of adaptive immune responses. This study aimed to comparatively analyze linear B- and $CD4^+$ T-cell epitopes in the G protein ectodomain of lyssaviruses from phylogroups I (RABV, EBLV-1, EBLV-2, DUVV, and ABLV) and II (LBV and MOKV) using bioinformatics tools. Protein sequences were obtained from GenBank, processed to isolate ectodomains, aligned for identity analysis, and used to generate consensus sequences. $CD4^+$ T-cell epitopes were predicted based on HLA-II binding affinity, while linear B-cell epitopes were identified using physicochemical properties and assessed for N-glycan masking. Amino acid identity ranged from 76.71% to 83.79% in phylogroup I and 82.72% in phylogroup II. Phylogroup I showed a higher density of HLA-II epitopes (0.22) than phylogroup II (0.18). Despite differences in antigenicity distribution, conserved linear B-cell epitopes in both phylogroups overlapped with peptides binding to HLA-II DRB1*15:01 and were not masked by N-glycans. These findings highlight putatively conserved antigenic regions identified through computational analysis and may support future studies focused on the development of improved vaccines and immunoprophylactic strategies against lyssaviruses.

Keywords: lyssavirus; viral proteins; epitopes; computational biology; immunity

Check for updates

Academic Editor: Sherif T.S. Hassan

Received: 23 April 2026
Revised: 21 May 2026
Accepted: 26 May 2026
Published: 29 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The genus *Lyssavirus* comprises a heterogeneous group of negative-sense single-stranded ribonucleic acid (ssRNA–) viruses within the family *Rhabdoviridae*, recognized for their ability to cause fatal encephalitis. Historically, rabies is one of the oldest recorded diseases, described as early as the 5th century BCE, with observations of abnormal behavior in animals following bites and early suggestions of zoonotic transmission [1]. Despite substantial advances in virological research, critical gaps remain in understanding host invasion mechanisms and viral adaptive strategies that enable persistence in mammalian hosts. Transmission occurs predominantly through bites from infected animals, although alternative routes—such as mucosal exposure, organ transplantation, and, rarely, aerosolization—have been reported. Disease progression is influenced by factors including inoculation site, viral load, and host immune status [2].

Within the genus, phylogroup I includes the classic Rabies virus (RABV) and related viruses such as European bat lyssavirus type 1 (EBLV-1), European bat lyssavirus type 2 (EBLV-2), Duvenhage virus (DUVV), and Australian bat lyssavirus (ABLV). RABV, the prototypical species, is responsible for approximately 59,000 human deaths annually and exhibits marked neurotropism, infecting a wide range of terrestrial mammals. Phylogroup II comprises members of zoonotic relevance such as Lagos bat virus (LBV) and Mokola virus (MOKV). LBV demonstrates remarkable genetic stability across temporal and geographic scales, suggesting long-term host adaptation, while MOKV remains poorly understood, with an unidentified reservoir and evidence suggesting preferential circulation among small terrestrial mammals [3].

Structurally, lyssaviruses display a bullet-shaped virion containing a helical nucleocapsid composed of the genomic RNA complexed with N, P, and L proteins, surrounded by a matrix layer and a lipid envelope embedded with the glycoprotein (G) [4]. Their replication cycle is initiated by the interaction of the trimeric G protein with host receptors such as nAChR, NCAM, and p75NTR, followed by endocytosis and pH-dependent membrane fusion. Viral transcription and replication occur in the cytoplasm, regulated by N protein availability, leading to the production of genomic RNA and viral proteins necessary for assembly. Newly formed virions are released via budding without inducing host cell lysis [5].

Immunologically, protective responses are primarily mediated by neutralizing antibodies targeting G protein; however, viral neurotropism and limited inflammation within the central nervous system contribute to immune evasion [6]. Although phylogroup I viruses share antigenic similarities, variations exist in pathogenicity, epidemiology, and host range, particularly among bat-associated lyssaviruses. On the other hand, phylogroup II viruses exhibit reduced susceptibility to neutralization by antibodies elicited by current rabies vaccines, reflecting significant antigenic divergence, particularly within G protein. These differences highlight important evolutionary and immunological distinctions between phylogroups, with implications for vaccine efficacy and cross-protection [7].

In this context, the human leukocyte antigen (HLA) system plays a central role in antigen presentation and adaptive immunity. Particularly, HLA class II molecules mediate the presentation of extracellular antigens to $CD4^+$ T cells, influencing the magnitude and specificity of B-cell responses [8]. Advances in immunoinformatics have enabled in silico prediction of B- and T-cell epitopes, integrating physicochemical properties and peptide–HLA binding affinity. Machine learning-based tools have significantly improved predictive accuracy, supporting the rational design of epitope-based vaccines [9]. In this context, the present study aimed to comparatively analyze the occurrence of linear B- and $CD4^+$ T-cell epitopes in the ectodomain of the G protein of lyssaviruses belonging to phylogroups I and II.

2. Materials and Methods

2.1. Retrieval of Amino Acid Sequences and Isolation of Ectodomain

The complete amino acid sequence of the G protein from lyssaviruses belonging to phylogroups I and II was obtained by conceptual translation in FASTA format from the reference genome sequence of each species, available in the GenBank database [10], under the following accession numbers: RABV (NP_056796.1), EBLV-1 (YP_001285391.1), EBLV-2 (YP_001285396.1), DUVV (YP_007641405.1), ABLV (NP_478342.1), LBV (YP_007641390.1), and MOKV (YP_142353.1). This activity was registered in the National System for the Management of Genetic Heritage and Associated Traditional Knowledge (SisGen) under number A83F17B, in compliance with Brazilian Law 13,123/2015 and its regulations.

Subsequently, the SignalP 6.0 [11] and TMHMM 2.0 [12] servers were used to identify residues corresponding to signal peptides and transmembrane helices, respectively, in the G protein of each member of phylogroups I and II. This approach enabled the isolation of only the amino acid sequence corresponding to the ectodomain for further analysis.

2.2. Determination of Percent Identity Matrix and Generation of Consensus Sequences

To construct the percent identity matrix, the amino acid sequences of the G protein ectodomain were subjected to multiple sequence alignment using Clustal Omega 1.2.4 [13], with the input order preserved and the output format set to include character counts. Subsequently, the aligned sequences were submitted to EMBOSS Cons 6.6.0 [14] to generate two consensus sequences, one for each lyssavirus phylogroup. The use of consensus sequences in the present study was intended as an exploratory strategy to identify antigenic patterns potentially shared among lyssaviruses within each phylogroup.

2.3. Prediction of Linear B- and CD4$^+$ T-Cell Epitopes

Consensus sequences were submitted to NetMHCIIpan 4.1 [15] for the prediction of binding affinity of peptides derived from the G protein ectodomain to HLA class II molecules. A peptide length of 15 residues was used, selecting the alleles DRB1*03:01, DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01, and DRB5*01:01, which are commonly employed in immunoinformatics studies as representative HLA class II molecules with broad relevance to antigen presentation across diverse human populations [16]. The resulting data were then processed using the Epitope-Evaluator server [17], considering a percentile rank threshold of 5%, as well as the positional distribution of each peptide along the primary structure of the G protein ectodomain.

For B-cell epitope prediction, hydrophilicity [18], accessibility [19], and flexibility [20] profiles of residues along the G protein ectodomain were evaluated using NPS@ PCProf [21], with a sliding window size of 7 amino acids. To enable direct comparison among parameters with different scales, all values were normalized to the 0–1 range. A composite score was then calculated as the simple arithmetic mean of the normalized values for each residue, assigning equal weight to hydrophilicity, accessibility, and flexibility. The three residues located at each N- and C-terminal region were excluded from the analysis to avoid edge effects inherent to sliding window calculations, which can produce artificially low or undefined values at sequence boundaries. Regions with higher antigenic potential were identified by applying a threshold corresponding to the top 50% of composite scores and considering a minimum length of 5 contiguous residues.

Finally, the identified epitopes were evaluated for potential masking by N- or O-glycosylation using the NetNGlyc 1.0 [22] and NetOGlyc 4.0 [23] servers, which employs artificial neural networks to detect N-X-S/T sequons and O-GalNAc (mucin-type) glycosylation sites, respectively, adopting a threshold of 0.5 for post-translational modification potential.

3. Results

3.1. Percent Identity Matrix and Consensus Sequences

Analysis of the percent identity matrix of the G protein ectodomain revealed moderate to high levels of conservation among the evaluated sequences from RABV, EBLV-1, EBLV-2, DUVV, ABLV, LBV, and MOKV. Identity values ranged from 61.24% (between DUVV and MOKV) to 83.79% (between EBLV-1 and EBLV-2), indicating that all sequences share significant homology, consistent with their classification within the *Lyssavirus* genus (Figure 1). Multiple sequence alignment of the G protein ectodomain for each species is shown in Figure S1.

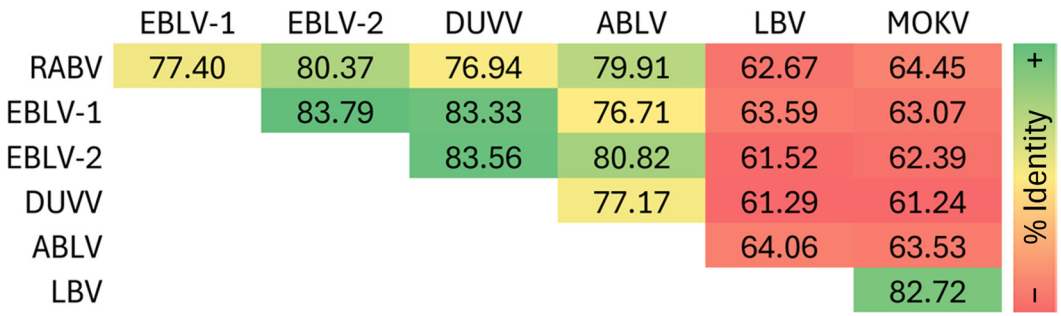


Figure 1. Percent identity matrix of the G protein ectodomain from lyssaviruses of phylogroups I and II.

Within phylogroup I, RABV showed the lowest identity when compared to DUVV (76.94%) and the highest identity when compared to EBLV-2 (80.37%). In phylogroup II, the matrix analysis revealed a moderate level of conservation between LBV and MOKV, with 82.72% identity. The consensus sequences derived from the multiple sequence alignment of the G protein ectodomain for each phylogroup are shown in Figure 2, comprising 439 residues for phylogroup I and 436 residues for phylogroup II.

```
>PG-I
KFPIYTIPDKLGPWSPIDIHHLSCPNNLVVEDEGCTTLTxFSYMELKVGYITTIKVNGFT
CTGVVTEAETYTNFVGYVTTTFKRKHFRPTPSACRDAYNWKIAGDPRYEESLHNPYPDSH
WLRTVKTTKESLLIISPSVADMDAYDKSLYSKMFPNGKCSGVSxxSxFCSTNHDYTIWMP
ENPNPGMSCDIFTxSKGKKATKDGKLCGFVDERGLYKSLKGACKLKLCGISGLRLMDGSW
VSLQxxNExKWCSPDQLVNIHDFHSDEIEHLVVEELVKKREECLDALESIITTKSISFRR
LSHLRKLVPGFGKAYTIINKTLMEADAHYKSVREWxEIIPSKGCLKAGGRCHPHYNGVFF
NGIILGPDGEVLIPEMQSALLQQHIELLESSVIPLMHPLADPSTVFKxDDEAEDFVEVHL
PDVQKQISGIDLGLPEWKx

>PG-II
DFPLYTIPERLnpWTPIDLIHLSCPNNLLSDaEGCSetSSFTYvELKTGFLAHQKVPGFT
CTGVINEAvTYTNFVGYVTTTFKRKHFKPTVSACRDAFNWKVSGDPRYEESLHTPYPDNS
WLRTVTTTKESLLIISPSIVEMDVYGRTLHSPMFPTGICSkfYPSIPSCaTNHDYTLWLP
EDPNLSLICDIFvTStGRKAMNGSRMCGFtDERGFYRTLKGACKLTLCGKPGLRLYDGTW
VSFTRPEIHVWCSPDQLVNVHNNRIDEIEHLIVDDLIRKREECLDTLETVLMSKSLSFRR
LSHFRKLVPGYGKAYTILNGSLMETNVHYKRVDnWvDILPSKGCLKVnnKCMEsdtGVfF
NGIIKGPDGRILIPEMQSglLKQHMDLLKAAVFPLRHPLIdKaSVFKKDGDADDFVDVHM
PDIqKlVSDVDLGLph
```

Figure 2. Consensus sequences of the G protein ectodomain from lyssaviruses of phylogroup I (PG-I) and phylogroup II (PG-II) in FASTA format.

3.2. Prediction of CD4⁺ T-Cell Epitopes

Prediction of HLA class II epitopes along the primary structure of the G protein ectodomain identified 97 peptides for lyssaviruses of phylogroup I and 78 peptides for lyssaviruses of phylogroup II, corresponding to densities of 0.22 and 0.18, respectively (left panel of Figure 3). For both phylogroups, the allele DRB3*02:02 exhibited the lowest epitope density—0.03 (14 peptides) in phylogroup I and 0.02 (9 peptides) in phylogroup II—whereas the allele DRB5*01:01 showed the highest epitope density—0.07 (32 peptides) in phylogroup I and 0.08 (33 peptides) in phylogroup II (right panel of Figure 3).

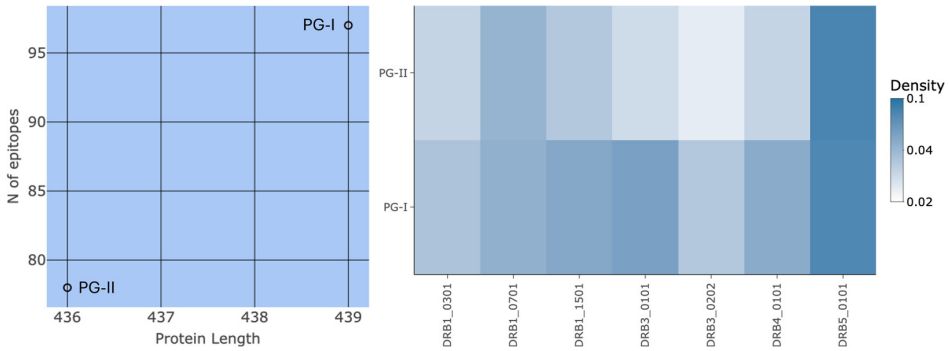


Figure 3. Epitope density for HLA class II in the G protein ectodomain of lyssaviruses from phylogroup I (PG-I) and phylogroup II (PG-II).

For lyssaviruses of phylogroup I, two overlapping peptides with binding affinity to four HLA class II alleles were identified. The first, located at positions 96–110, corresponded to the amino acid sequence DAYNWKIAGDPRYEE, while the second, located at positions 97–111, corresponded to AYNWKIAGDPRYEES. Both peptides exhibited affinity for the same alleles: DRB1*03:01, DRB3*01:01, DRB3*02:02, and DRB5*01:01 (red boxes in the upper panel of Figure 4). For phylogroup II lyssaviruses, two overlapping peptides with binding affinity to six HLA class II alleles were identified. The first, located at positions 253–267, corresponded to the sequence SPDQLVNVHNNRIDE, whereas the second, located at positions 254–268, corresponded to PDQLVNVHNNRIDEI. Both peptides also exhibited affinity for the same alleles: DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01, and DRB5*01:01 (red boxes in the lower panel of Figure 4).

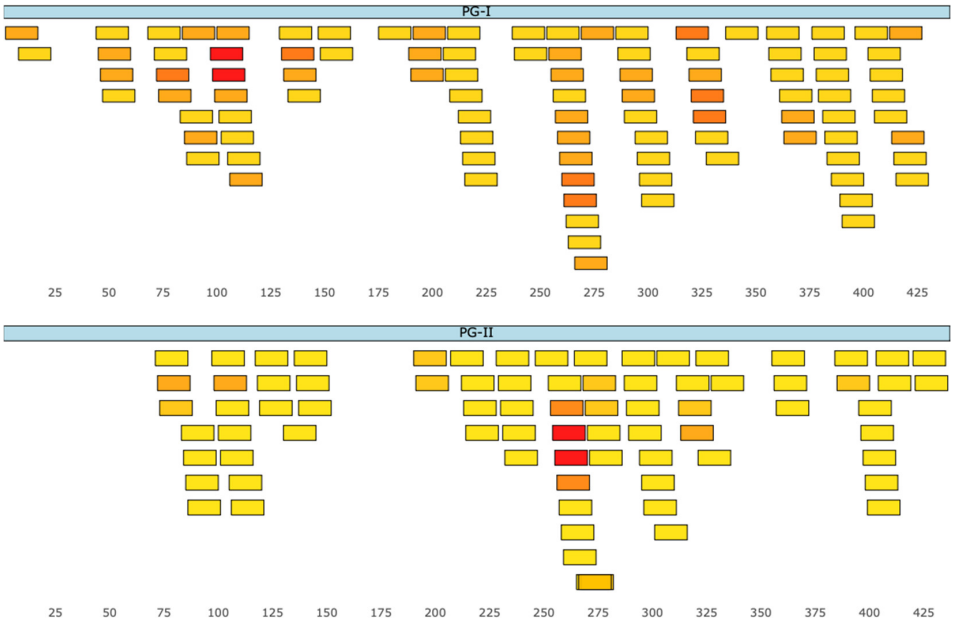


Figure 4. Prediction of $CD4^+$ T-cell epitopes for HLA class II along the primary structure (blue) of the G protein ectodomain of lyssaviruses from phylogroup I (PG-I) and phylogroup II (PG-II), colored according to allelic promiscuity in a gradient from yellow (lowest) to red (highest).

Analysis of epitope conservation between lyssavirus phylogroups revealed that, although most identified peptides were exclusive to phylogroups I (94 peptides) or II (75 peptides), three overlapping peptides were shared between both phylogroups: TYTNFVGYVTTTFKR (positions 70–84; allele DRB1*15:01), YTNFVGYVTTTFKRK (positions 71–85; alleles DRB1*07:01, DRB1*15:01, and DRB5*01:01), and TNFVGYVTTTFKRKH (positions 72–86; alleles DRB1*15:01 and DRB5*01:01) (Figure 5).

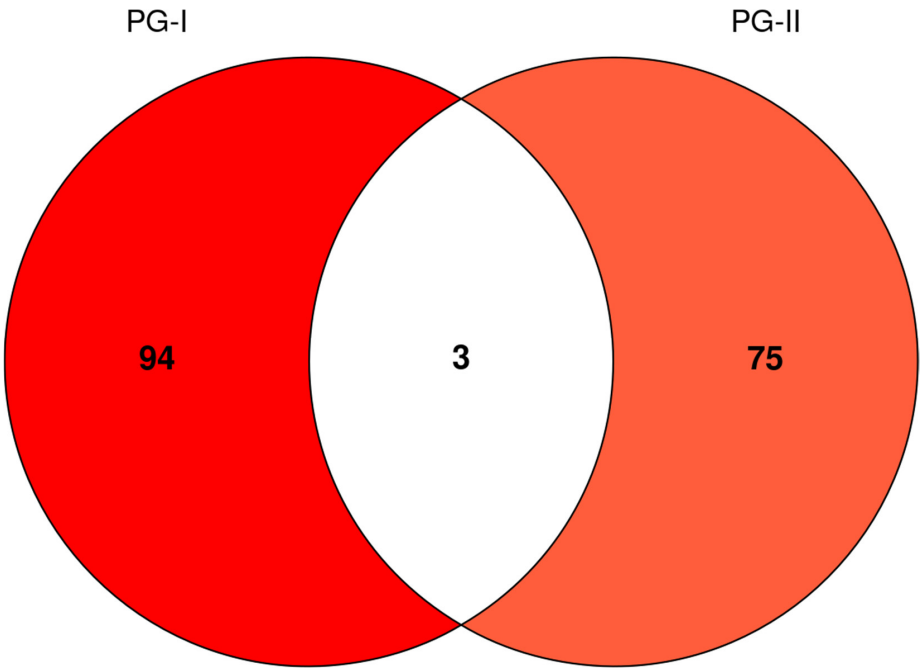


Figure 5. Conservation of HLA class II epitopes in the G protein ectodomain of lyssaviruses from phylogroup I (PG-I) and phylogroup II (PG-II).

3.3. Prediction of Linear B-Cell Epitopes

In phylogroup I lyssaviruses, the average antigenicity was higher in the N-terminal half of the G protein ectodomain, reaching the threshold at positions 82–87 (FKRKHF), 125–130 (VKTTKE), 195–205 (SKGKKATKDGK), 217–221 (KSLKG), 276–282 (LVKKREE), and 408–412 (XDDEA). In contrast, in phylogroup II lyssaviruses, the average antigenicity was higher in the C-terminal half of the G protein ectodomain, reaching the threshold at positions 81–88 (TFKRKHFK), 106–110 (PRYEE), 126–130 (TTTKE), 196–202 (TGRKAMN), 276–283 (LIRKREEC), and 404–412 (SVFKKDGDA) (Figure 6).

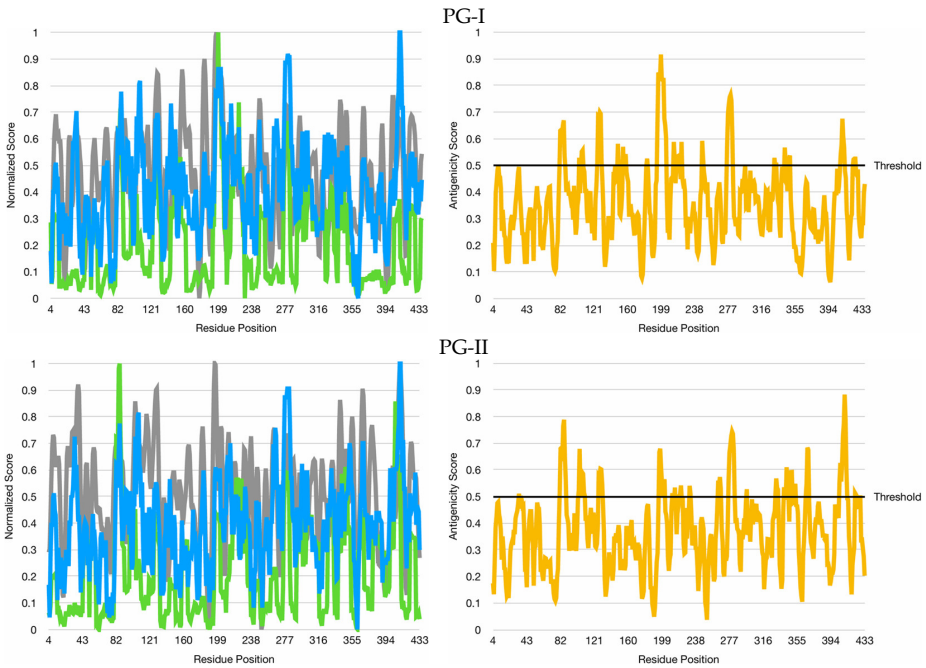


Figure 6. Antigenicity (yellow) in the G protein ectodomain of lyssaviruses from phylogroup I (PG-I) and phylogroup II (PG-II), calculated based on hydrophilicity (blue), accessibility (green), and flexibility (gray).

Regarding the N-glycosylation potential in the G protein ectodomain, one sequon was identified in phylogroup I lyssaviruses—at position 319 (NKTL)—and three sequons were identified in phylogroup II lyssaviruses—at positions 184 (NLSL), 202 (NGSR), and 319 (NGSL). Although mucin-type O-glycosylation sites were predicted in phylogroup I lyssaviruses at positions 295 and 403 (both serine residues), no such post-translational modification was predicted in phylogroup II lyssaviruses (Figure 7).

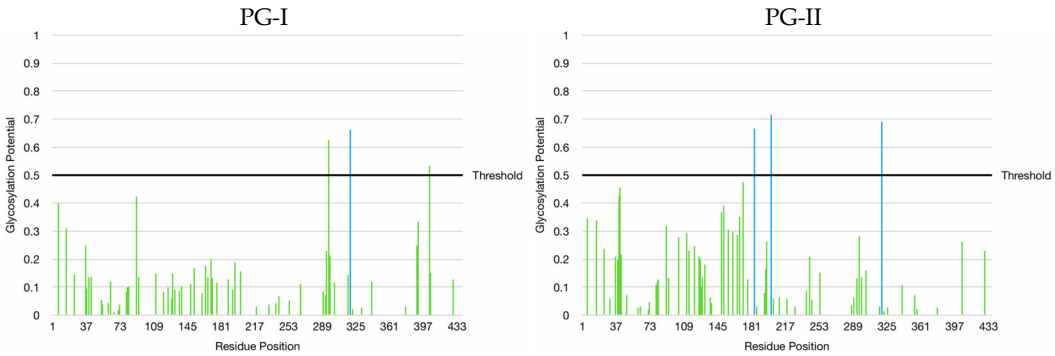


Figure 7. N-glycosylation (blue) and O-glycosylation (green) potentials in the G protein ectodomain of lyssaviruses from phylogroup I (PG-I) and phylogroup II (PG-II).

4. Discussion

The in silico prediction of HLA class II-restricted epitopes in the G protein ectodomain of lyssaviruses from phylogroups I and II revealed that the peptides with the highest allelic promiscuity are located at distinct positions within the molecule (96–110/97–111 and 253–267/254–268, respectively). Nevertheless, these peptides shared binding affinity to multiple HLA class II alleles, including DRB3*01:01, DRB3*02:02, and DRB5*01:01, as evidenced by the prediction algorithms employed. Additionally, the epitope density for HLA class II was slightly higher in phylogroup I (0.22) than in phylogroup II (0.18), suggesting that the G protein of the former constitutes a relatively more immunogenic target for $CD4^+$ T cell-mediated responses.

The allele set employed in the present study was selected according to a previously proposed empirical optimization strategy, which demonstrated that the combination of three DRB1 alleles with four DRB3/4/5 alleles provided the best predictive performance for identifying human HLA class II responses while covering the major HLA class II supertypes [16]. Importantly, the same study demonstrated that tailoring allele selection to specific ethnic cohorts resulted in only minor improvements in prediction performance, suggesting that this optimized allele set provides broadly representative HLA class II coverage despite population-level variability in allele frequencies. Nevertheless, additional analyses incorporating expanded HLA repertoires and formal population coverage assessments would further improve the translational applicability of the predicted epitopes across genetically diverse human populations exposed to lyssaviruses.

The identification of overlapping epitopes with binding affinity to a diverse set of HLA class II alleles is particularly relevant, as protein regions containing such epitopes tend to promote more efficient antigen presentation, increasing the likelihood of $CD4^+$ T cell activation and, consequently, enhancing both humoral and cell–mediated adaptive immune responses [24]. Moreover, the alleles identified are frequently reported in the literature as central players in the presentation of epitopes derived from medically relevant viruses, suggesting that lyssaviruses from phylogroups I and II share immunological determinants recognizable through classical pathways of the human adaptive immune system [25].

The prediction of candidate linear B-cell epitopes in the G protein ectodomain of lyssaviruses from phylogroups I and II was based on the combined analysis of hydrophilic-

ity, accessibility, and flexibility as proxies for antigenicity. Hydrophilic regions are more likely to be exposed on the protein surface and thus accessible to antibody recognition, while accessibility itself is a fundamental structural requirement for B-cell recognition, as only exposed residues can participate in antigen–antibody interactions. Furthermore, linear epitopes are often located in structurally permissive regions with higher conformational mobility, and such flexibility contributes to the epitope's ability to adapt to antibody binding [26]. Although the average antigenicity was higher in different portions of the G protein ectodomain in each phylogroup (i.e., the N-terminal half in phylogroup I and the C-terminal half in phylogroup II), only one of the six linear B-cell epitopes exceeding the threshold in each phylogroup did not exhibit sequence overlap (positions 217–221 in phylogroup I and 106–110 in phylogroup II). This finding suggests a differential propensity between the two phylogroups to initiate humoral adaptive immune responses.

Among the antigenic sites previously described for the RABV G protein, only antigenic site 1 (AS1) corresponds to a linear epitope—the remaining sites (AS2, AS3, AS4, and ASa) are conformational and discontinuous epitopes [27]. AS1 was characterized using a monoclonal antibody (CR57) capable of binding the residues KLCGVL (positions 226–231) [28]. In the present study, this region corresponds to KLCGIS in phylogroup I and TLCGKP in phylogroup II, but it did not exceed the antigenicity threshold established in our analysis. However, positions 195–205 (SKGKKATKDGK) in phylogroup I and 196–202 (TGRKAMN) in phylogroup II, which surpassed the threshold, overlap with positions 198–200 (KRA) of antigenic site 2 (AS2), which also includes positions 34–42 (GCTNLSGFS) [29]. Despite this overlap, complete sequence conservation was not observed for either site. Notably, AS1 and AS2 are both located within the pleckstrin homology domain (PHD) of RABV G protein [27], a complete β-barrel structure composed of two antiparallel β-sheets that connect on one side to the fusion domain (FD) and on the other side to the central domain (CD) [30].

Located within the FD of RABV G protein [31], the overlapping linear B-cell epitopes at positions 82–87 (FKRKHF) in phylogroup I and 81–88 (TFKRKHFK) in phylogroup II also share at least three residues (FKR) with the peptides TYTNFVGYVTTTFKR (positions 70–84), YTNFVGYVTTTFKRK (positions 71–85), and TNFVGYVTTTFKRKH (positions 72–86), which were commonly predicted to bind the HLA class II allele DRB1*15:01. This finding suggests potential functional relevance in $CD4^+$ T cell-dependent antibody responses, since partial overlap between predicted linear B-cell epitopes and HLA class II–binding peptides may favor linked recognition mechanisms involved in affinity maturation and durable humoral immunity [32,33]. The emphasis on DRB1*15:01 derives from its association with overlapping conserved peptides identified between phylogroups rather than from a specific epidemiological association with populations at risk for lyssavirus exposure. Additionally, DRB1*15:01 is frequently included in immunoinformatics analyses because of its documented high degree of repertoire sharing across HLA class II supertypes [34].

Furthermore, the above epitopes were not masked by the presence of predicted N- or O-glycans, indicating that this region of the G protein ectodomain may constitute a potentially relevant target for future immunobiological strategies against lyssaviruses from both phylogroups. Nevertheless, the absence of predicted N- or O-glycosylation motifs does not necessarily guarantee epitope accessibility in native viral particles, as glycan shielding is a dynamic structural phenomenon influenced by glycoprotein conformation, host–cell glycosylation pathways, and steric accessibility [35]. However, the present analyses were based exclusively on sequence-based prediction tools and therefore could not evaluate glycan dynamics. Hence, the influence of dynamic glycan shielding and host-dependent glycosylation variability on epitope accessibility remains to be experimentally determined by glycoproteomic and solvent accessibility analyses.

Although overlapping peptides shared between phylogroups I and II were identified, it is important to emphasize that the conservation analysis was based on consensus sequences derived from representative viral species rather than on exhaustive isolate-level variability analyses. Consequently, while these regions likely reflect relatively preserved antigenic patterns within the evaluated datasets, the present approach may overlook polymorphisms occurring among circulating strains. In addition, the identification of B- and T-cell epitopes relied exclusively on in silico prediction tools and thus does not provide direct evidence of antigen processing, HLA presentation, antibody recognition, or protective immune responses in vivo. Moreover, the present workflow focused primarily on linear epitopes and therefore may not adequately capture conformational antigenic determinants that are highly relevant for antibody-mediated neutralization.

Future investigations should expand upon the present findings through integrated structural and immunological validation approaches. Three-dimensional modeling of the G protein, combined with molecular docking and molecular dynamics simulations, may help determine whether the predicted epitopes remain exposed and structurally stable under native conformational states. Additionally, experimental validation through peptide–HLA binding assays, $CD4^+$ T-cell activation analyses, neutralization assessments, in vivo immunogenicity evaluations, and protection studies in appropriate animal models will also be essential to determine whether the identified epitopes can effectively support the development of broad-spectrum immunobiological strategies against lyssaviruses.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/immuno6020037/s1, Figure S1: Multiple sequence alignment of the G protein ectodomain for lyssaviruses of phylogroups I and II.

Author Contributions: Conceptualization, C.A.M.C.; methodology, C.A.M.C.; software, A.M.C.L.; validation, T.F.S.B.C.; formal analysis, A.M.C.L.; investigation, A.M.C.L.; resources, T.F.S.B.C.; data curation, A.M.C.L.; writing—original draft preparation, A.M.C.L.; writing—review and editing, T.F.S.B.C. and C.A.M.C.; visualization, A.M.C.L.; supervision, T.F.S.B.C. and C.A.M.C.; project administration, C.A.M.C.; funding acquisition, C.A.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), grant number 88887.965184/2024-00.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors are grateful to the DTU's Department of Health Technology, the EMBL's European Bioinformatics Institute, the BU's Fuxman Bass Lab, and the IBPC's Rhone Alpes Bioinformatic Pole for providing open access to the web servers used in this study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Schneider, M.C.; Santos-Burgoa, C. Tratamiento contra la rabia humana: Un poco de su historia. *Rev. Saúde Pública* **1994**, *28*, 454–463. [CrossRef]
2. Batista, H.B.C.R.; Franco, A.C.; Roehe, P.M. Raiva: Uma breve revisão. *Acta Sci. Vet.* **2007**, *35*, 125–144. [CrossRef]
3. Badrane, H.; Bahloul, C.; Perrin, P.; Tordo, N. Evidence of two Lyssavirus phylogroups with distinct pathogenicity and immunogenicity. *J. Virol.* **2001**, *75*, 3268–3276. [CrossRef]
4. Dhulipala, S.; Uversky, V.N. Looking at the Pathogenesis of the Rabies Lyssavirus Strain Pasteur Vaccins through a Prism of the Disorder-Based Bioinformatics. *Biomolecules* **2022**, *12*, 1436. [CrossRef]
5. Lafon, M. Rabies virus receptors. *J. Neurovirol.* **2005**, *11*, 82–87. [CrossRef] [PubMed]

6. Rupprecht, C.E.; Hanlon, C.A.; Hemachudha, T. Rabies re-examined. *Lancet Infect. Dis.* **2002**, *2*, 327–343. [CrossRef]
7. Nel, L.H. Vaccines for lyssaviruses other than rabies. *Expert Rev. Vaccines* **2005**, *4*, 533–540. [CrossRef] [PubMed]
8. Medhasi, S.; Chantratita, N. Human Leukocyte Antigen (HLA) System: Genetics and Association with Bacterial and Viral Infections. *J. Immunol. Res.* **2022**, *2022*, 9710376. [CrossRef]
9. Raoufi, E.; Hemmati, M.; Eftekhari, S.; Khaksaran, K.; Mahmodi, Z.; Farajollahi, M.M.; Mohsenzadegan, M. Epitope Prediction by Novel Immunoinformatics Approach: A State-of-the-art Review. *Int. J. Pept. Res. Ther.* **2020**, *26*, 1155–1163. [CrossRef] [PubMed]
10. Sayers, E.W.; Cavanaugh, M.; Frisse, L.; Pruitt, K.D.; Schneider, V.A.; Underwood, B.A.; Yankie, L.; Karsch-Mizrachi, I. GenBank 2025 update. *Nucleic Acids Res.* **2025**, *53*, D56–D61. [CrossRef]
11. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **2022**, *40*, 1023–1025. [CrossRef]
12. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [CrossRef] [PubMed]
13. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef]
14. Madeira, F.; Madhusoodanan, N.; Lee, J.; Eusebi, A.; Niewielska, A.; Tivey, A.R.N.; Lopez, R.; Butcher, S. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **2024**, *52*, W521–W525. [CrossRef]
15. Reynisson, B.; Alvarez, B.; Paul, S.; Peters, B.; Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **2020**, *48*, W449–W454. [CrossRef]
16. Paul, S.; Lindestam Arlehamn, C.S.; Scriba, T.J.; Dillon, M.B.; Oseroff, C.; Hinz, D.; McKinney, D.M.; Carrasco Pro, S.; Sidney, J.; Peters, B.; et al. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. *J. Immunol. Methods* **2015**, *422*, 28–34. [CrossRef]
17. Soto, L.F.; Requena, D.; Fuxman Bass, J.I. Epitope-Evaluator: An interactive web application to study predicted T-cell epitopes. *PLoS ONE* **2022**, *17*, e0273577. [CrossRef]
18. Hopp, T.P.; Woods, K.R. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 3824–3828. [CrossRef] [PubMed]
19. Janin, J. Surface and inside volumes in globular proteins. *Nature* **1979**, *277*, 491–492. [CrossRef] [PubMed]
20. Karplus, P.A.; Schulz, G.E. Prediction of chain flexibility in proteins. *Naturwissenschaften* **1985**, *72*, 212–213. [CrossRef]
21. Combet, C.; Blanchet, C.; Geourjon, C.; Deléage, G. NPS@: Network protein sequence analysis. *Trends Biochem. Sci.* **2000**, *25*, 147–150. [CrossRef]
22. Gupta, R.; Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* **2002**, *7*, 310–322.
23. Hansen, J.E.; Lund, O.; Tolstrup, N.; Gooley, A.A.; Williams, K.L.; Brunak, S. NetOglyc: Prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* **1998**, *15*, 115–130. [CrossRef] [PubMed]
24. Sidney, J.; Peters, B.; Sette, A. Epitope prediction and identification- adaptive T cell responses in humans. *Semin. Immunol.* **2020**, *50*, 101418. [CrossRef]
25. Sanchez-Mazas, A. A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. *Swiss Med. Wkly.* **2020**, *150*, w20214. [CrossRef] [PubMed]
26. El-Manzalawy, Y.; Dobbs, D.; Honavar, V.G. In Silico Prediction of Linear B-Cell Epitopes on Proteins. *Methods Mol. Biol.* **2017**, *1484*, 255–264.
27. Shi, C.; Sun, P.; Yang, P.; Liu, L.; Tian, L.; Liu, W.; Wang, M.; Zheng, X.; Zheng, W. Research progress on neutralizing epitopes and antibodies for the Rabies virus. *Infect. Med.* **2022**, *1*, 262–271. [CrossRef]
28. Marissen, W.E.; Kramer, R.A.; Rice, A.; Weldon, W.C.; Niezgoda, M.; Faber, M.; Slootstra, J.W.; Meloen, R.H.; Clijsters-van der Horst, M.; Visser, T.J.; et al. Novel rabies virus-neutralizing epitope recognized by human monoclonal antibody: Fine mapping and escape mutant analysis. *J. Virol.* **2005**, *79*, 4672–4678. [CrossRef]
29. Nagarajan, T.; Marissen, W.E.; Rupprecht, C.E. Monoclonal antibodies for the prevention of rabies: Theory and clinical practice. *Antib. Technol. J.* **2014**, *4*, 1–12. [CrossRef]
30. Yang, F.; Lin, S.; Ye, F.; Yang, J.; Qi, J.; Chen, Z.; Lin, X.; Wang, J.; Yue, D.; Cheng, Y.; et al. Structural Analysis of Rabies Virus Glycoprotein Reveals pH-Dependent Conformational Changes and Interactions with a Neutralizing Antibody. *Cell Host Microbe* **2020**, *27*, 441–453. [CrossRef]
31. Callaway, H.M.; Zyla, D.; Larrous, F.; de Melo, G.D.; Hastie, K.M.; Avalos, R.D.; Agarwal, A.; Corti, D.; Bourhy, H.; Saphire, E.O. Structure of the rabies virus glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Sci. Adv.* **2022**, *8*, eabp9151. [CrossRef] [PubMed]

32. Moss, C.X.; Tree, T.I.; Watts, C. Reconstruction of a pathway of antigen processing and class II MHC peptide capture. *EMBO J.* **2007**, *26*, 2137–2147. [CrossRef]
33. Barroso, M.; Tucker, H.; Drake, L.; Nichol, K.; Drake, J.R. Antigen-B Cell Receptor Complexes Associate with Intracellular major histocompatibility complex (MHC) Class II Molecules. *J. Biol. Chem.* **2015**, *290*, 27101–27112. [CrossRef] [PubMed]
34. Greenbaum, J.; Sidney, J.; Chung, J.; Brander, C.; Peters, B.; Sette, A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics* **2011**, *63*, 325–335. [CrossRef] [PubMed]
35. Watanabe, Y.; Bowden, T.A.; Wilson, I.A.; Crispin, M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta Gen. Subj.* **2019**, *1863*, 1480–1497. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.