

Article

In Silico Structural Analysis Predicting the Pathogenicity of PLP1 Mutations in Multiple Sclerosis

Antigoni Avramouli, Marios G. Krokidis *, Themis P. Exarchos and Panagiotis Vlamos 

Bioinformatics and Human Electrophysiology Laboratory, Department of Informatics, Ionian University, 491 00 Corfu, Greece

* Correspondence: mkrokidis@ionio.gr

Abstract: The X chromosome gene *PLP1* encodes myelin proteolipid protein (PLP), the most prevalent protein in the myelin sheath surrounding the central nervous system. X-linked dysmyelinating disorders such as Pelizaeus–Merzbacher disease (PMD) or spastic paraplegia type 2 (SPG2) are typically caused by point mutations in *PLP1*. Nevertheless, numerous case reports have shown individuals with *PLP1* missense point mutations which also presented clinical symptoms and indications that were consistent with the diagnostic criteria of multiple sclerosis (MS), a disabling disease of the brain and spinal cord with no current cure. Computational structural biology methods were used to assess the impact of these mutations on the stability and flexibility of PLP structure in order to determine the role of *PLP1* mutations in MS pathogenicity. The analysis showed that most of the variants can alter the functionality of the protein structure such as R137W variants which results in loss of helix and H140Y which alters the ordered protein interface. In silico genomic methods were also performed to predict the significance of these mutations associated with impairments in protein functionality and could suggest a better definition for therapeutic strategies and clinical application in MS patients.

Keywords: myelin proteolipid protein; protein structure prediction; functional analysis; multiple sclerosis



Citation: Avramouli, A.; Krokidis, M.G.; Exarchos, T.P.; Vlamos, P. In Silico Structural Analysis Predicting the Pathogenicity of PLP1 Mutations in Multiple Sclerosis. *Brain Sci.* **2023**, *13*, 42. <https://doi.org/10.3390/brainsci13010042>

Academic Editors: Woon-Man Kung and Dina Nur Anggraini Ningrum

Received: 2 December 2022

Revised: 19 December 2022

Accepted: 20 December 2022

Published: 24 December 2022



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Myelination is an important process of the CNS that provides electrical insulation to axons and facilitates the transmission of nerve impulses. This protective layer is formed by Schwann cells in the peripheral nervous system, while oligodendrocytes form the sheath in the CNS [1]. The myelin sheath is a multi-layered membrane composed of proteins and lipids (approximately 30% and 70%, respectively). The lipid composition contains high amounts of cholesterol, phospholipids and glycolipids [2]. PLP is one of the major myelin proteins which, together with the DM20 isoform resulting from alternative splicing, constitutes 50% of the total protein. PLP plays a crucial role in the formation and maintenance of proper myelin structure and stability in the CNS [3]. It is a transmembrane and hydrophobic protein, with 48% of its sequence being non-polar or aromatic amino acids including 14 cysteine residues, which either undergo post-translational modifications and bind to fatty acids or are involved in intramolecular disulfide bonds. It has been observed that patients suffering from MS have an increased population of T-cells specific for PLP peptides and increased levels of anti-PLP_{181–230} specific antibodies were found in serum levels compared to healthy individuals and patients with other neurological diseases [4]. Human and rodent PLP share several epitopes that are recognized by T cells [5]. Other main counterparts are myelin basic protein (MBP), which constitutes 30% of the total myelin protein in the CNS, myelin oligodendrocyte glycoprotein (MOG) and myelin-associated glycoprotein (MAG) [6–8]. Smaller percentage is occupied by alpha-beta crystallin, a small heat shock protein [9].

Multiple sclerosis (MS) is a chronic demyelinating inflammatory condition affecting the human central nervous system (CNS) [10]. It is unclear whether MS begins in the periphery,

through activation of immune cells that then penetrate the CNS and cause damage, or within the CNS through primary damage to myelin or oligodendrocyte [11,12]. This could be the result of mutations in molecules encoding essential myelin or oligodendrocyte components, even though genome-wide association studies have not indicated substantial associations with any of these in MS [13]. However, the possibility that mutations in genes encoding these components are present in some proportions of MS patients remains under consideration. To that end, missense mutations in *PLP1* gene have been described in patients with clinical symptoms consistent with an MS diagnosis, such as an amino acid substitution at residue 31 of PLP (L31V) in a female patient with primary progressive MS [14]. Mutations can largely affect protein functionality, hence analysis of potential alterations in protein tertiary structure can reveal new evidence for their effect on phenotype. A novel mutation in *PLP1* exon 2 that changed leucine to arginine (L31R) was reported in a mother and her son through sequencing of the *PLP1* gene [15]. In a recent study, a mother and daughter with a preliminary diagnosis of primary progressive MS carried a nonsense mutation at codon 210 T > G [16]. Moreover, a L31P mutation was also associated with severe PMD [17]. The transmembrane portion of PLP1/DM20 proteins could be disrupted by both mutations, affecting intracellular trafficking. Neuroinflammation and axonal neurodegeneration was also reported in mice carrying the R137W and L31R mutations by one year of age [18].

The structure and function of the native protein may be significantly altered by missense mutations, particularly those in the coding area that modify the amino acid configuration. In order to determine the impact of each nonsynonymous single nucleotide polymorphism (nsSNP) in a related protein, it is common practice to functionally compare mutant proteins with their wild-type counterparts associated with specific traits in vitro [19]. However, the experimental design for each mutational modification is time- and labor-intensive. Thus, it is feasible and cost-effective to perform data mining for mutational analysis and functional prediction on protein properties using computational methods [20]. The three-dimensional (3D) structure of a protein has a pivotal role in protein's functional characterization. There are many efficient structural biology algorithms for predicting tertiary protein structures based on their amino acid sequences. Therefore, considering the role of the *PLP1* gene in spontaneous myelin and axonal damage, we retrieved all mutations in the *PLP1* gene related to MS. Using in silico structural and functional analyses, this study aimed to describe potential disease-associated variants of the *PLP1* gene.

2. Materials and Methods

A variety of different computational approaches was used to screen out the functional effects of the variants in the *PLP1* gene related to MS. The methodology we followed is divided into four distinct levels, including (i) primary data collection, (ii) creation of the 3D protein structures, (iii) 3D protein structure comparison process (iv) variant functional analysis. A baseline method raised from pure bioinformatics approaches was utilized as a benchmark for validation.

2.1. Summary of Variants

Five variants of *PLP1* gene and specifically the association of three of them with multiple sclerosis were analyzed in our study. The transcript sequence and the protein encoded by the *PLP1* gene were retrieved from the Ensembl database [21]. Then, the UniProt ID for the amino acid (P60201-MYPR_HUMAN) was obtained from UniProt Protein Database.

2.2. In Silico Methods for Predicting Mutation Significance

dbNSFP database was used for functional prediction and annotation of potential non-synonymous single-nucleotide variants (nsSNVs). The current version of this high-performance variant annotation tool can be queried to extract prediction scores from 38 algorithms [22,23]. dbNSFP also provides conservation scores and supplementary

data, such as allele frequencies, functional gene descriptions, gene expression and gene interaction data, etc. MutPred2 machine learning-based approach was implemented to predict the pathogenicity of amino acid substitutions and their molecular mechanisms [24].

Further computational tools such as ANNOVAR (<https://annovar.openbioinformatics.org/en/latest/>, accessed on 16 September 2022), KGGSeq (<http://pmglab.top/kggseq/>, accessed on 16 September 2022), VarSome (<https://varsome.com>, accessed on 16 September 2022), UCSC Genome Browser's Variant Annotation Integrator (<http://genome.ucsc.edu/cgi-bin/hgVai>, accessed on 19 September 2022), Ensembl Variant Effect Predictor (<http://www.ensembl.org/info/docs/tools/vep/index.html>, accessed on 20 September 2022), SnpSift (<https://pcingola.github.io/SnpEff/>, accessed on 20 September 2022) and HGMD (<https://www.hgmd.cf.ac.uk/ac/index.php>, accessed on 22 September 2022) were used to strengthen the analysis, and the outcomes were validated using each platform separately. The algorithms utilized in this study are publicly accessible for all academic, non-commercial uses.

2.3. Protein Stability Correlation Analysis

The correlation between mutations and protein stability was analyzed based on a lesser decrease in free energy (ΔG or dG). Alterations in protein stability are determined by differences in free energy ($\Delta\Delta G$ or ddG) between wild-type and mutant proteins [25]. The DynaMut server was used to assess the effect of a single point mutation on protein stability, conformation, and flexibility, and to visualize protein dynamics [26]. DynaMut provides more accurate (p -value < 0.001) assessments of the effects of single mutations on protein stability than other well-established methods. In addition, the DynaMut server defines $\Delta\Delta G \geq 0$ as stabilizing and $\Delta\Delta G < 0$ for comparison purposes. In addition, Site Directed Mutator (SDM) server was used to estimate the change in protein stability following mutation [27].

2.4. Analysis of Protein Structural Conformation and Conservation

ConSurf server (<http://consurf.tau.ac.il/>, accessed on 7 October 2022) was utilized to identify highly conserved functional areas of the PLP1 gene-encoded protein [28].

2.5. Prediction of the Secondary Structure

Using the PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>, accessed on 10 October 2022), the secondary structure of PLP1 was predicted [29]. It is based on a two-stage neural network with position-specific scoring matrices derived from PSI-BLAST to predict the available secondary structures of a protein.

2.6. Homology Modeling

SWISS-MODEL was utilized to determine the three-dimensional structure of PLP1. The CAMEO system determines the precision of the generated model. SWISS-MODEL is based on evolutionary information and searches a high-throughput template library (SMTL) for the optimum sequence–template alignment to construct the model [30]. Phyre-2 server was used to predict the homology-based three-dimensional structure of the query amino acid sequence [31]. I-TASSER was selected for protein structure prediction and structure-based function annotation [32–34]. Initially, structural templates from the PDB are discovered using LOMETS, a multithreaded algorithm. With the templates as guides, full-length atomic models are then constructed using simulations of fragment assembly iterations. The 3D models are re-run through the BioLiP database of protein functions to gain insight into the target's function. C-I-TASSER, an enhanced version of I-TASSER designed to accurately predict protein structures and functions was used to generate inter-residue contact maps beginning with a query sequence [34]. The structural templates of the PDB are derived by the multithreaded method LOMETS, and their full-length atomic models are constructed using contact maps and replica exchange Monte Carlo simulations. Finally, COFACTOR uses the structural model to deduce the protein's biological functions.

C-I-TASSER produces significantly more accurate models than I-TASSER in large-scale benchmark tests.

2.7. Mutated Structure Prediction

Once the mutations were identified, the construction protein-based structures in PDB format were followed. I-TASSER and DynaMut servers were used to perform the transformation of the amino acid sequences to 3D protein models. Then, for each model, structural alignment was carried out and the structural similarity score was calculated. The TM-align and TM-score algorithms were selected for the alignment and the similarity score calculation, respectively [35].

2.8. Protein Three-Dimensional Model Verification

Three-dimensional structures were validated using Ramachandran plot analysis (<http://molprobity.biochem.duke.edu>, accessed on 19 October 2022). It provides the number of residues that are located in the allowed, favored, and outlier regions. If a significant fraction of residues resides in the allowed and favored region, it is projected that the model is accurate [36].

3. Results

3.1. PLP1 Variants Associated with MS

PLP consists of 276 amino acid residues and four hydrophobic transmembrane domains, and its expression is restricted to oligodendrocyte cells. The area of the *PLP1* gene that encodes residue 31 appears to be a hotspot for mutation, since it has been described in MS patients (L31R and L31V mutation) [14]. In cases of severe PMD, the L31P mutation has also been documented [14]. The idea of mutation hotspots in *PLP1* has been previously characterized in PMD patients, and numerous mutations have been detected in a number of amino acids [37]. R137W mutation has also been described in MS patients, while the H140Y one was selected because it is the closest known mutation to residue 137 [14]. Detailed information about the variants analyzed in the current study is shown in Table 1.

Table 1. PLP1 variants. This mutation was selected as the closest mutation to residue 137 known for *PLP1*.

Location	Codon Change	Amino Acid Position	Amino Acid Alteration	Description ¹
X:103785668	CTG/GTG	31	L/V	MS- like disease mutation
X:103785669	CTG/CCG	31	L/P	Severe PMD mutation
X:103785669	CTG/CGG	31	L/R	MS- like disease mutation
X 103786682	CGG/TGG	137	R/W	MS- like disease mutation
X 103786691	CAT/TAT	140	H/Y	Mild SPG2 mutation

¹ MS: multiple sclerosis; PLP: myelin proteolipid protein; PMD: Pelizaeus–Merzbacher disease; SPG2: spastic paraplegia type 2.

3.2. Variant Functional Analysis

There are numerous assessment strategies for missense variants and recent databases include results from a variety of techniques to assist the evaluation of the impact of variations predicted to modify the peptide sequence of a gene. Herein, using dbNSFP, we investigated the functional consequences of missense SNPs, including whether they are normal, disease-causing, or effective by chance. As Table 2 shows, functional analysis revealed that R137W, L31P, L31V, L31R are damaging from the most prediction tools with a high score, while the results for H140Y were different across the different methods.

Table 2. Functional analysis of PLP variants using dbNSFP.

Variant	L31P	L31V	L31R	R137W	H140Y	Range (Low to Damaging)
Polyphen2_HDIV_score	0.999	0.997	0.999	0.999	0.015	0.03061 to 0.91137
LRT_converted_rankscore	0.8433	0.8433	0.8433	0.4496	0.53742	0.00162 to 0.8433
MutationTaster_score	1	0.999999	1	0.627105	0.281663	0 to 1
MutationTaster_converted_rankscore	0.81001	0.58761	0.81001	0.81001	0.81001	0.08979 to 0.81001
MutationAssessor_score	2.71	2.36	2.71	1.24	0.69	−5.17 to 6.49
MetaLR_score	0.9794	0.9771	0.9794	0.9547	0.9047	0 to 1
MetaRNN_score	0.988245	0.946438	0.98141	0.892939	0.962404	0 to 1
MutPred_score	0.932	0.813	0.887	0.663	0.779	0 to 1
DEOGEN2_score	0.994664	0.910984	0.994756	0.707527	0.697598	0 to 1
ClinPred_score	0.996376	0.982601	0.997494	0.958003	0.509568	0 to 1

We strengthened our analysis using further computational tools such as ANNOVAR, KGGSeq and VarSome and the outcomes were validated using each platform separately.

The results of MutPred2 demonstrated that these variants may alter the function of protein structures (Table S1). MutPred2 provides a general score which represents the average of all neural network scores based on a ranked list of specific molecular alterations potentially affecting the phenotype, and therefore, this number indicates the probability that the amino acid substitution could be harmful. A score threshold of 0.50, if considered as a probability, could reveal pathogenicity. However, a threshold of 0.68 results in a false positive rate (fpr) of 10%, whereas a threshold of 0.80 results in an fpr of 5%. In our case, L31P (score 0.973), L31V (score 0.854), and L31R (score 0.972) mutations may result in an altered transmembrane protein (Table S1). The R137W variant (score 0.684) may lead to a loss of helix, whereas the H140Y variant (score 0.556) may result in a changed ordered interface or transmembrane protein.

3.3. Conformational Analysis and Alteration of Protein Stability upon Amino Acid Substitution

DynaMut predicts the change in stability by calculating the changes in unfolding Gibbs free energy ($\Delta\Delta G$), as summarized in Table 3. For comparison, $\Delta\Delta G$ predictions based on protein structure were also displayed including distinct approaches and assumptions. Parallel analysis was performed using Site Directed Mutator (SDM) computational methods to verify the molecular effect of the five variants (Table S2). Three of them (L31V, L31P L31R) revealed a diminution in stability by increasing the molecular flexibility of the wild-type proteins (Tables 3 and S2). On the contrary, H140Y variant enhanced the stability of the PLP1 protein. R137W variant revealed conflicting results. DynaMut demonstrated that the amino acid changes in R137W decrease stability, while SDM exhibited elevated stability (Table S2). ENCoM analysis was executed to calculate the vibrational entropy difference (ΔS) between wild-type and mutant structures as well as to explore protein conformational space and the effect of mutations on protein function and stability. As Figure 1 illustrates, the mutation causes a change in the vibrational entropy of the amino acid.

Table 3. Conformational Analysis of Protein's Stability Change upon Amino Acid Substitution.

Variant	$\Delta\Delta G$ (kcal/mol)	Outcome	$\Delta\Delta S_{VibENCoM}$ (kcal.mol ⁻¹ .K ⁻¹)	Outcome ¹
L31V	−0.133	Destabilizing	0.083	Increase
L31P	−1.011	Destabilizing	0.413	Increase
L31R	−0.256	Destabilizing	0.231	Increase
R137W	−0.400	Destabilizing	0.111	Increase
H140Y	0.519	Stabilizing	−0.063	Decrease

¹ Molecule flexibility.

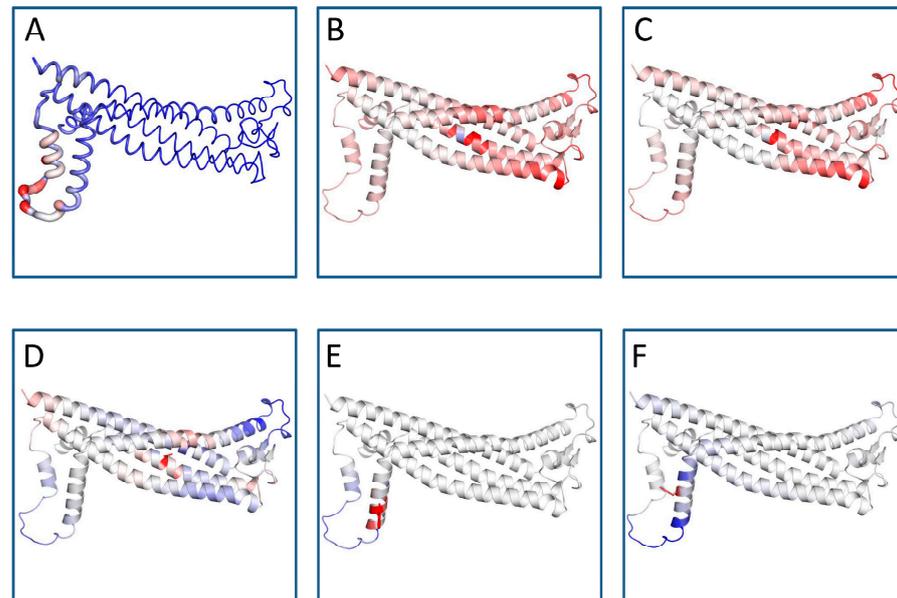


Figure 1. Protein flexible conformation based on the Vibrational Entropy difference ($\Delta\Delta S$) between wild-type and mutant structures on the structure of PLP1. A visual representation of the chain in which the mutation occurs is also mapped. Amino acids colored according to the vibrational entropy change upon mutation. Blue represents a rigidification of the structure and red represents a gain in flexibility. (A) Normal PLP1; (B) L31P mutant; (C) L31R mutant; (D) L31V mutant; (E) R137W mutant; (F) H140Y mutant. The image is illustrated by DynaMut. The positions of the point mutations are 31, 137 and 140. Abbreviations: L is leucine; P is proline; R is arginine; V is valine; W is tryptophan; H is histidine; Y is tyrosine.

3.4. Analysis of the Structural Conformation and Conservation of PLP1

According to ConSurf analysis, the variants located at position 31 (L31V, L31P, and L31R) were found in a highly conserved region with a conservation score of 9 (Figure 2). Based on this indication, we can estimate that these nsSNPs play a functional role on the protein conformation. On the contrary, R137 and H140 displayed a conservation score of 1 (Figure S1).

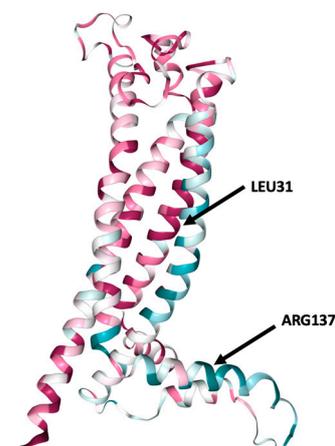


Figure 2. ConSurf analysis of conserved functional areas of the structural model of PLP1 gene-encoded protein. Amino acid at positions 31 (leucine) and 137 (arginine) are highlighted. Leucine in position 31 is a highly conserved region. Conservation score is presented in Figure S1.

3.5. PLP1 Protein Secondary Structure Prediction

The alpha helix, beta sheet distribution and coils for PLP1 were calculated according to PSIPRED protein structure prediction server. Among the exposed secondary structures, the highest in percentage was alpha helix (65%) followed by coils (30%) and no beta-sheet (0.0%) (Figure S2).

3.6. Prediction Software Benchmarking and Creation of Tertiary Protein Structures

Using four distinct homology modeling techniques, the three-dimensional (3D) structures of the PLP1-encoded protein were reconstructed. Since only 3% of residues 45–53 is represented on the protein data bank (<https://www.ebi.ac.uk/pdbe/pdbe-kb/>, accessed on 10 October 2022), there was no known crystal data of this protein of the appropriate length. Once the mutations were identified, the next step was to construct protein-based structures to represent these variants. Since the 3D protein feature view was not determined through experimental methodologies, established computational tools and databases such as Uniprot (UniProt Consortium, London, UK, 2015), Swiss-Model, Phyre-2, I-TASSER, C-I-Tasser, PDBeFold and Dynamut were evaluated for predicting the mutated structures and calculated the effect of these domain mutations on the 3D protein structure.

Based on the Hidden Markov approach, the Phyre-2 server was implemented to predict the homology-based three-dimensional structure of the query amino acid sequence. It incorporates five phases to construct a model: (1) collection of homologous sequences, (2) screening of fold library, (3) modeling of loops, (4) ab initio folding simulation Poing for multiple template modeling, and (5) placement of side chains [31]. Tertiary protein structures were formed based on the available structure prediction tools. Out of an extensive benchmarking of the structural predictive tools, we selected to retrieve the PLP1 target structure from the AlphaFold database. Comparison results revealed that AlphaFold reached the highest accuracy between predicted and experimental structure. In Figure 3, the visualization of the predicted 3D model of PLP1 protein is presented as performed by AlphaFold, Phyre-2, I-TASSER and C-I-Tasser, respectively, as these servers exhibited the highest accuracy in predicting the experimental structure. The model–template alignment by Swiss Model retrieved structures that did not include amino acids involved in the present outcomes.

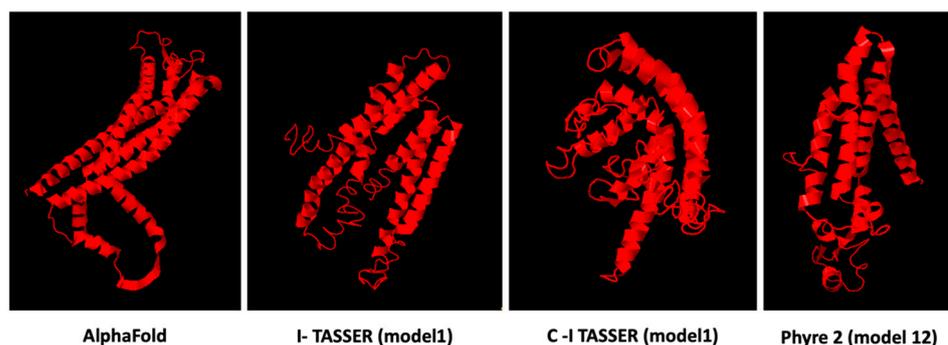


Figure 3. Protein structure prediction models of PLP1 calculated by different computational methodologies. AlphaFold model presented high confidence for the residue of the protein at position 31 (pLDDT > 90), but limited confidence for the PLP1 residue at position 137 (pLDDT < 50). Model 1 showed the highest C- score (−3.95) in I-TASSER and C-I-Tasser (−3.89) servers. Phyre2 model 12 with a confidence of 16.98% was the only one that included the residues analyzed in this study.

3.7. Variant Tertiary Protein Structures

The next step in our pipeline was to construct the structures for these mutations using I-TASSER and DynaMut [26]. The DynaMut provides a comprehensive evaluation and visualization of protein mobility and flexibility using two independent, well-established normal mode approaches to analyze protein dynamics by sampling conformations. In parallel, assessment of the effect of mutations on protein dynamics and stability due

to changes in vibrational entropy can be executed. The server combines graph-based signatures with normal mode dynamics to predict the influence of a selected mutation on protein stability. The predicted models were compared against the corresponding AlphaFold structure through the TM-align algorithm and a benchmarking of the structural predictive tools was accomplished. Comparative results revealed that DynaMut reached the highest accuracy between the predicted and experimental structure and was also used to verify the impact of mutations on protein conformation, flexibility and stability as well as to visualize protein dynamics. The TM-score is the metric that will lead to the selection of the ideal approach for producing the potential tertiary structures of a protein. TM-align generates an optimal residue-to-residue alignment based on structural similarity utilizing dynamic programming iterations for two protein structures of uncertain equivalence [38]. TM-score for the five mutated structures shows that they were approximately in the same fold with the normal protein. Figure 4 presents the predicted interatomic interactions calculated for the wild-type protein and the single point mutations. Both wild-type and mutant residues are colored in light green and depicted as sticks, along with domains participating in any interactions surrounding them.

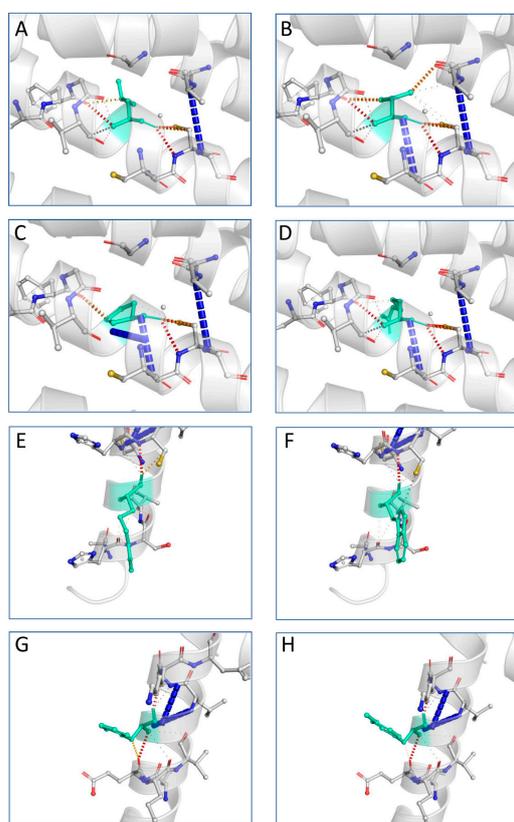


Figure 4. Interatomic interactions for wild-type and mutant PLP1. Both wild-type and mutant residues are colored in light green and depicted as sticks, along with domains participating in any interactions surrounding them. Leucine at position 31 is hydrophobic and highly conserved: (A) wild-type residue at position 31; (B) L31V (hydrophobic); (C) L31P (nonpolar); (D) L31R (polar). Arginine at position 137 and histidine at position 140 are polar: (E) wild-type residue at position 137; (F) R137W (aromatic); (G) wild-type residue at position 140; (H) H140Y (aromatic). The image is illustrated by DynaMut. A scale of color definition for each type of interaction is provided by software: red depicts hydrogen bonds; slight red depicts water-mediated hydrogen bonds; blue depicts halogen bonds; gold depicts ionic interactions; purple depicts metal complex interactions; light blue depicts aromatic contacts; green depicts hydrophobic contacts; pink depicts carbonyl contacts. The positions of the point mutations are 31, 137 and 140. Abbreviations: L is leucine; P is proline; R is arginine; V is valine; W is tryptophan; H is histidine; Y is tyrosine.

3.8. Validation of the Predicted Structures

The Ramachandran plot was used to examine the conformation of the protein's backbone. It represents an x-y plot of the phi/psi dihedral angles between NC-alpha and C-alpha-C bonds. The Ramachandran plot of the wild-type protein in the AlphaFold model revealed 259 residues (94.2%) in the favored regions, 268 (97.5%) in the allowed regions and seven residues in the outlier region (Figure S3). The mutant protein structures obtained by DynaMut demonstrated the same results with wild-type PLP1. On the contrary, I-TASSER structure prediction models display poor Ramachandran plots compared to other algorithms. I-TASSER generates a model by reassembling structural parts from various templates, hence the model occasionally features unfavorable Ramachandran plot regions (Figure S4). The homology models indicated that PLP1 protein models obtained by DynaMut were accurate and they are useful for conducting additional studies and gaining a deeper understanding of the biological activity of the studied protein.

4. Discussion

In this study, the majority of tools indicated that MS-associated PLP1 mutations would have a significant impact on the protein structure, stability and function. Our analysis employed several computational approaches to predict the effects of the PLP1 gene variants, and important results were obtained. Examination of the modified protein structure revealed a destabilizing effect and an increase in flexibility. Loss of protein thermodynamic stability can reduce the ability of its structure to perform normal functions. Furthermore, precise analysis using MutPred2 revealed that these variants may affect protein functionality and structure. We used this machine learning-based approach to integrate data to reason probabilistically about the pathogenicity of amino acid substitutions. The resulted predictions for L31P, L31V and L31R indicated that these mutations may lead to an altered transmembrane protein. The R137W variant may also cause loss of helix, while the H140Y variant may alter the ordered interface of transmembrane protein. The findings of this study provide important insights for future investigations aimed at determining the role of PLP1 in MS.

Missense mutations have a substantial effect on protein functionality. A comprehensive computational examination of the phenotypic characteristics associated with specific variants can reveal the vulnerabilities that interfere with the normal protein activity. This study suggests that mutations in myelin-related genes may play a role in the development of MS. There are two putative PLP1-related MS mechanisms: PLP1 mutations could damage oligodendrocytes [39], generating an inside-out disease process, or they could cause the expression of neoantigens that the immune system could target [14]. Both can occur concurrently, so PLP1 should be investigated further. Previous studies showed that a wide variety of PLP1 genetic alterations have been identified as the underlying causes of PMD and SPG2 [40,41]. Understanding the pathophysiology of the disorders illustrated by a genotype–phenotype correlation requires an understanding of their cellular and metabolic impacts. The consequences of pathological modifications of PLP1 gene were better understood than the physiological functions of the PLP1 protein [37]. After more than 50 years of research, most of the intracellular mechanisms related to PLP1 functionality are still unknown, although the remarkable level of sequence conservation suggests that many mutations could cause severe implications, including MS [14].

In the present study, for most of the known variants, the 3D structures of the proteins are not experimentally known, so there is a clear lack of experimental evaluations of variant effects. Prediction methods can help close the sequence-annotation gap, but with respect to deep annotations of function, *in silico* methods remain limited. These methods are mainly oriented towards intrinsically disordered proteins and clustered data are based on sequence identity thresholds, retaining a single representative sequence from each group. This approach results in models that resemble having learned a concept instead of a probability distribution. Well-defined theoretical support for this situation is an open

problem that will formalize and improve understanding of this long-standing practice in computational biology.

MS is a persistent autoimmune inflammatory disease of the human central nervous system (CNS). It is characterized by loss of motor and sensory function resulting from immune-mediated inflammation, demyelination and sequelae destruction of nerve axons. Along the axon, there are intermittent points that are not surrounded by myelin and are called junctions of Ranvier [42]. MS shows great diversity both at the point of disease onset and at the stage of developmental progression. Four main types of the disease are distinguished: Relapsing–Remitting MS (RRMS) that is characterized by clearly defined relapses of increased disease activity and the worsening of symptoms; Secondary Progressive MS (SPMS), the next step of the RRMS progress for the majority of patients; Primary Progressive MS (PPMS), presenting with symptoms that have been steadily worsening since onset of the disease, without relapses or remissions; and finally Progressive Relapsing MS (PRMS) that is progressive from onset with continuous worsening between relapses [43,44]. Myelin proteins are considered potential targets of the immune system in MS, and activated T-cells recognize specific myelin epitopes at sites of extensive demyelination. According to clinical, pathologic, imaging and electrophysiologic studies, it is not yet understood whether MS is beginning in the periphery, by stimulation of immune cells that thereafter penetrate the CNS and cause damage, or within the CNS through primary myelin or oligodendrocyte injury [45]. This could be the result of mutations in molecules encoding critical components of myelin or oligodendrocytes despite the fact that genome-wide association studies have not found significant links between them and multiple sclerosis. However, it remains possible that mutations in genes encoding these components may be present in a subset of MS patients. In this regard, missense mutations in PLP1 have been identified in patients exhibiting clinical symptoms consistent with a diagnosis of multiple sclerosis [14].

Although the pathogenesis of MS remains unclear, multiple genes, generally of poor penetrance, have been related to MS susceptibility, and their nature suggests autoimmunity causes disease development in most cases [46]. MS is a serious autoimmune disease, unfortunately without a cure; however, over the last three decades, there has been a rapid expansion of therapeutic approaches for the disorder including immunoprotective strategies, shingosine-1-phosphate receptor modulators and cell-based therapies [47]. Emphasis should be placed on early identification of risk factors for early therapeutic interventions. The disease has a different pathogenetic factor in each patient. PLP1 mutations L31V, L31R and R137W could impair PLP trafficking out of the ER and induce the unfolded protein response (UPR). The data imply that PLP1 mutations could have a harmful effect on oligodendrocyte functionality and consequently cause MS [14]. This is confirmed by recent finding in mice carrying the L31R and R137W mutations: they showed neuroinflammation, axonal degeneration, neuronal loss, and brain shrinkage by one year [18]. The same mutations and the loss of function of glial *PLP1* gene indicated a clinical scenario similar to MS in humans. The area of PLP1 gene encoding residue 31 appears to be a hotspot for mutation as L31P has been linked to severe PMD. The L31V mutation shows the least effects on PLP expression, trafficking, or UPR induction is a conservative mutation, as we already stressed [14]. It is not expected to have a significant impact on the hydrophobicity of the first transmembrane region in which it is located, as L and V are hydrophobic amino acids with similar structures and neutral side chains. An L31R mutation in the first transmembrane domain of PLP could affect the overall charge, hydrophobicity, and/or secondary structure of the transmembrane helix, disrupting PLP structure. The L31P mutation would force a stiff bend on the polypeptide and damage the transmembrane helix. R137W occurs in exon 3B, which is deleted in PLP DM20. DM20 is expressed before PLP during ontogenesis and may play a role in the development of new oligodendrocytes [48]. Several L31V-mutated peptides were expected to bind with higher affinity to some of the patient's HLA molecules than the native peptide, producing de novo epitopes and potentially inducing/activating a new group of autoreactive T cells [14,49]. Such responses depend on the presence of

proteases that can digest peptides and T cells that can recognize novel epitopes in the patient's T cell repertoire.

5. Conclusions

PLP1 plays an important role in myelin structure and stability, an insulating lipoprotein which helps transmit nerve impulses. Numerous computational tools were utilized in the present in silico analysis, which demonstrated that the amino acid changes L31V, L31R, and R137W of the PLP1 protein are functionally detrimental. The L31V and L31R variants of PLP1 reside in the conserved domain of the protein. To examine the stability of mutant and wild-type PLP1 proteins, we also calculated the changes in their free energies. Our findings provide evidence for the functional role of these three variations, which facilitates the establishment of accurate insights for drug targeting and future clinical application in patients with multiple sclerosis. Alteration of overall cellular activity often arises as a consequence of altered function of one or more individual proteins. Identification of more variants as specific targets may provide a better understanding of conformational dynamics for future studies, while molecular recognition specific to mutated proteins will play an important role in broadening the scope of intracellular mechanisms involved in inflammatory demyelinating diseases.

Key Points

1. The majority of computational tools indicated that MS-associated *PLP1* mutations would have a significant impact on the protein structure, stability and function.
2. Loss of protein thermodynamic stability can reduce the ability of its structure to perform normal functions.
3. The resulted predictions for L31P, L31V and L31R indicated that these mutations may lead to an altered transmembrane protein.
4. The R137W variant may also cause loss of helix, while the H140Y variant may alter the ordered interface of transmembrane protein.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/brainsci13010042/s1>, Table S1: Prediction of pathogenicity of amino acid substitutions of *PLP1* variants; Table S2: Analysis of missense mutations on protein stability; Figure S1: Prediction of evolutionary conserved amino acid residues of PLP1; Figure S2: Protein secondary structure predictions of PLP1; Figure S3 and Figure S4: Ramachandran plots.

Author Contributions: A.A., M.G.K., T.P.E. and P.V. contributed to conceptualization, methodology, data curation, writing—original draft preparation and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call Regional Excellence (Research Activity in the Ionian University, for the study of protein folding in neurodegenerative diseases) (FOLDIT) MIS 5047144.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Salzer, J.L. Schwann cell myelination. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a020529. [[CrossRef](#)] [[PubMed](#)]
2. Poitelon, Y.; Kopec, A.M.; Belin, S. Myelin fat facts: An overview of lipids and fatty acid metabolism. *Cells* **2020**, *9*, 812. [[CrossRef](#)] [[PubMed](#)]
3. Greer, J.M.; Pender, M.P. Myelin proteolipid protein: An effective autoantigen and target of autoimmunity in multiple sclerosis. *J. Autoimmun.* **2008**, *31*, 281–287. [[CrossRef](#)] [[PubMed](#)]

4. Greer, J.M.; Trifilieff, E.; Pender, M.P. Correlation between anti-myelin proteolipid protein (PLP) antibodies and disease severity in multiple sclerosis patients with PLP response-permissive HLA types. *Front. Immunol.* **2020**, *11*, 1891. [[CrossRef](#)] [[PubMed](#)]
5. Mangalam, A.K.; Khare, M.; Krco, C.; Rodriguez, M.; David, C. Identification of T cell epitopes on human proteolipid protein and induction of experimental autoimmune encephalomyelitis in HLA class II-transgenic mice. *Eur. J. Immunol.* **2004**, *34*, 280–290. [[CrossRef](#)]
6. Martinsen, V.; Kursula, P. Multiple sclerosis and myelin basic protein: Insights into protein disorder and disease. *Amino Acids* **2021**, *54*, 99–109. [[CrossRef](#)]
7. Peschl, P.; Bradl, M.; Höftberger, R.; Berger, T.; Reindl, M. Myelin oligodendrocyte glycoprotein: Deciphering a target in inflammatory demyelinating diseases. *Front. Immunol.* **2017**, *8*, 529. [[CrossRef](#)]
8. Quarles, R.H. Myelin-associated glycoprotein (MAG): Past, present and beyond. *J. Neurochem.* **2007**, *100*, 1431–1448. [[CrossRef](#)]
9. Rothbard, J.B.; Zhao, X.; Sharpe, O.; Strohm, M.J.; Kurnellas, M.; Mellins, E.D.; Robinson, W.H.; Steinman, L. Chaperone activity of α B-crystallin is responsible for its incorrect assignment as an autoantigen in multiple sclerosis. *J. Immunol.* **2011**, *186*, 4263–4268. [[CrossRef](#)]
10. Höftberger, R.; Lassmann, H. Inflammatory demyelinating diseases of the central nervous system. *Handb. Clin. Neurol.* **2018**, *145*, 263–283.
11. Mapunda, J.A.; Tibar, H.; Regragui, W.; Engelhardt, B. How Does the Immune System Enter the Brain? *Front. Immunol.* **2022**, *13*, 805657. [[CrossRef](#)] [[PubMed](#)]
12. Dhaiban, S.; Al-Ani, M.; Elemam, N.M.; Al-Aawad, M.H.; Al-Rawi, Z.; Maghazachi, A.A. Role of peripheral immune cells in multiple sclerosis and experimental autoimmune encephalomyelitis. *Sci* **2021**, *3*, 12. [[CrossRef](#)]
13. Cotsapas, C.; Mitrovic, M. Genome-wide association studies of multiple sclerosis. *Clin. Transl. Immunol.* **2018**, *7*, e1018. [[CrossRef](#)] [[PubMed](#)]
14. Cloake, N.C.; Yan, J.; Aminian, A.; Pender, M.P.; Greer, J.M. PLP1 mutations in patients with multiple sclerosis: Identification of a new mutation and potential pathogenicity of the mutations. *J. Clin. Med.* **2018**, *7*, 342. [[CrossRef](#)]
15. Warshawsky, I.; Rudick, R.A.; Staugaitis, S.M.; Natowicz, M.R. Primary progressive multiple sclerosis as a phenotype of a PLP1 gene mutation. *Ann. Neurol.* **2005**, *58*, 470–473. [[CrossRef](#)]
16. Rubegni, A.; Battisti, C.; Tessa, A.; Cerase, A.; Doccini, S.; Santorelli, F.M.; Federico, A. SPG2 mimicking multiple sclerosis in a family identified using next generation sequencing. *J. Neurol. Sci.* **2017**, *375*, 198–202. [[CrossRef](#)]
17. Cailloux, F.; Gauthier-Barichard, F.; Mimault, C.; Isabelle, V.; Courtois, V.; Giraud, G.; Dastugue, B.; Boespflug-Tanguy, O.; Clinical European Network on Brain Dismyelinating Disease. Genotype-phenotype correlation in inherited brain myelination defects due to proteolipid protein gene mutations. *Eur. J. Hum. Genet.* **2000**, *8*, 837–845. [[CrossRef](#)]
18. Groh, J.; Friedman, H.C.; Orel, N.; Ip, C.W.; Fischer, S.; Spahn, I.; Schäffner, E.; Hörner, M.; Stadler, D.; Buttmann, M.; et al. Pathogenic inflammation in the CNS of mice carrying human PLP1 mutations. *Hum. Mol. Genet.* **2016**, *25*, 4686–4702.
19. Robert, F.; Pelletier, J. Exploring the impact of single-nucleotide polymorphisms on translation. *Front. Genet.* **2018**, *9*, 507. [[CrossRef](#)]
20. Flores, S.C.; Alexiou, A.; Glaros, A. Mining the Protein Data Bank to improve prediction of changes in protein-protein binding. *PLoS ONE* **2021**, *16*, e0257614. [[CrossRef](#)]
21. Zerbino, D.R.; Achuthan, P.; Akanni, W. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46*, D754–D761. [[CrossRef](#)] [[PubMed](#)]
22. Liu, X.; Jian, X.; Boerwinkle, E. dbNSFP: A lightweight database of human non-synonymous SNPs and their functional predictions. *Hum. Mutat.* **2011**, *32*, 894–899. [[CrossRef](#)] [[PubMed](#)]
23. Liu, X.; Li, C.; Mou, C.; Dong, Y.; Tu, Y. dbNSFP v4: A comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med.* **2020**, *12*, 103. [[CrossRef](#)] [[PubMed](#)]
24. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.A.; Lin, G.N.; Nam, H.; Mort, M.; Cooper, D.N.; Sebat, J.; Iakoucheva, L.M.; et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat. Commun.* **2020**, *11*, 5918. [[CrossRef](#)] [[PubMed](#)]
25. Tokuriki, N.; Stricher, F.; Serrano, L.; Tawfik, D.S. How protein stability and new functions trade off. *PLoS Comput. Biol.* **2008**, *4*, e1000002. [[CrossRef](#)]
26. Rodrigues, C.H.; Pires, D.E.; Ascher, D.B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* **2018**, *46*, W350–W355. [[CrossRef](#)]
27. Worth, C.L.; Preissner, R.; Blundell, T.L. SDM—A server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* **2011**, *39*, 215–222. [[CrossRef](#)]
28. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, W344–W350. [[CrossRef](#)]
29. Buchan, D.W.; Minnici, F.; Nugent, T.C.; Bryson, K.; Jones, D.T. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res.* **2013**, *41*, W349–W357. [[CrossRef](#)]
30. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195–201. [[CrossRef](#)]
31. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)] [[PubMed](#)]

32. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]
33. Yang, J.; Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **2015**, *43*, W174–W181. [[CrossRef](#)] [[PubMed](#)]
34. Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Bell, E.W.; Zhang, Y. Folding non-homology proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Rep. Methods* **2021**, *1*, 100014. [[CrossRef](#)] [[PubMed](#)]
35. Pandit, S.B.; Skolnick, J. Fr-tm-align: A new protein structural alignment method based on fragment alignments and the tm-score. *BMC Bioinform.* **2008**, *9*, 531. [[CrossRef](#)]
36. 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. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293–315. [[CrossRef](#)]
37. Khalaf, G.; Mattern, C.; Begou, M.; Boespflug-Tanguy, O.; Massaad, C.; Massaad-Massade, L. Mutation of proteolipid protein 1 gene: From severe hypomyelinating leukodystrophy to inherited spastic paraplegia. *Biomedicines* **2022**, *10*, 1709. [[CrossRef](#)]
38. Zhang, Y.; Skolnick, J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **2005**, *33*, 2302–2309. [[CrossRef](#)]
39. Sima, A.A.; Pierson, C.R.; Woltjer, R.L.; Hobson, G.M.; Golden, J.A.; Kupsky, W.J.; Schauer, G.M.; Bird, T.D.; Skoff, R.P.; Garbern, J.Y. Neuronal loss in Pelizaeus–Merzbacher disease differs in various mutations of the proteolipid protein 1. *Acta Neuropathol.* **2009**, *118*, 531–539. [[CrossRef](#)]
40. Bonnet-Dupeyron, M.N.; Combes, P.; Santander, P.; Cailloux, F.; Boespflug-Tanguy, O.; Vours-Barrière, C. PLP1 splicing abnormalities identified in Pelizaeus–Merzbacher disease and SPG2 fibroblasts are associated with different types of mutations. *Hum. Mutat.* **2008**, *29*, 1028–1036. [[CrossRef](#)]
41. Grossi, S.; Regis, S.; Biancheri, R.; Mort, M.; Lualdi, S.; Bertini, E.; Uziel, G.; Boespflug-Tanguy, O.; Simonati, A.; Corsolini, F.; et al. Molecular genetic analysis of the PLP1 gene in 38 families with PLP1-related disorders: Identification and functional characterization of 11 novel PLP1 mutations. *Orphanet J. Rare Dis.* **2011**, *6*, 40. [[CrossRef](#)] [[PubMed](#)]
42. Lubetzki, C.; Sol-Foulon, N.; Desmazières, A. Nodes of Ranvier during development and repair in the CNS. *Nat. Rev. Neurol.* **2020**, *16*, 426–439. [[CrossRef](#)] [[PubMed](#)]
43. Klineova, S.; Lublin, F.D. Clinical course of multiple sclerosis. *Cold Spring Harb. Perspect. Med.* **2018**, *8*, a028928. [[CrossRef](#)] [[PubMed](#)]
44. Cree, B.A.; Arnold, D.L.; Chataway, J.; Chitnis, T.; Fox, R.J.; Ramajo, A.P.; Murphy, N.; Lassmann, H. Secondary progressive multiple sclerosis: New insights. *Neurology* **2021**, *97*, 378–388. [[CrossRef](#)] [[PubMed](#)]
45. Oudejans, E.; Luchicchi, A.; Strijbis, E.M.; Geurts, J.J.; van Dam, A.M. Is MS affecting the CNS only? Lessons from clinic to myelin pathophysiology. *Neurol. Neuroimmunol. Neuroinflamm.* **2021**, *8*, e914. [[CrossRef](#)] [[PubMed](#)]
46. Patsopoulos, N.A. Genetics of multiple sclerosis: An overview and new directions. *Cold Spring Harb. Perspect. Med.* **2018**, *8*, a028951. [[CrossRef](#)] [[PubMed](#)]
47. Yang, J.H.; Rempe, T.; Whitmire, N.; Dunn-Pirio, A.; Graves, J.S. Therapeutic advances in multiple sclerosis. *Front. Neurol.* **2022**, *13*, 824926. [[CrossRef](#)]
48. Spörkel, O.; Uschkureit, T.; Büssov, H.; Stoffel, W. Oligodendrocytes expressing exclusively the DM20 isoform of the proteolipid protein gene: Myelination and development. *Glia* **2002**, *37*, 19–30. [[CrossRef](#)]
49. Wang, P.; Sidney, J.; Dow, C.; Mothé, B.; Sette, A.; Peters, B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput. Biol.* **2008**, *4*, e1000048. [[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.