TLR4-Targeted Compounds as New Therapies for Alzheimer's Disease

Owen Qu

Received October 10, 2024
Accepted January 14, 2025
Electronic access January 31, 2025

The accumulation of the β-amyloid peptide (Aβ) in the brain, along with the deterioration of acetylcholine neurotransmitters, are hallmarks of Alzheimer's Disease, a progressive neurodegenerative disease that is the leading cause of dementia. While current FDA-approved therapies can delay and mitigate symptoms such as neuroinflammation and neuron death, there is no way to fully prevent people from getting Alzheimer's. The target protein of this experiment is TLR4, a protein causing many of the physical symptoms of Alzheimer's, including amyloid plaque buildup and neuroinflammation. Thus, researching drug candidates to mediate a reaction between it and TLR4 can help reduce and mitigate these symptoms. Using the 3D model of TLR4, binding site analysis tools such as Proteins Plus, FTSite, and Prankweb, are used to generate predictions regarding the best binding sites. Then, pharmacophore maps are generated using PocketQuery, which are then evaluated by ZINCPharmer to find 20 promising compounds. SwissDock is used to evaluate protein-ligand binding affinity based on 3D models of a compound and TLR4. SwissADME and ProTox are run to assess the drug-likeness of a compound based on absorption, distribution, metabolism, excretion, and toxicity properties. Numerous compounds were found with a change in Gibbs-free energy of less than -7.0, three of which had toxicity values below the average of current FDA-approved drugs. In addition, all compounds analyzed met criteria for absorption and permeability–properties that allow the drug to enter the human body. Ultimately, the study finds a promising compound with necessary properties well within industry standards that can be the topic of further study.

Introduction

Background

Alzheimer's Disease (AD), the leading cause of dementia, is a devastating irreversible neurodegenerative disease[1] most commonly affecting the elderly characterized by cognitive decline and memory loss[2]. Early stages of Alzheimer's involve mild memory loss, though this quickly progresses to severe impairment in language, problem-solving, and memory. The World Health Organization (WHO) estimates over 49 million people suffer from the disease worldwide[3]. With a patient developing AD every three seconds, it is estimated that the number of people affected by the condition will more than double in the next two decades[4]. Being an incurable disease, Alzheimer's has a high death rate as the neurological damage causes patients to lose basic body functions such as eating and drinking which leads to malnutrition and dehydration, complications that can result in death. Despite extensive research into the causes of Alzheimer's, costing 40 billion dollars, no foolproof therapy exists, highlighting the need for continuous work and progress into better treatments and, ultimately, a cure for this destructive disease.

Cholinergic Hypothesis

A theorized cause of cognitive dysfunction is the deterioration of neurons and loss of neurotransmission. Specifically, acetylcholine (ACh), a neurotransmitter with a role in memory, learning, and attention, is rapidly broken down by acetylcholinesterase (AChE)[5], interfering with neurotransmission. When this occurs, neurotransmission in the brain is disrupted, leading to memory and cognitive issues, hallmarks of an AD diagnosis. Observations that correlated cholinergic system issues with cognitive impairment[6] gave the cholinergic hypothesis widespread acceptance. In addition, other neurological processes related to AD damage and destroy cells that produce ACh[7], further limiting the function of neural cells. ACh is critical for the natural cognitive functions humans have as it also acts as a neuromodulator that can act on other

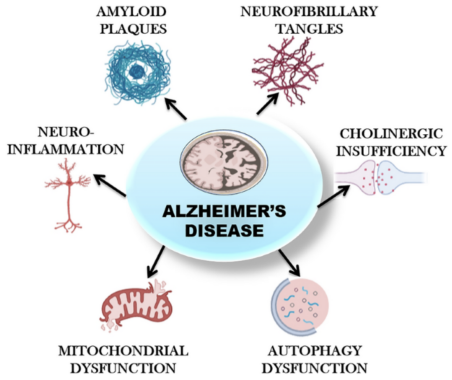


Fig. 1 A diagram of the 6 major causes of Alzheimer's Disease.

neurotransmitter systems controlling dopamine and serotonin. Thus, the breakdown process of ACh can cause detrimental effects on neurological functions and the innate immune system. Given this, an abundance of studies and research has been done on AChE inhibitors that prevent the breaking down of ACh, delaying the progression of AD into more severe stages. However, the scope for drug discovery has since broadened from this hypothesis, taking on a range of other different theories to research.

Amyloid Plaques and Neuroinflammation Theories

The second and third theories regarding the cause of AD involve the β-amyloid peptide (Aβ). A hallmark pathology required for diagnosing AD is the buildup of Aβ peptides, occurring when the amyloid precursor protein (APP) breaks down[8] This proteolysis forms Aβ peptides of different lengths, most notably, Aβ42, which is extremely prone to aggregation and forming plaques[9]. The accumulation of these amyloids disrupts cell functions involved with neurotransmission systems, and Aβ42, in particular, possesses neurotoxic properties, damaging the central nervous system[10]. In addition to forming plaques in the brain, Aβ can also cause neuroinflammation. Neurodegenerative areas in the brain show a tendency to release proinflammatory cytokines, proteins that regulate immune responses[11], contributing to neuroinflammation, which has been largely attributed to increased cytokine production[12]. Microglia, critical cells in the brain that regulate neural networks[13], lose their homeostasis functions when prolonged neuroinflammation occurs, leading to cognitive decline. Currently, studies are largely focused on creating Aβ inhibitors that prevent the assembly of the Aβ peptide by binding to them at specific places, thus reducing the likelihood of the overproduction of cytokines and neuroinflammation.

FDA-approved Therapies

Though there is no known cure for AD, many treatments are available that mitigate symptoms. The three drugs mainly prescribed for AD are all AChE inhibitors. Donepezil, the most common medication for AD, prevents ACh from being broken down by blocking acetylcholinesterase[14]. However, it cannot fully prevent the worsening of Alzheimer's, only being able to delay the symptoms of late-stage AD. In addition, ACh appears elsewhere in the body, allowing for skeletal muscle contractions[15]. Thus, taking Donepezil has many unwanted side effects, including muscle cramps[16]. Similarly, Rivastigmine is another common drug that mitigates the symptoms of Alzheimer's by blocking both AChE and Butyrylcholinesterase (BuChE)[17], a backup enzyme for AChE that identifies poisons that may inhibit AChE[18]. The third most common drug for AD is Galantamine, which is a much more potent AChE inhibitor compared to Donepezil and Rivastigmine[19]. It is also better at targeting Aβ plaques than the other two drugs, allowing it to reduce neuroinflammation[20]. Unfortunately, despite the medicinal abilities of these drugs, they still can only delay the symptoms of AD and are unable to offer full prevention against the neurodegenerative disease. Ultimately, we still don't know the root cause of AD. Outside of theories positing that Alzheimer's is caused by abnormal buildups of protein, scientists are unable to determine precisely what causes the illness, which is why we can currently only improve symptoms rather than fully preventing the progression of the disease.

TLR4

Toll-like receptors (TLRs) belong to the family of pattern recognition receptors (PRRs) which play critical roles in the innate immune system[21]. They regulate inflammatory reactions, recognize pathogens, and activate immune system responses to eliminate infection and bacteria[22]. TLR4 is a type of TLR responsible for detecting types of damage-associated molecular patterns (DAMPS)[23] that could contribute to pathological inflammation[24]. They are also responsible for dealing with a specific type of bacteria known as gram-negative bacteria (GNB), which are highly resistant to antibiotics[25]. The outer membrane of GNBs carries compounds that activate them and allow them to then carry out proinflammatory responses[26] through a series of activation pathways. Thus, TLR4 receptors naturally play a vital role in resisting otherwise difficult to manage diseases and illnesses by activating immune system responses.

TLR4 in Alzheimer's Disease

As previously discussed, TLR4 receptors detect different types of DAMPS and activate immune system responses. However, the response doesn't always fix the issue–the ability of the brain to self-regulate inflammatory signals is flawed. When TLR4 detects harmful DAMPS, it enacts a prolonged inflammatory response which induces neuroinflammation[27]. Since neuroinflammation is a known cause of AD as it damages neurons, TLR4 in fact "overreacts" and stimulates the progression of AD[28]. However, though TLR4 causes detrimental neuroinflammation, it also increases Aβ clearance in the early stages of the disease, allowing the brain to clear Aβ deposits and avoid the accumulation of neurotoxic Aβ42[29]. It is able to keep Aβ plaques from forming by activating microglia, which are able to internalize Aβ deposits[30]. However, as AD progresses into later stages, TLR signalling starts to become dysfunctional due to persistent stimulation by Aβ plaques, leading to decreased Aβ clearance and accelerated disease progression[31]. Thus, it is vital to search for drugs and medicine that keeps the beneficial reactions that TLR4 induces while regulating the prolonged

neuroinflammatory responses in order to prevent the root causes of AD. It must be stressed that Alzheimer's is a deeply complex disease–numerous factors play a role in the development of AD. While it's not the only pathway through which AD occurs, TLR4 remains a promising protein target that, with more research, could lead to the cure for Alzheimer's Disease.

Concerns with TLR4 Modulation

Clearly, inhibiting TLR4 presents a complex therapeutic trade-off in Alzheimer's disease treatment. While TLR4 plays a role in the inflammatory response induced by DAMPs detection, it also serves multiple crucial functions that complicate its targeting. First, TLR4 is highly important for the innate immune system in defending against pathogens and controlling bacterial infections, meaning its over-inhibition can cause increased susceptibility to other diseases. Additionally, TLR4 has a beneficial role in Aβ clearance, and inhibiting it could potentially interfere with this natural clearance mechanism and inadvertently worsen amyloid accumulation. This dual role creates a significant challenge for drug development, as reducing neuroinflammation through TLR4 inhibition might simultaneously compromise both immune defense and Aβ homeostasis. That being said, a controlled about of inhibition can safely isolate the benefits of the TLR4 compound while simultaneously preventing harmful effects.

The Study

While Alzheimer's Disease has clear devastating impacts on affected patients and their families, a cure is yet to be found. If promising compounds are discovered by this study, further advancements in compounds with a similar pharmacokinetic profile could lead to more effective treatment of AD. Thus, the goal of this research is to identify drug candidates that have high binding affinity with TLR4 using computational tools to potentially use in future physical screening. However, there are some limitations–the scope of the study is limited to theoretical factors, not experimental; this means that conducting physical tests on the compounds is needed. The contents of this study are based on predictions from a number of computational tools, not on experiments. While the computational tools used have high accuracy and precision, biophysical screening through MST, SPR, or ITC, is needed to test for more factors. In the long term, clinical trials would further assess the pharmacokinetics of a potential compound, and testing if the compound reduces neuroinflammation in genetically engineered mice showing symptoms of AD would be even more affirmative of its effectiveness. In addition, selectivity, or the drug's ability to have the desired effect, can be tested using other computational tools to ensure that it is able to act as a TLR4 inhibitor without having many unwanted side effects. The ultimate goal of the research is to identify promising compounds targeting the TLR4 receptor to further advance therapies of Alzheimer's Disease and to mitigate the shortcomings with current FDA-approved medicine. If promising compounds are discovered, further advancements in compounds with a similar pharmacokinetic profile could lead to more effective treatment of AD.

Methodology Overview

The research methodology employed included entering parameters and settings into specific computational tools to evaluate different factors of compounds and pharmacophore maps. These results are then used to create a comprehensive list of the compounds with the most suitable values across different properties.

Methodology

Analysis of Binding Sites using Proteins Plus

1. Go to the ProteinsPlus website: proteins.plus
2. Enter the PDB ID of 2Z64
3. Select chain A and leave other settings on default
4. Click "Calculate"

Analysis of Binding Sites using FTSite

1. Go to the FTSite website: ftsite.bu.edu
2. Enter the PDB ID of 2Z64
3. Enter "A" under "PDB Chain Ids"
4. Click "Find My Binding Site"

Analysis of Binding Sites using PrankWeb

1. Go to the PrankWeb website: prankweb.cz
2. Enter the PDB code of 2Z64
3. Deselect chain C, leaving chain A
4. Click "Submit"

Searching for Pharmacophore Maps using PocketQuery

1. Go to the PocketQuery website: pocketquery.csb.pitt.edu/pocket.html
2. Enter the PDB code of 2Z64
3. Enter "C" under chain ID
4. Click "Search"

Transferring a Pharmacophore Map to ZINCPharmer

1. Click on the desired pharmacophore map on PocketQuery
2. Select the "Export" tab in the bottom-left corner
3. Click "send to ZINCPharmer"

Analysis of Pharmacophore Maps using ZINCPharmer

1. Go to the tab labeled "Viewer"
2. Deselect "Visible" for both Ligands and Receptor Residues
3. Change pharmacophore classes to include by selecting and deselecting the "Enabled" button
4. Click "Submit Query"

Analysis of Protein-Ligand Binding Affinity using SwissDock

1. Go to the SwissDock website: swissdock.ch
2. Copy and paste the SMILES string from ZINC12 into SwissDock under "1. Submit a ligand"
3. Enter the desired box center coordinates and box dimensions under "Define search space"
4. Click "Check parameters" with the Number of RIC at 1
5. Click "START DOCKING"

Evaluating Drug-Likeness using SwissADME

1. Go to the SwissADME website: swissadme.ch
2. Enter the ZINC ID of the desired compound into the "Quick Search Bar" and click "Go"
3. Copy and paste the SMILES string from ZINC12 into SwissADME
4. Click "Run!"

Evaluating Toxicity using ProTox

1. Go to the ProTox website: https://tox.charite.de/protox3/index.php?site=home
2. Enter the ZINC ID of the desired compound into the "Quick Search Bar" and click "Go"
3. Copy and paste the SMILES string from ZINC12 into ProTox
4. Click "Start Tox-Prediction"

Procedural Concerns

In conducting this computational research, ethical standards were adhered to by ensuring transparency in the methodologies–this research did not involve any animal studies or human subjects and relied entirely on databases and tools already available from public sources. Additionally, procedures followed in this research project were validated with the known TLR4-MD-2 complex Eritoran and compared with studies evaluating docking procedures with the complex. The computational screening was further validated by testing it against a set of well-characterized TLR4 inhibitors. The validation set included established compounds such as TAK-242 (Resatorvid), CLI-095, and Eritoran, which have demonstrated TLR4 inhibitory activity in previous studies. The predicted binding poses for these known inhibitors aligned well with experimentally determined binding modes from previous studies, further supporting the robustness of our methodology.

Results and Discussion

Protein-Ligand Binding Sites

Proteins are some of the most important cells in our bodies–they act as structural building blocks, hormones, and enzymes, and initiators of cellular death[32]. They also participate in chemical bonding and mediate reactions to achieve important biological functions[33]. By binding with other molecules and ions, they can mediate reactions to achieve corresponding functions. Proteins are made up of strands of 20 different types of amino acids[34], which are required for the synthesis of, among other important compounds, neurotransmitters[35]. In order to bond with small molecules and ligands, specific locations of the protein are needed. These are typically crevice-like pockets on the surface of the protein known as ligand binding sites (LBSs). They are highly important for designing therapeutic compounds to moderate the functions of a protein. In the case of Alzheimer's, the target protein is TLR4, as its response to DAMPS is one of the main causes of AD. Finding potential LBSs on the surface of TLR4 is the first step to discovering new therapies for the neurodegenerative disease.

Protein Data Bank

The Protein Data Bank (PDB) is a worldwide archive of structural data of biomolecules. It is often used for computational analysis of macromolecules, proteins, and nucleic acids. This database is highly important for searching for binding sites, as it is able to provide the 3D structural model of the target protein, thus allowing tools to search for pockets that could act as LBSs. For AD, the required protein is a complex between TLR4 and MD-2. MD-2 is a molecule responsible for mediating the reaction of TLR4 when in contact with the LPS carried by

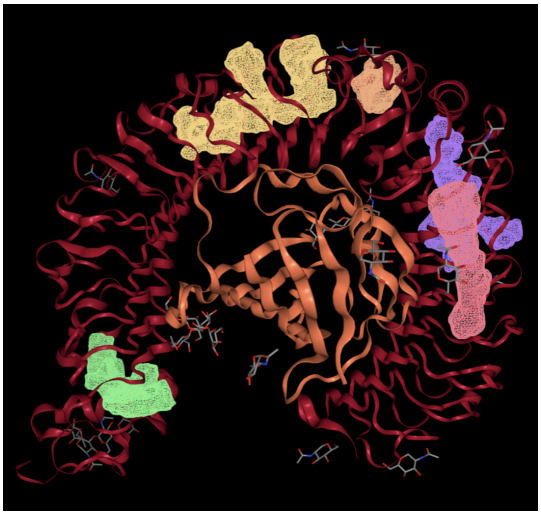


Fig. 2 3D model of potential binding sites in colored clumps. TLR4 is shown in red amino acid chains and MD-2 is shown in orange amino acid chains.

gram-negative bacteria[36]. Thus, MD-2 is required for the LPS pathway of TLR4 activation and is physically associated with TLR4 on a molecular level. Therefore, the molecular compound we use from PDB is the TLr4–MD-2 complex. Every molecular model in the PDB has a unique identification code, known as a PDB code. It is used so that tools can access the 3D model of a specific protein. The PDB code of the TLR4 and MD-2 complex is 2Z64, which is used by the software to collect the structural data to conduct binding site analysis.

Geometric Method

The first major factor when identifying potential binding sites is the shape and size. Small ligand binding occurs within cavities found on the surface of proteins that often tunnel deep within the protein. The opening of the hollow must be of adequate size–if it is too small, the ligand will not fit, and if it is too large, the ligand could be exposed to other cells that interfere with the binding. Therefore, the first characteristic to look for in binding site candidates is an appropriate size and shape. This method is implemented through the use of ProteinsPlus and DoGSiteScorer, a grid-based binding site detector solely focused on the 3D structure. The PDB code of 2Z64 is entered and the DoGSiteScorer Binding site detection software is selected. The 2Z64 protein is a complex between TLR4, corresponding to chain A, and MD-2, corresponding to chain C, but we only want binding sites candidates from the TLR4 subprotein. Thus, we select chain A, unselect chain C, and calculate with the other default settings. We obtain the following chart and 3D model:

A drug score, or druggability score, is a number between 0 and 1, with a higher druggability score indicating a better ability to bind with small molecules and ligands[37]. Note that

Table 1: Potential LBSs on the surface of TLR4 from DoGSiteScorer sorted by Drug Score.

Color Represented in Model	Volume ($Å^3$)	Surface ($Å^2$)	Drug Score
	519.51	592.53	0.89
	631.94	730.88	0.88
	382.81	492.8	0.79
	334.46	604.92	0.64
	249.12	396.74	0.62

the binding sites with the highest druggability score is not necessarily the one with the largest volume or surface area–a binding site that is too large will not adequately mediate protein-ligand reactions. In addition, the binding sites with the highest drug score were located near the MD-2 binding site, making our results consistent with how proteins actually bind. Finally, these binding site predictions were similar to other studies conducted on the TLR4-MD-2 complex, further confirming their validity.

Energy Method

Another critical factor to consider when searching for possible LBSs is the chemical property of the protein. Energy-based methods of searching for LBSs consider potential interactions between the protein and the macromolecule such as hydrogen bonding and electrostatic interactions. These tools use molecule probes, groups of atoms and molecules to study molecular structures[38], to predict the interaction energy between probes at specific points and the protein. One of the most accurate computational tools of this kind is FTSite, which uses 16 molecular probes to search for potential binding sites. For each molecule probe, the software calculates and generates clusters with the most favorable interactions, which are then ranked using complex formulas[39]. Each structural map from the 16 probes are layered on top of one another, and areas with many clusters from different probes are considered as potential LBSs. These potential binding sites are ranked depending on how many different clusters they contain. In figure 2.2, for example, four clusters from the four different molecular probes overlap near one another, creating a promising LBS that has favorable interactions with molecules of different moiety.

In order to locate potential binding sites on the TLR4–MD-2 complex, we submit the required information to the FTSite queue. The PDB ID will be 2Z64, and since we only want LBSs on the surface of the TLR4 protein, we enter “A” under PDB Chain IDs. We obtain the following result:

The three best binding sites based on the energy method are very close to one another. Though the MD-2 protein is not shown in the 3D model, it would be located extremely close to the binding sites. Thus, both the binding sites found by the geometric method and the energy method are in similar regions, indicating that binding sites next to the TLR4–MD-2 binding area will be the most promising. Amino acid residues within 5Å are also shown using a ball and stick representation. This includes amino

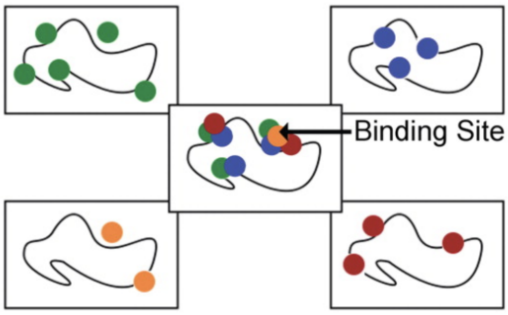


Fig. 3 Cluster maps from four molecular probes being combined to find a promising binding site

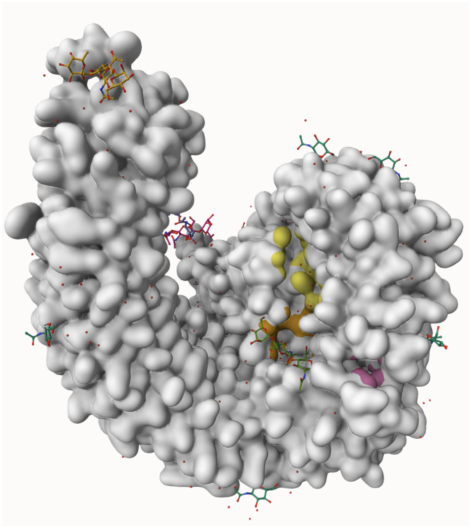


Fig. 5 3D structure of PrankWeb3 potential binding sites. TLR4 and MD-2 are both represented as white clusters, while potential binding sites are colored with various colors.

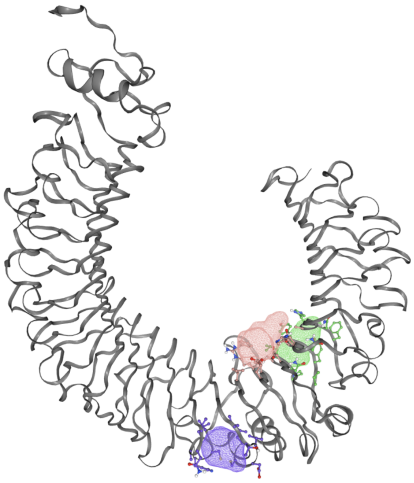


Fig. 4 FTSite 3D model result. TLR4 is shown in gray amino acid chains and potential LBSs are shown in colored clumps.

acids such as phenylamine, leucine, and glutamine.

Machine Learning Method

Previous methods for identifying potential LBSs have been categorized based on the specific properties they assess. However, many modern tools instead use a combination of different methods, taking into account many factors that could impact how promising a binding site is. One of these modern tools is PrankWeb 3, a machine learning ligand-binding site prediction tool that combines many approaches to evaluating protein binding sites. The backend computational tool of PrankWeb3 is P2Rank, whose method is used by PrankWeb to carry out analysis. The software generates a set of spaced points which is overlaid onto the protein surface. Then, using a machine learning model, it scores each point many times, each analyzing different properties. The most promising points of each analysis are overlaid, to create potential pockets. To use PrankWeb3, we enter the PDB code 2Z64. Since we only wish to look for binding sites on the TLR4 protein, we deselect "Use original structure" and deselect chain C. After submitting, we get the following results:

Table 2: Potential LBSs on the surface of TLR4 from PrankWeb3 sorted by Score.

Color Represented in Model	Score	Number of Residues	Volume (Å3)
	8.87	32	1117.4
	4.59	21	915.6
	1.72	15	356.8

The P2Rank software evaluates each binding site and gives it a score using complex formulas and equations utilizing the

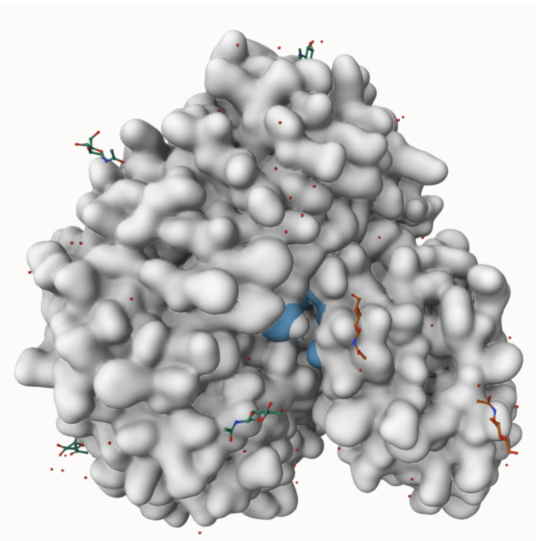


Fig. 6 Alternate view of PrankWeb3 results. The TLR4 protein is on the left, while the smaller MD-2 protein is on the right.

different properties it assessed. The number of residues refers to the number of amino acid residues the binding site borders. We found that the binding sites with more neighboring amino acid residues generally have a higher overall score. This is because the number of residues influences binding affinity–the strength of the interaction between the ligand and the binding site. Specific residues can create non-covalent interactions with the ligand, such as serine and asparagine forming hydrogen bonds or lysine forming ionic bonds[40]. In addition, once again, the most promising LBSs appear near the binding site between the TLR4 and MD-2 molecules. PrankWeb can perform additional tasks, such as finding the volume of pockets it identified. By selecting "Volume" and the rank of the pocket, we can create a task that calculates the volume of a specific binding site for additional research.

Pharmacophore Maps

All compounds have a specific and distinct set of features that dictate how they interact with biological targets like proteins and receptors. These features are a combination of chemical, structural, and physical attributes of a molecular structure[41]. A pharmacophore map is a 3D representation of these key features that are essential for a drug to bind to its target. Chemical features include hydrogen acceptors, hydrogen donors, anions, cations, hydrophobics, and aromatics, which are listed in detail in Table 3.3. Structural features depend on the locations of each pharmacophore class and the distances between classes and physical attributes change based on how many of each class there are in a given map. A pharmacophore map is made up of these features and is critical to finding compounds with specific features that can be used for receptor binding. By grouping functional groups with similar properties, such as hydrophobicity and charge, into pharmacophore classes, we can quickly screen for compounds that fit the constraints of a given map. Since each pharmacophore map has these specific features, computational tools can generate different maps, which can then be used to find unique compounds that can be used in ligand binding. This approach allows us to efficiently search through large databases of compounds to find those that might bind to our target of interest, significantly speeding up the drug discovery process

Table 3: Different pharmacophore classes and their representation in ZINCPharmer.

Image Representation	**Pharmacophore Class**	**Image Representation**	**Pharmacophore Class**
	Hydrogen Acceptor		Positive Ion
	Hydrogen Donor		Hydrophobic
	Negative Ion		Aromatic

Generating Pharmacophore Maps

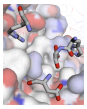
The computational tool used to generate pharmacophore maps based on a given 3D model of a protein is PocketQuery. By using a database of molecular analyses of proteins found in the Protein Data Bank, PocketQuery is able to quickly and efficiently create promising maps by adding atoms to improve bonding and interactions. After generating thousands of pharmacophore maps, it evaluates each of them based on a few properties, which include change of free energy and surface area. Critically, it provides a consensus score for each map using a machine learning model that looks for similarities between the structures of different pharmacophore maps to generate clusters of classes. The score is a value between 0 and 1, with a higher score indicating a better map. To generate this list of pharmacophore maps, we enter the PDB ID of 2Z64 into PocketQuery. We will add a criteria by clicking "Add Search Criteria", selecting "Chain", and entering "C". After clicking "Search", we obtain the following results:

Next, we can then export these pharmacophore maps to ZINCPharmer for further analysis and look for compounds that fit the given criterion.

Analyzing Pharmacophore Maps

In any given pharmacophore map, there could be millions of compounds that satisfy the given structural and chemical restrictions. Thus, we need to use software to search for them as well. ZINCPharmer is an interface to search for compounds

Table 4: Pharmacophore Maps binding to TLR4 from PocketQuery sorted by Score.

Map ID	Chain	Residues	Score	Figure
1	C	Histidine – #96 Histidine – #98 Aspartic Acid – #99 Aspartic Acid – #101	0.811032	
2	C	Histidine – #98 Aspartic Acid – #99 Aspartic Acid – #106 Arginine – #106	0.804953	
3	C	Histidine – #98 Aspartic Acid – #101	0.802837	
4	C	Histidine – #98 Aspartic Acid – #99 Aspartic Acid – #101	0.801867	
5	C	Histidine – #96 Histidine – #98 Aspartic Acid – #101	0.796931	
6	C	Histidine – #96 Aspartic Acid – #99 Aspartic Acid – #101	0.796187	
7	C	Histidine – #96 Aspartic Acid – #101	0.789858	
8	C	Aspartic Acid – #99 Aspartic Acid – #101	0.782479	

that match pharmacophore maps. By searching over 200 million conformations of 21 million compounds, ZINCPharmer is able to find the compounds that fit a given map the best. For every compound, ZINCPharmer calculates the RMSD, or root mean square deviation, a value used to determine how closely the compound matches with the pharmacophore map. The lower the RMSD, the closer a compound fits the constraints of a map. We can directly export the PocketQuery map to ZINCPharmer by going to the "Export" tab and clicking "Send to ZINCPharmer". Then, we can enable and disable different pharmacophore classes to query and find compounds with low RMSDs. We obtain the following results:

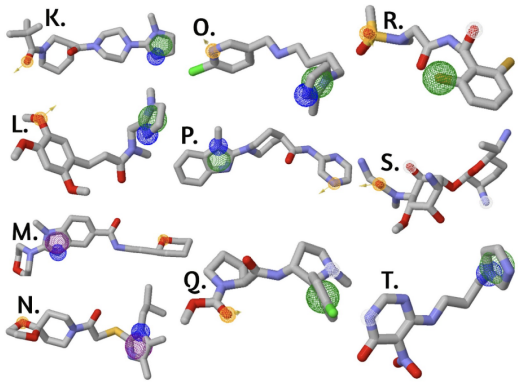


Fig. 7 10 more compounds with the lowest RMSD binding with TLR4 from different pharmacophore maps.

Table 5: 20 compounds with the lowest RMSD. The labels corresponding to the structure found in figures 3.6 and 3.7 are listed, along with the pharmacophore map ID, found in table 3.4, corresponding to the map that was used to find the compound.

Name	Compound Label	RMSD	Molecular Mass (Da)	Pharmacophore Map ID
ZINC01890537	A	0.001	334	4
ZINC14733825	B	0.002	459	1
ZINC14537809	C	0.002	334	1
ZINC19112940	D	0.002	441	1
ZINC14746137	E	0.002	364	1
ZINC14749573	F	0.002	437	1
ZINC86792785	G	0.002	220	2
ZINC79771686	H	0.002	350	3
ZINC58366224	I	0.002	334	6
ZINC91728012	J	0.002	328	6
ZINC84140047	K	0.002	362	6
ZINC61513047	L	0.002	346	6
ZINC95383080	M	0.003	373	2
ZINC80091168	N	0.003	369	2
ZINC83472120	O	0.003	281	3
ZINC76373299	P	0.003	337	3
ZINC72347649	Q	0.003	367	7
ZINC77634225	R	0.003	322	7
ZINC15657705	S	0.003	410	8
ZINC40659872	T	0.003	265	6

The 20 compounds with the lowest RMSD all have scores

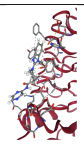
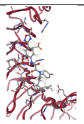
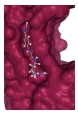
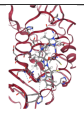
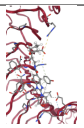
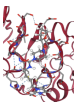
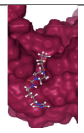
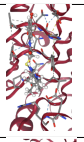
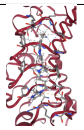
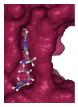
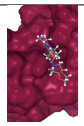
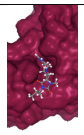
of 0.003 or less. This indicates that they almost perfectly match their corresponding pharmacophore maps, which are the ones that create the most suitable ligand-binding conditions. Compounds that are generated using the same map share many structural similarities other than the pharmacophore class, such as compounds A through E. In addition, some maps will be extremely similar, thus finding the same compounds that match with a low RMSD. The pharmacophore maps with IDs 2 and 5 were nearly identical, and the compounds found using map 2 were the same as those found using map 5. For this reason, map 5 is not listed as the map ID for any compound. Many of these similar top pharmacophore maps included hydrophobic and positive ion pharmacophore classes–for this reason, compounds derived from these maps all shared similar components. Notably, 16 of the compounds contained a positive ion, and 15 contained a hydrophobic class. Additionally, 16 of the compounds contained benzene rings, another common structural feature.

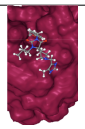
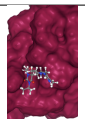
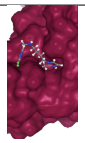
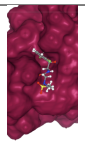
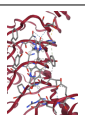
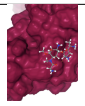
Molecular Docking

Molecular docking is a computational method that analyses the orientations of protein-ligand interactions and predicts the most favorable conformation to form stable complexes[42]. By generating a number of possible positions and ranking them based on scoring functions, molecular docking tools can evaluate the binding affinity of a specific compound with a protein. The scoring system is based on the change in Gibbs free energy, a thermodynamic measurement of the potential work done by a reaction. A negative change in Gibbs free energy indicates a spontaneous reaction that occurs without external energy. One of the most accurate softwares is SwissDock, a protein-ligand docking server that is able to predict these specific interactions. To input a compound, we input the Simple Molecular Input Line Entry System, or SMILES, which is obtained from the ZINC12 database. Then, by providing a target protein, in this case, TLR4, defining a search space, in this case, a box center of -18, -12, and -10, and a box size of 72, 38, and 32, it will give a list of the best clusters ranked by the change in Gibbs free energy. We obtain the following results:

These results are able to show which compounds will have the most energetic interactions with the TLR4 protein upon binding. While the results from section 3.8. are mere predictions of how good the interaction will be, the outcomes of the tests run through SwissDock show the actual binding affinity between the compound and TLR4. Although the most promising compound from section 3.8. had a lower RMSD than the other compounds, it didn't have the best actual interaction. Similarly, the best compound found by SwissDock based on change in Gibbs free energy did not have the lowest RMSD–ZINCPharmer gives predictions that may lead to the best results, but the actual compounds that mediate the most energetic reactions may not be the best predicted. This also emphasizes the importance of further testing–predictions made by screening tools are not always accurate. Thus physical experimentation in labs is needed to confirm the results reached.

Table 6: The best cluster member binding to TLR4 ranked by change in Gibbs free energy for the top 20 compounds identified in step 2.6.

Compound Label	Change in Gibbs Free Energy (kcal/mol)	Binding Diagram for the compound/protein interaction	Binding Diagram with Protein Surface
Q	-7.7162		
P	-7.6984		
F	-7.4735		
M	-7.4094		
B	-7.3778		
H	-7.2416		
I	-7.2128		
J	-7.1314		
N	-7.0838		
C	-7.0247		

Compound Label	Change in Gibbs Free Energy (kcal/mol)	Binding Diagram for the compound/protein interaction	Binding Diagram with Protein Surface
K	-6.9795		
D	-6.8994		
E	-6.8843		
O	-6.8807		
R	-6.8611		
L	-6.7557		
T	-6.7556		
G	-6.6725		
A	-6.46		
S	-6.1228		

In addition to calculating the change in Gibbs free energy, SwissDock gives the number of members in a cluster of docking locations, as well as their respective scores. The cluster sizes range from 1 member to 9 members, with the best compound having 8 members, all of which have a change in Gibbs free energy of -7.35 kcal/mol or less. There is no clear correlation between the number of members in a cluster and the overall druggability of a specific compound.

Lipinski's Rule of 5

Even if a compound has a great interaction with the target protein, it needs to reach the protein first in order to mediate a reaction. The property of compounds that measures how well they can enter the body is called absorption, and their ability to reach target compounds is known as permeability. Absorption is the ability to be absorbed into the bloodstream, and permeability is the ability to pass across biological membranes. Lipinski's "Rule of 5" is a computational method through which the solubility, membrane permeability, and efficacy are estimated. It outlines the specific restrictions that make a drug have good absorption and permeability–no more than five hydrogen bond donors, no more than 10 hydrogen bond acceptors, a calculated LogP of less than 5, and a molecular mass of less than 500 Da[43]. LogP is the partition coefficient, which is used to measure the balance between aqueous solubility and lipophilicity. Other factors of a druggable compound include gastrointestinal absorption, the ability to be absorbed in the GI tract[44], and blood-brain barrier permeation, the ability to cross the membrane between the blood and brain[45]. Existing FDA-approved drugs such as Donepezil and Galantamine satisfy all 5 rules along with having blood-brain barrier permeation. The tool used to assess these properties of compounds is called SwissADME. We add the list of SMILES strings of the 20 compounds and click "Run." We narrow the results to the following three top compounds:

Table 7: The three compounds that satisfy Lipinski's Rule of 5 and have both GI absorption and BBB permeability properties.

Compound Label	Number of Hydrogen Bond Donors	Number of Hydrogen Bond Acceptors	Molecular Weight (Da)	GI Absorption	BBB Permeant
J	2	2	328.46	Yes	Yes
L	1	4	346.4	Yes	Yes
O	2	1	280.8	Yes	Yes

Once again, the top compounds based on absorption and permeability turned out not to be the top compounds based on molecular docking computation. Out of the 20 compounds advanced to this stage, only one, compound S, didn't satisfy Lipinski's Rule of 5. While the other 19 did have good absorption and permeability, only the listed three have the BBB permeant property, which is critical specifically for Alzheimer's Disease, as TLR4 is mainly found within the brain. Therefore, this property allows for easier access of the drug to the brain once it disperses in the bloodstream. These results are similar to existing FDA-approved drugs. This can also be tested through physical screening rather than simply relying on predictions.

Molecular Toxicology

Molecular toxicology is a field concerning biological response to specific compounds[46]. It is highly important in preclinical drug discovery as even if a compound will bond well with the

target protein, it is useless if the body responds poorly to it. Thus, testing the toxicity of a particular compound is needed for it to move on to physical screening. The tool used to evaluate the toxicity of a given compound is called ProTox. By using a database of 38,000 unique compounds with known toxicity profiles and methods in analyzing the compound to find toxic fragments, ProTox is able to accurately predict the toxicity of a compound with a number of different values[47]. The first important value is the predicted LD50, or Lethal Dose 50. The LD50 value is the amount of the substance that is lethal for 50% of experimental units exposed to it[48]. It is measured in mg/kg, or the amount of substance per kilogram of body weight. The next important value is the predicted toxicity class, a metric to determine how dangerous the compound is for the human body. A toxicity class of 1 or 2 means the compound is fatal if swallowed; class 3 means toxic if swallowed; class 4 and 5 means harmful if swallowed; and class 6 means it is non-toxic. The average similarity is a percent given that represents how similar the input compound is to other compounds that bind at the particular target. Finally, the accuracy of the prediction is also calculated and given as a percent. ProTox also provides a toxicity radar chart, which compares the different probabilities of active toxicities a compound has with those of different toxicity models active within the training set. The results for the top compound, ZINC91728012, are shown below:

Table 8: Results from ProTox for the compound ZINC91728012.

Property	Value
Predicted LD50	500 mg/kg
Predicted Toxicity Class	4
Average Similarity	46.89%
Prediction Accuracy	54.26%

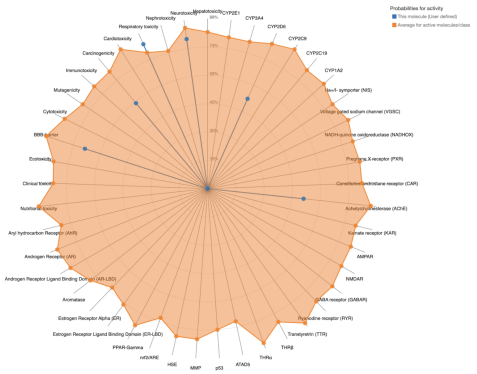


Fig. 8 Toxicity Radar Chart from ProTox for the compound ZINC91728012. The average probabilities of toxicity from active molecules for each type of toxicity are shown in orange, while the probabilities of toxicity for the given compound are shown in blue.

The compound ZINC91728012, according to the results, will be harmful if swallowed. However, a toxicity class of 4 is within the acceptable range, and the predicted LD50 is well above the lethal threshold. At an LD50 of 500 mg/kg, the compound is only moderately toxic, and is similar to many commonly prescribed drugs, such as Ibuprofen at 636 mg/kg and Aspirin at 200 mg/kg. Of the six types of toxicity detected, four have probabilities well below similar drugs in the database. Neurotoxicity is slightly lower while respiratory toxicity is slightly over by 5%, which is still within the acceptable range. We can conclude from these results that the top compound, ZINC91728012, is relatively safe to consume.

Advantages

While it is difficult to say specifically what advantages there are due to the intricacies of pharmacology, there are properties of the identified top compound that have certain benefits over existing FDA-approved drugs. For comparison, contrasts will be made against the most common Alzheimer's drugs–Donepezil and Galantamine–which have ZINC IDs of ZINC897251 and ZINC491073, respectively. The main advantage of the compound identified is its toxicity level–as mentioned in section 3.11, it has a toxicity class of 4, meaning it's only slightly toxic. However, compared to Galantamine, which has a toxicity class of 3, the compound is much less toxic.

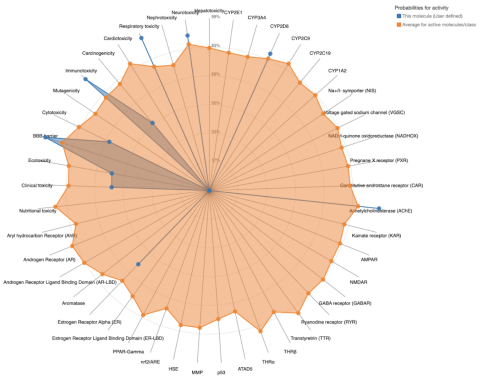


Fig. 9 Toxicity Radar Chart from ProTox for Donepezil.

Furthermore, when comparing the toxicity charts, more information is revealed. In the toxicity groups of BBB-barrier, immunotoxicity, and respiratory toxicity, Galantamine exceeds the average toxicity level of active drugs. Additionally, Donepezil is more toxic than industry standards in 6 different categories. On the other hand, the identified compound in this study only exceeds the average in respiratory toxicity–it also has lower neurotoxicity, toxicity that affects the nervous system, than Galantamine. While it is not harmful to utilize drugs with higher toxicities, being less toxic has a few advantages. Firstly,

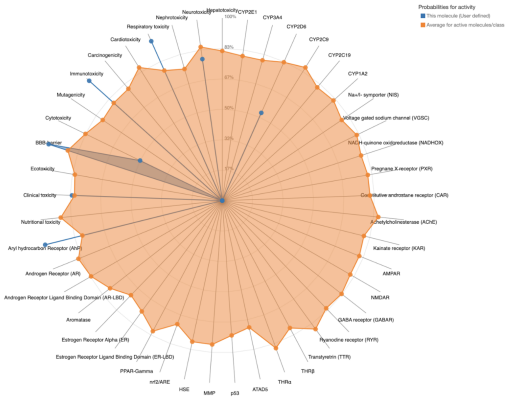


Fig. 10 Toxicity Radar Chart from ProTox for Galantamine.

there is a wider margin of safety in the case that people consume too much of the drug. Additionally, drugs with lower toxicity generally have fewer unwanted side effects, meaning that the new compound may have fewer ramifications.

Conclusion

The steps taken to identify compounds with high binding affinity with TLR4, a compound allowing signaling pathways to activate neuroinflammatory responses associated with Alzheimer's Disease, were successful. Further analysis of how the compound interacts with the human body in factors such as absorption, permeability, and toxicity, showed that the top compounds are promising candidates for physical screening. Ultimately, the research comes together to progress cures for Alzheimer's.

Acknowledgements

I would like to thank Dr. Moustafa Gabr from Cornell University for the guidance in the development of this research paper.

References

1 Z. R. Chen, J. B. Huang, S. L. Yang and F. F. Hong, *Role of cholinergic signaling in Alzheimer's Disease*, 2022.

2 L. M. Osborn, W. Kamphuis, W. J. Wadman and E. M. Hol, *Astrogliosis: an integral player in the pathogenesis of Alzheimer's Disease*, 2016.

3 I. Morales, C. Cerda-Troncoso, V. Andrade and R. B. Maccioni, *The natural product curcumin as a potential coadjuvant in Alzheimer's treatment*, 2016.

4 Y. Cai, Y. Chai, Y. Fu, Y. Wang, Y. Zhang, X. Zhang, L. Zhu, M. Miao and T. Yan, *Salisdroside ameliorates Alzheimer's disease by targeting NLRP3 inflammasome-mediated pyroptosis*, 2022.

5 V. I. Tsetlin, *Acetylcholine and Acetylcholine receptors: textbook knowledge and new data*, 2020.

6 E. K. Perry, B. E. Tomlinson, G. Blessed, K. Bergmann, P. H. Gibson and R. H. Perry, *Correlation of cholinergic abnormalities with senile plaques and mental test scores in senile dementia*, 1978.

7 M. B. Colović, D. Z. Krstić, T. D. Lazarević-Pašti, A. M. Bondžić and V. M. Vasić, *Acetylcholinesterase inhibitors: pharmacology and toxicology*, 2013.

8 R. J. O'Brien and P. C. Wong, *Amyloid precursor protein processing and Alzheimer's disease*, 2011.

9 G. Pfundstein, A. G. Nikonenko and V. Sytnyk, *Amyloid precursor protein (APP) and amyloid (A) interact with cell adhesion molecules: implications in Alzheimer's disease and normal physiology*, 2022.

10 C. Ma, F. Hong and S. Yang, *Amyloidosis in Alzheimer's disease: pathogeny, etiology, and related therapeutic directions*, 2022.

11 V. Rani, R. Verma, K. Kumar and R. Chawla, *Role of pro-inflammatory cytokines in Alzheimer's disease and neuroprotective effects of pegylated self-assembled nanoscaffolds*, 2022.

12 J. P. Konsman, *Cytokines in the brain and neuroinflammation*, 2022.

13 M. Colonna and O. Butovsky, *Microglia function in the central nervous system during health and neurodegeneration*, 2017.

14 Z. Breijyeh and R. Karaman, *Comprehensive review on Alzheimer's disease: causes and treatment*, 2020.

15 C. Sam and B. Bordoni, in *Physiology, acetylcholine*, StatPearls Publishing, Treasure Island (FL), 2024.

16 J. S. Birks and R. J. Harvey, *Donepezil for dementia due to Alzheimer's disease*, 2018.

17 R. Bullock, J. Touchon, H. Bergman, G. Gambina, Y. He, G. Rapatz and R. Lane, *Rivastigmine and donepezil treatment in moderate to moderately-severe Alzheimer's disease over a 2-year period*, 2005.

18 P. Masson and O. Lockridge, *Butyrylcholinesterase for protection from organophosphorus poisons: catalytic complexities and hysteretic behavior*, 2010.

19 C. B. da Silva, A. Pott, S. Elifio-Esposito, L. Dalarmi, K. F. do Nascimento, L. M. Burci, M. de Oliveira, J. de Fátima Gaspari Dias, Z. W. Zanin, O. G. Miguel and M. D. Miguel, *Effect of donepezil, tacrine, galantamine and rivastigmine on acetylcholinesterase inhibition in Dugesia tigrina*, 2016.

20 Y. H. Siddique, F. Naz and R. Varshney, *Comparative study of rivastigmine and galantamine on the transgenic Drosophila model of Alzheimer's disease*, 2022.

21 O. Takeuchi and S. Akira, *Pattern recognition receptors and inflammation*, 2010.

22 S. Akira, K. Takeda and T. Kaisho, *Toll-like receptors: critical proteins linking innate and acquired immunity*, 2001.

23 F. S. Lira, J. C. Rosa, G. D. Pimentel, V. A. F. Tarini, R. M. Arida, F. Faloppa, E. S. Alves, C. O. do Nascimento, L. M. Oyama, M. Seelaender, M. T. de Mello and R. V. T. Santos, *Toll-like receptors: critical proteins linking innate and acquired immunity*, 2001.

24 J. S. Roh and D. H. Sohn, *Damage-associated molecular patterns in inflammatory diseases*, 2018.

25 J. Oliveira and W. C. Reygaert, in *Gram-negative bacteria*, StatPearls Publishing, Treasure Island (FL), 2024, p. N/A.

26 A. Ciesielska, M. Matyjek and K. Kwiatkowska, *TLR4 and CD14 trafficking and its influence on LPS-induced pro-inflammatory signaling*, 2021.

27 L. B. Ivashkiv and L. T. Donlin, *Regulation of type I interferon responses*, 2014.

28 S. Liu, Y. Liu, W. Hao, L. Wolf, A. J. Kiliaan, B. Penke, C. E. Rube, J. Walter, M. T. Heneka, T. Hartmann, M. D. Menger and K. Fassbender, *TLR2 is a primary receptor for Alzheimer's amyloid beta peptide to trigger neuroinflammatory activation*, 2009.

29 K. Tahara, H. D. Kim, J. J. Jin, J. A. Maxwell, L. Li and K. Fukuchi, *Role of toll-like receptor signalling in Abeta uptake and clearance*, 2006.

30 C. Y. Lee and G. E. Landreth, *The role of microglia in amyloid clearance from the AD brain*, 2010.

31 M. Go, J. Kou, J. E. Lim, J. Yang and K. I. Fukuchi, *Microglial response to LPS increases in wild-type mice during aging but diminishes in an Alzheimer's mouse model: implication of TLR4 signaling in disease progression*, 2016.

32 A. LaPelusa and R. Kaushik, in *Physiology, proteins*, StatPearls Publishing, Treasure Island (FL), 2024, p. N/A.

33 B. Alberts, A. Johnson, J. Lewis and et al., *Molecular biology of the cell*, 4th edn, 2002, p. N/A.

34 M. J. Lopez and S. S. Mohiuddin, *Biochemistry, essential amino acids*, 2024.

35 G. Wu, *Amino acids: metabolism, functions, and nutrition*, 2009.

36 R. Shimazu, S. Akashi, H. Ogata, Y. Nagai, K. Fukudome, K. Miyake and M. Kimoto, *MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4*, 1999.

37 M. Michel, T. Visnes, E. J. Homan, B. Seashore-Ludlow, M. Hedenström, E. Wiita, K. Vallin, C. B. J. Paulin, J. Zhang, O. Wallner, M. Scobie, A. Schmidt, A. Jenmalm-Jensen, U. W. Berglund and T. Helleday, *Computational and Experimental Druggability Assessment of Human DNA Glycosylases*, 2019.

38 H. Lesch, J. Schlichter, J. Friedrich and J. M. Vanderkooi, *Molecular probes: what is the range of their interaction with the environment?*, 2004.

39 C. H. Ngan, D. R. Hall, B. Zerbe, L. E. Grove, D. Kozakov and S. Vajda, *FTSite: high accuracy detection of ligand binding sites on unbound protein structures*, 2012.

40 W. Y. Wan and E. J. Milner-White, *A recurring two-hydrogen-bond motif incorporating a serine or threonine residue is found both at -helical N termini and in other situations*, 1999.

41 S. K. Jain and A. Chincholikar, *Pharmacophore Mapping and Drug Design*, 2004.

42 P. H. M. Torres, A. C. R. Sodero, P. Jofily and F. P. Silva-Jr, *Key Topics in Molecular Docking for Drug Design*, 2019.

43 R. Roskoski, *Properties of FDA-approved small molecule protein kinase inhibitors*, 2019.

44 G. D. Anderson and R. P. Saneto, *Current oral and non-oral routes of antiepileptic drug delivery*, 2012.

45 A. K. Dotiwala, C. McCausland and N. S. Samra, in *Anatomy, Head and Neck: Blood Brain Barrier*, StatPearls Publishing, Treasure Island (FL), 2024.

46 K. S. Ramos, in *Introduction to Molecular Toxicology*, ed. C. A. McQueen, 2010, vol. 2.

47 M. N. Drwal, P. Banerjee, M. Dunkel, M. R. Wettig and R. Preissner, *ProTox: a web server for the in silico prediction of rodent oral toxicity*, 2014.

48 T. R. Lane, J. Harris, F. Urbina and S. Ekins, *Comparing LD50/LC50 Machine Learning Models for Multiple Species*, 2023.