POLITECNICO DI TORINO

Master’s Degree in Biomedical Engineering

Insights into Alpha-Synuclein/S100B Interactions in Parkinson’s Disease by Molecular Modelling

Thesis Supervisor

Prof. Marco Agostino Deriu

Co-supervisor

Marcello Miceli, PhD

Candidate

Greta Maria Grieco Nobile

ACADEMIC YEAR 2024-2025

Table of contents

Abstract

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor dysfunction and cognitive decline, primarily attributed to the abnormal accumulation of alpha-synuclein (α-syn). This protein tends to misfold and aggregate, forming Lewy bodies, which are pathological hallmarks of PD. The aggregation of α-syn disrupts neuronal function and viability, leading to the clinical manifestations of the disease. Previous studies have explored interactions between amyloidogenic proteins and S100 family members, suggesting the involvement of S100 proteins, a family of calcium-binding proteins, in neurodegenerative processes. Notably, interactions between S100A9 and α-synuclein have been documented, demonstrating that S100A9 influences the aggregation kinetics and fibril structure of α-syn, potentially worsening neurodegeneration. Similarly, the interaction between S100B and β-amyloid has been explored, revealing that S100B can inhibit β-amyloid aggregation, indicating a modulatory role in Alzheimer's disease. In the context of Parkinson's disease (PD), elevated levels of the S100B protein have been observed in the substantia nigra of patients, suggesting its potential involvement in the disease's development. However, the direct interaction between S100B and α-synuclein remains underexplored. This work aims to investigate the interaction between S100B and α-synuclein through computational modeling and molecular dynamics simulations. Since no experimentally determined structure of the complex was available in the literature, its structure was predicted using two different approaches. The first method employed AlphaFold, an AI-based structure prediction tool, while the second approach utilized HADDOCK for molecular docking. In the second case, due to the lack of experimental data regarding the active residues of the two proteins, these were selected based on similarity hypotheses with other protein interactions and insights from previous studies. Molecular dynamics simulations were performed on both predicted models to explore the stability of the predicted complexes and assess potential binding regions. In conclusion, this study provides a computational characterization of the S100B and α-synuclein interaction offering a foundation for further experimental investigations into its potential biological significance. Understanding this interaction could be relevant for determining whether S100B influences α-syn aggregation and toxicity, providing new insights into its role in neurodegenerative processes and its possible use as a therapeutic target.

1 Introduction

Neurodegenerative diseases represent a significant global health challenge, with Parkinson's disease (PD) being the second most prevalent, affecting over 10 million people worldwide. It is characterized by the progressive degeneration of dopaminergic neurons in the substantia nigra, a key brain region involved in motor control, leading to motor dysfunction and cognitive decline. Parkinson's disease is primarily attributed to the abnormal aggregation of α-synuclein (α-syn), which tends to misfold and accumulate into Lewy bodies, the pathological hallmarks of PD. Understanding the molecular mechanisms that regulate α-syn aggregation is crucial for identifying potential therapeutic strategies.

Previous studies have explored the interactions between amyloidogenic proteins and members of the S100 protein family, a group of calcium-binding proteins involved in numerous cellular processes, including neuroinflammation and neurodegeneration. Notably, it has been demonstrated that S100A9 interacts with α-synuclein and influences its aggregation kinetics, accelerating aggregation and promoting the formation of fibrillar α-syn structures. Similarly, studies on Alzheimer's disease have shown that S100B interacts with β-amyloid (Aβ) and can inhibit its aggregation, suggesting a potential regulatory role. Interestingly, elevated S100B levels have been detected in the substantia nigra of Parkinson's disease patients, indicating a possible link to α-syn pathology. However, the direct interaction between S100B and α-synuclein remains largely unexplored, leaving a significant gap in understanding its role in PD progression.

This work is part of the Unite! Seed Fund project titled "Multiscale investigation of the stabilization of neurodegeneration-linked amyloid proteins by metal ions and chaperones", a collaboration between TU Darmstadt, ULisboa, and Politecnico di Torino. The project investigates how metal ions and molecular chaperones influence the aggregation of neurodegeneration-related proteins, aiming to identify potential therapeutic targets. By integrating in vitro and in silico approaches, researchers focus on S100 proteins and their role in modulating amyloid formation.

Within this context, the present study focuses on the computational characterization of the S100B/α-synuclein interaction, providing insights into its structural and dynamic properties. Since no experimentally determined structures of the S100B–α-synuclein complex are available in the current literature, two different methodologies are employed to predict the structure of the complex: AlphaFold2-Multimer, an artificial intelligence-based tool for protein structure prediction, and HADDOCK, a molecular docking program that generates models by simulating protein-protein interactions based on known interaction patterns and computational hypotheses. Once structural models are obtained, the most promising complexes are selected and subjected to molecular dynamics (MD) simulations. These simulations provide insight into the stability of the predicted structures, potential binding interfaces and the dynamic behavior of the interaction.

This work is organized as follows:

Chapter 1 is the present introduction.

Chapter 2 provides a biological background on neurodegenerative diseases, with a specific focus on Parkinson's disease, α-synuclein aggregation and S100 proteins. Furthermore, previous studies on the interaction between amyloidogenic proteins and S100 proteins are discussed.

Chapter 3 is a description of the methods used in this study. It includes an overview of molecular modeling principles, covering the fundamentals of Molecular Mechanics and Molecular Dynamics. Additionally, the computational approaches employed for structure prediction, such as HADDOCK for molecular docking and AlphaFold2-Multimer for AI-based structure prediction, are introduced.

Chapter 4 presents the methods and results of the structural prediction of the S100B–α-synuclein complex, outlining the computational steps taken to generate and validate the models.

Chapter 5 describes the methods and results of molecular dynamics simulations performed on the predicted S100B–α-synuclein complexes. It details the simulation setup, analysis techniques, and findings regarding the stability and interaction dynamics of the predicted models.

Chapter 6 summarizes the key findings of the work, discussing their implications and outlining potential future research directions.

2 Biological Background

2.1 An Introduction to neurodegenerative diseases

The term “neurodegenerative diseases” refers to a wide range of debilitating conditions characterized by the progressive degeneration of neurons, leading to cognitive, motor, and behavioral impairments. Unlike acute neurological conditions, which result from sudden insults or injuries to the nervous system, neurodegenerative disorders develop over extended periods, gradually eroding neuronal function and connectivity [1], [2].

The progressive dysfunction and loss of neurons are associated with deposition of proteins showing altered physicochemical properties, also known as misfolded proteins, in the brain and peripheral organs. These abnormal protein aggregates result in the formation of amyloid fibrils [3]. Currently, more than 30 known proteins are associated with these disorders. The most frequently involved proteins in the pathogenesis of neurodegenerative diseases are amyloid-β, prion protein, tau, α-synuclein, TAR-DNA-binding protein 43 kDa, and fused-in sarcoma protein [4].

Clinical manifestations of neurodegenerative diseases vary widely, reflecting the heterogeneous nature of the pathologies and the diverse regions of the brain affected. Alzheimer's disease, the most prevalent form of dementia, is typified by memory deficits and the accumulation of amyloid-beta plaques and neurofibrillary tangles. Parkinson's disease manifests through motor symptoms like tremors and rigidity, accompanied by the aggregation of alpha-synuclein in Lewy bodies [5]. Other notable disorders include frontotemporal dementia, amyotrophic lateral sclerosis, and Huntington's disease, each with its distinct pathological hallmarks and clinical trajectories.

The global burden of neurodegenerative diseases is significant and steadily increasing, with approximately 50 million individuals affected worldwide, a number expected to rise to 152 million by 2060 [6]. Their profound impact on individuals and the healthcare system makes the study of amyloid protein aggregation very important. An incomplete knowledge about the structures of amyloid aggregation intermediates and pathways to target is the key factor limiting drug discovery against amyloid-related diseases [7].

2.2 *Parkinson's disease*

Parkinson's disease (PD) is a progressive neurological disorder characterized by a wide number of motor and non-motor features. Symptoms such as tremor at rest, bradykinesia, muscle rigidity, and postural instability are its key clinical signs. PD predominantly targets the nigrostriatal pathway, which consists of dopamine-producing neurons in the substantia nigra pars compacta (SNpc) that project to the putamen [8]. The loss of these neurons and the resulting damage to the striatal connections are hallmark features of this disease. However, many patients also experience non-motor symptoms, including gastrointestinal issues, autonomic dysfunction, and cognitive decline [9]. Like other neurodegenerative diseases, PD involves the abnormal accumulation of proteins within cells, specifically Lewy bodies (Figure 1), which are primarily composed of α-synuclein. Consequently, PD is recognized as the most common synucleinopathy and is also classified as a cerebral amyloid disorder.

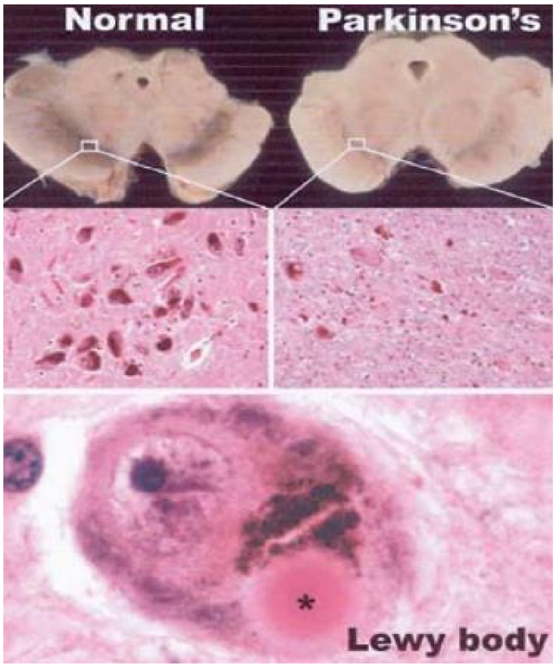


Figure 1: Comparison between the substantia nigra in a normal midbrain and in the midbrain of a patient with Parkinson disease. In a normal midbrain, the substantia nigra is darkly pigmented. In Parkinson's disease, it appears pale because of dopaminergic neuron loss, with some remaining neurons showing Lewy body inclusions [10] .

2.2.1 *Main clinical features*

The four main symptoms of Parkinson's disease (PD) can be summarized using the acronym TRAP (Tremor at rest, Rigidity, Akinesia and Postural instability) [11].

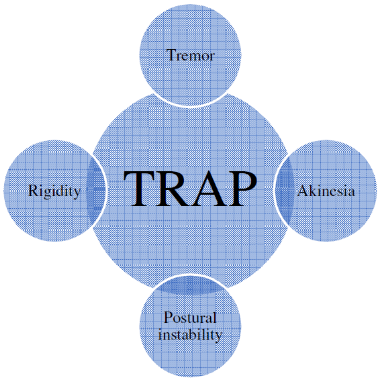


Figure 2: Motor symptoms of Parkinson’s disease [12].

- *Tremor at rest*

  Resting tremor is the most frequent and easily identifiable sign of Parkinson's disease (PD). This tremor typically occurs on one side of the body and has a frequency of 4–6 Hz. It is prominent in the distal parts of an extremity and, in PD patients, affects the hands, lips, chin, jaw, and legs. Unlike essential tremor, it seldom affects the neck, head, or voice. A defining characteristic of resting tremor is that it disappears with action and during sleep.

- *Rigidity*

  Although rigidity is not usually the main symptom early in Parkinson’s disease, it is a defining feature of PD. This condition is experienced as a stiffness of the arms or legs, which can occur on one or both sides of the body and contributes to a decreased range of motion.

- *Akinesia*

  Akinesia (or bradykinesia), defined as slowness of movement, represents the most prominent clinical feature of Parkinson’s disease (PD). As a hallmark of disorders affecting the basal ganglia, bradykinesia involves impairments in the ability to plan, initiate, and execute movements, as well as difficulties in performing sequential and simultaneous tasks. Bradykinesia often initially manifests as a reduction in speed when carrying out daily activities, accompanied by slower movements and prolonged reaction times.

- *Postural instability*:

  Postural instability, resulting from the loss of postural reflexes, can be present even at diagnosis but becomes more prevalent and worsens with disease progression.

  The pull test involves rapidly pulling patients backward or forward by the shoulders to assess their degree of retropulsion or propulsion. If the patient takes more than two steps backward or fails to exhibit any postural response, it suggests an abnormality in their postural reflexes.

2.2.2 Clinical diagnosis

The diagnosis of Parkinson's disease has evolved significantly over the past few decades, driven by improvements in clinical understanding and diagnostic precision. A crucial step in this process was the development of the Queen Square Brain Bank (QSBB) criteria, which have become widely used in both clinical practice and research.

The QSBB criteria follow a structured three-step approach [13].

- *Step 1: Identification of Parkinsonian Syndrome.*

  The presence of bradykinesia - characterized by “slowness of initiation of voluntary movement, with a progressive reduction in speed and amplitude of repetitive actions” - is identified as an obligatory criterion for the syndrome.
- *Step 2*: *Exclusion Criteria*

  Potential alternative diagnoses are excluded based on a patient's history, clinical findings, and disease progression.
- *Step 3*: *Supportive Criteria*

  The diagnosis is confirmed by supportive criteria, such as unilateral onset with ongoing asymmetry, rest tremor, symptom progression, consistent levodopa response (>70%), levodopa-induced chorea, levodopa response for more than 5 years, long clinical course (>10 years).

In 2015, the International Parkinson and Movement Disorder Society (MDS) introduced a revised set of criteria to further improve the accuracy of PD diagnosis and enhance the identification of early-stage cases [13]. The new MDS criteria maintained the fundamental diagnostic principles of the QSBB framework but introduced two different diagnostic categories:

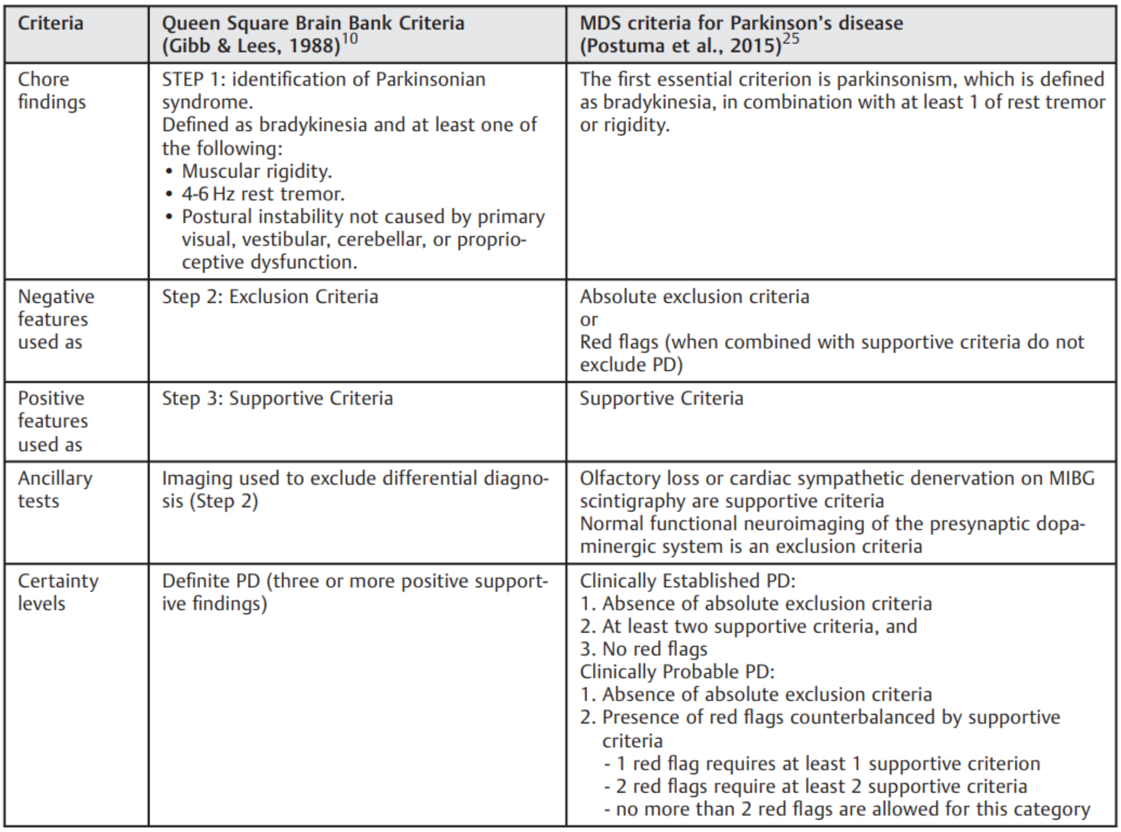
- *Clinically Established PD*, which emphasizes high specificity

- *Clinically Probable PD*, which balances sensitivity and specificity.

This differentiation allows clinicians to capture a broader spectrum of PD cases while minimizing the risk of false positives. Additionally, the MDS criteria introduced derivative standards for early-stage PD, particularly useful in clinical trials where early and accurate diagnosis is critical. The comparison between the QSBB and the new MDS criteria is shown in the **Error! Reference source not found.**.

Together, the QSBB and MDS criteria have set new standards for diagnosing Parkinson's disease, providing clinicians with robust tools to improve diagnostic accuracy at different stages of the disease and ensuring more personalized and effective care for patients.

Table 1: Comparison of the QSBB and the new MDS criteria [13].



Criteria	Queen Square Brain Bank Criteria (Gibb & Lees, 1988)[10]	MDS criteria for Parkinson's disease (Postuma et al., 2015)[25]
Chore findings	STEP 1: identification of Parkinsonian syndrome. Defined as bradykinesia and at least one of the following: • Muscular rigidity. • 4-6 Hz rest tremor. • Postural instability not caused by primary visual, vestibular, cerebellar, or proprioceptive dysfunction.	The first essential criterion is parkinsonism, which is defined as bradykinesia, in combination with at least 1 of rest tremor or rigidity.
Negative features used as	Step 2: Exclusion Criteria	Absolute exclusion criteria or Red flags (when combined with supportive criteria do not exclude PD)
Positive features used as	Step 3: Supportive Criteria	Supportive Criteria
Ancillary tests	Imaging used to exclude differential diagnosis (Step 2)	Olfactory loss or cardiac sympathetic denervation on MIBG scintigraphy are supportive criteria Normal functional neuroimaging of the presynaptic dopaminergic system is an exclusion criteria
Certainty levels	Definite PD (three or more positive supportive findings)	Clinically Established PD: 1. Absence of absolute exclusion criteria 2. At least two supportive criteria, and 3. No red flags Clinically Probable PD: 1. Absence of absolute exclusion criteria 2. Presence of red flags counterbalanced by supportive criteria - 1 red flag requires at least 1 supportive criterion - 2 red flags require at least 2 supportive criteria - no more than 2 red flags are allowed for this category

2.2.3 Treatments for motor symptoms of Parkinson's disease

The treatment of Parkinson's disease should begin with a detailed discussion about the diagnosis, disease progression and the available treatment options, focusing on maintaining functional independence and quality of life. The type of therapy and the timing of treatment initiation depend on the specific symptoms presented by the patient, as well as his or her lifestyle and personal situation [14].

From the beginning, patients should be encouraged to maintain good hydration and follow a high-fiber diet, ideally a Mediterranean diet [15], to help prevent constipation. Moreover, aerobic exercise is recommended where appropriate, evaluating individual's fitness level and physical fragility, to reduce the risk of injury.

2.2.4 Initial treatment of motor Parkinson's disease

Both the American Academy of Neurology and the National Institute for Health and Care Excellence published an evidence-based guideline stating that levodopa should be the initial pharmacological treatment of Parkinson's disease, with the caveat that dopamine agonists might be preferable for first-line treatment in patients with additional risk factors for dyskinesia [16]. Levodopa has been shown effectively relieve motor symptoms with minimal side effects (nausea, daytime sleepiness or postural hypotension may occur in a small percentage of patients) and starting levodopa therapy at low doses (less than 400 mg/day) is not expected to accelerate the onset of long-term complications like dyskinesia. However, prolonged use of levodopa may result in elevated plasma homocysteine levels, which are a risk factor for stroke, heart disease and dementia [17]; this can be prevented by a regular oral supplementation with folic acid and vitamin B12.

2.2.5 Adjunctive treatment of motor Parkinson's disease

In some cases, patients may not experience adequate symptom relief despite starting treatment, especially those with tremor-dominant forms of the disease; an increase in levodopa dosage might be required. However, if rigidity or bradykinesia are refractory to levodopa despite good compliance, then it may suggest the presence of atypical parkinsonism rather than classic Parkinson's disease.

In patients who have a good initial response to levodopa, as Parkinson's disease progresses, bothersome motor symptoms can re-emerge. In such situations, it is necessary to decide whether to increase the levodopa dose (monotherapy) or introduce adjunctive drugs. Studies have shown that a combination of drugs at low or moderate doses can provide optimal symptom control and reduce side-effects, often proving more effective than high doses of a single dopaminergic medication.

The progression of Parkinson's disease can lead to the onset of the complex phase, characterized by various motor fluctuations, such as:

- *early morning off symptoms* (re-emergence of slowness, stiffness, tremor, or dystonia), which can occur when dopamine levels drop too low;
- *end-of-dose wearing off;*
- *dose failures;*
- *delayed on time.*

Trial data support the use of dopamine agonists,27 catechol-O-methyltransferase (COMT) inhibitors (including entacapone or more recently, the once-daily opicapone23,24), or any of the MAO-BIs15 in reducing off time in patients with Parkinson's disease with motor fluctuations [16].

Furthermore, the oral extended-release formulation of levodopa, known as IPX066, can reduce "off" time by approximately one hour each day in comparison to immediate-release levodopa; however, it necessitates a meticulous dose conversion approach.

2.2.6 Non-oral therapies for the complex phase of Parkinson's disease

Non-oral therapies for the complex phase of Parkinson's disease aim to manage "off" periods, when medications are less effective, and dyskinesias [16].

- *Apomorphine*
  Daily apomorphine infusions are administered to patients with a less predictable response to oral medications. Moreover, apomorphine injections are also used as an emergency treatment to quickly address problems such as nighttime immobility, delayed movement initiation and the failure of individual oral doses.
- *Levodopa–carbidopa intestinal gel*
  Levodopa-carbidopa intestinal gel, administered through a percutaneous intrajejunal tube, has been an effective treatment for severe motor fluctuations since its introduction in Europe in 2004.
- *Novel formulations of levodopa*
  Levodopa inhalers provide rapid symptom relief, while subcutaneous formulations allow continuous levodopa–carbidopa infusions without the need for a jejunal tube. Clinical data show improvements in "off" time and a reduction in dyskinesias.
- *Deep brain stimulation (DBS)*

Subthalamic nucleus deep brain stimulation (STN-DBS), introduced in 1995, is an effective treatment for patients with Parkinson’s disease who experience motor fluctuations. It improves akinesia and dyskinesias.

- *Ablative neurosurgery*
  Despite the increasing use of DBS, the creation of precise lesions through ablative surgery may be preferable in some cases, such as for the treatment of Parkinsonian tremors. Techniques like focused ultrasound allow for lesion creation without invasive procedures.

2.3 α-Synuclein protein

Alpha-synuclein is a small, intrinsically disordered protein that belongs to the synuclein gene family, which includes α-syn, β-synuclein and γ-synuclein [18]. While α-syn is primarily expressed in the brain, its highest concentrations are found in regions such as the neocortex, hippocampus, substantia nigra (SN), and thalamus, with lower levels in the cerebellum. It is also worth mentioning that α-syn expressions have been observed in non-neuronal cells, including red blood cells and various immune cells.

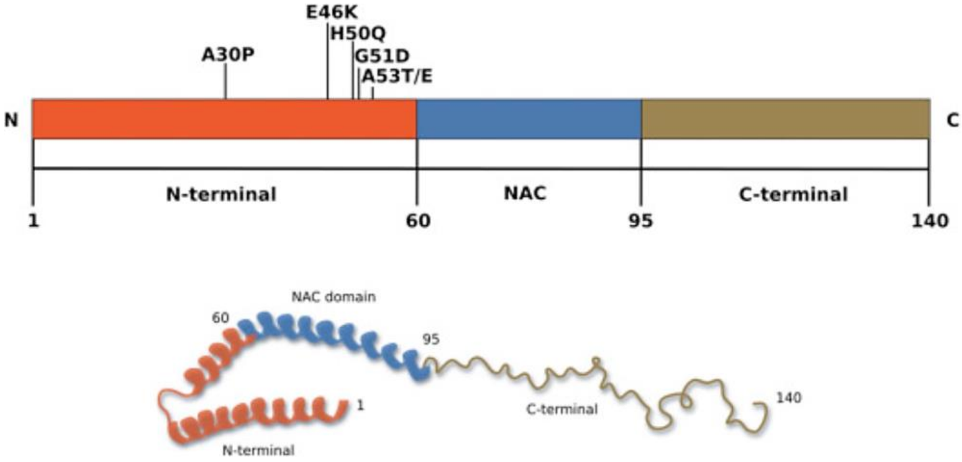


Figure 3: Schematic representation of a-synuclein, depicting the three main regions of the protein. Mutations associated with familial forms of PD are also indicated [19].

The α-synuclein protein consists of 140 amino acids and its primary structure can be divided into three distinct regions [7]:

- An **N-terminal region (amino acids 1-60)**, which adopts an α-helical conformation upon interacting with lipid membranes. This is the region where the familial mutations leading to early onset Parkinson’s disease reside.

- A **central non-amyloid component NAC (amino acids 61-95)**, which is prone to β-sheet formation and is critical in driving fibril assembly.
- A highly acidic **C-terminal region (amino acids 96-140)**, which modulates intramolecular interactions and binds divalent metal ions such as calcium.

Under normal conditions, α-synuclein contributes to the stabilization of synaptic vesicles and assists in neurotransmitter release, yet its dynamic, unstructured nature renders it susceptible to misfolding under certain environmental conditions, leading to the formation of pathological aggregates [18].

2.3.1 Alpha-synuclein aggregation and Lewy Bodies

Physiologically, α-synuclein monomers exist in equilibrium with the membrane-bound α-Syn or in the tetrameric form [20]. In these conditions, α-syn resists abnormal pathogenic aggregation. The disruption of this balance leads to monomer assembly into oligomers through a nucleation-dependent process. Alpha-synuclein can undergo liquid-liquid phase separation (LLPS), forming condensates that may facilitate nucleation of aggregates. Alternatively, monomers can directly aggregate into oligomeric species, which are critical intermediates in the transition toward fibril formation.

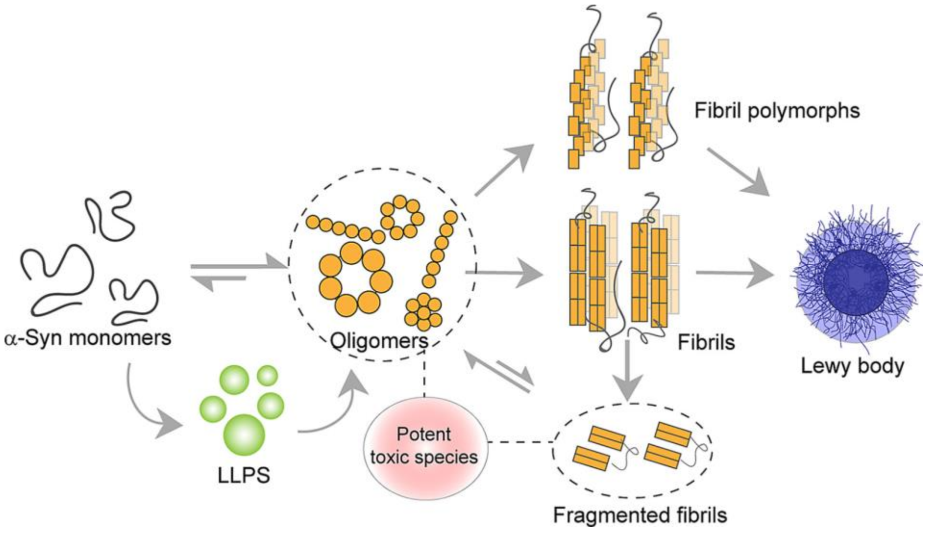


Figure 4: Mechanism of α-Syn aggregation and formation of Lewy bodies [20].

The transition converts the soluble, natively unfolded protein into β-sheet-rich amyloid structures. It proceeds through distinct phases, starting from a lag phase where unstable nuclei form, a rapid elongation as monomers add to these seeds and eventually reaching a saturation point as shown in Figure 5 [20].

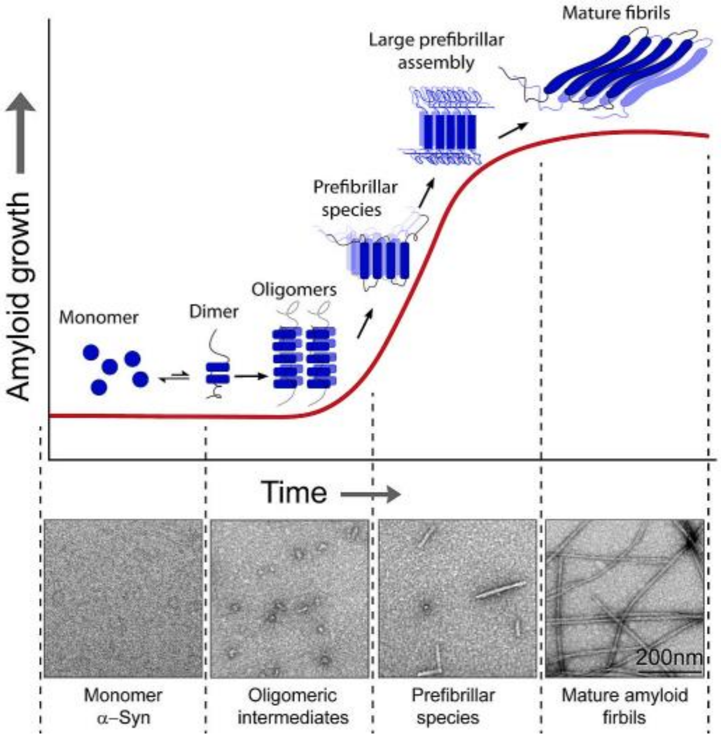


Figure 5: Schematic representing the sigmoidal growth curve for α-synuclein aggregation, representing lag, exponential and stationary phase with different conformational stages.

Fibrillar α-synuclein aggregates can accumulate and mature into Lewy bodies, the cytoplasmic inclusions characteristic of Parkinson's disease and related synucleinopathies, which are implicated in cellular dysfunction and neurodegeneration.

2.4 S100 proteins

In the context of neurodegenerative diseases, S100 proteins have long been the focus of considerable research and are still under investigation because of their role in the regulation of various processes associated with both Alzheimer’s and Parkinson’s diseases. However, their direct contribution and involvement in the pathophysiology of these disorders remain incompletely defined [21].

The S100 protein family is a group of small, acidic, calcium-binding proteins (approximately 10-12kDa), characterized by two conserved EF-hand motifs (helix E-loop-helix F) [22] as shown in Figure 6 [23]. While vertebrates possess many S100 genes, these genes are completely absent in invertebrates [24]. The name "S100" comes from their solubility in a 100%-saturated solution with ammonium sulphate at neutral pH.

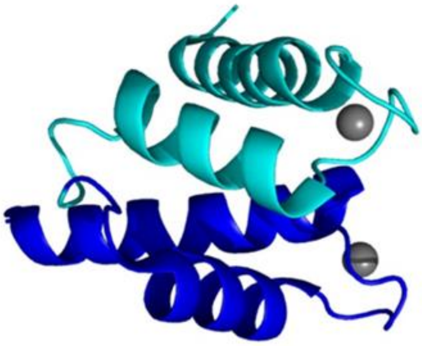


Figure 6: EF-hand domain composed of two helix-turn-helix EF-hand Ca^{2+}-binding motif.

Across a wide range of tissues, multifunctional proteins belonging to the S100 family regulate numerous intracellular and extracellular processes. Within cells, they control protein phosphorylation, modulate the dynamics of cytoskeletal components, maintain calcium homeostasis, regulate enzyme activities and transcription factors, and influence cell growth, differentiation, and the inflammatory response [25]. Regarding their roles outside the cell, there is growing evidence that some S100 proteins are released or secreted into the extracellular environment. While extracellular effects have been observed for proteins such as S100B, S100A1, S100A2, S100A4, S100A7, S100A8, S100A9, S100A10 and S100A12, secretion has been confirmed only for S100B, S100A8, and S100A9, with the precise secretion mechanism still not fully understood [26]. In the extracellular space, their effects vary with concentration—at lower levels they support cell growth, while at higher concentrations they become toxic. In addition, they serve as signals that attract leukocytes, influence cell proliferation rates and control macrophage activation [25].
This variety of functions is attributed to [27]:

The extensive diversification of S100 proteins, with 25 different members identified in humans.

- The unique metal ion-binding properties of each individual S100 protein.
- Their specific localization, either within particular intracellular compartments or in the extracellular space.
- Their ability to form non-covalent homodimers and heterodimers, which allows for the dynamic exchange of S100 subunits.

S100 proteins predominantly form homodimers, although some can also assemble into heterodimers, such as S100A8/A9, S100B/A1, and S100B/A6. In addition, these proteins can interconvert into various functional oligomers, including tetramers, hexamers, and octamers, a process facilitated by Ca^{2+} or Zn^{2+} binding. These oligomeric forms are associated with enhanced interactions with RAGE, support for microtubule assembly, stimulation of neurite outgrowth, and tumor suppression.

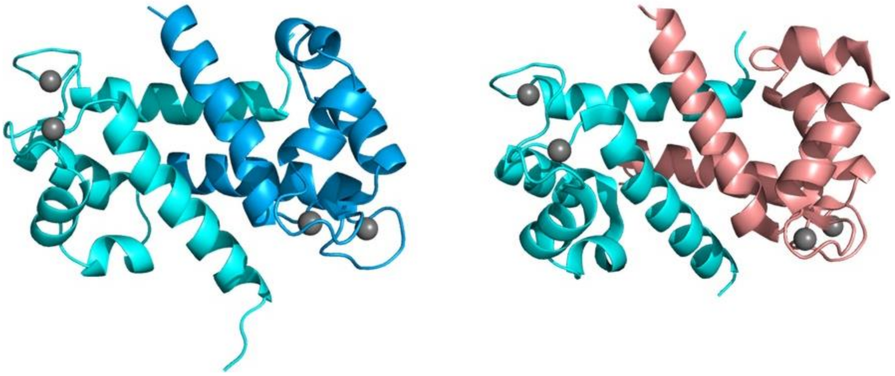


Figure 7: Integration of two EF-hand domains in the S100A8/S100A8 homodimer (left) and the S100A8/S100A9 heterodimer (right) [23] .

2.4.1 *Insights into S100B protein*

S100B was originally considered to be brain-specific but was later found in various extra-neural cells. In the nervous system, S100B is primarily found in astrocytes but is also expressed by oligodendrocytes, Schwann cells, ependymal cells, retinal Müller cells, and enteric glial cells, as well as in certain neuronal populations and peripheral ganglia. Beyond the nervous system, S100B has been detected in other cell types, including chondrocytes, melanocytes, various immune cells and others [28].

In the brain, S100B has different functions depending on its location and concentration, making it an important protein in neurobiology. Inside cells, S100B, like other S100 proteins, binds calcium and interacts with various molecular targets to regulate cellular processes including transcription, protein phosphorylation, cell growth, differentiation, and energy metabolism [29]. On the other hand, when S100B is released into the extracellular space, it exerts its effects through interactions with receptors, most notably the Receptor for Advanced Glycation End Products (RAGE).

- At *low (nanomolar) concentrations*, extracellular S100B supports normal physiological processes, acting as a neurotrophic factor that promotes neurite outgrowth and neuronal survival. These trophic effects contribute to proper neuronal development and regeneration, and are thought to play a protective role in brain function maintenance [30].

- In contrast, at *higher (micromolar) concentrations*, S100B can act as Damage-Associated Molecular Pattern (DAMP) molecule, triggering inflammatory signaling cascades. This shift from trophic to toxic effects is believed to be mediated by the enhanced activation of RAGE, which subsequently initiates pathways leading to the production of pro-inflammatory cytokines and oxidative stress. Such inflammatory responses may contribute to the progression of neurodegenerative processes and exacerbate tissue damage following injury [30], [31].

2.4.2 Computational studies on the interaction between Amyloidogenic Proteins and S100 Proteins

Several computational studies have shed light on the molecular interactions between amyloidogenic proteins and members of the S100 family, providing insights that complement experimental observations and deepen the understanding of their roles in neurodegenerative diseases.

Rodrigues et al. (2021) [32] performed a detailed study employing both computational and experimental approaches, revealing that the chaperone protein S100B dynamically interacts with the amyloid beta Aβ42 peptide, particularly through the Aβ25–35 fragment. Docking results, performed using HADDOCK, identified a binding site at the S100B dimer interface where key Aβ residues, notably Lys28, play a crucial role. Long-term MD simulations confirmed the complex's stability, demonstrating that the Aβ25–35 fragment maintains an almost complete α-helical structure throughout the simulation and interacts persistently with S100B. Energy analysis using the MM-PBSA method revealed that the complex is energetically favorable (approximately *−35 kcal/mol*), with a significant stabilizing component from electrostatic interactions (around *−80 kcal/mol*). In particular, the interaction between *Lys28* of Aβ and a triad composed of *Met79, Thr82 and Glu86* of S100B is decisive for maintaining the complex. These computational findings were supported by in vitro experiments, which showed that the anti-aggregation activity of S100B on Aβ42 significantly decreases under high ionic strength conditions (250 mM NaCl). This effect indicates that Coulombic interactions are crucial for stabilizing the S100B/Aβ42 complex, thereby contributing to the holdase-like activity of S100B that delays amyloid aggregation.

In line with this research, **Coelho et al. (2024)** [33] investigated the impact of S100B oxidation on its protective role against Aβ aggregation by mimicking oxidative extracellular conditions. A controlled oxidation process was applied to S100B, resulting in methionine oxidation, as confirmed by mass spectrometry, without evidence of cysteine-mediated crosslinking. Structural analysis showed that oxidized S100B retained its folding, overall architecture, and stability, including its quaternary structure. However, kinetic studies of Aβ aggregation demonstrated that oxidized S100B was more effective in preventing aggregation, possibly due to the oxidation of Met residues within the S100/Aβ binding region, which enhances molecular interactions.

In parallel, the interaction between S100A9 and α-synuclein has also been investigated using different approaches.

Horvath et al. (2018) [34] provided strong ex vivo and in vitro evidence that S100A9 co-localizes with α-synuclein in Lewy bodies and neuronal cells of Parkinson's disease patients. Ex vivo imaging techniques were used to map the localization and aggregation of S100A9 and α-synuclein in Parkinson's disease brain tissues. In vitro studies revealed that S100A9 interacts with the C-terminal region of α-synuclein ($Kd \approx 5\ \mu M$) and that their co-aggregation occurs more rapidly than when each protein assembles individually, forming larger amyloid aggregates. These co-localized aggregates were found in approximately 20% of Lewy bodies and 77% of neuronal cells in the substantia nigra, with the co-aggregation mitigating the toxicity of S100A9 oligomers.

Additionally, **Toleikis et al. (2021)** [35] revealed that in the absence of S100A9, alpha-synuclein aggregated into three distinct fibril types, each with unique fluorescence properties, aggregation kinetics, and structural characteristics, as determined by FTIR and AFM analyses. When even low concentrations of S100A9 were introduced, the aggregation behavior of alpha-synuclein changed markedly. The presence of S100A9 reduced the heterogeneity of the aggregation process by stabilizing a specific fibril structure, similar to one naturally occurring species observed in its absence. However, while fluorescence and FTIR data indicated a convergence toward this stable structure, AFM images showed that the fibrils became thicker, suggesting an enhanced lateral association of fibril strands. Furthermore, the study examined the role of ionic strength by varying NaCl concentrations, demonstrating that the interaction between S100A9 and alpha-synuclein was sensitive to ionic conditions. This sensitivity underscored the importance of ion-pair interactions in modulating the aggregation process. Overall, the authors suggested that S100A9 not only accelerated the nucleation phase of alpha-synuclein aggregation but also directed the formation of specific fibril structures.

Complementing these observations, **Toleikis et al. (2022)** [7] used ^{19}F and 2D ^{15}N–^{1}H HSQC NMR spectroscopy in combination with molecular dynamics simulations to map the interaction interface between S100A9 and α-synuclein. Their computational work identified specific regions (*the N-terminal segment, specifically from Val3 to Ser9, the region spanning from Glu35 to Gly41 and, to a lesser extent, the C-terminal segment from Glu123 to Tyr133*) that are crucial for both binding to S100A9 and initiating aggregation. These findings suggest that S100A9 may influence the early nucleation events of α-synuclein aggregation, potentially altering the pathway and kinetics of fibril formation.

Collectively, these computational studies highlight that S100B appears to function as a chaperone that can inhibit Aβ aggregation through a mechanism highly dependent on calcium binding and its oxidation state. On the other hand, S100A9 seems to promote the aggregation of α-synuclein, potentially exacerbating the neurodegenerative process in Parkinson's disease.

Building on these studies, this thesis explores a crossover between these two lines of research by investigating the potential interaction between S100B and α-synuclein, which has not yet been reported. This work aims to determine whether S100B, known for its chaperone-like role in inhibiting Aβ aggregation, may also influence α-synuclein dynamics, thereby providing new insights into its broader neuroprotective or pathogenic implications.

3 Materials and Methods

3.1 Molecular Modeling

Molecular Modeling is a set of theoretical and computational approaches that can be used to study the molecular structure of various biological systems, allowing the simulation and analysis of the behavior of molecules.

Molecular Modeling techniques allow for the examination of a wide range of molecular systems, from small chemical compounds to large biological macromolecules, with a particular focus on protein structures and complex biological systems [36].

These analyses investigate features such as the number and types of atoms, bond characteristics, including their nature, lengths, and angles, as well as dihedral angles, energy profiles, geometric optimization, enthalpy, and vibrational frequencies [37]. Furthermore, they evaluate properties like nucleophilicity, electrophilicity, and electrostatic potential, thereby facilitating the prediction of molecular and biological characteristics that underpin structure–activity relationships. The fundamental idea is to simplify the intrinsic complexity of molecular systems, sometimes neglecting electronic details in order to concentrate on the interactions between atomic nuclei and macroscopically observable properties.

Molecular modeling is rapidly evolving, as are its applications in many research fields. Some of its applications include computational biology, computational chemistry, drug discovery and design, biomaterials development, emerging material research, and spectroscopy.

3.2 Molecular Mechanics

The term Molecular Mechanics (MM) was coined in the 1970s and refers to a computational approach that models molecular systems by representing them in terms of the positions of their atomic nuclei, thereby bypassing the need to explicitly account for electron motion. This simplification is made possible by the Born–Oppenheimer approximation which assumes that, owing to their significantly lower mass, electrons adjust almost instantaneously to any changes in the positions of the much heavier nuclei.

By treating the nuclei as classical particles that follow Newton's laws, MM reduces the complexity of the system, enabling the simulation of large molecular structures that can contain thousands of atoms [38].

Molecular Mechanics is also commonly known as the force field method (sometimes referred to as Westheimer's method), where a force field is defined as the collection of mathematical functions and parameters used to describe the potential energy of a molecular system.

3.2.1 Potential Energy Function

In Molecular Mechanics, the potential energy function (PEF) describes the energy of a molecular system as a function of the positions of its constituent atoms. For a system composed of *N* atoms, where each atom is represented by a position vector r^i, the overall potential energy, often referred to as the potential energy surface (PES), can be expressed as:

$$\mathcal{V}(r^N) = \mathcal{V}_{bond}(r^N) + \mathcal{V}_{non-bond}(r^N) \tag{1}$$

This summation provides a complete description of how atoms, seen as masses interconnected by springs, interact within the system.

- $\mathcal{V}_{bond}(r^N)$ represents the bonded interactions involving atoms connected by chemical bonds, describing phenomena such as bond stretching, angle bending, and torsional rotations.
- $\mathcal{V}_{non\text{-}bond}(r^N)$ represents the non-bond interactions, such as van der Waals attractions and electrostatic repulsions.

3.2.2 Bond Interactions

Bonded interactions in Molecular Mechanics refer to the energy contributions resulting from atoms connected by covalent bonds. They are typically divided into three categories: bond stretching (involving two atoms), angle bending (involving three atoms), and torsional or dihedral interactions (involving four atoms) [39].

$$\mathcal{V}_{bond}(r^N) = \sum_{bonds} \mathcal{V}_{bonds}(r^N) + \sum_{angles} \mathcal{V}_{angles}(r^N) + \sum_{torsion} \mathcal{V}_{torsion}(r^N) \tag{2}$$

Different formulations have been proposed to describe each of these contributions to the overall bond potential energy function.

The energy associated with bond stretching (Figure 8A) is most commonly modeled using Hooke's law, which introduces a penalty when the bond length deviates from the reference value l_0. Alternative approaches such as cubic or Morse potentials may also be used to capture anharmonic effects more accurately. The harmonic model of the bond term is:

$$\mathcal{V}_{bonds}(l) = \frac{k_i}{2}\left(l_i - l_{i,0}\right)^2 \tag{3}$$

where k is the stretching constant of the bond and l_0 is the reference bond length between the two particles interacting.

The angles term (Figure 8B) describes the energy associated with deviations from the ideal valence angle defined by three atoms (i.e., atoms i, j and k, with i and k both bonded to j). Similarly, it is often represented by a harmonic potential characterized by a force constant h and a reference value of the valence angle θ_0. Less energy is needed to distort an angle than to stretch or compress a bond.

$$\mathcal{V}_{angles}(\theta) = \frac{h_i}{2}\left(\theta_i - \theta_{i,0}\right)^2 \tag{4}$$

The torsional or dihedral term (Figure 8) describes the energy change that occurs during the rotation around a bond involving four atoms. It is usually modeled by a sinusoidal function, which is parameterized by a barrier term V_n reflecting the energy required for rotation, a multiplicity n indicating the number of minima encountered in a full 360° rotation, and a phase factor γ defining the position of the minima.

$$\mathcal{V}_{torsion}(\phi) = \frac{V_n}{2}\left(1 + \cos(n\phi - \gamma)\right) \tag{5}$$

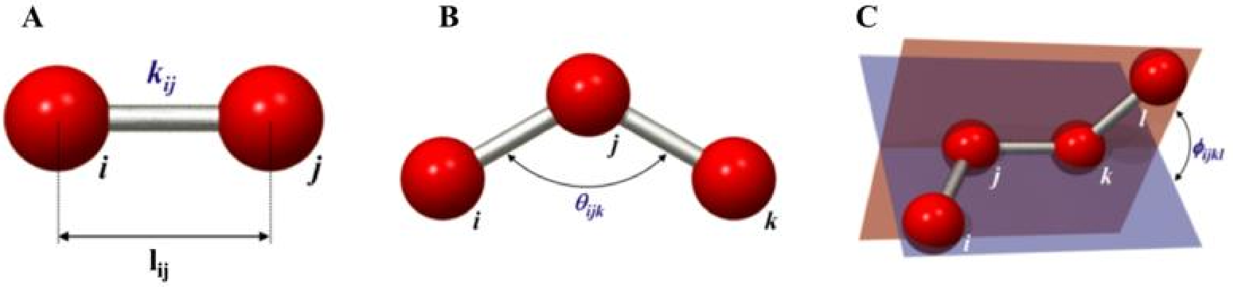


Figure 8: Representation of bond interactions. (A) Bond length. (B) Bond angle. (C) Bond torsion.

3.2.3 Non-Bond Interactions

Non-bond interactions describe the forces between atoms or molecules that are not directly connected by covalent bonds and are typically modeled as functions inversely proportional to the distance between the interacting particles.

The inter-molecular contribution to the potential energy function can be expressed as:

$$\mathcal{V}_{non-bond}(r^N) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \mathcal{V}_{electrostatic}(r_{ij}) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \mathcal{V}_{vdW}(r_{ij}) \qquad (6)$$

Electrostatic forces are long-range interactions that extend up to about 10 nm. They result from differences in electronegativity between atoms, which cause an unequal distribution of charges within molecules. The interaction energy can be calculated using Coulomb's law as follows:

$$\mathcal{V}_{electrostatic}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad (7)$$

The energy associated to two atoms with net atomic charges q_i and q_j at a distance r_{ij} is proportional to the product of their charges and inversely proportional to the distance between them.

Van der Waals forces are short-range interactions and become significant when atoms are close enough to interact (typically within about 1 nm). At short distances, the forces are repulsive and grows exponentially; while at longer distances, attractive forces - primarily London dispersion forces, which arise from transient induced dipoles - contribute to a weak attraction between atoms. The Lennard-

Jones 12-6 potential is the most commonly used mathematical model to describe van der Waals interactions:

$$\mathcal{V}_{vdW}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{8}$$

It is characterized by two parameters as shown in Figure 9: the collision diameter σ, representing the distance at which the potential energy is zero, and the well depth ϵ, which indicates the minimum potential energy at equilibrium.

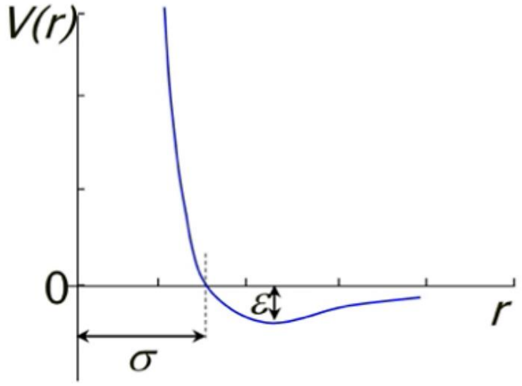


Figure 9: Potential energy of van der Waals interaction modelled with Lennard-Jones 12-6 potential.

Because every charged particle interacts with all others in the system, the number of non-bond interactions increases quadratically with the number of particles. To reduce the computational cost of these calculations, various techniques such as distance cut-offs, potential switching functions, and the Particle Mesh Ewald method are employed.

3.2.4 Periodic Boundary Conditions

In molecular dynamics simulations, only a finite number of particles can be considered due to computational constraints. They are usually inserted in a simulation box whose geometry depends on the system being studied. The most common geometries include the cubic box, the truncated octahedron and the rhombic dodecahedron, each offering specific advantages. For example, the hexagonal prism is often used for elongated molecules, such as DNA, while the rhombic dodecahedron approximates a spherical geometry, optimizing space utilization and minimizing empty volume.

To avoid edge effects that could artificially influence the system's properties, Periodic Boundary Conditions (PBC) are commonly implemented (Figure 10). This approach involves replicating the simulation box in all directions, creating an infinite grid of periodic replicas [40]. Particles located at the edges of the primary box interact with periodic images of the particles on the opposite sides, thus eliminating the artificial boundary and allowing the system to behave as if it were infinite in size. Therefore, the application of PBC ensures that the total number of particles in the box remains constant.

However, the use of PBC introduces the necessity to respect the *minimum image convention*: the cut-off radius for non-bonded interactions must be less than half the smallest side length of the box.

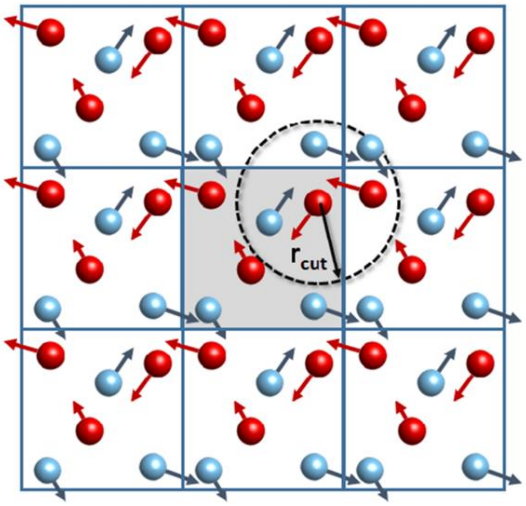


Figure 10: Representation of PBC in a cubic box, illustrating the cut-off radius required to satisfy the minimum image convention.

3.2.5 Minimization of the potential energy function

The PES (Potential Energy Surface) is a multidimensional function of the molecular system's coordinates. A system of N atoms can be described using 3N Cartesian coordinates, which specify the position of each atom in three-dimensional space, or 3N-6 internal coordinates, which describe the system's structure in terms of bond lengths, bond angles, and torsional angles. The Potential Energy Surface is characterized by various minima, corresponding to stable molecular configurations. *Local minima* refer to points that are lower than the surrounding regions, while the *global minimum* refers to the lowest energy point on the entire surface, representing the most stable possible configuration.

Several approaches have been proposed in the literature to locate minima on the PES. However, most algorithms can only identify the closest local minimum from the starting point rather than the global minimum, which often does not correspond to the molecule's active state. Nonetheless, reaching the nearest minimum is still relevant, as the relationship between force and potential energy implies that being at a minimum results in lower atomic forces and a more stable structure. The energy minimization methods are divided into *derivative methods*, which use the energy derivatives, and *non-derivative methods*.

Non-derivative methods include the *simplex method* and the *sequential univariate method*.

Derivative methods are divided into first- derivative methods and second-derivative methods. The former rely on the gradient of the potential energy, as the direction of the first derivative indicates where the minimum is located. Among first-order derivative methods, the most well-known are *steepest descent, conjugate gradients minimization*, and *line search in one dimension*.

Second-derivative methods, on the other hand, are based on the second derivatives, which provide information about the curvature of the function, allowing for a prediction of where the function will change direction. Among them are the *Newton-Raphson method* and the *Broyden-Fletcher-Goldfarb-Shanno (BFGS) method*.

There is no universally best method. The choice of the minimization method depends on various factors, including available memory capacity, robustness, desired accuracy, and the ability to compute the second derivative.

3.3 Molecular Dynamics

To analyze complex phenomena involving the temporal evolution of molecules, such as conformational changes or dynamic interactions, it is essential to adopt methods that consider the system's time dimension. In this context, molecular dynamics (MD) emerges as a fundamental tool, as it is a simulation technique capable of modeling the time-dependent behavior of complex molecular systems, thereby providing a detailed and comprehensive view of dynamic processes at the atomic level.

MD is based on the numerical integration of Newton's equations of motion. For a system composed of N particles, Newton's second law states that the force F_i acting on a particle is equal to the product of its mass m_i and its acceleration (the second derivative of its position r_i):

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} = \boldsymbol{F}_i \tag{9}$$

The force acting on each particle can also be expressed as the gradient of the system's potential energy E(R):

$$\boldsymbol{F}_i = -\nabla_{r_i} E(R) \tag{10}$$

where ∇_{r_i} denotes the gradient with respect to the position of the i-th particle, and $R(t)=(r_1(t), r_2(t), \ldots, r_N(t))$ is the vector of the positions of the system's particles. By combining these two relations, the following equation is obtained:

$$m_i \frac{\partial^2 r_i}{\partial t^2} = -\nabla_{r_i} E(R) \tag{11}$$

This differential equation must be solved for all particles in the system to determine the time evolution of their positions and velocities. Since these equations are usually not solvable analytically for larger systems, numerical integration methods are employed. Many methods have been proposed in literature such Verlet scheme, Velocity Verlet, Leap frog.
The choice of the integration timestep Δt is crucial and largely depends on the oscillation frequency of the phenomena being studied. It must be sufficiently small to accurately capture the system's dynamics but large enough to make the simulation computationally efficient. Typically, Δt is on the order of femtoseconds (10^{-15} seconds).

3.3.1 Statical ensembles

In statistical mechanics, the principal aim of molecular dynamics is to determine the average properties of a system by sampling the accessible microstates within phase space.

For a system of N atoms, *phase space* is a multidimensional space with 6N dimensions, representing all possible configurations of the system. Each point in this space, called a *microstate*, represents a specific configuration defined by 3N positional coordinates and 3N momentum components. A collection of microstates sharing identical macroscopic properties, such as temperature or pressure, constitutes a *macrostate*. A *statistical ensemble* consists of all points in phase space that have the same macroscopic or thermodynamic state.

In Molecular Dynamics, the main statistical ensembles are:

- The *micro-canonical ensemble (NVE),* which describes an isolated system characterized by a fixed number of particles (N), assigned volume (V) and constant energy (E);
- The *canonical ensemble (NVT),* which describes a closed system with fixed number of particles (N), assigned volume (V), and coupled with a thermostat to maintain a constant temperature (T);
- The *isothemal-isobaric ensemble (NPT),* which describes a closed system with fixed number of particles (N) coupled both with a thermostat and a barostat to maintain constant temperature (T) and pressure (P);
- The *grand-canonical ensemble (μVT),* which describes an open system with fixed chemical potential (μ), volume (V), and temperature (T).

Since exploring every possible microstate accessible to the system is impractical, the ergodic hypothesis can be employed to calculate the representative average properties of a system. By allowing the system to evolve over a sufficiently long period, it is assumed that the system will pass through all accessible states. Consequently, the ensemble average of a property A corresponds to its time average:

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \tag{12}$$

3.3.2 Procedural Steps in Molecular Dynamics

In this work, molecular dynamics (MD) simulations are conducted following a general protocol:

1. *Simulation Box Setup:* Define a dodecahedral box to encapsulate the system under investigation, ensuring sufficient space to prevent unwanted interactions between periodic images.

2. *Solvation*: Solvent molecules, typically water, are added to the simulation box to fully immerse the solute. This step is crucial for replicating the system's natural environmental conditions.

3. *Ions Addition:* To neutralize the system's total charge and simulate realistic ionic conditions, ions such as sodium (Na^+) or chloride (Cl^-) are introduced into the solution.

4. *Energy Minimization:* The system's structure is optimized by minimizing its potential energy.

5. *Equilibration at constant temperature (NVT) and equilibration at constant pressure (NPT):* The system is gradually brought to the target thermodynamic conditions, such as a specific temperature and pressure.

6. *Production MD:* After equilibration, the production simulation is performed, during which particle trajectories are recorded over time, collecting data on their positions and velocities.

7. *Removal of artifacts caused by periodic boundary conditions (PBC)*

8. *Analysis of results*

3.4 Molecular Docking

Molecular docking is a computational technique widely used in structural biology and drug discovery to simulate and predict the preferred orientation of a molecule when it binds to another molecule, forming a complex. This in silico approach provides an in-depth view of molecular binding phenomena through sophisticated algorithms and scoring functions, enabling the assessment of binding affinity, the prediction of binding modes, and the generation of crucial insights for optimizing drug development processes. [41].

The term *ligand-protein docking* refers to the interaction where a ligand, which can be a small molecule such as a drug or metabolite, binds to a target protein. In contrast, *protein-protein docking* involves interactions between two proteins [42]. The molecular docking process involves two main stages:

- *Conformational sampling*: In this phase, various spatial configurations of the involved molecules are generated to explore possible interaction modes.
- *Evaluation using Scoring Functions*: The generated conformations are analyzed using scoring functions to determine which configuration represents the most stable and probable interaction.

Ideally, sampling algorithms should accurately reproduce the experimental binding mode, while scoring functions should rank this conformation as the most favorable among all generated poses [43]. There are various types of scoring functions used in molecular docking, including force field-based, empirical, knowledge-based and machine learning-based methods [44].

Different computational tools and algorithms are available for molecular docking techniques, including both commercial and free-of-charge options. Some of the most commonly used docking programs include AutoDock, GOLD, and DOCK, each recognized for their effectiveness in optimizing lead compounds, facilitating virtual screening, and elucidating the intricate details of protein-ligand interactions. [41], [45].

The process of obtaining the optimal binding conformation, which minimizes the system's free energy, involves several stages, as illustrated in Figure 11 [46]. Molecular docking is structure-based and requires a high-resolution 3D representation of the target protein obtained through techniques like X-ray crystallography, Nuclear Magnetic Resonance Spectroscopy, or Cryo-Electron Microscopy. Thus, the process typically begins with obtaining the three-dimensional structures of both the target protein and ligand. Subsequently, protonation states and partial charges are assigned to these molecules. If not previously known, the binding site on the target is detected or a blind docking approach may be employed. The docking procedure itself comprises two primary steps: generating possible ligand poses within the target site and evaluating these poses using scoring functions, resulting in a ranked list of potential target-ligand complexes.

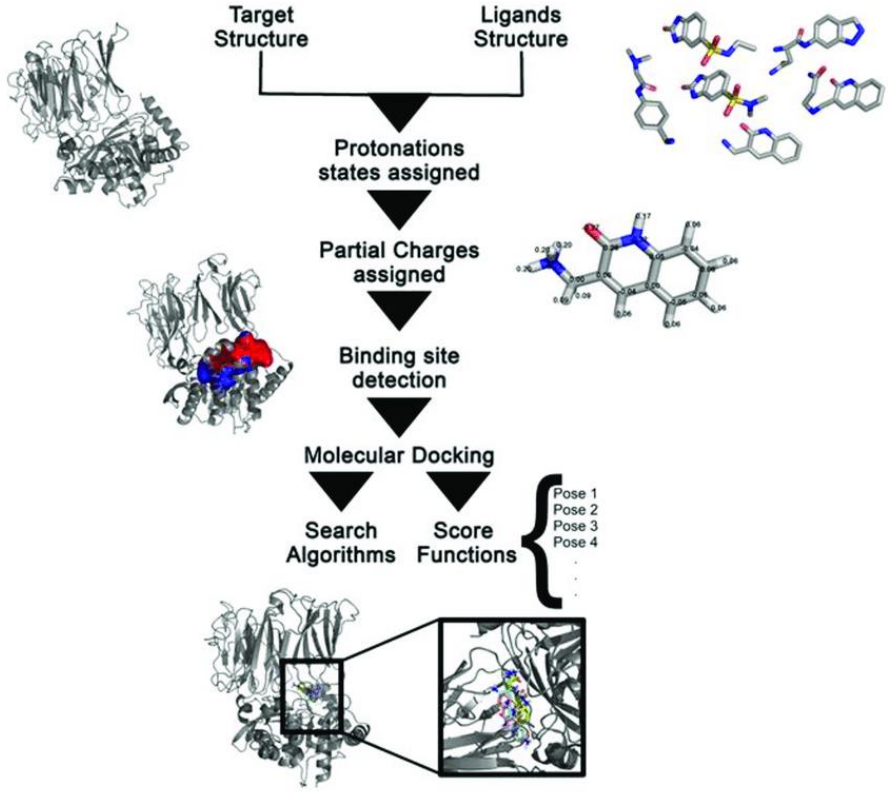


Figure 11: General workflow of molecular docking calculations [46] .

3.4.1 HADDOCK

HADDOCK (High Ambiguity Driven DOCKing) is an integrative modeling software designed to predict the three-dimensional structures of macromolecular complexes, such as protein-peptide interactions. This platform distinguishes itself from others by incorporating experimental and bioinformatics data to guide the docking process [47]. It utilizes information from various experimental techniques, including NMR chemical shift perturbations and biochemical interaction data, to define Ambiguous Interaction Restraints (AIRs). These restraints focus the conformational search on the most probable interaction interfaces, enhancing the accuracy of the predicted models, and categorize residues into [48]:

- *Active Residues*: solvent-exposed residues directly participating in the interaction between molecules and identified through experimental data. They are "enforced" to be at the interface.
- *Passive Residues:* solvent-accessible residues adjacent to active residues. They are incorporated to overcome the limitation of experimental data, which often provide incomplete

coverage of the actual binding interface. Passive residues may be at the interface but are not penalized if absent.

HADDOCK's docking process comprises three main stages [48]:

1. **Rigid-Body Docking (It0)**: Initial positioning of molecules based on AIRs, treating them as rigid entities to generate preliminary complex conformations.
2. **Semi-Flexible Refinement (It1)**: Introduction of flexibility at the interface, allowing side chains and backbone adjustments through simulated annealing to refine interactions.
3. **Explicit Solvent Refinement**: Final optimization of the complex in an explicit solvent environment, enhancing the realism and stability of the predicted model.

3.5 Alphafold2-Multimer

AlphaFold2-Multimer is an extension of the AlphaFold2 framework developed by DeepMind, specifically designed to predict the structures of protein complexes [49]. While traditional molecular docking methods simulate the binding interactions between molecules, often requiring predefined binding sites and relying on scoring functions to evaluate binding affinities, AlphaFold2-Multimer approaches the problem differently. It predicts the three-dimensional structures of protein complexes directly from amino acid sequences, without necessitating prior knowledge of interaction interfaces [49], [50]. The methodology of AlphaFold2-Multimer is rooted in deep learning techniques. It utilizes multiple sequence alignments (MSAs) and evolutionary information to model inter-protein contacts and interfaces accurately. By integrating co-evolutionary data, the model can infer which residues across different protein chains are likely to be in proximity, thereby predicting the quaternary structure of protein assemblies with high precision. The neural network architecture of AlphaFold2-Multimer is adapted from AlphaFold2 but incorporates additional modifications to handle multi-chain predictions, enabling accurate modeling of homomeric and heteromeric complexes [50].

One of the most innovative aspects of AlphaFold2-Multimer is its ability to generate plausible protein structures even in the absence of experimental data, modeling the entire complex structure de novo and providing an extremely powerful predictive tool for structural research. However, since the model does not directly compute binding energies or assess conformational stability through explicit simulations, the generated structures could require further refinement for optimal application in docking or molecular dynamics contexts [51].

4 Construction of the α-Synuclein/S100B Complex Structure through Molecular Docking and AI-Based Prediction

4.1 Introduction

Parkinson's disease and Alzheimer's disease are the most common neurodegenerative disorders, characterized by a progressive decline in neurological functions. In Parkinson's disease, neuronal degeneration in the substantia nigra within the basal ganglia leads to motor impairments. In Alzheimer's disease, on the other hand, damage to neurons in the hippocampus and cerebral cortex results in a gradual deterioration of cognitive abilities and memory. In both cases, the progressive and irreversible loss of neurons is associated with the presence of misfolded protein aggregates in the brain, known as amyloid fibrils, which are formed by β-amyloid accumulation in Alzheimer's disease and α-synuclein accumulation in Parkinson's disease [52].

Several studies have been conducted to investigate the mechanisms of protein aggregation in neurodegenerative diseases, with the aim of understanding their dynamics and causes, as well as identifying molecules capable of modulating their kinetics. As part of this research, the Unite! Seed Fund project, titled "*Multiscale investigation on the stabilization of amyloid proteins associated with neurodegeneration by metal ions and chaperones*", is a collaboration between TU Darmstadt, ULisboa, and Politecnico di Torino. The project aims to improve the understanding of the molecular mechanisms underlying protein aggregation in neurodegenerative diseases and to identify the factors that regulate these pathological processes in the human brain. In particular, through a combined in vitro and in silico approach, the study analyzes the interactions between aggregation-prone proteins, metal ions and chaperone proteins. Among these chaperones, special attention is given to S100 proteins, which play a crucial role in the dynamic regulation of proteins and metal ions to prevent amyloid aggregation. Currently, the understanding of these mechanisms remains limited, representing a significant obstacle both in studying the pathogenesis of neurodegenerative diseases and in developing effective therapies. The Unite! Seed Fund project therefore aims to bridge this gap, providing new fundamental insights for the design of innovative therapeutic strategies.

Studies in the literature on the interaction between Aβ and S100B have suggested that S100B may function as a chaperone, inhibiting Aβ aggregation [32], [33]. On the other hand, research on the interaction between α-synuclein and S100A9 has shown that S100A9 appears to promote α-synuclein aggregation [7], [34], [35]. Building on this evidence, this thesis investigates the intersection between these two lines of research by examining the potential interaction between S100B and α-synuclein, a relationship that has not yet been characterized.

This chapter focuses on the construction of the α-synuclein/S100B complexes, which are used in the molecular dynamics simulations described in the following chapter. Since no experimental structures of the α-synuclein/S100B complex are available, multiple computational approaches have been employed to generate them.

4.2 Materials and methods

Molecular docking techniques using HADDOCK, along with AI-based methods such as AlphaFold, have been applied to generate structural models of the α-synuclein/S100B complex.

4.2.1 Molecular Docking Model Construction Using HADDOCK

To generate structural models of the α-synuclein/S100B complex, molecular docking calculations were performed using the HADDOCK 2.4 web server with standard parameters. HADDOCK requires specific input data, including the 3D structures of both molecules and the definition of active and passive residues for each protein. Passive residues were automatically assigned by the software. However, since no experimental data were available for the interaction between α-synuclein and S100B, the selection of active residues was performed following an approach guided from insight on S100B amyloid interaction and S100A9 α-syn interaction reported in previous studies [7] [32].

Initially, a comparative sequence analysis was performed using MOE to evaluate the sequence identity and similarity between α-synuclein and β-amyloid. To ensure consistency with the approach used by Toleikis et al. [7], only the first 50 amino acids of α-synuclein were considered. Therefore, starting from the α-syn PDB structure 1XQ8 and using MOE, residues 51 to 140 were removed. Similarly, for β-amyloid, only the $A\beta_{25\text{-}35}$ fragment was considered, as suggested in the study by Rodrigues et al. [32]. The structure was obtained from the PDB entry 1IYT, and residues 25 to 35 were isolated using MOE.

Subsequently, a structural comparison was conducted between S100B and S100A9. For this purpose, the S100B dimer was obtained from the Protein Data Bank (PDB: 3D0Y) and processed using PyMOL to remove water molecules, Zn^{2+} and Ca^{2+} ions, and heteroatoms (PG4 and FME). Similarly, S100A9 was obtained from its crystal structure 5I8N in the PDB and processed in PyMOL by removing water molecules and Ca^{2+} ions.

In both comparisons, the two structures were aligned in two different ways: first based only on sequence, and then using both structure and sequence. They were then superposed and compared. The best results in terms of similarity and identity were obtained when the alignment was based only on sequence. The sequence and structural similarity comparisons conducted aimed at evaluating the following aspects:

- Whether the active residues of S100A9, identified in its interaction with α-synuclein by Toleikis et al. [7], could also be applicable to S100B when it forms a complex with α-synuclein.
- Whether the active residues of α-synuclein involved in its interaction with S100B could correspond to those identified for β-amyloid in its interaction with S100B by Rodrigues et al. [32].
- Whether the active residues of S100B in its interaction with α-synuclein could be the same as those found for S100A9 in its interaction with α-synuclein by Toleikis et al. [7].

Two docking calculations were performed using the previously mentioned structure of S100B and a fragment of alpha-synuclein built in PyMOL, comprising residues 31 to 41. Each calculation employed a distinct set of active residues, while in both cases, the passive residues were automatically assigned by the software based on their spatial proximity to the active residues.

For the first docking calculation, the active residues were defined as follows (Table 2):

- **S100B active residues**: the residues of S100B that, in the sequence alignment between S100B and S100A9, corresponded to the S100A9 residues experimentally identified as active in its interaction with α-synuclein (as reported by Toleikis et al. [7]).
- **α-synuclein active residues**: the residues of α-synuclein that, in the sequence alignment between β-amyloid and α-synuclein, corresponded to the β-amyloid residues reported as active in its interaction with S100B (as indicated in the study by Rodrigues et al. [32]).

Table 2: Set of active residues used in the first docking calculation.

Active residues of S100B	***Active residues of α-synuclein***
Leu3, Glu4, Lys5, Ala6, Met7, Val8, Ala9, Leu10, Ile11, Asp12, Val13, Phe14, Leu35, Phe43, Leu44, Glu45, - , Glu46, Val53, Asp54, Lys55, Val56, Met57, Glu58, Thr59, Leu60, Asp61, Asn62, Asp63, Gly64, Asp65, Gly66, Asp69, Val80, Thr81, Thr82, Ala83, Cys84 Leu93, Glu94, Lys95, Ala96, Met97, Val98, Ala99, Leu100, Ile101, Asp102, Val103, Phe104, Leu125, Phe133, Leu134, Glu135, -, Glu136, Val143, Asp144, Lys145, Val146, Met147, Glu148, Thr149, Leu150, Asp151, Asn152, Asp153, Gly154, Asp155, Gly156, Asp159,Val170, Thr171, Thr172, Ala173, Cys174	Gly31, Lys32, Thr33, Lys34, Glu35, Gly36, Val37, Leu38, Tyr39, Val40, Gly41

For the second docking calculation, the selection of active residues was as follows (**Error! Reference source not found.**):

- **S100B active residues**: the active residues of S100B identified in its interaction with β-amyloid, as reported in the study by Rodrigues et al. [32].
- **α-synuclein active residues**: the same α-synuclein residues of the first calculation.

Active residues of S100B	***Active residues of α-synuclein***
Asn62, Gln71, Met74, Ala75, Ala78, Met79, Thr82, Ala83, Gln161, Glu162, Ala165, Ala168, Met169, Thr172, Glu176	Gly31, Lys32, Thr33, Lys34, Glu35, Gly36, Val37, Leu38, Tyr39, Val40, Gly41

Table 3: Set of active residues used in the second docking calculation.

Finally, the HADDOCK solutions were then evaluated and compared using the HADDOCK score to identify the most relevant binding mode to be studied through molecular dynamics simulations.

4.2.2 Building the AlphaFold2-Multimer Structural Model

To further explore potential interaction modes without imposing prior assumptions on binding sites, structural models of the α-synuclein/S100B complex were also generated using an AI-based approach, called AlphaFold2-Multimer [53]. This was implemented through ColabFold, a cloud-based adaptation of AlphaFold2, optimized for multimeric protein structure prediction. ColabFold integrates MMseqs2 for fast multiple sequence alignment (MSA) generation and applies deep learning models trained on extensive structural datasets to infer potential protein-protein interactions [54].

The primary amino acid sequences used as input for modeling the complex consisted of the first 50 residues of α-synuclein (*UniProtKB: P37840*) and the S100B dimer (*UniProtKB: P04271*). The formatted query sequence was reported in Table 4.

Table 4: Primary sequence for α-synuclein and S100B dimer

Protein name	**Primary sequence**
α-synuclein	MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVH:
S100B dimer	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETLDNDGDGECDFQEFMAFVAMVTTACHEFFEHE:MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETLDNDGDGECDFQEFMAFVAMVTTACHEFFEHE

A colon (":") was used to specify inter-protein chain breaks, distinguishing the three chains that compose the complex: the first 50 residues of α-synuclein, the first monomer of S100B and the second monomer of S100B.

Multiple computational trials were performed to explore different configurations of the complex, systematically varying key parameters. A summary of all the computational trials carried out is presented in Table 5, which reports the combination of parameters used for each prediction.

Table 5: Summary of AlphaFold2-Multimer trials performed for the α-synuclein/S100B complex.

Trial	*template_mode*	*num_recycles*	*num_relax*	*relax_max_iterations*
1	*none*	*24*	*0*	*-*
2	*pdb100*	*48*	*0*	*-*
3	*none*	*48*	*0*	*-*
4	*none*	*48*	*1*	*200*
5	*pdb100*	*24*	*0*	*-*

Across all trials, the following parameters were kept constant:

- *msa_mode*: mmseqs2_uniref_env, to enhance sequence diversity by searching both UniRef and environmental databases.
- *pair_mode*: unpaired_paired, allowing the algorithm to select pairs of sequences from the same species while simultaneously retaining an unpaired MSA for broader evolutionary coverage.
- *model_type*: auto, automatically applying alphafold2_multimer_v3 for complex predictions.
- *recycle_early_stop_tolerance*: auto, meaning that alphafold2_multimer_v3 was applied for complex predictions, with a default threshold of 0.5 when model_type = alphafold2_multimer_v3.
- *pairing strategy*: greedy, enabling the algorithm to pair taxonomically related sequences to improve evolutionary-based predictions.
- *ranked outputs*: For each trial, the top five ranked models (rank_1 to rank_5) were selected for visualization, and the lDDT (Intrinsic Distance Difference Test) coloring scheme was applied to assess local confidence scores.

The main variables tested across the computational trials were:

- *template_mode*: set either to none (no structural templates used, relying entirely on sequence-based predictions) or to pdb100 (enabling the search for homologous structures in the PDB100 database to incorporate template-derived information).
- *num_recycles*: set to either 24 or 48, controlling the number of recycling steps to refine predictions and potentially improve model quality.
- *Amber relaxation*: most models were generated with relaxation disabled (num_relax = 0), meaning no steric optimization was applied. However, in one specific trial, relaxation was enabled (num_relax = 1) with a maximum of 200 iterations (relax_max_iterations = 200). Despite this additional optimization step, no significant structural changes were observed, suggesting that the predicted models were already well-formed without further refinement.

4.2.3 Binding Affinity Evaluation using PRODIGY

To quantitatively assess the predicted α-synuclein/S100B complexes, a comparative evaluation of the HADDOCK-based docking models and AlphaFold2-Multimer-generated structures was performed. While HADDOCK inherently provides a scoring function (the HADDOCK score), AlphaFold2 does not include a built-in scoring metric to directly estimate the stability of protein-protein interactions. To overcome this limitation and ensure a consistent comparison across all predicted models, the PRODIGY (PROtein binDIng enerGY prediction) web server was used. PRODIGY is a widely adopted tool for estimating binding affinity ΔG and the dissociation constant Kd of protein-protein complexes based on structural parameters such as interfacial contacts *(*ICs*)* and non-interacting surface (NIS) areas [55].

Specifically, PRODIGY calculates the following key parameters:

- **Binding free energy** (ΔG, in kcal/mol): indicates the thermodynamic favorability of the interaction, where more negative values correspond to stronger binding.
- **Dissociation constant** (Kd, in M): reflects the stability of the complex, with lower values indicating higher interaction strength.
- **Interfacial contacts** (ICs): describe the nature of the interactions at the interface, distinguishing between electrostatic (charged-charged, charged-polar, charged-apolar) and hydrophobic (polar-apolar, apolar-apolar) contributions, which are essential for protein-protein stabilization.
- **Non-interacting surface** (NIS): quantifies the solvent-exposed surface areas of the complex, with lower values generally suggesting a more compact and stable interface.

A PRODIGY calculation performed on all models allowed for a standardized and comparative assessment of the α-synuclein/S100B interaction models, supporting the identification of the most energetically favorable configurations to be considered for further analysis.

4.3 Results

Results were analyzed and compared to identify the most stable interaction models.

4.3.1 PRODIGY-Based Comparison of α-Synuclein/S100B Interaction Models

PRODIGY calculations were carried out on the two HADDOCK docking models and all AlphaFold2-Multimer-generated structures. Among the AlphaFold2 models, the most favorable PRODIGY-predicted binding energy was obtained for the structure generated without relaxation, with template mode = pdb100 and num_recycles = 48.

The results of the PRODIGY analysis for the two HADDOCK docking solutions and the best-performing AlphaFold2-Multimer prediction, are summarized in Table 6.

Table 6: Summary of PRODIGY-based binding affinity (ΔG, Kd), interfacial contacts (ICs), and non-interacting surface (NIS) for α-Synuclein/S100B complex models

Model	ΔG (kcal/mol)	Kd (M)	ICs Charged-Charged	ICs Charged-Polar	ICs Charged-Apolar	ICs Polar-Polar	ICs Polar-Apolar	ICs Apolar-Apolar	NIS Charged (%)	NIS Apolar (%)
HADDOCK Model 1	- 8.4	7.0×10^{-7}	7	6	14	1	11	14	42.45	32.37
HADDOCK Model 2	- 8.2	9.6×10^{-7}	4	6	6	2	15	14	44.12	30.15
AlphaFold2 Multimer Best Model	- 14.2	3.8×10^{-11}	12	8	20	2	33	57	44.51	31.1

The first HADDOCK docking model exhibited moderate binding affinity, with ΔG = -8.4 kcal/mol and Kd = 7.0×10^{-7} M. The interfacial contacts were well-balanced, with 14 charged-apolar and 14 apolar-apolar interactions, contributing to a moderately compact interface (NIS: 42.45% charged, 32.37% apolar).

The second HADDOCK docking model showed a comparable binding affinity (ΔG = -8.2 kcal/mol, Kd = 9.6×10^{-7} M), with a minor shift in interfacial contacts, characterized by fewer charged-apolar interactions and an increase in polar-polar and polar-apolar contacts. The NIS values exhibited small variations, indicating subtle differences in interface compactness.

The best AlphaFold2 model displayed a lower ΔG (-14.2 kcal/mol) and a much lower Kd (3.8×10^{-11} M), indicating a much stronger predicted interaction. This model featured a higher number of hydrophobic contacts (57 apolar-apolar, 20 charged-apolar) and a similar degree of interface compactness (NIS: 44.51% charged, 31.1% apolar), suggesting that hydrophobic interactions dominate its predicted binding mode.

4.3.2 Analysis of the Molecular Docking Results from HADDOCK

The docking solutions were evaluated not only based on PRODIGY but also on the HADDOCK score, which accounts for various energy terms, including van der Waals, electrostatic, desolvation and restraint violation energies, providing a comprehensive evaluation of the predicted binding affinities [55].

Table 7: HADDOCK Docking Results Summary

Parameter	HADDOCK Model 1	HADDOCK Model 2
HADDOCK Score	-30.4 ± 17.7	-75.9 ± 4.8
Cluster Size	8	18
RMSD (Å)	0.8 ± 0.6	0.6 ± 0.4
Van der Waals Energy	-27.2 ± 9.1	-23.1 ± 3.3
Electrostatic Energy	-376.7 ± 116.2	-295.3 ± 8.2
Desolvation Energy	4.2 ± 5.4	3.0 ± 1.5
Restraints Violation Energy	680.1 ± 97.4	32.3 ± 29.8
Buried Surface Area	1219.1 ± 162.0	1195.9 ± 49.0
Z-Score	-1.3	-1.8

As shown in Table 7, the first docking model obtained a HADDOCK score of -30.4 ± 17.7, with a small cluster size and higher restraint violation energy, indicating deviations from the docking constraints. Despite strong electrostatic interactions, the higher desolvation energy and moderate buried surface area suggest a less stable interface.

In contrast, the second docking model achieved a significantly lower HADDOCK score (-75.9 ± 4.8), indicating a more favorable binding affinity. It formed a larger cluster, had a lower RMSD, and showed reduced restraint violations, suggesting a more consistent and reliable docking solution. Although the van der Waals contribution was slightly less favorable, the improved electrostatic interactions and lower desolvation energy contributed to better overall stability.

These results suggest that the second docking model provides a more stable and energetically favorable interaction according to HADDOCK scoring criteria, with better agreement with docking constraints and a more well-defined binding interface.

A structural comparison of the two docking models, visualized in Figure 12, highlights the differences in their predicted conformations. The first docking model, shown in pink, and the second docking model, shown in cyan, exhibit variations in the binding orientation of α-synuclein relative to S100B.

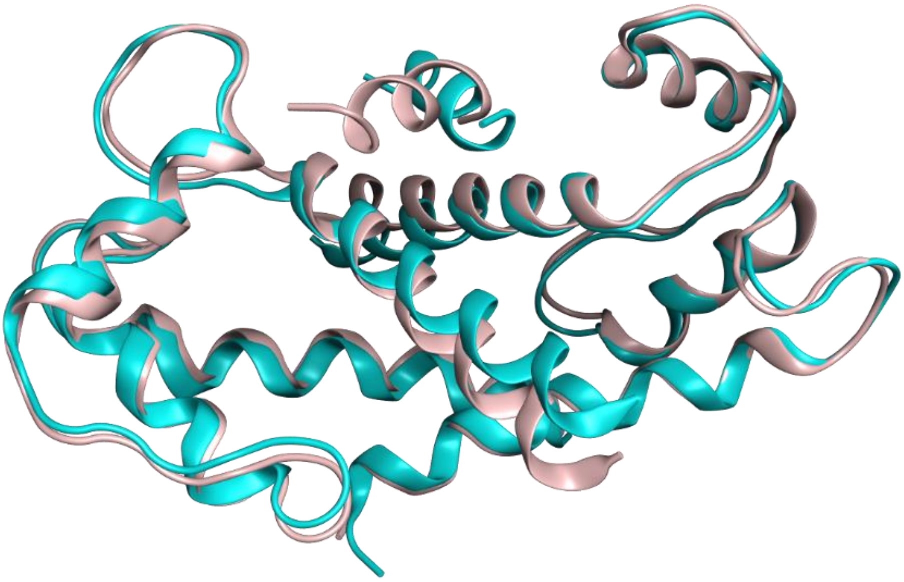


Figure 12: Structural Superposition of the First (Pink) and Second (Cyan) HADDOCK Docking Models

4.3.3 Analysis of the Predicted Complex Structure from AlphaFold2 Multimer

Table 8 reports the main results obtained from the five AlphaFold2-Multimer trials performed for the α-synuclein/S100B complex.

Table 8: Binding free energy and confidence metrics from AlphaFold2-Multimer predictions of the α-synuclein/S100B complex

Trial	*ΔG (kcal/mol)*	*pTM*	*ipTM*	*Average pLDDT*
1	-11.4	0.81	0.74	76.84
2	-14.2	0.80	0.72	77.07
3	-11.4	0.81	0.74	77.84
4	-13.7	0.80	0.72	76.22
5	-11.1	0.80	0.73	76.88

For each trial, the binding free energy (ΔG), calculated with PRODIGY, is provided together with the internal quality indicators from AlphaFold2: the predicted TM-score (pTM), reflecting the overall accuracy of the predicted complex; the interfacial predicted TM-score (ipTM), which focuses on the confidence of the interfacial regions between the interacting proteins; and the average per-residue confidence (pLDDT), expressing the local confidence of the predicted structure. Based on these results, Trial 2 (template mode = pdb100, num_recycles = 48, num_relax = 0) was selected as the reference AlphaFold2 model for further analysis, as it showed the most favorable binding free energy (ΔG = –14.2 kcal/mol) according to PRODIGY, along with consistent confidence indicators from AlphaFold2.

The following analysis focuses on Trial 2, examining in detail the sequence coverage, structural confidence, and interaction reliability of the predicted α-synuclein/S100B complex.

The **Sequence Coverage Plot** shown in Figure 13 provides insight into the quantity and quality of homologous sequences retrieved from databases during the multiple sequence alignment (MSA) process, which directly impacts the reliability of the structural prediction [56]. A well-represented protein in evolutionary databases generally leads to a more accurate model, whereas a lower number of homologous sequences introduces uncertainty. In this plot, the x-axis represents the residue positions, with 0-50 corresponding to α-synuclein and the remaining positions to S100B. The y-axis indicates the number of homologous sequences found in the database. The color gradient reflects sequence identity to the query (blue = high, green-yellow = moderate, red = low), while the black line represents the amount of sequence data available for each residue.

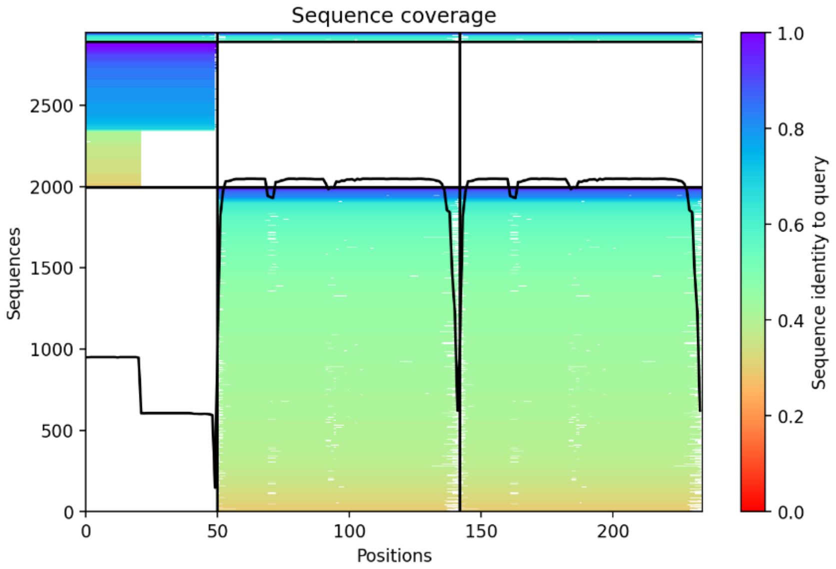


Figure 13: Sequence Coverage Plot, providing insight into the quantity and quality of homologous sequences retrieved from databases during the MSA process.

As for α-synuclein (0-50), the color of some retrieved sequences tends to be blue, indicating a good sequence identity, but the black line drops rapidly, indicating a low number of homologous sequences. This suggests that α-synuclein is poorly represented, resulting in greater uncertainty in its predicted structure. For S100B (50+), the sequence identity is moderate to high (green-blue color) and the black line remains high and stable, indicating a large number of homologous sequences. This strong coverage provides AlphaFold2 with sufficient evolutionary information to generate a more reliable structural model.

Figure 14 presents the **Predicted Aligned Error (PAE) matrices** for the top five ranked models, providing insight into the confidence of residue positioning [57] within the predicted α-synuclein/S100B complex. The x- and y-axes represent residue positions, with α-synuclein (A: 0-50 residues) and S100B monomers (B: 50-150, C: 150-250 residues). The color scale indicates prediction confidence, where blue represents high confidence and red signifies high uncertainty in spatial alignment. The PAE plots reveal a strong contrast between the well-structured S100B dimer and the flexible, uncertain nature of α-synuclein's interaction with S100B.

Regarding Intra-Protein Confidence (Diagonal Regions), diagonal regions corresponding to the S100B dimer are predominantly blue, confirming high confidence in the folding and dimerization of S100B. In contrast, the A-A diagonal region exhibits a red zone, reinforcing the difficulty in predicting a stable conformation for this protein. This aligns with α-synuclein's intrinsically disordered nature.

For Inter-Protein Confidence (Off-Diagonal Regions), the A-B and A-C regions, representing the α-synuclein/S100B interaction interface, contain significant red areas, suggesting high uncertainty in the predicted binding mode. This aligns with the low sequence coverage for α-synuclein, previously observed in the Sequence Coverage Plot, indicating that AlphaFold2 lacks sufficient evolutionary data to confidently define its interaction with S100B. In contrast, the B-C region remains blue, confirming high inter-protein confidence in the dimerization of S100B.

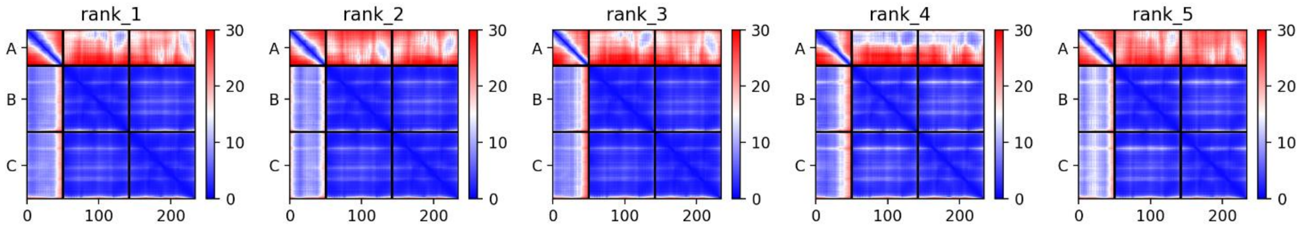


Figure 14: Predicted Aligned Error (PAE) Matrices for the Top 5 Ranked Models, showing confidence in residue positioning within the α-synuclein/S100B complex across five independent predictions.

Figure 15A presents the **Predicted Local Distance Difference Test (pLDDT)** per position, which assesses the local confidence [56] of the predicted α-synuclein/S100B complex across the top five ranked models. This information is further complemented by Figure 15B, which shows the 3D structure of the Rank 1 model, color-coded based on pLDDT scores.

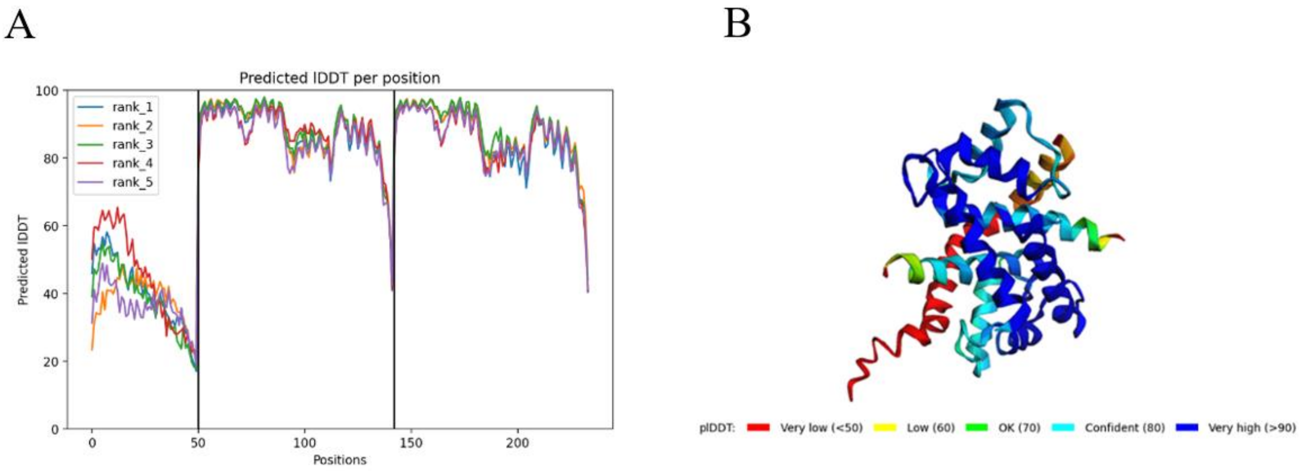


Figure 15: (A). Predicted Local Distance Difference Test (pLDDT) per Position for the Top 5 Ranked Model. (B). 3D Structural Representation of the Rank 1 Model, color-coded by pLDDT scores to visualize confidence levels in different regions of the α-synuclein/S100B complex.

In Figure 15A, the x-axis represents residue positions, while the y-axis shows the pLDDT score, ranging from 0 to 100, where higher values indicate greater structural confidence.

The first 50 residues of α-syn exhibit consistently low pLDDT scores (<60) across all ranked models, indicating high structural variability and low confidence in the predicted conformation. This is further highlighted b*y* **Error! Reference source not found.**B, where the N-terminal region of α-synuclein appears predominantly red. In contrast, S100B displays consistently high pLDDT scores (80-90), indicating strong confidence in its structural integrity. This observation aligns with findings from the PAE analysis.

As observed, there is a slight decrease in pLDDT values at the α-synuclein/S100B interaction interface, particularly near the 50-residue boundary, which corresponds to high-uncertainty regions in the PAE matrices. This suggests that the binding mode of α-synuclein may not be rigidly defined, and that this protein could interact with S100B in a dynamic and transient manner. Therefore, molecular dynamics simulations are needed to determine whether this interaction stabilizes over time.

4.4 Discussion

In this chapter, since no experimental structures of the α-synuclein/S100B complex are currently available, two complementary computational strategies were applied to explore plausible interaction modes between the two proteins: molecular docking with HADDOCK and structure prediction with AlphaFold2-Multimer. These methods were used to build reliable structural models of the complex, providing starting points for subsequent molecular dynamics (MD) simulations.

For the docking approach, active residues were selected based on sequence and structural comparisons with homologous systems reported in the literature, particularly considering studies on the interactions between α-synuclein and S100A9, and between S100B and β-amyloid. Among the two docking solutions obtained, the second was identified as the most reliable, as it achieved a significantly better HADDOCK score, indicating a more favorable binding mode, despite comparable PRODIGY-predicted binding energies between the two models.

Regarding AlphaFold2-Multimer, multiple predictions were generated by varying key parameters. The model built without relaxation, using pdb100 templates and 48 recycles, was selected as the most promising, showing the lowest binding free energy according to PRODIGY.

In conclusion, this chapter has provided a comprehensive computational characterization of the α-synuclein/S100B complex through two distinct modeling strategies. The resulting structures represent plausible interaction models that, in the absence of experimental data, offer a solid foundation for subsequent molecular dynamics simulations, which will be essential to investigate stability and validate the biological relevance of the predicted interactions.

5 Molecular Dynamics Simulations of the α-Synuclein/S100B Complexes for Stability and Interaction Analysis

5.1 Introduction

As discussed in the previous chapter, the structural models of the α-synuclein/S100B complex were obtained through two distinct computational strategies: molecular docking with HADDOCK and artificial intelligence-based prediction with AlphaFold2-Multimer. In the case of HADDOCK, the absence of experimental data on the interaction interface required the definition of active residues based on analogy with previously characterized systems. Specifically, residues were selected by comparing the sequences of α-synuclein and β-amyloid, as well as S100B and S100A9, assuming that homologous regions involved in similar amyloid interactions might also contribute to α-synuclein/S100B association. In contrast, AlphaFold2-Multimer predicted the complex structure de novo, without the need for prior information about the binding site.

To evaluate the dynamic behavior of these models and investigate their interaction characteristics, molecular dynamics (MD) simulations were performed under the same conditions for both complexes. The purpose of this phase was to track the time evolution of the predicted structures and determine whether the binding interfaces remained stable throughout the simulations.

By comparing the results obtained from the molecular dynamics simulations of the models generated via HADDOCK and AlphaFold2-Multimer, this chapter aims to highlight both similarities and differences between the two predictions and to investigate whether the independently generated complexes exhibit consistent interaction patterns, thereby increasing confidence in the plausibility of the proposed binding modes. Particular attention was also given to identifying the residues involved in the interaction and evaluating their behavior throughout the simulations, comparing these findings with the hypotheses initially formulated during the modeling phase. Together, these observations provide a basis for future experimental validation and for exploring the potential biological significance of the α-synuclein/S100B interaction.

5.2 Materials and methods

This section describes the molecular dynamics (MD) protocol employed to investigate the stability and interaction properties of the α-synuclein/S100B complexes generated by HADDOCK and AlphaFold2-Multimer. All simulations were performed under identical conditions to allow a direct comparison between the two models.

The setup of the systems and the molecular mechanics settings adopted for the simulations are first described. This is followed by the presentation of the computational analyses performed on the MD trajectories, with a focus on structural stability, interaction interfaces, and binding energetics.

5.2.1 Molecular Mechanics/Molecular Dynamics Settings

All molecular dynamics (MD) simulations were carried out using GROMACS v.2022.5, and employing a in house modified version of the Amber19SB force field [58].

As a preparatory step for the HADDOCK-derived complex, it was necessary to reconstruct the α-synuclein fragment comprising residues 1–50, since the initial docking calculations had been performed using only the 31–41 region of the protein. The missing N-terminal residues (1–30) and C-terminal residues (42–50) of the α-syn1–50 fragment were modeled using PyMOL, adopting an α-helical conformation consistent with structural insights described in experimental studies of the N-terminal region of full-length α-synuclein [7]. Conversely, in the AlphaFold2-Multimer complex, the α-synuclein structure already included the entire 1–50 fragment; therefore, no modifications were required prior to the simulation setup.

Each complex was placed in a dodecahedral periodic simulation box, ensuring a minimum distance of 1.2 nm between the protein surface and the limits of the box. Solvation was performed using the TIP4P water model and, to neutralize the system's net positive charge of +5e, a total of five chloride ions (Cl^-) were added.

The systems were energy minimized in two sequential stages, each consisting of 30,000 steps, using the steepest descent algorithm. In the first stage, constraints were applied only to hydrogen bonds, while in the second stage all bonds were constrained using the LINCS algorithm. The SETTLE algorithm was used to constrain all water molecules.

After minimization, systems were equilibrated in the NVT and then NPT ensembles. This procedure was conducted following a three-step protocol, designed to gradually relax the system while preserving the integrity of the complex. The initial step consisted of a 100 ps NVT MD simulation with initial velocities generated from a Maxwell velocity distribution at 310 K. Temperature coupling was handled using a modified Berendsen thermostat (v-rescale) with independent coupling applied to solute and solvent groups and a relaxation time of 0.1 ps. During this phase, position restraints of 1000 kJ mol^{-1} nm^{-2} were applied to the Cα atoms of both α-synuclein and S100B. This step was followed by two consecutive 200 ps NPT simulations to stabilize the system pressure: the first with positional restraints reduced to 100 kJ mol^{-1} nm^{-2}, and the second with weaker restraints of 10 kJ mol^{-1} nm^{-2}. Throughout these stages, isotropic pressure coupling was applied using the Parrinello–Rahman barostat, keeping the pressure at 1 bar, with an isothermal compressibility of 4.5×10^{-5} bar^{-1} and a relaxation time of 0.5 ps

A pre-equilibration MD simulation of 200 ns was performed to refine the interaction interface, with distance restraints applied to selected residues of α-synuclein and S100B. For the HADDOCK-derived model, the restrained residues corresponded to the active sites specified during the docking setup, targeting the Cα atoms of the selected residues from both proteins. In the AlphaFold2-Multimer complex, active residues were identified based on the analysis of Predicted Aligned Error (PAE) matrices, selecting regions with the lowest interfacial prediction error to define the most reliable contact sites and applying distance restraints to their Cα atoms as well.

At the end of the pre-equilibration phase, the replica corresponding to the frame at 100 ns was extracted from the trajectory and used as the starting configuration for a 500 ns production MD simulation.

The MD simulations were performed using the leap-frog integration scheme with a time step of 1 fs, applied consistently throughout both equilibration and production phases. Periodic boundary conditions were imposed in all directions (xyz). Long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) method, with a cutoff distance of 1.2 nm. Both Lennard–Jones and Coulomb interactions were explicitly calculated up to 1.2 nm. During the production MD simulations, system coordinates were recorded every 100 ps, yielding a total of 5000 frames over the 500 ns trajectory.

5.2.2 Computational Analysis

To examine the stability and interaction characteristics of the α-synuclein/S100B complexes, a series of computational analyses was conducted on the molecular dynamics (MD) trajectories. All procedures were applied to both systems to facilitate direct comparison.

The first step of the analysis was to assess the structural stability of the complexes by calculating the root-mean-square deviation (RMSD) of the Cα atoms over time. RMSD was computed for the entire complex, S100B and α-synuclein separately, allowing the evaluation of the global conformational changes in the system, as well as the independent behavior of each protein during the simulation. These calculations were performed over the entire trajectory to verify whether the system reached equilibrium [59]. Ensuring equilibrium was essential so that some of the subsequent analyses were conducted on stable portions of the simulation. The flexibility of the complexes was evaluated by calculating the root-mean-square fluctuation (RMSF) of each residue within α-synuclein and S100B. RMSF was computed over the entire 500 ns after fitting the structures to the Cα atoms of the complex, allowing the identification of the most mobile regions and highlighting possible differences in fluctuations between the two proteins.

To investigate the conformational dynamics of the complexes, the secondary structure evolution of α-synuclein and S100B was analyzed using the DSSP method (Dictionary of Secondary Structure of Proteins). The DSSP program automatically assigns the secondary structure of each residue to one of the predefined states based on the protein's 3D coordinates [60]. The simplified three-state classification (helix, strand, coil) was applied to monitor possible structural rearrangements and the persistence of secondary structure elements throughout the trajectory.

To characterize the interaction interface between α-synuclein and S100B, solvent-accessible surface area (SASA) calculations were performed on the last 100 ns of the trajectories, offering a measure of the interface's solvent accessibility [61]. This analysis allowed the quantification of the buried surface area (BSA) between α-synuclein and S100B using the following formula:

$$BSA = \frac{SASA_{\alpha-synuclein} + SASA_{S100B} - SASA_{complex}}{2} \tag{13}$$

where $SASA_{\alpha-synuclein}$ and $SASA_{S100B}$ denote the solvent-accessible surface areas of the individual proteins calculated within the context of the complex, while $SASA_{complex}$ refers to the total SASA of the full complex. This approach allowed for the quantification of the size of the interface in the protein– protein complex, with larger BSA values indicative of more extensive interfacial interactions [62]. To compare the two complexes, the mean and variance of the BSA were computed over the analysis window.

To identify which residues are most involved in the interaction, contact frequency analyses were performed for both S100B and α-synuclein. For each residue, the percentage of frames in which at least one atom was within a cutoff distance of 0.35 nm from the partner protein was calculated over the last 100 ns of the trajectory. The contact percentage was obtained by dividing the number of frames in which contact occurred by the total number of frames in the analysis window. This method enabled the detection of the residues remaining in contact for a significant portion of the simulation, suggesting their relevance in complex stabilization.

Finally, the binding free energy between α-synuclein and S100B was estimated through the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method. This technique combines molecular mechanics energy terms with solvation effects modeled through a generalized Born approximation and non-polar contributions derived from solvent-accessible surface area [63]. The analysis was carried out by dividing the trajectory into multiple equally spaced sub-windows, in which the binding free energy was calculated independently and then averaged to obtain a final estimate. Additionally, a per-residue energy decomposition was performed to determine each residue's specific energetic impact on the overall binding energy, thereby identifying key residues that favor complex stability [64].

5.2.3 Plots and Figures

Three-dimensional molecular representations of the α-synuclein/S100B complexes were generated using Visual Molecular Dynamics (VMD). All structural and energetic analyses, including RMSD, RMSF, solvent-accessible surface area (SASA), secondary structure content, and contact frequencies, were graphically rendered using Python scripts based on Matplotlib [65].

5.3 Results

5.3.1 Structural Stability Analysis

The RMSD of the Cα atoms was monitored over time for the entire complex (Figure 16A), α-synuclein (Figure 16B) and S100B separately (Figure 16C). The results highlight stability differences between the HADDOCK and AlphaFold2 models.

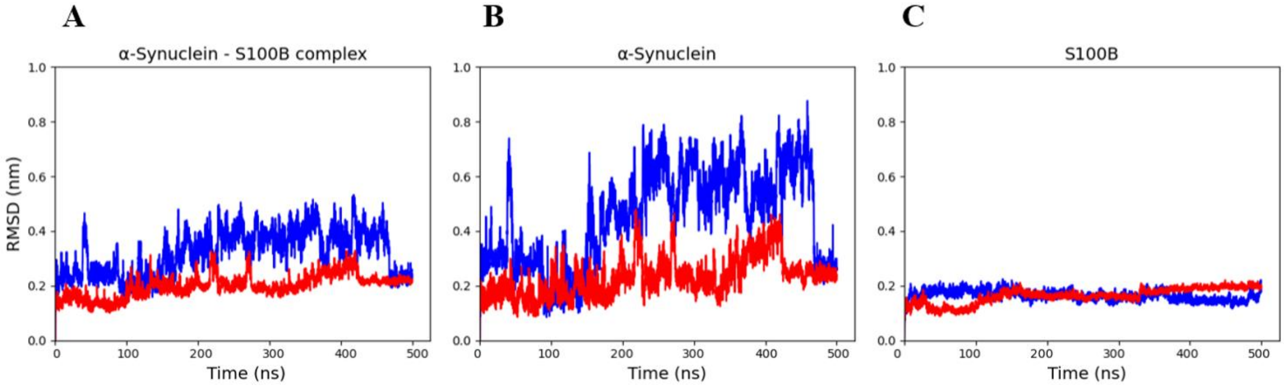


Figure 16: RMSD Comparison Between HADDOCK (blue) and AlphaFold (red) Models. (A). RMSD of Cα atoms in the α-synuclein/S100B complex. (B). RMSD of Cα atoms in α-synuclein. (C). RMSD of Cα atoms in S100B.

As for the α-synuclein/S100B complex, the AlphaFold2-predicted model exhibits a lower and more stable RMSD (~0.2–0.3 nm) throughout the 500 ns simulation , while the HADDOCK-derived model shows greater fluctuations and higher RMSD values. A deeper analysis of the individual components highlights that α-synuclein contributes most to these differences rather than S100B. The RMSD of S100B remains low (~0.2 nm) and stable in both models, indicating that this protein retains its conformation throughout the simulation. In contrast, α-synuclein exhibits notable fluctuations, particularly in the HADDOCK model, where its RMSD reaches up to 0.6 nm. This suggests a greater deviation from its initial configuration, and more importantly, a lack of convergence to a stable RMSD value, indicating that its interaction with S100B remains dynamic and less constrained.

To ensure that the following analyses focus on stable portions of the simulation, the final 100 ns of each trajectory were selected to compute average properties. This approach reduces the impact of early transient structural adjustments.

The Root Mean Square Fluctuation (RMSF) analysis was conducted over the full 500 ns simulation to assess the flexibility of individual residues of α-synuclein (Figure 17A) and S100B (Figure 17B).

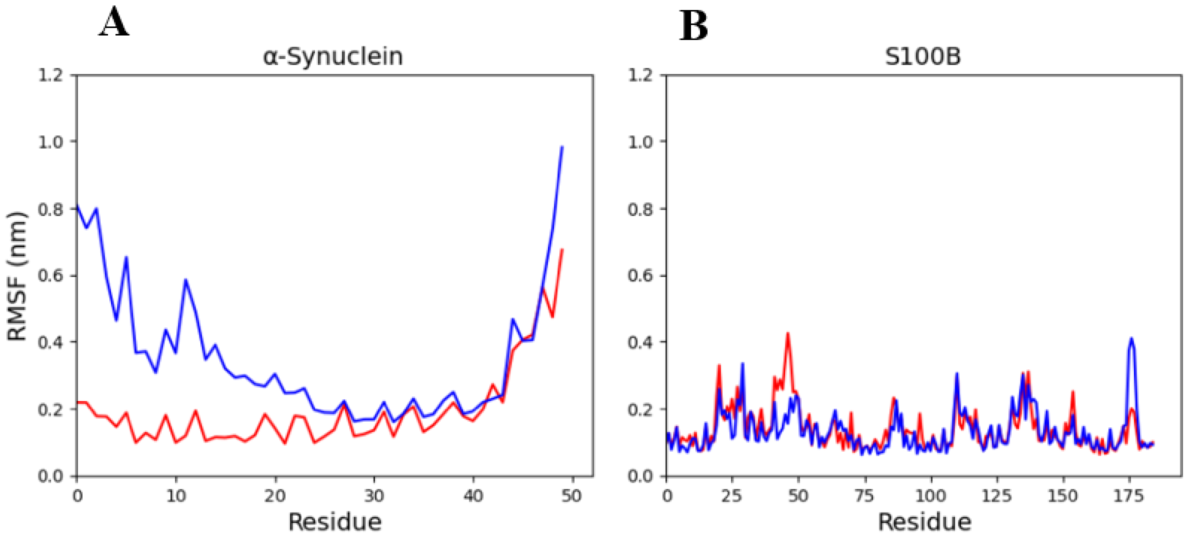


Figure 17: RMSF Comparison Between HADDOCK (blue) and AlphaFold (red) Models. (A). RMSF of α-Synuclein (B). RMSF of S100B.

The results indicate that the overall fluctuation patterns are similar between the two models, with some localized differences. In the case of α-synuclein, the N-terminal residues display greater fluctuations in the HADDOCK model, reaching RMSF values up to ~0.8 nm, whereas in the AlphaFold2 model these fluctuations remain lower (~0.1–0.2 nm). This suggests that in the HADDOCK complex, the N-terminal segment of α-synuclein remains highly flexible and undergoes larger conformational variations. For S100B, its C-terminal region displays slightly higher fluctuations in the HADDOCK model compared to AlphaFold2. For the rest, the fluctuation patterns are almost identical. These findings align with the RMSD results and further support the notion that α-synuclein is the main source of structural differences between the two models.

The secondary structure evolution of α-synuclein and S100B over the full 500 ns simulation was examined for both the HADDOCK and the AlphaFold model (Figure 18).

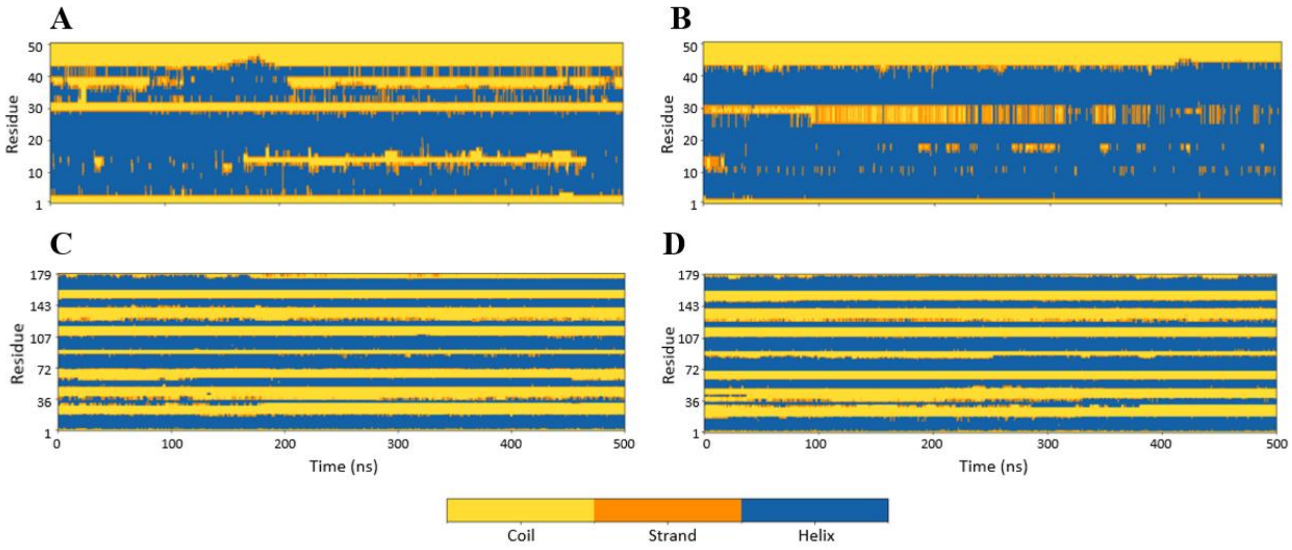


Figure 18: Evolution of the Secondary Structure over time. (A) Secondary structure of α-synuclein in the HADDOCK model. (B) Secondary structure of α-synuclein in the AlphaFoldl model. (C) Secondary structure of S100B in the HADDOCK model. (D) Secondary structure of S100B in the AlphaFold model.

For S100B, the structure remains highly stable in both models. Its α-helical regions are well-maintained throughout the entire simulation, confirming the high structural stability of S100B.

In contrast, α-synuclein exhibits structural differences between the two models. In the AlphaFold2 model, the protein retains a higher fraction of α-helix, suggesting that its interaction with S100B stabilizes this conformation. This helical stabilization could also result from the distance restraints applied to the identified interaction sites, comprising residues 5–14 and 20, during the pre-equilibration phase, potentially reinforcing the structural integrity of these regions in the final simulation. Instead, in the HADDOCK model, the N-terminal region initially loses its helical structure before recovering it later in the simulation.

The secondary structure in the HADDOCK model appears more variable, with a tendency to lose and reform helical segments over time. Notably, α-synuclein in the HADDOCK model was initially reconstructed in PyMOL with an imposed α-helical conformation. However, during the simulation, the secondary structure shows variability, particularly in the C-terminal portion of the α-synuclein fragment (residues 31–41), where the initial α-helix transitions into a coil conformation.

In conclusion, while α-synuclein in the AlphaFold2 complex exhibits an increase in α-helical content, in the HADDOCK model, the C-terminal portion tends to lose its helical structure, especially in the region including residues 31–41.

5.3.2 Interaction Interface and Contact Analysis

The Buried Surface Area (BSA) analysis was performed to measure the extent of the contact interface between α-synuclein and S100B in both models (Figure 19).

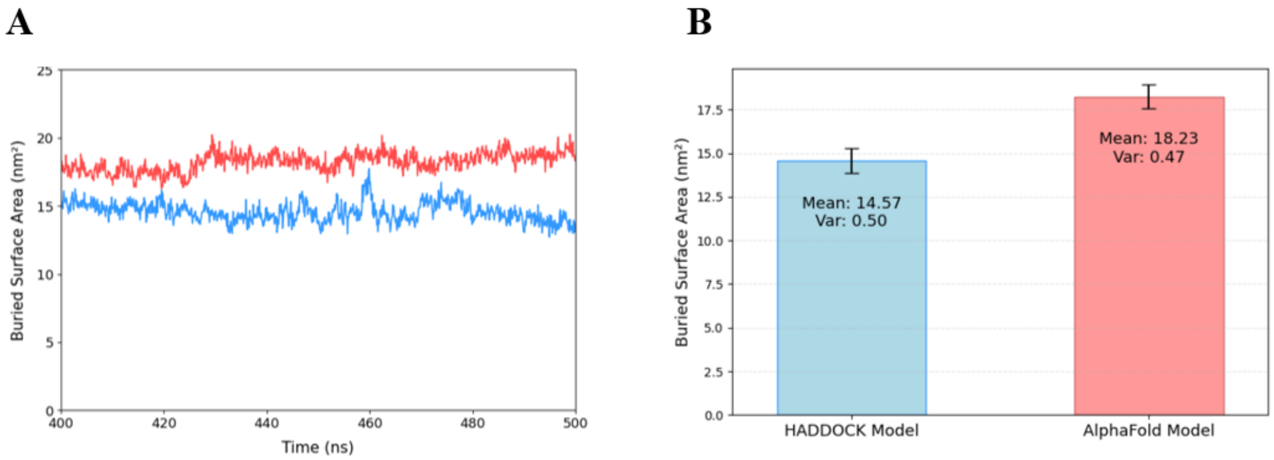


Figure 19: Comparison of Buried Surface Area (BSA) Between HADDOCK (blue) and AlphaFold2 (red) Models. (A). Time evolution of the buried surface area over the last 100 ns of the simulation. (B). Mean and variance of the BSA for both models.

The time evolution plot (Figure 19A) shows that the BSA of the α-synuclein/S100B complex remains stable over the last 100 ns of the simulation, without significant fluctuations. This suggests that, once the equilibrium is achieved, the interface remains largely unchanged. Figure 19B provides a quantitative comparison of the average BSA and its variance for the two complexes. The AlphaFold model exhibits a higher average BSA (18.23 nm²) related to the HADDOCK model (14.57 nm²). Regarding BSA variance, the two models present comparable values.

The contact frequency analysis provides further insights into the interaction interface between α-synuclein and S100B, highlighting residues that maintain persistent interactions across the last 100 ns of the simulation (Figure 20, Figure 22).

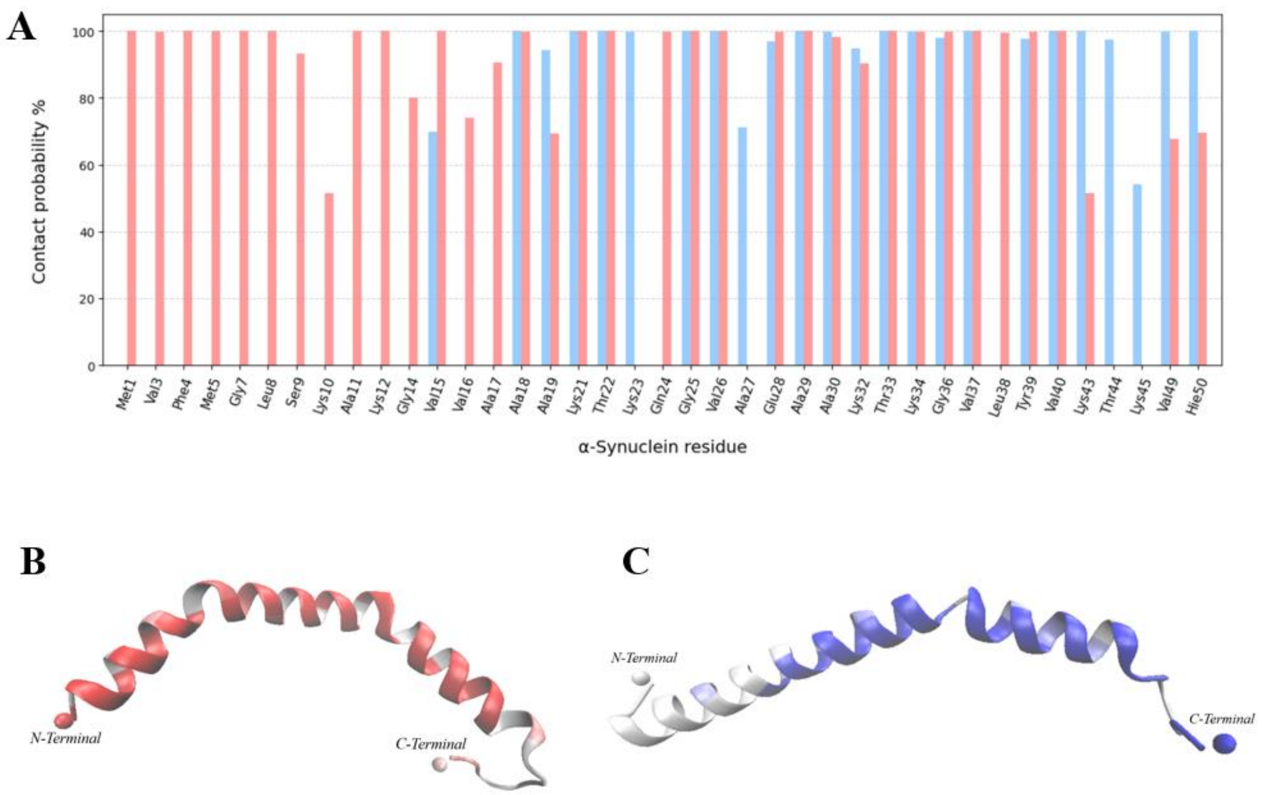


Figure 20: Contact probability of α-synuclein residues. (A) Contact probability of each α-synuclein residue with S100B over the last 100 ns of simulation for both HADDOCK (blue) and AlphaFold (red). (B) Structural representation of α-synuclein in the AlphaFold model, colored according to contact probability (white to red). (C) Structural representation of α-synuclein in the HADDOCK model, colored according to contact probability (white to blue).

In Figure 20A, the plot reports, for each model, the percentage of simulation frames where each α-synuclein residue is in contact with S100B, considering only residues with a contact frequency exceeding 50%. The results reveal both similarities and differences in the interaction patterns, which can also be visually examined in Figure 20B and Figure 20C, where contact probability values are mapped onto the α-synuclein structure.

Notably, several residues exhibit a high contact frequency in both models, suggesting a common interaction interface. Among these, *Ala18, Ala21, Thr22, Gly25, Gly26, Ala29, Ala33, Ala34, Val37* and *Val40* stand out, maintaining contact with S100B for 100% of the last 100 ns of the simulation in both cases. This agreement between models reinforces the likelihood that these residues play a significant role in the α-synuclein/S100B binding.

However, some differences also emerge in the interaction patterns. In the AlphaFold model, the contact frequency is high for Met1, Val3, Phe4, Met5, Gly7, and Leu8, as well as for residues spanning Ser9–Ala17, whereas in the HADDOCK model, interactions in this region are almost entirely absent. Conversely, in the HADDOCK model, additional contacts or higher contact frequencies are observed in the C-terminal region, particularly around Thr43, Thr44, Val45, Val49 and His50. This suggests that while both models agree on certain key binding residues, primarily located in the central region, the AlphaFold model predicts a broader interaction interface, incorporating more extensive N-terminal contacts.

Figure 21 highlights the contact probability of the α31–41 segment in the AlphaFold (Figure 21A) and HADDOCK (Figure 21B) models. In the HADDOCK model, most residues within the 31–41 segment, defined as active during docking, remained in contact with S100B. The high contact probability in this region for both models supports the initial hypothesis used in HADDOCK, suggesting that these residues are likely to be key binding residues.

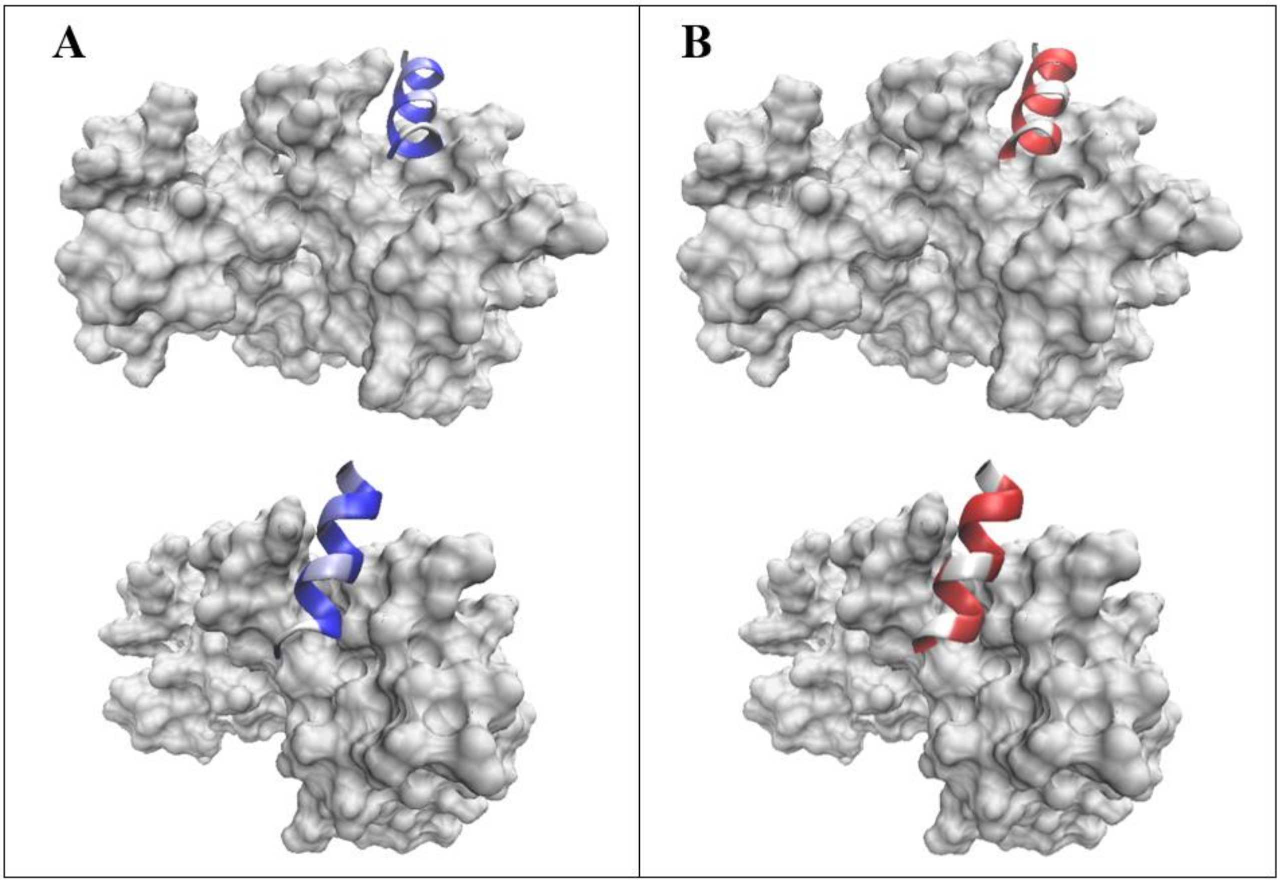


Figure 21: Contact Probability Mapping for α $_{31–41}$ in the α-Synuclein/S100B Complex. (A) AlphaFold model, where residues 31–41 of α-synuclein are colored according to contact probability (white to red). (B) HADDOCK model, where the same residues are colored on a white-to-blue scale.

S100B residues exhibiting a contact frequency greater than 50% with α-synuclein in at least one of the two models are reported in Figure 22A.

Comparing the curves, it is evident that in the AlphaFold model, S100B engages a more extensive interaction interface than in HADDOCK. The red bars, corresponding to AlphaFold, maintain higher values along most of the x-axis, indicating that AlphaFold predicts a higher contact frequency. For instance, numerous red peaks are observed in the central and C-terminal regions of the second S100B monomer (approximately spanning residues Ser131–Val146 and Cys174–Phe178), where the contact probability in the HADDOCK model is completely absent for these residues. Conversely, only a few S100B residues, such as Ile47, Glu58, and Gln71, simultaneously exhibit both a high contact probability in the HADDOCK model and the absence of interactions in the AlphaFold model. While differences are observed, several residues exhibit a high contact frequency in both models. Among them, residues such as Val52, Thr59, Ala75, and Met79 maintain a contact probability of 100% in both cases. This consistency suggests that HADDOCK and AlphaFold reliably predict these S100B sites as key interaction points with α-synuclein.

The results presented in Figure 22A are further supported by the structural representations in Figure 22B and Figure 22C, where the contact probability is mapped onto the S100B structure.

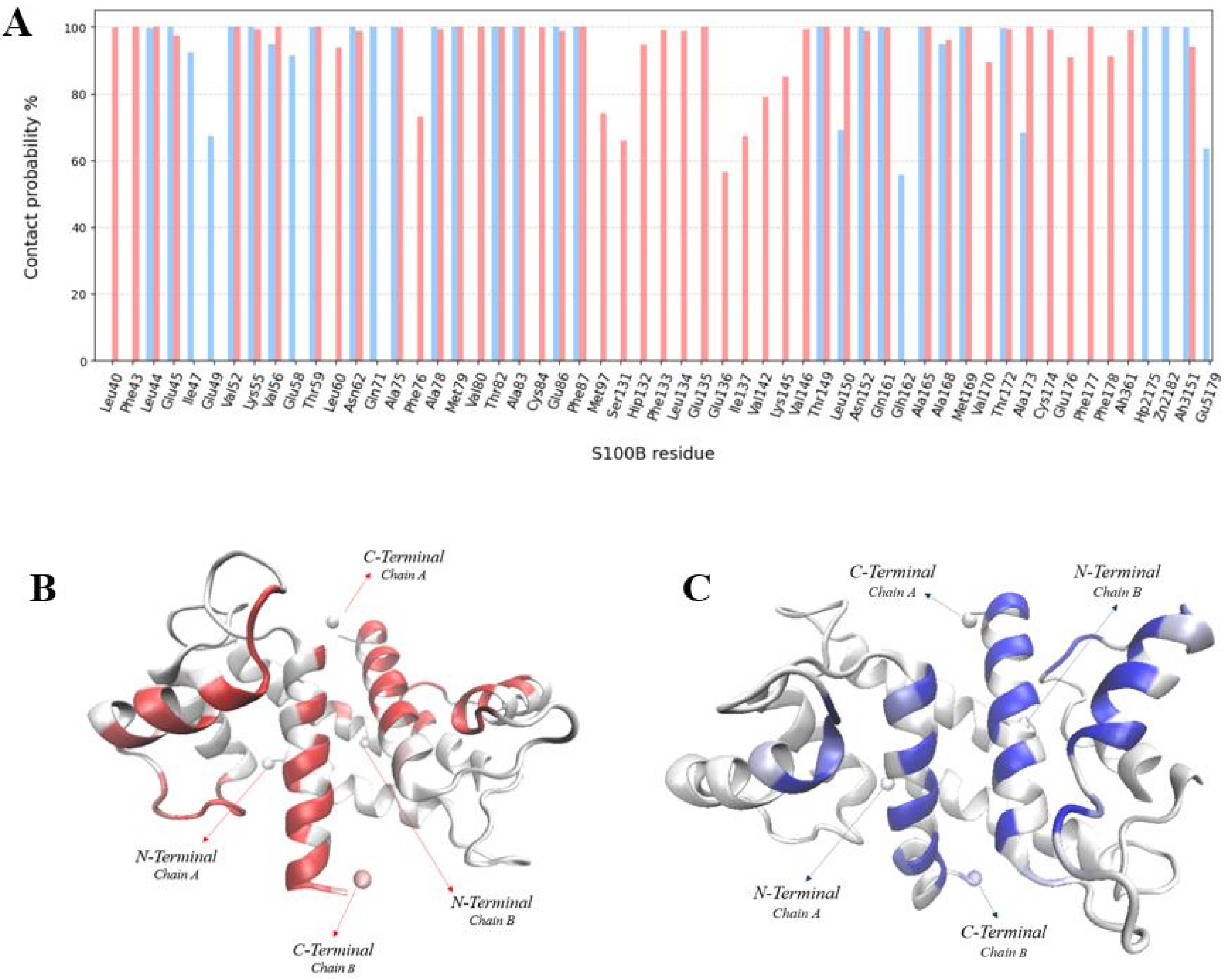


Figure 22: Contact probability of S100B residues. (A) Contact probability of each S100B residue with α-synuclein over the last 100 ns of simulation for both HADDOCK (blue) and AlphaFold (red). (B) Structural representation of S100B in the AlphaFold model, colored according to contact probability (white to red). (C) Structural representation of S100B in the HADDOCK model, colored according to contact probability (white to blue).

Additionally, when comparing the S100B active residues used as input for docking with the residues identified in the HADDOCK contact frequency analysis (Figure 22A), only 20% of the initially hypothesized residues remained in contact with α-synuclein for at least 50% of the frames in the last 100 ns of the simulation. This suggests that, while some of the predefined active residues maintained stable interactions, a significant fraction did not, indicating possible deviations from the initial docking constraints during the molecular dynamics simulations.

5.3.3 Binding Energy Analysis and Per-Residue Contributions

The binding energy analysis was performed using the MM/GBSA approach to evaluate the energetic contributions stabilizing the α-synuclein/S100B complex in both the HADDOCK and AlphaFold models (Figure 23).

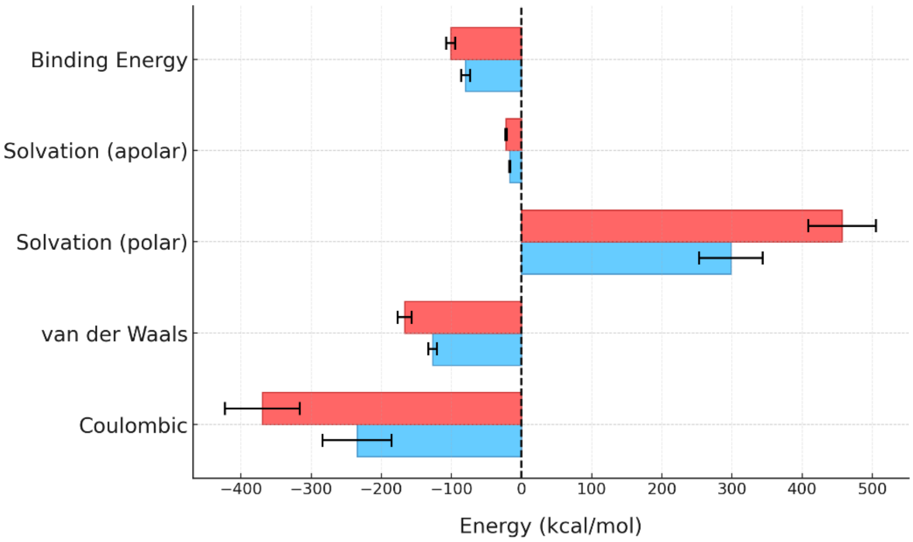


Figure 23: Decomposition of the Binding Energy Components for the α-Synuclein/S100B Complex. in HADDOCK (blue) and AlphaFold (red) Models.

The results confirm that interaction is energetically favorable in both models, with negative binding energy values. For AlphaFold, the average binding energy over the last 100 ns of simulation is -101.11 kcal/mol, while for HADDOCK, it is -80.12 kcal/mol. In both cases, Coulombic interactions provide the strongest stabilization, particularly in the AlphaFold model, where this term reaches approximately -370 kcal/mol, indicating a slightly more pronounced electrostatic contribution compared to HADDOCK. Van der Waals interactions also contribute significantly to binding stabilization, with slightly more favorable values in the AlphaFold model. Meanwhile, polar solvation effects oppose binding, with the polar solvation term being more unfavorable in AlphaFold, suggesting a higher desolvation cost to sustain the interaction. In contrast, the apolar solvation component provides a small stabilizing contribution, though its impact remains negligible compared to the other energetic terms.

The analysis of the binding free energy decomposition for α-synuclein residues over the last 100 ns of the simulation is presented in Figure 24, which displays only the residues with the most significant stabilizing contributions. *Met1, Asp2, Glu13, Glu20, Gln24, Glu28, Glu35, Glu46* and *Hie50* emerge as the most energetically favorable residues, maintaining strong stabilizing contributions in both HADDOCK and AlphaFold. In particular, *Met1* exhibits the lowest energy value in both models, with a slightly stronger effect in HADDOCK (-84.07 kcal/mol) compared to AlphaFold (-79.54 kcal/mol). Overall, these results indicate that both HADDOCK and AlphaFold predict a similar set of key stabilizing residues.

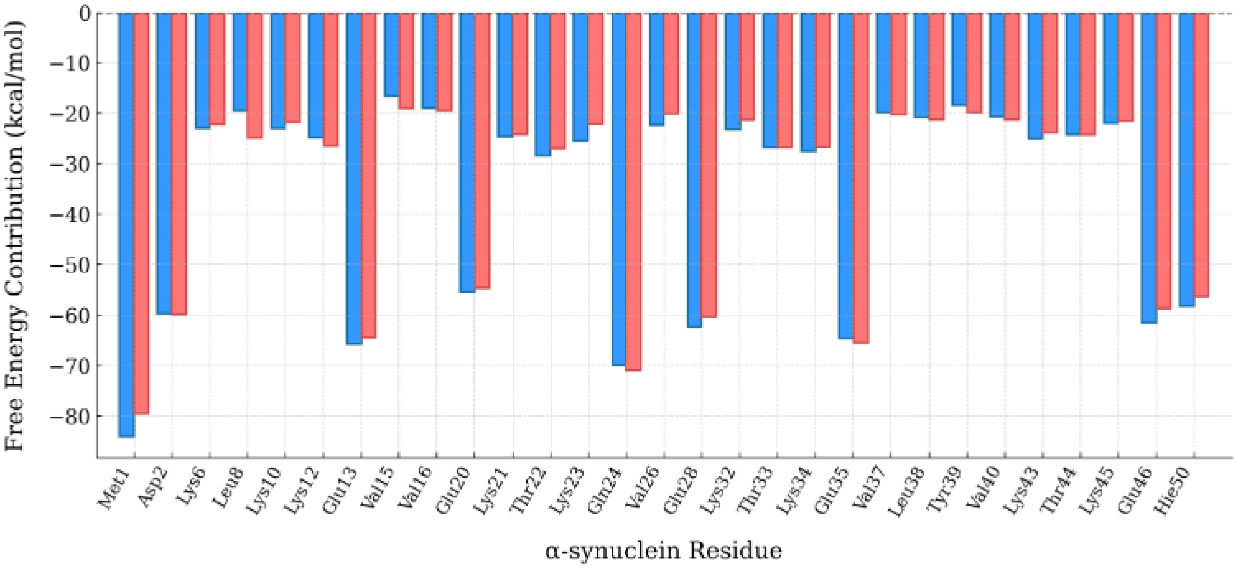


Figure 24: Per-Residue Decomposition of Binding Free Energy for α-Synuclein in the HADDOCK (blue) and AlphaFold (red) Models.

A corresponding per-residue free energy decomposition was conducted for S100B over the last 100 ns of the simulation, as shown in Figure 25. As observed for α-synuclein, S100B residues also exhibit largely comparable energetic contributions between the two models, with some residues standing out for their pronounced stabilizing role. Notably, Arg20 and Arg110 emerge as the most favorable contributors in both HADDOCK and AlphaFold, suggesting a key role in the interaction.

Figure 25: Per-Residue Decomposition of Binding Free Energy for S100B in the HADDOCK (blue) and AlphaFold (red) Models.

While Figure 24 and Figure 25 focused on residues with negative free energy contributions, which contribute to the stabilization of the α-synuclein/S100B complex, Figure 26 highlights residues that exhibit a positive free energy contribution in at least one of the two models. A positive value indicates a destabilizing effect on the complex, meaning that these residues oppose binding rather than favoring it. Figure 26A illustrates the per-residue decomposition of the free energy contributions for α-synuclein, where Phe4 and Met5 show a destabilizing effect in at least one model. Similarly, Figure 26B presents the S100B residues with at least one positive contribution, highlighting positions such as His42, Ile47, Cys68, Cys84, and Ile137, which contribute unfavorably to binding in both HADDOCK and AlphaFold, with comparable free energy values across the models.

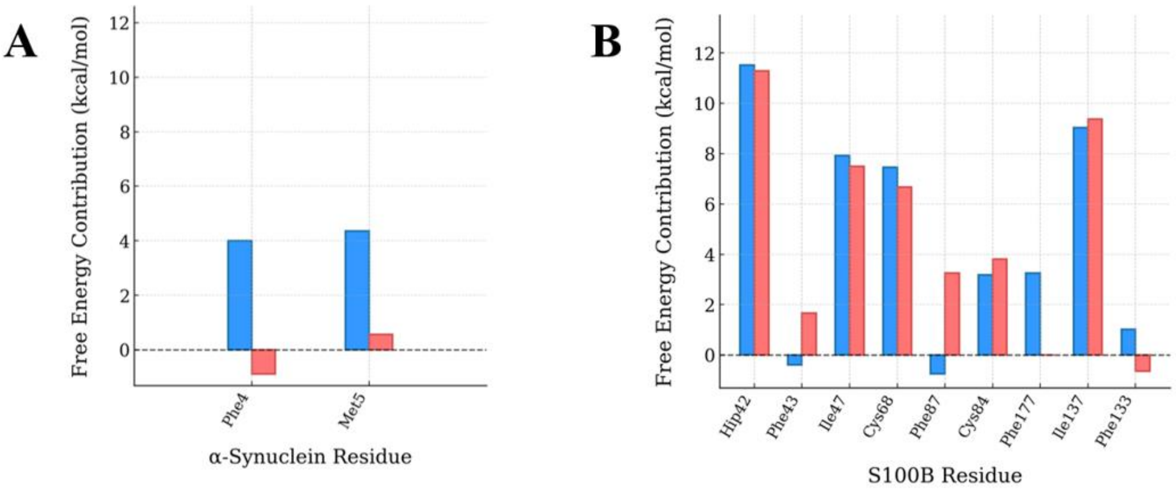


Figure 26: Residues with Positive Free Energy Contribution in α-Synuclein and S100B. (A) Per-residue decomposition of binding free energy for α-synuclein, displaying residues with at least one positive contribution in either HADDOCK (blue) or AlphaFold (red). (B) Per-residue decomposition of binding free energy for S100B, highlighting residues with at least one positive contribution in either HADDOCK (blue) or AlphaFold (red).

5.4 Discussion

In this chapter, the structural models generated by AlphaFold2-Multimer and HADDOCK were employed as a starting point for molecular dynamics simulations to explore the interaction dynamics between α-synuclein and S100B. The simulation results were then analyzed with the objective of assessing the stability of the complexes over time, their conformational changes and the molecular interactions occurring at the interface. Despite the differences in how these models were generated, the results demonstrated a good degree of consistency in the identified interaction patterns.

The RMSD analysis showed that both complexes eventually reached equilibrium after an initial phase of structural rearrangement. Although the AlphaFold2-Multimer model maintained a more constrained conformation compared to the HADDOCK model, neither complex underwent dissociation. Both models agreed that complex flexibility was mainly attributed to α-synuclein, which displayed greater conformational dynamics than S100B, especially in the HADDOCK model. This behavior aligns with its intrinsically disordered nature, as under physiological conditions, α-synuclein does not adopt a well-defined secondary or tertiary structure but exists as a highly dynamic ensemble of conformations. Notably, the N-terminal region of the 1-50 fragment of α-synuclein in the HADDOCK model exhibited higher fluctuations compared to the AlphaFold2-Multimer model, which displayed more restricted motion in this region. Additionally, some fluctuations were observed in the C-terminal region of α-synuclein in both models. In contrast, S100B remained highly stable throughout the simulations.

The analysis of secondary structure further supports these observations. S100B remained highly stable in both models, with well-preserved α-helical regions throughout the simulation. In contrast, α-synuclein exhibited model-dependent structural variability. In the AlphaFold2-Multimer model, a higher proportion of α-helix was retained. In the HADDOCK model, however, the N-terminal α-helix was partially lost before reforming later in the simulation, while the C-terminal region showed a greater tendency to transition into a coil conformation. These differences further highlight the dynamic nature of α-synuclein and its sensitivity to the initial structural modeling approach.

The Buried Surface Area (BSA) remained stable over the final 100 ns of the simulation in both models, indicating that once equilibrium was reached, the interaction interface exhibited no significant changes. However, the AlphaFold2-Multimer model showed a higher average BSA than the HADDOCK model, suggesting a more extensive interaction surface. This is further supported by contact probability analysis, which revealed that a greater number of residues remained in contact with the binding partner for at least 50% of the final phase of the simulation in the AlphaFold model, both for S100B and α-synuclein. This analysis also highlighted both conserved and distinct binding regions between the two models. Several residues, particularly in the central region of α-synuclein, maintained persistent interactions with S100B in both models, reinforcing their role in stabilizing the complex. However, some differences also emerged: in the AlphaFold model, the N-terminal region of α-synuclein displayed a greater number of contacts compared to the HADDOCK model, whereas in the HADDOCK model, additional interactions were observed in the C-terminal region. An RMSF analysis conducted over the last 100 ns of the simulation reveals that the N-terminal segment of α-synuclein exhibits higher RMSF values and greater fluctuations in the HADDOCK model compared to AlphaFold, suggesting more pronounced conformational changes. Conversely, the C-terminal region shows the opposite trend. This behavior may explain the observed contact frequency patterns. In the AlphaFold model, S100B exhibited a broader interaction interface than in HADDOCK, particularly involving the central and C-terminal regions of the second monomer. Despite these differences, both models highlighted a group of highly conserved residues that may be crucial for complex stabilization.

The MM/GBSA binding energy analysis confirmed that the α-synuclein/S100B interaction is energetically favorable in both models, as indicated by negative binding energy values. The

AlphaFold2-Multimer model showed a slightly stronger binding affinity (-101.11 kcal/mol) than HADDOCK (-80.12 kcal/mol), primarily due to more pronounced electrostatic stabilization. Coulombic interactions provided the largest stabilizing contribution in both models, while van der Waals interactions also played a significant role in maintaining complex stability.

The per-residue energy decomposition analysis revealed key stabilizing residues in both α-synuclein and S100B. Among them, Met1, Asp2, Glu13, Glu20, Gln24, Glu28, Glu35, Glu46, and His50 in α-synuclein displayed the strongest stabilizing effects, with Met1 contributing the most in both models. Similarly, for S100B, Arg20 and Arg110 emerged as the most favorable contributors, indicating their importance in complex stabilization. In contrast, Phe4 and Met5 in α-synuclein, as well as His42, Ile47, Cys68, Cys84, and Ile137 in S100B, exhibited positive free energy contributions, opposing binding. For S100B, these destabilizing residues were the same in both HADDOCK and AlphaFold, with comparable energy values, suggesting a consistent pattern of unfavorable interactions across the two models.

Comparing the present findings with previous studies on S100B/β-amyloid and S100A9/α-synuclein interactions allows for an evaluation of potential similarities in binding mechanisms.

Rodrigues et al. (2021) [32] investigated how S100B interacts with amyloid beta (Aβ42), with a particular focus on the Aβ25–35 fragment, using molecular docking and molecular dynamics simulations. Their analysis revealed a strong electrostatic component in the HADDOCK docking score, a characteristic also observed in this study for the HADDOCK-derived α-Syn50/S100B complex, which exhibited an electrostatic energy contribution of -295.3 ± 8.2 kcal/mol. After molecular dynamics simulations, the analysis of energy contributions further confirmed that coulombic interactions play a major role in stabilizing both the Aβ42/S100B and α-syn50/S100B complexes.

Furthermore, Rodrigues et al. identified Aβ-Lys28 as a crucial stabilizing residue, which formed preferential contacts with Met79, Thr82, and Glu86 of S100B. In the present study, when constructing the HADDOCK model, the active residues of α-synuclein were selected based on sequence similarity to β-amyloid25–35. Specifically, the 11-residue fragment α-syn31–41 was chosen as it exhibited the highest resemblance to Aβ25–35. Within this region, it was observed that a lysine residue (Lys34) occupies position 4, mirroring the presence of Lys28 in β-amyloid. In both HADDOCK and AlphaFold2-Multimer models, Lys34 displayed a contact probability close to 100% and provided a significant stabilizing contribution in terms of binding energy. Following the methodology of Rodrigues et al., an analysis was conducted to determine which S100B residues predominantly interacted with Lys34 over the last 100 ns of simulation. The results revealed distinct interaction patterns, highlighting differences not only between the HADDOCK and AlphaFold models of α-syn50/S100B but also when compared to the Aβ42/S100B complex. Specifically, in the HADDOCK model, Lys34 formed stable contacts with Glu58, Thr59, and Asn62 of S100B, whereas in the AlphaFold-derived structure, its main interacting partners were Ala173, Glu176, and Phe177.

As for the secondary structure evolution, Rodrigues et al. modeled Aβ42 in an α-helical conformation, which remained stable in the Aβ25–35 segment but underwent structural changes in the N-terminal region (1–24) during equilibration. Similarly, in this study, α-synuclein was initialized in an α-helical conformation. However, while the AlphaFold model retained a higher fraction of α-helical structure, particularly within residues 31–41, the HADDOCK-derived complex exhibited greater structural variability, with the α-helix in this region transitioning into a coil conformation. Despite these variations, the secondary structure of S100B remained highly stable throughout the simulations, consistent with previous findings. Additionally, in both cases, the interaction interfaces were well preserved.

The study by Toleikis et al. (2022) [7] on S100A9-α-synuclein interactions offers additional context for interpreting these results. In both their study and the present one, molecular dynamics simulations were performed exclusively on the first 50 residues of the protein.

A parallel can be drawn between their findings and the present study, as both indicate that S100 proteins maintain a stable secondary structure during simulations, while α-synuclein remains highly flexible. However, in their study, α-synuclein retained α-helical content only in the N-terminal region (residues 1–10), whereas in the present study with S100B, helical structures were also preserved in other regions. Similarly, their simulations showed that the complex exhibited a certain degree of flexibility, likely attributable to α-synuclein dynamics.

Toleikis et al. confirmed through molecular dynamics simulations that residues 1–9 in the N-terminal region of α-synuclein, along with residues 35–41, acted as key binding sites as initially detected via NMR experiments. The present study is generally in agreement with these findings, as most of these residues provided a stabilizing contribution in terms of binding energy in both the HADDOCK and

AlphaFold models. The only exceptions were Phe4 and Met5, which in this study exhibited a destabilizing effect rather than a stabilizing role, as observed in S100A9 binding.

When analyzing contact frequency, several residues within the 35–41 region exhibited strong interactions in both models. However, for the 1–9 segment, contacts were observed only in the AlphaFold-derived structure.

Overall, these comparisons suggest that S100B interacts with α-synuclein through mechanisms that share some similarities with both S100A9 and β-amyloid interactions, while also exhibiting distinct features.

In conclusion, the results suggest that α-synuclein and S100B can form a stable complex, as indicated by the persistence of key binding residues, favorable binding energies and consistent interaction patterns across the models. However, since these findings are based only on molecular modeling without direct experimental validation from techniques such as NMR or X-ray crystallography, they represent a plausible hypothesis rather than definitive confirmation of the interaction under physiological conditions. Further experimental studies would be necessary to validate these computational predictions and assess the biological relevance of the interaction.

6 Conclusions

Neurodegenerative diseases such as Parkinson's and Alzheimer's are associated with protein misfolding and aggregation. In Parkinson's disease, the accumulation of α-synuclein leads to neuronal dysfunction, making it essential to investigate the mechanisms underlying protein aggregation and identify potential modulators of its kinetics. Although growing evidence suggests that members of the S100 protein family interact with amyloidogenic proteins, the direct interaction between α-synuclein and S100B remains largely unexplored.

Since no experimentally resolved structures were available in the literature, HADDOCK and AlphaFold2-Multimer were used to predict the complex structure. The resulting models were then analyzed through molecular dynamics simulations to evaluate their stability, binding interfaces, and interaction dynamics over time. The results suggest that α-synuclein and S100B establish stable interaction interfaces. The interaction involving the α-synuclein 31-41 segment and S100B was found to be energetically favorable and a quantitative decomposition of the binding energy highlighted electrostatic interactions as the main driving force stabilizing the complex. Despite differences between the HADDOCK and AlphaFold2-Multimer models in terms of structural fluctuations and binding orientations, both approaches identified a common subset of residues emerging as key candidates for stabilizing interactions.

One of the main limitations of this study is the lack of direct experimental data on the α-synuclein/S100B interaction, which required the selection of active residues based on hypotheses derived from studies on related systems.

Further studies could employ techniques such as NMR spectroscopy to experimentally verify the identified binding sites. Beyond structural validation, further investigations should explore the biological implications of this interaction, particularly in relation to α-synuclein aggregation. Understanding whether S100B stabilizes or alters α-synuclein's structure could provide valuable insights into its role in Parkinson's disease and contribute to the development of more effective therapeutic strategies.

7 Bibliography

[1] K. K. Moodley, A. Nitkunan, e A. C. Pereira, «Acute neurology: a suggested approach», *Clin. Med.*, vol. 18, fasc. 5, pp. 418–421, ott. 2018, doi: 10.7861/clinmedicine.18-5-418.

[2] C. P. Palanisamy *et al.*, «New strategies of neurodegenerative disease treatment with extracellular vesicles (EVs) derived from mesenchymal stem cells (MSCs)», *Theranostics*, vol. 13, fasc. 12, pp. 4138–4165, 2023, doi: 10.7150/thno.83066.

[3] R. N. Rambaran e L. C. Serpell, «Amyloid fibrils: Abnormal protein assembly», *Prion*, vol. 2, fasc. 3, pp. 112–117, lug. 2008, doi: 10.4161/pri.2.3.7488.

[4] G. G. Kovacs, «Concepts and classification of neurodegenerative diseases», in *Handbook of Clinical Neurology*, vol. 145, Elsevier, 2018, pp. 301–307. doi: 10.1016/B978-0-12-802395-2.00021-3.

[5] B. N. Dugger e D. W. Dickson, «Pathology of Neurodegenerative Diseases», *Cold Spring Harb. Perspect. Biol.*, vol. 9, fasc. 7, p. a028035, lug. 2017, doi: 10.1101/cshperspect.a028035.

[6] R. Armstrong, «What causes neurodegenerative disease?», *Folia Neuropathol.*, vol. 58, fasc. 2, pp. 93–112, 2020, doi: 10.5114/fn.2020.96707.

[7] Z. Toleikis *et al.*, «Interactions between S100A9 and Alpha-Synuclein: Insight from NMR Spectroscopy», *Int. J. Mol. Sci.*, vol. 23, fasc. 12, p. 6781, giu. 2022, doi: 10.3390/ijms23126781.

[8] T. R. Mhyre, J. T. Boyd, R. W. Hamill, e K. A. Maguire-Zeiss, «Parkinson's Disease», in *Protein Aggregation and Fibrillogenesis in Cerebral and Systemic Amyloid Disease*, vol. 65, J. R. Harris, A c. di, in Subcellular Biochemistry, vol. 65. ,

Dordrecht: Springer Netherlands, 2012, pp. 389–455. doi: 10.1007/978-94-007-5416-4_16.

[9] G. DeMaagd e A. Philip, «Parkinson's Disease and Its Management».

[10] C. Frank e F. G. Pari, «Approach to diagnosis of Parkinson disease».

[11] J. Jankovic, «Parkinson's disease: clinical features and diagnosis», *J. Neurol. Neurosurg. Psychiatry*, vol. 79, fasc. 4, pp. 368–376, apr. 2008, doi: 10.1136/jnnp.2007.131045.

[12] K. preetKaur, N. Khurana, N. Sharma, N. Sharma, e N. Sharma, «PHYTOCHEMICALS AS FUTURE DRUGS FOR PARKINSON'S DISEASE: A REVIEW», *PLANT Arch.*, vol. 21, fasc. 1, dic. 2020, doi: 10.51470/PLANTARCHIVES.2021.v21.S1.384.

[13] R. P. Munhoz, V. Tumas, J. L. Pedroso, e L. Silveira-Moriyama, «The clinical diagnosis of Parkinson's disease», *Arq. Neuropsiquiatr.*, vol. 82, fasc. 06, pp. 001–010, giu. 2024, doi: 10.1055/s-0043-1777775.

[14] C. Carrarini *et al.*, «A Stage-Based Approach to Therapy in Parkinson's Disease», *Biomolecules*, vol. 9, fasc. 8, p. 388, ago. 2019, doi: 10.3390/biom9080388.

[15] M. Bisaglia, «Mediterranean Diet and Parkinson's Disease», *Int. J. Mol. Sci.*, vol. 24, fasc. 1, p. 42, dic. 2022, doi: 10.3390/ijms24010042.

[16] T. Foltynie, V. Bruno, S. Fox, A. A. Kühn, F. Lindop, e A. J. Lees, «Medical, surgical, and physical treatments for Parkinson's disease», *The Lancet*, vol. 403, fasc. 10423, pp. 305–324, gen. 2024, doi: 10.1016/S0140-6736(23)01429-0.

[17] J. D. Rogers, A. Sanchez-Saffon, A. B. Frol, e R. Diaz-Arrastia, «Elevated Plasma Homocysteine Levels in Patients Treated With Levodopa: Association With Vascular Disease», *Arch. Neurol.*, vol. 60, fasc. 1, p. 59, gen. 2003, doi: 10.1001/archneur.60.1.59.

[18] D. L. Heiden, B. Monogue, M. D. H. Ali, e J. D. Beckham, «A functional role for alpha-synuclein in neuroimmune responses», *J. Neuroimmunol.*, vol. 376, p. 578047, mar. 2023, doi: 10.1016/j.jneuroim.2023.578047.

[19] S. Gallegos, C. Pacheco, C. Peters, C. M. Opazo, e L. G. Aguayo, «Features of alpha-synuclein that could explain the progression and irreversibility of Parkinson's disease», *Front. Neurosci.*, vol. 9, mar. 2015, doi: 10.3389/fnins.2015.00059.

[20] L. Gadhe *et al.*, «Intermediates of α-synuclein aggregation: Implications in Parkinson's disease pathogenesis», *Biophys. Chem.*, vol. 281, p. 106736, feb. 2022, doi: 10.1016/j.bpc.2021.106736.

[21] J. S. Cristóvão e C. M. Gomes, «S100 Proteins in Alzheimer's Disease», *Front. Neurosci.*, vol. 13, p. 463, mag. 2019, doi: 10.3389/fnins.2019.00463.

[22] F. Michetti *et al.*, «The S100B Protein: A Multifaceted Pathogenic Factor More Than a Biomarker», *Int. J. Mol. Sci.*, vol. 24, fasc. 11, p. 9605, mag. 2023, doi: 10.3390/ijms24119605.

[23] V. Garcia e W. J. Chazin, «A new approach to discovery of S100 protein heterodimers», *FEBS J.*, vol. 286, fasc. 10, pp. 1838–1840, mag. 2019, doi: 10.1111/febs.14851.

[24] F. Sedaghat e A. Notopoulos, «S100 protein family and its application in clinical practice», *Hippokratia*, vol. 12, fasc. 4, pp. 198–204, 2008.

[25] R. Donato, «Functional roles of S100 proteins, calcium-binding proteins of the EF-hand type», *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, vol. 1450, fasc. 3, pp. 191–231, lug. 1999, doi: 10.1016/S0167-4889(99)00058-0.

[26] R. Donato, «Intracellular and extracellular roles of S100 proteins», *Microsc. Res. Tech.*, vol. 60, fasc. 6, pp. 540–551, apr. 2003, doi: 10.1002/jemt.10296.

[27] «hippokratia-12-198».

[28] P. Kd e F. L, «Structural Basis for S100B Interaction with its Target Proteins», *J. Mol. Genet. Med.*, vol. 12, fasc. 3, 2018, doi: 10.4172/1747-0862.1000366.

[29] R. Bianchi, C. Adami, I. Giambanco, e R. Donato, «S100B binding to RAGE in microglia stimulates COX-2 expression», *J. Leukoc. Biol.*, vol. 81, fasc. 1, pp. 108–118, gen. 2007, doi: 10.1189/jlb.0306198.

[30] M. Rothermundt, M. Peters, J. H. M. Prehn, e V. Arolt, «S100B in brain damage and neurodegeneration», *Microsc. Res. Tech.*, vol. 60, fasc. 6, pp. 614–632, apr. 2003, doi: 10.1002/jemt.10303.

[31] G. Sorci *et al.*, «S100B Protein, a Damage-Associated Molecular Pattern Protein in the Brain and Heart, and Beyond», *Cardiovasc. Psychiatry Neurol.*, vol. 2010, pp. 1–13, ago. 2010, doi: 10.1155/2010/656481.

[32] F. E. P. Rodrigues, A. J. Figueira, C. M. Gomes, e M. Machuqueiro, «Computational Analysis of the Interactions between the S100B Extracellular Chaperone and Its Amyloid β Peptide Client», *Int. J. Mol. Sci.*, vol. 22, fasc. 7, p. 3629, mar. 2021, doi: 10.3390/ijms22073629.

[33] R. Coelho *et al.*, «Secondary Modification of S100B Influences Anti Amyloid-β Aggregation Activity and Alzheimer's Disease Pathology», *Int. J. Mol. Sci.*, vol. 25, fasc. 3, p. 1787, feb. 2024, doi: 10.3390/ijms25031787.

[34] I. Horvath *et al.*, «Co-aggregation of pro-inflammatory S100A9 with α-synuclein in Parkinson's disease: ex vivo and in vitro studies», *J. Neuroinflammation*, vol. 15, fasc. 1, p. 172, dic. 2018, doi: 10.1186/s12974-018-1210-9.

[35] Z. Toleikis, M. Ziaunys, L. Baranauskiene, V. Petrauskas, K. Jaudzems, e V. Smirnovas, «S100A9 Alters the Pathway of Alpha-Synuclein Amyloid Aggregation», *Int. J. Mol. Sci.*, vol. 22, fasc. 15, p. 7972, lug. 2021, doi: 10.3390/ijms22157972.

[36] J. C. A. Boeyens, «Fundamentals of Molecular Modelling», in *Fundamental Principles of Molecular Modeling*, W. Gans, A. Amann, e J. C. A. Boeyens, A c. di, Boston, MA: Springer US, 1996, pp. 1–9. doi: 10.1007/978-1-4899-0212-2_1.

[37] N. A. Saleh, H. Elhaes, e M. Ibrahim, «Design and Development of Some Viral Protease Inhibitors by QSAR and Molecular Modeling Studies», in *Viral Proteases and Their Inhibitors*, Elsevier, 2017, pp. 25–58. doi: 10.1016/B978-0-12-809712-0.00002-2.

[38] G. A. Worth e L. S. Cederbaum, «BEYOND BORN-OPPENHEIMER: Molecular Dynamics Through a Conical Intersection», *Annu. Rev. Phys. Chem.*, vol. 55, fasc. 1, pp. 127–158, giu. 2004, doi: 10.1146/annurev.physchem.55.091602.094335.

[39] J. Crabbe, «Molecular modelling: Principles and applications», *Comput. Chem.*, vol. 21, fasc. 3, p. 185, gen. 1997, doi: 10.1016/S0097-8485(96)00029-0.

[40] S. A. Gorbunov, A. E. Volkov, e R. A. Voronkov, «Periodic boundary conditions effects on atomic dynamics analysis», *Comput. Phys. Commun.*, vol. 279, p. 108454, ott. 2022, doi: 10.1016/j.cpc.2022.108454.

[41] P. C. Agu *et al.*, «Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management», *Sci. Rep.*, vol. 13, fasc. 1, p. 13398, ago. 2023, doi: 10.1038/s41598-023-40160-2.

[42] M. Mursal, M. Ahmad, S. Hussain, e M. Faraz Khan, «Navigating the Computational Seas: A Comprehensive Overview of Molecular Docking Software in Drug Discovery», in *Biomedical Engineering*, vol. 27, Č. Podlipnik, A c. di, IntechOpen, 2024. doi: 10.5772/intechopen.1004802.

[43] X.-Y. Meng, H.-X. Zhang, M. Mezei, e M. Cui, «Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery», *Curr. Comput. Aided-Drug Des.*, vol. 7, fasc. 2, pp. 146–157, giu. 2011, doi: 10.2174/157340911795677602.

[44] L. Zheng *et al.*, «Improving protein–ligand docking and screening accuracies by incorporating a scoring function correction term», *Brief. Bioinform.*, vol. 23, fasc. 3, p. bbac051, mag. 2022, doi: 10.1093/bib/bbac051.

[45] «A Comprehensive Review on the Top 10 Molecular Docking Softwares».

[46] P. H. M. Torres, A. C. R. Sodero, P. Jofily, e F. P. Silva-Jr, «Key Topics in Molecular Docking for Drug Design», *Int. J. Mol. Sci.*, vol. 20, fasc. 18, p. 4574, set. 2019, doi: 10.3390/ijms20184574.

[47] E. Karaca, A. S. J. Melquiond, S. J. De Vries, P. L. Kastritis, e A. M. J. J. Bonvin, «Building Macromolecular Assemblies by Information-driven Docking», *Mol. Cell. Proteomics*, vol. 9, fasc. 8, pp. 1784–1794, ago. 2010, doi: 10.1074/mcp.M000051-MCP201.

[48] A. Saponaro, V. Maione, A. Bonvin, e F. Cantini, «Understanding Docking Complexes of Macromolecules Using HADDOCK: The Synergy between Experimental Data and Computations», *BIO-Protoc.*, vol. 10, fasc. 20, 2020, doi: 10.21769/BioProtoc.3793.

[49] J. Jumper *et al.*, «Highly accurate protein structure prediction with AlphaFold», *Nature*, vol. 596, fasc. 7873, pp. 583–589, ago. 2021, doi: 10.1038/s41586-021-03819-2.

[50] P. Bryant e F. Noé, «Improved protein complex prediction with AlphaFold-multimer by denoising the MSA profile», *PLOS Comput. Biol.*, vol. 20, fasc. 7, p. e1012253, lug. 2024, doi: 10.1371/journal.pcbi.1012253.

[51] V. Scardino, J. I. Di Filippo, e C. N. Cavasotto, «How good are AlphaFold models for docking-based virtual screening?», *iScience*, vol. 26, fasc. 1, p. 105920, gen. 2023, doi: 10.1016/j.isci.2022.105920.

[52] R. N. L. Lamptey, B. Chaulagain, R. Trivedi, A. Gothwal, B. Layek, e J. Singh, «A Review of the Common Neurodegenerative Disorders: Current Therapeutic Approaches and the Potential Role of Nanotherapeutics», *Int. J. Mol. Sci.*, vol. 23, fasc. 3, p. 1851, feb. 2022, doi: 10.3390/ijms23031851.

[53] Z. Yang, X. Zeng, Y. Zhao, e R. Chen, «AlphaFold2 and its applications in the fields of biology and medicine», *Signal Transduct. Target. Ther.*, vol. 8, fasc. 1, p. 115, mar. 2023, doi: 10.1038/s41392-023-01381-z.

[54] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, e M. Steinegger, «ColabFold: making protein folding accessible to all», *Nat. Methods*, vol. 19, fasc. 6, pp. 679–682, giu. 2022, doi: 10.1038/s41592-022-01488-1.

[55] A. Vangone *et al.*, «Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1», *Proteins Struct. Funct. Bioinforma.*, vol. 85, fasc. 3, pp. 417–423, mar. 2017, doi: 10.1002/prot.25198.

[56] G. Kim *et al.*, «Easy and accurate protein structure prediction using ColabFold», 1 dicembre 2023, *Protocol Exchange*. doi: 10.21203/rs.3.pex-2490/v1.

[57] M. Varadi *et al.*, «AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models», *Nucleic Acids Res.*, vol. 50, fasc. D1, pp. D439–D444, gen. 2022, doi: 10.1093/nar/gkab1061.

[58] S. Scalera, «Computational insights on the role of metal ions as stabilizers for chaperone-amyloid interactions», Politecnico di Torino, 2024. [Online]. Disponibile su: https://webthesis.biblio.polito.it/31937/

[59] F. Ormeño e I. J. General, «Convergence and equilibrium in molecular dynamics simulations», *Commun. Chem.*, vol. 7, fasc. 1, p. 26, feb. 2024, doi: 10.1038/s42004-024-01114-5.

[60] P. Carter, «DSSPcont: continuous secondary structure assignments for proteins», *Nucleic Acids Res.*, vol. 31, fasc. 13, pp. 3293–3295, lug. 2003, doi: 10.1093/nar/gkg626.

[61] S. Ali, Md. Hassan, A. Islam, e F. Ahmad, «A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States», *Curr. Protein Pept. Sci.*, vol. 15, fasc. 5, pp. 456–476, mag. 2014, doi: 10.2174/1389203715666140327114232.

[62] D. Chakravarty, M. Guharoy, C. H. Robert, P. Chakrabarti, e J. Janin, «Reassessing buried surface areas in protein–protein complexes», *Protein Sci.*, vol. 22, fasc. 10, pp. 1453–1457, ott. 2013, doi: 10.1002/pro.2330.

[63] S. Genheden e U. Ryde, «The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities», *Expert Opin. Drug Discov.*, vol. 10, fasc. 5, pp. 449–461, mag. 2015, doi: 10.1517/17460441.2015.1032936.

[64] M. Tian *et al.*, «Application of per-Residue Energy Decomposition to Design Peptide Inhibitors of PSD95 GK Domain», *Front. Mol. Biosci.*, vol. 9, p. 848353, mar. 2022, doi: 10.3389/fmolb.2022.848353.

[65] J. D. Hunter, «Matplotlib: A 2D Graphics Environment», *Comput. Sci. Eng.*, vol. 9, fasc. 3, pp. 90–95, 2007, doi: 10.1109/MCSE.2007.55.