MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN
Faculty of Chemical and Biopharmaceutical Technologies
Department of Biotechnology, Leather and Fur

QUALIFICATION THESIS

on the topic **Research on Enhancing the Activity of *Streptomyces noursei* Polylysine Synthetase through Rational Design Methods**
First (Bachelor's) level of higher education
Specialty 162 "Biotechnology and Bioengineering"
Educational and professional program "Biotechnology"

Completed: student of group BEBT-20
Su LIQI

Scientific supervisor
Ihor HRETSKYI, Ph.D., As. prof.

Reviewer
Liubov ZELENA, Ph.D., As. prof.

Kyiv 2024

KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN

Faculty: Chemical and Biopharmaceutical Technologies
Department: Biotechnology, Leather and Fur
First (Bachelor's) level of higher education
Specialty: 162 Biotechnology and Bioengineering
Educational and professional program Biotechnology

APPROVE
Head of Department of Biotechnology, Leather and Fur, Professor,
Doctor of Technical Science
Olena MOKROUSOVA
_____
«___»_____2024

ASSIGNMENTS
FOR THE QUALIFICATION THESIS
Su Liqi

1. Thesis topic **Research on Enhancing the Activity of *Streptomyces noursei* Polylysine Synthetase through Rational Design Methods**
Scientific supervisor Ihor Hretskyi, Ph.D., As. prof

approved by the order of KNUTD "__"_____2024, №____

2. Initial data for work: assignments for qualification thesis, scientific literature on the topic of qualification thesis, materials of Pre-graduation practice

3. Content of the thesis (list of questions to be developed): literature review; object, purpose, and methods of the study; experimental part; conclusions

4. Date of issuance of the assignments_____

EXECUTION SCHEDULE

№	The name of the stages of the qualification thesis	Terms of performance of stage	Note on performance
1	Introduction	From 01 April 2024 to 11 April 2024	
2	Section 1 Literature review	From 06 April 2024 to 20 April 2024	
3	Section 2 Object, purpose, and methods of the research	From 21 April 2024 to 30 April 2024	
4	Section 3 Experimental part	From 01 May 2024 to 10 May 2024	
5	Conclusions	From 07 May 2024 to 12 May 2024	
6	Draw up a bachelor's thesis (final version)	From 12 May 2024 to 24 May 2024	
7	Submission of qualification work to the supervisor for feedback (14 days before the defense)	From 24 May 2024 to 10 June 2024	
8	Submission of bachelor's thesis to the department for review (12 days before the defense)	13 June 2024	
9	Checking the bachelor's thesis for signs of plagiarism (10 days before the defense)	15 June 2024	
10	Submission of bachelor's thesis for approval by the head of the department (from 7 days before the defense)	17 June 2024	

I am familiar with the task:

Student _____ Su LIQI

Scientific supervisor _____ Ihor HRETSKYI

SUMMARY

Liqi Su. Research on Enhancing the Activity of Streptomyces noursei Polylysine Synthetase through Rational Design Methods. - Manuscript.

Qualification thesis on the specialty 162 «Biotechnology and Bioengineering». – Kyiv National University of Technologies and Design, Kyiv, 2024.

ε-Polylysine synthetase (ε-Poly-L-lysine synthetase, Pls) is a novel single-module non-ribosomal The activity of peptide synthetase (NRPS) is an important factor affecting the synthesis degree of ε-polylysine (ε-Poly-L-lysine, ε-PL). The laboratory obtained Pls from *Streptomyces nourans*, *S. albulus* NBRC 14147, but its Less active. This project is based on rational design methods and explores the purpose of locating key site mutations through structural analysis to improve its activity.

The experiment uses the homology modeling tool SWISS-MODEL to construct the three-dimensional structure of the original and mutant Pls; through Pymol. The visualization software observes and analyzes the structural characteristics of the key regions of Pls to determine potential mutation sites; it uses molecular docking methods and molecular dynamics simulation methods to test the theoretical effects of mutations. Key region structure analysis found that there are many amino acid residues with larger side chains on the Pls active region channel, which has a greater impact on the entry and exit of small substrate molecules into the active region. steric hindrance. Based on this discovery, this project intends to improve its activity by reducing the steric hindrance near the reaction channel. After further analysis and screening, the four key residues Arg370, Arg401, Asp404 and Lys499 were finally determined as mutations location. Molecular docking and molecular dynamics analysis results show that in addition to the Lys499 site, the binding ability of other site mutants to substrate molecules has been improved to a certain extent, among which Asp404 mutant has the strongest binding ability to substrate, Arg401 mutant comes second, Arg370 Mutants third.

The results of this paper deepen our understanding of the catalytic mechanism of ε-polylysine synthase, and provide a theory for the laboratory engineering of polylysine synthase protein. Guidance to facilitate the industrial production of ε-polylysine.

Key words: Polylysine synthase; computational biology; bioinformatics; gene mutation; homology modeling; molecular docking; molecular dynamics

TABLE OF CONTENTS

INTRODUCTION

ε-Poly-L-lysine synthetase (ε-PLs) is a novel single-module non-ribosomal peptide synthetase (NRPS) that catalyzes the synthesis of ε-poly-L-lysine (ε-PL). As a biopolymer with various potential applications, ε-PL is recognized for its edibility, degradability, and non-toxic nature to both humans and the environment, presenting broad prospects for use in the food, pharmaceutical, and materials industries. However, the low activity of the Pls obtained in the laboratory has limited its application in industrial production. This study employed the homology modeling tool SWISS-MODEL to construct the three-dimensional structures of the wild-type and mutant Pls. Key structural features of the enzyme were observed and analyzed using PyMOL software to identify potential sites for mutation. The theoretical effects of mutations were then examined using molecular docking and molecular dynamics simulation methods. The research commenced by retrieving the amino acid sequence of Pls from the *S. albulus* NBRC 14147 strain through the NCBI database, followed by homology modeling using structural information from the PDB database. Subsequently, the quality of the protein model was assessed using a Ramachandran plot analysis. Based on this, the location of the active site pocket and mutation sites were determined, and key amino acid residues were subjected to gene mutation. The study found that there were numerous amino acid residues with bulky side chains in the active site channel of Pls, which caused significant steric hindrance to the entry and exit of substrate molecules. By reducing the steric hindrance near the reaction

channel, the research team identified four key residues (Arg370, Arg401, Asp404, and Lys499) as targets for mutation. Molecular docking and molecular dynamics analysis showed that, except for the Lys499 site, the binding ability of other mutant enzymes to substrate molecules had improved, with the Asp404 mutant exhibiting the strongest binding capacity.The findings of this study have not only deepened the understanding of the catalytic mechanism of ε-poly-L-lysine synthetase but also provided theoretical guidance for protein engineering of poly-lysine synthetase in the laboratory, aiding in the industrial production of ε-poly-L-lysine. Although this is a preliminary study on the activity of Pls, the experimental conclusions have potential practical application value and offer new directions for future research.

The relevance of the topic is Utilizing tools such as homology modeling, molecular docking, and molecular dynamics simulations in bioinformatics for gene modification.

The purpose of the study is Exploring the factors that affect the activity of poly-L-lysine synthetase.

The objectives of the study obtain a strain of Pls (Poly-L-lysine synthetase) with higher activity.

The object of the study S. albulus NBRC 14147 strain

The subject of conducted on the genes and structure of the strain

Research methods Literature review method and Computational Biology

The scientific novelty Recent advances in the field of gene research using bioinformatics.

The practical Obtaining a high-yielding strain that can be widely applied across various industries, which can reduce its usage costs and enhance the safety of its application.

CHAPTER 1

LITERATURE REVIEW

1.1 Introduction to polylysine synthase

ε-Poly-L-lysine synthetase (Pls) is a transmembrane protein and a new single-module non-ribosomal peptide synthetase (NRPS) [1], non-ribosomal peptide synthetase has multiple semi-autonomous functional domains, and multiple L-lysines can form peptide chains after passing through the polylysine synthase-related regions and keep extending. Polylysine synthasehas amino acid ligase-like activity and can catalyze the formation of peptide bonds. Current research on the synthesis mechanism of polylysine has confirmed that polylysine synthase is a key factor affecting the degree of polylysine synthesis. Polylysine (ε-Poly-L-lysine,ε-PL) is a biopolymer material that is edible, degradable and harmless to people and the environment. The advantages of poisoning are widely used in food, medicine and materials [2].

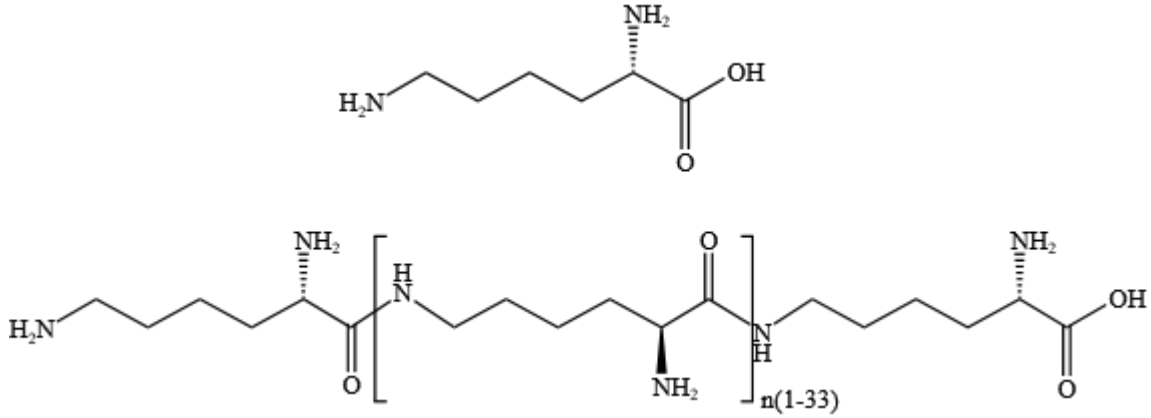


Figure 1.1- Polylysine molecular structure

1.2 Structure and catalytic mechanism of polylysine synthase

Yamanaka et al. isolated a transmembrane protein [2] **Ошибка! Источник ссылки не найден.** from the cell membrane of *S. albulus* NBRC14147, and used this transmembrane protein to conduct lysine synthesis experiments in vitro, and finally in the product The ε-PL component was detected, confirming that the transmembrane protein is ε-PL synthase (Pls). Kewai et al. synthesized ε-PL [3] in vitro using cell membrane fragments through the isotope-labeled glucose method. The results showed that ε-PL synthase is a non-ribosomal peptide synthetase. (NRPS), in the system of non-ribosomal peptide synthetase (NRPS), there are three key active regions: adenylation domain (A domain), thiolation domain (T domain) and condensation domain (C domain). The function of the A domain is to activate the amino acid substrate into aminoacyl-O-AMP, while the T domain is responsible for transferring the activated substrate to a specific location. The C domain performs a dehydration condensation reaction to connect the activated substrates to form a polypeptide chain. However, *ε*-polylysine synthetase (PLS) differs from typical NRPS enzymes in that it lacks the C domain in its structure and instead possesses three functionally similar C domains. regions, labeled C1, C2, and C3 domains respectively. Pls does not contain the traditional dehydration condensation domain (C domain) and thioesterase (TE domain). In contrast, it has a domain consisting of six transmembrane segments surrounding three soluble domains arranged in series.

In the traditional mechanism of non-ribosomal peptide synthetase (NRPS), A domain has the ability to recognize specific amino acids and activate ATP to

synthesize the corresponding aminoacyl adenylate. After this activation process, the amino acid is anchored to a highly conserved serine residue in the T domain, which is linked to 4'-phosphopantetheine. Next, the T domain passes this activated structural unit to the C domain, which is responsible for catalyzing the formation of the peptide bond between the two activated amino acids. Finally, the synthesized peptide chain is cleaved and released at the thioesterase (TE) domain of the terminal module. By comparing the C domain sequences of Pls with those of traditional NRPS, In the three consecutive regions of Pls, each region shows similarities with the histidine motif common to all conventional C domains in non-ribosomal peptide synthetases (NRPS) has highly similar sequences. The amino acid sequences of these continuous regions and their predicted three-dimensional structures have certain structural similarities with the C domain of NRPS, and they work synergistically to promote the condensation reaction between substrates, Thus Plays a key role in the synthesis of peptide chains.

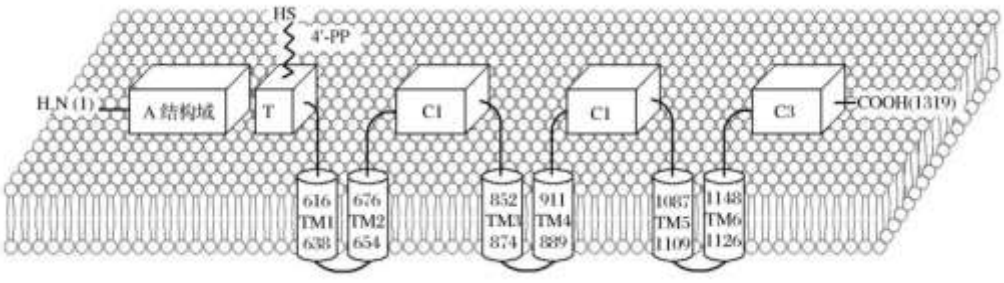


Figure 1.2 - Domain composition of Pls [4].

Pls (polylysine synthetase) in ε-PL(*ε*-poly plays a catalytic role in the

biosynthesis of lysine). This process is similar to the catalytic mechanism of NRPS (non-ribosomal peptide synthetase), starting from L- The lysine monomer serves as an elongation unit, Pls A domain and is subsequently transferred to the phosphopantetheine arm of the T domain. At the Pls C end,C1, C2, C3 domains are responsible for catalyzing the L-lysine activated and transferred by the AT domain with as the starting unit An isopeptide bond between L-lysine is formed, resulting in aL-lysine dimer. This dimer then participates in the subsequent polymerization reaction as a freely diffusing substrate. Due to the lack of Pls NRPSTE catalytic product release and termination reactions common domain, Pls can continuously catalyze the polymerization of L-lysine to produce different Degree of polymerization ε-PL.As the polymerization reaction proceeds, it terminates Pls catalytic activity.

1.3 Synthetic pathway of polylysine

In the biosynthesis process of polylysine, the formation of ε-polylysine (ε-PL) depends on the presence of adenosine triphosphate (ATP). Its synthesis starts from L-lysine monomer, which One process is catalyzed by polylysine synthase (Pls), and Pls is encoded by the Pls gene in microorganisms [5]. In microorganisms that produceε-PL, the biosynthetic pathway of L-lysine is similar to that of bacteria, mainly following the diaminopimelic acid pathway (DAP). The starting material of this pathway is aspartate, which first undergoes phosphorylation and subsequent conversion to aspartate-β-semialdehyde. The synthesis pathway then branches into

those for lysine, threonine, isoleucine or methionine. Under the catalysis of Pls, L-lysine monomer undergoes continuous polymerization reaction to finally form*ε*-polylysine.

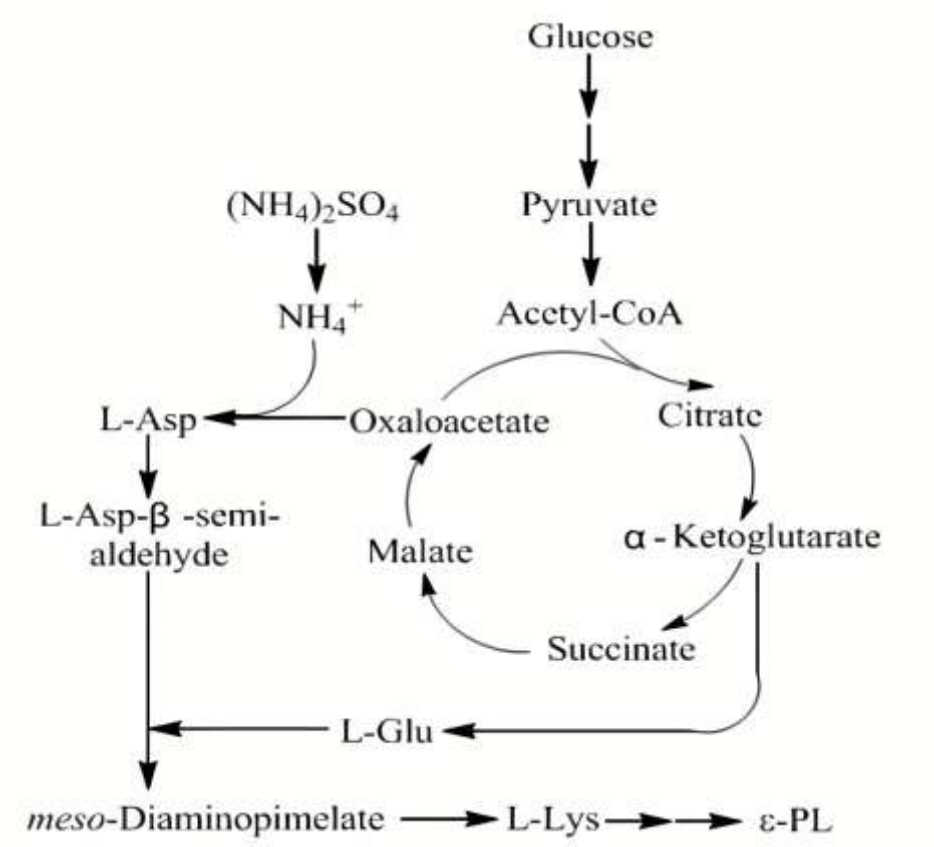


Figure 1.3 - Synthetic pathway of ε-PL

1.4 Discovery and application of polylysine synthase

1.4.1 Discovery and screening of polylysine synthetase

In 1977, Shima et al. first discovered ε-poly-L-lysine (ε-PL) in the culture medium of *Streptomyces albulus* [6]. At that time, there was a lack of specific technology to detect and identify *ε*-PL-producing bacteria, so the researchers performed liquid cultures of individual colonies and then analyzed the metabolites in

their culture media. This screening method is not only inefficient, but also requires a huge amount of work. Since Shima and his team first screened *S. albulus*346 strains capable of producing ε-PL in 1977, there has been no research on the screening of ε-PL production for a long time. Bacteria reports.

Based on the positively charged characteristics of ε-polylysine (ε-PL) and its ability to react with basic or acidic dyes, Nishikawa and Ogawa developed a simple and sensitive method for screening ε-PL producing bacteria. New methods [7]. This technology has significantly improved the screening efficiency ofε-PL producing bacteria, and researchers have successfully screened a variety of newε-PL producing strains. Currently known ε-PL producing bacteria are mainly concentrated in the *Streptomycetaceae* and Ergot fungi families, especially members of the *Streptomycetaceae* family.

Table 1.1 - **Some published Pls sequences**

Strains	DNA number	Pls number
Streptomyces albulus NBRC 14147	AB385841	BAG68864
Streptomyces albulus NK660	CP007574	AIA07462
Streptomyces albulus PD-1	AXDB00000000	EXU90606
Streptomyces albulus CCRC 11814	AROY00000000	EPY92338
Streptomyces roseoverticillatus MN-10	AB477240	BAH85292

1.4.2 Application of polylysine synthase in food

ε-Polylysine (*ε*-PL) is used in the food industry because of its broad antibacterial effect, excellent water solubility, thermal stability and safety to the human body. broadly used. Initially, *ε*-PL was developed and applied as a biopreservative in Japan and was approved for use in food by the Japanese government in the 1980s. Since then, the United States and South Korea have successively approved the use of ε-PL as a food additive [8]. China officially approved ε-PL as a food preservative in 2014. Due to its strong antibacterial ability, only a small amount of addition can effectively extend the shelf life of food. It is worth noting that*ε*-PL itself has a bitter taste, so its use in food needs to be carefully controlled to avoid affecting the taste of the food. *ε*-PL is widely used in the preparation of foods such as noodles and rice. ε-PL is also used as a preservative for vegetables and has a significant effect in inhibiting the growth of bacterial colonies [9]**Ошибка! Источник ссылки не найден.**. Because of its ability to inhibit lipase activity and reduce fat absorption [10], ε-PL is also used as a natural obesity inhibitor [11], providing a healthy choice for some consumers who want to control their weight.

1.4.3 Application of polylysine synthase in biomedicine

ε-Polylysine (ε-PL) has attracted much attention due to its excellent biocompatibility and degradability, and because its molecular structure is positively

charged, it can adsorb negatively charged molecules, making it a potential drug carrier material [12]. ε-PL has sustained-release properties and can be combined with sodium alginate and other substances to form sustained-release capsules for wrapping small molecule proteins and anti-cancer drugs, thereby achieving slow release of drugs and improving therapeutic effects [13]. An important application of ε-PL is as a non-viral gene carrier for DNA transportation [14]. It is a highly branched polymer that combines polylysine and tree polymer-like advantages without causing an immune response. Due to the compactness of its molecular structure, ε-PL can embed a variety of molecular agents within its molecular framework, which makes it have broad application prospects in the field of drug development, including but not limited to drug delivery systems and gene therapy vectors, development of diagnostic imaging agents, biosensors, and specific cancer treatment strategies. As further research into this substance continues, it is expected that more of its potential for medical applications will be revealed.

1.4.4 Application of polylysine synthase in biological materials

ε-Polylysine (ε-PL) is mixed with highly water-absorbent polymers such as propyleneglycol alginate ester to form a high-strength , high swelling and highly absorbent hydrogel. When ε-PL is treated with gamma ray irradiation, a cross-linking reaction occurs, resulting in a super absorbent polymer (SAP) material with a water absorption capacity of more than 200 times [15]. ε-PL's biodegradability and antibacterial properties make it an ideal material for making sanitary products. This

polymer gel with strong water absorption properties is also very suitable for application in agriculture, food and medicine. The electrostatic self-assembly function of polylysine enables it to combine with biomolecules such as oligonucleotides, antibodies, and proteins to form a uniform coating layer on the surface. This characteristic makes it potentially valuable in sensor applications in the fields of biomedicine, biological protection, and environmental monitoring [16]. *ε*-Polylysine can also be used to build high-tech products such as biochips and integrated circuits.

1.5 Research progress and significance of polylysine synthase

1.5.1 Research progress on polylysine synthase

With the innovation of screening methods for*ε*-polylysine (*ε*-PL) production strains, more and more ε-polylysine (*ε*-PL) production strains have been gradually discovered, but the effects of ε-polylysine synthase (Pls) of different strains on the degree of polymerization of lysine vary. Currently, *S. albulus.* The research on NBRC 14147 is relatively in-depth. With the development of bioinformatics, sequence alignment and prediction of *ε*-polylysine synthetase (Pls) can better help us conduct in-depth research on other strains. This method is based on Pls. The mechanism of action takes into account the substrate-specific recognition, binding and condensation processes through domain recognition and the identification of amino acid residues that determine substrate specificity. The predicted polylysine (Pls) is highly consistent in structure and similarity with known polylysine, including

their specific domain composition. Although this method requires further validation of its accuracy and reliability through experiments, it shows great potential in discovering new strains.

Scientists around the world are committed to using diverse strategies to cultivate microbial strains that can efficiently synthesize ε-polylysine (ε-PL). Since biotechnological means, such as physical and chemical mutagenesis, will have extensive and complex effects on strains at multiple levels such as the gene level, metabolic processes, and protein expression, therefore, relying only on a single gene locus or a single metabolic pathway It is difficult for studies to fully reveal the high-yield mechanisms behind high-yield mutant strains. Modern biological metabolism research is beginning to adopt a more comprehensive and multidimensional approach. This methodology involves the application of a series of advanced "omics" technologies, including genomics, metabolomics, transcriptomics and proteomics, combined with verification methods of molecular biology, to systematically analyze the biological basis of high-yield strains. It provides a powerful tool for in-depth understanding of the biological characteristics of high-yielding strains. Multi-dimensional research methods provide theoretical basis and technical support for the genetic improvement of strains, helping to cultivate more efficient and stable ε-PL production strains.

1.5.2 Research significance of polylysine synthase

With economic growth and improvement in quality of life, people pay more

attention to food safety. In this context, ε-polylysine (ε-PL) has shown significant advantages in the field of food preservation due to its unique position in nature. ε-PL not only has extensive antibacterial activity, excellent water solubility and excellent thermal stability, but more importantly, it can be naturally decomposed in the human body and converted into lysine, one of the essential amino acids for the human body. acid, which provides a strong guarantee for its biological safety. Thanks to these properties, ε-PL has been recognized and officially approved for food preservation by many countries, including Japan, South Korea, the United States, and China, demonstrating its potential in the food industry. Huge application potential. With the continuous deepening of research on ε-PL and the continuous innovation of application technology, it is of great significance to improve the level of food safety and promote the healthy development of the food industry. ε-polylysine is also very important to study the safety of ε-polylysine. Conduct in-depth research on its metabolic pathways, toxicological properties and long-term intake effects in the human body to ensure its safe use in various fields. ε- The research of polylysine involves biology, chemistry, materials science, medicine and other disciplines. Its interdisciplinary characteristics provide a platform for cooperation and communication between scientists in different fields, which is important for promoting related The technological progress of the industry is of great significance and has a profound impact on promoting human health, environmental protection, and agricultural development.

1.6 Research objectives

This topic aims to explore the four key site pairs *S. albulus* NBRC14147 (NCBI ID:WP_020931051.1). Acid synthase activity. We mutated four key sites and, through structural analysis and modification, aimed to enhance the activity of Pls and develop Pls mutants with higher activity.

CHAPTER 2

OBJECT, PURPOSE, AND METHODS OF THE STUDY

2.1 Pls Sequence source

Search the *S. albulus* NBRC 14147 strain Pls sequence (NCBI ID: WP_020931051.1) through the NCBI database (https://www.ncbi.nlm.nih.gov/) and through the Protein Data Bank (PDB, https:/ /www.rcsb.org/) database retrieved ε-Structure of the poly-L-lysine synthase adenylation domain information (PBB ID: 7WEW).

2.2 Homology modeling of Pls

2.2.1 Pls Homology modeling

This experiment uses the homology modeling method to construct a three-dimensional model. SWISS-MODEL uses advanced algorithms to predict protein three-dimensional structures based on experimentally determined protein structure data. It is currently the most credible homology modeling tool. Simply provide the amino acid sequence of a target protein and submit a modeling job, and the platform automates a complex set of steps including homology search, template selection, model construction, and model evaluation. SWISS-MODEL is tightly integrated with the Protein Data Bank (PDB) database, which contains a large amount of experimentally determined protein structure data. SWISS-MODEL uses these data as modeling templates. SWISS-MODEL uses advanced homology search algorithms,

such as PSI-BLAST and SPARKS-X. These algorithms can efficiently identify and select the most suitable template structure and apply a variety of modeling techniques to ensure that the constructed protein model has structural integrity. It is as close to the real state as possible and provides model quality assessment. Through a comprehensive and systematic homology modeling process, a more reliable three-dimensional structure can be obtained.

2.2.2 Plots Ramachandran Plots

Ramachandran Plots are used to analyze amino acid residues in protein molecules. φ(phi, the angle between N-Cα-C atoms) and ψ (psi, that is, the angle between Cα-C-N atoms) graphical tool for dihedral angle distribution, proposed by Gopalasamudram Narayana Ramachandran and Vasudevan Sasisekharan in 1963, so it is named after them. Ramachandran plots were used to verify the quality of the protein model. If the residue dihedral angles in a protein model are within the allowed region, it usually means that the structure of the model is reasonable. If the φ and ψ corners of the residue are within the disallowed region, there may be an error in the protein model. The Ramachandran diagram can also be used to understand different secondary structures in the protein (e.g. Typical dihedral angle ranges of α-helix, β-sheet) to understand the secondary structure and folding rules of proteins. This laboratory uses the SAVES v6.0 (https://saves.mbi.ucla.edu/) tool developed by UCLA-DOE Laboratory which provides including Ramachandran diagram A variety of analysis functions can comprehensively evaluate the stereochemical quality of the

protein model by analyzing important data in the figure to ensure the accuracy and reliability of the model.

2.3 Determination of active pocket position and mutation site

2.3.1 Determine the active pocket position and mutation site position

The research team of Tomoya Hino and others analyzed the three-dimensional structure of the adenylation domain of L-lysine-activated Pls protein [17]. In the experiment, the Hino team used X-ray crystallography to crystallize the complex of the adenylation domain of L-Lys-activated Pls and adenylate (LAD) binding, with a high resolution of 2.3 angstroms (Å). Resolution resolved its structure, and the results of this work are recorded in the PDB database (PDB ID: 7WEW). This study demonstrates the precise mode of L-Lysine binding to the adenylation domain of Pls, revealing the intermediate state of adenylate (LAD) during activation. Through high-resolution structural images, it is possible to observe the interactions between amino acid residues and how they work together to bind and activate adenylate, providing a basis for understanding how L-lysine is activated through the adenylation domain. Pls proteins provide molecular-level insights.

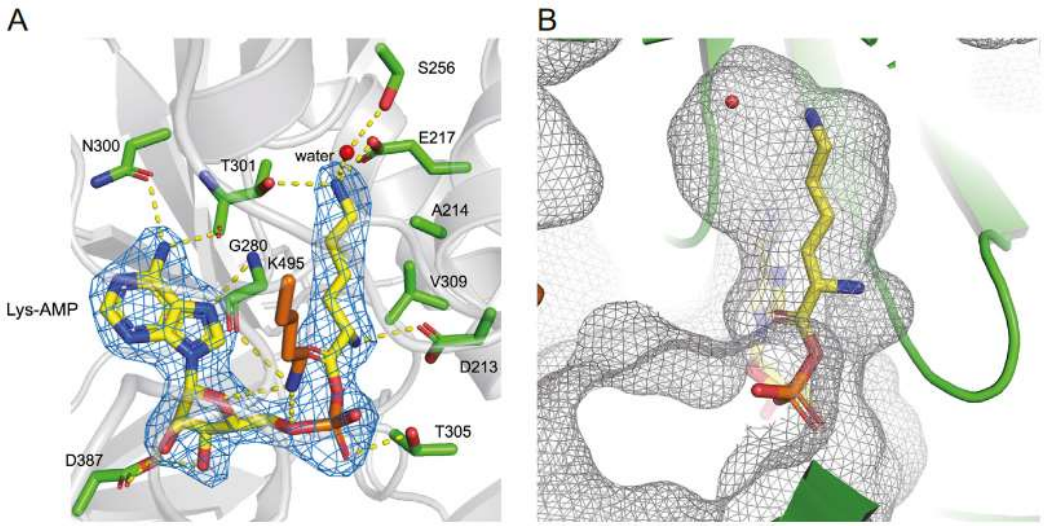


Figure 2.1 - L-lysine binding pocket structure [18]

NOTE: (A) Structure of the substrate binding pocket. Ly5-AMP is shown in stick model format. Amino acid residue side chains that form hydrogen bonds with Lys-AMP and residues within 4 Å of Ly5-AMP are shown in stick model format. These residues are stained green in the N-terminal domain and brown in the C-terminal domain. The water molecules that mediate the interaction between LyS-AMP and Ser256 appear in the shape of red spheres. Hydrogen bonds are represented by dashed lines. The hydrogen bond distance ranges from 2.6 -3.4Å. Blue net shows the palisade omission pattern formed by LyS-AMP at 3.0σ. (B) Grid representation of the substrate binding pocket. The bound Ly5-AMP and water molecules are shown in (A). (used to explain references to color in the figure)

In order to improve the efficiency of substrate molecules leading to the active site, this experiment identified four key sites that play an important role in the reaction channel, namely Arg401, Asp404, Arg370 and Lys499. The amino acid

residues at these sites cause certain obstacles to the smooth entry of substrate molecules into the reaction channel due to their larger side chain structures. The amino acids at the four positions of Arg401, Asp404, Arg370 and Lys499 were replaced with glycine. Due to its small side chain, glycine is a commonly used amino acid in protein engineering and plays a role in reducing steric hindrance and improving flexibility. This solution can effectively promote the transport of substrate molecules and thereby improve enzyme activity.

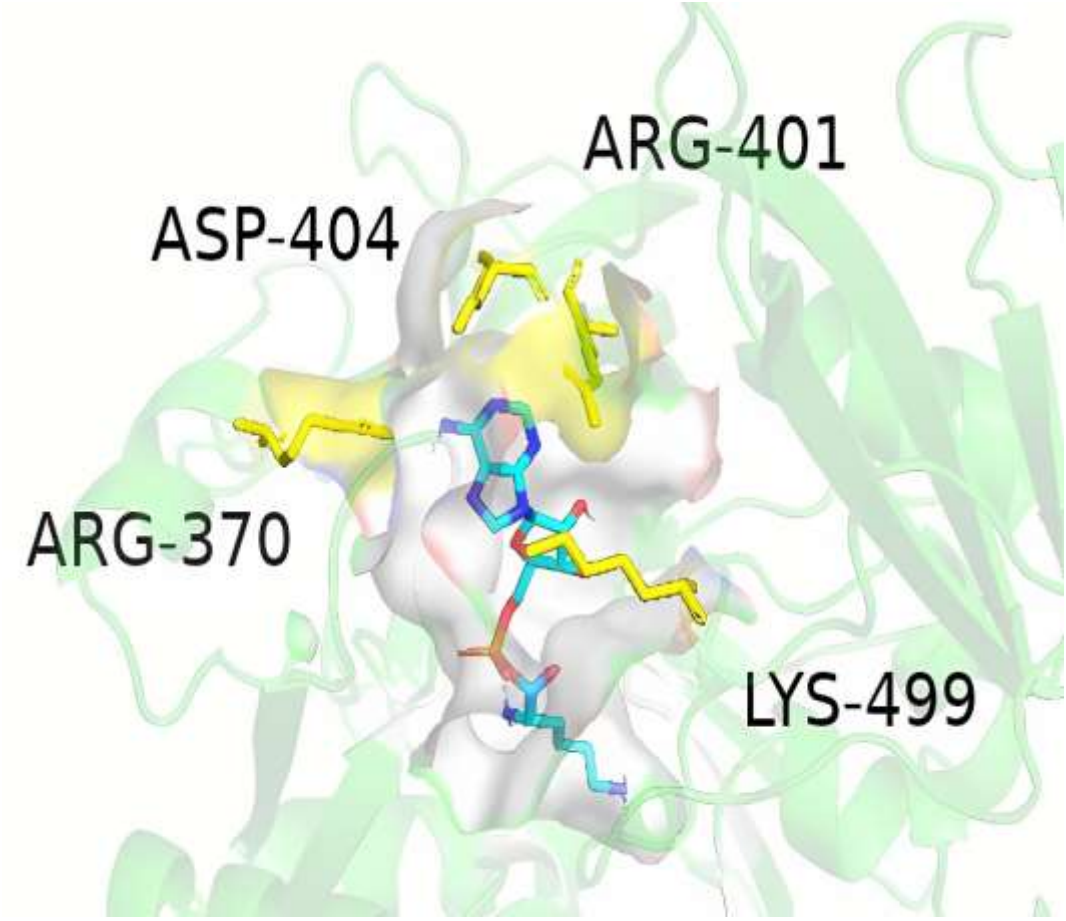
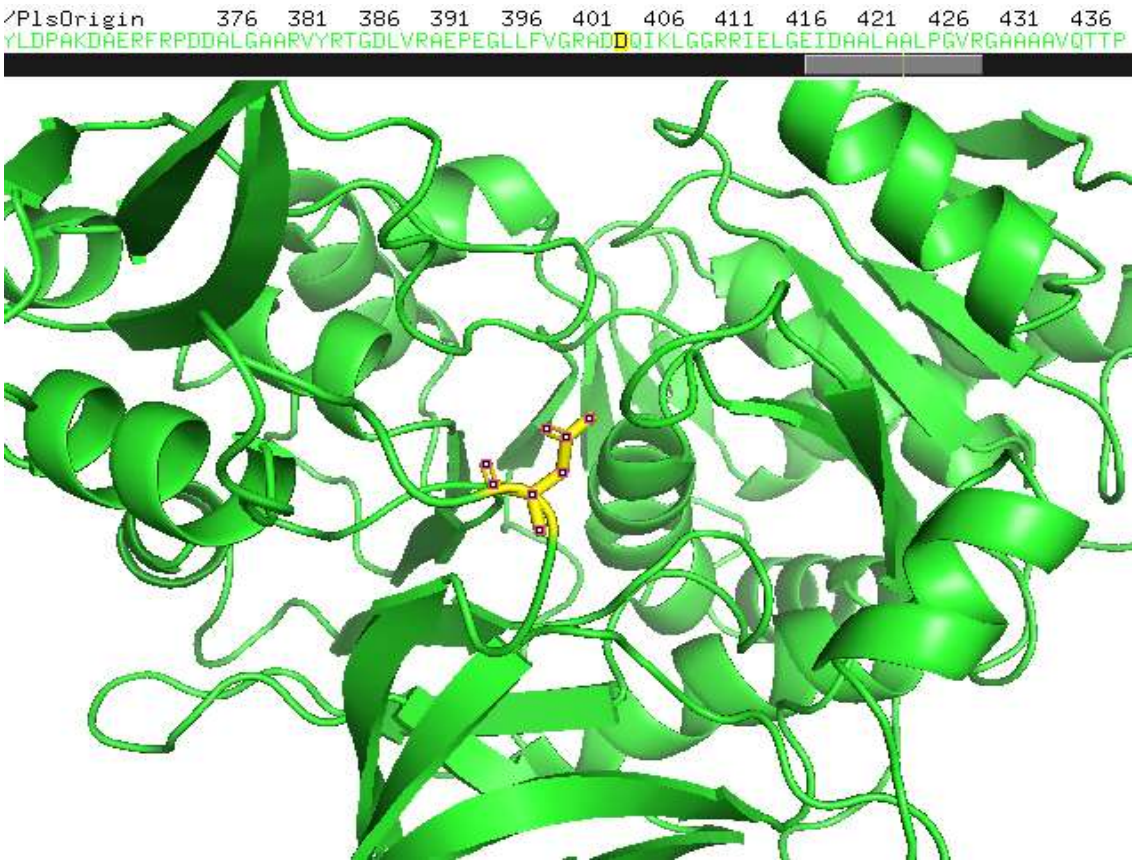


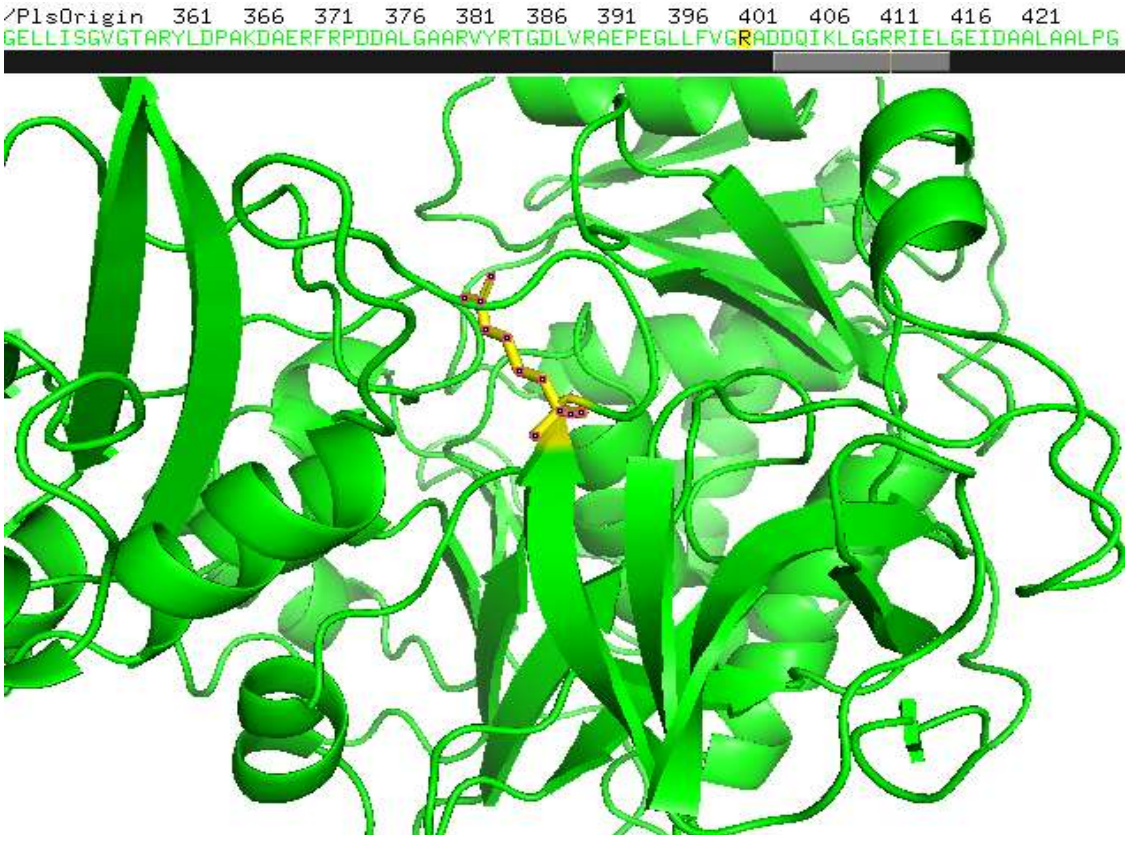
Figure 2.2 - Schematic diagram of the side chain position near the reaction channel

Note: The yellow part in the figure represents the position of the residue.

2.3.2 Gene mutation method

Enter the selection command in the PyMOL command execution interface to locate the position of key residues in the three-dimensional structure. After selecting the display sequence information command, you can determine the position of these sites in the amino acid sequence and edit the amino acid sequence. The amino acid at a key position was mutated to glycine. After mutation is completed, the new amino acid sequence will be used as the mutant amino acid sequence. Homology modeling via SWISS-MODEL.


/PlsOrigin 376 381 386 391 396 401 406 411 416 421 426 431 436
YLDPAKDAERFRPDDALGAARVYRTGDLVRAEPEGLLFVGRADDQIKLGGRRIELGEIDAALAALPGVRGAAAAVQTTP


/PlsOrigin 361 366 371 376 381 386 391 396 401 406 411 416 421
GELLISGVGTARYLDPAKDAERFRPDDALGAARVYRTGDLVRAEPEGLLFVGRADDQIKLGGRRIELGEIDAALAALPG

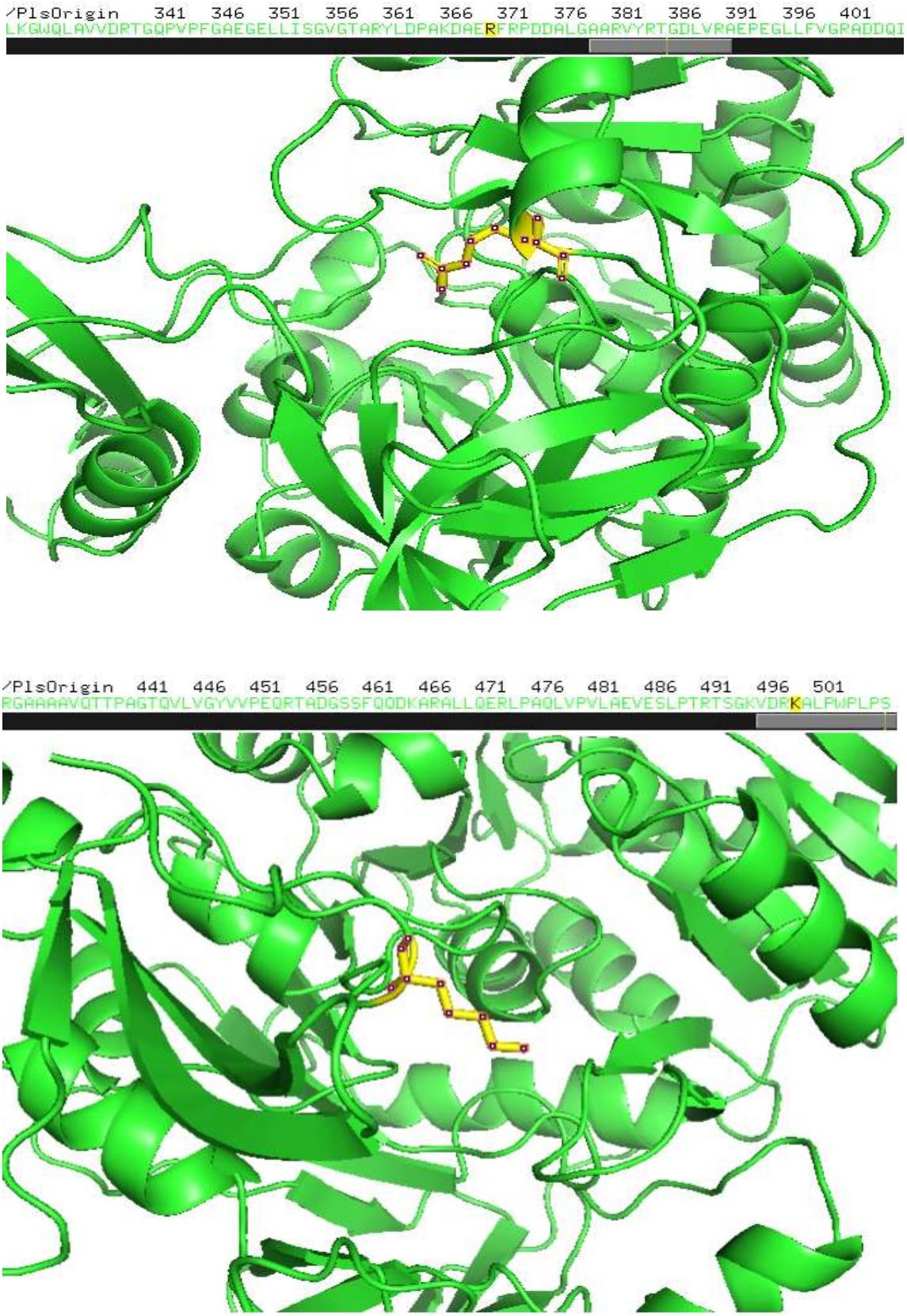


Figure 2.3 - Schematic diagram of pre-mutation site

2.4 Pls Molecular docking with substrate

Molecular docking technology uses computational methods to predict and analyze how one molecule (usually a small molecule ligand or drug candidate) interacts with another molecule (such as a protein or nucleic acid). These interactions include hydrogen bonding, hydrophobic effects, van der Waals forces, and ionic bonds etc. Quantitative calculation of the energies of these interactions can predict whether molecules can form stable complexes and their possible binding positions and affinities. As an advanced docking tool, CB-Dock2 incorporates a new molecular docking technology - Fit Dock, which identifies and extracts similar docking patterns from existing protein-ligand complex databases, and applies these patterns to During the docking process of target protein and ligand [18]. The molecular docking process of CB-Dock2 is as follows: after importing the precise three-dimensional structure of the ligand and receptor, the molecular docking work is performed. First, the active site or binding pocket that may interact with the ligand is identified in the receptor. Taking into account the possible flexibility of the ligand and receptor, CB-Dock2 performs a conformational search to identify possible binding modes. Once identified, computational methods are used to place the ligands in the active site of the receptor, simulating their interactions. A variety of software tools used in the process of performing molecular docking include but are not limited to AutoDock, DOCK, Glide, and Maestro and Gold in the Schrödinger software suite. They provide different algorithms and user interfaces to meet different research needs. With the rapid growth of structural information in protein databases, docking simulation

technology has been significantly improved by drawing on known protein-ligand complex structural information to increase the scientific nature and accuracy of docking results [19]. This experiment uses CB-Dock2 for molecular docking to analyze and discuss the Vina score and Cavity volume data changes of the four mutants obtained.

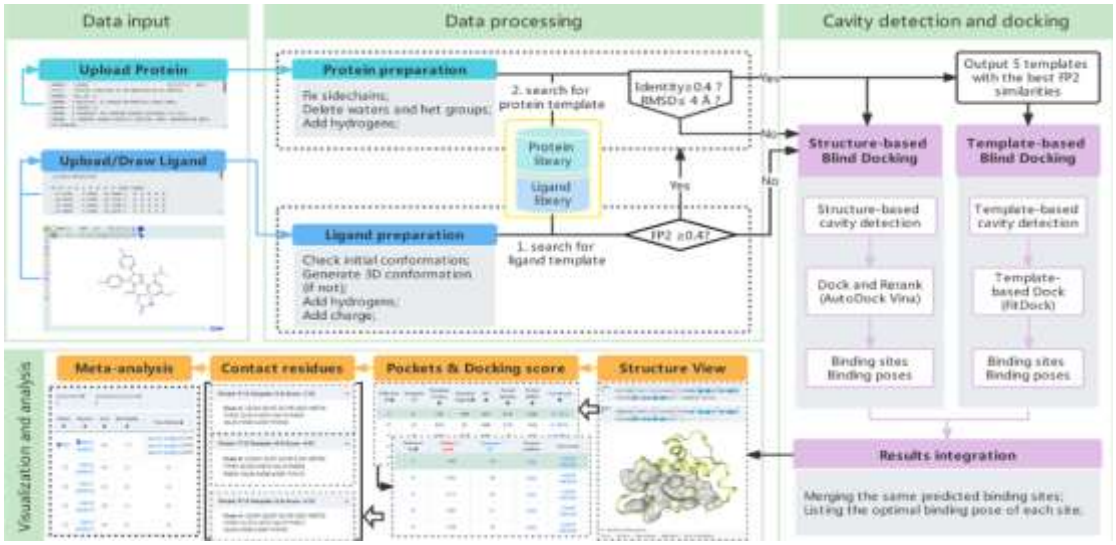


Figure 2.4 - CB-Dock2Overall operation logic diagram [20]

2.5 Molecular Dynamics Simulation

GROMACS is a powerful molecular dynamics simulation software, mainly used to study the dynamic behavior of biological macromolecules such as proteins, lipids, and nucleic acids in solution [21]. The software can not only perform classic atomic-level molecular dynamics simulations, but also handle larger-scale systems, such as coarse-grained models, and perform longer simulations. It has optimized its algorithm in continuous updates to Supports high-performance computing

environments, enabling efficient parallel processing on multi-core processors and GPUs, significantly speeding up simulations [22]. The software supports a variety of force fields, allowing researchers to Target the appropriate model for simulation; provide flexible input file formats to easily define complex biomolecular systems, and output data in multiple formats at the same time to facilitate subsequent analysis. In addition to the simulation itself, GROMACS also has a variety of built-in analysis tools for calculating energy, trajectory, structure and other properties to help researchers understand molecular behavior in depth. GROMACS is widely used in biophysics, structural biology, drug design and other fields, providing valuable tools for scientific research [23]. The software R&D team regularly releases new versions, introducing new features and improvements to keep the software cutting-edge and competitive [24].

2.5.1 Construction of initial system of simulated reaction solution

In the initial stage of molecular dynamics simulation, the initial conformation of the complex plays a decisive role in the accuracy and scientificity of subsequent simulations. In this regard, we have carried out the following work:

1) Protein processing: Use CHARMM36 force field to process proteins. This force field is specially designed for biological macromolecules and is widely used to simulate proteins, nucleic acids, lipids, etc.

2) Small molecule substrate processing: First use Avogadro software to hydrogenate the small molecule substrate and convert it to mol2 format. Since the

CHARMM36 force field does not contain predefined parameters for small molecules, we generated the required force field parameters through the CGenFF website(https://app.cgenff.com/). CGenFF is part of the CHARMM force field series, which focuses on the parameterization of small molecules and uses chemical topology methods to predict force field parameters.

3) Simulation box construction: Choose a cubic-shaped simulation box and ensure that the minimum distance between the molecule and the box boundary is 1.0 nanometers to avoid restricted molecular motion or collision with the boundary, thus improving the reliability of the simulation.

4) Water model selection: Use the SPC216 water model, which can effectively simulate the structure and dynamic characteristics of water while maintaining high computational efficiency and is suitable for large-scale simulations.

5) Ion addition: Add an appropriate amount of ions to the system in order to maintain the electrical neutrality of the system or adjust the required ion concentration.

6) Energy minimization: After the simulation system is constructed, energy minimization is performed to eliminate high-energy defects in the initial structure and obtain a more reasonable and stable starting conformation. Adopt conservative parameter settings to ensure that the architecture is optimized within a reasonable amount of time.

2.5.2 Molecular Dynamics Simulation

In order to ensure that the system successfully completes the molecular dynamics simulation operation, the system needs to be energy minimized, that is, to maintain a rational configuration so that the distance between atoms in the model is not too close. The system is first energy minimized using the steepest gradient method and a convergence criterion of 10.0 kj/mol. Then the restricted dynamics of 1000 ps NVT temperature control and NPT pressure control were carried out respectively, and finally 5000 pwas carried out using the leap frog integration algorithm. Dynamic simulation calculation of finished product. After the simulation is completed, you will get many result files name dnpt-nopr. Among these files, the trajectory file (trr) is particularly important. It is a file that saves various parameters. It is crucial for subsequent analysis of results. The trjconv command is used to compress the trr file, which not only saves hard disk space, but also increases the calculation time by using the xtc file for calculation. After the enzyme and substrate reach equilibrium under simulated conditions, the system parameters are analyzed. The steps for molecular dynamics simulation are as follows:

1) File type conversion: Use the program pdb2gmx to convert the PDB coordinate file into a gromacs file to generate three files, a molecular topology file, a position restriction file and a post-structure production file.

2) Define the simulation box: Use the program editconf to define the size of the simulation box, and place the protein in the center of the rectangular box so that the distance between the atoms on the protein surface and the wall of the box is at least

1.0 nm.

3) Add solvent molecules: Use the program genbox to addSPC216type to the defined box water molecule. This program requires a run file with the extension.tpr, which is generated by the grompp command. The command integrates the parameters specified in the mdp type file to generate a structure and topology file.tprfile.

4) Add ions to ensure electrical neutrality: Use the program genion to add counterions (sodium ionsor chloride ions).

5) Energy minimization: Use the programs gromppandmdrunto minimize the energy of the system, making the position between water molecules and protein structures in the system more reasonable. First run1000step steepest descent method to make the system energy reach the threshold1000 KJ/(mol·nm). On this basis, perform3000step conjugate gradient method.

6) Constrained molecular dynamics simulation: Constrain the protein structure in the system, allowing the movement of solvent molecules, allowing the solvent molecules to enter the interior of the molecule. Run10000 steps in total,20 ps in total, electrostatic interaction adopts Particle Mesh Ewald(PME) algorithm, the van der Waals force constraint interception radius is1nm.

7) Free molecular dynamics simulation: run a total of 500000 steps with step size 2 fs, totaling1ns. The electrostatic interaction adopts the Particle Mesh Ewald(PME) algorithm, and the van der Waals force constraint interception radius is 1 nm.

8) Data analysis: Use the calculation tools provided by Gromacs to analyze the

simulation trajectory. Use the programg_rmsd to calculate the root mean square deviation of the trajectory(root mean square deviation, RMSD)value; in terms of free molecules The structure before dynamics simulation is used as the reference structure; g_hbond analyze the presence of hydrogen bonds in the molecular dynamics simulation trajectory.

CHAPTER 3

EXPERIMENTAL PART

3.1 Pls amino acid sequence and Acquisition of adenylation domain information

Select the protein amino acid sequence displayed in FASTA format from the NCBI search results.

Figure 3.1 - Amino acid sequence of Pls

Search the PDB database for structural information of the Pls adenylation domain

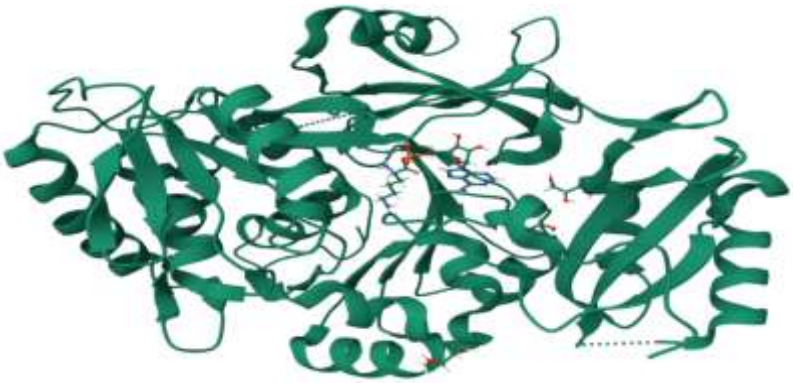


Figure 3.2 - PlsAdenylation domain

3.2 Pls Homology modeling result analysis

3.2.1 Homology modeling of Pls through SWISS-MODEL

Select the modeling model with Epsilon-poly-L-lysine synthase (PDB ID: 7WEW) as the template in the modeling results, because the source of structural information in the PDB database is in recent years. The latest research progress and its high structural resolution and experimental verification, the modeling results are reliable.

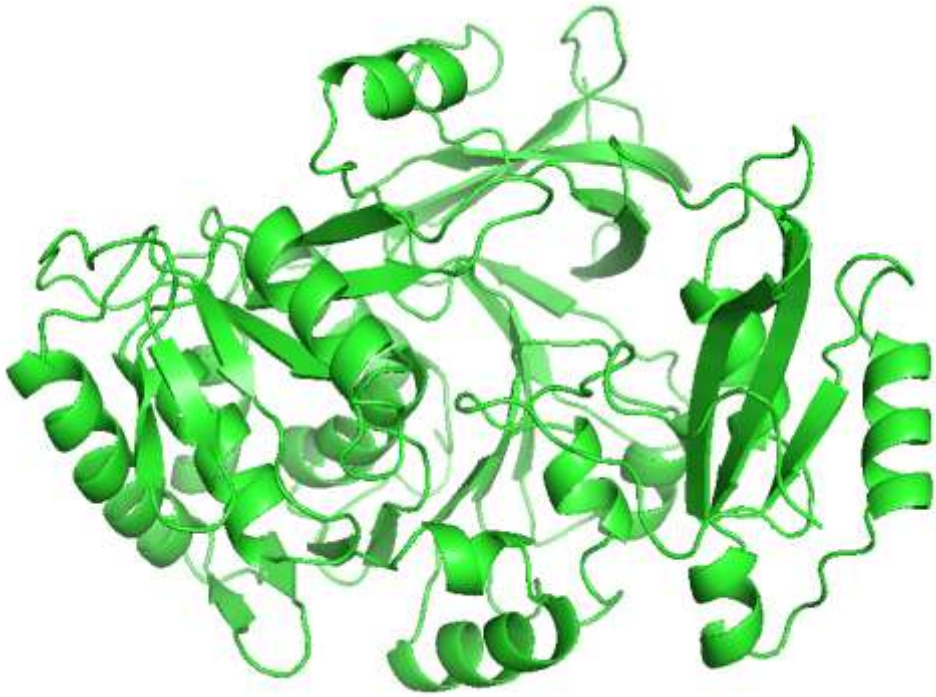


Figure 3.3 – Pls Modeling results

3.2.2 Ramachandran Plots Data Analysis

The allowed area in the Ramachandran diagram will show the conformational areas where amino acid residues in the protein structure can theoretically appear, and the disallowed area will show the theoretically impossible or unstable conformational areas of the amino acid residues. By analyzing the protein structure The φ and ψ angles of all residues can provide an understanding of the distribution of various amino acids in the protein. Among them, 371 residues are located in the most popular regions. These regions are represented by [A, B, L], accounting for 93.7% of the number of non-glycine and non-proline residues, indicating that the φ-ψ angle of most residues falls within Being within the statistically most common range is a sign of good quality protein structure models. 21 residues are located in the additional

allowed regions, which are represented by [a, b, l, p], accounting for 5.3%. These regions are acceptable, although not the most common, and often appear in protein structures. Three residues are located in the generously allowed region, represented by [~a,~b,~l,~p], accounting for 0.8%. Residues in these regions may represent some structural variation or flexibility. One residue is located in the disallowed region, accounting for 0.3%, which may indicate that the φ-ψ angle configuration of this residue is rare in normal protein structures and may be a sign of model errors or structural abnormalities. The number of non-glycine and non-proline residues is 396. Adding glycine and proline residues, the total number of residues is 490. After excluding glycine and proline, the number of terminal residues is 2. Terminal residues lack the constraints of adjacent residues and may have a wider conformational distribution. Glycine residues number 50 and are usually represented by triangles. Due to the simplicity of the side chains, they often appear in flexible regions of protein structures. Proline has 42 residues and contains rigid side chains, which can cause the protein chain to bend or twist. Overall, most of the model residues are located in the most popular regions, indicating that the model quality is high.

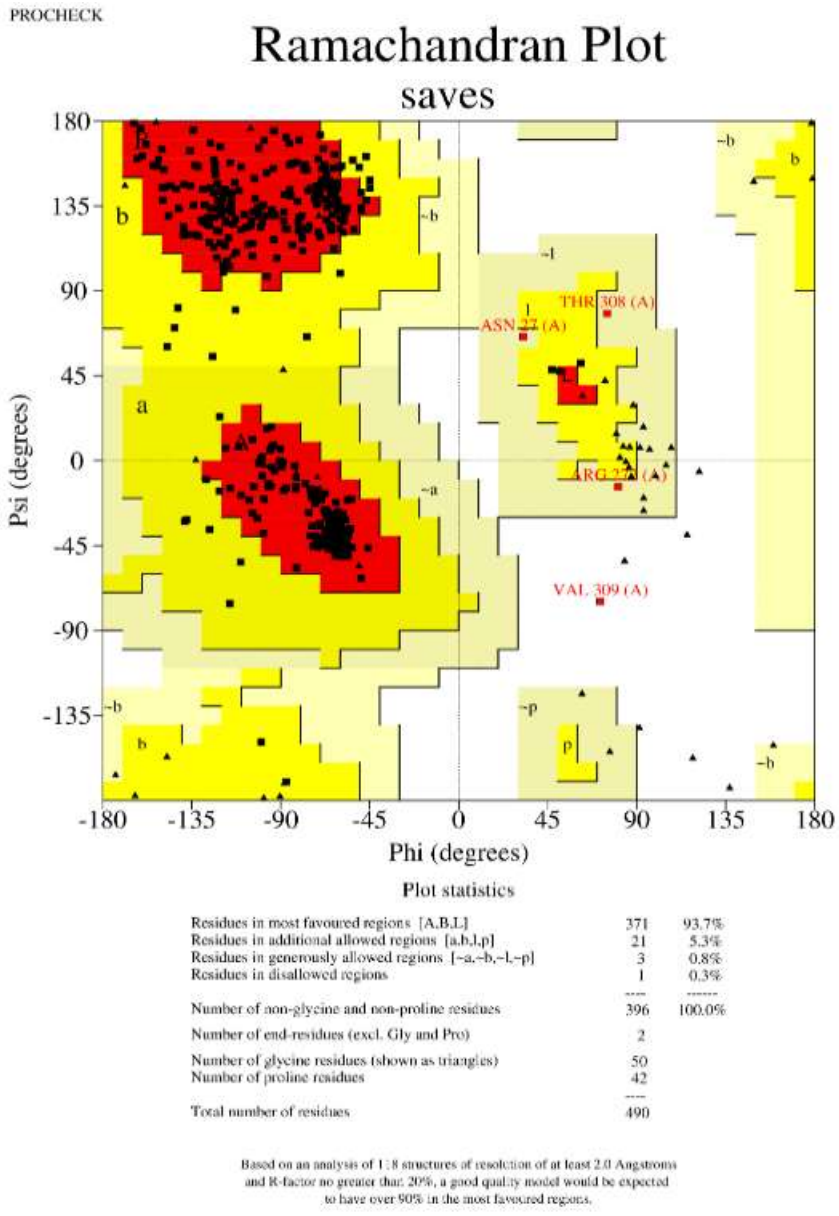
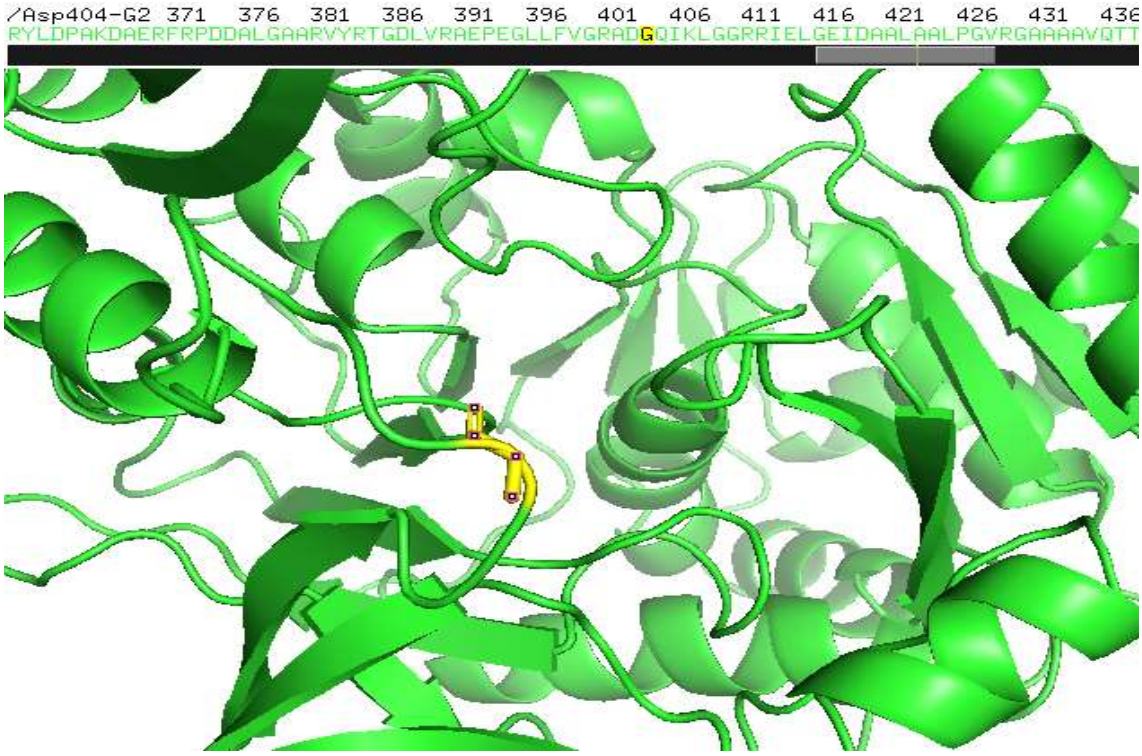
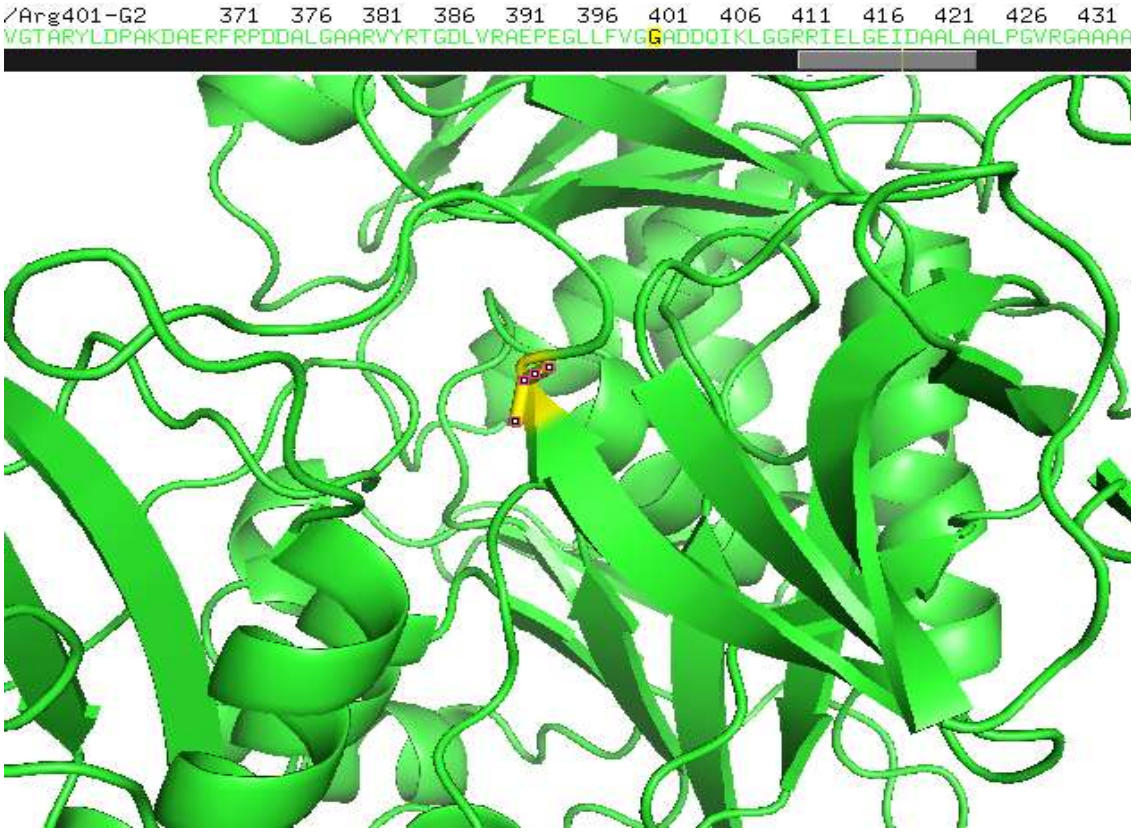


Figure 3.4 - Ramachandran diagram of Pls

3.3 Modeling results of Pls mutants and data analysis of Ramachandran Plots

3.3.1Modeling results of mutants


/Asp404-G2 371 376 381 386 391 396 401 406 411 416 421 426 431 436
RYLDPAKDAERFRPDDALGAARVYRTGDLVRAEPEGLLFVGRADGQIKLGGRRIELGEIDAALAALPGVRGAAAAVQTT


/Arg401-G2 371 376 381 386 391 396 401 406 411 416 421 426 431
VGTARYLDPAKDAERFRPDDALGAARVYRTGDLVRAEPEGLLFVGGADDQIKLGGRRIELGEIDAALAALPGVRGAAAA

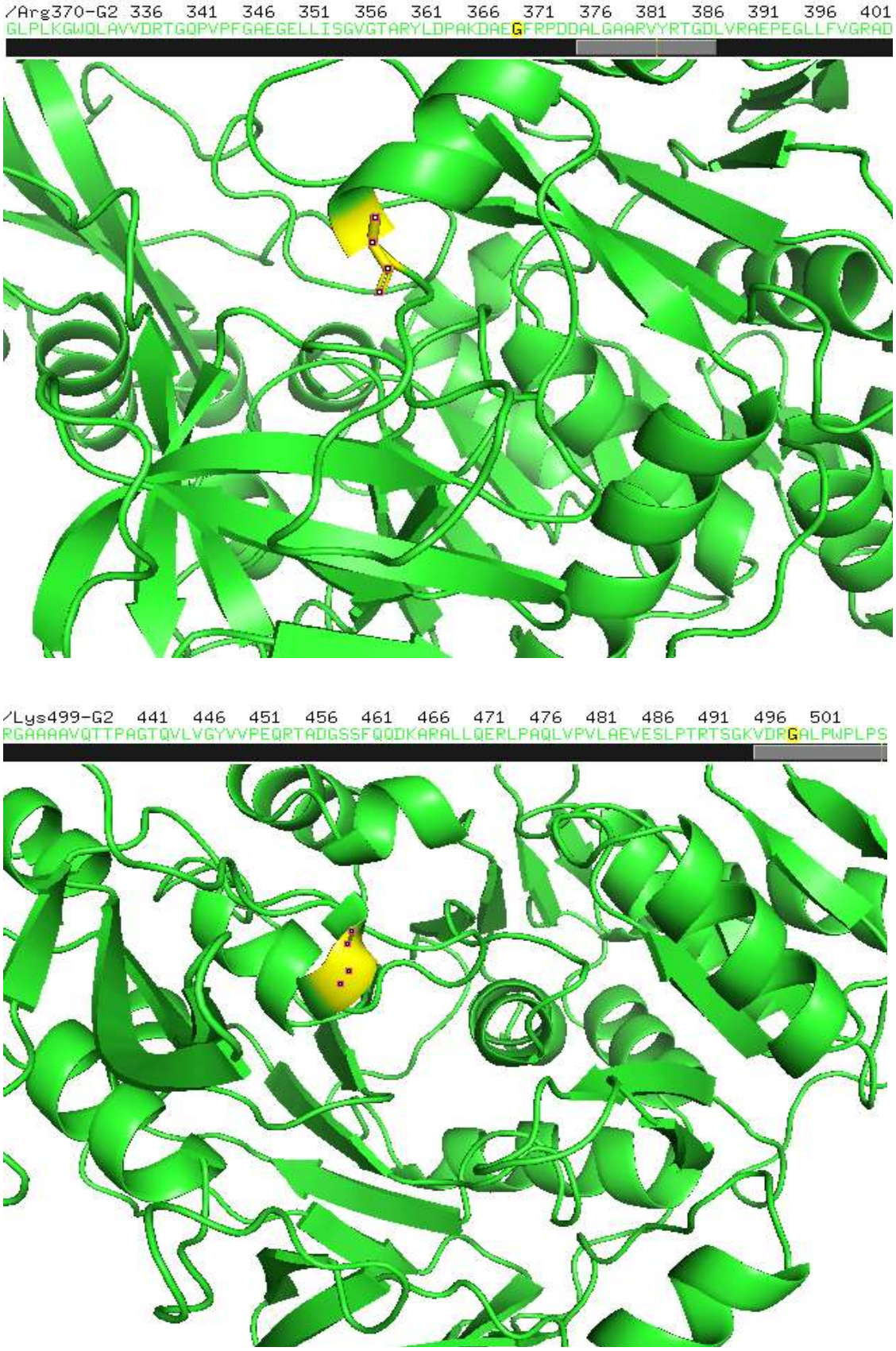


Figure 3.5 - Schematic diagram of the mutation site

By comparing the structure of the key site residues before mutation, we can intuitively find that the steric hindrance caused by the key site residues after mutation is greatly reduced, which helps to increase the rate of substrate molecules entering and exiting the active pocket.

3.3.2 Mutant Ramachandran PlotsAnalysis

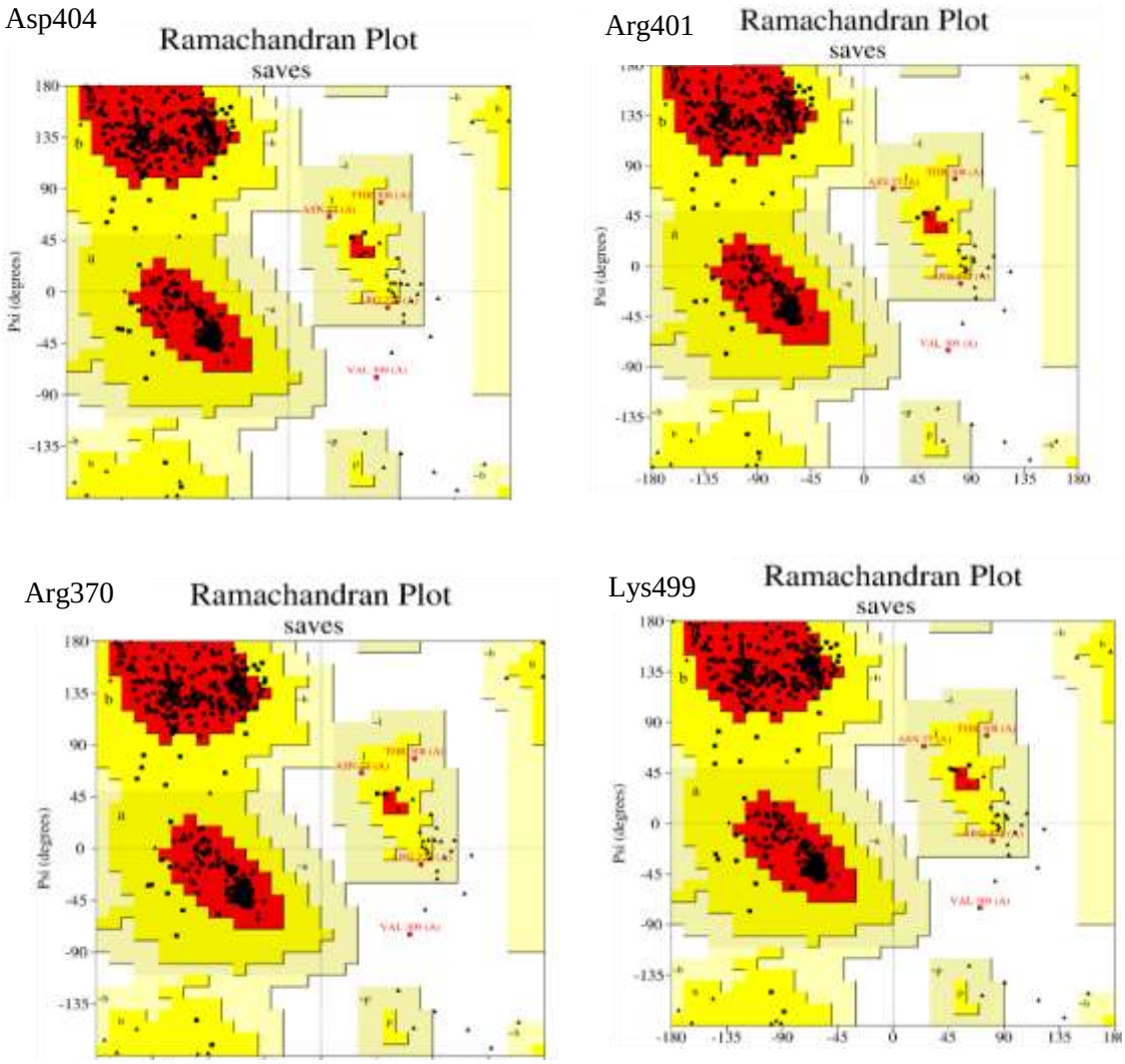


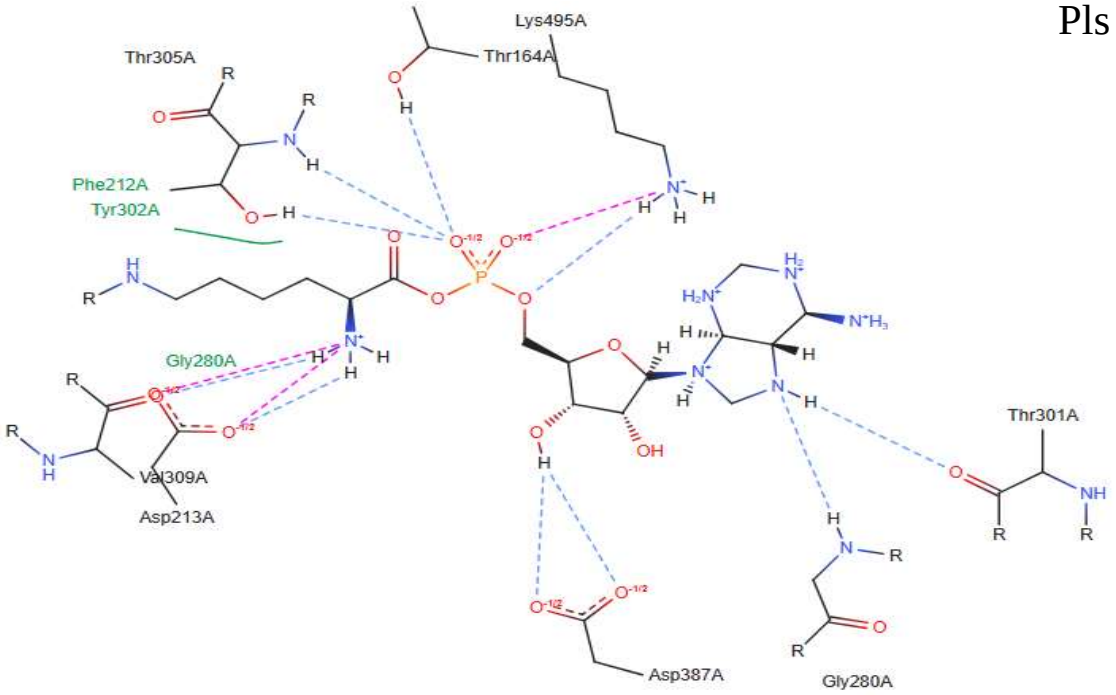
Figure 3.6 - Ramachandran plot of four mutants

Through data analysis of the Ramachandran plots of the four mutants, it was concluded that most of the amino acid residues of the four mutants are located in the most popular regions, accounting for non-glycine and non-proline residues The number of is above 90%, indicating that the φ-ψ angle of the residues falls within the statistically most common range. both have5.3% of residues are in the extra allowed region, which is not the most common but acceptable. The amino acid residues located in the generous allowed region account for 0.8%, and the amino acids located in the disallowed region are only1, accounting for 0.3%. Overall, most of the residues of the four mutant models are located in the most popular regions, and the quality of the constructed mutant models is high.

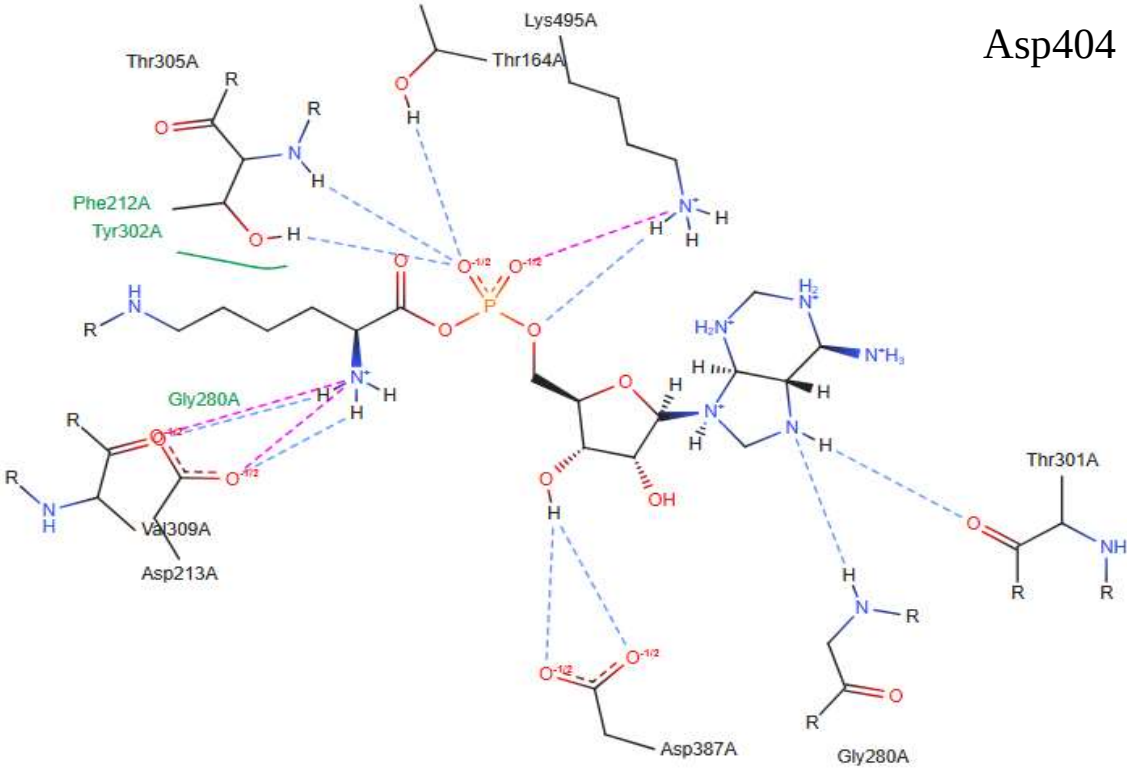
3.4 Analysis of molecular docking results

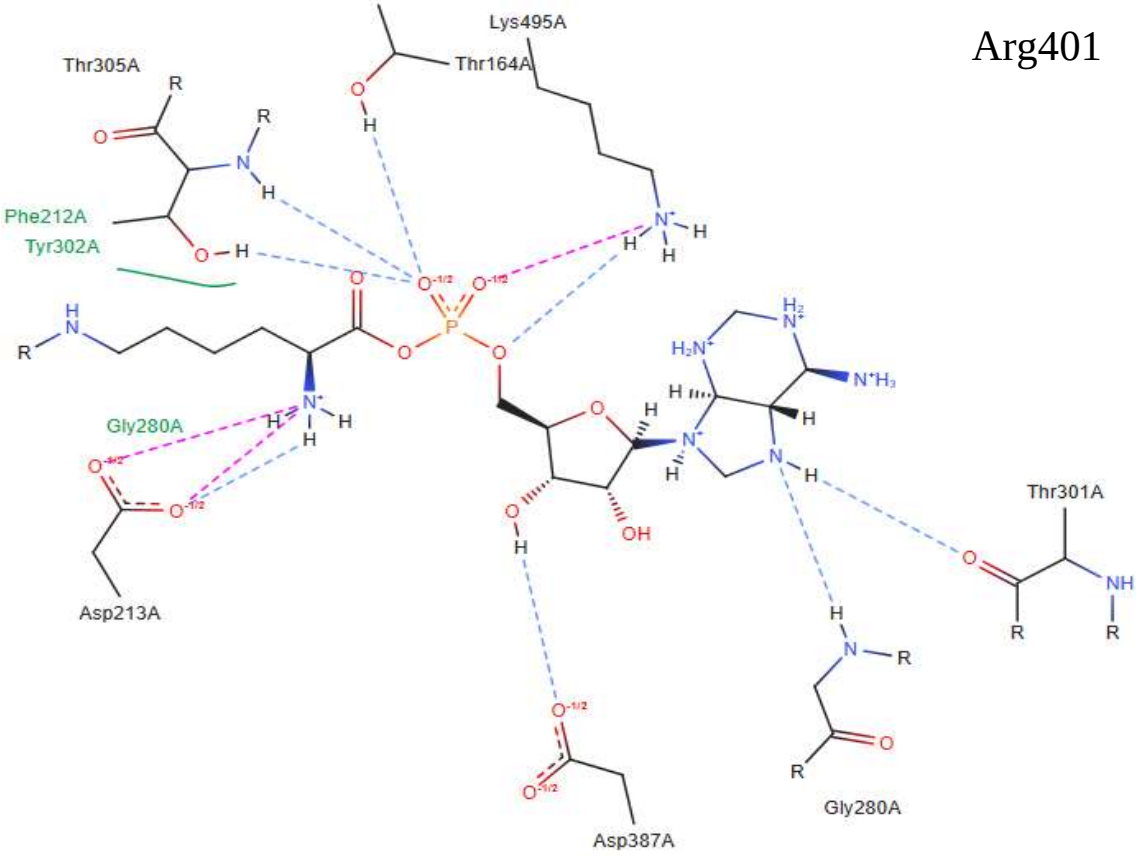
3.4.1 Protein-ligand interaction diagram

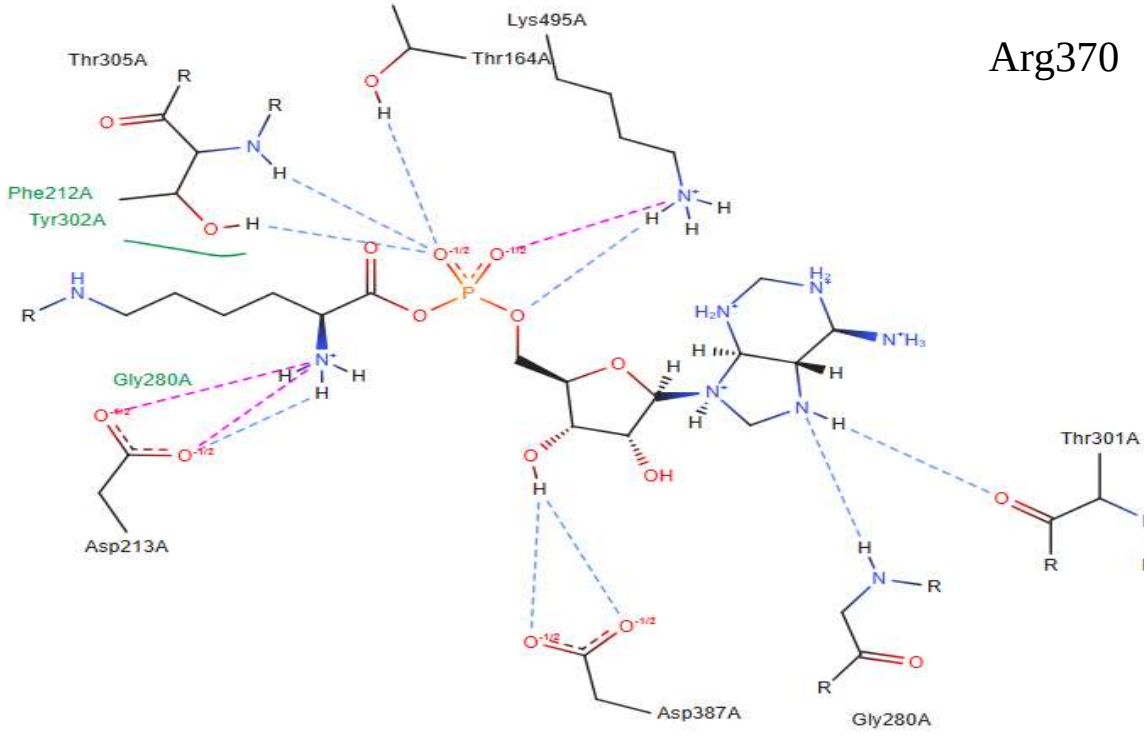
Pls



Asp404




Arg401
Lys495A
Thr164A
Thr305A
Phe212A
Tyr302A
Gly280A
Asp213A
Thr301A
Asp387A
Gly280A


Arg370
Lys495A
Thr164A
Thr305A
Phe212A
Tyr302A
Gly280A
Asp213A
Thr301A
Asp387A
Gly280A

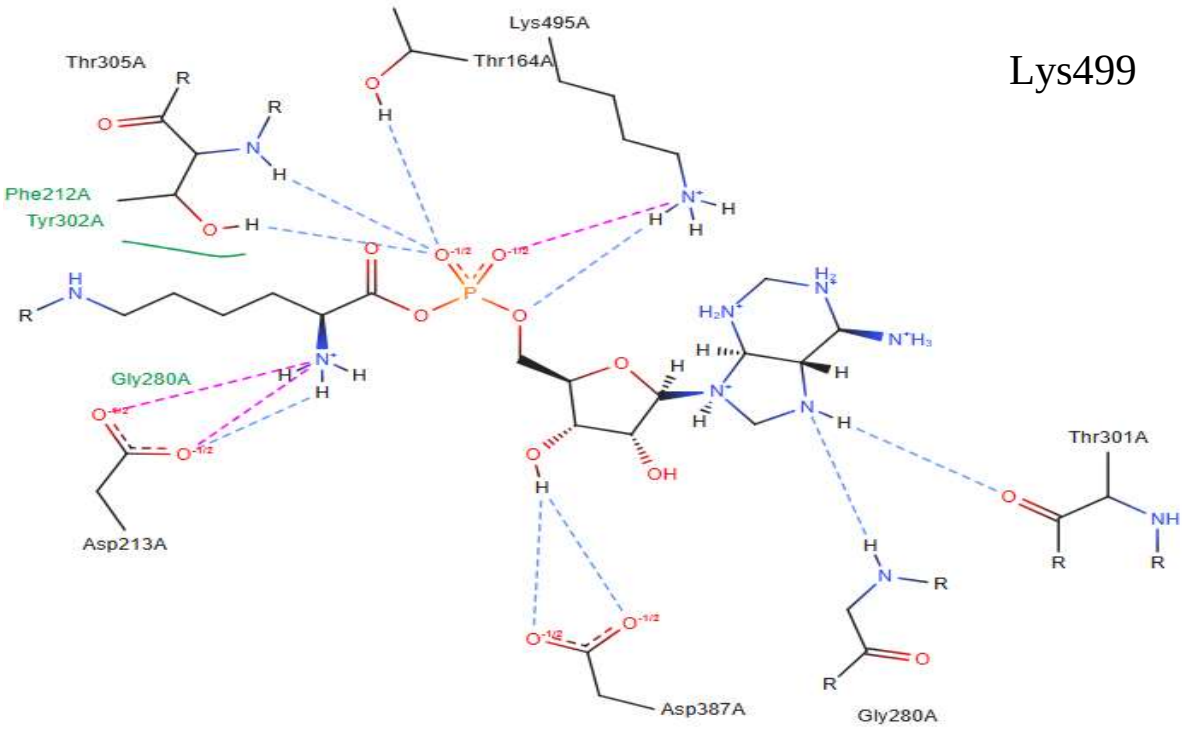
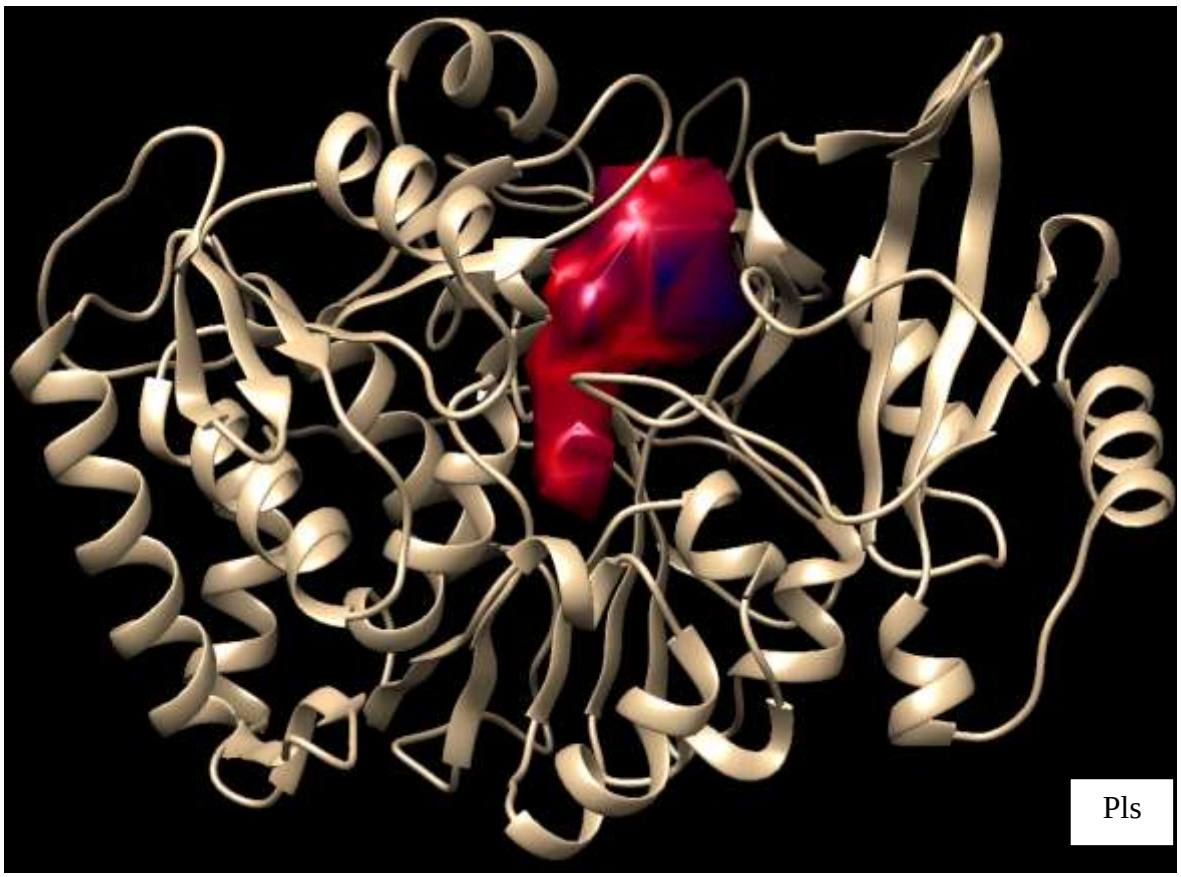
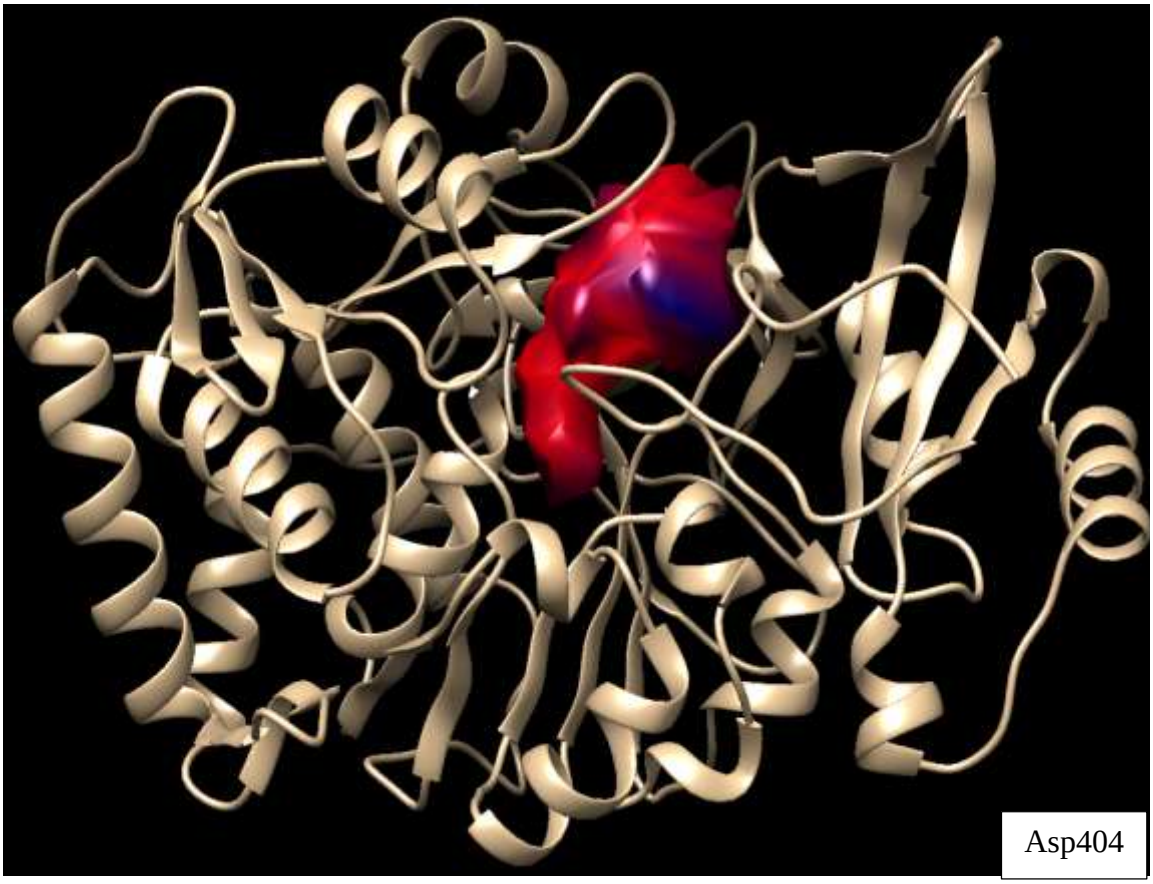
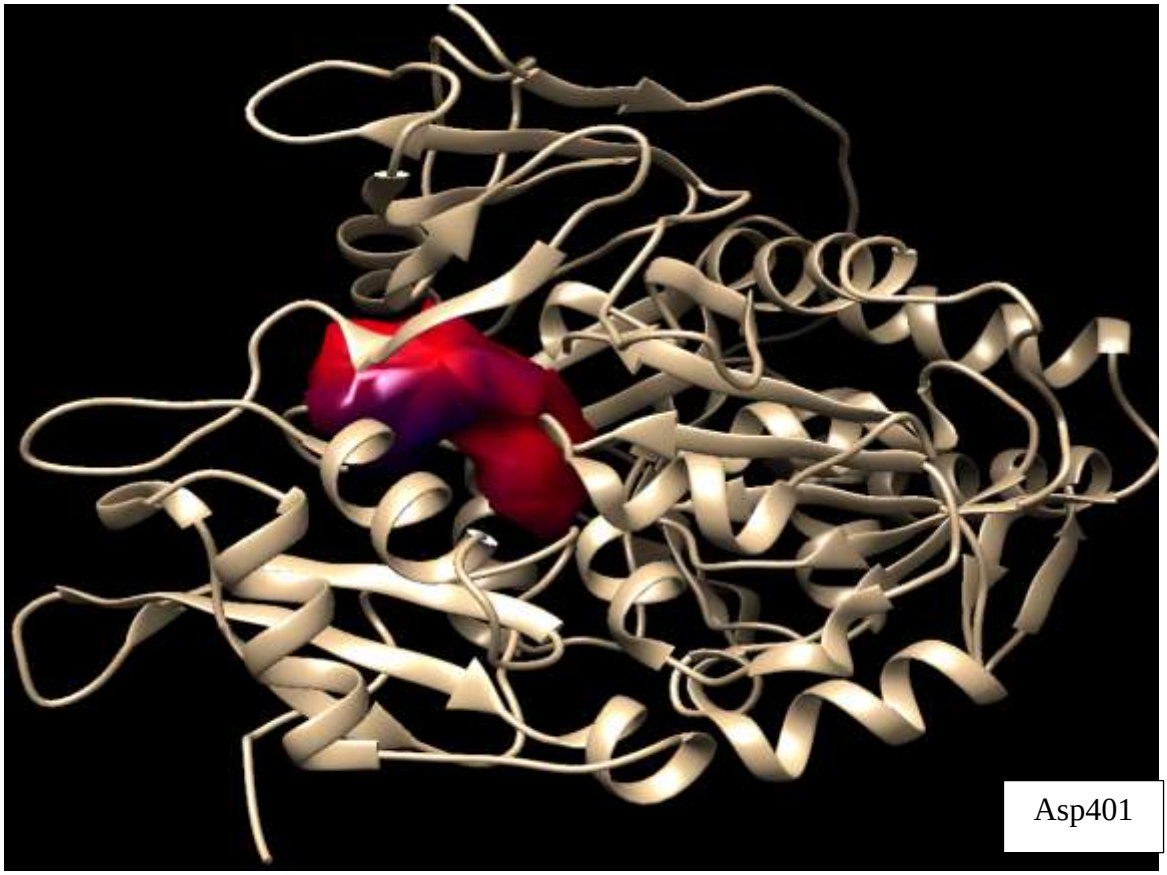


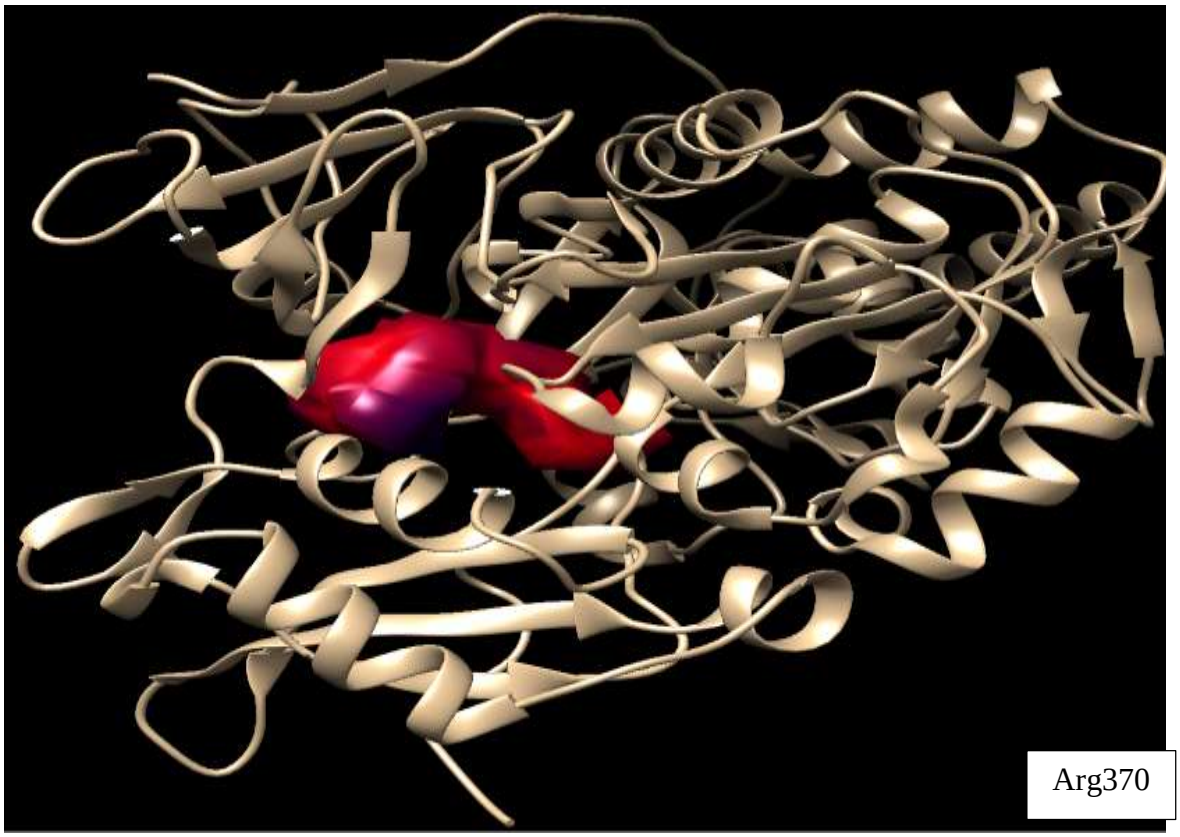
Figure 3.7 - Protein-ligand interaction diagram

3.4.2 Analysis of active pocket volume and binding energy changes




Asp404


Asp401



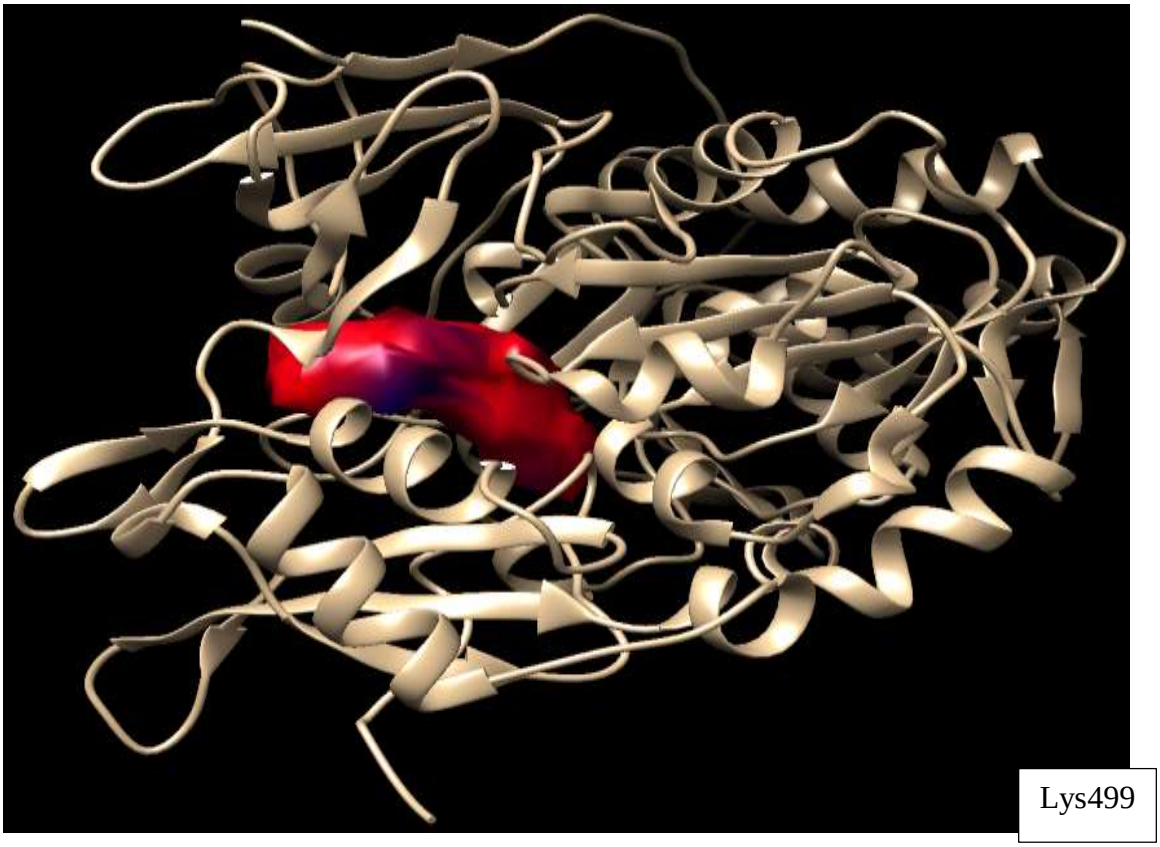


Figure 3.8 - Active pockets of different mutants

Table 3.1 - **Molecular docking affinity score and cavity volume**

Type of enzyme	Docking affinity score	Cavity volume
Types of enzymes	Vina score /kcal.mol^{-1}	Cavity volume /Å^3
Pls	-7.4	5043
Asp404	-7.9	5047
Arg401	-7.9	3864
Arg370	-7.8	4737
Lys499	-7.4	4932

In molecular docking, the lower the Vina score, the stronger the binding force and the higher the affinity between the ligand and the receptor. This score is obtained by evaluating different conformations of the ligand in the active site of the receptor, aiming to predict the stability and strength of the ligand binding to the receptor; Cavity volume refers to the void space inside a protein or other biomolecule. The volume of a cavity or pocket. These cavities are usually formed by the side chains of amino acid residues. They form specific spaces in the three-dimensional structure of the protein and can accommodate small molecule ligands, ions or other molecules. The size and shape of the cavity affect its interaction with the ligand. With regard to the binding capacity, a larger cavity may be able to accommodate more ligand molecules, thereby affecting enzyme activity.

Except for the mutant at the Lys499 site, which did not show significant enhancement compared with the original state, the mutants at other sites showed an increase in binding capacity compared with the original state. The Vina Score of the Asp404, Arg401 and Arg370 mutants The values were reduced by 0.5, 0.5 and 0.4 respectively. It is worth mentioning that the volume of the active pocket of the Asp404 mutant has increased compared with the original enzyme, but this change has

not been seen in other mutants. This shows that the mutation of the Asp404 point may be more conducive to the entry and exit of small substrate molecules, thereby Increase the rate of reaction. In order to further verify the reliability of these observations, we performed molecular dynamics simulations of the system and verified the effects of mutations through more dynamic parameter analysis.

3.5 Molecular dynamics data analysis

3.5.1 RMSD Analysis

In molecular dynamics simulations, RMSD (Root Mean Square Deviation) is an important parameter that measures the stability of proteins or other molecular structures during the simulation process. RMSD is used to quantify the difference between the structures of two molecules, or the structural differences of the same molecule at different points in time. RMSD calculates the deviation of the position of all atoms in a molecule at a certain point in time from its position at a reference time point (usually the starting point of the simulation) and finds its root mean square value. The smaller the RMSD value, the smaller the degree of deviation of the atoms from their initial position, so it can be considered that the molecule maintains a higher structural stability during the simulation process. On the contrary, large RMSD values may indicate that the molecular structure has changed significantly during the simulation, or that the simulation has not yet reached equilibrium.

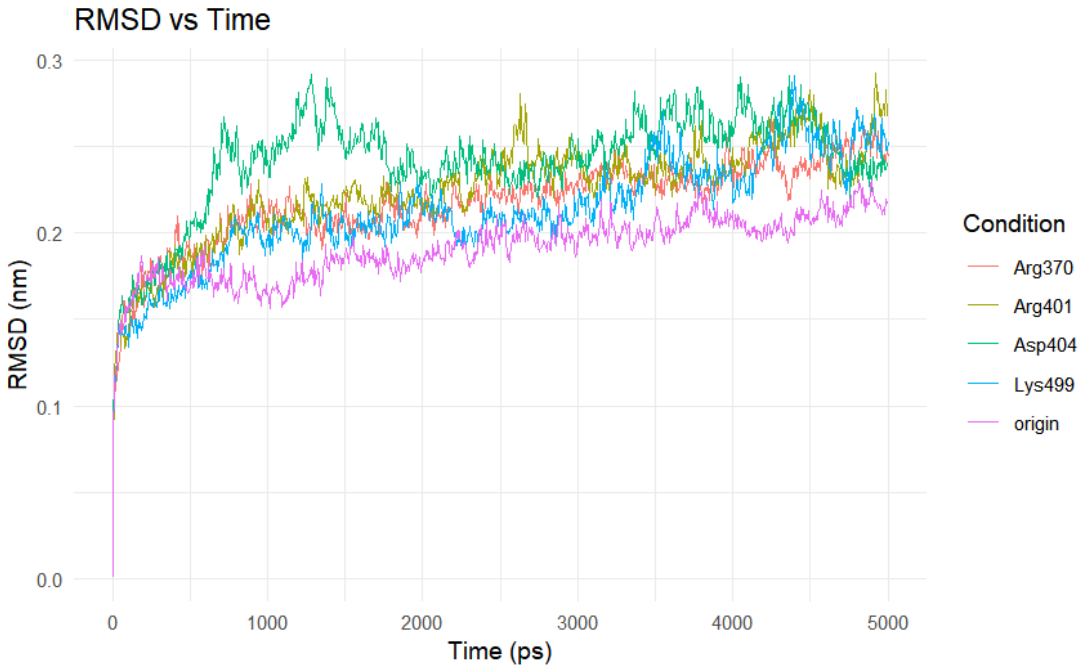


Figure 3.9 – Mutant RMSD Data comparison graph

By comparing the RMSD data of each mutant and the mutant, we can clearly see that during the entire catalytic process, the overall fluctuations and peaks of Asp404 have been most significantly improved, and its peak reached about 0.28nm, which shows that its atoms The position changes during the catalytic process are very strong, that is, the overall structure of the Asp404 mutant has undergone major changes, indicating that the activity of the Asp404 mutant has been greatly improved during the catalytic process, and it has the best performance in the entire catalytic process. The performance of the Arg401 mutant in the entire catalytic process is second only to that of the Asp404 mutant. The overall fluctuations and peak increases in the RMSD data of Arg401 in the early stage of the catalytic process are not as dramatic as those of the Asp404 mutant, but there is a certain improvement compared with before the mutation. In the catalytic process, The performance in the middle stage of the process is the most outstanding, second only to Asp404 in the overall reaction process. The performance of the Arg370 mutant in the entire catalytic process was relatively stable, with an overall improvement compared to before the mutation, but there was a certain gap with the Asp404 and Arg401 mutants, and the

overall performance was inferior to the Asp404 and Arg401 mutants. Compared with the pre-mutation catalytic process, the overall RMSD fluctuation and peak value of the Lys mutant decreased slightly in the early stage but improved in the middle and late stages. Compared with the other three mutants, the Lys mutant ranked last in the early and middle stages of the reaction, but in the late stage of the reaction The performance is slightly higher than that of Arg401. Overall, the Lys mutant is only better than before the mutation, but inferior to the other three mutants. In summary, the performance of each mutant in the entire catalytic reaction process is: Asp404 mutant>Arg401 mutant>Arg370 mutant>Lys mutant.

3.5.2 Hydrogen BondsAnalysis

Hydrogen Bonds are a special type of intermolecular force that play an important role in the structure and function of biological molecules, such as water and proteins. In molecular dynamics simulations, the identification and analysis of hydrogen bonds can help Researchers understand protein folding, stability, and interactions with other molecules, such as ligands or solvent molecules. The software provides tools to identify and quantify hydrogen bonds. These tools can calculate the number, lifetime and geometric parameters of hydrogen bonds, etc., thus providing important information for understanding molecular behavior in complex biological systems. This experiment focuses on analyzing the data of RMSD and Hydrogen Bonds to reveal the impact on enzyme activity before and after mutation.

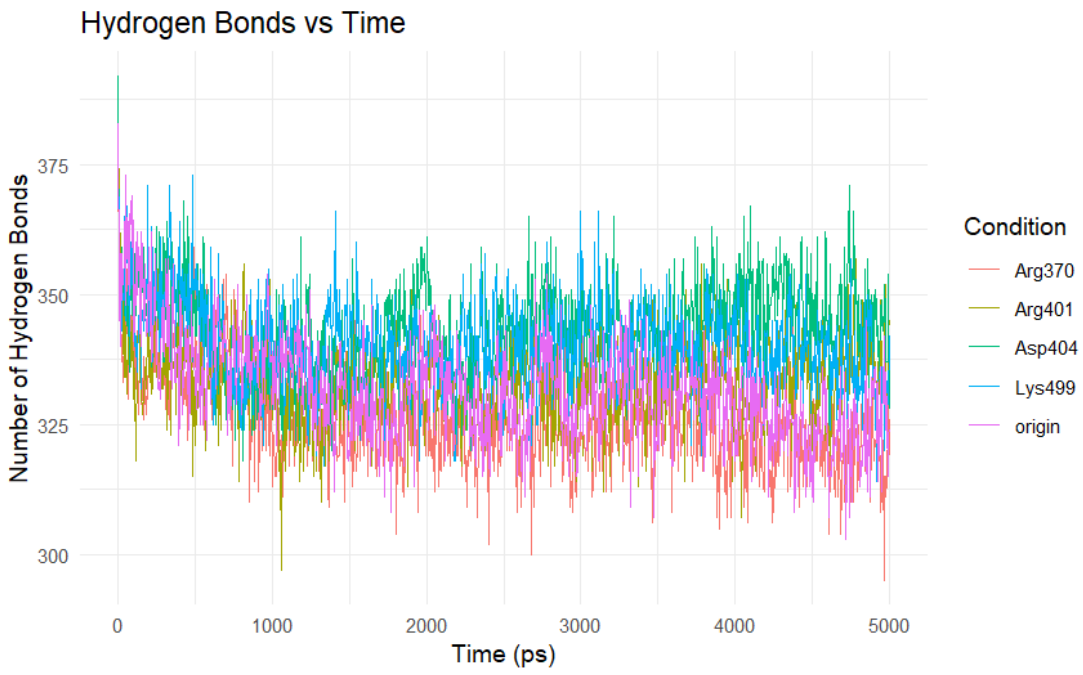


Figure 3.10 - Comparison of changes in the number of hydrogen bonds in mutants during the catalytic process

Breaking of hydrogen bonds often leads to structural changes, and structural changes affect enzyme activity. When the number of hydrogen bonds decreases, it means that the structure of the enzyme is changing, allowing the enzyme to play a role in the reaction, which is a manifestation of increased enzyme activity; maintaining a stable number of hydrogen bonds means that the structure of the enzyme has been in a stable state, relatively The enzyme activity is lower. Through comparison, we can clearly see that during the entire catalytic process, the number of hydrogen bonds in the Asp404 mutant changes more drastically, and its enzyme activity ranks first, followed by the Arg401 mutant, and Arg370 is less mutated. There was a certain improvement before the mutation but there was a gap with the Asp404 and Arg401 mutants, and its enzyme activity ranked third. The Lys499 mutant had a certain improvement compared with before the mutation but there was a gap with the other three mutants, and its enzyme activity ranked third. end. In summary, the enzymatic activity of each mutant is as follows: Asp404 mutant>Arg401 mutant>Arg370 mutant>Lys499 mutant.

After analysis of RMSD and Hydrogen Bonds molecular dynamics data, we can

draw the following conclusions: The enzyme activities of WP_020931051.1 and its four mutants in descending order are Asp404 mutant, Arg401 mutant, Arg370 mutant, and Lys499 mutant. , WP_020931051.1.

CONCLUSIONS

This topic focuses on the polylysine synthase(Pls) produced by the *S. albulus* NBRC 14147 strain, with the aim of improving the efficiency of Pls. active. Based on the existing research progress, a research plan was formulated: by mutating the amino acid residues of larger side chains near the reaction channel to reduce the resistance of the substrate when passing through the reaction channel, thereby increasing the binding rate of the substrate and the enzyme and increasing the activity of the enzyme. At the experimental level, various bioinformatics tools are used to obtain the required information, and computational biology methods are used to solve a series of research problems.

The entire experimental process can be summarized as: (1) Retrieve the amino acid sequence information of the enzyme studied in the experiment (2) Mutation of the sequence to obtain the amino acid sequence information of the mutant (3) Homology modeling of the ontology and the mutant and study its structure (4) Molecularly dock the ontology and mutants with substrates and study the molecular docking results. (5) Simulate the catalytic process through molecular dynamics, study molecular dynamics data, and verify the docking results.

After the above experimental process and analysis of the experimental results and experimental data obtained in each link, we can draw the following conclusion: In this experiment, the enzyme activity of the Asp404 mutant ranked first, and the enzyme activity of the Arg401 mutant ranked second, the enzyme activity of the Arg mutant ranks third, the enzyme activity of the Lys499 mutant ranks fourth, and the enzyme activity ranked last is WP_020931051.1.

This conclusion has a certain guiding role in the production practice of Pls. Pls has a wide range of uses, involving various fields such as food, medicine, materials, etc. The improvement of the activity of Pls can help reduce the use cost of Pls and improve production in the production links of related fields. Efficiency, thereby further improving productivity, has certain economic value.

This topic is only a preliminary study on the activity of Pls, and the experimental conclusions need to be further verified by more practical experiments.

LIST OF REFERENCES

1. Oppermann-Sanio, F. B., & Steinbüchel, A. (2002). Occurrence, functions and biosynthesis of polyamides in microorganisms and biotechnological production. *Naturwissenschaften, 89*(1), 11-22.

2. Jia, S., Liu, Y., Zhuang, S., et al. (2019). Effect of ε-polylysine and ice storage on microbiota composition and quality of Pacific white shrimp (Litopenaeus vannamei) stored at 0°C. *Food Microbiology, 83*(OCT.), 27-35.

3. Yamanaka, K., Maruyama, C., Takagi, H., et al. (2009). Poly-L-lysine dispersity is controlled by a highly unusual nonribosomal peptide synthetase. *Nature Chemical Biology, 4*(12), 766-772.

4. Liu, Q. (2014). Isolation, purification and enzymatic properties study of Streptomyces sp.M-Z18 ε-polylysine degrading enzyme [D]. *Jiangnan University*.

5. Yamanaka, K., Maruyama, C., Takagi, H., et al. (2008). Epsilon-poly-L-lysine dispersity is controlled by a highly unusual nonribosomal peptide synthetase. *Nat Chem Biol, 4*(12), 766-772.

6. Yamanaka, K., Kito, N., Imokawa, Y., et al. (2010). Mechanism of epsilon-poly-L-lysine production and accumulation revealed by identification and analysis of an epsilon-poly-L-lysine-degrading enzyme. *Applied and Environmental Microbiology, 76*(17), 5669-5675.

7. Shima, S., & Sakai, H. (1977). Polylysine produced by Streptomyces. *Agricultural and Biological Chemistry, 41*(9), 1807-1809.

8. Nishikawa, M., & Ogawa, K. (2002). Distribution of microbes producing antimicrobial ε-poly-L-lysine polymers in soil microflora determined by a novel method. *Applied and Environmental Microbiology, 68*(7), 3575-3581.

9. El-Sersy, N. A., Abdelwahab, A. E., Abouelkhiir, S. S., et al. (2012). Antibacterial and anticancer activity of ε-poly-L-lysine (ε-PL) produced by a marine Bacillus subtilis sp. *Journal of Basic Microbiology, 52*(5), 513-522.

10. Hiraki, J., Ichikawa, T., Ninomiya, S., Seki, H., Uohama, K., Seki, H., Kimura, S., Yanagimoto, Y., & Barnett, J. W. (2003). Use of ADME studies to confirm the safety of ε-polylysine as a preservative in food. *Regul. Toxicol. Pharmacol., 37*, 328–340.

11. Chang, S., Lu, W. W., Park, S., & Kang, D. (2010). Control of foodborne pathogens on ready-to-eat roast beef slurry by ε-polylysine. *Int. J. Food Microbiol., 141*(3), 236–241.

12. Li, Y., Han, Q., Feng, J., Tian, W., & Mo, H. (2014). Antibacterial characteristics and mechanisms of ε-poly-lysine against Escherichia coli and Staphylococcus aureus. *Food Control, 43*, 22–27.

13. Gao, C., Yan, T., Du, J., He, F., Luo, H., & Wan, Y. (2014). Introduction of broad spectrum antibacterial properties to bacterial cellulose nanofibers via immobilizing ε-polylysine nanocoatings. *Food Hydrocoll., 36*, 204–211.

14. Hiraki, J., Ichikawa, T., Ninomiya, S. I., et al. (2003). Use of ADME studies to confirm the safety of ε-polylysine as a preservative in food. *Regulatory Toxicology and Pharmacology, 37*(2), 328-340.

15. Shih, I. L., Shen, M. H., & Van, Y. T. (2006). Microbial synthesis of poly(ε-lysine) and its various applications. *Bioresource Technology, 97*(9), 1148-1159.

16. Kunioka, M., & Choi, H. J. (1995). Properties of biodegradable hydrogels prepared by γ irradiation of microbial poly(ε-lysine) aqueous solutions. *Journal of Applied Polymer Science, 58*(4), 801-806.

17. Nishikawa, M., & Ogawa, K. (2002). Inhibition of Epsilon-Poly-L-Lysine Biosynthesis in Streptomycetaceae Bacteria by Short-Chain Polyols. *Applied & Environmental Microbiology, 72*(4), 2306-2312.

18. Tsujita, T., Takaichi, H., Takaku, T., Aoyama, S., & Hiraki, J. (2006). Antiobesity Action of Ɛ-Polylysine, A Potent Inhibitor of Pancreatic Lipase. *J. Lipid Res., 47*(8), 1852–1858.

19. Tsujita, T., & Takaku, T. (2009). Inhibition by ε-Polylysine of Fat Digestion in the Stomach and Intestine of Rats. *Biosci. Biotechnol. Biochem., 73*(3), 536–542.

20. Chen, S., Huang, S., Li, Y., & Zhou, C. (2021). Recent Advances in Epsilon-PolyL-Lysine and L-Lysine-Based Dendrimer Synthesis, Modification, and Biomedical Applications. *Front. Chem., 9*, 659304.

21. Yang, B., Lu, P., Liao, R., et al. (2018). ε-polylysine-polyethylenimine-βCyclodextrin polymer and its preparation method and application: China, CN201610307704.2 [P]. December 14.

22. Shen, Y., Wu, B., Xiang, J., et al. (2019). Polyethylene glycol-polylysine-isothiocyanate conjugate and its application as a drug carrier: China, CN201810045332.XP1. July 23.

23. Okamoto, T., Yamanaka, K., Hamano, Y., et al. (2022). Crystal structure of the adenylation domain from an ε-poly-l-lysine synthetase provides molecular mechanism for substrate specificity. *Biochemical and Biophysical Research Communications, 596*, 43-48.

24. Liu, Y., Yang, X., Gan, J., Chen, S., Xiao, Z.-X., & Cao, Y. (2022). CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Research, 50*(W1), W159–W164.

25. Paggi, J. M., Belk, J. A., Hollingsworth, S. A., Villanueva, N., Powers, A. S., Clark, M. J., Chemparathy, A. G., Tynan, J. E., Lau, T. K., Sunahara, R. K., et al. (2021). Leveraging nonstructural data to predict structures and affinities of protein–ligand complexes. *Proc. Natl. Acad. Sci. U.S.A., 118*, e2112621118.

26. Zhang, W., Bell, E. W., Yin, M., & Zhang, Y. (2020). EDock: blind protein–ligand docking by replica-exchange monte carlo simulation. *J. Cheminform., 12*, 37.

27. Yang, X., Liu, Y., Gan, J., Xiao, Z.-X., & Cao, Y. (2022). FitDock: protein–ligand docking by template fitting. *Brief. Bioinform.* https://doi.org/10.1093/bib/bbac087

28. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., & Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics, 29*, 845–854.

29. Páll, S., Abraham, M. J., Kutzner, C., Hess, B., & Lindahl, E. (2015). Tackling exascale software challenges in molecular dynamics simulations with GROMACS. In S. Markidis & E. Laure (Eds.), *EASC 2014: Solving Software Challenges for Exascale* (pp. 3–27). Lecture Notes in Computer Science, 8759. Springer, Cham.
30. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1–2*, 19–25.
31. Páll, S., Zhmurov, A., Bauer, P., Abraham, M., Lundborg, M., Gray, A., Hess, B., & Lindahl, E. (2020). Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *The Journal of Chemical Physics, 153*(13), 134110.