STUDY OF THE INTERACTION OF COPPER(II) IONS WITH SILK FIBROIN USING THEORETICAL CALCULATIONS

Khushnudbek Eshchanov
PhD., Department of Chemistry, Faculty of Natural and Agricultural sciences, Urgench State University, Kh.Alimjan 14 h., Urgench, 220100, Uzbekistan
*e-mail: xeshchanov77@gmail.com
Gulinur Polvonova
Teacher, 2nd school with classes where some subjects are studied in depth, Ravnak 126 h., 220200, Bagat, Uzbekistan.
Amir Shigabutdinov
Doctoral student, Department of Chemistry, Faculty of Natural and Agricultural sciences, Urgench State University, Kh. Alimjan 14 h., Urgench, 220100, Uzbekistan
e-mail: amir.shigabutdinov.96@mail.ru
Mukhabbat Baltaeva
DSc., Department of Chemistry, Faculty of Natural and Agricultural sciences, Urgench State University, Kh.Alimjan 14 h., Urgench, 220100, Uzbekistan.
e-mail: bmuhabbat@rambler.ru
Risolat Esomurodova
Teacher, Department of Algebra and mathematical engineering, Faculty of Physics and mathematics, Urgench State University, Kh. Alimjan 14 h., Urgench, 220100, Uzbekistan
e-mail: risolat@urdu.uz

ABSTRACT

In addition to being the primary raw material for the silk textile industry, materials based on silk fibroin and its compounds with various metals are also of significant importance. Studies have been conducted on the sorption of certain metal ions by silk fibroin and the formation of complex compounds through interactions with these ions. However, many scientific studies have not fully explored which amino acid sequences in silk fibroin bind with copper(II) ions. Using quantum chemical calculations, molecular docking methods, and other theoretical calculation methods, it was studied which amino acid sequences in silk fibroin are bound to by copper(II) ions. Through these calculations, an attempt was made to explain the structure of copper-fibroin. It was observed that the physical parameters related to the composition of the copper-fibroin complex determined based on calculations were consistent with the experimental results.

Keywords: *Copper(II) ion, IR spectrum, quantum chemical calculation, silk fibroin, UV-Vis spectrum.*

INTRODUCTION

The characterization of materials using quantum chemical calculations began mainly in the 1970s. However, with the improvement of computational methods and the development of computer technology, developments in this field continue rapidly [1, 2].

Silk consists mainly of fibroin, a natural strong, flexible, and biodegradable fibrous protein. The unique structure and properties of fibroin make it a valuable material for various applications. In the medical field, fibroin is used in areas such as tissue engineering and drug delivery systems. These properties of silk fibroin (SF) protein encourage theoretical and practical research on it. The main component of SF is fibroin protein, small amounts of lipids and polysaccharides. SF has a modular hydrophobic structure of high molecular weight, which is separated by small hydrophilic groups. SF consists of two main chains: a hydrophobic heavy (H-) chain and a hydrophilic light (L-) chain. These two chains are linked by a disulfide bond (HL). The P25 molecule is a hydrophobic glycoprotein that is bound to HL by intermolecular forces and hydrogen bonds and plays an important role in maintaining structural integrity.

The most widely produced silk fiber is obtained from the silkworm *Bombyx mori*. SF in silk from *Bombyx mori* consists of a heavy chain (H-chain, 360-390 kDa) and a light chain (L-chain, 27 kDa) linked by disulfide bonds, and they bind to a glycoprotein called P25. The bound state is cross-linked by non-covalent interactions in a molar ratio of 6:6:1, respectively [3, 4].

The H-chain is composed mainly of glycine (43-46%, G), alanine (30%, A), serine (12%, S), tyrosine (5.3% Y) and other amino acids residues. The H-chain of SF can be viewed as a repeating natural block copolymer consisting of 12 domains forming the crystal parts separated by 11 less organized domains consisting of a non-repeating primary sequence. This block copolymer structure of the H-chain accounts for the characteristic mechanical properties of SF [5].

The H-chain has a unique repeating sequence. This structure includes a highly repetitive sequence of Gly-Ala-Gly-Ala-Gly-Ser (GAGAGS), a relatively less repetitive sequence of hydrophobic or aromatic residues, and amorphous regions that include negatively charges, polar, bulky hydrophobic or aromatic residues [6].

NMR is one of the most suitable methods for studying the structure of SF, as it can determine the molecular structure at the atomic level of various sample forms, from solutions to solids. The structural properties of *Bombyx mori* fibroin can be

conveniently studied by using the $(AG)_n$ peptide as a model for the crystal region. The solution structure of liquid silk, especially the repeating sequences of $(GAGXGA)_n$ (X=Ser, Tyr or Val) and GAASGA, has been determined by NMR [7].

The SF structure was first proposed to have an antiparallel β-sheet structure by X-ray diffraction studies. Later, Takahashi et al. reported some internal structural distortions in this antiparallel β-sheet model of fibroin. In addition, the Ala-specific peak in the solid-state NMR spectrum of *Bombyx mori* silk fiber was found to be broad and asymmetric, reflecting the heterogeneous structure of the silk fiber. It was shown that conformation-dependent NMR chemical shifts can be used to easily and selectively determine the proportion of mixed structures. In addition, the atomic-level structure can be obtained by NMR with stable isotope targets of peptides. The sequential peptide model of the crystal region $(AGSGAG)_5$ was confirmed by solid-state NMR [8, 9].

The L-chain and P25 participate in the formation of the amorphous part of SF. They consist of α-helix and random structures. The distance between the same type of atoms (helix layer) in the α-helix chain is 1.5 Å. The angle between the helices is 26°. One turn of the helix consists of 3.6 amino acid residues. If these residues are transferred from the helical state in the linear form, this distance will be 5.4 Å. The orderly arrangement (crystallization) of the fibroin polypeptide chains increases with a decrease in the moisture content. In addition, thermal heating of the fiber in an inert environment also leads to an increase in the degree of crystallization [10].

O. Kratky and S. Kuriyama, based on X-ray diffraction analysis of crystal fragments of *Bombyx mori* SF, showed that the unit cell parameters are a=9.29 Å, b=9.44 Å, c=6.95 Å and that it has an orthorhombic form [11]. Y. Takahashi reported in his scientific publications that the unit cell parameters of SF are a=9.38 Å, b=9.49 Å, c=6.98 Å and that it is monoclinic [12]. The unit cell parameters of SF obtained from silk fiber waste are a=10.88 Å, b=7.62 Å, c=7.79 Å, α=90°, β=90°, γ=105° and it is monoclinic [13].

Quantum chemical calculations of silk proteins involve the use of computational methods to simulate the electronic structure and properties of the protein. These calculations can provide insight into the chemical, physical, and mechanical properties of the protein, as well as its interactions with water or other molecules. The most commonly used quantum chemical methods for studying proteins are based on density functional theory (DFT) and molecular mechanics (MM). DFT calculations provide more precise information about the electronic structure and energies of the protein, while MM methods are used to describe the molecular mechanics of the

protein, including bond lengths, angles, and torsion angles. In addition, semi-empirical methods are also used [14, 15].

Before studying SF through theoretical calculations, it is necessary to have a good understanding of its molecular structure. There are many databases that contain information on the structures and other properties of proteins. The most widely used and reliable database of this kind is UniProt. UniProt is an open database for studying protein sequences.

UniProt is a freely available database of amino acid residue sequences and functional groups in proteins, with much of the information coming from genome projects. It also contains a large amount of information on the biological functions of proteins from the research literature [16].

Using these databases, the molecular structure of SF can be studied to a sufficient extent. If we pay attention to the information about fibroin in the UniProt and Swiss databases, we can see that the heavy chain (H) of SF consists of 5263 amino acid residues. The molecular mass of the H-chain is 391593 Da [17].

The crystal part of the H-chain consists of β-structure chains. β-structure chains play a very important role in the formation of crystal parts of proteins. If the β-structure chains are disrupted, the crystal parts decrease. It is of great importance to study the physical and chemical properties of silk using theoretical calculations.

Calculations have been carried out to determine the interactions between graphite and SF using molecular dynamics (MD) simulations. It has been considered that SF-layered nanostructures have strong orientational and spatial compatibility with graphite, which leads to hydrophobic effects, destabilizing solvent-protein interactions, and stabilizing protein-protein and protein-graphite interactions [18].

Curcumin has been studied by researchers to show its very promising medicinal properties when loaded within an SF matrix. Molecular-level characterization using molecular dynamics simulations showed that the drug entrapment efficiency was in perfect agreement with the interaction energy models derived from MM/PBSA calculations [19].

Fibroin-chitosan composites, especially those containing nanohydroxyapatite, show potential for bone regeneration. The physicochemical properties of these biocomposites depend on the compatibility between their components. Studies have been conducted to analyze the intermolecular interactions of fibroin and chitosan using a molecular dynamics approach. It has been shown that β-sheet fragments in fibroin participate in the formation of numerous hydrogen bonds and ionic bonds with chitosan [20].

Since copper(II) ions have antibacterial properties, the interaction of SF with copper(II) ions has prompted research into the production of new materials with antibacterial properties. Such materials could be used in medicine for wound dressings, hygiene products, and antibacterial agents [21].

Until now, the coordination and structure of copper(II) ions in aqueous solutions have been a highly controversial topic. Many theoretical and experimental studies have been carried out on this issue. For example, in the studies conducted by Katia Julia de Almeida and her team, they tried to prove the structure of water-bound copper(II) ions in the six- and five-coordinate states through quantum chemical calculations (density functional theory) [22].

Studies have also shown that copper(II) ions in copper(II) sulfate solution and crystals are coordinated by six water molecules using other physical methods. Aqueous solutions of copper(II) sulfate have been found to have a distorted octahedral shape with four water molecules located closer to the cation coordination sphere and two water molecules located further away [23].

An examination of the total intensity of X-ray scattering on copper(II) chloride showed that the copper(II) ion in a copper(II) chloride solution with a concentration of 3.18 mol/L is covered by 2.7 oxygen atoms and 3.3 chloride ions at a distance of 0.193±0.003 nm. A 4.35 mol/L aqueous solution of copper(II) chloride has 2.4 oxygen atoms and 3.6 chloride ions at the same distance. An X-ray structural analysis of a 2.95 mol/L aqueous solution of copper(II) chloride showed that the four equatorial sites in the disordered complex of the copper(II) ion are occupied by an average of 2.8 water molecules and 1.2 chloride ions at a distance of 0.195 nm. In dilute solutions, copper(II) ions are bound to five or six water molecules [24, 25].

X-ray diffraction analyses of copper(II) nitrate solutions showed that the copper(II) ions were in five or six coordinated states with water molecules. EXAFS and XANES studies of aqueous solutions of copper(II) perchlorate at various concentrations (from 0.3 mol/L to saturated solutions) showed that the copper hexaaqua complex was stable [26-28].

Most copper(II) complexes are highly catalytically active compounds. They can be used in chemical syntheses and various other chemical processes. Copper(II) ions interact with SF to form copper-fibroin composite materials. Copper(II) ions can be formed by absorbing or adding copper(II) ions to SF solutions. Copper(II) ions confer bioactivity and antibacterial properties to SF, while SF is a biocompatible and biodegradable product that supports cell growth and tissue regeneration. SF contains groups that react with copper(II) ions, but the binding and reaction mechanisms of these groups have not been fully studied.

Bonds can be formed as a result of ionic interactions between negatively charged functional groups in SF and positively charged metal ions. The SF protein has negatively charged polyfunctional groups, including -COOH, -NH_2, -OH, -CO-NH-, -CO- and -S-S-. These negatively charged functional groups provide binding sites for metal ions. In addition, these polar functional groups contain atoms with lone pairs, including O and N. These atoms tend to chelate with electron-deficient metal ions, forming coordination bonds [29, 30].

Xiao-Hong, Zong Ping Zhou et al. studied the effects of Cu(II) ions and pH on the conformational changes of SF recovered from *Bombyx mori* fibers using EPR and ^{13}C-NMR methods. Raman spectroscopy analysis results showed that when the solution pH was lowered from 8.0 to 4.0, the Cu(II) coordination atoms in SF changed from four nitrogens to two nitrogens and two oxygens, and also to one nitrogen and three oxygens. It was suggested that the imidazole Np in the basic and neutral AHGGYSGY polypeptide in SF could form a 1:1 complex with Cu(II) ions. It was shown that its residues were bound to two deprotonated main chain nitrogens from two glycines and one nitrogen or oxygen from serine [31].

W. X. Chen et al. proposed a new alternative method for ammonia removal from gas-phase systems by coordination exchange reaction with metal-SF complex fibers and attempted to characterize the complex composition. Electron spin resonance spectra (ESR) and atomic absorption spectrometry in this work showed that when the copper(II) content was low and under acidic conditions, the $Cu(O)_4$ coordination complex was formed on the SF fiber. When the copper(II) content was high, the cuprammonium complex was formed. The experimental results showed that the $Cu(O)_4$ complex was more effective than the $Cu(N)_4$ complex in ammonia adsorption [32].

Increasing the amount of copper(II) chloride in SF weakens silk I crystallization and Cu(II) ion chelation, which makes SF scaffolds unstable in water. Copper chloride dihydrate above 5.0 % has been shown to inhibit silk I crystallization [33, 34]. Based on the above information, it has not been systematically studied how to coordinate copper(II) ions with which amino acids and which parts of the SF chain. Therefore, it is appropriate to determine the optimal structure of the copper-fibroin complex through theoretical calculations.

THEORETICAL PART

Computational methods for protein structure determination

The most widely used force fields for studying protein structures are the CHARMM force fields. These force fields include CHARMM19, CHARMM22, and

its two-dimensional potential-corrected variant CHARMM22/CMAP, as well as later versions such as CHARMM27, CHARMM36, and various modifications such as CHARMM36m and CHARMM36IDPSFF [35]. The CHARMM22 force field takes into account the partial charges of the atoms of the protein, the interactions between the model compounds, and water, in quantum chemical calculations. In addition, CHARMM22 is fully parameterized for the TIP3P exact water model [36]. The CHARMM22 force field has the following potential energy function [37]:

$$V = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi [1 + \cos(n\phi - \delta)] + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u (u - u_0)^2 + \\ + \sum_{nonbonded} \left(\varepsilon_{ij} \left[\left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon_r r_{ij}} \right) \quad (1)$$

The CHARMM program allows for the creation and analysis of a wide range of molecular simulations. The most basic types of simulations are the minimization of the production processes of a given structure and molecular dynamics trajectory. Advanced features include free energy perturbation (FEP), quasi-harmonic entropy estimation, correlation analysis, and combined quantum and quantum mechanics-molecular mechanics (KM/MM) methods [38].

There are also many programs that predict the structure of proteins based on computational methods such as the above. One of the most reliable of these is the Robetta program. This program includes relatively fast and accurate deep learning-based methods, RoseTTAFold and TrRosetta, and an interactive submission interface that allows for customizing procedures for homology modelling, constraints, etc. It can model multi-chain complexes using RoseTTAFold or comparative modelling (CM) and provides a large sampling capacity. The CM method uses the PDB100 template database, the Evolutionary Model Database (EMD) and also provides an option for custom templates [39]. The best options with the lowest error rate can be selected and downloaded.

Based on the above, the CHARMM22 computational method and the Robetta program were used to study the structure of the SF molecule chains.

Determination of the ability of copper(II) ions to bind to SF using molecular docking

In the field of molecular modelling, docking techniques can be used to determine the possibility of a ligand and a receptor forming a stable complex and the preferred orientation of one molecule towards the other when bound. The preferred orientation can in turn be used to predict the binding or binding affinity between two molecules using scoring functions. Many scoring functions are based on molecular

mechanics force fields, which estimate the pose energy at the binding site. The various contributions of binding can be written in the following equation:

$$\Delta G = \Delta G_{sol} + \Delta G_{konf} + \Delta G_{\text{int}} + \Delta G_{rot} + \Delta G_{t/t} + \Delta G_{vib} \quad (2)$$

There are many programs available for performing molecular docking calculations. Among them, the online platform SeamDock combines various docking tools into a common system. The platform devotes considerable effort to 3D visualization of ligands, receptors, and docking poses and their interactions with receptors [40]. Taking into account the above capabilities and methods, the SeamDock platform was used to determine which parts of SF molecules and which amino acid residues strongly bind copper(II) aqua complex ions. ProteinSol and CamSol programs were also used to calculate some properties of the fibroin macromolecule [41, 42].

Determination of the coordination binding possibilities of copper(II) ions with SF molecules using quantum chemical calculations

Using the SeamDock platform, it was determined which amino acid residues of SF molecules formed strong bonds with copper(II) aqua complex ions and the possibilities of forming coordination bonds were determined using the CHARMM22 and PM3 calculation methods.

To do this, models of variants that can form coordination bonds were created and their full energies were determined. The model variant with the lowest full energy was determined and the composition of the resulting complex was theoretically determined. Based on the models of the identified variants, their total IR spectrum was calculated. In addition, UV spectra were also calculated. The IR and UV spectra obtained based on theoretical calculations were compared with the ATR-IR and UV-Vis spectra of the copper-fibroin complex obtained in practice.

EXPERIMENTAL

Obtaining the copper-fibroin complex

0.5 g of SF was added to 100 ml of 0.1 M $CuCl_2$ solution and mixed thoroughly for 10 minute. Then, the mixing was stopped, the mouth of the container was closed, and the mixture was left for 7 days. After 7 days, the color of the SF sample turned yellow. The sample was filtered and washed with hot distilled water until no excess copper(II) ions remained in the sample. The yellow sample was dried, and a portion of it was dissolved in a 1:2:8 molar mixture of $CaCl_2$, ethanol, and water. The UV spectrum was obtained on a UV-1800 Shimadzu spectrophotometer, and the ATR-

FTIR spectrum of the remaining portion was studied using a FT/IR-JASCO 4600 spectrometer.

RESULTS AND DISCUSSION

The UniProt database provides accurate scientific information on the structure and chemical composition of the *Bombyx mori* SF H-, L-chains and P25 macromolecule. The amino acid residues with the largest share in the H-chain are Ser (12.1 %), Ala (30.3 %) and Gly (45.9 %). Determining which parts of the SF macromolecule form primary bonds of copper(II) ions in an aqueous solution is a problematic issue.

Initially, the binding possibilities of copper(II) ions to the L-chain and the P25 molecule were determined on the SeamDock platform. For this, a model of the copper(II) aqua complex ($[Cu(H_2O)_6]^{2+}$) was created, and the model was saved in sdf and pdb formats. The SeamDock platform accepts sdf, pdb and mol2 formats for the ligand file.

According to the results of the molecular docking calculation of the L-chain with the $[Cu(H_2O)_6]^{2+}$ ion in the Smina program on the SeamDock platform, a binding corresponding to the smallest value of the Gibbs energy (ΔG= -4.6 kcal/mol) was found (Figure 1).

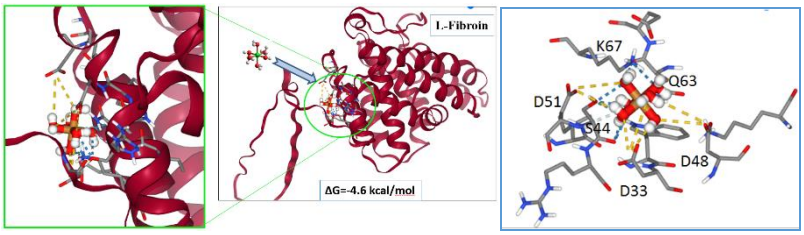

Fig.1. Binding of $[Cu(H_2O)_6]^{2+}$ ion to the L-chain of SF.

It was found that the $[Cu(H_2O)_6]^{2+}$ ion was strongly bound to the amino acid residues Asp33, Asp51, Asp48, Ser44, Lys67 and Gln63 in the L-chain. Table 1 below shows which functional groups of the amino acid residues it binds to.

Table 1. Types of $[Cu(H_2O)_6]^{2+}$ ion binding to amino acid residues in the L-chain

Parts that form a bond	Chemical bond type
D33-COO^-..... $[Cu(H_2O)_6]^{2+}$	Ionic bonding
D51-COO^-..... $[Cu(H_2O)_6]^{2+}$	Ionic bonding
D48-COO^-..... $[Cu(H_2O)_6]^{2+}$	Ionic bonding
S44-OH..... $[[Cu(H_2O)_6]^{2+}$	hydrogen bond
S44-C=O..... $[Cu(H_2O)_6]^{2+}$	hydrogen bond

K67-NH_2..... $[Cu(H_2O)_6]^{2+}$	hydrogen bond
Q63-$CONH_2$..... $[Cu(H_2O)_6]^{2+}$	hydrogen bond

The placement of the $[Cu(H_2O)_6]^{2+}$ ion in the part of the L-chain where the above-mentioned amino acids are located can be explained by another reason. The negative potential field of the part where these amino acids are located is considered to have a relatively strong effect. This is also shown by the results obtained by calculating using the ProteinSol program. The probability of $[Cu(H_2O)_6]^{2+}$ ion going to the part with the high negative potential field is very high (Figure 2).

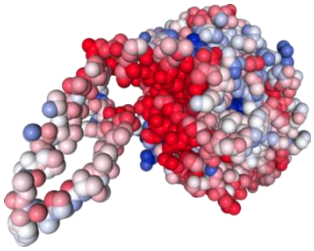

Fig.2. The potential field of the L-chain: Negative potential field in red, positive potential field in blue.

Therefore, it is observed that $[Cu(H_2O)_6]^{2+}$ ions are located in the part of the L-chain where the above-mentioned amino acids are located.

Considering the determined binding states, the coordination binding of the copper(II) ion with the functional groups of the above amino acid residues and the structure was determined using the PM3, UHF and CHARMM22 calculation methods. Calculations showed that among the considered model variants, the coordination binding of the copper(II) ion with the C=O and –OH groups of Ser44, with the $-COO^-$ group of Asp33 and Asp 48, with the $-(CH_2)_2NH_2$ group of Gln63, and with the $-(CH_2)_4NH_2$ group of Lys67 has the smallest value of the total energy determined by the CHARMM22 method, i.e. -943.865 kcal/mol. The total energy value determined by PM3 was equal to -199334 kcal/mol (Figure 3).

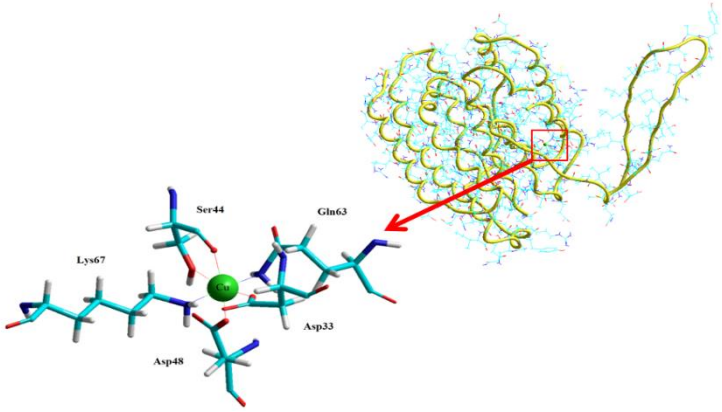

Fig.3. Coordinative binding of copper(II) ion to amino acid residues in the L-chain of SF.

In the Smina program on the SeamDock platform, calculations were performed using the molecular docking method of the P25 molecule of SF with the $[Cu(H_2O)_6]^{2+}$ ion. As a result, it was observed that the $[Cu(H_2O)_6]^{2+}$ ion forms bonds with the His220, Tyr56, Trp203, Ser207 and Arg211 amino acid residues of the P25 molecule. In this case, the change in Gibbs energy was -4.2 kcal/mol. This indicator was found to be the smallest value among all poses (Figure 4).

The localization of the $[Cu(H_2O)_6]^{2+}$ ion in the part of the P25 molecule where the above amino acids are located can be explained by the potential field effect. The negative potential field of the part where the His220, Tyr56, Trp203, Ser207 and Arg211 amino acid residues are located is considered to have a relatively strong effect. The results of the calculations also confirm this data. The model of the P25 molecule, generated based on the results obtained through calculations in the ProteinSol program, is illustrated below with a potential field.

The parts shown in red represent the negative potential field. Therefore, the $[Cu(H_2O)_6]^{2+}$ ion has a high probability of binding to this part of the P25 molecule (Figure 5, Table 2).

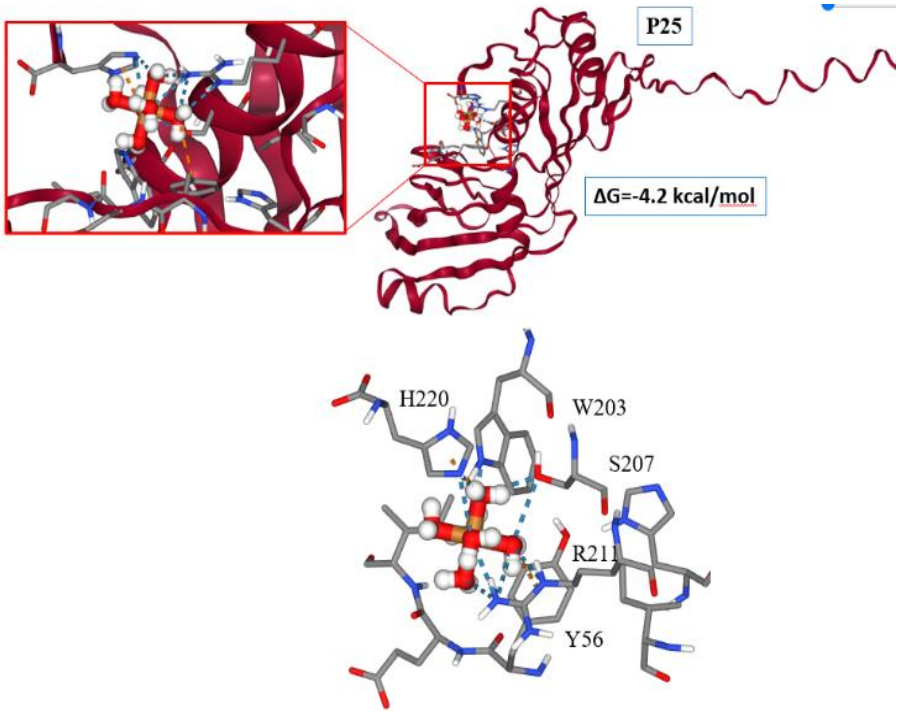

Fig.4. General and specific forms of binding of the $[Cu(H_2O)_6]^{2+}$ ion to the P25 molecule of SF.

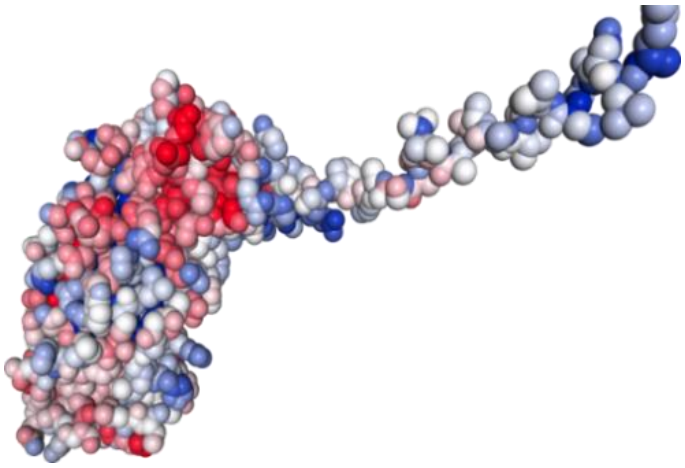

Fig.5. The potential field of the P25 molecular.

Table 2. Types of $[Cu(H_2O)_6]^{2+}$ ion binding to amino acid residues in the P25 molecule

Parts that form a bond	Chemical bond type
H220-C_2NH_3NH….....….$[Cu(H_2O)_6]^{2+}$	Ionic bonding
Y56-CH_2-C_6H_4-OH……..…...$[Cu(H_2O)_6]^{2+}$	Ionic bonding
H220-C_2NH_3NH…...…….$[Cu(H_2O)_6]^{2+}$	hydrogen bond
W203-$CH_2C_8H_5NH$.....$[Cu(H_2O)_6]^{2+}$	hydrogen bond
S207-OH.....$[Cu(H_2O)_6]^{2+}$	hydrogen bond
R211-$(CH_2)_3NHCNH_2NH_2$.....$[Cu(H_2O)_6]^{2+}$	hydrogen bond

Based on the above result, the possibilities of forming coordination bonds of the copper(II) ion to the functional groups of amino acid residues in the P25 molecule were determined using the PM3 and CHARMM22 calculation methods.

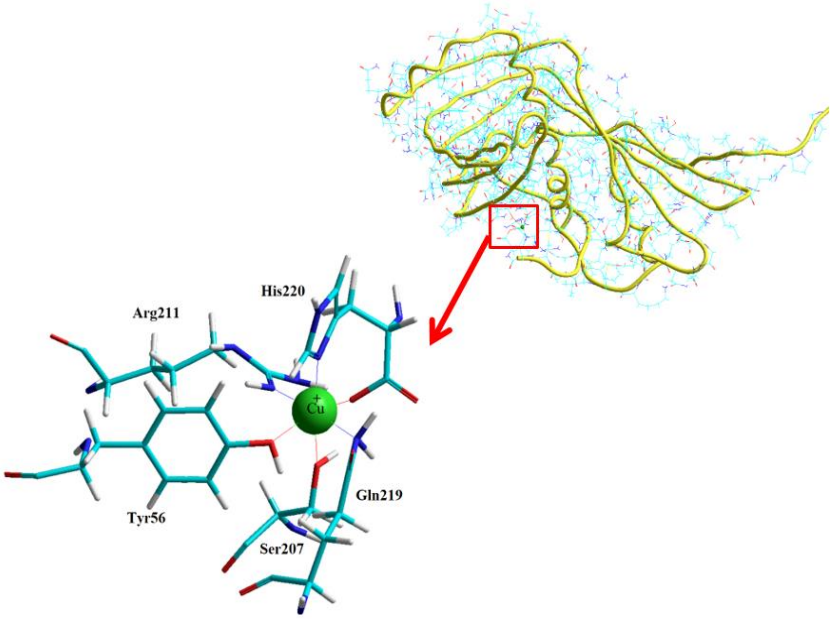

Fig.6. Coordinative binding of copper(II) ion to amino acid residues in the P25 molecule of SF.

It was found that the coordination bond state of the copper(II) ion with the N and -COO^- of the side group of Y220, the -$(CH_2)_3NHCNHNH_2$ group of Arg211, the -CH_2-C_6H_4-OH group of Tyr56, the -$(CH_2)_2NH_2$ group of Gln219 and the -OH group of Ser207 is the most favorable in terms of energy. In this case, the smallest value of the total energy was -1361.83 kcal/mol according to the CHARMM22 method and -222494 kcal/mol according to PM3.

The H-chain consists of 5263 amino acid residues, and there are 11 chains that separate the crystal parts from each other. The N-terminal domain and C-terminal domains in the chain also have random and α-structure states, which are also part of the amorphous parts.

N-terminal domain

MRVKTFVILCCALQYVAYTNANINDFDEDYFGSDVTVQSSNTTDEIIRD ASGAVIEEQITTKKMQRKNKNHGILGKNEKMIKTFVITTDSDGNESIVEEDVL MKTLSDGTVAQSYVAADAGAYSQSGPYVSNSGYSTHQGYTSDFSTSAAV

Chains between crystal parts

A1) TGSSGFGPYVANGGYSRSDGYEYAWSSDFGT
A2) TGSSGFGPYVAHGGYSGYEYAWSSESDFGT
A3) TGSSGFGPYVANGGYSGYEYAWSSESDFGT
A4) TGSSGFGPYVAHGGYSGYEYAWSSESDFGT
A5) TGSSGFGPYVAHGGYSGYEYAWSSESDFGT
A6) TGSSGFGPYVANGGYSGYEYAWSSESDFGT
A7) TGSSGFGPYVANGGYSGYEYAWSSESDFGT
A8) TGSSGFGPYVANGGYSGYEYAWSSESDFGT
A9) TGSSGFGPYVNGGYSGYEYAWSSESDFGT
A10) TGSSGFGPYVANGGYSGYEYAWSSESDFGT
A11) TGSSGFGPYVANGGYSRREGYEYAWSSKSDFET

C-terminal domain

RGYGQGAGSAASSVSSASSRSYDYSRRNVRKNCGIPRRQLVVKFRALPC VNC

Since the H-chain is very large, it requires studying the chain in parts. Initially, models of the chains forming the amorphous parts were created using the Robetta program. Modeling is based on the amino acid sequence. The structures of the N- and

C-terminal domain chains and 11 chains located between the crystalline parts created in the Robetta program are presented below (Figure 7).

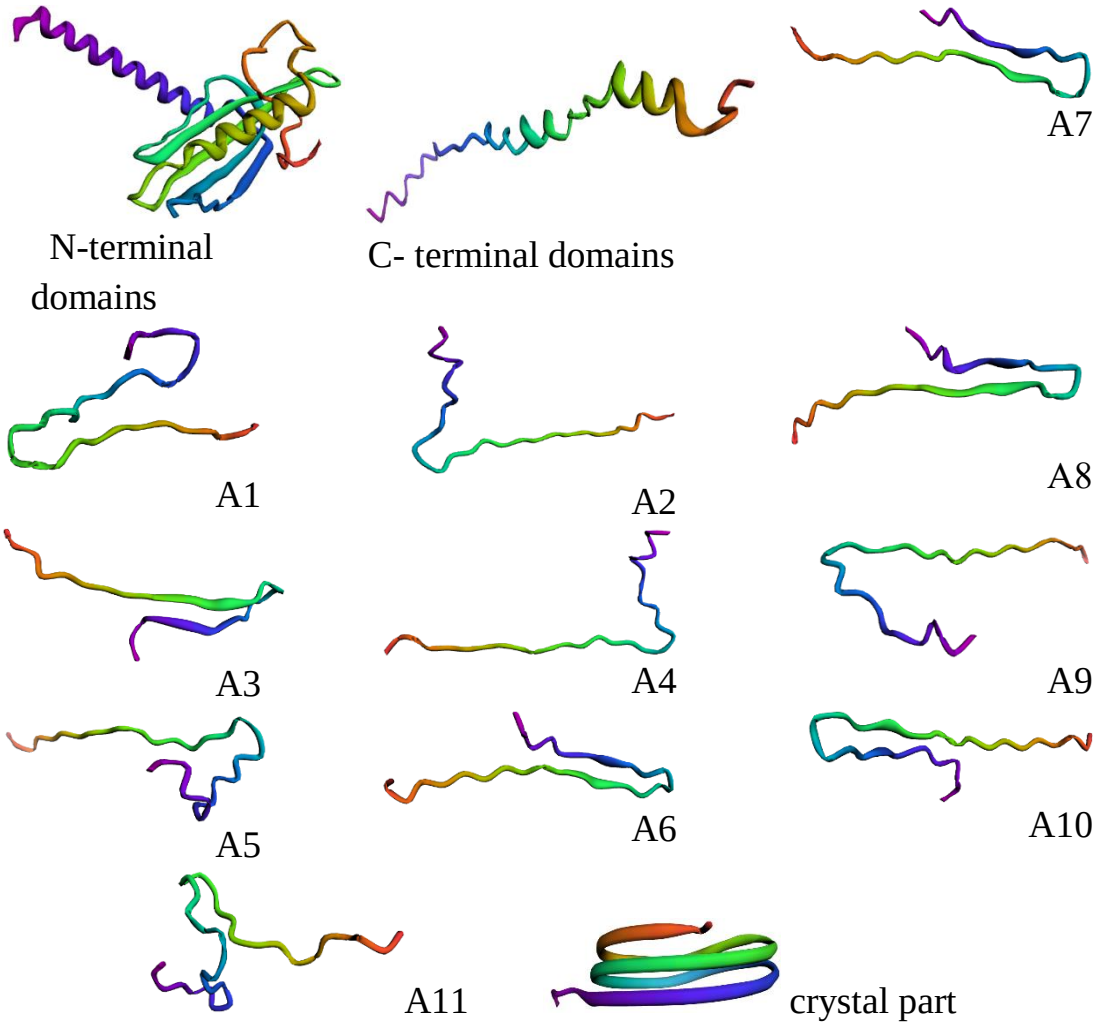

Fig.7. Structures of 11 chains located between the N- and C-terminal domains chains and crystal fragments

Using the Smina program on the SeamDock platform, the binding possibilities of the $[Cu(H_2O)_6]^{2+}$ ion to the chains created above for the structural models were studied and the change in the Gibbs energy value was determined. The values of the Gibbs energy representing the binding of the $[Cu(H_2O)_6]^{2+}$ ion to the chains forming the amorphous parts are presented in the graph below (Figure 8).

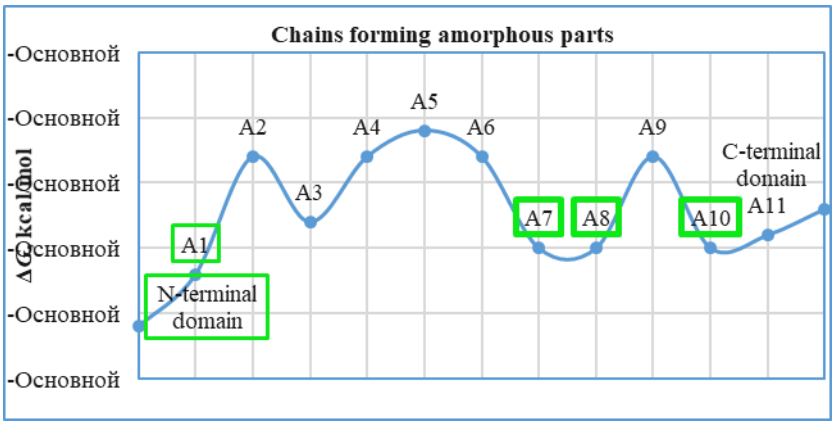

Fig. 8. Gibbs energy change values for the binding of $[Cu(H_2O)_6]^{2+}$ ion to chains forming amorphous parts in the H-chain.

It can be seen from the graph that the value of change of Gibbs energy for binding of $[Cu(H_2O)_6]^{2+}$ ion to N-terminal domain, A1, A7, A8 and A10 chains is the smallest. The Gibbs energy change value for the binding of $[Cu(H_2O)_6]^{2+}$ ions to the chains forming the crystal part is relatively large, -3.6 kcal/mol. Therefore, the probability of $[Cu(H_2O)_6]^{2+}$ ions being located in the N-terminal domain, A1, A7, A8 and A10 chains of the H-chain of SF is high. It can be concluded that the probability of binding of $[Cu(H_2O)_6]^{2+}$ ions to the chains forming the crystal part is very low (Figure 9).

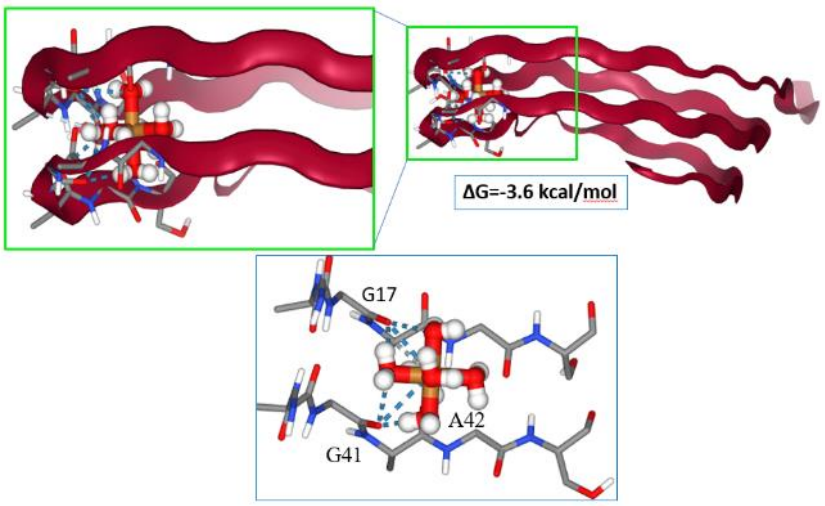

Fig.9. Binding of the $[Cu(H_2O)_6]^{2+}$ ion to the chain forming the crystal part

Figures 10 and 11 fully illustrate the binding of the $[Cu(H_2O)_6]^{2+}$ ion to amino acid residues in the N- terminal domain, A1, A7, A8, and A10 chains, and the Gibbs energy values.

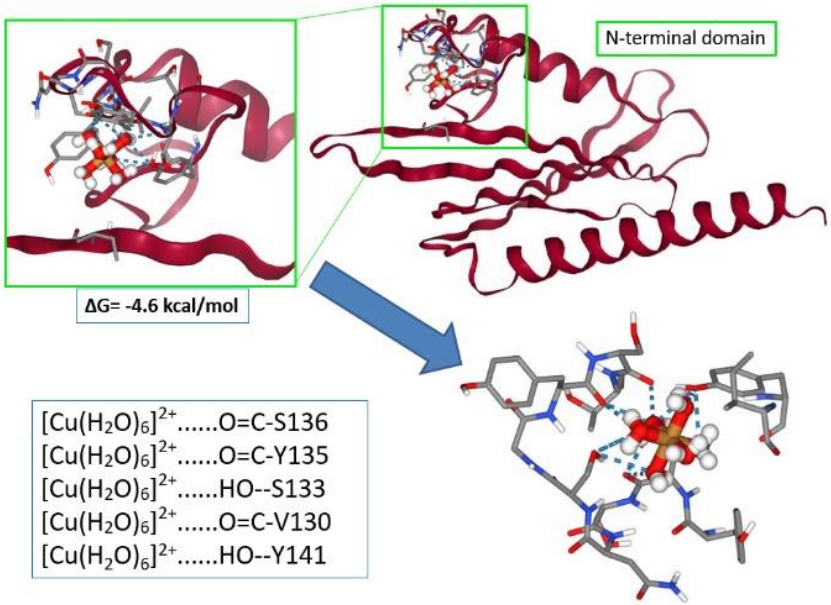

Fig.10. Binding of $[Cu(H_2O)_6]^{2+}$ ions to the N-terminal domain.

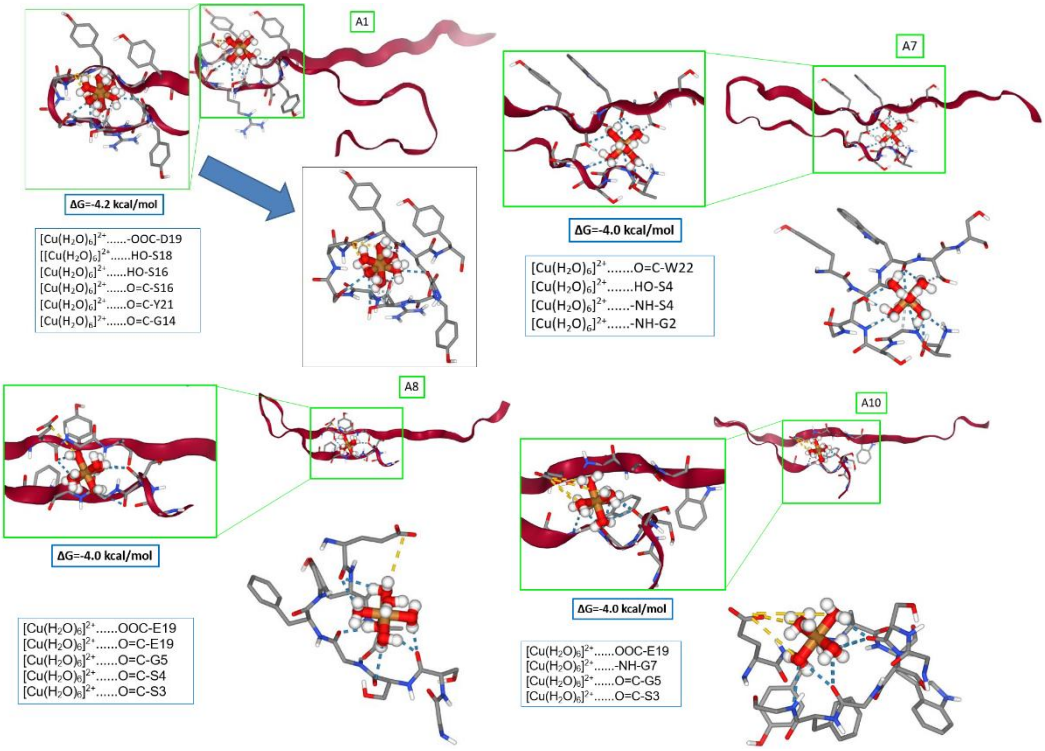

Fig.11. Binding of $[Cu(H_2O)_6]^{2+}$ ion to chains A1, A7, A8 and A10.

Taking into account the results of the molecular docking calculation method, the coordination binding of $[Cu(H_2O)_6]^{2+}$ ions to amino acid residues in the N-terminal domain, A1, A7, A8, and A10 chains of the H-chain was determined using the PM3 and CHARMM22 calculation methods.

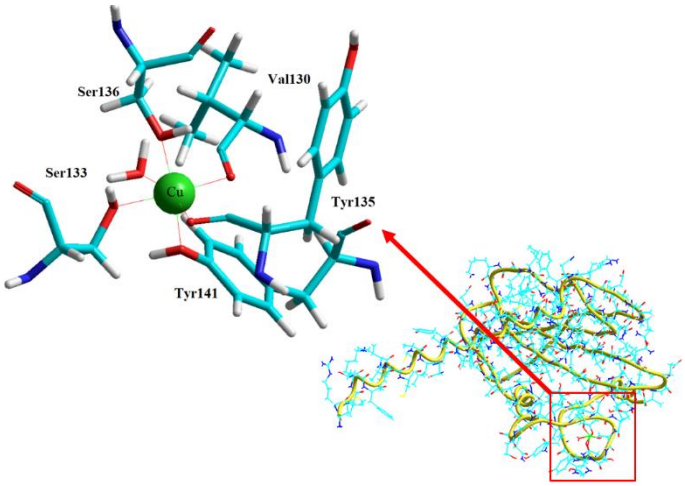

Fig.12. Coordination binding of copper(II) ions to the N-terminal domain in the H-chain

The probability of forming a coordination compound with the N-terminal domain in the H-chain of copper(II) ions in the form depicted in Figure 12 was determined. In this case, the total energy value was equal to -74.65 kcal/mol according to CHARMM22, and -205852 kcal/mol according to PM3. However, it turned out that the probabilities of coordination bonds of copper(II) ions with chains A1, A7, A8 and A10 are energetically very low.

Based on the results of theoretical calculation methods, it turns out that the probabilities of forming coordination bonds of copper(II) ions with the L-chain of SF, the N-terminal domain in the H-chain and amino acid residues in the P25 molecule are the highest. It can be assumed that copper(II) ions form complexes by binding to the functional groups of the last amino acid residues His220, Arg211, Gln219, Ser207 and Tyr56 of the P25 molecule.

Next, copper(II) ions can form complexes with the functional groups of the amino acid residues Ser44, Asp33, Asp 48, Gln63 and Lys67 in the L-chain. It can be considered that it can also form a complex with the amino acids in the N-terminal domain chain in the H-chain. The probability of copper(II) ions binding to the functional groups of the amino acid residues Val130, Ser133, Tyr135, Ser136 and Tyr141 in the N- terminal domain is high.

The table below compares the total energy values determined using the PM3 semi-empirical calculation method for models of possible complexes (Table 3).

Table 3. Copper(II) ions are coordination bonded parts

Copper(II) ions are coordination bonded part	ΔG, kcal/mol	E, kcal/mol
P25-Cu	-222585	-222494

L-chain-Cu	-199392	-199334
N-terminal domen-Cu	-205940	-205852

The calculation results showed that the probability of coordination of copper(II) ions with the functional groups of amino acid residues in the crystal parts is very low.

It can be assumed that the penetration of metal ions in an aqueous solution between SF molecules occurs most rapidly in the regions of fibroin molecules with high water affinity. Therefore, knowing which parts of SF have high water affinity (soluble parts) is very important. Theoretical calculations were performed to confirm these data.

CamSol software was used to determine the H- and L-chains of SF and the water affinity of P25 molecule. In the calculation, parts with a score greater than 1 indicate highly water-affected parts, and a score less than -1 indicate parts with very low water affinity. The overall result helps to determine the overall intrinsic solubility of the protein. Below are diagrams showing the results of calculations performed in the CamSol program for the L-chain of SF and the P25 molecule (Figure 13 and 14).

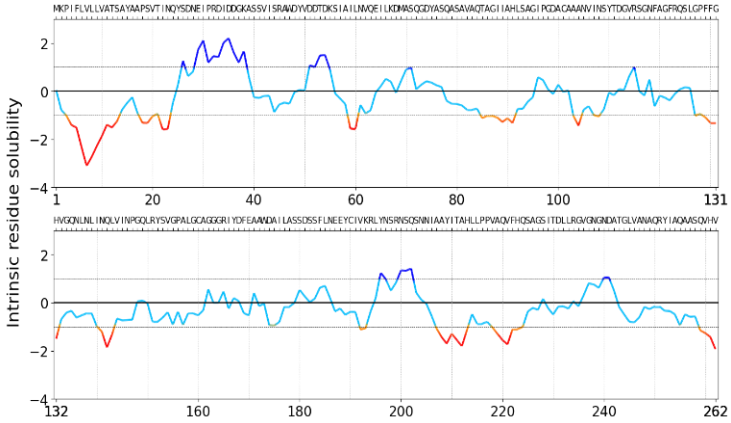

Fig.13. Map of the water-soluble parts of the L-chain.

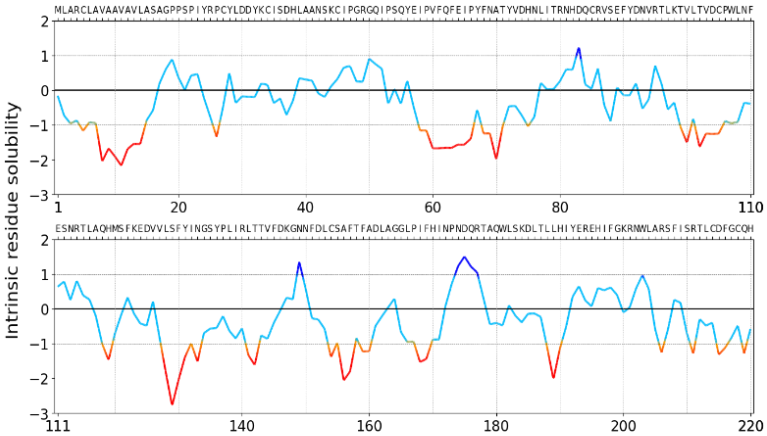

Fig.14. Map of the water-prone (soluble) parts of the P25 molecule of SF.

Since the H-chain of SF is very large, it was analyzed in parts. The N-terminal domain, the first crystal part and the A1 chain were selected for this. The analysis results showed that the highest value of water affinity was observed in the N-terminal domain. The A1 chain had a slightly lower value than the N-terminal domain. The lowest value was determined in the chain forming the crystal part (Figure 15-17).

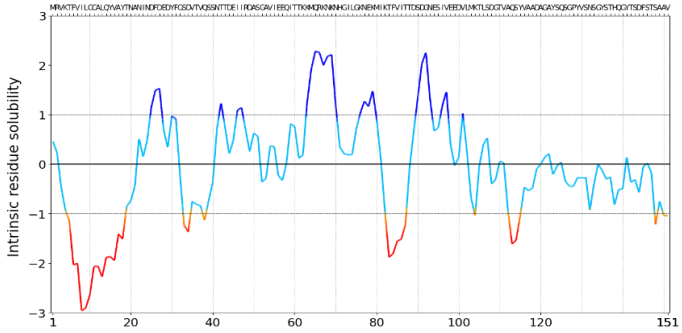

Fig.15. Map of the water-soluble parts of the N-terminal domain.

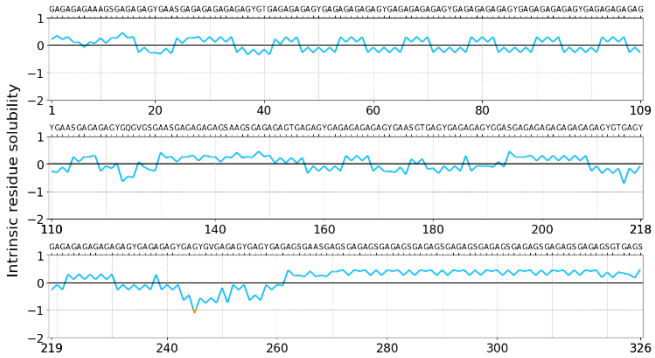

Fig.16. Map of the water-soluble parts of the crystal part after the N-terminal domain.

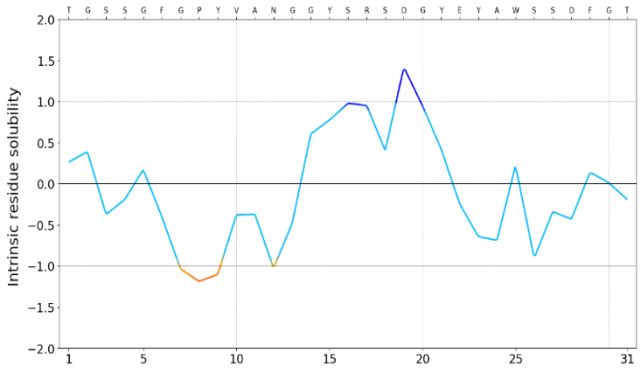

Fig.17. Map of the water-soluble parts of the A1 chain

Based on the calculation results, water penetration is intensive in the L-chain and the P25 molecule. In the H-chain, water penetration is intensive in the parts where the N-terminal domain and A1 chains are located. Therefore, the probability of $[Cu(H_2O)_6]^{2+}$ ions in the aqueous solution penetrating these parts and forming a complex compound was high.

An experiment was carried out to obtain a copper-fibroin complex. When an SF sample was placed in a solution containing copper(II) ions ($CuCl_2$) and stored for 7 days, it was observed that the SF sample turned from white to yellow. The resulting yellow sample was filtered and washed with hot distilled water until the excess copper(II) ions were completely washed out. The yellow color of the sample did not disappear. The observation of a color change can be considered an initial indication that SF and copper(II) ions formed a compound.

This sample was dried in a drying cabinet at a temperature of 70ºC. The yellow color was retained even after the drying process was completed. ATR-IR spectroscopy analysis was performed to determine which functional groups of this sample had changed (Figure 18).

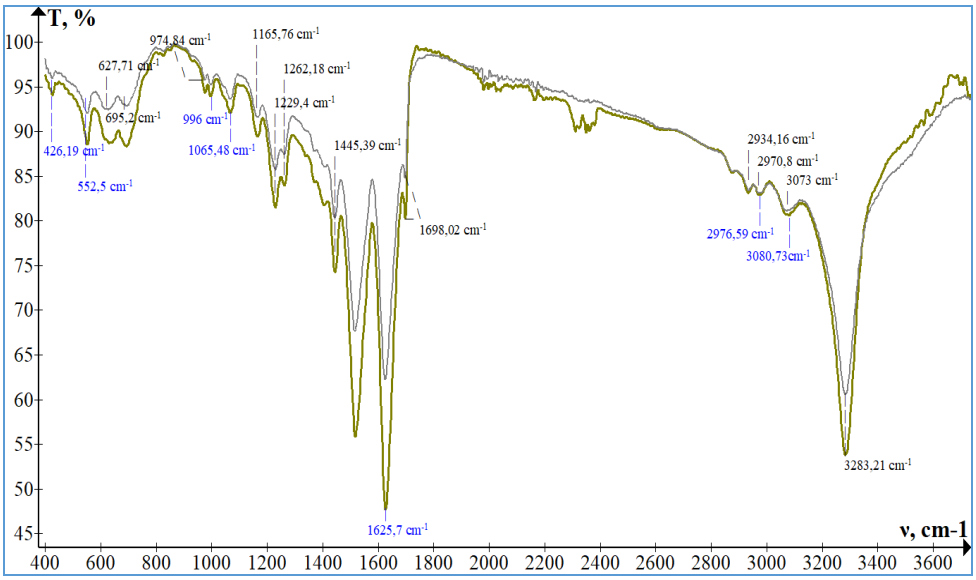

Fig.18. ATR-IR spectrum of the copper-fibroin complex obtained in the experiment.

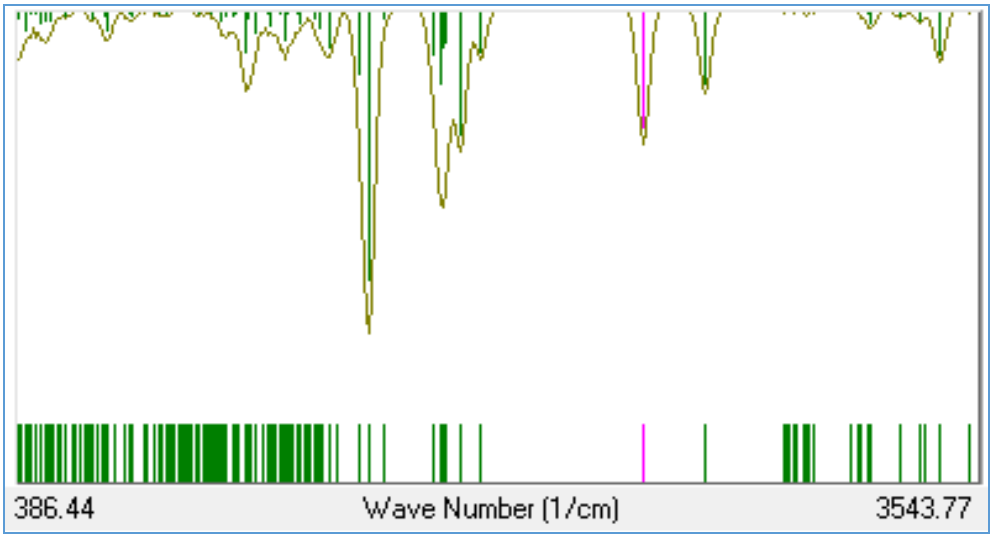

Fig.19. The IR spectrum of the Cu-fibroin complex obtained through theoretical calculations.

An increase in the intensity of the absorption at 426.19 cm^{-1} in the ATR-IR spectrum of the SF sample containing copper was observed. It is known that in the IR spectra of proteins bound by copper coordination bonds, absorptions related to Cu-N, and Cu-O bonds are observed in the regions 470-350 cm^{-1}. The absorption in the 426.19 cm^{-1} region in the ATR-IR spectrum indicates the formation of Cu-N and Cu-O bonds [43].

IR spectra were calculated by theoretical calculations based on models of the states in which copper(II) ions can form complexes with SF (Figure 19) [44, 45]. The table below compares the results of spectroscopic analysis obtained in practice and determined through theoretical calculations (Table 4).

Table 4. IR spectrum results of the copper-fibroin complex obtained by theoretical calculation and in practice

Calculation-derived IR spectrum results	Practical FTIR spectrum results	Copper-bound functional group
3283,2 cm^{-1}, 3293,7 cm^{-1}	3283,21 cm^{-1}	Cu-NH, Cu-O=C
1626,4 cm^{-1}	1625,7 cm^{-1}	Cu-O=C-Gly, Cu-NH-Gly
427,11 cm^{-1}	426,19 cm^{-1}	Cu-O=C, Cu-NH-

The results of the analysis showed that the theoretically calculated IR spectrum parameters were found to be in agreement with the ATR-FTIR spectrum parameters of the sample obtained in practice. These results confirm that the structural models determined on the basis of the calculations are correct.

When the yellow Cu-fibroin complex obtained in the experiment was dissolved in a 1:2:8 molar mixture of calcium chloride, ethanol and water, and the UV-Vis spectrum was obtained, maximum absorption was observed in the 275 and 326 nm regions. The absorption in the 275 nm region belongs to the fibroin protein. We considered the absorption in the 326 nm region to be related to the complex compounds formed by copper(II) ions with proteins. The maximum absorption is observed in the 540-550 nm regions in the complex compounds formed by copper(II) ions with proteins under alkaline conditions. The reaction of copper(II) ions with SF in an aqueous solution results in the formation of a complex with a specific composition, as indicated by the color of the sample and the UV-Vis spectrum (326 nm) (Figure 20).

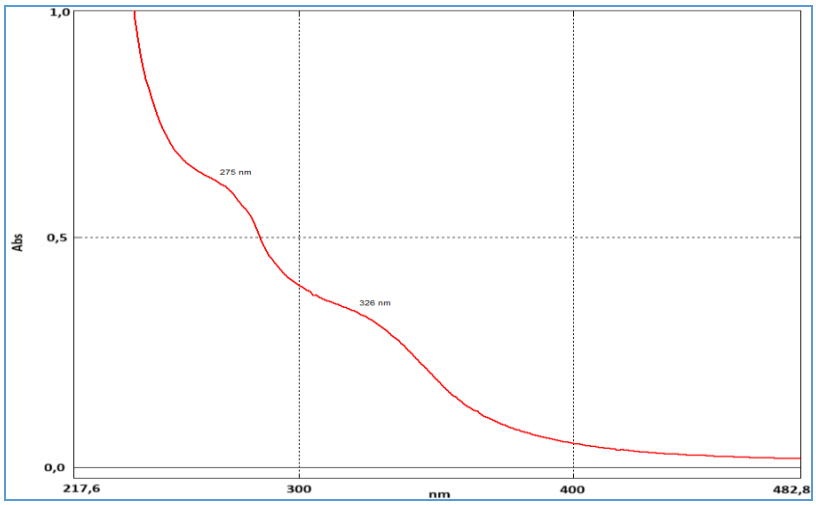

Fig.20. UV-Vis spectrum of the Cu-fibroin complex obtained in the experiment: absorptions were observed in the wavelength regions of 275 and 326 nm.

Based on the above models, their UV spectra were determined by quantum chemical calculations. The calculations were performed using the ZINDO/S semi-empirical method (Figure 21-23).

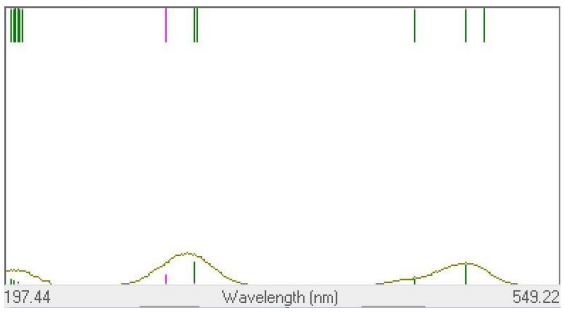

Fig.21. UV spectrum of the copper-fibroin complex formed in the L-chain.

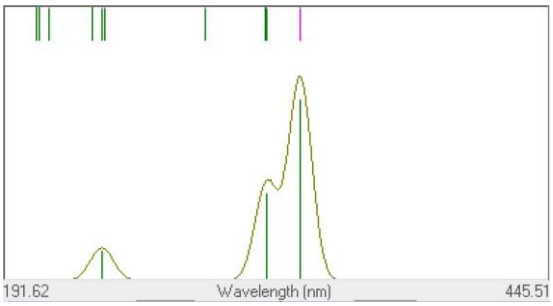

Fig.22. UV spectrum of the copper-fibroin complex formed with the P25 molecule.

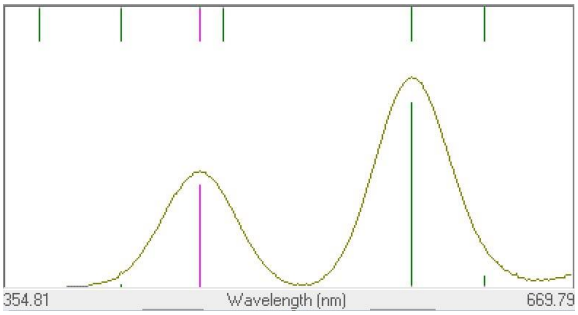

Fig.23. UV spectrum of the complex formed in the N- terminal domain of the H-chain.

If we pay attention to the UV spectrum results obtained through theoretical calculations, the absorption for the copper-fibroin complex formed in the L-chain was observed at 308 nm, for the complex formed in the P25 molecule at 313.7 and 329.7 nm, and for the complex formed in the N- terminal domain at 462.5 nm. It can be seen that the UV-Vis spectrum of the copper-fibroin complex obtained in the experiment is relatively consistent with the theoretically calculated UV spectrum results of the complex formed in the L-chain and especially in the P25 molecule. Based on the results of experimental physical studies and theoretical calculations, it can be assumed that copper(II) ions in solution form copper-fibroin complexes with amino acid residues in the L-chain of SF and the P25 molecule.

CONCLUSION

Theoretical calculations revealed that the coordination bonds of copper (II) ions with the C=O and –OH groups of Ser44, with the $–COO^-$ group of the Asp33 and Asp48, with the $–(CH_2)_2NH_2$ group of Gln63, and with the $–(CH_2)_4NH_2$ group of Lys67 in L-fibroin are energetically most favourable. It was found that the copper (II) ion forms relatively stable coordination bonds with the N and $–COO^-$ of the side group of His220 in the P25 molecule, with the $–(CH_2)_3NHCNHNH_2$ group of Arg211, with the $–CH_2-C_6H_4-OH$ group of Tyr56, with the $–(CH_2)_2NH_2$ group of Gln219 and

with the –OH group of Ser207. It was found that copper (II) ions form coordination bonds only with the N-domain in the heavy chain. The results of the calculated IR and UV spectra of the copper-fibroin complex determined by quantum chemical calculations are consistent with the ATR-IR and UV-Vis spectra of the experimentally obtained sample, confirming the results of the above study. This study may contribute to the study of the interaction of other metal ions with silk fibroin to some extent.

Author contributions

KE author designed the research; KE and GP authors performed the experiments; KE and MB authors analysed the data; KE and RE authors wrote the manuscript draft; KE and AS authors revised the manuscript. All authors approved the final version of the manuscript.

REFERENCES:

1. Cook D. B. (1998) *Handbook of computational quantum chemistry*, Oxford University Press.

2. Bocharnikova, E. N., Tchaikovskaya, O. N., Bazyl, O. K., Artyukhov, V. Y., & Mayer, G. V. (2020) Theoretical study of bisphenol A photolysis. *Adv. Quantum Chem.*, 81, 191-217.

3. Hakimi O, Knight D.P, Vollrath F, Vadgama P (2007) Spider and mulberry silkworm silks as compatible biomaterials. *Composites. Part B. Engineering.*, 38, 324–337.

4. Inoue S, Tanaka K, Arisaka F, Kimura S, Ohtomo K, Mizuno S. (2000) Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular mass elementary unit consisting of H-chain, L-chain, and P25, with a 6:6:1 molar ratio. *J. Biol. Chem.*, 275, 40517–40528.

5. Moreno-Tortolero, R. O., Luo, Y., Parmeggiani, F., Skaer, N., Walker, R., *et al*. (2024) Molecular organization of fibroin heavy chain and mechanism of fibre formation in Bombyx mori. *Commun. Biol.*, 7, 786.

6. Valluzzi R, Gido S.P, Muller W, Kaplan D.L. (1999) Orientation of silk III at the air-water interface". *Int. J. Biol. Macromol.*, 24, 237–242.

7. Suzuki, Y. (2016) Structures of silk fibroin before and after spinning and biomedical applications. *Polym. J.*, 48, 1039-1044.

8. Takahashi, Y., Gehoh, M., & Yuzuriha, K. (1999) Structure refinement and diffuse streak scattering of silk (Bombyx mori). *Int. J. Biol. Macromol.*, 24, 127-138.

9. Suzuki, Y., & Asakura, T. (2010) Local conformation of serine residues in a silk model peptide,(Ala–Gly–Ser–Gly–Ala–Gly) 5, studied with solid-state NMR: REDOR. *Polym. J.*, 42, 354-356.

10. Magoshi, J., Magoshi, Y., Becker, M. A., Kato, M., Han, Z., *et al*. (2000) Crystallization of silk fibroin from solution. *Thermochim. Acta*., 352, 165-169.

11. Warwicker, J. O. (1956) The crystal structure of silk fibroins. *Trans. Faraday Soc.*, 52, 554-557.

12. Takahashi Y., Gehoh M. and Yuzuriha K. (1991) Crystal structure of silk (Bombyx mori). *J. Polym. Sci., Part B:Polym. Phys*., 29, 889-891.

13. Sarymsakov, A. A., Baltaeva, M. M., & Eshchanov, K. O. (2024) Preparation and study of properties of sorbents based on the fibrous waste of “Bombyx mori” natural silk. *Lat. Am. Appl. Res*., 54, 531-538.

14. Ryde U. (2016) QM/MM Calculations on Proteins. *Methods Enzymol.*, 577, 119-158.

15. Schmitz, S., Seibert, J., Ostermeir, K., Hansen, A., Göller, A. H., *et al*. (2020) Quantum chemical calculation of molecular and periodic peptide and protein structures. *J. Phys. Chem. B.*, 124, 3636-3646.

16. Žurovec, M., Vašková, M., Kodrík, D., Sehnal, F., & Kumaran, A. K. (1995) Light-chain fibroin of Galleria mellonella L. *Mol. Gen. Genet*., 247, 1-6.

17. Teramoto, H., & Kojima, K. (2015) Incorporation of methionine analogues into Bombyx mori silk fibroin for click modifications. *Macromol. Biosci*., 15, 719-727.

18. Zorman, M., Phillips, C., Shi, C., Zhang, S., De Yoreo, et al. (2024) Thermodynamic Analysis of Silk Fibroin–Graphite Hybrid Materials and Their Morphology. *J. Phys. Chem. B.*, 128, 2371-2380.

19. Montalban, M. G., Chakraborty, S., Pena-Garcia, J., Verli, H., Villora, G., et al. (2020) Molecular insight into silk fibroin based delivery vehicle for amphiphilic drugs: Synthesis, characterization and molecular dynamics studies. *J. Mol. Liq*., 299, 112156.

20. Przybyłek M., Tuwalska A., Ledziński D., Śmigiel S., Sionkowska A., et al. (2024) Effect of Nanohydroxyapatite on Silk Fibroin–Chitosan Interactions—Molecular Dynamics Study, *Appl. Sci.*, 14, 4131.

21. Leal Egaña, A., & Scheibel, T. (2010) Silk-based materials for biomedical applications. *Biotechnol. Appl. Biochem*., 55, 155-167.

22. de Almeida, K. J., Rinkevicius, Z., Hugosson, H. W., Ferreira, A. C., & Ågren, H. (2007) Modeling of EPR parameters of copper (II) aqua complexes. *Chem. Phys*., 332, 176-187.

23. Musinu, A., Paschina, G., Piccaluga, G., and Magini, M. (1983) Short-range order in a NASIGLAS sample by X-ray diffraction. *Inorg. Chem*., 22, 1184.

24. Ansell, S., Tromp, R. H., & Neilson, G. W. (1995) The solute and aquaion structure in a concentrated aqueous solution of copper (II) chloride. *J. Phys. Condens. Matter*., 7, 1513-1524.

25. Texler, N. R., Holdway, S., Neilson, G. W., & Rode, B. M. (1998) Monte Carlo simulations and neutron diffraction studies of the peptide forming system 0.5 mol kg− 1 CuCl 2–5mol kg− 1 NaCl–H 2 Oat 293 and 353 K. *J. Chem. Soc., Faraday Trans.*, 94, 59-65.

26. Sham, T.K., Hastings, J.B., and Perlman, M.L., Chem. Phys. Lett., 1981, vol. 83, no. 2, p. 391.

27. Neilson, D.L., Newsome, J.R., and Sandstrom, M., J. Chem. Soc., Faraday Trans., 1981, vol. 77, no. 7, p.1245.

28. Tajiri, Y. and Wakita, H., Bull. Chem. Soc. Japan., 1986, vol. 59, no. 7, p. 2285.

29. Li, Y.; Yao, L.; Zhang, L.; Zhang, Y.; Zheng, T.; Liu, L.; Zhang, L. Enhanced physicochemical stabilities of cyanidin-3-O-glucoside via combination with silk fibroin. Food Chem. 2021, 355, 129479.

30. Yao, L.; Xu, J.; Zhang, L.; Liu, L.; Zhang, L. Nanoencapsulation of Anthocyanin by an Amphiphilic Peptide for Stability Enhancement. Food Hydrocoll. 2021, 118, 106741.

31. Chen, W. X., Lu, S. F., Yao, Y. Y., Pan, Y., & Shen, Z. Q. (2005). Copper (II)-silk fibroin complex fibers as air-purifying materials for removing ammonia. Textile research journal, 75(4), 326-330.

32. Hua, J., Yang, H., Li, X., Xiao, J., Zhou, S., You, R., & Ma, L. (2022). Cu (II)-functionalized silk fibroin films for the catalytic generation of nitric oxide. Biointerphases, 17(3), 031001.

33. MacKerell AD Jr, Feig M, Brooks III CL (2004). "Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations". J Comput Chem. 25 (11): 1400–1415.

34. Brooks CL, Chen J, Im W (2006). "Balancing solvation and intramolecular interactions: toward a consistent generalized born force field (CMAP opt. for GBSW)". J Am Chem Soc. 128 (11): 3728–3736.

35. Vanommeslaeghe, K.; MacKerell, A. D. (May 2015). "CHARMM additive and polarizable force fields for biophysics and computer-aided drug design". Biochimica et Biophysica Acta (BBA) - General Subjects. 1850 (5): 861–871.

36. Mostashari-Rad T, Arian R, Mehridehnavi A, Fassihi A, Ghasemi F (June 13, 2019). "Study of CXCR4 chemokine receptor inhibitors using QSPR andmolecular docking methodologies". Journal of Theoretical and Computational Chemistry. 178 (4).

37. Kitchen DB, Decornez H, Furr JR, Bajorath J (Nov 2004). "Docking and scoring in virtual screening for drug discovery: methods and applications". Nature Reviews. Drug Discovery. 3 (11): 935–49. doi:10.1038/nrd1549.

38. Hwang, W., Mallis, R. J., Lang, M. J., & Reinherz, E. L. (2020). The αβ TCR mechanosensor exploits dynamic ectodomain allostery to optimize its ligand recognition site. Proceedings of the National Academy of Sciences, 117(35), 21336-21345.

39. Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., & Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. Proceedings of the National Academy of Sciences, 117(3), 1496-1503.

40. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 30(16), 2785-2791.

41. Hebditch, M., & Warwicker, J. (2019). Web-based display of protein surface and pH-dependent properties for assessing the developability of biotherapeutics. Scientific reports, 9(1), 1969.

42. Sormanni, P., Aprile, F. A., & Vendruscolo, M. (2015). The CamSol method of rational design of protein mutants with enhanced solubility. Journal of molecular biology, 427(2), 478-490.

43. Nakamoto, K. Infrared and Raman spectra of inorganic and coordination compounds. Part A: theory and applications in inorganic chemistry; Part B: application in coordination, organometallic, and bioinorganic chemistry, John Wiley & Sons, Inc., 2009.

44. Hayward S, Kitao A, Go N. Harmonic and anharmonic aspects in the dynamics of BPTI—a normal-mode analysis and principal component analysis. Protein Science, 3, (1994) 936–943.

45. Mott, A. J., & Rez, P. Calculation of the infrared spectra of proteins. European biophysics journal, 44(3), (2015) 103–112.