EXPERIMENTAL ARTICLES

UDC 579.22:577.322 https://doi.org/10.15407/biotech18.03.014

APPLICATION OF REVERSE MOLECULAR DOCKING FOR THE IDENTIFICATION OF PROTEIN TARGETS OF S-ETHYLTHIOSULFANYLATE INVOLVED IN BIOSURFACTANT BIOSYNTHESIS

Yanvarov Ye.B.
Havryliak V.V.

Lviv Polytechnic National University, Ukraine

E-mail: yehor.b.yanvarov@lpnu.ua

Received 2025/03/19
Revised 2025/04/13
Accepted 2025/06/30

Aim. The study is focused on determining the effect of the ligand S-ethylthiosulfanylate on protein targets involved in the synthesis of biosurfactants and evaluating their potential interaction.

Materials and Methods. A reverse docking approach was employed to investigate the interaction of a single ligand with 8 protein targets. Molecular docking was performed using AutoDock Vina with the Vina scoring function. The preparation of the ligand and protein targets was carried out using AutoDockTools from the MGLTools package. Visualization of the results was accomplished using ChimeraX and BIOVA Discovery Studio.

Results. Docking the ligand with 8 protein targets enabled the identification of three promising targets — 3RKY, 2B4Q, and 8IK2 — with affinities lower than −5.5 kcal/mol. Predominantly, hydrogen bonds and hydrophobic interactions were observed, indicating the stability of ligand binding within the active sites of these proteins.

Conclusions. The study confirmed the effectiveness of reverse docking for identifying potential protein targets, demonstrating that the ligand can influence biosurfactant biosynthesis through specific interactions with proteins 3RKY, 2B4Q, and 8IK2.

Key words: biosurfactants, S-ethylthiosulfanylate, reverse molecular docking, ligand-protein interaction, affinity, target prediction.

In today's rapidly evolving computing environment, with new algorithms and the development of artificial intelligence, molecular docking continues to be a powerful tool for studying biomolecular interactions. It is especially relevant to study the interaction of a particular ligand with different proteins, as it allows us to investigate the mechanisms of ligand action and discover new protein targets. The use of reverse docking will enable us to evaluate the interaction of one ligand with a group of potential proteins, which significantly speeds up the discovery of new interactions.

Recent advances in protein structure prediction, including the AlphaFold neural network, have significantly expanded the database of available protein structures [1–3], so reverse docking has gained a new impetus. The authors of AlphaFold received the Nobel Prize in Chemistry in 2024 for developing this algorithm. Modern approaches allow for the study of a large number of proteins, which opens up prospects for the discovery of unconventional targets.

Citation: Yanvarov, Y. B., Havryliak, V. V. (2025). Application of reverse molecular docking for the identification of protein targets of S-ethylthiosulphanilate involved in biosurfactant biosynthesis. *Biotechnologia Acta*, 18(3), 14–22. https://doi.org/10.15407/biotech18.03.014

Recent studies on reverse docking have shown that this method is an effective tool for screening many protein targets, allowing for the rapid identification of potential ligand targets [4–7]. The main algorithms used in these studies include numerous methods for energy estimation and complex configuration optimization, including algorithms integrated into software packages such as AutoDock, Glide, and DOCK.

The main goal of our study was to identify protein targets involved in the biosurfactants' biosynthesis and to evaluate the potential interaction of the synthesized ligand S-ethylthiosulfanylate (ETS) (Fig. 1) [8] with target proteins.

Using reverse docking will allow us to identify proteins with which the ligand under study may have the highest binding energy.

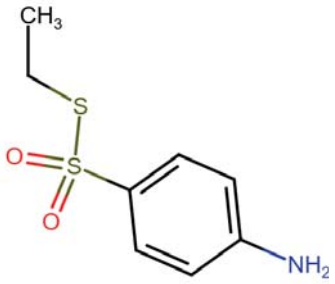


Fig. 1. **Structural formula of the ligand S-ethylthiosulfanylate**

The results can form the basis for a hypothesis regarding the mechanisms of interaction of the ligand with key enzymes of metabolic pathways. They will help in the experimental study of its interaction.

It is known that RhlA and RhlG proteins play a key role in the biosynthesis of rhamnolipids in *Pseudomonas aeruginosa* [9, 10]. Analysis of their potential interactions with ligands can provide new knowledge about the regulation mechanisms of this process. This approach allows us to identify the target proteins with which the ligand forms the most stable complexes and narrow the direction of further experimental studies accordingly.

Materials and Methods

We studied the structures of eight protein targets: 1ABE, 1JMK, 2B4Q, 2CBG, 3RKY, 4MRT, 8F7F, and 8IK2 (Table 1), which were downloaded from the RCSB Protein Data Bank (RCSB PDB). These proteins were chosen based on their involvement in metabolic pathways that regulate the biosynthesis of microbial biosurfactants.

Preprocessing of the protein structures was performed using MGLTools 1.5.7. Solvent molecules, heteroatoms, ions, and ligands in the original PDB files were removed. Missing amino acid side chains of proteins were restored, hydrogen atoms were added, and charges were calculated using AutoDockTools software.

Table. **Structures of the target proteins**

No,	PDB ID	Description	Organism	Reference
1	1ABE	NOVEL STEREOSPECIFICITY OF THE L-ARABINOSE-BINDING PROTEIN	*Escherichia coli*	[11]
2	1JMK	Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE	*Bacillus subtilis*	[12]
3	2B4Q	*Pseudomonas aeruginosa* RhlG/NADP active-site complex	*Pseudomonas aeruginosa*	[13]
4	2CBG	Crystal structure of the PMSF-inhibited thioesterase domain of the fengycin biosynthesis cluster	*Bacillus subtilis*	[14]
5	3RKY	Structural characterisation of *Staphylococcus aureus* biotin protein ligase	*Staphylococcus aureus*	[15]
6	4MRT	Structure of the Phosphopantetheine Transferase Sfp in Complex with Coenzyme A and a Peptidyl Carrier Protein	*Brevibacillus parabrevis*	[16]
7	8F7F	The condensation domain of surfactin A synthetase C in space group P43212	*Bacillus subtilis*	[17]
8	8IK2	RhlA exhibits dual thioesterase and acyltransferase activities during rhamnolipid biosynthesis	*Pseudomonas aeruginosa*	[18]

Binding sites were determined for each protein based on the natural ligands present in that structure. If the binding site contained several alternative positions, all conformational changes of the active site were considered.

The S-ethylthiosulfanylate ligand was obtained as a SMILES file and converted to PDB format using OpenBabel 3.1.1. Further preparation of the ETS ligand was carried out using AutoDockTools software, which included the following steps: addition of polar hydrogen atoms, calculation of charges, and torsional degrees of freedom.

Reverse docking was performed using AutoDock Vina [19]. The search for an accurate conformation was confirmed by the following parameters: num_modes = 20, exhaustiveness = 64. The search window was defined around the previously identified binding sites. In cases where proteins had several known active sites, docking was performed on each of them separately.

The resulting protein-ligand complexes were ranked according to the interaction affinity calculated by AutoDock Vina. Only complexes with a binding energy below −5,5 kcal/mol were selected for further analysis. The top three protein targets meeting this criterion were subjected to additional visual inspection using ChimeraX 1.9 and BIOVA Discovery Studio 2025.

ChimeraX 1.9 and BIOVA Discovery Studio 2025 were used for visual analysis of the obtained molecular complexes. The geometric parameters of binding, conformational stability of the complex, and the nature of interactions (hydrogen bonds, hydrophobic bonds, π-π stacking interactions) were evaluated.

Results and Discussion

The results of reverse molecular docking (Fig. 2), performed using AutoDock Vina, revealed varying degrees of affinity of the ligand ETS toward the selected protein targets. The lowest binding energy values were observed for proteins 3RKY (−6,364 kcal/mol), 2B4Q (−5,885 kcal/mol), and 8IK2 (−5,51 kcal/mol), indicating their potential capability to interact with the ligand.

The strongest binding affinity of ETS was observed for protein 3RKY, which exhibited the most favorable binding energy (−6,364 kcal/mol), suggesting a highly stable complex conformation. Proteins 2B4Q (−5,885 kcal/mol) and 8IK2 (−5,51 kcal/mol) also showed high affinity toward the ligand, making them promising candidates for further investigation. In contrast, proteins 1ABE (−3,081 kcal/mol) and 1JMK (−3,626 kcal/mol) demonstrated considerably weaker binding energies, implying low

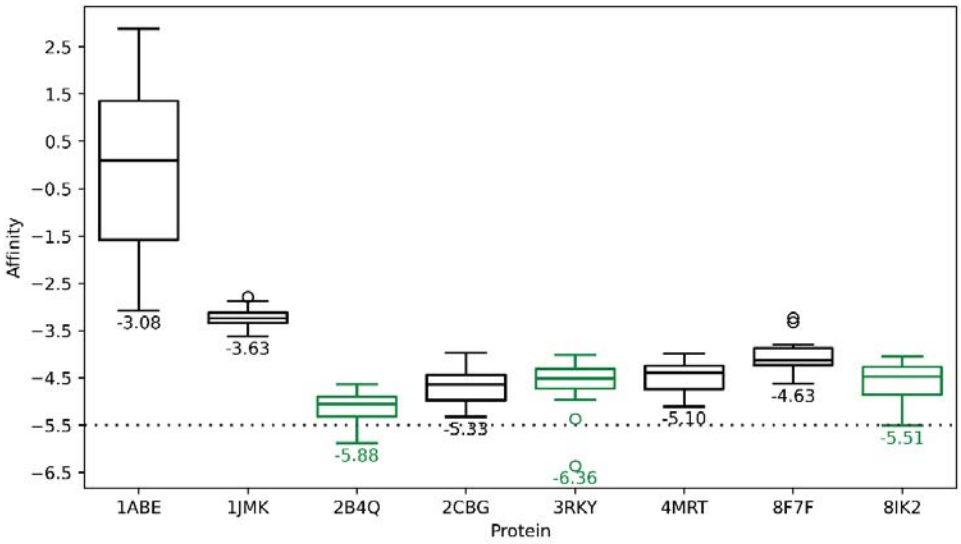


Fig. 2. **Distribution of binding affinity values (kcal/mol) of the ligand ETS with the target proteins 1ABE, 1JMK, 2B4Q, 2CBG, 3RKY, 4MRT, 8F7F, and 8IK2**

For each protein, 20 binding affinity values obtained from molecular docking simulations are presented. The black horizontal dashed line represents the affinity threshold of −5.5 kcal/mol. Proteins exhibiting affinity values below this threshold are highlighted in green, while those above the threshold are shown in black. The minimum binding energy for each protein is indicated numerically below the corresponding boxplot.

interaction strength and, consequently, limited biological relevance.

Conformational stability was assessed by analyzing the interquartile range of binding affinity values, specifically the difference between the first and third quartiles. For protein 3RKY, a narrow affinity range was observed, indicating overall stability in ligand positioning, along with the lowest binding energy, which may suggest optimal anchoring of the primary ligand within the binding site. A similar pattern was noted for proteins 2B4Q and 8IK2, implying the presence of well-defined binding pockets. In contrast, proteins 1ABE and 1JMK exhibited wider ranges of high binding energy values, which may reflect the flexibility or instability of ligand binding within their active sites.

Subsequent visualization of ligand-protein interactions using ChimeraX and BIOVIA Discovery Studio enabled the identification of key amino acid residues involved in complex formation (Fig. 3). In the case of protein 3RKY, the primary interactions included hydrogen bonds and hydrophobic contacts, which contributed to its high binding affinity. A similar interaction pattern was observed for proteins 2B4Q and 8IK2, where the ligand established stable non-covalent interactions within the binding site.

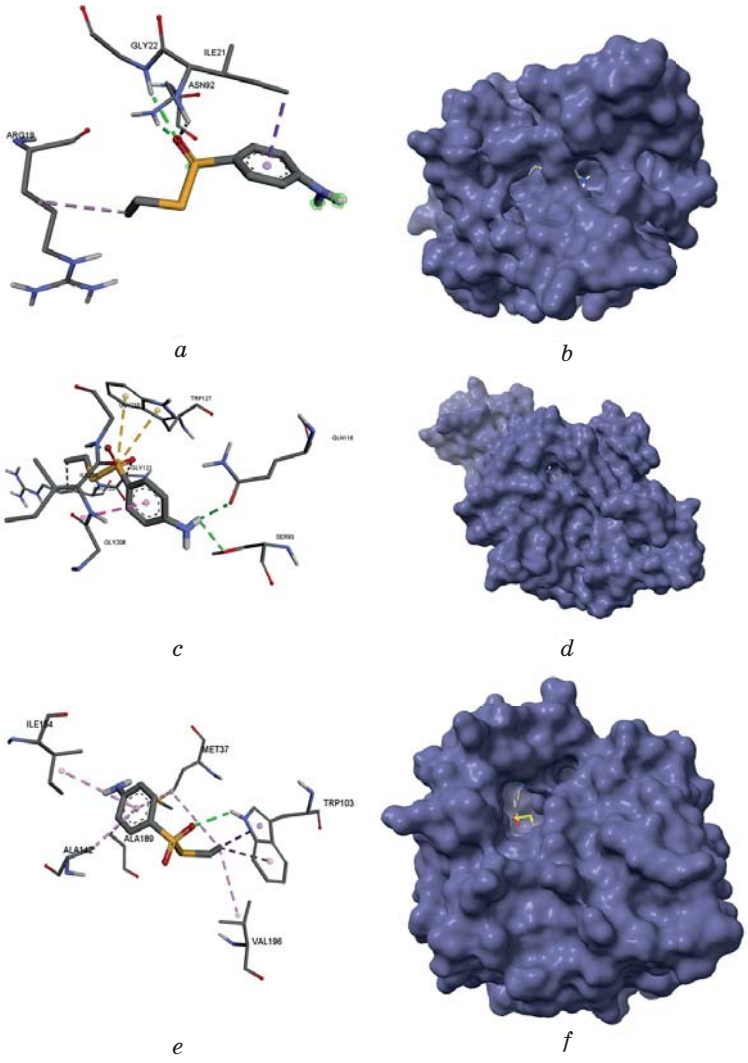


Fig. 3. **Visualization of the best binding conformation of ETS within the binding sites of three proteins — 2B4Q, 3RKY, and 8IK2 — with binding affinities of −5,88, −6,36, and −5,51 kcal/mol, respectively:** *a* — ETS in the active site of protein 2B4Q with annotated amino acid residues involved in the interaction; *b* — ETS positioned within the 3D surface representation of protein 2B4Q; *c* — ETS in the active site of protein 3RKY with annotated interacting residues; *d* — ETS within the 3D surface context of protein 3RKY; *e* — ETS in the active site of protein 8IK2 with annotated amino acid residues involved in the interaction; *f* — ETS within the 3D surface representation of protein 8IK2.

Analysis of the ETS–2B4Q protein complex (Fig. 4) revealed three conventional hydrogen bonds formed between the carbonyl group of the ligand (acting as a proton acceptor) and the amide groups of residues Ile21, Gly22, and Asn92, with bond lengths ranging from 1.97 to 2.25 Å. These hydrogen bonds displayed favorable angles (99.26–165.13°), indicative of high geometric stability. Additionally, a hydrophobic π–σ interaction was observed between the side chain of Ile21 (CD1) and the π-system of the ligand at a distance of 3.50 Å, along with an alkyl interaction between a carbon atom of the ligand and Arg19 at 4.97 Å.

Analysis of the ETS–3RKY complex (Fig. 5) revealed several hydrogen and hydrophobic interactions that contribute to the stable binding of the ligand within the protein active site. Specifically, three conventional hydrogen bonds (ranging from 3.39 to 5.19 Å) were formed between the protonated groups of the ligand and residues Gly210, Ser93, and Gln116. Additionally, a C–H···O interaction involving Gly121 was observed at a distance of 3.81 Å. The favorable bond angles (approximately 100–136°) suggest a high degree of geometric complementarity in the hydrogen bonding network.

In addition to hydrogen bonding, the complex structure revealed π–S interactions (4.11–4.24 Å) between the sulfur-containing group of the ligand and the aromatic residue

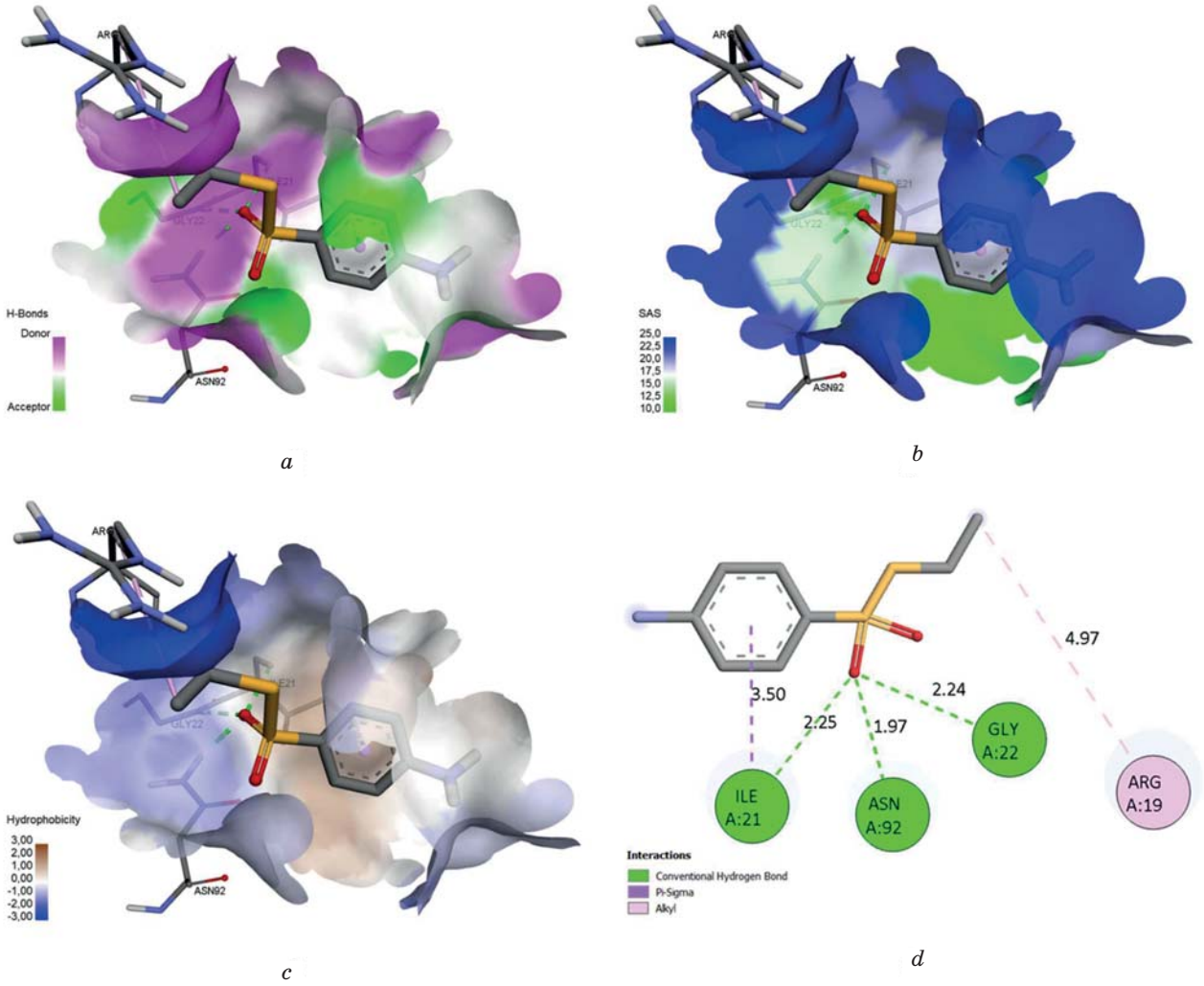
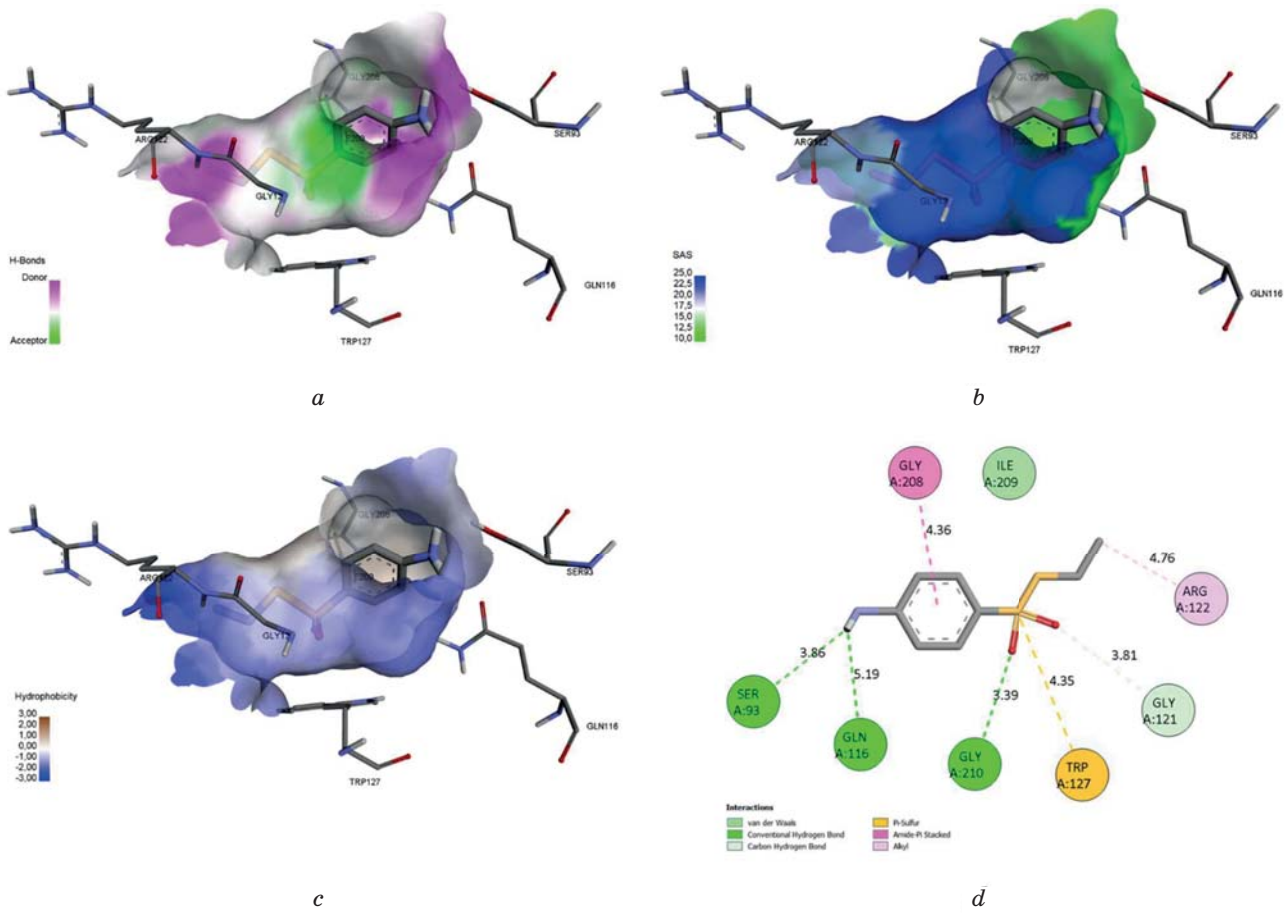


Fig. 4. **Interaction profile of the best binding conformation of ETS within the binding pocket of protein 2B4Q**

a — contribution of hydrogen bonding interactions; *b* — solvent accessibility of the binding pocket; *c* — hydrophobic properties of the binding site; *d* — 2D interaction diagram summarizing ligand-protein interactions.



a *b*

c *d*

Fig. 5. **Interaction profile of the best binding conformation of ETS within the binding pocket of protein 3RKY:**
a — contribution of hydrogen bonding interactions; *b* — solvent accessibility of the binding pocket; *c* — hydrophobic properties of the binding site; *d* — 2D interaction diagram summarizing ligand-protein interactions.

Trp127, as well as an amide–π stacking interaction involving the amide moieties of Gly208/Ile209 (4.07 Å). An alkyl interaction with Arg122 (4.17 Å) was also identified, further contributing to the stabilization of the complex.

Analysis of the ETS–8IK2 complex (Fig. 6) revealed a number of interactions that contribute to the effective binding of the ligand within the active site. In particular, a conventional hydrogen bond (2.50 Å) was identified between the protonated hydrogen atom of residue Trp103 (HE1) and the carbonyl group of the ligand. The favorable geometric parameters (DHA angle ~120° and HAY angle ~163°) indicate a high degree of stability for this interaction.

In addition to the hydrogen bond, the ligand engages in hydrophobic interactions with several residues. Specifically, a π–σ interaction (3.90 Å) is observed between one of the ligand's carbon atoms and the π-system of Trp103, while π–alkyl contacts (4.45–5.33 Å) are formed with Met37, Ala142, Ile154, and Ala189. Alkyl interactions are also detected with Val196 and Met37 at distances of 4.41 Å and 5.50 Å, respectively.

The combined contribution of hydrogen bonds and hydrophobic interactions (π–σ, π–alkyl, and alkyl) underscores the cooperative role of both polar and nonpolar contacts in stabilizing the complex. The observed spatial complementarity between the ligand and the 8IK2 binding pocket is consistent with the high binding affinity predicted by docking analysis.

Conclusions

The application of reverse molecular docking for S-ethylthiosulfanylate against eight protein targets revealed several

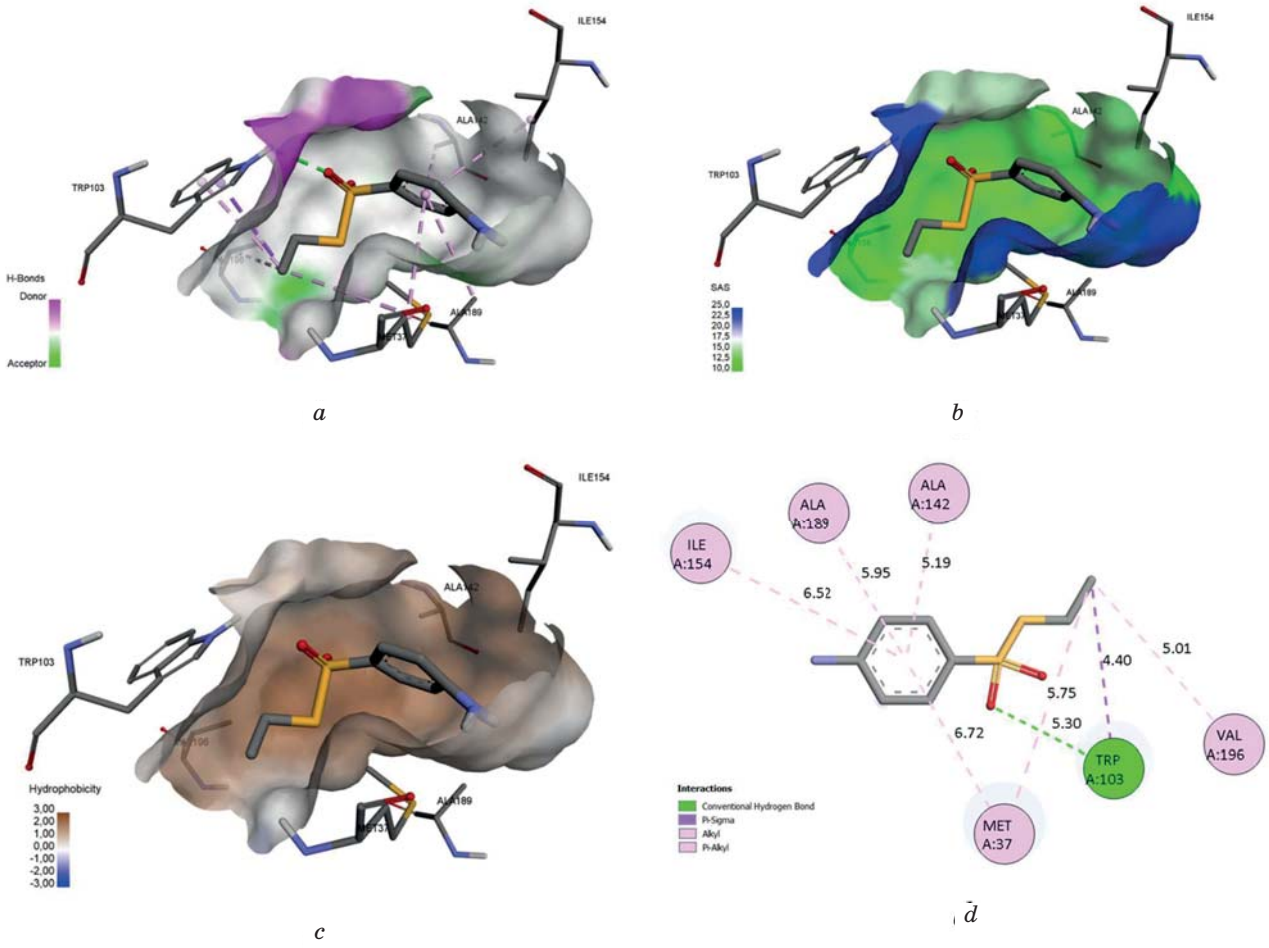


***Fig. 6*. Interaction profile of the best binding conformation of ETS within the binding pocket of protein 8IK2:**
a — contribution of hydrogen bonding interactions; *b* — solvent accessibility of the binding pocket; *c* — hydrophobic properties of the binding site; *d* — 2D interaction diagram summarizing ligand-protein interactions.

structures exhibiting high binding affinity. Notably, proteins 3RKY, 2B4Q, and 8IK2 demonstrated the most stable ligand binding within their respective active sites.

Detailed analysis of the docking results and subsequent visualization of the complexes confirmed the presence of hydrogen bonds, hydrophobic (π–σ, π–alkyl), and alkyl interactions between the ligand and amino acid residues of the target proteins. These interactions contribute to the spatial complementarity and overall stability of the ligand-protein complexes.

Given that the investigated proteins are likely involved in biosurfactant biosynthesis, the findings imply a potential role of the ligand in modulating or regulating their function.

While molecular docking provides a valuable first-line screening method, experimental validation — such as molecular dynamics simulations or biophysical assays — is necessary to confirm the ligand's biological activity and to clarify its binding mechanisms.

Funding
The present research was conducted without any external financial support.

Conflicts of Interest
The authors declare no conflicts of interest.

Author Contributions
All authors contributed equally to the conception and design of the study, selection of protein targets, ligand and protein preparation, molecular docking, and data interpretation. All authors participated in the drafting, revising, and final approval of the manuscript.

REFERENCES

1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., ..., Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
2. Jumper, J., Hassabis, D. (2022). Protein structure predictions to atomic accuracy with AlphaFold. *Nature Methods*, 19(1), 11–12. https://doi.org/10.1038/s41592-021-01362-6
3. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L, ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w
4. Krause, F., Voigt, K., Di Ventura, B., Ozturk, M. A. (2023). ReverseDock: A web server for blind docking of a single ligand to multiple protein targets using AutoDock Vina. *Frontiers in Molecular Biosciences*, 10. https://doi.org/10.3389/fmolb.2023.1243970
5. Wang, H., He, H., Zhang, T., Jiang, J. (2022). Application of Reverse Docking in the Research of Small Molecule Drugs and Traditional Chinese Medicine. *Biological and Pharmaceutical Bulletin*, 45(1), 19-26. https://doi.org/10.1248/bpb.b21-00324
6. Luo, Q., Wang, S., Li, H. Y., Zheng, L., Mu, Y., Guo, J. (2024). Benchmarking reverse docking through AlphaFold2 human proteome. *Protein Science*, 33(10), e5167. https://doi.org/10.1002/pro.5167
7. Kharkar, P. S., Warrier, S., Gaud, R. S. (2014). Reverse Docking: A Powerful Tool for Drug Repositioning and Drug Rescue. *Future Medicinal Chemistry*, 6(3), 333–342. https://doi.org/10.4155/fmc.13.207
8. Lubenets, V. I., Baranovych, D. B., Yarish, M. Ye., Voloshynets, V. A., Novikov, V. P. (2001). Synthesis of S-[2-(arylsulfonyl)ethyl]thiosulfanates and their hydrolysis. *Ukr. khym. zhurn*, 67(11–12), 103–109. (In Ukrainian)
9. Deziel, E., Lepine, F., Milot, S., Villemur, R. (2003). rhlA is required for the production of a novel biosurfactant promoting swarming motility in *Pseudomonas aeruginosa*: 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs), the precursors of rhamnolipids. *Microbiology*, 149(8), 2005–2013. https://doi.org/10.1099/mic.0.26154-0
10. Campos-Garcia, J., Caro, A. D., Najera, R., Miller-Maier, R. M., Al-Tahhan, R. A., Soberon-Chavez, G. (1998). The *Pseudomonas aeruginosa* rhlG gene encodes an NADPH-dependent beta-ketoacyl reductase which is specifically involved in rhamnolipid synthesis. *Journal of Bacteriology*, 180(17), 4442–4451. https://doi.org/10.1128/jb.180.17.4442-4451.1998
11. Quiocho, F. A., Vyas, N. K. (1984). Novel stereospecificity of the L-arabinose-binding protein. *Nature*, 310(5976), 381–386. https://doi.org/10.1038/310381a0
12. Bruner, S. D., Weber, T., Kohli, R. M., Schwarzer, D., Marahiel, M. A., Walsh, C. T., Stubbs, M. T. (2002). Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure*, 10(3), 301–310. https://doi.org/10.1016/s0969-2126(02)00716-5
13. Miller, D. J., Zhang, Y.-M., Rock, C. O., White, S. W. (2006). Structure of RhlG, an Essential β-Ketoacyl Reductase in the Rhamnolipid Biosynthetic Pathway of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry*, 281(26), 18025–18032. https://doi.org/10.1074/jbc.m601687200
14. Samel, S. A., Wagner, B., Marahiel, M. A., Essen, L.-O. (2006). The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-ribosomal Lipopeptide. *Journal of Molecular Biology*, 359(4), 876–889. https://doi.org/10.1016/j.jmb.2006.03.062
15. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., Wilce, M. C. J. (2013). Structural characterization of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Science*, 22(6), 762–773. http://doi.org/10.1002/pro.2262
16. Tufar, P., Rahighi, S., Kraas, F. I., Kirchner, D. K., Lohr, F., Henrich, E., Kopke, J., Dikic, I., ..., Dötsch, V. (2014). Crystal Structure of a PCP/Sfp Complex Reveals the Structural Basis for Carrier Protein Posttranslational Modification. *Chemistry & Biology*, 21(4), 552–562. https://doi.org/10.1016/j.chembiol.2014.02.014
17. Folger, I. B., Frota, N. F., Pistofidis, A., Niquille, D. L., Hansen, D. A., Schmeing, T. M., Hilvert, D. (2024). High-throughput reprogramming of an NRPS condensation domain. *Nature Chemical Biology*, 20(6), 761–769. https://doi.org/10.1038/s41589-023-01532-x
18. Tang, T., Fu, L., Xie, W., Luo, Y., Zhang, Y., Zhang, J., Si, T. (2023). RhlA Exhibits Dual Thioesterase and Acyltransferase Activities during Rhamnolipid Biosynthesis. *ACS Catalysis*, 13(8), 5759–5766. https://doi.org/10.1021/acscatal.3c00046
19. Eberhardt, J., Santos-Martins, D., Tillack, A. F., Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, 61(8), 389103898. http://doi.org/10.1021/acs.jcim.1c00203

ЗАСТОСУВАННЯ ЗВОРОТНЬОГО МОЛЕКУЛЯРНОГО ДОКІНГУ ДЛЯ ІДЕНТИФІКАЦІЇ ПРОТЕЇНОВИХ МІШЕНЕЙ S-ЕТИЛТІОСУЛЬФАНІЛАТУ, ЗАЛУЧЕНИХ ДО СИНТЕЗУ БІОСУРФАКТАНТІВ

Январьов Є.Б., Гавриляк В.В.

Національний університет «Львівська політехніка», Україна

E-mail: yehor.b.yanvarov@lpnu.ua

Мета. Визначити вплив ліганда S-етилтіосульфанілату на протеїнові мішені, що залучені до синтезу біосурфактантів, та оцінити їхню потенційну взаємодію.

Матеріали й методи. Використано метод зворотного докінгу, який дозволяє дослідити взаємодію одного ліганда із 8 протеїнами -мішенями. Молекулярний докінг проводили в *AutoDock Vina* зі скоринговою функцією *Vina*. Підготовку ліганду та протеїнових мішеней проводили за допомогою *AutoDockTools* з пакету *MGLTools*. Візуалізацію результатів здійснювали за допомогою *ChimeraX* та *BIOVA Discovery Studio*.

Результати. Проведення докінгу для ліганда з 8 протеїнами -мішенями дозволило ідентифікувати три перспективні мішені 3RKY, 2B4Q та 8IK2 з афінністю нижчою за −5,5 ккал/моль. Виявлено домінування водневих та гідрофобних зв'язків, що свідчить про стабільність зв'язування ліганда в активних центрах протеїнів.

Висновки. Дослідження підтвердило ефективність зворотного докінгу для пошуку потенційних протеїнових мішеней, демонструючи, що ліганд може впливати на біосинтез біосурфактантів через специфічні взаємодії з протеїнами 3RKY, 2B4Q та 8IK2.

Ключові слова: біосурфактанти, S-етилтіосульфанілат, зворотний молекулярний докінг, ліганд-протеїнова взаємодія, енергія зв'язування, прогнозування мішені.