

ISSN 2220-637X

Вісник
Харківського
Національного
Університету
імені В. Н. Каразіна

СЕРІЯ «ХІМІЯ»
Вип. 43 (66)

Kharkiv University Bulletin
Chemical series. Issue 42 (65)

Заснований 1935 року як
“Труди інституту хемії при
Харківському державному
університеті”

Published since 1935; initially
under the title "Proceedings of
the Institute of Chemistry at
Kharkiv State University"

Харків Kharkiv
2024

УДК 54

Фахове видання з хімічних наук, категорія "Б"

Засновник: Харківський національний університет імені В. Н. Каразіна. Рік заснування: 1935

Публікує статті, присвячені різним аспектам теоретичної хімії, хімічного аналізу, органічної хімії, спектроскопії, фізико-хімії розчинів та поверхневих явищ, електрохімії, хімічного матеріалознавства. Вісник включено до переліку фахових видань, в яких можуть публікуватися основні результати дисертаційних робіт на здобуття наукових ступенів доктора та кандидата хімічних наук. (Наказ Міністерства освіти і науки України No 1643 від 28.12.2019 року).
Періодичність 2 випуски на рік

Затверджено до друку рішенням Вченої ради Харківського національного університету імені В.Н. Каразіна (протокол № 26 від 23 грудня 2024 р.)

Головний редактор

О.І. Коробов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Редактори

А.О. Дорошенко

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

М.О. Мchedlov-Петросян

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна,
член-кореспондент НАН України

Технічний редактор

А.Б. Захаров

к.х.н., ХНУ імені В.Н. Каразіна, Україна

Редакційна рада

В.В. Іванов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.О. Черановський

д.ф.-м.н., проф., ХНУ імені В.Н. Каразіна, Україна

С.А. Шаповалов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

О.І. Юрченко

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Міжнародна консультативна рада

П. Едловскі

професор, Університет Карой Естерхази, Угорщина

А. Ідриссі

професор, Університет Лилля, Франція

О.М. Калугін

к.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

А.Ю. Назаренко

PhD, Prof., Buffalo State College, USA

О.В. Преждо

PhD, Prof., University of Southern California, USA

В.А. Чебанов

д.х.н., проф., НТК «Інститут монокристалів», Україна,
член-кореспондент НАН України

Редактори консультанти

І.М. В'юник

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.І. Ларін

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.І. Лебідь

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Адреса редакційної колегії: Україна, 61022, Харків, майдан Свободи, 4,
ХНУ імені В.Н. Каразіна, хімічний факультет; тел.: +38 057 707 51 29.
E-mail: chembull@karazin.ua a.korobov@karazin.ua

Статті пройшли внутрішнє та зовнішнє рецензування.

Ідентифікатор медіа у Реєстрі суб'єктів у сфері медіа: R30-04479 (Рішення № 1538 від 09.05.2024 р. Національної ради України з питань телебачення і радіомовлення. Протокол № 15)

© Харківський національний університет
імені В.Н. Каразіна, оформлення, 2024

UDK 54

Professional publication in chemical sciences, category "B"

Founder: V. N. Karazin Kharkiv National University. Year of foundation: 1935

Publishes papers devoted to various aspects of theoretical chemistry, chemical analysis, organic chemistry, inorganic chemistry, physical chemistry of solutions and surface phenomena, electrochemistry, materials chemistry. The bulletin is officially authorized by the Highest Attestation Commission of Ukraine to publish results of research submitted for PhD and ScD degrees. (Order of the Ministry of Education and Science of Ukraine No. 1643 of December 28, 2019). 2 issues per year

Publication of this issue is approved by the Academic Council of V.N. Karazin Kharkiv National University (protocol № 26 from 23.12.2024).

Editor-in-chief

Alexander Korobov

ScD, Prof., V.N. Karazin Kharkiv National University

Editors

Andrey Doroshenko

ScD, Prof., V.N. Karazin Kharkiv National University

Nikolay Mchedlov-Petrosyan

ScD, Prof., V.N. Karazin Kharkiv National University,
corresponding member of NAS of Ukraine

Managing editor

Anton Zakharov

PhD, V.N. Karazin Kharkiv National University

Editorial board

Vladimir Ivanov

ScD, Prof., V.N. Karazin Kharkiv National University

Vladislav Cheranovsky

ScD, Prof., V.N. Karazin Kharkiv National University

Sergey Shapovalov

ScD, Prof., V.N. Karazin Kharkiv National University

Oleg Yurchenko

ScD, Prof., V.N. Karazin Kharkiv National University

International advisory board

Pál Jedlovsky

Dr. Prof., Eszterházy Károly University, Hungary

Abdenacer Idrissi

Prof., Université Lille, France

Oleg Kalugin

PhD, Prof., V.N. Karazin Kharkiv National University

Alexander Nazarenko

PhD, Prof., Buffalo State College, USA

Oleg Prezhdo

PhD, Prof., University of Southern California, USA

Valentyn Chebanov

ScD, Prof., SSI "Institute for Single Crystals", Ukraine
corresponding member of NAS of Ukraine

Consulting editors

Ivan Vyunnik

ScD, Prof., V.N. Karazin Kharkiv National University

Vasyl Larin

ScD, Prof., V.N. Karazin Kharkiv National University

Valentyn Lebed

ScD, Prof., V.N. Karazin Kharkiv National University

Address of editorial team: 4 Svobody Sq., Kharkiv, 61022, Ukraine, V.N. Karazin Kharkiv National University, School of Chemistry; tel.: +38 057 707 51 29.

E-mail: chembull@karazin.ua a.korobov@karazin.ua

All articles have been reviewed.

Media identifier in the Register of the field of Media Entities: R30-04479 (Decision № 1538 dated May 9, 2024 of the National Council of Television and Radio Broadcasting of Ukraine, Protocol № 15)

© V.N. Karazin Kharkiv National University,
Design, 2024

ЗМІСТ

- 6 Порівняння ефективності прогнозування 3D-структури білка з використанням AlphaFold 3 від Google DeepMind. **Є. Дума, О. Кириченко.**
- 26 Еволюційна оптимізація будови енсирелвіру як нековалентного інгібітора основної протеази SARS-COV-2 M^{pro}. **К.О. Логачова, А. С. Святенко, О. В. Кириченко, О. М. Калугін.**
- 38 Структура та транспортні властивості розчинів NaPF₆ у сумішах етиленкарбонату з диметилкарбонатом для натрій-іонних акумуляторів: МД моделювання. **А.В. Боровик, Я.В. Колесник, О.М. Калугін.**
- 48 Методи захисту/зняття захисту гідроксильних груп при синтезі полігідроксифлавонолів. **О. Демидов, О. Рошаль.**
- 56 Синтез моно- та дифеніл заміщених саліцилових альдегідів, важливих білдинг-блоків для синтезу флуоресцентних барвників та барвників-сенсibilізаторів для DSSC. **П. В. Тросянко, В. В. Пашко, С. М. Коваленко.**

CONTENTS

- 6 Benchmarking Google DeepMind's AlphaFold 3 Performance for Protein 3D-Structure Prediction. **Y. Duma, A. Kyrychenko.**
- 26 Evolutionary Structure Optimization of Ensitrelvir as Non-Covalent Inhibitor of SARS-CoV-2 Main Protease M^{pro}. **K. O. Lohachova, A. S. Sviatenko, A. V. Kyrychenko, O. N. Kalugin.**
- 38 Structure and Transport Properties of NaPF₆ Solutions in Mixtures of Ethylene Carbonate with Dimethyl Carbonate for Sodium-Ion Batteries: MD Simulation. **A.V. Borovyk, Y.V. Kolesnik, O.M. Kalugin.**
- 48 Methods of Protection/Deprotection of Hydroxy Groups in the Synthesis of Polyhydroxy Flavonols. **O. Demidov, A. Roshal.**
- 56 Synthesis of Mono- and Diphenyl Substituted Salicylic Aldehydes, Important Building Blocks for the Synthesis of Fluorescent Dyes and Dyes-Sensitizers for DSSC. **P. V. Trostianko, V. V. Pashko, S. M. Kovalenko.**

BENCHMARKING GOOGLE DEEPMIND'S AlphaFold 3 PERFORMANCE FOR PROTEIN 3D-STRUCTURE PREDICTION

Y. Duma^a, A. Kyrychenko^b

V. N. Karazin Kharkiv National University, School of Chemistry, 4 Svobody sq., Kharkiv, 61022 Ukraine

a) ✉ yelyzaveta.duma@student.karazin.ua
b) ✉ a.v.kyrychenko@karazin.ua

 <https://orcid.org/0009-0004-0873-6490>
 <https://orcid.org/0000-0002-6223-0990>

The 3D structure of proteins is directly linked to their function, making its determination crucial for understanding biological processes and addressing issues related to human health and life sciences. Despite the continuous experimental acquisition of new protein structures, there remains a significant gap between the number of protein sequences available and those that have an established experimental high-resolution tertiary structure. Several computational approaches have focused on predicting protein structures using either templates or empirical force field modeling. In recent years, various methods have been combined to address the individual limitations of these approaches, leading to the development of AlphaFold 3 (AF3) by Google DeepMind. AF3 enables prediction of 3D protein structures with high accuracy based on its amino acid sequence. In this study, we benchmarked applicability, performance, and limitations of AF3 for predicting 3D structure of a broad series of proteins, including SARS-CoV-2 coronavirus proteins, other bacterial and viral proteins, as well as some plant enzymes. We found that AlphaFold 3 could capture the overall backbone features of the most examined proteins in terms of small deviation from available X-ray structures. Some minor miss-folding of N- and C-terminal segments were common, which, often, did not affect biological roles of the studied proteins. In cases involving protein dimers or higher-order oligomers, there are notable differences between the predicted AF3 models of a single-chain monomer and their corresponding experimental structures. These discrepancies are particularly evident in regions related to protein dimerization, assembly, and binding interfaces. Ultimately, while capturing the overall fold, predicting the complex structure of the Spike glycoprotein is still beyond the current capabilities of AF3.

Keywords: protein folding, enzyme, 3D structure, drug design, artificial intelligence, AlphaFold 3.

Introduction

Tertiary protein structures are typically determined using techniques, such as X-ray crystallography, electron cryomicroscopy (cryo-EM), and nuclear magnetic resonance (NMR). However, these methods are complex, time-consuming, and expensive, and they often do not capture the protein in its native form. Given these challenges, it is not surprising that the number of proteins with resolved tertiary structures is relatively small — only ~202000 entries in the Protein Data Bank (PDB) — compared to the vast number of known protein sequences, which stands at 240 million entries in UniProt DataBase. This disparity presents an open challenge, highlighting the need for innovative approaches in protein structure prediction, particularly methods that utilize computational models to predict 3D protein structures from polypeptide sequences.

Several algorithms and web servers have been developed to enhance protein structure prediction. The significance of these efforts is highlighted by their impact on various applications, including rational drug design, mutational studies, structural comparisons, evolutionary analysis, and folding studies. However, despite significant advancements, leading to gradual improvements in accuracy, the fundamental problem of predicting protein structures remained largely unsolved for decades. Since the introduction of AlphaFold by Google DeepMind in 2018 and AlphaFold 1 in 2019, these systems have rapidly outperformed other computational methods in predicting protein structures. Following this, AlphaFold 2 was released in 2020 and revolutionized the field of protein structure prediction. The latest version, AlphaFold 3, was presented in May 2024 and represents a significant advancement. The major improvements in AlphaFold 3 stem from a substantial evolution in its network architecture and training procedures, allowing it to represent a wider range of chemical structures and enhancing the efficiency of learning from available data. Additionally, AlphaFold 3 incorporates a diffusion model to predict the coordinates of each atom, replacing the structure module used in AI-

phaFold 2, which operated at the amino acid level. So, recent studies have demonstrated that AF3 can achieve experimental accuracy in an increasing number of cases. Finally, considering all these efforts, AlphaFold's developers, David Baker, Demis Hassabis and John Jumper, have been awarded the Nobel Prize for Chemistry 2024 for their work on computational protein design and protein structure prediction.

In this study, we benchmarked AlphaFold 3's applicability and limitations for predicting 3D tertiary structure of a broad series of proteins, including SARS-CoV-2 virus enzymes, bacterial and viral proteins, as well as some plant enzymes.

Results and Discussion

To benchmark performance of AF3 for protein structure prediction we selected diverse families of functional proteins and enzymes, such as SARS-CoV-2 proteins, some proteins from *Mycobacterium tuberculosis*, as well as other crucial enzymes. Selection criteria were based on biological importance and availability of experimental high-resolution 3D structures. The selected proteins are composed of various conservative secondary structures, such as α -helices and β -pleated sheets, offering some computational challenge for their structure evaluations.

Key proteins of the SARS-CoV-2 coronavirus

Coronavirus disease 2019 (COVID-19) is caused by a novel member of the *Coronaviridae* family known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The genome of this virus contains both structural and non-structural proteins (NSps). The structural proteins include Spike (S), membrane (M), hemagglutinin-esterase (H), and envelope (E) proteins, while the NSPs consist of accessory proteins and replicase proteins (Fig. 1).

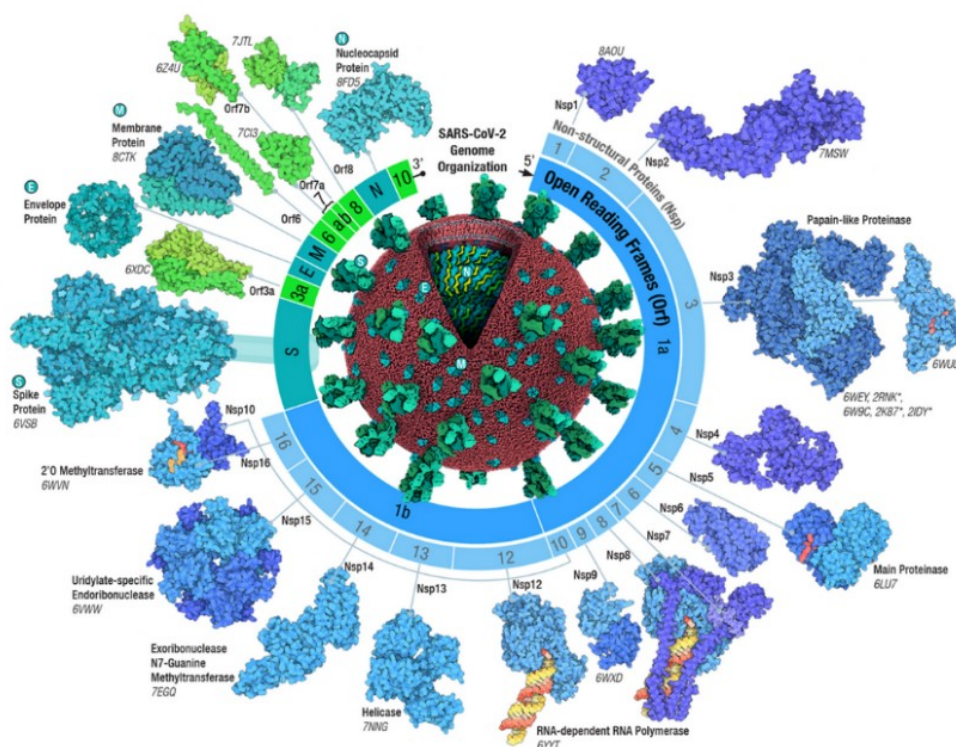


Figure 1. Architecture of the SARS-CoV-2 genome and proteome, demonstrating key enzymes of the coronavirus. Adopted with permission from . Copyright © 2021 Wiley Periodicals LLC.

The structural and Nsp components of SARS-CoV-2 are crucial for its infectivity, and some may play significant roles in the development of chronic diseases such as cancer, coagulation disorders, neurodegenerative disorders, and cardiovascular diseases. Additionally, the proteins of SARS-CoV-2 interact with targets like the angiotensin-converting enzyme 2 (ACE2) receptor.

A structural comparison of several essential SARS-CoV-2 proteins, including ten variants of the Spike protein and cysteine proteases, was recently conducted using an earlier version of AlphaFold 2 . In this study, some of these structures were reanalyzed using the newer version, AlphaFold 3.

The first set of the studied proteins were selected from SARS-CoV-2 coronavirus' lifecycle. SARS-CoV-2 has a genome that consists of a long RNA strand, making it one of the largest among all RNA viruses. When the virus infects a cell, this genome acts like messenger RNA (mRNA) and directs the synthesis of long polyproteins (Fig. 1). These polyproteins contain the components necessary for the virus to replicate and produce new viruses. The proteins include a replication/transcription complex that generates more RNA, several structural proteins that assemble new virions, and two cysteine proteases. These proteases are crucial as they cut the polyproteins into the functional pieces needed for the virus's lifecycle. In this set, the main protease M^{pro} (PDB: 6LU7), papain-like protease PL^{pro} (PDB: 7SGW), Nsp12-Nsp7-Nsp8 complex of non-structural proteins (Nsp) (PDB: 7BV1) were considered (Fig. 2). Additionally, Spike protein of recent Omicron variant of the virus is also considered.

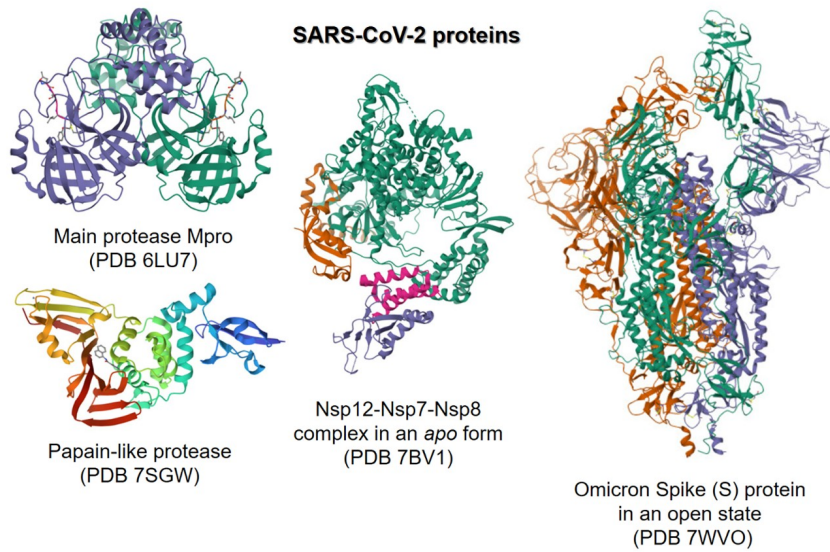


Figure 2. The X-ray or cryo-EM structures of some essential proteins of the SARS-CoV-2 virus. The corresponding PDB codes are given in brackets.

Some recent studies utilizing X-ray crystallography cryo-electron microscopy have reported the high-resolution structure of some essential SARS-CoV-2 proteases, such as main and papain-like proteases, Nsp complex as well as Spike-proteins, as summarized in Fig. 2. Therefore, these selected proteins of the SARS-CoV-2 family were subjected to the AF3 evaluation using their FASTA sequences (Table 1).

Table 1. Amino acid sequences of the M^{pro} and PL^{pro} proteases used for AF3-prediction.

Protein, PDB Code	FASTA Sequence
Main pro- tease M ^{pro} , 6LU7	SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDML- NPNY EDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTP- KYKFVRIQPGQ TFSVLACYNGSPSGVYQCAMRPNFTIKGS- FLNGSCGSVGFNIDYDCVSFCYMHM ELPTGVHAGTDLEGNFYGPFVDRQ- TAQAAGTDTTITVNLAWLYAAVINGDRWFLN RFTTTLNDFNL- VAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMN GRTIL- GSALLEDEFTPFDDVVRQCSGVTFQ
Papain-like protease PL ^{pro} ,	SNAEVRTIKVFTTVDNINLHTQVVDMSTYQQFGPTYLDGADVTKIKPHN- SHEGK TFYVLPNDDTLRVEAFEYYHTTDP SFLGRYM- SALNHTKKWKYPQVNGLT SIKWADN NCYLATALLTLQQIELKFNPALQ-

7SGW	DAYYRARAGEAANFCALILAYCNKTVGELGDVR ETMSYLFQHANLDSCK- RVLNVVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQI PCTCGKQATKYLQQESPFVMMMSAPPAQYELKHGTFTCASEYT- GNYQCGHYKHITS KETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIK
------	--

Main protease M^{pro} . The main protease, known as M^{pro} (a 3C-like protease), is responsible for cleaving eleven specific sites on the two SARS-CoV-2 polyproteins. M^{pro} is comprised of three domains: a chymotrypsin-like domain (Domain I), a 3C-protease-like domain (Domain II), and a third domain (Domain III) that contains five α -helices. The binding pocket of M^{pro} is situated between Domains I and II and features a catalytic Cys-His dyad, consisting of Cys145 and His41. Therefore, M^{pro} has attracted essential attention and is considered an ideal target for developing antiviral agents.

Fig. 3 shows the overlap of the AF3-predicted structure (*green*) of the M^{pro} protease with its native X-ray structure (PDB 6LU7, *blue*) in two different views. We estimate quantitative measure of AF3 performance as the C_α root mean square deviation (C_α RMSD) between the predicted AF3 model and experimental X-ray structure for entire protein, calculated by PyMOL tool *super*. The AF3 predicted models show an average global C_α RMSD of 0.411 Å from the experimental structures, suggesting that the global structure properties of M^{pro} have been captured by AlphaFold3. In addition, the RMSD of main chain atoms between the predicted AF3 model and experimental X-ray structure for M^{pro} were also calculated by Discovery Studio software and found to be 2.796 Å, as summarized in Table 2. Some key metrics of AF3 confidence measures for its structural predictions are also given in Table 2.

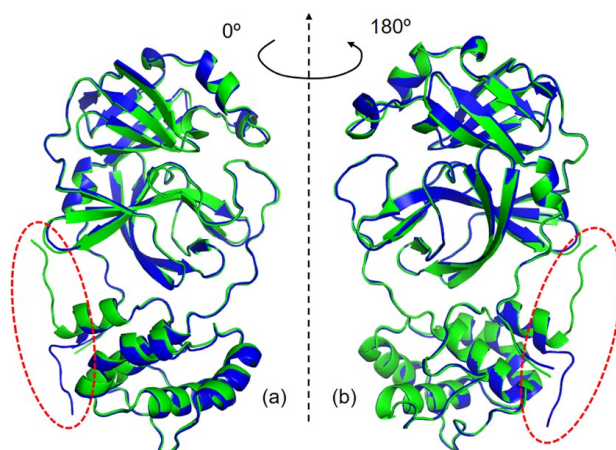


Figure 3. Overlay of the native X-ray structure of the M^{pro} protease (PDB 6LU7, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Some disagreement between the AF3-predicted model and the experimental structure are highlighted in *red* circles.

Papain-like protease PL^{pro} . In SARS-CoV-2, the Nsp3 protein consists of 1,945 residues (approximately 212 kDa), and PL^{pro} is a domain of this large, multidomain protein that plays a crucial role in the replication-transcription complex (RTC). PL^{pro} is located in Nsp3, positioned between the SARS unique domain (SUD/HVR) and a nucleic acid-binding domain (NAB). This enzyme is highly conserved and is found in all coronaviruses.

PL^{pro} exhibits multiple functions, including proteolytic activity, processing of Pp1a/Pp1ab, and deubiquitination, which involves efficiently disassembling mono-ubiquitin, polyubiquitin chains, and branched polyubiquitin chains. It features a canonical cysteine protease catalytic triad, making it a target for anti-COVID-19 drug development.

Fig. 4 demonstrates comparison of the AF3-predicted structure (*green*) of the PL^{pro} protease with its native X-ray structure (PDB 7SGW, *blue*) in two different views. The RMSD between the predicted AF3 model and experimental X-ray structure for PL^{pro} was found to be 0.456 Å (PyMOL) and

0.522 Å (Discover Studio), suggesting that AlphaFold3 maintains its excellent performance in predicting overall 3D structures of monomeric single-chain proteins.

Table 2. Comparison of AF3-predicted models for the studied proteins, focusing on AF3 confidence metrics and their deviation from experimental results.

Protein, PDB	AF3 confidence metrics			RMSD (Å)	
	Chain PTM ¹	Chain pair PAE min ²	Ranking Score ³	C _α (PyMOL)	Main chain (Discovery Studio)
Main protease M ^{pro} , 6LU7	0.93	0.76	0.93	0.411	2.796
Papain-like protease PL ^{pro} , 7SGW	0.92	0.76	0.92	0.456	0.522
Nsp12-Nsp7-Nsp8 complex, 7BV1	0.94	0.76	0.95	0.614	1.303
Omicron Spike (S) protein, 7WVO	0.76	0.76	0.81	2.046	9.488
Oxidoreductase DprE1, 4FF6	0.90	0.76	0.94	0.236	1.355
Dihydrofolate reductase DHFR, 1DF7	0.94	0.76	0.94	0.408	0.486
Pantothenate kinase PanK, 4BFT	0.86	0.76	0.89	0.629	2.560
Enoyl-ACP reductase InhA, 4TRO	0.95	0.76	0.95	0.385	1.421
Sirtuin SIRT1, 4I5I	0.83	0.76	0.87	1.279	3.301
Sirtuin SIRT2, 4RMG	0.92	0.76	0.92	0.586	1.938
Diphtheria toxin, 7K7E	0.89	0.76	0.89	0.792	16.034
<i>Paenibacillus polymyxa</i> β-Glucosidase B, 2O9R	0.96	0.76	0.97	0.371	0.736

¹ – A per-chain pTM score, which suggests confidence for the individual chain structure.

² – Minimum predicted aligned error (PAE) for chain pairs, indicating reliability of structural alignment.

³ – A general confidence score, which summarizes different predictions.

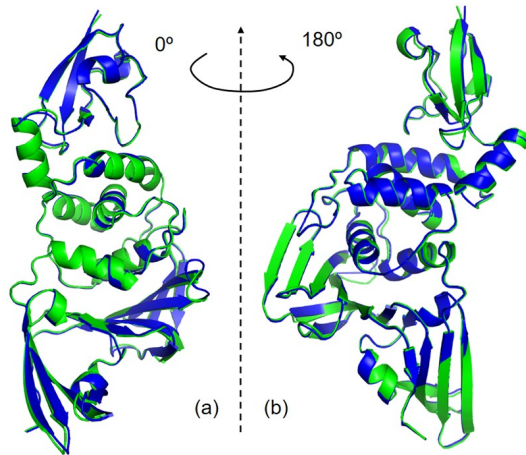


Figure 4. Overlay of the native Cryo-EM structure (PDB 7SGW, blue) of the PL^{pro} protease and the AF3-predicted structure (green) in a side (a) and 180°-rotated (b) views, respectively.

Nsp12-Nsp7-Nsp8 complex. The genome organization of SARS-CoV-2 comprises sixteen non-structural proteins, Nsp1–16, which are highly conserved among coronaviruses (Fig. 1) . Non-structural proteins from this family present essential challenge for 3D structure prediction.

Fig. 5 illustrates the AF3-predicted model of the Nsp12-Nsp7-Nsp8 complex, which is part of the RNA-dependent RNA polymerase (RdRp) enzyme of SARS-CoV-2. When compared to the experimental Cryo-EM structure, which was measured at a resolution of 2.8 Å, the AF3 model displayed a more compact fold, with two terminal protein segments predicted to differ in several aspects. While the C_α RMSD measured using PyMOL was relatively small at 0.614 Å, the main chain RMSD showed a deviation of 1.303 Å (Table 2). This suggests that predicting the structures of viral nonstructural proteins remains a significant computational challenge.

Spike (S) protein. The SARS-CoV-2 virus is covered by a large number of glycosylated Spike (S) proteins, which bind to the host cell receptor known as angiotensin-converting enzyme 2 (ACE2). This binding mediates the entry of the virus into host cells (see Fig. 1) . The S protein is crucial for receptor recognition and the process of cell membrane fusion, consisting of two subunits: S1 and S2. The S1 subunit contains a receptor-binding domain that recognizes and attaches to ACE2, while the S2 subunit facilitates the fusion of the viral and host cell membranes by forming a six-helical bundle through its two-heptad repeat domain . When the S protein binds to the receptor, it is activated by the host cell's TM protease serine 2 (TMPRSS2), a type 2 transmembrane serine protease. This activation promotes the entry of the virus into the cell. Once inside, the viral RNA is released, and the cell translates polyproteins from the RNA genome. The replication and transcription of the viral RNA genome occur via protein cleavage and the assembly of the replicase–transcriptase complex. Consequently, the design of antibodies or small-molecule entry blockers targeting S proteins is highly significant for preventing the virus and treating COVID-19 .

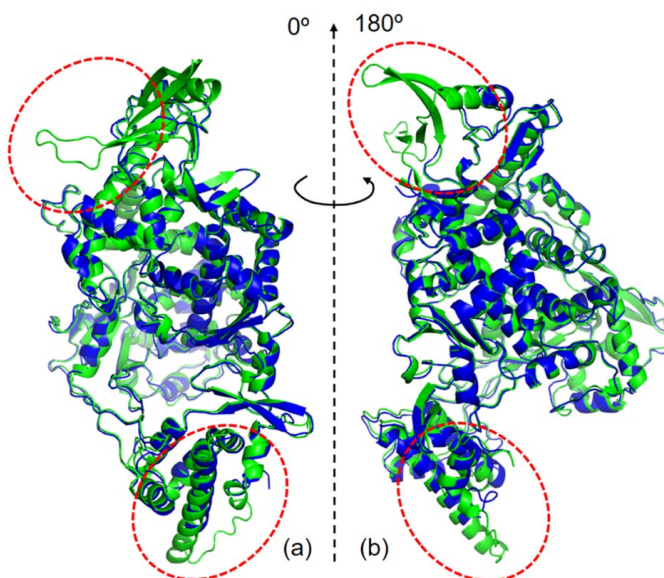


Figure 5. Overlay of the native Cryo-EM structure of the Nsp12-Nsp7-Nsp8 RdRp complex in an *apo*-form (PDB 7BV1, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted circles.

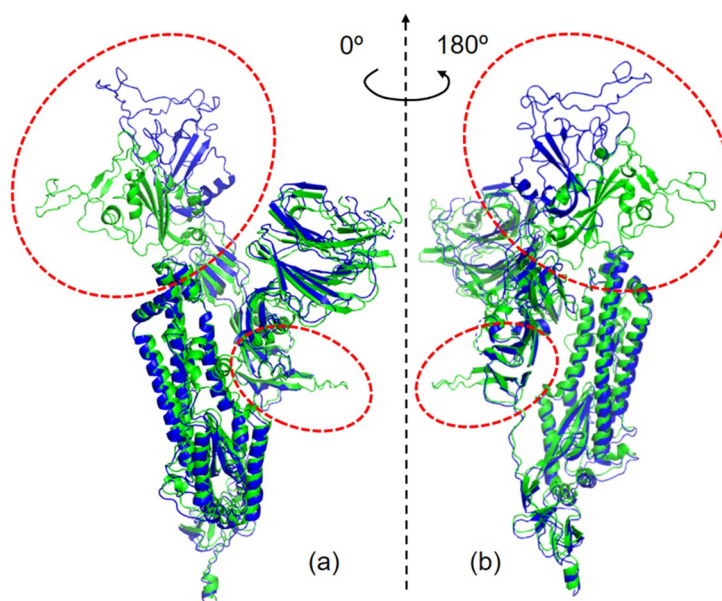


Figure 6. Overlay of the native X-ray structure of the Omicron Spike (S) protein in an open state (PDB 7WVO, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Considering the complex structure of the S-protein, which consists of 1,258 amino acids (Fig. 2), this case represents one of the most challenging tests for AF3 to date. Figure 6 illustrates the model of the S-protein predicted by AF3 and compares it with the corresponding cryo-EM structure. It is evident that AF3 successfully captured most secondary structure features, such as α -helices and β -sheets. However, some crucial segments of the S-protein, particularly the receptor-binding domain (RBD) that plays a central role in ACE2 recognition, were predicted with unexpectedly low accuracy (Table 2). Therefore, we can conclude that predicting the structure of the Spike glycoprotein remains beyond the capabilities of AF3 at this time.

Essential proteins from *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (Mtb), the bacterium responsible for tuberculosis (TB), is a resilient pathogen that has latently infected about one-third of the world's population. However, traditional TB treatment regimens are increasingly inadequate in addressing the rising threat of drug resistance. This situation has prompted the development of innovative anti-tuberculosis agents, particularly those targeting new proteins. The Mtb genome contains approximately 4,000 predicted proteins, many of which are enzymes involved in various cellular metabolic processes. Notably, over 200 of these proteins play a role in fatty acid biosynthesis. This process is crucial for building the cell envelope and is closely linked to the pathogenesis and drug resistance of mycobacteria.

Therefore, the second set of the proteins were selected from *Mycobacterium tuberculosis* (Mtb) family, as summarized in Fig. 7.

Oxidoreductase DprE1. Decaprenylphosphoryl- β -D-ribose 2'-oxidoreductase (DprE1) is a highly conserved enzyme that contains FAD. Together with its downstream NADH-dependent reductase, DprE2, it facilitates the conversion of decaprenylphosphoryl- β -D-ribofuranose (DPR) into decaprenylphosphoryl- β -D-arabinofuranose (DPA). DprE1 is a promising target for the development of new anti-tubercular drugs, and researchers are actively investigating various inhibitors of DprE1.

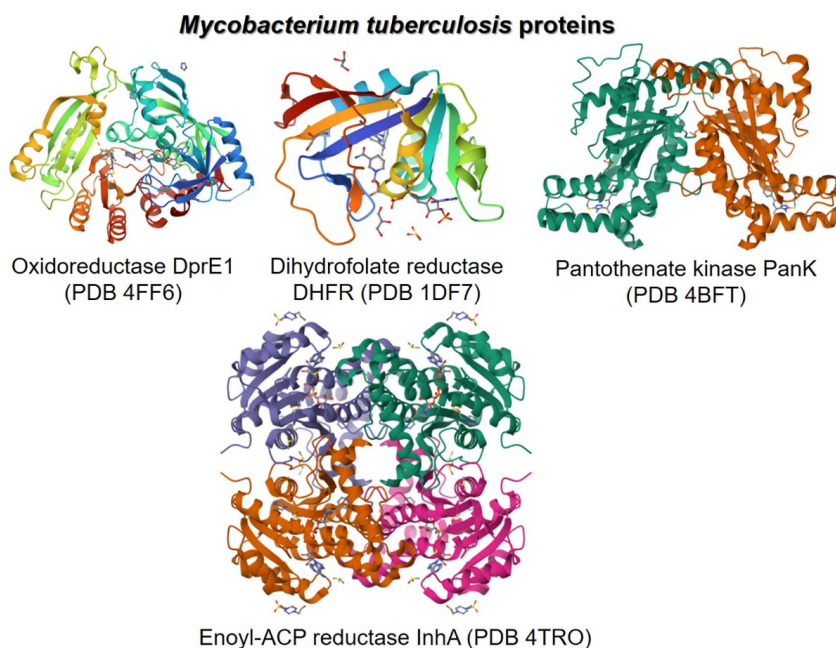


Figure 7. The X-ray or cryo-EM structures of some essential proteins of the *Mycobacterium tuberculosis*. The corresponding PDB codes are given in brackets.

Fig. 8 shows the AF3-predicted model of DprE1 enzyme that revealed the very good agreement with the corresponding X-ray structure. The PyMOL C_α RMS deviation was found to be 0.236 Å only. Some larger RMSD of 1.355 Å for the main chain DprE1 is because of contribution of non-structured loops (Fig. 8a-b). It should also be noted that the AF3-predicted model was able to reconstruct a large protein segment of 269-283, missed in the original X-ray structure.

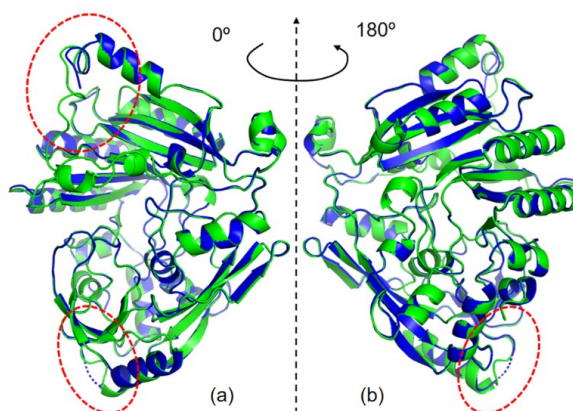


Figure 8. Overlay of the native X-ray structure of the oxidoreductase DprE1 measured with resolution of 2.6 Å (PDB 4FF6, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Some disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Dihydrofolate reductase (DHFR). The dihydrofolate reductase (DHFR) enzyme catalyzes the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate. This reaction produces a precursor for cofactors necessary for synthesizing thymidylate, purine nucleotides, methionine, serine, and glycine, which are required for DNA, RNA, and protein synthesis. Therefore, DHFR is a promising target for developing novel antitubercular agents.

Fig. 9 demonstrates the strong correlation between the AF3-predicted model of DHFR and the available X-ray structure. The calculated RMSDs of 0.408 and 0.486 Å rank DHFR among the best predicted models analyzed in Table 2.

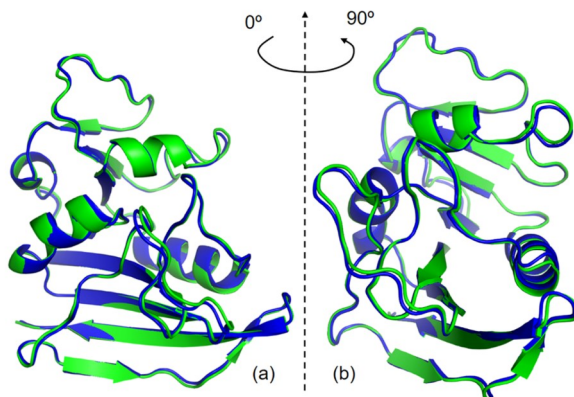


Figure 9. Overlay of the native X-ray structure of the dihydrofolate reductase DHFR (PDB 1DF7, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 90°-rotated (b) views, respectively.

Pantothenate kinase PanK. Pantothenate kinase (PanK) is responsible for the phosphorylation of pantothenate, which is the first committed and rate-limiting step in the biosynthesis of coenzyme A (CoA). Research has shown that the type I isoform, encoded by the *coaA* gene, is an essential pantothenate kinase in *Mycobacterium tuberculosis*. This crucial information has been utilized to screen large libraries for the identification of different classes of PanK inhibitors that operate through various mechanisms.

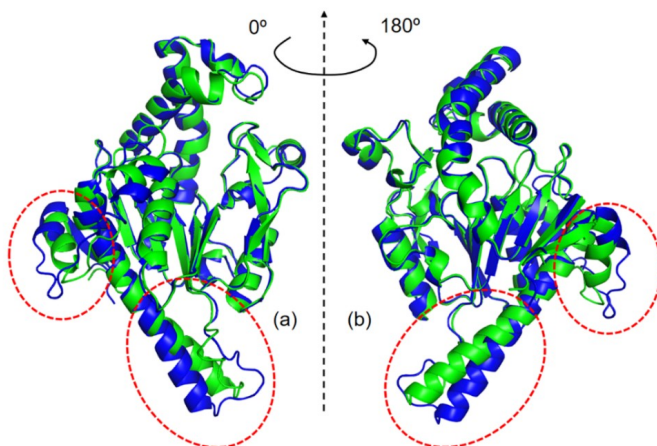


Figure 10. Overlay of the native X-ray structure of the Pantothenate kinase PanK measured with resolution of 2.29 Å (PDB – 4BFT), *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Fig. 10 illustrates the strong correlation between the AF3-predicted model of PanK and the available X-ray structure in most segments, with the exception of α -helix 4. In the X-ray structure, α -helix 4 is crucial for binding between PanK chains A and B, as it forms a helix-helix dimer (see Figure 7). Consequently, the proper conformation of α -helix 4 is maintained in the PanK dimer, which is why this protein segment is not accurately predicted single-chain PanK by AF3. The calculated $C\alpha$ RMSD is 0.629 Å, while the main chain RMSD of 2.560 Å indicates this deviation (refer to Table 2).

Enoyl-acyl carrier protein reductase InhA. Enoyl-ACP reductase (InhA) is a vital biosynthetic enzyme involved in the production of mycolic acids (MAs). It catalyzes the NADH-dependent reduction of long-chain trans-2-enoyl-ACP in the type II fatty acid synthesis of *Mycobacterium tuberculosis* (*Mtb*). Importantly, *Mtb* InhA does not have a human counterpart, which may reduce the risk of toxicity from inhibitors targeting this enzyme. As a result, InhA has been established as a reliable

target for the design and discovery of novel agents against *Mtb*, particularly for frontline and second-line antitubercular drugs such as isoniazid and ethionamide .

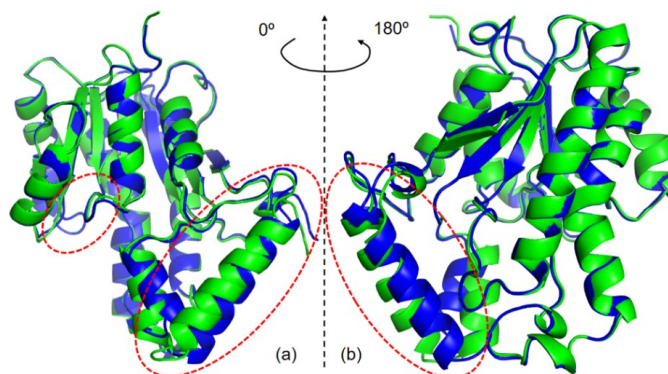


Figure 11. Overlay of the native X-ray structure of the Enoyl-ACP reductase InhA measured with resolution of 1.40 Å (PDB – 4TRO), *blue* and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Fig. 11 illustrates the AF3-predicted model of InhA, which resemble some prediction features of the above mentioned PanK. In particular, AF3 performs well in most protein segments, with the exception of α -helices 6 and 7. These α -helices are crucial for assembling the quaternary structure of InhA composed of four chains A-C, as shown in Fig. 7. Consequently, this case further demonstrate some limitation of AF3 in reproducing 3D-structure of a single-chain protein when it is a part of complex quaternary structure, such as dimer or tetramer. The AF3 confidence score of 0.95 for InhA is also among the best, as seen in Table 2.

Results of calculations

In our recent studies, we also considered several important enzymes, which are now selected for benchmarking using AlphaFold3. These enzymes are promising targets for drug design and development, making the prediction of their structure and function critically important. Among them are silent information regulators, a plant hydrolase, and the full-length diphtheria toxin (Fig. 12).

Sirtuins. Silent Information Regulator Two Homologue One (SIRT1) is the most extensively studied member of the Sirtuin family, which is related to the class III histone deacetylases and utilizes NAD^+ as its cofactor. SIRT1 is one of seven human Sirtuins, which are conserved from archaea to mammals . These proteins play key roles in regulating transcription, genome stability, longevity, and metabolism . SIRT1 influences transcription by deacetylating various transcription factors, including PPAR γ , NF κ B, and the tumor suppressor protein p53. Research indicates that SIRT1, along with its histone and non-histone targets, is involved in a wide range of biological processes. These include metabolic diseases, age-related conditions, viral infections, inflammation, and the growth and metastasis of tumor cells. Consequently, modulating SIRT1 expression could offer new insights into the development of therapeutics, whether they are derived from natural sources or synthetic origins .

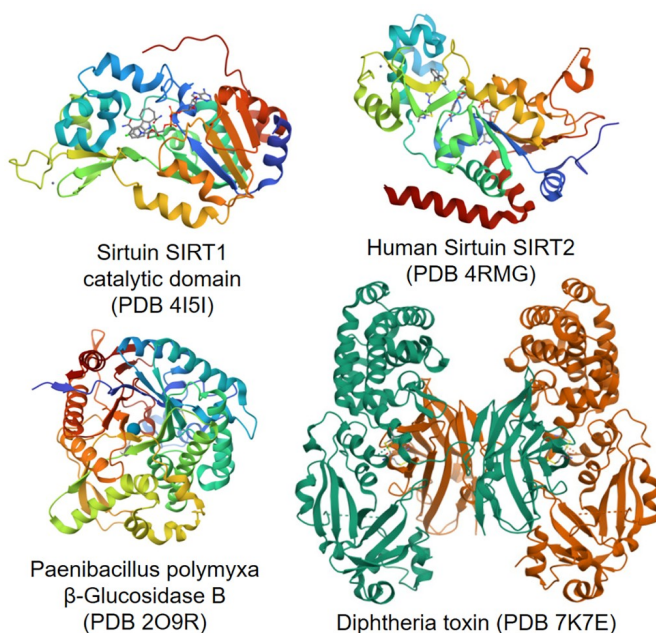


Figure 12. The X-ray or cryo-electron microscopy structures of SIRT1, SIRT2, diphtheria toxin and β -glucosidase, respectively. The corresponding PDB codes are given in brackets.

Despite the relatively simple folding of SIRT1's catalytic domain, predicting its structure with AF3 faced several systematic challenges. Fig. 13 compares the structures predicted by AF3 with those obtained through X-ray crystallography. However, there are significant deviations and inconsistencies between the two.

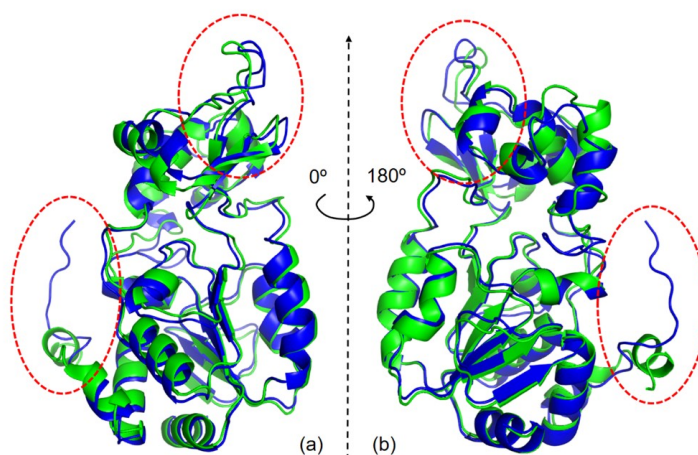


Figure 13. Overlay of the native X-ray structure of the catalytic domain of Sirtuin SIRT1 (PDB 4I5I, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Most of these discrepancies occur in the flexible, unstructured loops that connect various secondary structural elements, such as α -helices. So, SIRT1 is characterized by the largest RMSD of 1.279 Å and 3.301 Å observed among all the studied proteins, as detailed in Table 2.

AF3 faced similar systematic challenges when predicting the 3D structure of human Sirtuin SIRT2, as summarized in Fig. 14. Several unstructured loops were not accurately reproduced. Additionally, the experimental X-ray structure omitted some sequence segments, specifically 47-51 and 379-381. As a result, the reconstruction of these unstructured segments by AF3 produced an artificial prediction of α -helices.

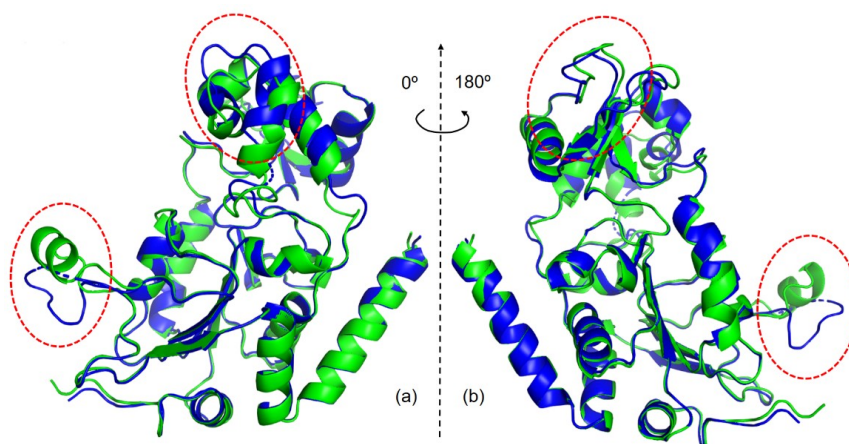


Figure 14. Overlay of the native X-ray structure of the human Sirtuin SIRT2 (PDB 4RMG, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

Diphtheria toxin. Diphtheria toxin is an exotoxin produced by *Corynebacterium* that causes disease in humans by inhibiting protein synthesis. It enters the cell through a process called receptor-mediated endocytosis. Once inside, the acidification of the endosome triggers a series of conformational changes. This results in the refolding and insertion of the translocation (T) domain into the membrane, ultimately allowing the catalytic domain to be transported into the cytoplasm. Additionally, the T-domain has become a convenient model for studying pH-triggered protein-membrane interactions.

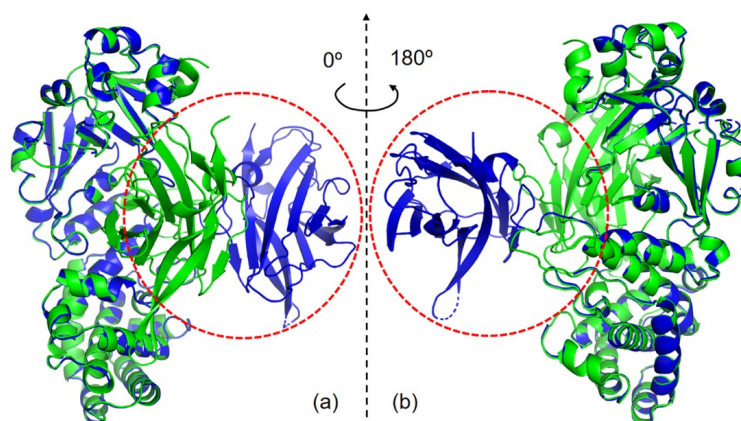


Figure 15. Overlay of the native X-ray structure of the diphtheria toxin (PDB 7K7E, *blue*) and the AF3-predicted structure (*green*) in a side (a) and 180°-rotated (b) views, respectively. Essential disagreements between the AF3-predicted model and the experimental structure are highlighted in *red* dotted areas.

In its crystalline form, the diphtheria toxin exists as a homodimer, with its quaternary structure maintained by strongly bound interfacial segments (see Fig. 15). Predicting the 3D structure of a single chain monomer of the toxin presented expected challenges with AF3, as illustrated in Fig. 15. AF3 predicts a compact fold for the toxin; however, the C-terminal domain in its native conformation unfolds towards another part of the homodimer (see Fig. 15). All these deviations are highlighted in Table 2.

***Paenibacillus polymyxa* β -Glucosidase B.** Bacterial species that degrade cellulosic substrates produce various enzymes. *Paenibacillus polymyxa* encodes β -glucosidase B (BglB), a monomeric enzyme that functions as an exo-beta-glucosidase, hydrolyzing cellobiose and higher-degree cellodextrins. It is a 1,4- β -glucosidase. β -Glucosidase is an important target in developing substrate-specific

inhibitors, passive binders and studying protein-ligand interactions . The X-ray structure of BglB (2O9R) is available making it possible to do direct comparison to the AF3 prediction.

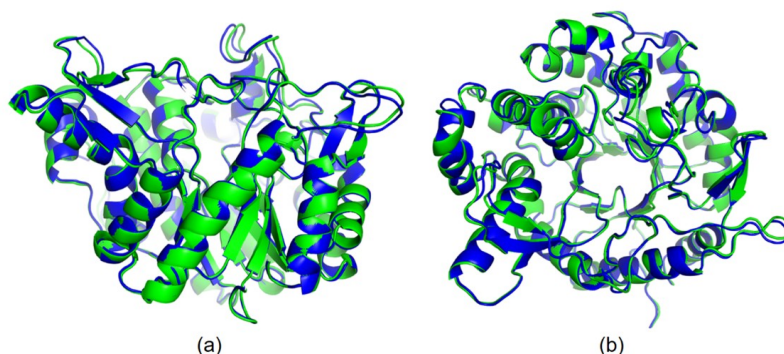


Figure 16. Overlay of the native X-ray structure of *Paenibacillus polymyxa* β -Glucosidase B (PDB 2O9R, *blue*) and the AF3-predicted structure (*green*) in a side (a) and top (b) views, respectively.

Fig. 16 demonstrates the impressive performance of AF3 in predicting the 3D structure of the water-soluble single-chain enzyme, such as β -glucosidase (BglB). This enzyme consists of a bundle of nine α -helices that encircle a deeply buried catalytic site. AF3 successfully predicted the 3D structure of BglB with exceptional accuracy, as indicated by a small C_{α} RMSD (Root Mean Square Deviation) of 0.371 Å. Even with the presence of numerous connecting loops, AF3 surprisingly reproduced their structure well, achieving a main chain RMSD of only 0.736 Å (Table 2). The AF3 confidence score of 0.97 for this protein is also among the best, as seen in Table 2.

Conclusion

Computational prediction of protein structure from its amino acid sequences has been a central problem in molecular biology since the 1960s, when it was demonstrated that a protein's amino acid sequence determines its 3D-structure . Beyond its theoretical significance in understanding the fundamental principles governing biological macromolecules, solving this problem offers some practical implications in protein design for various medical and industrial applications .

In this study, we benchmarked applicability of AlphaFold3 for predicting 3D structure of some proteins of varying sizes and functions, including SARS-CoV-2 coronavirus proteins, bacterial and viral proteins, as well as plant enzymes. We found that AF3 could capture the overall backbone features of the most examined proteins with small RMS deviation <0.5 Å from available X-ray structures. Some common deviation issues were observed for N- and C-terminal segments, which, however, did not affect biological roles of the studied proteins. In cases involving protein dimers or higher-order oligomers, we observed notable differences between the predicted AF3 models of a single-chain monomer and their corresponding experimental structures. These structural discrepancies are particularly apparent in regions associated with protein dimerization, assembly, and binding interfaces. Ultimately, while capturing the overall fold, predicting the complex structure of the Spike glycoprotein remains beyond the current capabilities of AF3.

From a practical standpoint, AF3 is becoming a highly promising computational tool for predicting the three-dimensional structure of proteins, especially when available X-ray or Cryo-EM structures are poorly resolved or have missing sequence segments (as demonstrated in Figs. 8 and 14) . Additionally, there is a new, encouraging direction in the virtual screening of protein mutations to evaluate their impact on protein folding, protein-protein interactions, and overall activity .

Finally, the appearance of AlphaFold 3 has spurred rapid development throughout the field of Computational Biology, inspiring the creation of competing and complementary tools, such as RoseTTAFold by the Baker Lab . OpenFold is an important open-source implementation of AlphaFold, facilitating further customization and innovation, which is expected to accelerate the field in the coming years .

Acknowledgement

A. K. acknowledges the grant № 87/0062 (2021.01/0062) “Molecular design, synthesis and screening of new potential antiviral pharmaceutical ingredients for the treatment of infectious diseases COVID-19” from the National Research Foundation of Ukraine.

References

1. Dill K. A., Ozkan S. B., Weikl T. R., Chodera J. D., Voelz V. A. The protein folding problem: When will it be solved? *Curr Opin Struct Biol* **2007**, 17 (3), 342-346. <https://doi.org/10.1016/j.sbi.2007.06.001>
2. Nassar R., Dignon G. L., Razban R. M., Dill K. A. The protein folding problem: The role of theory. *J. Mol. Biol.* **2021**, 433 (20), 167126. <https://doi.org/10.1016/j.jmb.2021.167126>
3. Weller J. A., Rohs R. Structure-based drug design with a deep hierarchical generative model. *J. Chem. Inf. Model.* **2024**, 64 (16), 6450-6463. <https://doi.org/10.1021/acs.jcim.4c01193>
4. Chen L., Li Q., Nasif K. F., Xie Y., Deng B., Niu S., Pouriyeh S., Dai Z., Chen J., Xie C. Y. AI-driven deep learning techniques in protein structure prediction. *Int. J. Mol. Sci.* **2024**, 25 (15), 8426. <https://doi.org/10.3390/ijms25158426>
5. Bryant P., Pozzati G., Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* **2022**, 13 (1), 1265. <https://doi.org/10.1038/s41467-022-28865-w>
6. Huang P.-S., Boyken S. E., Baker D. The coming of age of *de novo* protein design. *Nature* **2016**, 537 (7620), 320-327. <https://doi.org/10.1038/nature19946>
7. Zonta F., Pantano S. From sequence to mechanobiology? Promises and challenges for AlphaFold 3. *Mechanobiol. Med.* **2024**, 2 (3), 100083. doi: <https://doi.org/10.1016/j.mbm.2024.100083>
8. Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Židek A., Potapenko A., Bridgland A., Meyer C., Kohl S. A. A., Ballard A. J., Cowie A., Romera-Paredes B., Nikolov S., Jain R., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstern S., Silver D., Vinyals O., Senior A. W., Kavukcuoglu K., Kohli P., Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, 596 (7873), 583-589. <https://doi.org/10.1038/s41586-021-03819-2>
9. Nunes-Alves A., Merz K. AlphaFold2 in molecular discovery. *J. Chem. Inf. Model.* **2023**, 63 (19), 5947-5949. <https://doi.org/10.1021/acs.jcim.3c01459>
10. Abramson J., Adler J., Dunger J., Evans R., Green T., Pritzel A., Ronneberger O., Willmore L., Ballard A. J., Bambrick J., Bodenstern S. W., Evans D. A., Hung C.-C., O'Neill M., Reiman D., Tunyasuvunakool K., Wu Z., Žemgulytė A., Arvaniti E., Beattie C., Bertolli O., Bridgland A., Cherepanov A., Congreve M., Cowen-Rivers A. I., Cowie A., Figurnov M., Fuchs F. B., Gladman H., Jain R., Khan Y. A., Low C. M. R., Perlin K., Potapenko A., Savy P., Singh S., Stecula A., Thillaisundaram A., Tong C., Yakneen S., Zhong E. D., Zielinski M., Židek A., Bapst V., Kohli P., Jaderberg M., Hassabis D., Jumper J. M. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, 630 (8016), 493-500. <https://doi.org/10.1038/s41586-024-07487-w>
11. McDonnell R. T., Henderson A. N., Elcock A. H. Structure prediction of large RNAs with AlphaFold 3 highlights its capabilities and limitations. *J. Mol. Biol.* **2024**, 436 (22), 168816. <https://doi.org/10.1016/j.jmb.2024.168816>
12. Kondo H. X., Takano Y. Structure comparison of heme-binding sites in heme protein predicted by AlphaFold3 and AlphaFold2. *Chem. Lett.* **2024**, 53 (8), upae148. <https://doi.org/10.1093/chemle/upae148>
13. He X.-h., Li J.-r., Shen S.-y., Xu H. E. AlphaFold3 versus experimental structures: Assessment of the accuracy in ligand-bound G protein-coupled receptors. *Acta Pharm. Sinica* **2024**, doi: 10.1038/s41401-024-01429-y. <https://doi.org/10.1038/s41401-024-01429-y>
14. Lynch M. L., Snell E. H., Bowman S. E. J. Structural biology in the time of COVID-19: Perspectives on methods and milestones. *IUCrJ* **2021**, 8 335-341. <https://doi.org/10.1107/S2052252521003948>
15. Lubin J. H., Zardecki C., Dolan E. M., Lu C., Shen Z., Dutta S., Westbrook J. D., Hudson B. P., Goodsell D. S., Williams J. K., Voigt M., Sarma V., Xie L., Venkatachalam T., Arnold S., Alfaro

- Alvarado L. H., Catalfano K., Khan A., McCarthy E., Staggers S., Tinsley B., Trudeau A., Singh J., Whitmore L., Zheng H., Benedek M., Currier J., Dresel M., Duvvuru A., Dyszel B., Fingar E., Hennen E. M., Kirsch M., Khan A. A., Labrie-Cleary C., Laporte S., Lenkeit E., Martin K., Orellana M., Ortiz-Alvarez de la Campa M., Paredes I., Wheeler B., Rupert A., Sam A., See K., Soto Zapata S., Craig P. A., Hall B. L., Jiang J., Koeppe J. R., Mills S. A., Pikaart M. J., Roberts R., Bromberg Y., Hoyer J. S., Duffy S., Tischfield J., Ruiz F. X., Arnold E., Baum J., Sandberg J., Brannigan G., Khare S. D., Burley S. K. Evolution of the SARS-CoV-2 proteome in three dimensions (3D) during the first 6 months of the COVID-19 pandemic. *Proteins: Struct., Funct., Bioinf.* **2022**, 90 (5), 1054-1080. <https://doi.org/10.1002/prot.26250>
16. Yang Q., Jian X., Syed A. A. S., Fahira A., Zheng C., Zhu Z., Wang K., Zhang J., Wen Y., Li Z., Pan D., Lu T., Wang Z., Shi Y. Structural comparison and drug screening of spike proteins of ten SARS-CoV-2 variants. *Research* **2022**, 2022, 9781758. <https://doi.org/10.34133/2022/9781758>
 17. Parigger L., Krassnigg A., Grabuschnig S., Gruber K., Steinkellner G., Gruber Christian C. AI-assisted structural consensus-proteome prediction of human monkeypox viruses isolated within a year after the 2022 multi-country outbreak. *Microbiology Spectrum* **2023**, 11 (6), e02315-23. <https://doi.org/10.1128/spectrum.02315-23>
 18. Lubin J. H., Martinusen S. G., Zardecki C., Olivas C., Bacorn M., Balogun M., Slaton E. W., Wu A. W., Sakeer S., Hudson B. P., Denard C. A., Burley S. K., Khare S. D. A comprehensive survey of coronaviral main protease active site diversity in 3D: Identifying and analyzing drug discovery targets in search of broad specificity inhibitors for the next coronavirus pandemic. *bioRxiv* **2023**, 2023.01.30.526101. <https://doi.org/10.1101/2023.01.30.526101>
 19. Jin Z., Du X., Xu Y., Deng Y., Liu M., Zhao Y., Zhang B., Li X., Zhang L., Peng C., Duan Y., Yu J., Wang L., Yang K., Liu F., Jiang R., Yang X., You T., Liu X., Yang X., Bai F., Liu H., Liu X., Guddat L. W., Xu W., Xiao G., Qin C., Shi Z., Jiang H., Rao Z., Yang H. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature*. **2020**, 582 (7811), 289-293. <https://doi.org/10.1038/s41586-020-2223-y>
 20. Rut W., Lv Z., Zmudzinski M., Patchett S., Nayak D., Snipas S. J., El Oualid F., Huang T. T., Bekes M., Drag M., Olsen S. K. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. *Sci. Adv.* **2020**, 6 (42), eabd4596. <https://doi.org/10.1126/sciadv.abd4596>
 21. Yin W., Mao C., Luan X., Shen D.-D., Shen Q., Su H., Wang X., Zhou F., Zhao W., Gao M., Chang S., Xie Y.-C., Tian G., Jiang H.-W., Tao S.-C., Shen J., Jiang Y., Jiang H., Xu Y., Zhang S., Zhang Y., Xu H. E. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, 368 (6498), 1499. <https://doi.org/10.1126/science.abc1560>
 22. Tan B., Joyce R., Tan H., Hu Y., Wang J. SARS-CoV-2 main protease drug design, assay development, and drug resistance studies. *Acc. Chem. Res.* **2023**, 56 (2), 157-168. <https://doi.org/10.1021/acs.accounts.2c00735>
 23. Yevsieieva L. V., Lohachova K. O., Kyrychenko A., Kovalenko S. M., Ivanov V. V., Kalugin O. N. Main and papain-like proteases as prospective targets for pharmacological treatment of coronavirus SARS-CoV-2. *RSC Adv.* **2023**, 13 (50), 35500-35524. <https://doi.org/10.1039/d3ra06479d>
 24. Hu Q., Xiong Y., Zhu G.-H., Zhang Y.-N., Zhang Y.-W., Huang P., Ge G.-B. The SARS-CoV-2 main protease (Mpro): Structure, function, and emerging therapies for COVID-19. *MedComm* **2022**, 3 (3), e151. doi: <https://doi.org/10.1002/mco2.151>
 25. Lohachova K. O., Sviatenko A. S., Kyrychenko A., Ivanov V. V., Langer T., Kovalenko S. M., Kalugin O. N. Computer-aided drug design of novel nirmatrelvir analogs inhibiting main protease of coronavirus SARS-CoV-2. *J. Appl. Pharm. Sci.* **2024**, 14 (5), 232-239. <https://doi.org/10.7324/JAPS.2024.158114>
 26. Zagórska A., Czopek A., Fryc M., Jończyk J. Inhibitors of SARS-CoV-2 main protease (Mpro) as anti-coronavirus agents. *Biomolecules*. **2024**, 14 (7), 797. <https://doi.org/10.3390/biom14070797>
 27. Yang Y., Luo Y.-D., Zhang C.-B., Xiang Y., Bai X.-Y., Zhang D., Fu Z.-Y., Hao R.-B., Liu X.-L. Progress in research on inhibitors targeting SARS-CoV-2 main protease (Mpro). *ACS Omega* **2024**, 9 (32), 34196-34219. <https://doi.org/10.1021/acsomega.4c03023>

28. Xiao Y.-Q., Long J., Zhang S.-S., Zhu Y.-Y., Gu S.-X. Non-peptidic inhibitors targeting SARS-CoV-2 main protease: A review. *Bioorg. Chem.* **2024**, *147* 107380. doi: <https://doi.org/10.1016/j.bioorg.2024.107380>
29. Yevsieieva L., Trostianko P., Kyrychenko A., Ivanov V., Kovalenko S., Kalugin O. Design of non-covalent dual-acting inhibitors for proteases M^{pro} and PL^{pro} of coronavirus SARS-CoV-2 through evolutionary library generation, pharmacophore profile matching, and molecular docking calculations. *Sci. Rise. Pharm. Sci.* **2024**, (6(52)), 15-26. <https://doi.org/10.15587/2519-4852.2024.313808>
30. Ullrich S., Nitsche C. SARS-CoV-2 papain-like protease: Structure, function and inhibition. *ChemBioChem* **2022**, *23* (19), e202200327. <https://doi.org/10.1002/cbic.202200327>
31. Gao X., Qin B., Chen P., Zhu K., Hou P., Wojdyla J. A., Wang M., Cui S. Crystal structure of SARS-CoV-2 papain-like protease. *Acta Pharmac. Sinica B* **2021**, *11* (1), 237-245. <https://doi.org/10.1016/j.apsb.2020.08.014>
32. Osipiuk J., Azizi S.-A., Dvorkin S., Endres M., Jedrzejczak R., Jones K. A., Kang S., Kathayat R. S., Kim Y., Lisnyak V. G., Maki S. L., Nicolaescu V., Taylor C. A., Tesar C., Zhang Y.-A., Zhou Z., Randall G., Michalska K., Snyder S. A., Dickinson B. C., Joachimiak A. Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nat. Commun.* **2021**, *12* (1), 743. <https://doi.org/10.1038/s41467-021-21060-3>
33. Taylor A. J., Ampornnanai K., Rietz T. A., Zhao B., Thiruvaipati A., Wei Q., South T. M., Crow M. M., Apakama C., Sensintaffar J. L., Phan J., Lee T., Fesik S. W. Fragment-based screen of SARS-CoV-2 papain-like protease (PL^{pro}). *ACS Med. Chem. Lett.* **2024**, *15* (8), 1351-1357. <https://doi.org/10.1021/acsmchemlett.4c00238>
34. Magwaza N. N., Mushebenge A. G., Ugbaja S. C., Mbatha N. A., Khan R. B., Kumalo H. M. Mechanistic insights into targeting SARS-CoV-2 papain-like protease in the evolution and management of COVID-19. *BioChem* **2024**, *4* (3), 268-299. <https://doi.org/10.3390/biochem4030014>
35. Han H., Gracia A. V., Røise J. J., Boike L., Leon K., Schulze-Gahmen U., Stentzel M. R., Bajaj T., Chen D., Li I. C., He M., Behrouzi K., Khodabakhshi Z., Nomura D. K., Mofrad M. R. K., Kumar G. R., Ott M., Murthy N. A covalent inhibitor targeting the papain-like protease from SARS-CoV-2 inhibits viral replication. *RSC Adv.* **2023**, *13* (16), 10636-10641. <https://doi.org/10.1039/D3RA00426K>
36. Stasiulewicz A., Maksymiuk A. W., Nguyen M. L., Bełza B., Sulkowska J. I. SARS-CoV-2 papain-like protease potential inhibitors — *in silico* quantitative assessment. *Int. J. Mol. Sci.* **2021**, *22* (8), 3957. <https://doi.org/10.3390/ijms22083957>
37. Shan H., Liu J., Shen J., Dai J., Xu G., Lu K., Han C., Wang Y., Xu X., Tong Y., Xiang H., Ai Z., Zhuang G., Hu J., Zhang Z., Li Y., Pan L., Tan L. Development of potent and selective inhibitors targeting the papain-like protease of SARS-CoV-2. *Cell Chem Biol* **2021**, *28* (6), 855-865.e9. <https://doi.org/10.1016/j.chembiol.2021.04.020>
38. Ma C., Sacco M. D., Xia Z., Lambrinidis G., Townsend J. A., Hu Y., Meng X., Szeto T., Ba M., Zhang X., Gongora M., Zhang F., Marty M. T., Xiang Y., Kolocouris A., Chen Y., Wang J. Discovery of SARS-CoV-2 papain-like protease inhibitors through a combination of high-throughput screening and a FlipGFP-based reporter assay. *ACS Centrl. Sci.* **2021**, *7* (7), 1245-1260. <https://doi.org/10.1021/acscentsci.1c00519>
39. Huynh T., Cornell W., Luan B. *In silico* exploration of inhibitors for SARS-CoV-2's papain-like protease. *Front. Chem.* **2021**, *8* 624163. <https://doi.org/10.3389/fchem.2020.624163>
40. Kakavandi S., Zare I., VaezJalali M., Dadashi M., Azarian M., Akbari A., Ramezani Farani M., Zalpoor H., Hajikhani B. Structural and non-structural proteins in SARS-CoV-2: Potential aspects to COVID-19 treatment or prevention of progression of related diseases. *Cell Comm. Signal.* **2023**, *21* (1), 110. <https://doi.org/10.1186/s12964-023-01104-5>
41. Huang Y., Yang C., Xu X.-f., Xu W., Liu S.-W. Structural and functional properties of SARS-CoV-2 Spike protein: Potential antiviral drug development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41* (9), 1141-1149. <https://doi.org/10.1038/s41401-020-0485-4>
42. Wang L., Wu Y., Yao S., Ge H., Zhu Y., Chen K., Chen W.-Z., Zhang Y., Zhu W., Wang H.-Y., Guo Y., Ma P.-X., Ren P.-X., Zhang X.-l., Li H.-Q., Ali M. A., Xu W.-Q., Jiang H.-l., Zhang L.-K., Zhu L.-l., Ye Y., Shang W.-j., Bai F. Discovery of potential small molecular SARS-CoV-2

- entry blockers targeting the Spike protein. *Acta Pharmacol. Sin.* **2022**, *43* (4), 788-796. <https://doi.org/10.1038/s41401-021-00735-z>
43. Hong Q., Han W., Li J., Xu S., Wang Y., Xu C., Li Z., Wang Y., Zhang C., Huang Z., Cong Y. Molecular basis of receptor binding and antibody neutralization of Omicron. *Nature* **2022**, *604* (7906), 546-552. <https://doi.org/10.1038/s41586-022-04581-9>
 44. Yan W., Zheng Y., Dou C., Zhang G., Arnaout T., Cheng W. The pathogenic mechanism of *Mycobacterium tuberculosis*: Implication for new drug development. *Mol. Biomed.* **2022**, *3* (1), 48. <https://doi.org/10.1186/s43556-022-00106-y>
 45. Bose P., Harit A. K., Das R., Sau S., Iyer A. K., Kashaw S. K. Tuberculosis: Current scenario, drug targets, and future prospects. *Med. Chem. Res.* **2021**, *30* (4), 807-833. <https://doi.org/10.1007/s00044-020-02691-5>
 46. Liu L., Kong C., Fumagalli M., Savková K., Xu Y., Huszár S., Sammartino J. C., Fan D., Chiarelli L. R., Mikušová K., Sun Z., Qiao C. Design, synthesis and evaluation of covalent inhibitors of DpRe1 as antitubercular agents. *Eur. J. Med. Chem.* **2020**, *208* 112773. doi: <https://doi.org/10.1016/j.ejmech.2020.112773>
 47. Snizhko A. D., Kyrychenko A. V., Gladkov E. S. Synthesis of novel derivatives of 5,6,7,8-tetrahydro-quinazolines using of α -aminoamidines and *in silico* screening of their biological activity. *Int. J. Mol. Sci.* **2022**, *23* (7), 3781. <https://doi.org/10.3390/ijms23073781>
 48. Hong W., Wang Y., Chang Z., Yang Y., Pu J., Sun T., Kaur S., Sacchetti J. C., Jung H., Lin Wong W., Fah Yap L., Fong Ngeow Y., Paterson I. C., Wang H. The identification of novel *Mycobacterium tuberculosis* DHFR inhibitors and the investigation of their binding preferences by using molecular modelling. *Sci. Rep.* **2015**, *5* (1), 15328. <https://doi.org/10.1038/srep15328>
 49. Dal Molin M., Selchow P., Schäfle D., Tschumi A., Ryckmans T., Laage-Witt S., Sander P. Identification of novel scaffolds targeting *Mycobacterium tuberculosis*. *J. Mol. Med.* **2019**, *97* (11), 1601-1613. <https://doi.org/10.1007/s00109-019-01840-7>
 50. Reddy B. K. K., Landge S., Ravishankar S., Patil V., Shinde V., Tantry S., Kale M., Raichurkar A., Menasinakai S., Mudugal Naina V., Ambady A., Ghosh A., Tunduguru R., Kaur P., Singh R., Kumar N., Bharath S., Sundaram A., Bhat J., Sambandamurthy Vasanth K., Björkelid C., Jones T. A., Das K., Bhandarkar B., Malolanarasimhan K., Mukherjee K., Ramachandran V. Assessment of *Mycobacterium tuberculosis* pantothenate kinase vulnerability through target knockdown and mechanistically diverse inhibitors. *Antimicrob. Agents Chemother.* **2014**, *58* (6), 3312-3326. <https://doi.org/10.1128/aac.00140-14>
 51. Björkelid C., Bergfors T., Raichurkar A. K. V., Mukherjee K., Malolanarasimhan K., Bhandarkar B., Jones T. A. Structural and biochemical characterization of compounds inhibiting *Mycobacterium tuberculosis* pantothenate kinase. *J. Biol. Chem.* **2013**, *288* (25), 18260-18270. <https://doi.org/10.1074/jbc.M113.476473>
 52. Chollet A., Mourey L., Lherbet C., Delbot A., Julien S., Baltas M., Bernadou J., Pratviel G., Maveyraud L., Bernardes-Génisson V. Crystal structure of the enoyl-ACP reductase of *Mycobacterium tuberculosis* (InhA) in the apo-form and in complex with the active metabolite of isoniazid pre-formed by a biomimetic approach. *J. Struct. Biol.* **2015**, *190* (3), 328-337. <https://doi.org/10.1016/j.jsb.2015.04.008>
 53. Marchenko K. I., Kyrychenko A. V., Kolos N. M. Synthesis and modification of 7-aryloxy derivatives of 4,7-dihydro-[1,2,4]triazolo-[1,5-a]-pyrimidine as potent inhibitors of Sirtuin-2. *Funct. Mater.* **2024**, *31* (2), 260-268. <https://doi.org/10.15407/fm31.02.260>
 54. Zhao X., Allison D., Condon B., Zhang F., Gheyi T., Zhang A., Ashok S., Russell M., MacEwan I., Qian Y., Jamison J. A., Luz J. G. The 2.5 Å crystal structure of the Sirt1 catalytic domain bound to nicotinamide adenine dinucleotide (NAD⁺) and an indole (ex527 analogue) reveals a novel mechanism of histone deacetylase inhibition. *J. Med. Chem.* **2013**, *56* (3), 963-969. <https://doi.org/10.1021/jm301431y>
 55. Kumar A., Chauhan S. How much successful are the medicinal chemists in modulation of Sirt1: A critical review. *Eur. J. Med. Chem.* **2016**, *119* 45-69. <https://doi.org/10.1016/j.ejmech.2016.04.063>
 56. Lipson V. V., Yaremenko F. G., Vakula V. M., Kovalenko S. V., Kyrychenko A. V., Desenko S. M., Musatov V. I., Borysko P. O., Zozulya S. O. Imidazole derivatives as potent inhibitors of Sirtuin-1. *Funct. Mater.* **2023**, *30* (4), 486-493. <https://doi.org/10.15407/fm30.04.486>

57. Lipson V. V., Yaremenko F. G., Vakula V. M., Kovalenko S. V., Kyrychenko A. V., Desenko S. M., Borysko P. O., Zozulya S. O. Discovery of novel N-acylhydrazones as potent inhibitors of Sirtuin-1. *SynOpen* **2024**, 08 (02), 100-108. <https://doi.org/10.1055/s-0043-1763747>
58. Kyrychenko A. V., Ladokhin A. S. Fluorescence tools for studies of membrane protein insertion. *Biopolym. Cell.* **2018**, 34 (4), 251-271. <https://doi.org/10.7124/bc.00097F>
59. Ladokhin A. S., Kyrychenko A., Rodnin M. V., Vasquez-Montes V. Conformational switching, refolding and membrane insertion of the diphtheria toxin translocation domain. In *Methods Enzymol.*, Academic Press: 2021; Vol. 649, pp 341-370. <https://doi.org/10.1016/bs.mie.2020.12.016>
60. Kyrychenko A., Posokhov Y. O., Rodnin M. V., Ladokhin A. S. Kinetic intermediate reveals staggered pH-dependent transitions along the membrane insertion pathway of the diphtheria toxin T-domain. *Biochemistry*. **2009**, 48 (32), 7584-7594. <https://doi.org/10.1021/bi9009264>
61. Rodnin M. V., Kyrychenko A., Kienker P., Sharma O., Posokhov Y. O., Collier R. J., Finkelstein A., Ladokhin A. S. Conformational switching of the diphtheria toxin T domain. *J. Mol. Biol.* **2010**, 402 (1), 1-7. <https://doi.org/10.1016/j.jmb.2010.07.024>
62. Rodnin Mykola V., Kyrychenko A., Kienker P., Sharma O., Vargas-Urbe M., Collier R. J., Finkelstein A., Ladokhin Alexey S. Replacement of C-terminal histidines uncouples membrane insertion and translocation in diphtheria toxin T-domain. *Biophys. J.* **2011**, 101 (10), L41-L43. <https://doi.org/10.1016/j.bpj.2011.10.018>
63. Kurnikov I. V., Kyrychenko A., Flores-Canales J. C., Rodnin M. V., Simakov N., Vargas-Urbe M., Posokhov Y. O., Kurnikova M., Ladokhin A. S. pH-triggered conformational switching of the diphtheria toxin T-domain: The roles of N-terminal histidines. *J. Mol. Biol.* **2013**, 425 (15), 2752-2764. <https://doi.org/10.1016/j.jmb.2013.04.030>
64. Vargas-Urbe M., Rodnin M. V., Öjemalm K., Holgado A., Kyrychenko A., Nilsson I., Posokhov Y. O., Makhatadze G., von Heijne G., Ladokhin A. S. Thermodynamics of membrane insertion and refolding of the diphtheria toxin T-domain. *J Membr Biol* **2015**, 248 (3), 383-394. <https://doi.org/10.1007/s00232-014-9734-0>
65. Kyrychenko A., Posokhov Y. O., Vargas-Urbe M., Ghatak C., Rodnin M. V., Ladokhin A. S. Fluorescence applications for structural and thermodynamic studies of membrane protein insertion. In *Reviews in fluorescence 2016*, Geddes, C. D., Ed. Springer International Publishing: Cham, Switzerland, 2017; pp 243-274. https://doi.org/10.1007/978-3-319-48260-6_10
66. Rodnin M. V., Vasques-Montes V., Kyrychenko A., Oliveira N. F. B., Kashipathy M. M., Battaile K. P., Douglas J., Lovell S., Machuqueiro M., Ladokhin A. S. Histidine protonation and conformational switching in diphtheria toxin translocation domain. *Toxins* **2023**, 15 (7), Art. num. 410. <https://doi.org/10.3390/toxins15070410>
67. Rodnin M. V., Kashipathy M. M., Kyrychenko A., Battaile K. P., Lovell S., Ladokhin A. S. Structure of the diphtheria toxin at acidic pH: Implications for the conformational switching of the translocation domain. *Toxins* **2020**, 12 (11), Art. num. 704. <https://doi.org/10.3390/toxins12110704>
68. Isorna P., Polaina J., Latorre-García L., Cañada F. J., González B., Sanz-Aparicio J. Crystal structures of paenibacillus polymyxa β -glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases. *J. Mol. Biol.* **2007**, 371 (5), 1204-1218. <https://doi.org/10.1016/j.jmb.2007.05.082>
69. Chepeleva L. V., Tarasenko D. O., Chumak A. Y., Demidov O. O., Snizhko A. D., Kolomoitsev O. O., Kotliar V. M., Gladkov E. S., Tatarts A. L., Kyrychenko A. V., Roshal A. D. 4'-benzyloxyflavonol glucoside as fluorescent indicator for β -glucosidase activity. *Funct. Mater.* **2023**, 30 (4), 494-505. <https://doi.org/10.15407/fm30.04.494>
70. Chepeleva L. V., Demidov O. O., Snizhko A. D., Tarasenko D. O., Chumak A. Y., Kolomoitsev O. O., Kotliar V. M., Gladkov E. S., Kyrychenko A., Roshal A. D. Binding interactions of hydrophobically-modified flavonols with β -glucosidase: Fluorescence spectroscopy and molecular modelling study. *RSC Adv.* **2023**, 13 (48), 34107-34121. <https://doi.org/10.1039/D3RA06276G>
71. Demidov O. O., Gladkov E. S., Kyrychenko A. V., Roshal A. D. Synthetic and natural flavonols as promising fluorescence probes for β -glucosidase activity screening *Funct. Mater.* **2022**, 29 (2), 252-262. <https://doi.org/10.15407/fm29.02.252>
72. Zhang Y., Vass M., Shi D., Abualrous E., Chambers J. M., Chopra N., Higgs C., Kasavajhala K., Li H., Nandekar P., Sato H., Miller E. B., Repasky M. P., Jerome S. V. Benchmarking refined and

- unrefined AlphaFold2 structures for hit discovery. *J. Chem. Inf. Model.* **2023**, 63 (6), 1656-1667. <https://doi.org/10.1021/acs.jcim.2c01219>
73. Díaz-Rovira A. M., Martín H., Beuming T., Díaz L., Guallar V., Ray S. S. Are deep learning structural models sufficiently accurate for virtual screening? Application of docking algorithms to AlphaFold2 predicted structures. *J. Chem. Inf. Model.* **2023**, 63 (6), 1668-1674. <https://doi.org/10.1021/acs.jcim.2c01270>
 74. Beuming T., Martín H., Díaz-Rovira A. M., Díaz L., Guallar V., Ray S. S. Are deep learning structural models sufficiently accurate for free-energy calculations? Application of FEP⁺ to AlphaFold2-predicted structures. *J. Chem. Inf. Model.* **2022**, 62 (18), 4351-4360. <https://doi.org/10.1021/acs.jcim.2c00796>
 75. Bernatavicius A., Šicho M., Janssen A. P. A., Hassen A. K., Preuss M., van Westen G. J. P. AlphaFold meets *de novo* drug design: Leveraging structural protein information in multitarget molecular generative models. *J. Chem. Inf. Model.* **2024**, 64 (21), 8113-8122. <https://doi.org/10.1021/acs.jcim.4c00309>
 76. Gu S., Yang Y., Zhao Y., Qiu J., Wang X., Tong H. H. Y., Liu L., Wan X., Liu H., Hou T., Kang Y. Evaluation of AlphaFold2 structures for hit identification across multiple scenarios. *J. Chem. Inf. Model.* **2024**, 64 (9), 3630-3639. <https://doi.org/10.1021/acs.jcim.3c01976>
 77. Pei J., Andreeva A., Chuguransky S., Lázaro Pinto B., Paysan-Lafosse T., Dustin Schaeffer R., Bateman A., Cong Q., Grishin N. V. Bridging the gap between sequence and structure classifications of proteins with AlphaFold models. *J. Mol. Biol.* **2024**, 436 (22), 168764. <https://doi.org/10.1016/j.jmb.2024.168764>
 78. Wee J., Wei G.-W. Evaluation of AlphaFold 3's protein-protein complexes for predicting binding free energy changes upon mutation. *J. Chem. Inf. Model.* **2024**, 64 (16), 6676-6683. <https://doi.org/10.1021/acs.jcim.4c00976>
 79. Baek M., DiMaio F., Anishchenko I., Dauparas J., Ovchinnikov S., Lee G. R., Wang J., Cong Q., Kinch L. N., Schaeffer R. D., Millán C., Park H., Adams C., Glassman C. R., DeGiovanni A., Pereira J. H., Rodrigues A. V., van Dijk A. A., Ebrecht A. C., Opperman D. J., Sagmeister T., Buhlheller C., Pavkov-Keller T., Rathinaswamy M. K., Dalwadi U., Yip C. K., Burke J. E., Garcia K. C., Grishin N. V., Adams P. D., Read R. J., Baker D. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, 373 (6557), 871-876. <https://doi.org/10.1126/science.abj8754>
 80. Ahdriz G., Bouatta N., Floristean C., Kadyan S., Xia Q., Gerecke W., O'Donnell T. J., Berenberg D., Fisk I., Zanichelli N., Zhang B., Nowaczynski A., Wang B., Stepniewska-Dziubinska M. M., Zhang S., Ojewole A., Guney M. E., Biderman S., Watkins A. M., Ra S., Lorenzo P. R., Nivon L., Weitzner B., Ban Y.-E. A., Chen S., Zhang M., Li C., Song S. L., He Y., Sorger P. K., Mostaque E., Zhang Z., Bonneau R., AlQuraishi M. Openfold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nat. Methods* **2024**, 21 (8), 1514-1524. <https://doi.org/10.1038/s41592-024-02272-z>

Received 11.09.2024

Accepted 19.12.2024

Є. В. Дума, О. В. Кириченко. Порівняння ефективності прогнозування 3D-структури білка з використанням AlphaFold 3 від Google DeepMind.

Харківський національний університет імені В.Н. Каразіна, хімічний факультет, майдан Свободи, 4, Харків, 61022, Україна

Тривимірна структура білків безпосередньо пов'язана з їхньою функцією, що робить її визначення вирішальним для розуміння біологічних процесів і вирішення проблем, пов'язаних зі здоров'ям людини та науками про життя. Незважаючи на постійне експериментальне отримання нових білкових структур, залишається значний розрив між кількістю доступних білкових послідовностей та експериментальних третинних структур.

Останні дослідження зосереджені на прогнозуванні білкових структур за допомогою обчислювальних методів, заснованих або на шаблонах, або на моделюванні емпіричного силового поля. В останні роки різні методи були об'єднані для вирішення окремих обмежень цих підходів, що призвело до розробки AlphaFold3 Google DeepMind. Цей інноваційний інструмент дозволяє прогнозувати тривимірні структури білка з високою точністю на основі його амінокислотної послідовності. У цьому дослідженні ми перевірили застосовність, продуктивність і обмеження AlphaFold3 (AF3) для прогнозування тривимірної структури ши-

рокого ряду білків, включаючи білки коронавірусу SARS-CoV-2, інші бактеріальні та вірусні білки, а також рослинні ферменти. Ми виявили, що AlphaFold3 добре відтворює 3D-будову більшості з досліджених білків з точки зору відхилення від доступних рентгенівських структур. Деякі незначні помилки згортання N- і C-кінцевих сегментів були поширеними, що часто не впливало на біологічну роль досліджуваних білків. У випадках, пов'язаних із білковими димерами або олігомерами вищого порядку, спостерігалися помітні відмінності між прогнозованими моделями AF3 для одноланцюгових мономерів та їх відповідними експериментальними структурами. Ці розбіжності були особливо очевидні в областях, пов'язаних з димеризацією білка, білковою асоціацією і на границі стикування білків. Зрештою, незважаючи на відтворення глобального фолдингу, прогнозування складної структури Spike глікопротеїнів все ще виходить за межі поточних можливостей AF3.

Ключові слова: згортання білка, фермент, тривимірна структура, дизайн ліків, штучний інтелект, AlphaFold3

Надіслано до редакції 11.09.2024

Прийнято до друку 19.12.2024

Kharkiv University Bulletin. Chemical Series. Issue 43 (66), 2024

EVOLUTIONARY STRUCTURE OPTIMIZATION OF ENSITRELVIR AS NON-COVALENT INHIBITOR OF SARS-COV-2 MAIN PROTEASE M^{PRO}**K. O. Lohachova^a, A. S. Sviatenko^b, A. V. Kyrychenko^c, O. N. Kalugin^d***V. N. Karazin Kharkiv National University, School of Chemistry, 4 Svobody sqr., Kharkiv, 61022 Ukraine*

- a) ✉ ekaterinalogatcheva@gmail.com
b) ✉ anastasiia.sviatenko@student.karazin.ua
c) ✉ a.v.kyrychenko@karazin.ua
d) ✉ onkalugin@gmail.com

-  <https://orcid.org/0000-0001-7826-8320>
 <https://orcid.org/0009-0007-2452-1168>
 <https://orcid.org/0000-0002-6223-0990>
 <https://orcid.org/0000-0003-3273-9259>

Ensirelvir is a non-covalent, non-peptide inhibitor of the SARS-CoV-2 main protease, M^{PRO}. It has demonstrated effective antiviral activity against various coronavirus variants *in vitro*, along with favorable drug metabolism and pharmacokinetic profiles suitable for oral treatment. Thus, developing novel analogues of ensirelvir is of great importance. In this study, we conducted *in silico* design of its analogues by employing evolutionary structure optimization of the parent ensirelvir scaffold. In the first stage, we generated a virtual evolutionary library consisting of 6334 new analogues based on a series of fitness criteria, including molecular weight (M_w), cLogP, polar surface area, structural and conformational similarity, flexibility, and molecular shape. Next, we filtered the evolutionary library using a 3D pharmacophore model created from the available X-ray structure of the co-crystallized complex of ensirelvir and M^{PRO}. We then performed molecular docking calculations to rank the selected candidates according to their binding affinity and selectivity for the M^{PRO} receptor. This binding score ranking allowed us to identify ten analogues of ensirelvir that exhibit superior binding affinity to the protease M^{PRO} compared to the original ensirelvir inhibitor. Our evolutionary structure optimization indicates that the primary structural modifications that enhance the overall antiviral effect of ensirelvir are found in the 1-methyl-1*H*-1,2,4-triazole and 6-chloro-2-methyl-2*H*-indazole fragments.

Keywords: coronavirus, COVID-19, heterocyclic compounds, M^{PRO}, evolutionary library, molecular docking.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which emerged in late 2019, led to a global epidemic and was declared a pandemic by the World Health Organization (WHO) on March 11, 2020 [1]. The 2019 coronavirus disease (COVID-19) pandemic has left an irreversible impact on the world, with more than 18 million deaths and significant disruptions to the global economy that continue to be felt today [2]. The rapid introduction of vaccines has effectively reduced the risk of severe illness from the disease [3-4]. COVID-19, caused by the highly contagious and mutating SARS-CoV-2, saw the Omicron variant replace the Delta variant, and it continues to evolve into various new strains. This ongoing evolution presents substantial challenges for healthcare systems and public health infrastructure [5-7]. Overall statistics indicate that vaccination does not provide long-lasting immunity against the SARS-CoV-2 virus. Although the outbreak of COVID-19 has transitioned to a more normalized phase, ongoing research into inhibitors remains crucial for the prevention of future coronavirus pandemics [8-9].

COVID-19 is a respiratory disease caused by the SARS-CoV-2 virus, which is a positive single-stranded RNA virus [10]. This virus contains several structural proteins, including the nucleocapsid protein (N), spike protein (S), envelope protein (E), and membrane glycoprotein (M), along with its RNA genome. SARS-CoV-2 is classified as part of the β -coronavirus genus, which also includes SARS-CoV, HCoV-OC43, HCoV-HKU1, and MERS-CoV.

Structural proteins, particularly spike (S) proteins, are crucial in viral pathogenesis as they facilitate receptor recognition and membrane fusion. Although E proteins are the smallest structural proteins, they are essential for viral assembly, budding, envelope formation, and overall pathogenicity. M proteins play a key role in the assembly of the virus, while N proteins have multiple functions, including the packaging of the viral RNA genome. Nonstructural proteins of SARS-CoV-2, such as the main protease (M^{PRO}) and RNA-dependent RNA polymerase (RdRp), are highly conserved across different virus strains. This conservation makes them attractive targets for drug development, as targeting these

proteins can minimize interference with host cell physiological processes [11-13]. Importantly, the major protease (M^{pro}) of SARS-CoV-2 exhibits unique substrate specificity by cleaving peptides exclusively after glutamine residues, which reduces the risk of host cell toxicity from inhibitors targeting M^{pro} . Inhibition of SARS-CoV-2 M^{pro} has become a promising approach for the development of therapeutics for the treatment of COVID-19 [1, 14] (Fig. 1). M^{pro} is characterized as a three-domain cysteine protease with a catalytic dyad consisting of Cys145–His41 located in the cleft between Domains I and II [1, 15].

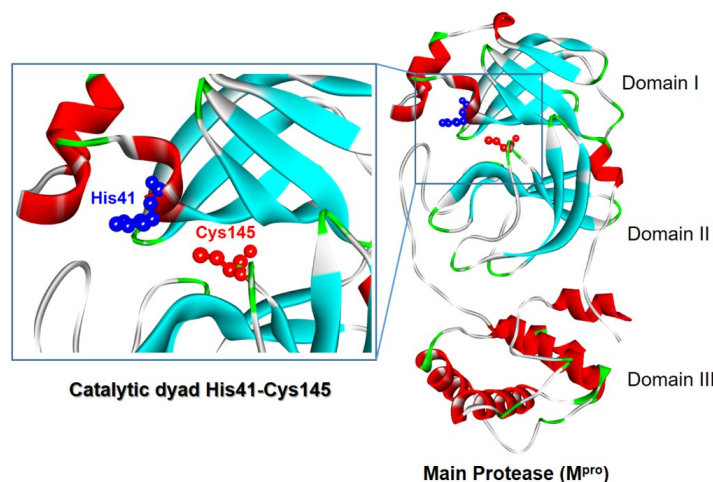


Figure 1. The X-ray structure and a binding site of the main cysteine protease M^{pro} of SARS-CoV-2 (PDB: 8HEF) [16]. An insert shows the M^{pro} binding site and two crucial catalytic residues of His41 and Cys145, are rendered as blue and red sticks, respectively.

Shionogi Pharmaceutical Company, in collaboration with Hokkaido University, has developed a non-peptide oral drug called ensitrelvir (experimental name S-217622, systematic name 1-(2,4,5-trifluorobenzyl)-3-[(1-methyl-1H-1,2,4-triazol-3-yl)methyl]-(6E)-6-[(6-chloro-2-methyl-2H-indazol-5-yl)-imino]-1,3,5-triazinane-2,4-dione, see Fig. 2a). This drug received approval for sale in Japan on November 22, 2022 [17]. The molecule was discovered through a process that included virtual screening, biological analysis, and optimization of the hit compound using a structure-based drug design approach. The X-ray crystal structure of the ensitrelvir complex with the M^{pro} enzyme reveals that the 1-methyl-1H-1,2,4-triazole link at the S1 position forms a H-bond of 2.99 Å with the side chain NH of H163 (see Fig. 2b).

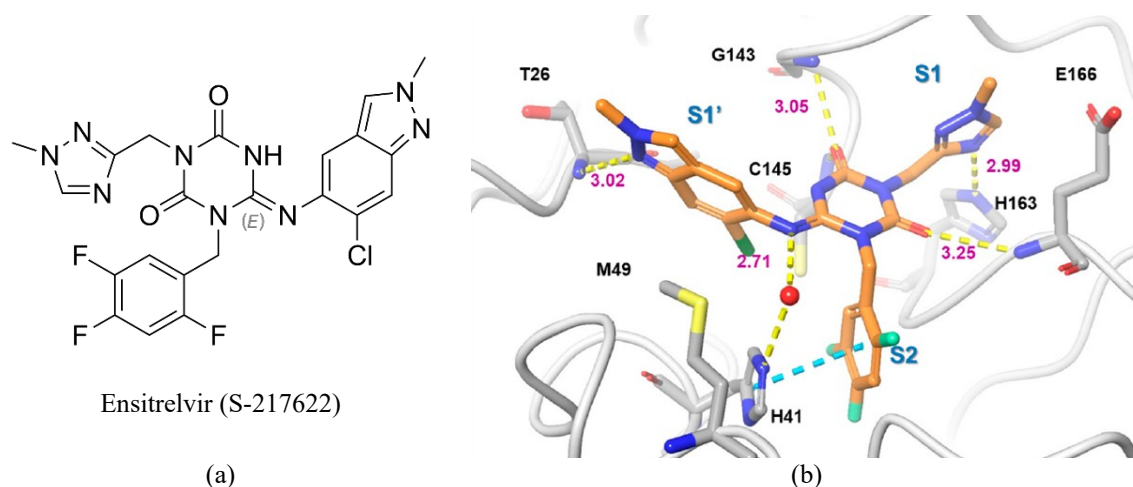


Figure 2. (a) Chemical structure of ensitrelvir. (b) The X-ray crystal structure of the ensitrelvir complex with M^{pro} (PDB: 8HEF). Adapted with permission by [16]. Copyright © 2022. The Authors. Published by American Chemical Society.

In addition, the trifluorobenzyl group of ensitrelvir occupies the S2 hydrophobic pocket, aligning parallel to the histidine H41 ring, which enhances contact. The 6-chloro-2-methyl-2H-indazole fragment, located in the S1' pocket, forms a H-bond of 3.02 Å with the NH of the main chain at T26 and maintains hydrophobic interactions with M49. Finally, one of the carbonyl groups of the 1,3,5-triazine-2,4-dione moiety forms a H-bond of 3.25 Å with the NH of E166.

Ensitrelvir showed significant inhibitory activity against SARS-CoV-2 M^{pro} as a non-covalent non-peptide inhibitor, with an IC₅₀ value of 0.013 µM and the effective antiviral activity with an EC₅₀ value of 0.37 µM [17, 18]. It has demonstrated antiviral activity against a number of SARS-CoV-2 variants and coronaviruses *in vitro*, as well as favorable drug metabolism and pharmacokinetic profiles for oral dosing [21-24].

Therefore, the purpose of this research is to investigate the non-covalent interactions between ligands and proteins. This study will focus on a virtual evolutionary library created based on the drug ensitrelvir and a leading compound that exhibits better affinity for the main M^{pro} protease of the SARS-CoV-2 virus.

Evolutionary Algorithm

To generate the evolutionary library of ensitrelvir analogues we used OSIRIS DataWarrior software [25]. The evolutionary algorithm mimics natural evolution and starts with a small initial set of molecules (first parent derivative) [26]. In the next step, several derivatives, new but similar molecules, are created by applying a small random structural modification (the first progeny generation). Modifications such as single atom replacement, atom insertion, bond reordering, substituent migration, ring aromatization, chiral center inversion, etc. are applied (Fig. 3). Each modification is then evaluated on how well it meets the target criteria.

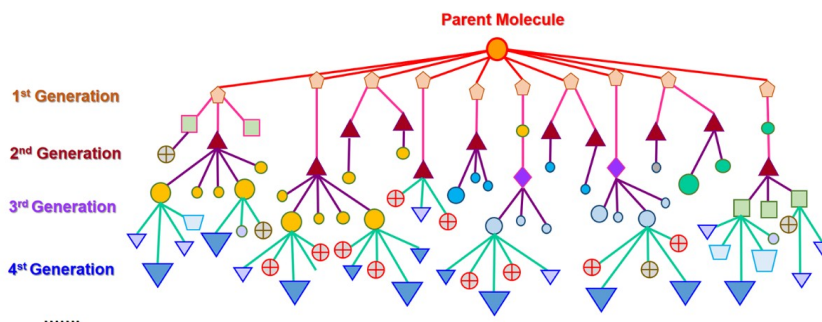


Figure 3. General scheme for generating the Evolutionary Library Tree. Bigger size symbols represent the 1st, 2nd, 3rd, and 4th survived generations, while smaller symbols illustrate some unsuccessful branches that did not evolve into the next generation.

In a next step, the algorithm applies customizable fitness criteria to rank the new generation's molecules according to these criteria. The highest ranking molecules from this generation are selected to survive and form the parent molecules for the next generation. Typically, after about one or a few hundred molecule generations DataWarrior has arrived at structures that optimally match the defined fitness criteria..

Molecular docking setup

Molecular docking was carried out for all the synthesized compounds against the main protease M^{pro} (PDB: 8HEF) of SARS-CoV-2 [17]. The preparation of the receptor and ligands were carried out with the LigandScout software, version 4.5 [27]. Molecular docking calculations were performed with the built-in version of the AutoDock Vina 1.1.2 software [28]. We carried out the semi-flexible docking, so that the receptor was kept rigid and the ligand molecules were conformational flexible. In the Autodock Vina, the Lamarckian genetic algorithm was used as a research parameter. For each ligand, three independent runs were performed using different random seeds. The best docking mode corresponds to the largest ligand-binding affinity. Molecular graphics and visualization were performed using VMD 1.9.3 [23].

Results and Discussion

Generation of Evolutionary Library of *Ensitrelvir Analogues*

To create a library of new molecules that are structurally similar to the original structure of ensitrelvir, several approaches can be employed [30]. One promising method involves generating a virtual library of ensitrelvir analogs using DataWarrior software [25]. This process utilizes an evolutionary algorithm that mimics the natural evolution of drug-like substances. The creation of new molecules follows a specific procedure: it begins with the ensitrelvir molecule (Fig. 2a), referred to as the "first generation" (Fig. 3). From this starting point, a series of similar derivatives are generated by applying minor random structural modifications to the parent structure. Each new modification is then assessed based on its adherence to nine selected suitability criteria, including molecular weight (M_w), logarithm of the partition coefficient (cLogP), polar surface area (PSA) [31], molecular shape, flexibility, and others (Table 1). The selection principles for these fitness criteria are discussed in more details in our recent works [26, 32].

Table 1. Selected fitness criteria and their parameters used to create a virtual evolutionary library of ensitrelvir analogues.

Fitness criteria	Settings
Number of generations	25
Compounds per generation	2048
Compounds survive a generation	256
Structure similarity	algorithm OrgFunctions target - ensitrelvir
Conformational similarity	algorithm PheSA
Molecular weight (M_w)	400-500 g/mol
cLogP	2-4
Polar surface area	80-120 Å ²
Molecular flexibility	0.3-0.7
Molecular shape	0.3-0.7
H-donors	3-5
H-acceptors	5-10
Rotatable bond count	4-8

Structural mutations that meet these criteria are more likely to occur. In the current generation, the new counterparts were organized according to these criteria, allowing the molecules with the highest ranking to be selected for survival and to serve as parent molecules for the next generation. Following this evolutionary algorithm, a total of 25 generations were produced, resulting in 6334 virtual ensitrelvir analogs.

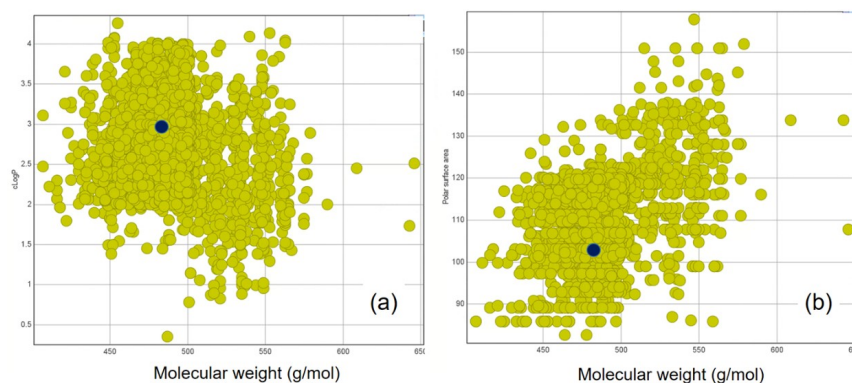


Figure 4. Two-dimensional graph of molecular weight distribution versus cLogP (a) and PSA (b) parameters for the created evolutionary library. The blue points indicate the parent ensitrelvir molecule.

The evolutionary library of ensitrelvir derivatives was generated based on the principle of structural variation, as well as by applying specific fitness criteria (see Table 1). These criteria included molecular weight ($400 \text{ g/mol} < M_w < 500 \text{ g/mol}$), polar surface area ($80 \text{ \AA}^2 < \text{PSA} < 120 \text{ \AA}^2$), and cLogP values ($2 < \text{cLogP} < 4$). Figure 4 illustrates the two-dimensional distribution of these parameters within the generated library, which exhibits a symmetric distribution.

3D Pharmacophore Screening

The created evolutionary library is large enough to verify its direction by the method of molecular docking. Therefore, at first the library was further analyzed and filtered *in silico* according to the criterion of the correspondence of each new molecule to the 3D model of the ensitrelvir pharmacophore [33]. The pharmacophore model was created using the LigandScout software package [27] (Fig. 5) based on the first available crystal structure of ensitrelvir in complex with M^{pro} (PDB: 8HEF) [16]. The screening procedure relative to the ensitrelvir 3D-pharmacophore model made it possible to reduce the chemical space of the library from 6334 virtual molecules to 41 active ensitrelvir derivatives that have pharmacophore structural similarity to the initial structure. In this pharmacophore filtering, a pharmacophore-fit score cut-off of 75 or higher was utilized.

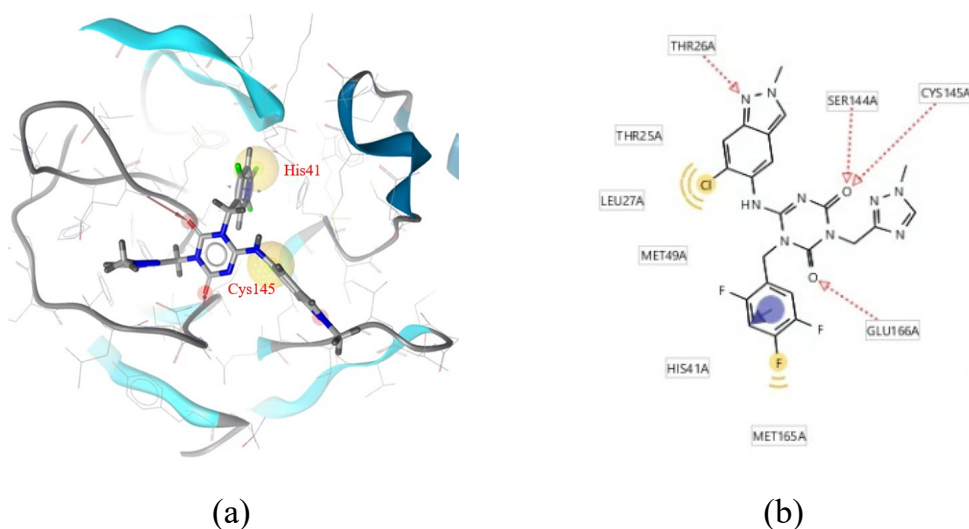


Figure 5. (a) 3D model of the ensitrelvir pharmacophore constructed from its crystal structure in complex with M^{pro} (PDB: 8HEF) [16]. The labeled His41 and Cys145 residues are the active site of M^{pro} . (b) Scheme of intermolecular interactions of ensitrelvir with key surrounding residues of M^{pro} .

Molecular Docking of Evolutionary Library

To identify potential leader compounds from the evolutionary library, we evaluated the binding affinities of 41 selected active candidates by performing molecular docking with M^{pro} using LigandScout software [27]. LigandScout incorporates a built-in version of the widely used AutoDock Vina molecular docking method. To calibrate and validate the ligand-protease docking algorithm, we conducted a docking study of ensitrelvir and compared its geometry to the corresponding X-ray structural data.

Molecular docking was conducted for the selected inhibitors of the M^{pro} main protease of the SARS-CoV-2 coronavirus. The spatial structure of the main protease M^{pro} (PDB ID: 8HEF) was obtained from the Brookhaven Protein Data Bank (Fig. 6a-b) [16]. The results of the molecular docking for ensitrelvir, along with the experimental crystallographic data of its complex with M^{pro} , are in good agreement (Fig. 6c). This consistency allows us to use this method for the quantitative analysis of the affinity of new derivatives of ensitrelvir to M^{pro} .

The molecular docking results for the selected 41 active derivatives revealed that 10 analogues demonstrated a higher binding affinity to M^{pro} than the parent ensitrelvir. Figure 7 illustrates the struc-

tures of the top five leading ligands, along with their energetic characteristics related to the protease-ligand interaction.

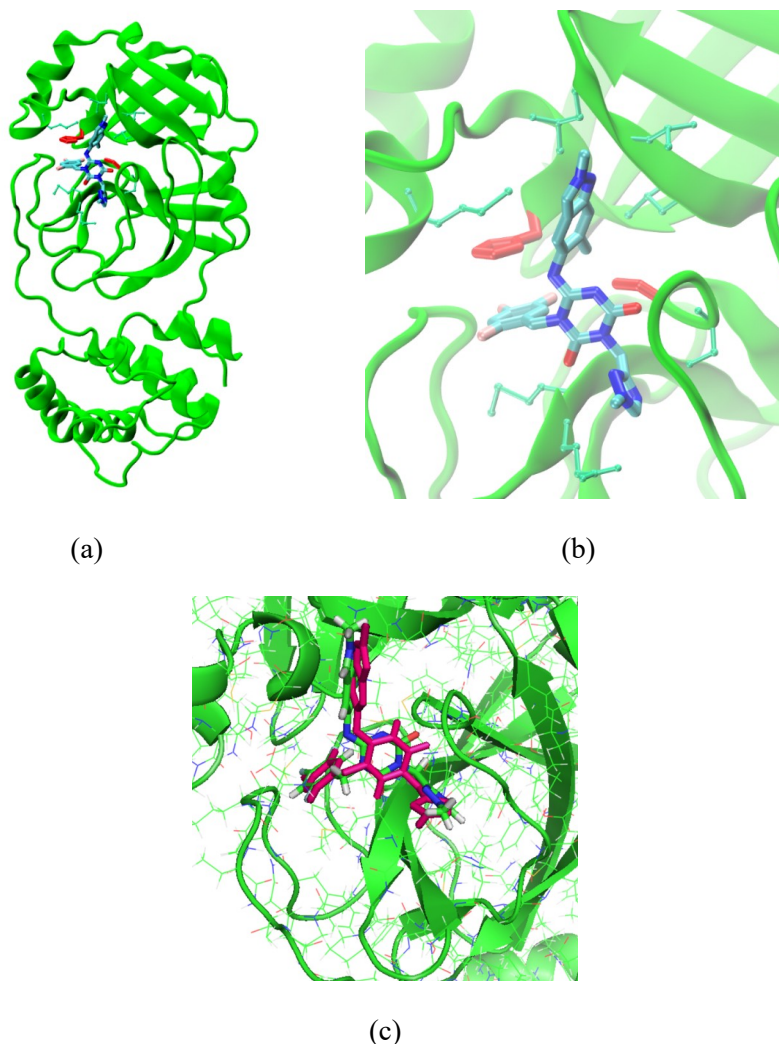


Figure 6. (a, b) Configuration of ensitrelvir incorporation into the active site of M^{pro} protease of SARS-CoV-2 virus (PDB: 8HEF) [16]. The catalytic amino acid residues His41 and Cys145 are shown in red. (c) Comparison of the molecular docking results of ensitrelvir with its crystallographic structure in complex with M^{pro}. The ligand configuration estimated by molecular docking is shown in pink.

The following parameters were considered as binding energy criteria:

1. Binding affinity as determined by the AutoDock Vina docking algorithm.
2. Binding Affinity Score (BAS) calculated using the LigandScout software package.

The BAS score accounts for both the protein-ligand binding enthalpy and the solvent effects. It is regarded as the most representative indicator and was therefore chosen for the selection of lead ligands.

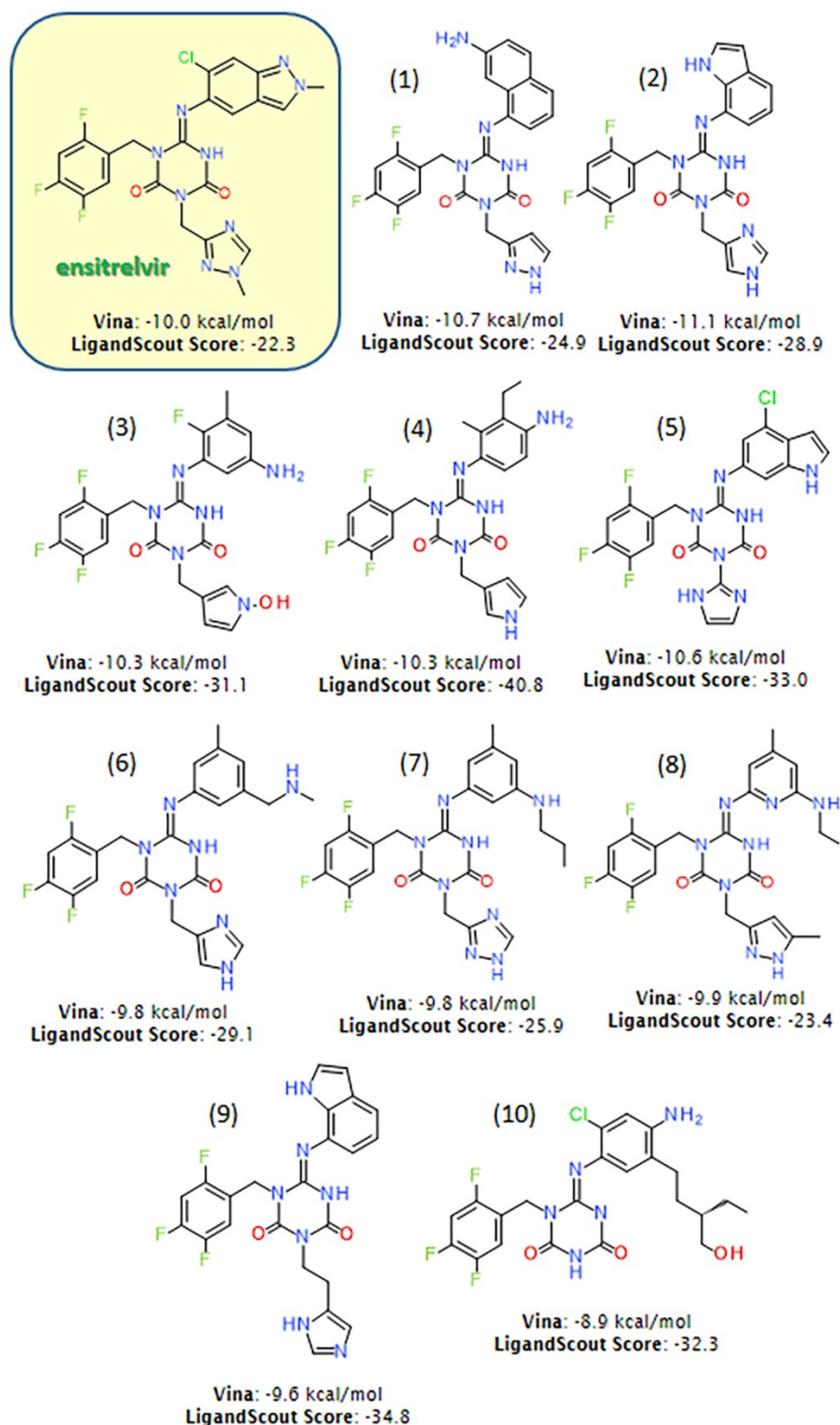


Figure 7. Molecular structure of ensitrelvir and its analogues (1-10), which show the higher binding affinity to M^{pro} compared to the parent ensitrelvir.

A key structural feature of the binding and docking of the selected inhibitors with the active site of M^{pro} is the maintenance of the configuration of the trifluorobenzyl fragment, as seen in derivatives 1-10 illustrated in Figures 7 and 8, respectively. Efforts to replace the 2,4,5-trifluorobenzyl moiety of ensitrelvir did not enhance the binding affinity. The overall binding affinity of these inhibitors

increases due to structural modifications in other parts of the molecule. The primary structural modifications observed in inhibitors **1-10** occur in two specific fragments of the ensitrelvir molecule: the 1-methyl-1H-1,2,4-triazole and the 6-chloro-2-methyl-2H-indazole fragments (see Fig. 7).

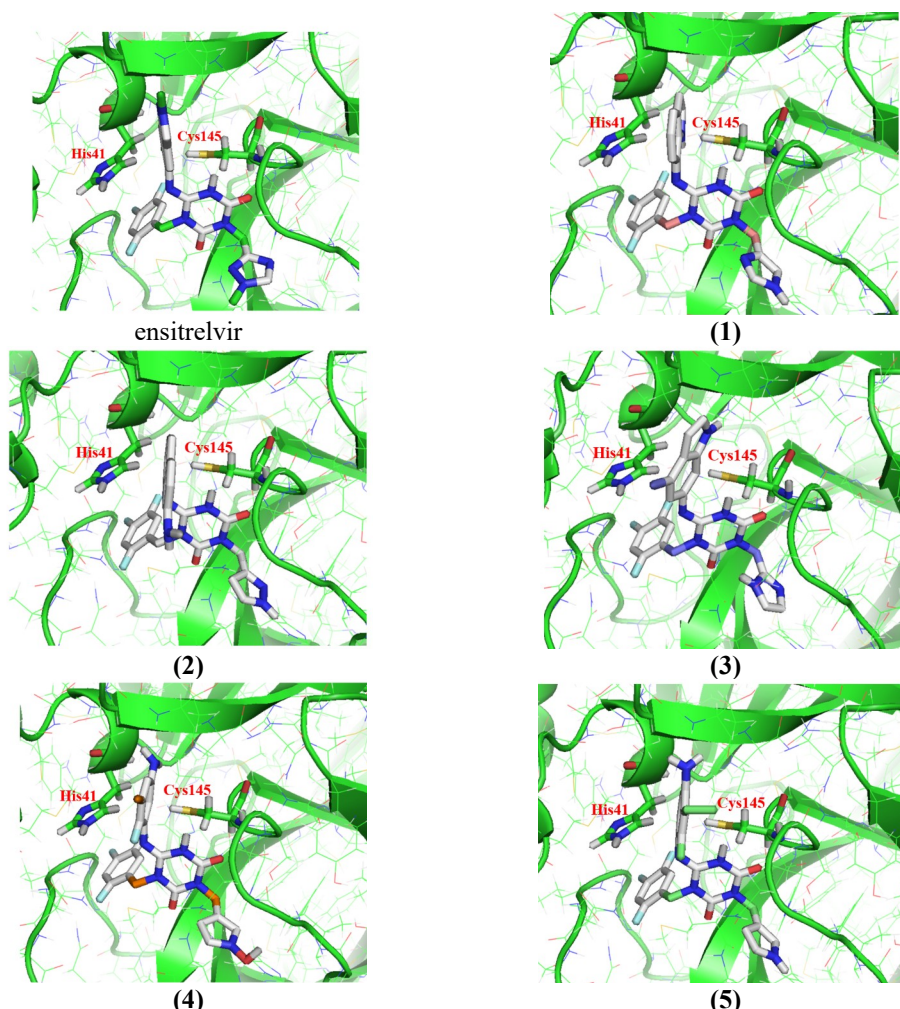


Figure 8. The best binding modes of ensitrelvir and its analogues **1-5** with the active center of M_{pro} protease of the SARS-CoV-2 virus (PDB: 8HEF).

The results presented in Figures 7 and 8 indicate that the increase in the binding energy and the enhancement of the overall affinity score result from paired structural modifications of the initial structure of ensitrelvir (**1**). These changes include: replacing the methyl-1,2,4-triazole and 6-chloro-2-methyl-2H-indazole with a pyrazole ring and aminonaphthol (ligand **1**), an imidazole ring and an indole ring (ligand **2**), chlorindole and an imidazole ring (ligand **5**), fluoromethylaniline and a hydroxypyrrole ring (ligand **3**), and ethylmethylaniline and a pyrrole ring (ligand **4**). These modifications significantly contribute to enhancing the antiviral efficacy of the ensitrelvir derivatives.

Conclusions

Modern methods for discovering new potential drugs increasingly rely on computer modeling techniques [14, 26, 30, 34-36]. In this study, design of novel analogues of ensitrelvir was conducted by employing evolutionary structure optimization of the parent ensitrelvir scaffold as summarized in Fig. 9.

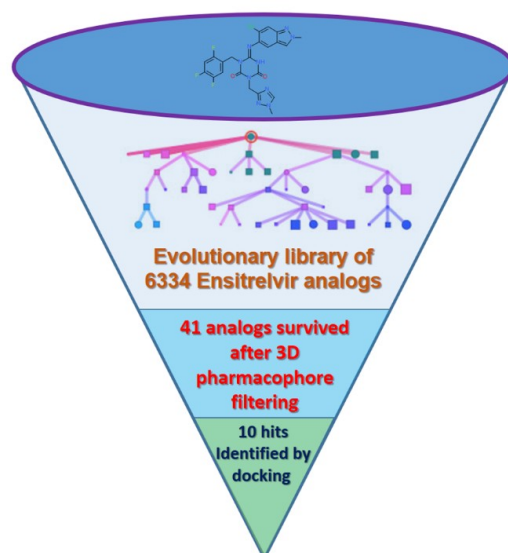


Figure 9. Workflow of discovery of novel ensitrelvir analogs

A library of 6,334 structurally similar ensitrelvir derivatives was created using a set of fitness criteria, including molecular weight (M_w), cLogP, polar surface area, structural and conformational similarity, flexibility, and molecular shape. Further filtering of this library based on its correspondence to a 3D pharmacophore model of ensitrelvir led to the identification of 41 promising derivatives. Subsequently, a detailed structural and energetic analysis of the binding affinity of these candidate derivatives to the M^{pro} protease was conducted using molecular docking methods. This analysis allowed us to select 10 ensitrelvir analogues that exhibit a stronger binding affinity to the M^{pro} protease compared to the original ensitrelvir inhibitor. Finally, we demonstrated for the first time that the key structural modifications responsible for enhancing the overall antiviral effectiveness of the inhibitor are found in the 1-methyl-1H-1,2,4-triazole and 6-chloro-2-methyl-2H-indazole fragments of ensitrelvir. We believe that our proposed workflow for evolutionary optimization of existing scaffolds of approved drugs is a promising, cost-effective strategy for developing new antiviral agents against COVID-19 [37].

Acknowledgement

The authors acknowledge the grant 2021.01/0062 “*Molecular design, synthesis and screening of new potential antiviral pharmaceutical ingredients for the treatment of infectious diseases COVID-19*” from the National Research Foundation of Ukraine.

The authors thank Prof. T. Langer from University of Vienna for giving us the opportunity to use LigandScout 4.4.9 suite.

References

1. Yevsieieva L. V., Lohachova K. O., Kyrychenko A., Kovalenko S. M., Ivanov V. V., Kalugin O. N. Main and papain-like proteases as prospective targets for pharmacological treatment of coronavirus SARS-CoV-2. *RSC Adv.* **2023**, *13* (50), 35500–35524. <https://doi.org/10.1039/d3ra06479d>
2. Polatoğlu I., Oncu-Oner T., Dalman I., Ozdogan S. COVID-19 in early 2023: Structure, replication mechanism, variants of SARS-CoV-2, diagnostic tests, and vaccine & drug development studies. *MedComm* **2023**, *4* (2), e228. <https://doi.org/10.1002/mco2.228>
3. Lau J. J., Cheng S. M. S., Leung K., Lee C. K., Hachim A., Tsang L. C. H., Yam K. W. H., Chaothai S., Kwan K. K. H., Chai Z. Y. H., Lo T. H. K., Mori M., Wu C., Valkenburg S. A., Amarasinghe G. K., Lau E. H. Y., Hui D. S. C., Leung G. M., Peiris M., Wu J. T. Real-world COVID-19 vaccine effectiveness against the Omicron Ba.2 variant in a SARS-CoV-2 infection-naïve population. *Nat. Med.* **2023**, *29* (2), 348–357. <https://doi.org/10.1038/s41591-023-02219-5>
4. Fischer C., Willscher E., Paschold L., Gottschick C., Klee B., Diexer S., Bosurgi L., Dutzmann J., Sedding D., Frese T., Girndt M., Hoell J. I., Gekle M., Addo M. M., Schulze zur Wiesch J.,

- Mikolajczyk R., Binder M., Schultheiß C. SARS-CoV-2 vaccination may mitigate dysregulation of IL-1/IL-18 and gastrointestinal symptoms of the post-COVID-19 condition. *npj Vaccines* **2024**, 9 (1), 23. <https://doi.org/10.1038/s41541-024-00815-1>
5. Carabelli A. M., Peacock T. P., Thorne L. G., Harvey W. T., Hughes J., de Silva T. I., Peacock S. J., Barclay W. S., de Silva T. I., Towers G. J., Robertson D. L., Consortium C.-G. U. SARS-CoV-2 variant biology: Immune escape, transmission and fitness. *Nat. Rev. Microbiol.* **2023**, 21 (3), 162-177. <https://doi.org/10.1038/s41579-022-00841-7>
 6. Markov P. V., Ghafari M., Beer M., Lythgoe K., Simmonds P., Stilianakis N. I., Katzourakis A. The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **2023**, 21 (6), 361-379. <https://doi.org/10.1038/s41579-023-00878-2>
 7. Uriu K., Ito J., Zahradnik J., Fujita S., Kosugi Y., Schreiber G., Sato K. Enhanced transmissibility, infectivity, and immune resistance of the SARS-CoV-2 omicron Xbb.1.5 variant. *Lancet. Infect. Dis.* **2023**, 23 (3), 280-281. [https://doi.org/10.1016/S1473-3099\(23\)00051-8](https://doi.org/10.1016/S1473-3099(23)00051-8)
 8. Wang Z., Yang L. Post-acute sequelae of SARS-CoV-2 infection: A neglected public health issue. *Front. Public Health* **2022**, 10, 908757. <https://doi.org/10.3389/fpubh.2022.908757>
 9. Davis H. E., McCorkell L., Vogel J. M., Topol E. J. Long COVID: Major findings, mechanisms and recommendations. *Nat. Rev. Microbiol.* **2023**, 21 (3), 133-146. <https://doi.org/10.1038/s41579-022-00846-2>
 10. Blomberg B., Mohn K. G.-I., Brokstad K. A., Zhou F., Linchausen D. W., Hansen B.-A., Lartey S., Onyango T. B., Kuwelker K., Sævik M., Bartsch H., Tøndel C., Kittang B. R., Madsen A., Bredholt G., Vahokoski J., Fjellveit E. B., Bansal A., Trieu M. C., Ljostveit S., Olofsson J. S., Ertesvåg N., Sandnes H. H., Corydon A., Søyland H., Eidsheim M., Jakobsen K., Guldseth N., Hauge S., Cox R. J., Langeland N., Bergen C.-R. G. Long COVID in a prospective cohort of home-isolated patients. *Nat. Med.* **2021**, 27 (9), 1607-1613. <https://doi.org/10.1038/s41591-021-01433-3>
 11. Walker A. P., Fan H., Keown J. R., Knight M. L., Grimes Jonathan M., Fodor E. The SARS-CoV-2 RNA polymerase is a viral rna capping enzyme. *Nucleic Acids Res.* **2021**, 49 (22), 13019-13030. <https://doi.org/10.1093/nar/gkab1160>
 12. Yin W., Mao C., Luan X., Shen D.-D., Shen Q., Su H., Wang X., Zhou F., Zhao W., Gao M., Chang S., Xie Y.-C., Tian G., Jiang H.-W., Tao S.-C., Shen J., Jiang Y., Jiang H., Xu Y., Zhang S., Zhang Y., Xu H. E. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by Remdesivir. *Science* **2020**, 368 (6498), 1499. <https://doi.org/10.1126/science.abc1560>
 13. Poduri R., Joshi G., Jagadeesh G. Drugs targeting various stages of the SARS-CoV-2 life cycle: Exploring promising drugs for the treatment of COVID-19. *Cellular Signalling* **2020**, 74 109721. <https://doi.org/10.1016/j.cellsig.2020.109721>
 14. Zakharov A. B., Kyrpa M., Kyrychenko A. V., Kovalenko S. M., Kalugin O. N., Ivanov V. V., Adamowicz L. Towards the computational design of organic molecules with specified properties. *Struct. Chem.* **2025**, 36 (2), 723-738. <https://doi.org/10.1007/s11224-024-02441-y>
 15. , Lagarias P., Gao A., Townsend J. A., Meng X., Dube P., Zhang X., Hu Y., Kitamura N., Hurst B., Tarbet B., Marty M. T., Kolocouris A., Xiang Y., Chen Y., Wang J. Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and Cathepsin L. *Sci. Adv.* **2020**, 6 (50), eabe0751. <https://doi.org/10.1126/sciadv.abe0751>
 17. Unoh Y., Uehara S., Nakahara K., Nobori H., Yamatsu Y., Yamamoto S., Maruyama Y., Taoda Y., Kasamatsu K., Suto T., Kouki K., Nakahashi A., Kawashima S., Sanaki T., Toba S., Uemura K., Mizutare T., Ando S., Sasaki M., Orba Y., Sawa H., Sato A., Sato T., Kato T., Tachibana Y. Discovery of S-217622, a noncovalent oral SARS-CoV-2 3CL protease inhibitor clinical candidate for treating COVID-19. *J. Med. Chem.* **2022**, 65 (9), 6499-6512. <https://doi.org/10.1021/ac-sjmedchem.2c00117>
 18. Noske G. D., de Souza Silva E., de Godoy M. O., Dolci I., Fernandes R. S., Guido R. V. C., Sjö P., Oliva G., Godoy A. S. Structural basis of Nirmatrelvir and Ensitretevir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease. *J. Biol. Chem.* **2023**, 299 (3). <https://doi.org/10.1016/j.jbc.2023.103004>
 19. Duan Y., Zhou H., Liu X., Iketani S., Lin M., Zhang X., Bian Q., Wang H., Sun H., Hong S. J., Culbertson B., Mohri H., Luck M. I., Zhu Y., Liu X., Lu Y., Yang X., Yang K., Sabo Y., Chavez

- A., Goff S. P., Rao Z., Ho D. D., Yang H. Molecular mechanisms of SARS-CoV-2 resistance to Nirmatrelvir. *Nature* **2023**, 622 (7982), 376-382. <https://doi.org/10.1038/s41586-023-06609-0>
20. Kovalevsky A., Aniana A., Ghirlando R., Coates L., Drago V. N., Wear L., Gerlits O., Nashed N. T., Louis J. M. Effects of SARS-CoV-2 main protease mutations at positions L50, E166, and L167 rendering resistance to covalent and noncovalent inhibitors. *J. Med. Chem.* **2024**, 67 (20), 18478-18490. <https://doi.org/10.1021/acs.jmedchem.4c01781>
21. Lopez U. M., Hasan M. M., Havranek B., Islam S. M. SARS-CoV-2 resistance to small molecule inhibitors. *Curr. Clin. Micro Rpt.* **2024**, 11 (3), 127-139. <https://doi.org/10.1007/s40588-024-00229-6>
22. Krismer L., Schöppe H., Rauch S., Bante D., Sprenger B., Naschberger A., Costacurta F., Fürst A., Sauerwein A., Rupp B., Kaserer T., von Laer D., Heilmann E. Study of key residues in MERS-CoV and SARS-CoV-2 main proteases for resistance against clinically applied inhibitors nirmatrelvir and ensitrelvir. *npj Viruses* **2024**, 2 (1), 23. <https://doi.org/10.1038/s44298-024-00028-2>
23. Lin C., Jiang H., Li W., Zeng P., Zhou X., Zhang J., Li J. Structural basis for the inhibition of coronaviral main proteases by Ensitrelvir. *Structure* **2023**, 31 (9), 1016-1024.e3. <https://doi.org/10.1016/j.str.2023.06.010>
24. Kiso M., Yamayoshi S., Iida S., Furusawa Y., Hirata Y., Uraki R., Imai M., Suzuki T., Kawaoka Y. *In vitro* and *in vivo* characterization of SARS-CoV-2 resistance to Ensitrelvir. *Nat. Commun.* **2023**, 14 (1), 4231. <https://doi.org/10.1038/s41467-023-40018-1>
25. Sander T., Freyss J., von Korff M., Rufener C. Datawarrior: An open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model.* **2015**, 55 (2), 460-473. <https://doi.org/10.1021/ci500588j>
26. Lohachova K. O., Sviatenko A. S., Kyrychenko A., Ivanov V. V., Langer T., Kovalenko S. M., Kalugin O. N. Computer-aided drug design of novel nirmatrelvir analogs inhibiting main protease of coronavirus SARS-COV-2. *J. Appl. Pharm. Sci.* **2024**, 14 (5), 232-239. <https://doi.org/10.7324/JAPS.2024.158114>
27. Wolber G., Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* **2005**, 45 (1), 160-169. <https://doi.org/10.1021/ci049885e>
28. Trott O., Olson A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, 31 (2), 455-461. doi: <https://doi.org/10.1002/jcc.21334>
29. Humphrey W., Dalke A., Schulten K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, 14 (1), 33-38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
30. Ivanov V., Lohachova K., Kolesnik Y., Zakharov A., Yevsieieva L., Kyrychenko A., Langer T., Kovalenko S. M., Kalugin O. M. Recent advances in computational drug discovery for therapy against coronavirus SARS-CoV-2. *ScienceRise: Pharm. Sci.* **2023**, 6(46), 4-24. <https://doi.org/10.15587/2519-4852.2023.290318>
31. Ertl P., Rohde B., Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, 43 (20), 3714-3717. <https://doi.org/10.1021/jm000942e>
32. Yevsieieva L., Trostianko P., Kyrychenko A., Ivanov V., Kovalenko S., Kalugin O. Design of non-covalent dual-acting inhibitors for proteases M^{pro} and PL^{pro} of coronavirus SARS-CoV-2 through evolutionary library generation, pharmacophore profile matching, and molecular docking calculations. *Sci. Rise. Pharm. Sci.* **2024**, 6(52), 15-26. <https://doi.org/10.15587/2519-4852.2024.313808>
33. Ruiz-Moreno A. J., Cedillo-González R., Cordova-Bahena L., An Z., Medina-Franco J. L., Velasco-Velázquez M. A. Consensus pharmacophore strategy for identifying novel SARS-CoV-2 M^{pro} inhibitors from large chemical libraries. *J. Chem. Inf. Model.* **2024**, 64 (6), 1984-1995. <https://doi.org/10.1021/acs.jcim.3c01439>
34. Anokhin D., Kovalenko S., Trostianko P., Kyrychenko A., Zakharov A., Zubatiuk T., Ivanov V., Kalugin O. Towards the discovery of molecules with anti-COVID-19 activity: Relationships between screening and docking results. *Kharkiv University Bulletin. Chemical Series* **2024**, (42), 6-14. <https://doi.org/10.26565/2220-637X-2024-42-01>

35. Kyrychenko A., Bylov I., Geleverya A., Kovalenko S., Zhuravel I., Fetyukhin V., Langer T. Computer-aided rational design and synthesis of new potential antihypertensive agents among 1,2,3-triazole-containing nifedipine analogs. *SciRise: Pharm. Sci.* **2024**, 49 (3), 4-12. <https://doi.org/10.15587/2519-4852.2024.291626>
36. Zahrychuk H. Y., Gladkov E. S., Kyrychenko A. V., Poliovyi D. O., Zahrychuk O. M., Kucher T. V., Logoyda L. S. Structure-based rational design and virtual screening of valsartan drug analogs towards developing novel inhibitors of angiotensin II type 1 receptor. *Biointerface Res. Appl. Chem.* **2023**, 13 (5), 440. <https://doi.org/10.33263/BRIAC135.440>
37. Vincenzi M., Mercurio A. F., Leone M. Looking for SARS-CoV-2 therapeutics through computational approaches. *Curr. Med. Chem.* **2023**, 30 (28), 3158-3214. <http://dx.doi.org/10.2174/0929867329666221004104430>

Received 27.09.2024

Accepted 19.12.2024

К.О. Логачова, А. С. Святенко, О. В. Кириченко, О. М. Калугін. Еволюційна оптимізація будови енсїтрелвіру як нековалентного інгібітора основної протеази SARS-COV-2 M^{pro}.

Харківський національний університет імені В.Н. Каразіна, хімічний факультет, майдан Свободи, 4, Харків, 61022, Україна

Енсїтрелвір є нековалентним непептидним інгібітором основної протеази M^{pro} вірусу SARS-CoV-2. Він продемонстрував ефективну противірусну дію проти різних варіантів коронавірусу *in vitro*, а також сприятливий метаболізм і фармакокінетичні профілі, придатні для перорального лікування. Таким чином, розробка нових аналогів енсїтрелвіру має велике значення. У цьому дослідженні ми провели *in silico* дизайн його аналогів, використовуючи еволюційну оптимізацію структури батьківського скелета енсїтрелвіру. На першому етапі ми створили віртуальну еволюційну бібліотеку, що складається з 6334 нових аналогів на основі ряду критеріїв придатності, включаючи молекулярну масу (M_w), cLogP, площу полярної поверхні, структурну та конформаційну подібність, гнучкість і молекулярну форму. Далі ми відфільтрували еволюційну бібліотеку за допомогою 3D-моделі фармакофора, створеної з доступної рентгенівської структури спільно кристалізованого комплексу енсїтрелвіру та M^{pro}. Потім ми виконали розрахунки молекулярного докінгу, щоб класифікувати обраних кандидатів відповідно до їх афінності зв'язування та селективності до рецептора M^{pro}. Цей рейтинг зв'язування дозволив нам ідентифікувати десять аналогів енсїтрелвіру, які виявляють кращу афінність зв'язування з протеазою M^{pro} порівняно з вихідним інгібітором енсїтрелвіром. Наша еволюційна оптимізація структури вказує на те, що первинні структурні модифікації, які посилюють загальний противірусний ефект енсїтрелвіру, розташовані у структурних фрагментах 1-метил-1H-1,2,4-тріазолу та 6-хлор-2-метил-2H-індазолу.

Ключові слова: коронавірус, КОВІД-19, гетероциклічні сполуки, M^{pro}, еволюційна бібліотека, молекулярний докінг

Надіслано до редакції 27.09.2024

Прийнято до друку 19.12.2024

Kharkiv University Bulletin. Chemical Series. Issue 43 (66), 2024

СТРУКТУРА ТА ТРАНСПОРТНІ ВЛАСТИВОСТІ РОЗЧИНІВ NaPF_6 У СУМІШАХ ЕТИЛЕНКАРБОНАТУ З ДИМЕТИЛКАРБОНАТОМ ДЛЯ НАТРІЙ-ІОННИХ АКУМУЛЯТОРІВ: МД МОДЕЛЮВАННЯ

А.В. Боровик^а, Я.В. Колесник^б, О.М. Калугін^с

Харківський національний університет імені В.Н. Каразіна, хімічний факультет, кафедра неорганічної хімії, майдан Свободи, 4, Харків, 61022, Україна

a) ✉ nassty.ja@gmail.com

b) ✉ ykolesnik@karazin.ua

c) ✉ onkalugin@gmail.com

 <https://orcid.org/0009-0001-0095-4424>

 <https://orcid.org/0000-0002-9569-4556>

 <https://orcid.org/0000-0003-3273-9259>

Натрій-іонні акумулятори (SIBs) потенційно можуть стати новими ефективними пристроями зберігання електричної енергії. Однак, на даний момент основною проблемою є відсутність чіткої технології їх виготовлення. Для промислового виробництва SIBs необхідно розробити катодні та анодні матеріали, а також підібрати оптимальний склад електроліту. З цією метою за допомогою методів молекулярно-динамічного моделювання нами було розраховано густину, в'язкість, електропровідність та коефіцієнти дифузії для систем NaPF_6 у бінарному розчиннику EC:DMC (15:85 wt%, 30:70 wt% та 50:50 wt%), а також розглянуто їх структурні властивості.

Структура сольватної оболонки катіонів та аніонів була вивчена в рамках функцій радіального розподілу та поточних координатних чисел. Результати вказують на більш структуровану сольватну оболонку катіонів Na^+ , аніж аніонів PF_6^- .

Дослідження транспортних властивостей показало, що найбільш підходящими електролітами для виготовлення натрій-іонних акумуляторів є системи, у яких EC:DMC=15:85 wt%. Це пов'язано з тим, що електроліт саме цього складу виявив найменші величини в'язкості в області всіх концентрацій, а також найвищі значення електропровідності. Коефіцієнти дифузії Na^+ для цієї системи також набувають найвищих значень у порівнянні з електролітами іншого складу, що є переконливим аргументом у її майбутньому застосуванні у натрій-іонних акумуляторах.

Ключові слова: молекулярно-динамічне моделювання, натрій-іонні акумулятори, натрій гексафторофосфат, етиленкарбонат, диметилкарбонат, структурні властивості, транспортні властивості, MDNAES.

Вступ

На сьогоднішній день наука розвивається дуже стрімко, що призводить до винаходу передових технологій у різних галузях. Зокрема, велику увагу привернула методика виробництва та зберігання енергії для повсякденного застосування. За останні роки набули популярності літій-іонні акумулятори (LIBs), які завдяки високій енергетичній щільності, невеликій вазі та малому саморозрядженню широко застосовують у більшості мобільних електронних пристроях та транспортних засобах з електричним двигуном. Однак, одним з недоліків цього виду акумуляторів є обмежена доступність Літію, наслідком чого є зростання цін на LIBs. Таким чином, постає актуальне питання щодо створення альтернативних пристроїв зберігання електричної енергії, які б задовольняли потреби споживачів на високому рівні.

Натрій-іонні акумулятори (SIBs) – це нове покоління акумуляторів, яке наразі вважається найкращою альтернативою LIBs, оскільки Натрій широко доступний та має схожі хімічні властивості з Літієм. Велика кількість ресурсів і набагато нижча вартість мінеральної сировини, з якої виробляється карбонат натрію, порівняно з карбонатом літію, є переконливим обґрунтуванням для розвитку SIBs. Також слід зазначити, що, на відміну від літій-іонних акумуляторів, SIBs можна зберігати у повністю розрядженому стані, що зведе до мінімуму небезпеку їх транспортування. На сьогоднішній день немає загально визнаних рекомендацій щодо матеріалів для виготовлення натрій-іонних акумуляторів, хоча, в основному, компоненти батареї та механізм накопичення електроенергії SIBs та LIBs є однаковими, за винятком носіїв іонів.

По-перше, актуальним питанням для промислового виробництва натрій-іонних акумуляторів залишається правильний вибір матеріалу електродів. По-друге, необхідно підібрати тип

електроліту так, щоби досягти максимальної практичності у поєднанні з безпекою акумуляторів. Все це потребує досить багато часу та масштабного фінансування.

Молекулярно-динамічне моделювання є альтернативним вирішенням вище поставленої проблеми, оскільки має ряд переваг над дослідженнями у лабораторіях, такі як швидкість виконання експерименту, а також його дешевизна.

У цій статті розглядається питання створення оптимального складу органічного рідкого електроліту, який потенційно можна використовувати у натрій-іонних акумуляторах. У якості об'єктів дослідження нами було обрано три системи на основі солі натрій гексафторофосфату(V), розчиненій у суміші етиленкарбонату та диметилкарбонату у різних співвідношеннях. Наш вибір даних систем можна пояснити тим, що у випадку LIBs електроліт саме такого складу є найбільш популярним, завдяки його високій іонній провідності при досить низьких значеннях в'язкості. Тому вивчення об'єкту, у якому Літій замінено на Натрій викликає сильний науковий інтерес. Отже, основною метою даної роботи було прогнозування електричної провідності розчинів NaPF₆ у змішаних розчинниках.

Методологія

Деталі молекулярно-динамічного моделювання. Для МД моделювання NaPF₆ у бінарній суміші EC:DMC було вибрано системи різних концентрацій солі із співвідношенням молекул розчинників 50:50, 30:70 та 15:85 за масою. Кількість частинок катіону, аніону та молекул розчинника наведено у Таблиці 1. Кількість молекул EC та DMC для кожної системи розраховувалась відповідно до заданих масових співвідношень, із сумарною кількістю близько 250 молекул.

Всі розрахунки реалізовані у програмному пакеті MDNAES. Моделювання було проведено у *NPT* ансамблі при температурі 298.15 K та тиску 1 атм за допомогою термостату і баростату Берендсена. Час релаксації становив $\tau_T = 5$ пс та $\tau_P/\beta = 2 \cdot 10^6$ пс·атм, відповідно. Тривалість підготовчого запуску моделювання складала 10 пс. Далі врівноваження кожної системи проводилось протягом 1 нс з кроком інтегрування 2 фс (для внутрішньомолекулярних потенціалів 0.25 фс). Розрахунок властивостей систем на одному етапі тривав 2 нс, крок інтегрування – 2 фс. Для зниження випадкової похибки визначення кількісних характеристик модельованих систем для кожної із них було проведено 50 незалежних послідовних моделювань сумарною тривалістю 100 нс.

У таблиці 1 наведено експериментальні величини густини, розраховані з даних статті [12] шляхом інтерполяції, яка була використана в роботі для валідації використаних моделей.

Таблиця 1. Склад модельованих систем NaPF₆ у бінарній суміші EC:DMC
Table 1. Composition of modeled systems of NaPF₆ in binary mixtures of EC:DMC

c(NaPF ₆), моль/л	Кількість Na ⁺	Кількість PF ₆ ⁻	Кількість EC	Кількість DMC	Експериментальна густина ρ , кг/м ³
0.25	5	5	EC:DMC = 50:50 wt% 126	123	1230
0.50	10	10			1250
0.75	15	15			1270
1.00	20	20	EC:DMC = 30:70 wt% 76	173	1289
1.25	25	25			1309
1.50	30	30			1329
1.75	35	35	EC:DMC = 15:85 wt% 40	210	1349
2.00	40	40			1368
2.25	45	45			1388

Структурні властивості NaPF₆ у бінарних сумішах EC:DMC були проаналізовані за допомогою функцій радіального розподілу (ФРП) та поточного координаційного числа (ПКЧ). Дані функції були розраховані за рівняннями 1 та 2 відповідно.

$$g_{ij}(r) = \frac{V}{N} \frac{n_j(r)}{4\pi r^2 \Delta r} \quad (1)$$

$$n_{ij}(r) = \frac{N_j - \delta_{ij}}{V} \int_0^r 4\pi r^2 g_{ij}(r) dr \quad (2)$$

Де $n_{ij}(r)$ — середня кількість атомів сорту j , що знаходяться на відстані $r..r+\Delta r$ від центрального атому (сорту i).

У роботі були розраховані електрична провідність, в'язкість та коефіцієнти дифузії для розчинів.

Електричну провідність обчислювали за рівнянням Гріна-Кубо (рівняння 3) з автокореляційної функції миттєвого струму для рівноважної системи.

В'язкість обчислювали за рівнянням Гріна-Кубо (рівняння 4) з автокореляційної функції позадіагональних елементів тензора напружки для рівноважної системи.

Коефіцієнти дифузії були розраховані за рівнянням Ейнштейна (рівняння 5) на основі функцій середньоквадратичного зміщення.

$$\kappa = \frac{1}{3Vk_B T} \int_0^\infty C_{jj}(t) dt \quad (3)$$

$$\eta = \frac{1}{Vk_B T} \int_0^\infty dt \langle \mathbf{J}^{\alpha\beta}(0) \mathbf{J}^{\alpha\beta}(t) \rangle \quad (4)$$

$$\lim_{t \rightarrow \infty} \langle (r(0) - r(t))^2 \rangle = 6Dt \quad (5)$$

Силкові поля. Структуру аніона PF_6^- та молекул розчинника ЕС та DMC представлено на Рисунку 1.

Для катіону натрію Na^+ було використано модель з OPLS-AA.

Для аніону PF_6^- параметри були взяті з моделі, запропонованої С. Lopes та ін.

Для молекул етиленкарбонату та диметилкарбонату силкові поля були взяті із роботи Дударева зі співавторами. Задля покращення результатів моделювання властивостей відповідних розчинників автори роботи оптимізували параметри узяті з моделі OPLS-AA.

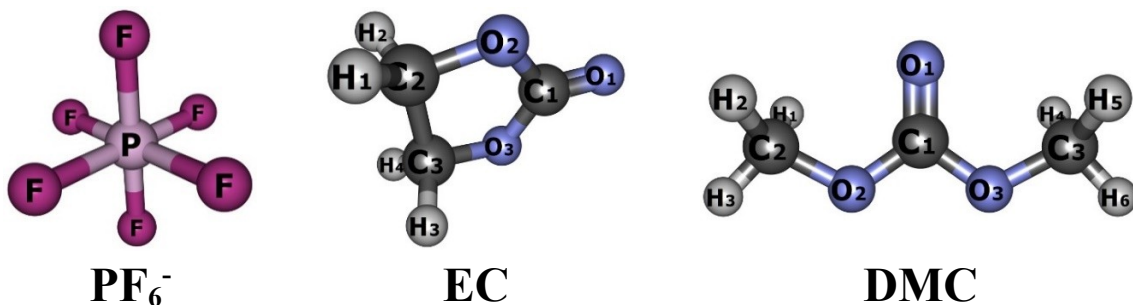


Рисунок 1. Структура аніона PF_6^- та молекул розчинника ЕС та DMC
Figure 1. Structure of the PF_6^- anion and solvent molecules EC and DMC

Для валідації обраних моделей силових полів нами була використана густина. З Рисунку 2 бачимо добре співпадіння величин, отриманих за допомогою МД; моделювання з експериментальними для системи ЕС:DMC=50:50. Спираючись на це, ми можемо говорити про успішне подальше моделювання без змін потенціальних полів досліджуваних частинок.

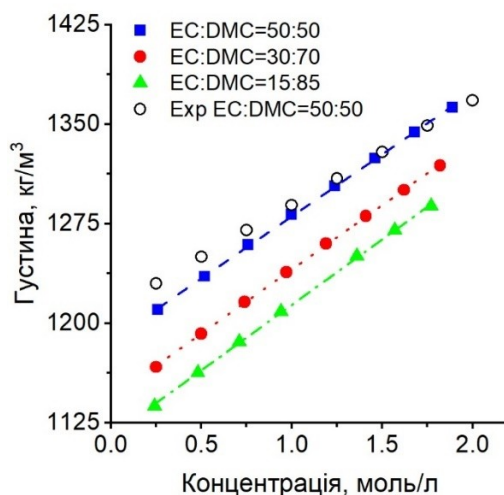


Рисунок 2. Концентраційна залежність густини розчинів NaPF_6 у EC:DMC у порівнянні з експериментальною

Figure 2. Concentration dependence of the density of the solutions of NaPF_6 in EC:DMC in comparison with the experimental one

Результати та обговорення

Структурні властивості. Структурні властивості NaPF_6 у бінарних сумішах EC:DMC були проаналізовані за допомогою функцій радіального розподілу (ФРР) та поточного координаційного числа (ПКЧ).

Для спрощення аналізували тільки ФРР між координаційними центрами іонів та молекул. На Рисунок 3 показано координаційні центри етиленкарбонату та диметилкарбонату. При аналізі ФРР пари катіон-розчинник у якості координаційного центру EC та DMC виступає атом O_1 , оскільки він має найбільш негативний заряд. В іншому випадку, для пари аніон-розчинник, координаційними центрами слугують атоми C_2/C_3 з найбільшим позитивним зарядом (дані атоми симетричні). Координаційним центром PF_6^- є атом Фосфору.

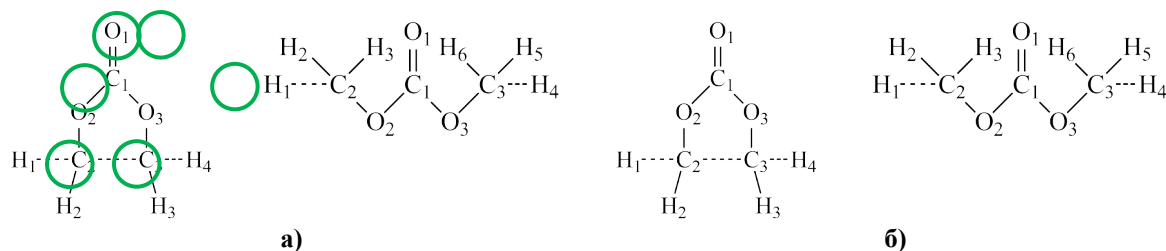


Рисунок 3. Координаційні центри EC та DMC для ФРР і ПКЧ а) катіон-розчинник; б) аніон-розчинник
Figure 3. EC and DMC coordination centers for RDF and RCN of a) cation-solvent; b) anion-solvent

ФРР та ПКЧ між Na^+ та атомами O_1 EC та DMC для всіх досліджуваних систем складу EC:DMC = 50:50 wt% представлені на Рисунок 4. Для систем із співвідношенням розчинників 30:70 та 15:85 були отримані такі самі результати.

З ФРР бачимо, що положення максимумів кривих, а також ширина піків залишаються однаковими при різних концентраціях солі. Також спостерігаємо високі та вузькі піки, що вказують на досить міцний зв'язок катіона з молекулами розчинника, про що також свідчить наявність плато на ПКЧ.

Дослідивши перший максимум ФРР ми можемо передбачити відстань між обраними двома атомами у першій сольватній оболонці. Так, для пари $\text{Na}-\text{O}_{1,\text{EC}}$ та $\text{Na}-\text{O}_{1,\text{DMC}}$ максимум функції спостерігається при 0.238 нм.

Положення першого мінімуму вказує на розмір першої сольватної оболонки і становить 0.361 у випадку пари $\text{Na}-\text{O}_{1,\text{EC}}$ та 0.371 для $\text{Na}-\text{O}_{1,\text{DMC}}$.

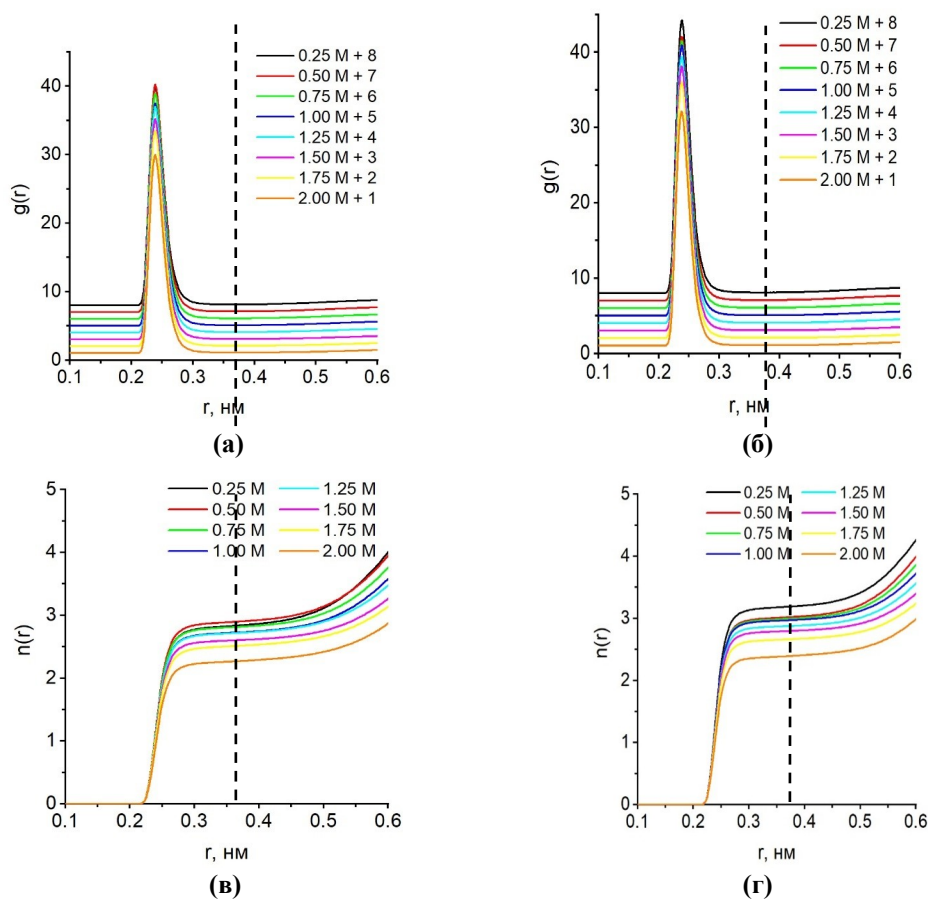


Рисунок 4. Функції радіального розподілу катіон-розчинник а) Na-O_{1,EC}, б) Na-O_{1,DMS} та поточні координаційні числа в) Na-O_{1,EC}, г) Na-O_{1,DMS}. Вертикальна пунктирна лінія відповідає першій сольватній оболонці.

Figure 4. Radial distribution functions of cation-solvent а) Na-O_{1,EC}, б) Na-O_{1,DMS} and running coordination numbers с) Na-O_{1,EC}, д) Na-O_{1,DMS}. The vertical dashed line corresponds to the first solvation shell.

За допомогою графіків поточного координаційного числа було визначено координаційні числа у парах іон-атом розчиннику. Для ФРР з диметилкарбонатом результати координаційних чисел вийшли більшими та становлять від 3.2 до 2.4, у той час як для ФРР з етиленкарбонатом КЧ варіюються від 2.9 до 2.3 зі збільшенням концентрації.

На Рисунку 5 представлені ФРР та ПКЧ між атомами Фосфору аніону PF₆⁻ та атомами C₂ ЕС та DMS для всіх досліджуваних систем зі складом розчинника 50:50.

Варто зазначити, що у цьому випадку як відстані між атомами, так і координаційні числа приймають більші значення. Останні варіюються у межах 5.1 - 4.1 для ФРР з ЕС та 3.4 – 2.9 для ФРР з DMS.

Проте, на відміну від пар катіон-розчинник, на ФРР аніона з молекулами ЕС та DMS піки приблизно в три рази нижче, а також більш широкі. Це говорить про те, що аніони дуже слабо сольватовані цими розчинниками, про що свідчить і відсутність плато на ПКЧ.

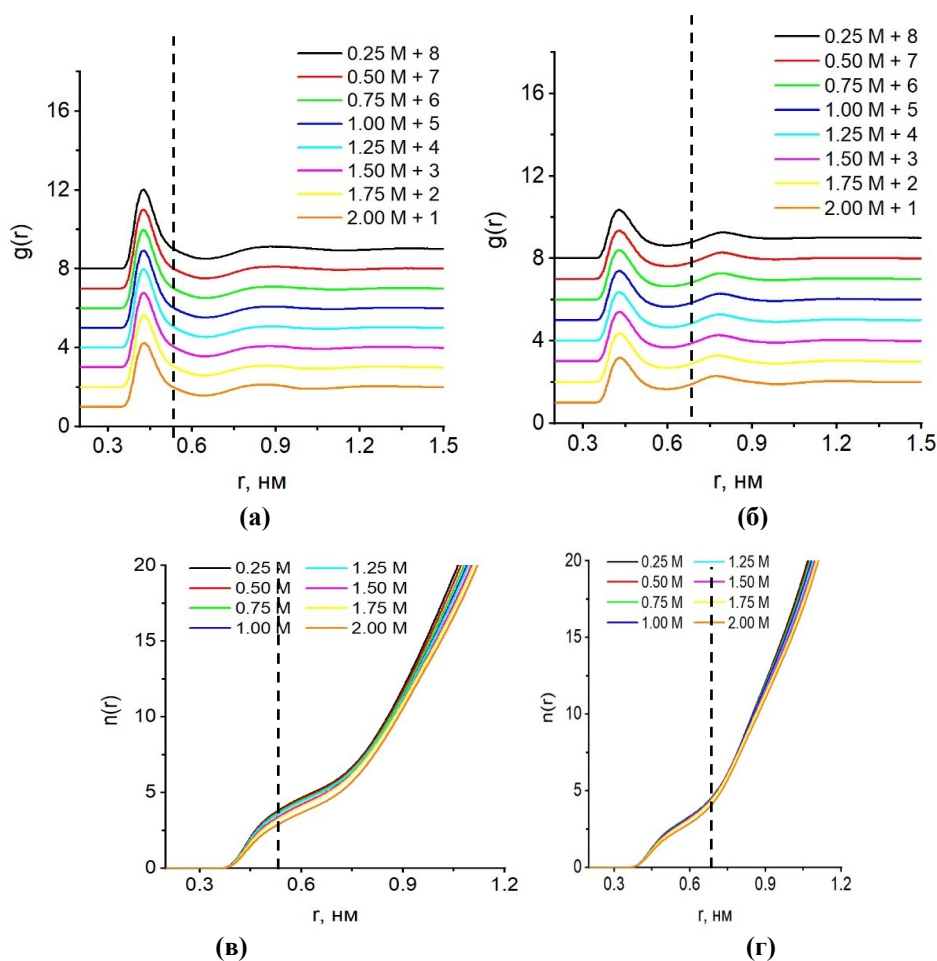


Рисунок 5. Функції радіального розподілу аніон-розчинник а) $\text{P-C}_{2,\text{EC}}$, б) $\text{P-C}_{2,\text{DMC}}$ та поточні координаційні числа в) $\text{P-C}_{2,\text{EC}}$, г) $\text{P-C}_{2,\text{DMC}}$. Вертикальна пунктирна лінія відповідає першій сольватній оболонці.
Figure 5. Radial distribution functions of anion-solvent a) $\text{P-C}_{2,\text{EC}}$, b) $\text{P-C}_{2,\text{DMC}}$ and running coordination numbers c) $\text{P-C}_{2,\text{EC}}$, d) $\text{P-C}_{2,\text{DMC}}$. The vertical dashed line corresponds to the first solvation shell.

Транспортні властивості. На першому етапі у якості транспортних властивостей було вивчено в'язкість та електропровідність (Рисунок 6 та 7).

Як ми бачимо, експериментально визначені величини в'язкості добре узгоджуються з розрахованими за допомогою МД моделювання, та їх зміна зі збільшенням концентрації відбувається майже експоненційно. У області всіх концентрацій нижчу в'язкість, має розчин зі співвідношенням $\text{EC:DMC}=15:85$ (wt %), а найвищу – $\text{EC:DMC}=50:50$ (wt %).

Зі збільшенням концентрації збільшується в'язкість, що перешкоджає збільшенню електропровідності за рахунок збільшення концентрації носіїв струму. У результаті спостерігається екстремальна залежність питомої провідності (яка прямо пропорційна концентрації носіїв струму) зі зміною концентрації. Збільшення концентрації носіїв струму та збільшення в'язкості фактично обумовлює різке падіння електропровідності в області концентрацій $\sim$ більше 1 моль/л.

Електропровідність – це величина, яка визначається за результатами МД моделювання із досить великою випадковою похибкою. З урахуванням рівня випадкової похибки у визначенні електропровідності отримані дані були описані найпростішою кривою з екстремумом – параболою, відповідно для всіх концентрацій максимум електропровідності приходить приблизно на концентрацію 1 моль/л.

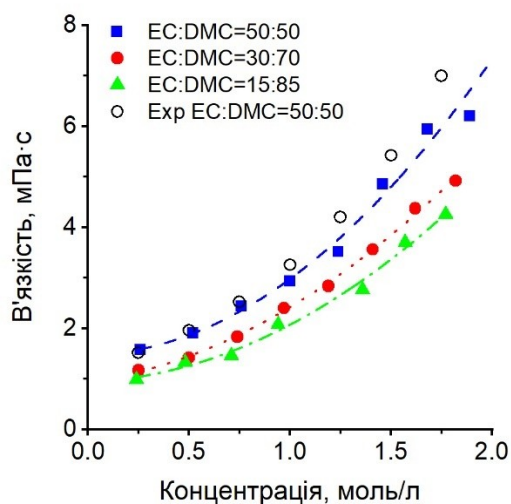


Рисунок 6. Концентраційна залежність в'язкості розчинів NaPF_6 у EC:DMC у порівнянні з експериментальною

Figure 6. Concentration dependence of viscosity of the solutions of NaPF_6 in EC:DMC in comparison with the experimental one¹²

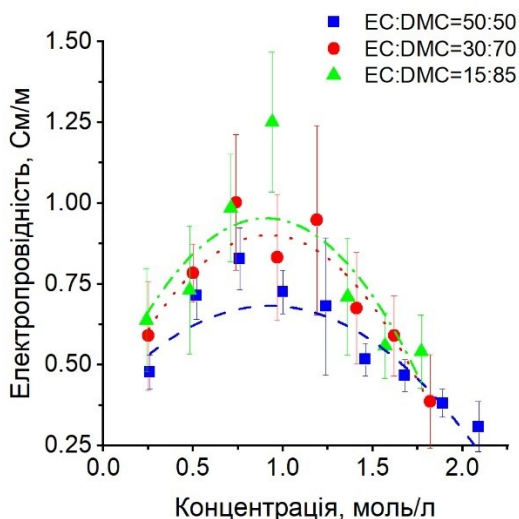


Рисунок 7. Концентраційна залежність питомої електропровідності розчинів NaPF_6 у EC:DMC

Figure 7. Concentration dependence of specific conductivity of the solutions of NaPF_6 in EC:DMC

Також, для вивчення транспортних властивостей досліджуваних розчинів, у ході молекулярно-динамічного моделювання були розраховані коефіцієнти дифузії (Рисунок 8).

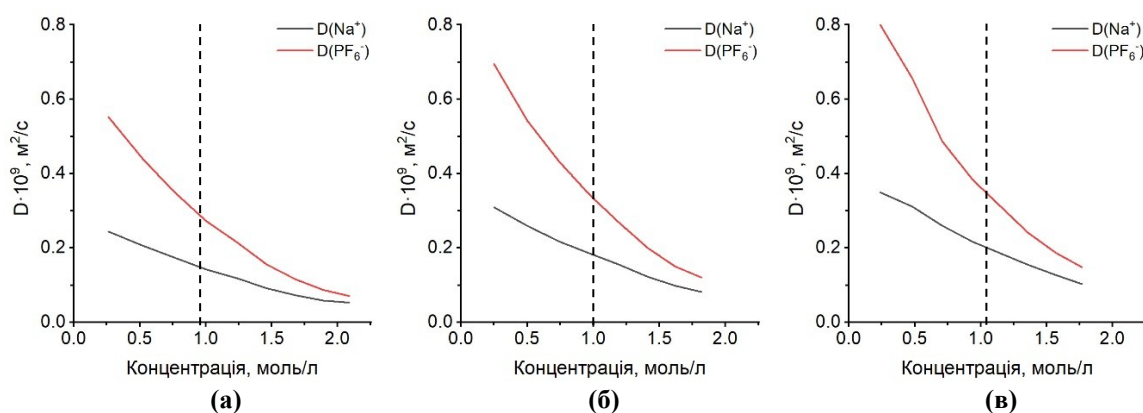


Рисунок 8. Залежність коефіцієнтів дифузії від концентрації NaPF_6 у EC:DMC а) 50:50 wt%; б) 30:70 wt%; в) 15:85 wt%.

Figure 8. Dependence of diffusion coefficients of NaPF_6 on concentration in EC:DMC а) 50:50 wt%; б) 30:70 wt%; в) 15:85 wt%.

Коефіцієнти дифузії прямо пропорційні рухливості іонів і чим вище коефіцієнт дифузії катіона в електролітному розчині, тим більше його число переносу, тобто тим краще використовувати такий електроліт в акумуляторі. У Таблиці 2 наведено результати розрахунку коефіцієнтів дифузії з випадковими похибками. Також нами було досліджено коефіцієнти дифузії Na^+ при концентрації 1 моль/л, оскільки саме ця концентрація відповідає максимуму електропровідності для кожної системи. З отриманих результатів бачимо, що у розчинах EC:DMC=30:70 та EC:DMC=15:85 коефіцієнт дифузії вище, ніж у EC:DMC=50:50, який традиційно використовується в акумуляторах.

Таблиця 2. Коефіцієнти дифузії іонів у системах NaPF₆ у EC:DMC
Table 2. Ion diffusion coefficients in systems of NaPF₆ in EC:DMC

Концентрація, моль/л	$D(\text{Na}^+) \cdot 10^9, \text{ м}^2/\text{с}$	$D(\text{PF}_6^-) \cdot 10^9, \text{ м}^2/\text{с}$
EC:DMC (50:50 wt%)		
0.26	0.245±0.009	0.552±0.019
0.52	0.206±0.006	0.442±0.015
0.76	0.175±0.005	0.352±0.010
1.00	0.143±0.004	0.273±0.007
1.24	0.119±0.006	0.215±0.011
1.46	0.092±0.003	0.156±0.003
1.68	0.073±0.002	0.117±0.004
1.89	0.059±0.002	0.087±0.007
2.09	0.053±0.004	0.071±0.007
EC:DMC (30:70 wt%)		
0.25	0.309±0.024	0.695±0.048
0.50	0.260±0.006	0.548±0.012
0.74	0.218±0.016	0.433±0.027
0.97	0.186±0.005	0.344±0.012
1.19	0.155±0.010	0.270±0.008
1.41	0.123±0.008	0.200±0.011
1.62	0.098±0.007	0.150±0.009
1.82	0.082±0.003	0.120±0.007
EC:DMC (15:85 wt%)		
0.24	0.350±0.024	0.800±0.049
0.48	0.312±0.021	0.658±0.032
0.71	0.260±0.014	0.486±0.028
0.94	0.217±0.010	0.384±0.021
1.36	0.155±0.004	0.241±0.013
1.57	0.127±0.006	0.187±0.008
1.77	0.103±0.004	0.148±0.009

Висновки

За допомогою молекулярно-динамічного моделювання у роботі було досліджено три системи сіль – бінарний розчинник, які потенційно можна використовувати у натрій-іонних акумуляторах. Обрані моделі молекул розчинника та іонів добре відтворюють експериментальну залежність величини густини та в'язкості.

Аналіз функцій радіального розподілу та поточних координаційних чисел показав, що, на відміну від аніонів, катіони добре сольватовані молекулами розчинників. Катіон натрію краще сольватується диметилкарбонатом, а поточні координаційні числа варіюються у межах від 2.4 до 3.2.

Дослідження транспортних властивостей показало, що максимальна електропровідність електроліту NaPF₆ – EC – DMC відповідає 1М розчину NaPF₆ у змішаному розчиннику EC:DMC у межах від 30:70 до 15:85 (wt %) та становить 0.90-0.95 См/м. Коефіцієнт дифузії Na⁺ у 1М розчині електроліту досягає більших значень для системи EC:DMC=15:85 та дорівнює $0.22 \cdot 10^{-9} \text{ м}^2/\text{с}$.

Подяки

Автори вдячні МОН України за фінансову підтримку в межах базового фінансування за НДР № 0121U112886. Комп'ютерні розрахунки були проведені з використанням розрахункового ресурсу *Dell EMC PowerEdge R740 supercomputer* Центру колективного використання наукового обладнання "Лабораторія мікро- та нано-систем, нових матеріалів та технологій" МОН України в Харківському національному університеті імені В.Н. Каразіна а також *Azzurra HPC center, Université Côte d'Azur, Nice, France*.

References

1. Kim, T.; Song, W.; Son, D.-Y.; Ono, L. K.; Qi, Y. Lithium-ion batteries: outlook on present, future, and hybridized technologies. *Journal of Materials Chemistry A* 2019, 7 (7), 2942-2964. <https://doi.org/10.1039/C8TA10513H>
2. Li, M.; Lu, J.; Chen, Z.; Amine, K. 30 years of lithium-ion batteries. *Advanced Materials* 2018, 30 (33), 1800561. <https://doi.org/10.1002/adma.201800561>
3. Hwang, J.-Y.; Myung, S.-T.; Sun, Y.-K. Sodium-ion batteries: present and future. *Chemical Society Reviews* 2017, 46 (12), 3529-3614. <https://doi.org/10.1039/C6CS00776G>
4. Adelhelm, P.; Hartmann, P.; Bender, C. L.; Busche, M.; Eufinger, C.; Janek, J. From lithium to sodium: cell chemistry of room temperature sodium–air and sodium–sulfur batteries. *Beilstein Journal of Nanotechnology* 2015, 6 (1), 1016-1055. <https://doi.org/10.3762/bjnano.6.105>
5. Hong, S. Y.; Kim, Y.; Park, Y.; Choi, A.; Choi, N.-S.; Lee, K. T. Charge carriers in rechargeable batteries: Na ions vs. Li ions. *Energy & Environmental Science* 2013, 6 (7), 2067-2081. <https://doi.org/10.1039/C3EE40811F>
6. Vaalma, C.; Buchholz, D.; Weil, M.; Passerini, S. A cost and resource analysis of sodium-ion batteries. *Nature Reviews Materials* 2018, 3 (4), 1-11. <https://doi.org/10.1038/natrevmats.2018.13>
7. Delmas, C. Sodium and sodium-ion batteries: 50 years of research. *Advanced Energy Materials* 2018, 8 (17), 1703137. <https://doi.org/10.1002/aenm.201703137>
8. Ponrouch, A.; Monti, D.; Boschini, A.; Steen, B.; Johansson, P.; Palacín, M. R. Non-aqueous electrolytes for sodium-ion batteries. *Journal of Materials Chemistry A* 2015, 3 (1), 22-42. <https://doi.org/10.1039/C4TA04428B>
9. Kim, H.; Kim, H.; Ding, Z.; Lee, M. H.; Lim, K.; Yoon, G.; Kang, K. Recent progress in electrode materials for sodium-ion batteries. *Advanced Energy Materials* 2016, 6 (19), 1600943. <https://doi.org/10.1002/aenm.201600943>
10. Kim, S. W.; Seo, D. H.; Ma, X.; Ceder, G.; Kang, K. Electrode materials for rechargeable sodium-ion batteries: potential alternatives to current lithium-ion batteries. *Advanced Energy Materials* 2012, 2 (7), 710-721. <https://doi.org/10.1002/aenm.201200026>
11. Wang, L. P.; Yu, L.; Wang, X.; Srinivasan, M.; Xu, Z. J. Recent developments in electrode materials for sodium-ion batteries. *Journal of Materials Chemistry A* 2015, 3 (18), 9353-9378. <https://doi.org/10.1039/C4TA06467D>
12. Monti, D.; Jónsson, E.; Boschini, A.; Palacín, M. R.; Ponrouch, A.; Johansson, P. Towards standard electrolytes for sodium-ion batteries: physical properties, ion solvation and ion-pairing in alkyl carbonate solvents. *Physical Chemistry Chemical Physics* 2020, 22 (39), 22768-22777. <https://doi.org/10.1039/D0CP03639K>
13. Chayambuka, K.; Cardinaels, R.; Gering, K. L.; Raijmakers, L.; Mulder, G.; Danilov, D. L.; Notten, P. H. An experimental and modeling study of sodium-ion battery electrolytes. *Journal of Power Sources* 2021, 516, 230658. <https://doi.org/10.1016/j.jpowsour.2021.230658>
14. Hakim, L.; Ishii, Y.; Matsumoto, K.; Hagiwara, R.; Ohara, K.; Umebayashi, Y.; Matubayashi, N. Transport properties of ionic liquid and sodium salt mixtures for sodium-ion battery electrolytes from molecular dynamics simulation with a self-consistent atomic charge determination. *The Journal of Physical Chemistry B* 2020, 124 (33), 7291-7305. <https://doi.org/10.1021/acs.jpcb.0c04078>
15. Xu, K. Nonaqueous liquid electrolytes for lithium-based rechargeable batteries. *Chemical Reviews* 2004, 104 (10), 4303-4418. <https://doi.org/10.1021/cr030203g>
16. Vignarooban, K.; Kushagra, R.; Elango, A.; Badami, P.; Mellander, B.-E.; Xu, X.; Tucker, T.; Nam, C.; Kannan, A. M. Current trends and future challenges of electrolytes for sodium-ion batteries. *International Journal of Hydrogen Energy* 2016, 41 (4), 2829-2846. <https://doi.org/10.1016/j.ijhydene.2015.12.090>
17. Ponrouch, A.; Marchante, E.; Courty, M.; Tarascon, J.-M.; Palacín, M. R. In search of an optimized electrolyte for Na-ion batteries. *Energy & Environmental Science* 2012, 5 (9), 8572-8583. <https://doi.org/10.1039/C2EE22258B>
18. Kalhoff, J.; Eshetu, G. G.; Bresser, D.; Passerini, S. Safer electrolytes for lithium-ion batteries: state of the art and perspectives. *ChemSusChem* 2015, 8 (13), 2154-2175. <https://doi.org/10.1002/cssc.201500284>

19. Landesfeind, J.; Hosaka, T.; Graf, M.; Kubota, K.; Komaba, S.; Gasteiger, H. A. Comparison of ionic transport properties of non-aqueous lithium and sodium hexafluorophosphate electrolytes. *Journal of The Electrochemical Society* 2021, 168 (4), 040538. <http://dx.doi.org/10.1149/1945-7111/abf8d9>
20. Kalugin, O.; Volobuev, M.; Kolesnik, Y. V. MDNAES: the program set for computer modelling of ion-molecular systems by using molecular dynamics method. *Kharkov University Bulletin* 1999, 454, 58-79.
21. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* 2007, 126 (1). <https://doi.org/10.1063/1.2408420>
22. Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 1984, 81 (8), 3684-3690. <https://doi.org/10.1063/1.448118>
23. Haile, J. M. *Molecular dynamics simulation: elementary methods*; John Wiley & Sons, Inc., 1992. ISBN:978-0-471-81966-0
24. Haberlandt, R.; Fritzsche, S.; Peinel, G. *Grundlagen und Anwendungen*. Wiesbaden: Friedr. Vieweg & Sohn Verlagsgesellschaft mbH 1995, 19. <https://doi.org/10.1007/978-3-322-90870-4>
25. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* 1996, 118 (45), 11225-11236. <https://doi.org/10.1021/ja9621760>
26. Canongia Lopes, J. N.; Deschamps, J.; Pádua, A. A. Modeling ionic liquids using a systematic all-atom force field. *The Journal of Physical Chemistry B* 2004, 108 (6), 2038-2047. <https://doi.org/10.1021/jp0362133>
27. Dudarev, D.; Holubenko, Y.; Jallah, R.; Kalugin, O. Microstructure and transport properties of lithium hexafluorophosphate solutions in binary mixture of dimethyl carbonate with ethylene carbonate from molecular dynamics simulation. *Kharkiv University Bulletin. Chemical Series* 2024, (42), 23-37. <https://doi.org/10.26565/2220-637X-2024-42-03>

Надіслано до редакції 03.10.2024

Прийнято до друку 19.12.2024

A.V. Borovyk, Y.V. Kolesnik, O.M. Kalugin. Structure and Transport Properties of NaPF_6 Solutions in Mixtures of Ethylene Carbonate with Dimethyl Carbonate for Sodium-Ion Batteries: MD Simulation

V.N. Karazin Kharkiv National University, School of Chemistry, Inorganic chemistry department, 4 Svobody sq., 61022 Kharkiv, Ukraine

Sodium-ion batteries (SIBs) have the potential to become new efficient electrical energy storage devices. However, at the moment, the main problem is the lack of a clear technology for their production. For the industrial production of SIBs, it is necessary to develop cathode and anode materials, as well as to choose the optimal composition of the electrolyte. For this purpose, using molecular dynamics modeling methods, we calculated the density, viscosity, electrical conductivity, and diffusion coefficients for NaPF_6 systems in the binary solvent EC:DMC (15:85 wt%, 30:70 wt%, and 50:50 wt%), and their structural properties were also considered.

The structure of the solvation shell of cations and anions was studied within the framework of radial distribution functions and current coordination numbers. The results indicate a more structured solvation shell of Na^+ cations than of PF_6^- anions.

The study of transport properties showed that the most suitable electrolytes for the production of sodium-ion batteries are systems in which EC:DMC=15:85 wt%. This is due to the fact that the electrolyte of this particular composition showed the lowest viscosity values in the region of all concentrations, as well as the highest values of electrical conductivity. The Na^+ diffusion coefficients for this system also reach the highest values compared to electrolytes of other compositions, which is a convincing argument for its future use in sodium-ion batteries.

Key words: molecular dynamics simulations, sodium-ion batteries, sodium hexafluorophosphate, ethylene carbonate, dimethyl carbonate, structural properties, transport properties, MDNAES.

Received 03.10.2024

Accepted 19.12.2024

Kharkiv University Bulletin. Chemical Series. Issue 43 (66), 2024

METHODS OF PROTECTION/DEPROTECTION OF HYDROXY GROUPS IN THE SYNTHESIS OF POLYHYDROXY FLAVONOLS

O. Demidov^a, A. Roshal^b

V. N. Karazin Kharkiv National University, The Research Institute of Chemistry, 4 Svobody sq., Kharkiv, 61022 Ukraine

a) ✉ alex.demidov2019@gmail.com

b) ✉ alexandre.d.rochal@karazin.ua

 <https://orcid.org/0000-0003-4657-6884>

 <https://orcid.org/0000-0003-1537-9044>

The article represents a review of methods for obtaining polyhydroxy flavonols without protection of hydroxy groups, as well as syntheses using methylation, alkylation and benzylation of the initial reagents and, accordingly, demethylation, dealkylation and debenylation of the final flavonols. It is shown that the most convenient for the synthesis of natural polyhydroxy flavonols and their analogues is the debenylation reaction using a Pd/C catalyst in tetrahydrofuran, which allows to obtain flavonols containing both hydroxy and methoxy groups. Syntheses using benzylation/debenylation reactions are easily scaled up, which allows to obtain of large quantities of polyhydroxy flavonols, in addition, the latter do not contain impurities of hydrogen halides, which makes it possible to use the obtained flavonols in the pharmaceutical and food industries.

The syntheses of hydroxy flavonols with a pyrogallol-like structure of the side phenyl ring were carried out, and the natural flavonol fisetin, a promising medicinal product and component of food additives, was obtained through benzylation/debenylation reactions. effect of ensitrelvir are found in the 1-methyl-1*H*-1,2,4-triazole and 6-chloro-2-methyl-2*H*-indazole fragments.

Keywords: Flavonols, flavonol synthesis, Algar-Flynn-Oyamada reaction, hydroxy group protection.

Introduction

Flavonoids are biologically active substances widespread in plants and fungi [1]. Of greatest interest are the derivatives of 3-hydroxy-2-phenylbenzo- γ -pyrone – flavonols (Figure 1), which are powerful antioxidants, carcinostatics, and have anti-inflammatory properties [2]. Flavonols are also used as complexing agents in the fluorescence analysis of some metal ions; the use of flavonols increases the efficiency of metal ion extraction [3, 4]. In addition, flavonols are convenient models for studying the excited-state intramolecular proton-transfer reaction [5].

The greatest biological activity is demonstrated by polyhydroxy derivatives of flavonols containing two to five hydroxy groups. Thus, 5,7,3',4'-tetrahydroxyflavonol (quercetin) is used as a medicine in cardiology and phlebology, the diglucoside (rutinoside) of this compound - rutin is used as a medicinal form of vitamin P [6, 7]. Currently, 7,3',4'-trihydroxyflavonol - fisetin, demonstrating neurotropic activity that supports the survival, differentiation, and functional maintenance of brain cells, is beginning to be used as a medicine [8].

Currently, flavonols for medicinal use are extracted from plant raw materials. The processes of extraction and purification of these compounds are labor-intensive, which determines the high cost of the final product. Therefore, the development of effective and cheap methods for synthesizing polyhydroxyflavonols is currently very relevant.

There are several routes to synthesizing flavonols, the most commonly used being the oxidative cyclization method. Typically, the first step involves synthesizing 2'-hydroxychalcone, which is then followed by the Algar-Flynn-Oyamada cyclization reaction in the presence of hydrogen peroxide [9, 10]. This method is widely used to obtain synthetic derivatives of flavonols containing amino groups, nitro groups, halogens, and to obtain heteroaromatic analogs of flavonols whose side ring is substituted with pyridine, benzadiazole, or thiazole fragments. However, the synthesis of natural flavonols containing a large number of hydroxyl groups is associated with certain difficulties. Condensation of the corresponding hydroxy acetophenones and hydroxy aldehydes is most often carried out in a highly alkaline medium or in a basic solvent. This leads to the dissociation of hydroxy groups, forming unstable anionic forms that are oxidized by atmospheric oxygen and are less active in aldol condensations when obtaining intermediate hydroxy chalcones. The Algar-Flynn-Oyamada reaction is also carried out in an alkaline medium, where the resulting hydroxychalcone anions are oxidized by hydrogen per-

© Demidov O., Roshal A., 2024



[This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0.](https://creativecommons.org/licenses/by/4.0/)

oxide, which is used as a mandatory reagent. In this case, protection of the hydroxy groups is logical, but this protection must occur selectively and not affect the 2'-hydroxy group of acetophenone and the chalcone derivative.

In this article, we briefly reviewed methods for the synthesis of polyhydroxy flavonols containing the pyrogallol-like 3',4'-dihydroxy groups in the side benzene ring, which are most susceptible to oxidation. We also discuss the possibility of obtaining partially methylated derivatives typical for many natural compounds. Basing the information obtained, we carried out regioselective syntheses of 3',4'-R-hydroxyflavonols and the natural flavonol fisetin using the debenzylation reaction. The structures of the obtained flavonols are shown in Table 1.

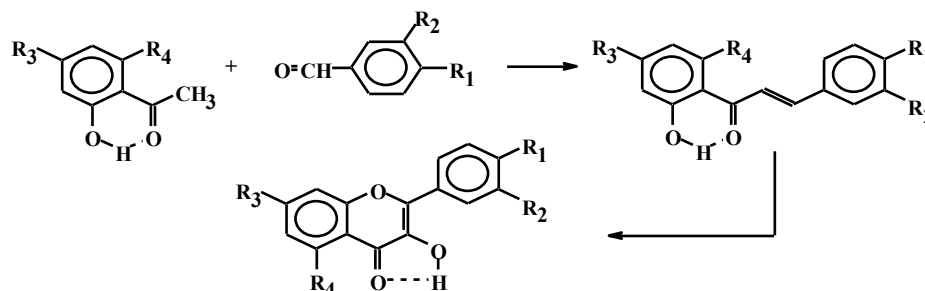


Figure 1. Scheme of flavonols' synthesis according to Algar-Flynn-Oyamada reaction

Table 1. Structures of synthesized flavonol derivatives*

	Substituents			
	R ₁ (C4')	R ₂ (C3')	R ₃ (C7)	R ₄ (C5)
flavonol	H	H	H	H
1	OBn	OBn	H	H
2	OH	OH	H	H
3	OBn	OMe	H	H
4	OH	OMe	H	H
5	OMe	OBn	H	H
6	OMe	OH	H	H
7	OBn	OBn	OBn	H
fisetin	OH	OH	OH	H

* OBn – benzyloxy-, OMe – methoxy- substituents.

Materials and Methods

Commercial reagents were used for the synthesis and physicochemical studies of the flavonol structure. The purification degree of intermediate and final reaction products was controlled using an Agilent 1100 high-performance liquid chromatograph with a SUPELCO Ascentis Express C18 chromatographic column 2.7 μ m 4.6 mm x 15 cm.

Identification of compounds was conducted by mass spectrometry using an Agilent LC/MSD SL mass-selective detector. ¹H NMR spectra were recorded using Unity Inova 400, Bruker Avance DRX 500 and Bruker Avance III 400 MHz spectrometers in DMSO-d₆. ¹³C NMR spectra were recorded on Bruker Avance DRX 500 and Agilent ProPulse 500 MHz spectrometers at a working frequency of 126 MHz in DMSO-d₆. Chemical shifts are presented in δ (ppm) scale. The reaction progress and the individuality of the obtained substances were monitored by TLC on silica gel-coated Polychrome SI F254 plates with a fluorescent detector in a hexane-ethyl acetate 2:1 system. Melting points are performed using Hanon Instruments MP450 Automatic Melting Temperature Controller.

Results and Discussion

Analysis of scientific publications allowed us to identify two strategies for the synthesis of polyhydroxy flavonols (PHF): obtaining final products from the corresponding acetophenones and aldehydes without the protection of hydroxy groups, as well as synthesis with preliminary protection of most hydroxy groups of the starting reagents (except for the 2-hydroxy group of acetophenones). Methylation and benzylation were most frequently used to protect hydroxyl groups.

As noted above, carrying out the reaction in an alkaline medium leads to partial oxidation of both the initial reagents – hydroxy benzaldehydes and acetophenones, the intermediate products – polyhydroxy chalcones, as well as the final flavonols initially formed in the anionic form. Nevertheless, the authors of works [11-13] attempted direct synthesis of hydroxy flavonols without protecting hydroxy groups. The synthesis of flavonols with hydroxy groups in the C3' and C4' positions of the side benzene ring was described in [11]. To prevent oxidation, the synthesis was carried out in a nitrogen atmosphere. The yield of the final compounds was 40-48%. The yield of 3'-hydroxy flavonol obtained in the same synthetical way but in the presence of air decreased to 33% [12]. A compound with an additional 7-hydroxy group was obtained with a lower yield. The authors of [13] performed syntheses of various flavonols in one and two stages, but they failed to reproduce the syntheses of polyhydroxy flavonols described in [11] and [12]. We also failed to reproduce the synthesis of 4'-hydroxy flavonol: the compound was obtained in low yield (<10%), and a high degree of resinification of the final flavonol was observed when carrying out the Algar-Flynn-Oyamada reaction. Also, when purifying the latter by column chromatography on silica gel, the flavonol partially decomposes during the separation process upon contact with the sorbent.

The second method of synthesis is based on selective methylation, ethylation or butylation of hydroxy groups in the initial acetophenones and aldehydes, obtaining alkoxy derivatives of 2'-hydroxychalcone in the first stage of synthesis, with subsequent cyclization of the latter into polyalkoxy flavonols. Depending on the position of the alkoxy groups, the final yield of flavonols is within 15-40%. The last stage of the PHF synthesis is the removal of protective alkyl groups. In this case, a 57% aqueous solution of hydroiodic acid is frequently used for dealkylation of flavonols [14]. The yield of polyhydroxy flavonols in the HI-assisted dealkylation reaction is 40-60%. Demethylation of flavonols with hydroiodic acid is also carried out in solutions of acetic acid [15] or acetic anhydride [16, 17] in the presence of phenol [18] followed by extraction of the reaction mixture with ethyl acetate. In this case, the demethylation yields of reactions are in the range of 50–75%.

Demethylation in aqueous HBr solutions occurs with a yield of approximately 20% [19], in glacial acetic acid and in an inert atmosphere the reaction yield increases to 65% [20]. Demethylation of 3'- and 4'-hydroxy groups of flavonols using HCl in glacial hydrochloric acid has also been described [21], in that case, polyhydroxy flavonols were obtained with a yield of 70–72%.

A less common method used to remove protective alkyl groups is dealkylation with BBr_3 in dry dichloromethane, under an inert gas atmosphere and at -78°C [21] or by heating to 50°C [22]. The yield of dealkylated compounds depends on the location of the protected groups and the presence of other substituents and generally ranges from 50–70%.

All methods for obtaining hydroxy flavonols from initial methylated (alkylated) reagents with subsequent demethylation (dealkylation) of the final compound are characterized by a number of disadvantages – first of all, the difficult purification of flavonols from traces of acetic acid and hydrogen halides. The presence of traces of the latter in polyhydroxyflavonols limits their use as pharmaceuticals and components of food additives. In addition, none of the listed methods allows regioselective demethylation of flavonols. Since natural flavonoids can simultaneously have hydroxy and methoxy groups, it is not possible to obtain such compounds using the above dealkylation methods. Even if the molecule contains different protective fragments, such as methoxy, butoxy or benzyloxy groups, the use of the above methods leads to their complete and non-selective removal.

Comparison of the synthetical methods showed that the most convenient is the protection of hydroxy groups of the initial reagents with benzyl residues, as well as the debenzylation of finally obtained benzyloxy flavonols. As noted above, the benzyl group can be removed with hydrohalic acids in concentrated acetic acid. For weaker acid HCl, it is necessary to heat the reaction mixture to boiling [21, 23], the reaction with stronger HBr occurs at room temperature [24]. The yields of the debenzylation reaction are 45-70% and 60-80%, respectively, depending on the location of the benzyloxy groups in the molecule. Benzyloxy groups are also removed with TiCl_4 [25] with a hydroxyflavonols' yield of 64%, and trifluoroacetic acid in the presence of thioanisole as a catalyst [26, 27] with yields of debenzylated compounds of 30-60%.

The use of the BBr_3 reagent in dichloromethane [21] is non-selective and results in the removal of not only benzyloxy but also any other alkoxy substituents. In addition, the yield of debenzylated derivatives for various compounds can vary widely from 30 to 95%.

Recently, the removal of benzyl fragments by catalytic hydrogenation in the presence of a Pd/C catalyst, metallic palladium adsorbed on a carbon carrier, has become widespread [28, 29]. The use of palladium compounds as a catalyst, for example, Pd(OH)₂/C, is undesirable in the case of polyhydroxyflavonols since it leads to the formation of colored flavonol-palladium complexes [30]. Pd/C catalysts are renewable, non-toxic, used in the food and pharmaceutical industries, and therefore are most suitable for the synthesis of analogs of natural flavonols.

Tetrahydrofuran, methanol, ethanol, ethyl acetate, or their mixtures in various ratios are used as solvents in debenzylation reactions [31–34]. However, when scaling up reactions to 1–10 grams, the reproducibility of the methods with MeOH [35, 36], EtOAc [37], and EtOH [38, 39] turned out to be low, with yields of 30–40%. This is due to the extremely low solubility of benzyloxy derivatives of flavonols in the indicated solvents. The best reproducibility during the scaling up of syntheses was noted when using tetrahydrofuran [40] as a solvent. In our experiments, the use of 10–20 volume equivalents of tetrahydrofuran relative to the starting materials provided high yields in the range of 75–95%.

Taking into account the possibility of carrying out regioselective synthesis using debenzylation reactions of flavonols, we obtained compounds with different arrangements of hydroxy and methoxy groups and also synthesized the natural flavonol fisetin with the protection of hydroxyls by benzyl groups with subsequent debenzylation reaction. Typical methods of flavonol synthesis are described below. The results of physicochemical studies of the structure of the obtained compounds are given in Tables 2–4.

General Procedure for the Synthesis of Chalcones. Under a nitrogen atmosphere, a suspension of NaH (6.5 mmol, 60% dispersion in mineral oil) in DMF (20 mL) was cooled to 0 °C. A solution of 2'-hydroxyacetophenone (5 mmol) was then added dropwise, and the mixture was stirred at 0 °C for 10 minutes. Subsequently, a solution of the corresponding benzaldehyde (5 mmol) in DMF (10 mL) was added dropwise, and the reaction was allowed to proceed at room temperature for 2 hours. The reaction mixture was then gradually acidified with glacial acetic acid until pH 4. The resulting precipitate was collected by filtration, washed with MeOH (10 mL), and used in the next stage without further purification.

Table 2. Melting points and mass-spectra of synthesized flavonols, yields of flavonol synthesis stages.

flavonols	Mass spectra (m/z, %)	T _{m.p.} , °C	Yield, %			
			Claisen-Schmidt condensation	AFO*	Debenzylation	All stages
1	451 ([M+H] ⁺ , 74), 473 ([M+Na] ⁺ , 26)	146-158	90 %	74 %	–	67 %
2	271 ([M+H] ⁺ , 88), 293 ([M+Na] ⁺ , 4)	253-291	90 %	74 %	86 %	57 %
3	375 ([M+H] ⁺ , 100)	153-155	98 %	67 %	–	66 %
4	285 ([M+H] ⁺ , 100)	242-253	98 %	67 %	90 %	59 %
5	375 ([M+H] ⁺ , 88), 397 ([M+Na] ⁺ , 12)	197-208	96 %	68 %	–	65 %
6	285 ([M+H] ⁺ , 94), 307 ([M+Na] ⁺ , 6)	207-208	96 %	68 %	90 %	59 %
7	557 ([M+H] ⁺ , 100)	158-159	88%	65%	–	57%
fisetin	287 ([M+H] ⁺ , 100)	329-331	88%	65%	85%	49%

* AFO – the yield of Flynn-Algar-Oyamada reaction

General Procedure for the Synthesis of Flavonols. A suspension of chalcone (6 mmol) in MeOH (50 mL) was treated with a 15% solution of KOH (15 mmol) and stirred at -15 °C for 10 minutes. A 30% H₂O₂ solution (18 mmol) was then added dropwise, and the reaction was allowed to proceed at room temperature for 2.5 hours. The reaction mixture was subsequently acidified with glacial acetic acid until pH 2. The precipitate was collected by filtration, and washed with MeOH (20 mL) to afford the pure product.

General Procedure for Debenzylation. The debenzylation reaction was carried out by dissolving the target compound in THF (10–20 volEq.) in the presence of 10% Pd/C (0.1 weq.). The mixture was stirred under a hydrogen atmosphere at room temperature for 12 hours. Upon completion, the catalyst

was removed by filtration through a silica gel, and the solvent was evaporated under reduced pressure to yield the deprotected product.

Conclusions

As shown by the analysis of the methods for hydroxy group protection in the synthesis of flavonols, the best results are obtained using benzylation of the initial reagents and subsequent debenzilation of the reaction products.

Table 3. ^{13}C -NMR signals of synthesized flavonol

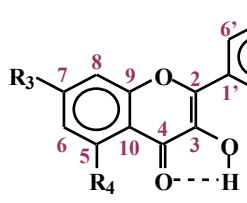
Atom numbering in ^{13}C NMR spectra	Comp.	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10
	1	145.2	138.4	172.6	124.7	124.0	133.4	118.3	154.3	121.9
	2	146.6	138.4	172.9	125.2	124.9	133.8	118.6	154.8	121.8
	3	145.3	138.3	172.6	124.7	124.5	133.4	118.3	154.3	121.2
	4	145.9	138.0	172.5	124.4	124.6	133.3	118.3	154.3	121.8
	5	145.3	138.3	172.6	113.0	124.7	133.4	118.3	154.4	121.3
	6	145.9	138.0	172.5	124.4	124.6	133.3	118.3	154.3	121.8
	7	146.3	137.8	172.8	124.0	125.1	133.1	115.0	155.0	158.2
	fisetin	145.5	137.6	172.4	122.9	124.3	134.1	116.0	156.7	162.7
Comp.	C-1'	C-2'	C-3'	C-4'	C-5'	C-6'	(CH ₃)	C (CH ₂)	C (Ph*)	
1	121.2	113.7	147.7	149.8	113.7	124.4	–	70.4, 69.9	137.0, 136.8, 128.4, 127.9, 127.8, 127.6	
2	122.8	116.1	145.6	148.1	115.8	120.5	–	–	–	
3	123.5	112.7	147.4	150.6	111.7	121.9	55.6	70.2	136.9, 128.4, 128.1, 128.9	
4	122.3	111.7	147.4	148.7	115.5	121.3	55.7	–	–	
5	121.2	111.3	148.7	149.3	124.4	123.9	55.7	69.8	136.7, 128.4, 127.9, 127.8	
6	122.1	111.5	149.7	149.2	114.7	121.0	–	71.2	128.0, 127.7, 127.1	
fisetin	122.8	115.3	145.5	146.7	115.3	116.5	–	–	–	

Table 4. ^1H -NMR signals of synthesized flavonols

Flavonols	H-5	H-6	H-7	H-8	H-2'	H-5'	H-6'	H (OH)	H (CH ₂)	H (Ph*)	H (CH ₃)
1	8.09 (d)	7.83 – 7.70 (m)	7.83 – 7.70 (m)	7.54 – 7.28 (m)	7.94 (s)	7.25 (d)	7.87 (d)	9.52 (s)	5.22 (d)	7.54 – 7.28 (m)	–
2	7.9 (dd)	7.43 (t)	7.72 – 7.65 (m)	7.79 – 7.72 (m)	7.79 – 7.72 (m)	6.91 (d)	7.61 (dd)	9.39 (m)	–	–	–
3	8.09 (d)	7.51 – 7.36 (m)	7.51 – 7.36 (m)	7.80 – 7.73 (m)	7.51 – 7.36 (m)	7.22 (d)	7.80 – 7.73 (m)	9.48 (s)	5.17 (s)	7.88 – 7.80 (m), 7.51 – 7.36 (m)	3.85 (s)
4	8.09 (d)	7.44 (t)	7.80 – 7.73 (m)	7.80 – 7.73 (m)	7.83 (s)	6.96 (d)	7.80 – 7.73 (m)	9.55 (m)	–	–	3.86 (s)
5	8.10 (d)	7.94 – 7.87 (m) and 7.84 – 7.72 (m)	7.84 – 7.72 (m)	7.84 – 7.72 (m)	7.49 – 7.38 (m)	7.17 (d)	7.94 – 7.87 (m)	9.50 (s)	5.18 (s)	7.49 – 7.38 (m), 7.34 (t)	3.86 (s)
6	8.11 (dd)	7.45 (t)	7.73 – 7.60 (m)	7.83 – 7.73 (m)	7.83 – 7.73 (m)	7.10 (d)	7.73 – 7.60 (m)	9.29 – 9.21 (m)	–	–	3.86 (s)
7	8.05	7.15 (d)	–	7.22 (s)	6.94 (s)	7.12 (d)	7.35 – 7.39 (m)	10.3	5.26	7.48 – 7.30 (m)	–
fisetin	7.90 (d)	7.52 (t)	–	7.67 – 6.64 (m)	7.62 – 6.58 (m)	6.95 (d)	6.85 – 6.87 (m)	9.49 (s) 10.72 (s)	–	–	–

* Phenyl ring of benzyl moiety

Debenzylation using a Pd/C catalyst in tetrahydrofuran provides a high yield of polyhydroxy flavonols – 86-90%; the synthesis method is successfully scaled up, the catalyst is easily regenerated, and the resulting final products do not contain traces of halogens, which allows them to be used in the pharmaceutical and food industries. Debenzylation also allows one to obtain flavonols with different arrangements of hydroxy- and methoxy- groups, which is important in the synthesis of natural compounds. The results presented in the article confirm the possibility of synthesizing flavonols with various substituents in the side phenyl fragment, as well as the possibility of successfully synthesizing fisetin using benzylation/debenzylation reactions.

Acknowledgement

The authors acknowledge the grant 2021.01/0062 “*Molecular design, synthesis and screening of new potential antiviral pharmaceutical ingredients for the treatment of infectious diseases COVID-19*” from the National Research Foundation of Ukraine.

The authors thank Prof. T. Langer from University of Vienna for giving us the opportunity to use LigandScout 4.4.9 suite.

References

1. Hao, B.; Yang, Zh.; Liu, H.; Liu, Yu; Wang Sh. Advances in Flavonoid Research: Sources, Biological Activities, and Developmental Prospectives. *Current Issues in Molecular Biology*. 2024, 46(4), 2884-2925. <https://doi.org/10.3390/cimb46040181>.
2. Panche, A.N.; Diwan, A.D.; Chandra, S.R. Flavonoids: an overview. *Journal of Nutritional Science*. 2016, 5, art.num e47. <https://doi.org/10.1017/jns.2016.41>
3. Roshal, A.D. Complexation of flavonoids: Spectral phenomena, regioselectivity, interplay with charge and proton transfer. *The Chemical Record*. 2024, 4 (2), art. num. e202300249. <https://doi.org/10.1002/tcr.202300249>.
4. Demidov, O.O.; Krasnopyorova, A.V.; Yukhno, G.D.; Efimova, N.V.; Roshal, A.D. Flavonol assisted extraction of divalent and trivalent metal ions. *Functional Materials*. 2024, 31(4), 601–608, <https://doi.org/10.15407/fm31.04.601>.
5. Roshal, A.D.; Organero, J.A.; Douhal, A. Tuning the mechanism of proton-transfer in a hydroxyflavone derivative. *Chemical Physics Letters*, 2003, 379, 53-59. <https://doi.org/10.1016/j.cplett.2003.08.008>
6. Yang, D.; Wang, T.; Long, M.; Li, P. Quercetin: Its Main Pharmacological Activity and Potential Application in Clinical Medicine. *Oxidative Medicine and Cellular Longevity*, 2020, art. num. 2020:8825387. <https://doi.org/10.1155/2020/8825387>.
7. Ganeshpurkar, A.; Saluja, A.K. The Pharmacological Potential of Rutin. *Saudi Pharmaceutical Journal*, 2016, 25(2), 149–164. <https://doi.org/10.1016/j.jsps.2016.04.025>.
8. Elwan, A.H.; El-Masry, S.M.; Habib, D.A.; Zewail, M. An insight into fisetin, the miraculous multifaceted flavonol: Paving the road for enhanced delivery through promising pharmaceutical nano-formulations. *Journal of Drug Delivery Science and Technology*, 2024, 101 (Part B), art. num. 106292, <https://doi.org/10.1016/j.jddst.2024.106292>.
9. Algar, J.; Flynn, J.P. A New Method for the Synthesis of Flavonols. *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science*. 1934, XLII, 1-8.
10. Oyamada, T.J. A new general method for the synthesis of the derivatives of flavonol. *Bulletin of the Chemical Society of Japan*, 1934, 55, 1256–1261. <https://doi.org/10.1246/bcsj.10.182>.
11. Ma, M.-L.; Li, M.; Gou, J.-J.; Ruan, T.-Y.; et al. Design, synthesis and biological activity of flavonoid derivatives as selective agonists for neuromedin U 2 receptor. *Bioorganic & Medicinal Chemistry*. 2014, 22(21), 6117–6123. <https://doi.org/10.1016/j.bmc.2014.08.038>.
12. Ferrari, G.V.; Pappano, N.B.; Montaña, M.P.; Garcia, N.A.; Debattista, N.B. Synthesis of 3,3'-Dihydroxyflavone and Apparent Formation Constants of Flavonoid–Ga(III) Complexes. *Journal of Chemical & Engineering Data*, 2010, 55(9), 3080–3083. <https://doi.org/10.1021/jc901091f>.
13. Gunduz, S.; Goren, A.C.; Ozturk T. Facile Syntheses of 3-Hydroxyflavones. *Organic Letters*. 2012, 14(6), 1576–1579. <https://doi.org/10.1021/ol300310e>.
14. Sobottka, A.M.; Werner, W.; Blaschke, G.; Kiefer, W.; et al. Effect of Flavonol Derivatives on the Carrageenin-Induced Paw Edema in the Rat and Inhibition of Cyclooxygenase-1 and 5-

- Lipoxygenase in Vitro. *Archiv der Pharmazie*. 2020, 333 (7), 205–210 [https://doi.org/10.1002/1521-4184\(20007\)333:7<205::aid-ardp205>3.0.co;2-y](https://doi.org/10.1002/1521-4184(20007)333:7<205::aid-ardp205>3.0.co;2-y)
15. Shaw, B.L.; Simpson, T.H. Chelate systems. Part II. *Journal of the Chemical Society* 1952, 5027–5032. <https://doi.org/10.1039/jr9520005027>
 16. Gupta, S.R.; Seshadri T.R. Survey of anthoxanthins. Part VI. Colouring matter of tamarix troupii. Constitution of the aglycone and its synthesis. *Journal of the Chemical Society*. 1954, 3063–3065. <https://doi.org/10.1039/jr9540003063>
 17. Ahluwalia, V.K.; Seshadri T.R. Synthetic experiments in the benzopyrone series. *Proceedings of the Indian Academy of Sciences – Section A*. 1954, 39 (6), 296–300. <https://doi.org/10.1007/bf03048703>
 18. Sagareishvili, T.G.; Alaniya, M.D.; Tsitsishvili, V.G.; Kemertelidze, E.P. Micranthoside – A new glycoside from *Eupatorium micranthum*. *Chemistry of Natural Compounds*. 1981, 17, 225–230. <https://doi.org/10.1007/BF00568507>
 19. Shih, T.-L.; Chou, C.-E.; Liao, W.-Y.; Hsiao C.-A. Copper-mediated trimethylsilyl azide in amination of bromoflavonoids to synthesize unique aminoflavonoids. *Tetrahedron*. 2014, 70 (23), 3657–3664, <https://doi.org/10.1016/j.tet.2014.04.022>
 20. Lindel, Th.; Mende S. Synthesis of morin and morin derivatives. US Patent 2020/0308131 A1, 2020, C07D 311/28.
 21. Sousa, J.L.C.; Proença, C.; Freitas, M.; Fernandes, E.; Silva, A.M.S. New polyhydroxylated flavon-3-ols and 3-hydroxy-2-styrylchromones: synthesis and ROS/RNS scavenging activities. *European Journal of Medicinal Chemistry*. 2016, 119, 250–259. <https://doi.org/10.1016/j.ejmech.2016.04.057>
 22. Ahn, M.; Park, S.E.; Choi, J.; Choi, J. et al. Synthesis and biological evaluation of flavonoid-based IP6K2 inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 2023, 38 (1), art.# 2193866, <https://doi.org/10.1080/14756366.2023.2193866>
 23. Qin, C.X.; Chen, X.; Hughes, R.A.; Williams, S.J.; Woodman, O.L. Understanding the Cardioprotective Effects of Flavonols: Discovery of Relaxant Flavonols without Antioxidant Activity. *Journal of Medicinal Chemistry*. 2008, 51 (6), 1874–1884. <https://doi.org/10.1021/jm070352h>
 24. Serdiuk, I.E.; Roshal, A.D.; Błażejowski, J. Origin of Spectral Features and Acid–Base Properties of 3,7-Dihydroxyflavone and Its Monofunctional Derivatives in the Ground and Excited States. *The Journal of Physical Chemistry A*. 2016, 120 (25), 4325–4337. <https://doi.org/10.1021/acs.jpca.6b03290>
 25. Ranga Rao, R.; Tiwari, A. K.; Prabhakar Reddy, P.; Suresh Babu, K.; et al. Synthesis of antihyperglycemic, α -glucosidase inhibitory, and DPPH free radical scavenging furanochalcones. *Medicinal Chemistry Research*. 2011, 21 (6), 760–774. <https://doi.org/10.1007/s00044-011-9583-7>
 26. Venkateswararao, E.; Son, M.-J.; Sharma, N.; Manickam, M.; et al. Exploration of Pharmacophore in Chrysosplenol C as Activator in Ventricular Myocyte Contraction. *ACS Medicinal Chemistry Letters*. 2015, 6 (7), 758–763 <https://doi.org/10.1021/acsmedchemlett.5b00043>
 27. Yap, S.; Woodman, O. L.; Crack, P. J.; Williams, S. J. Synthesis of a hypoxia-targeted conjugate of the cardioprotective agent 3',4'-dihydroxyflavonol and evaluation of its ability to reduce ischaemia/reperfusion injury. *Bioorganic & Medicinal Chemistry Letters*. 2011, 21 (17), 5102–5106. <https://doi.org/10.1016/j.bmcl.2011.03.040>
 28. Xie, J.; Xu, H.; Zhang, Q.; Wu, Z.; et al. Semi-Synthesis of Flavonoid Glycosides and Their Anti-Inflammatory and Antitumor Activities towards Triple Negative Breast Cancer. *Chemistry & Biodiversity*. 2023, 20 (2), art. num e202200899. <https://doi.org/10.1002/cbdv.202200899>
 29. Jian, J.; Fan, J.; Yang, H.; Lan, P.; et al. Total Synthesis of the Flavonoid Natural Product Houttuynoid A. *Journal of Natural Products*. 2018, 81 (2), 371–377. <https://doi.org/10.1021/acs.jnatprod.7b00791>
 30. Docampo-Palacios, M. L.; Alvarez-Hernández, A.; Adiji, O.; Gamiotea-Turro, D.; Valerino-Diaz, A. B.; et al. Glucuronidation of Methylated Quercetin Derivatives: Chemical and Biochemical Approaches. *Journal of Agricultural and Food Chemistry*. 2020, 68 (50), 14790–14807. <https://doi.org/10.1021/acs.jafc.0c04500>
 31. Jian, J.; Fan, J.; Yang, H.; Lan, P.; et al. Total Synthesis of the Flavonoid Natural Product Houttuynoid A. *Journal of Natural Products*. 2018, 81 (2), 371–377. <https://doi.org/10.1021/acs.jnatprod.7b00791>

32. Kim, S.; Li, Y.; Lin, L.; Sayasith, P. R.; et al. Synthesis and Biological Evaluation of 4'-Substituted Kaempfer-3-ols. *The Journal of Organic Chemistry*. 2020, 85 (6), 4279–4288. <https://doi.org/10.1021/acs.joc.9b03461>
33. Horie, T.; Tsukayama, M.; Kawamura, Y.; Seno, M.; et al. Studies of the Selective O-Alkylation and Dealkylation of Flavonoids. XI. A New Convenient Method for Synthesizing 3,5,7-Trihydroxy-8-methoxyflavones from 7-Hydroxy-3,5,8-trimethoxyflavones. *Bulletin of the Chemical Society of Japan*. 1988, 61 (2), 441–447. <https://doi.org/10.1246/bcsj.61.44>
34. Chiruta, C.; Schubert, D.; Dargusch, R.; Maher, P. Chemical Modification of the Multitarget Neuroprotective Compound Fisetin. *Journal of Medicinal Chemistry*. 2011, 55 (1), 378–389. <https://doi.org/10.1021/jm2012563>
35. Mei, Q.; Wang, C.; Zhao, Z.; Yuan, W.; et al. Synthesis of icariin from kaempferol through regioselective methylation and para-Claisen–Cope rearrangement. *Beilstein Journal of Organic Chemistry*. 2015, 11, 1220–1225. <https://doi.org/10.3762/bjoc.11.135>
36. Nguyen, V.-S.; Shi, L.; Li, Y.; Wang, Q.-A. Total Synthesis of Icaritin via Microwave-assistance Claisen Rearrangement. *Letters in Organic Chemistry*. 2014, 11 (9), 677–681. <https://doi.org/10.2174/157017861109140903103927>
37. Sheng, X.; Jia, X.-Y.; Tang, F.; Wang, Y.; et al. The total synthesis of (±)-sanggenol F. *Tetrahedron*. 2017, 73 (25), 3485–3491. <https://doi.org/10.1016/j.tet.2017.05.022>
38. Kan, T.; Hiza, A.; Tsukaguchi, Y.; Ogawa, T.; et al. Synthetic Studies of Fisetin, Myricetin and Nobiletin Analogs and Related Probe Molecules. *Heterocycles*. 2014, 88 (2), 1371–1396. [https://doi.org/10.3987/com-13-s\(s\)107](https://doi.org/10.3987/com-13-s(s)107)
39. He, L.; Zhou, Z.; Fang, Z.; Jin, H.; et al. Selective Monomethylation of Quercetin. *Synthesis*. 2010, 23, 3980–3986. <https://doi.org/10.1055/s-0030-1258310>
40. Estévez-Sarmiento, F.; Said, M.; Brouard, I.; León, F.; et al. 3'-Hydroxy-3,4'-dimethoxyflavone blocks tubulin polymerization and is a potent apoptotic inducer in human SK-MEL-1 melanoma cells. *Bioorganic & Medicinal Chemistry*. 2017, 25 (21), 6060–6070. <https://doi.org/10.1016/j.bmc.2017.09.043>

Received 29.10.2024

Accepted 19.12.2024

О. Демидов, О. Рошаль. Методи захисту/зняття захисту гідроксильних груп при синтезі полігідроксифлавонолів

Харківський національний університет імені В.Н. Каразіна, НДІ хімії, майдан Свободи, 4, Харків, 61022, Україна

У статті представлено огляд методів одержання полігідроксифлавонолів без захисту гідроксигруп, а також синтезів із застосуванням метилювання, алкілювання та бензилювання вихідних реагентів і, відповідно, деметилювання, деалкілювання та дебензилювання кінцевих флавонолів. Показано, що найбільш зручною для синтезу природних полігідроксифлавонолів та їх аналогів є реакція дебензилювання з використанням Pd/C каталізатора в тетрагідрофурані, що дозволяє отримати флавоноли, що містять як гідрокси-, так і метоксигрупи. Синтези з використанням реакцій бензилювання/дебензилювання легко масштабуються, що дозволяє отримувати великі кількості полігідроксифлавонолів, крім того, останні не містять домішок галогеноводнів, що дає можливість використовувати отримані флавоноли у фармацевтичній та харчовій промисловості.

Проведено синтези гідроксифлавонолів з пірогаллол-подібною структурою бічного фенільного циклу. З використанням реакцій бензилювання/дебензилювання отримано природний флавонол фізетин – перспективний лікарський засіб та компонент харчових добавок.

Ключові слова: флавоноли, синтез флавонолів, реакція Алгара-Флінна-Оямади, захист гідроксигруп.

Надіслано до редакції 29.10.2024

Прийнято до друку 19.12.2024

СИНТЕЗ МОНО- ТА ДИФЕНІЛ ЗАМІЩЕНИХ САЛІЦИЛОВИХ АЛЬДЕГІДІВ, ВАЖЛИВИХ БІЛДИНГ-БЛОКІВ ДЛЯ СИНТЕЗУ ФЛУОРЕСЦЕНТНИХ БАРВНИКІВ ТА БАРВНИКІВ-СЕНСИБІЛІЗАТОРІВ ДЛЯ DSSC

П. В. Тростьянко^{*,a}, В. В. Пашко^{†,b}, С. М. Коваленко^{*,c}

^{*}V. N. Karazin Kharkiv National University, School of Chemistry, 4 Svobody sqr., Kharkiv, 61022 Ukraine

[†]ТОВ «Донау Лаб Україна», вул. Стратегічне шосе, 16, Київ, 03028, Україна

a) ✉ trostianko.p.v@gmail.com

 <https://orcid.org/0000-0002-1333-9375>

b) ✉ pashko@dlu.com.ua

 <https://orcid.org/0009-0002-2140-9166>

c) ✉ kovalenko.sergiy.m@gmail.com

 <https://orcid.org/0000-0003-2222-8180>

Саліцилові альдегіди становлять інтерес для синтезу великої кількості біологічно активних сполук, ліганд комплексуютьовачів, флуоресцентних зондів та функціональних барвників-сенсibilізаторів для DSSC тощо.

В даній роботі нами розроблено метод синтезу моно- та дифеніл заміщених саліцилових альдегідів за реакцією крос-сполучення Судзукі-Міяури відповідних бромсаліцилових альдегідів з фенілбороновою кислотою, з метою дослідження можливостей модифікації будови функціональних барвників. Виконано серію експериментів з варіації умов синтезу, заміни розчинників та каталізаторів як при конвекційному нагріванні при атмосферному тиску, так із використанням реактора кондуктивного нагріву з герметичним резервуаром Anton Paar Monowave 50.

В процесі пошуку оптимальних умов проведення реакції приділено велику увагу сучасним вимогам «зеленої хімії». Найкращий результат одержано при проведенні реакції крос-сполучення Судзукі-Міяури бромсаліцилових альдегідів з фенілбороновою кислотою у воді з використанням каталізатора $\text{Pd}(\text{PPh}_3)_4$. Цей метод вдалося вдосконалити, використовуючи як каталізатор 10% Pd/C , що має ряд переваг, а саме: більша доступність, простота зберігання та використання, менша вартість та екологічна безпечність. Також каталізатор 10% Pd/C є більш стійким до впливу навколишнього середовища, що дозволило проводити реакції без використання інертної атмосфери в реакторі кондуктивного нагріву з герметичним резервуаром Anton Paar Monowave 50.

Ключові слова: органічний синтез, саліцилові альдегіди, крос-сполучення, реакція Судзукі-Міяури, зелена хімія, екологічно чистий синтез, паладієвий каталізатор, реактор Anton Paar Monowave 50, флуоресцентні барвники, DSSC.

Вступ

Саліцилові альдегіди є відомими білдинг-блоками з великим синтетичним потенціалом, що використовують в синтезі флуоресцентних барвників [1], лігандів комплексуютьовачів [2], флуоресцентних зондів [3], металевих комплексів барвників-сенсibilізаторів для DSSC [4,5], як вихідні сполуки в синтезі різноманітних похідних кумаринів [6] тощо.

Ефективність барвників-сенсibilізаторів для DSSC тісно пов'язана з довжиною π -спряженого ланцюга. Так, одним з найпоширеніших методів подовження ланцюга π -спряження органічних молекул є використання реакції крос-сполучення Судзукі-Міяури [7]. З метою оптимізації синтетичного шляху, нами було проведено дослідження методів синтезу та модифікації саліцилових альдегідів, як вихідних білдинг-блоків для барвників, що містять кумаринову ланку.

В роботі [8] описано синтез фенілсаліцилових альдегідів за класичною реакцією Судзукі-Міяури, з використанням ацетату паладію та лігандів трифенілфосфіну для каталізу реакції. Відомий також синтез похідних 4-фенілсаліцилового альдегіду, через крос-сполучення проміжної алільної похідної з використанням каталізатора $\text{Pd}(\text{PPh}_3)_4$ та подальшим її окисненням до альдегіду [9].

В сучасній органічній хімії активно досліджують можливості "озеленення" реакції Судзукі-Міяури. Так, велику увагу приділяють застосуванню екологічно безпечних розчинників, таких як H_2O , MeOH , EtOH на заміну небезпечних DMF, DXN, THF тощо. Також застосовують більш екологічні та придатні до відновлення та повторного застосування каталізатори, такі як

нанодисперсний паладій на активованому вугіллі [10, 11], складні фторовані паладієві каталізатори, придатні до переробки [12], та методи мікрохвильового синтезу в комбінації з вище наведеними підходами [13].

Матеріали та методи

Вихідні реагенти були придбані у Sigma-Aldrich (USA) та використані без додаткової очистки. Контроль реакцій проведено з використанням тонкошарової хроматографії на пластинках "Sorbfil UV-254" з використанням елюенту етил ацетат-гексан 1:2, v/v. ^1H та ^{13}C ЯМР-спектри були виміряні на Varian Gemini 400 MHz спектрометрі у дейтерованому диметилсульфоксиді ($\text{DMSO-}d_6$) як розчиннику. LC/MS спектри одержані на PE SCIEX API 150EX рідинному хроматографі з УФ детектором (λ_{max} 215 та 254 нм) та колонкою Luna-C18. Елюювання починали водою і закінчували сумішшю ацетонітрil/вода (95:5, v/v) з використанням лінійного градієнта зі швидкістю потоку 0,15 мл/хв і тривалістю циклу аналізу 25 хв.

Реактор кондуктивного нагріву з герметичним резервуаром Anton Paar Monowave 50

Monowave 50 – це реактор для синтезу з конвекційним підігрівом, спеціально створений для навчальних лабораторій та стандартних процедур експериментальної хімії (рис. 1). Цей прилад забезпечує контрольований нагрів до 250°C та контроль тиску до 20 бар, що дозволяє досягти продуктивності мікрохвильового реактора для проведення різноманітних реакцій органічного синтезу [14].

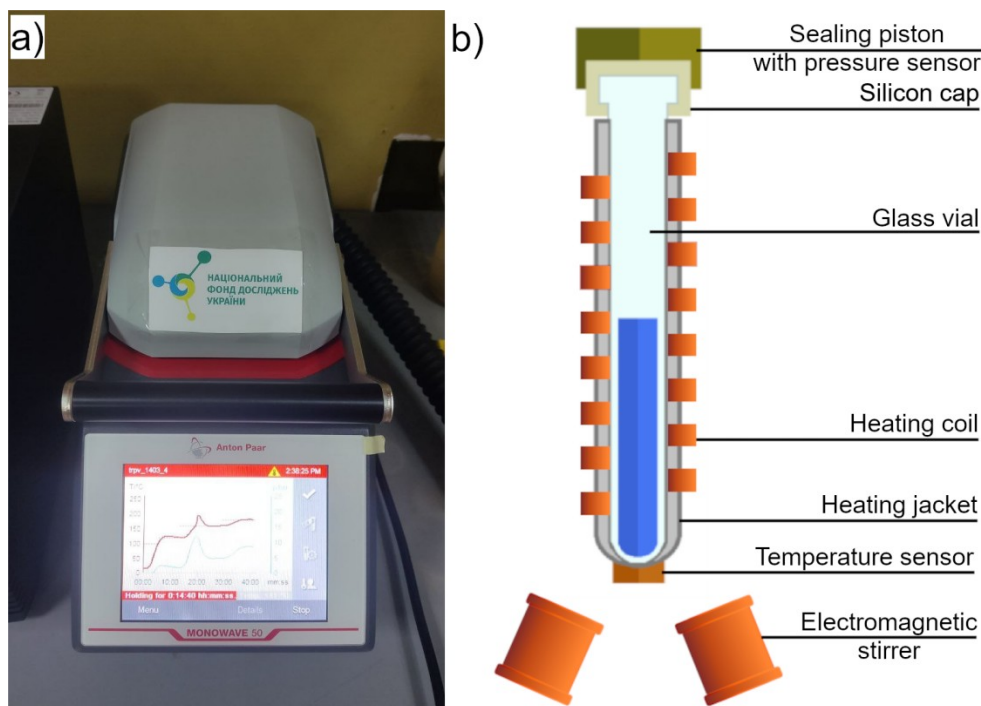


Рис. 1. а) Реактор Anton Paar Monowave 50 з датчиком тиску, б) схематичне зображення пристрою кондуктивного нагріву.

Fig. 1. a) Anton Paar Monowave 50 reactor with a pressure sensor, b) schematic representation of the conductive heating device.

Пристрій кондуктивного нагріву Monowave 50 складається з нагрівальної оболонки, в яку розміщують скляну віалу об'ємом 10 мл, споряджену магнітним якорем та силіконовою кришкою з ПТФЕ прокладкою. Рівномірне нагрівання пристрою відбувається за допомогою спіралі, що регулюється вбудованим датчиком температури, з подальшим охолодженням струмом повітря. Ущільнювальний поршень з вбудованим датчиком тиску забезпечує герметичність реакційного середовища та контроль тиску до 20 бар. Пристрій оснащений системою електромагнітів для перемішування середовища до 1200 RPM.

Реактор має ряд переваг над аналогами, а саме: програмований покроковий режим нагрівання, незалежність від здатності різних розчинників до поглинання енергії електромагнітного випромінювання, характеризується більшою безпечністю та зручністю у використанні.

Результати та їх обговорення

В даній роботі ми дослідили методи синтезу моно- та дифеніл заміщених саліцилових альдегідів **2a-e** з використанням реактора кондуктивного нагріву з герметичним резервуаром Anton Paar Monowave 50, дотримуючись підходів “зеленої хімії”.

Для синтезу цільових продуктів була використана реакція крос-сполучення Судзукі-Міяури бромопохідних саліцилового альдегіду **1a-e** з фенілбороною кислотою, в присутності паладієвого каталізатора та основи K_2CO_3 (рис. 2) [8].

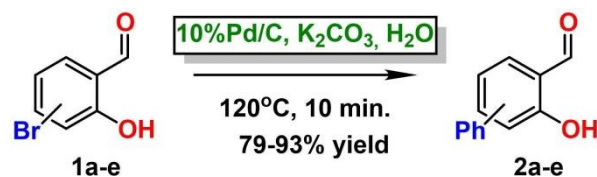


Рис. 2. Схема синтезу моно- та дифеніл заміщених саліцилових альдегідів.

Fig. 2. Scheme of the synthesis of mono- and diphenyl substituted salicylic aldehydes.

Нами проведено серію експериментів з оптимізації умов реакції крос-сполучення, починаючи з підбору параметрів роботи реактора, розчинників для реакційного середовища, каталізатора та температури проведення реакції.

Отже, обрано три температурні програми для застосування у реакторі – повільний нагрів до заданої температури ($T_{\text{макс}}$ = а) 100°C ; б) 120°C ; в) 150°C) протягом 5 хв та утримання при цій температурі ще 10 хв. На рис. 3 наведено відповідні профілі температури (червоний графік) та тиску (синій графік). Реактор забезпечує рівномірний нагрів та підтримання температури та тиску без значних стрибків параметрів (незначний перегрів на початку реакції пов’язаний з процесом розчинення/розплавлення реагентів у суміші).

Результати експерименту наведено на рис. 3 та в таблиці 1.

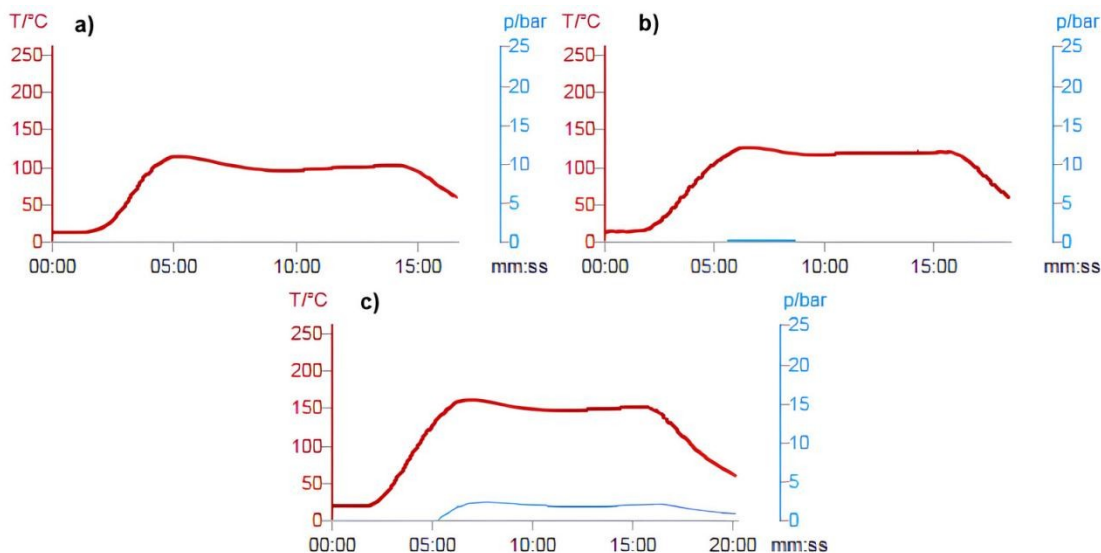


Рис. 3. Профіль температури та тиску зразка реакції у воді а) $T_{\text{макс}} = 100^{\circ}\text{C}$, t утримання = 10 хв; б) $T_{\text{макс}} = 120^{\circ}\text{C}$, t утримання = 10 хв; в) $T_{\text{макс}} = 150^{\circ}\text{C}$, t утримання = 10 хв.

Fig. 3. Temperature and pressure profile of the reaction sample in water а) $T_{\text{max}} = 100^{\circ}\text{C}$, hold time = 10 min; б) $T_{\text{max}} = 120^{\circ}\text{C}$, hold time = 10 min; в) $T_{\text{max}} = 150^{\circ}\text{C}$, hold time = 10 min.

Таблиця 1. Результати експерименту з синтезу 5-фенілсаліцилового альдегіду **2a**.**Table 1.** Results of the experiment on the synthesis of 5-phenylsalicylic aldehyde **2a**.

№	Розчинник	Каталізатор ^{a*}	T, °C	t	Вихід, %	Чистота, % ^{b*}
1	DXN-H ₂ O 10:1	Pd(PPh ₃) ₄	80	12 г	76	99
2	EtOH-H ₂ O 2:1	Pd(PPh ₃) ₄	60	12 г	82	100
3	EtOH-H ₂ O 2:1	10% Pd/C	60	12 г	80	100
4	H ₂ O	Pd(PPh ₃) ₄	80	12 г	74	96
5	H ₂ O	10% Pd/C	80	12 г	71	93
6	H ₂ O	10% Pd/C	к.т.	24 г	60	94
7	DXN-H ₂ O 10:1	Pd(PPh ₃) ₄	120	10 хв	84	99
8	EtOH-H ₂ O 2:1	Pd(PPh ₃) ₄	120	10 хв	94	100
9	H ₂ O	Pd(PPh ₃) ₄	100	10 хв	89	100
10	H ₂ O	Pd(PPh ₃) ₄	120	10 хв	96	100
11	H ₂ O	10% Pd/C	120	10 хв	92	100
12	H ₂ O	10% Pd/C	150	10 хв	90	100

a* кількість використаного каталізатора – 1мг на 500мг вихідної сполуки, що відповідає $7,57 \cdot 10^{-3}$ % mol. Pd(0) для Pd(PPh₃)₄, та $3,78 \cdot 10^{-2}$ % mol. для 10% Pd/C.

b* чистоту зразка визначали за результатом LC/MS спектроскопії після очистки методом флеш-хроматографії, елюент хлороформ.

За результатами проведеної серії експериментів, було визначено, що при проведенні реакції в класичних умовах використання каталізатора Pd(PPh₃)₄ дає дещо кращий результат за 10% Pd/C, але в надкритичних умовах ця різниця майже нівелюється, що надає значну перевагу останньому у зв'язку з екологічною безпечністю, значно нижчою вартістю, придатністю до переробки та повторного використання.

Підвищення температури реакції дозволяє ефективно використовувати воду як розчинник через часткову розчинність у ній саліцилових альдегідів та їх відносно низьку температуру плавлення (36 – 102°C).

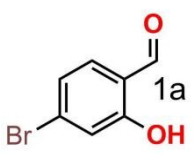
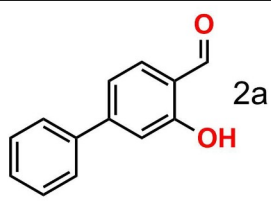
В свою чергу, підвищення температури вище 120°C не покращує результати синтезу, окрім теоретичної можливості скоротити тривалість реакції, втім це не значно зменшує загальні витрати часу на проведення експерименту.

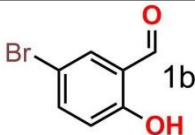
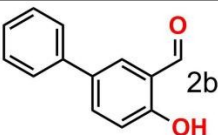
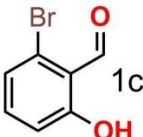
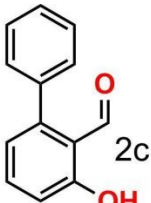
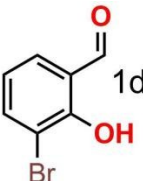
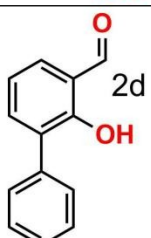
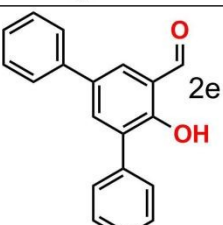
Запропонований нами варіант реакції крос-сполучення фенілсаліцилових альдегідів не потребує використання каталізаторів міжфазного переносу, наприклад, таких як додецилсульфат натрію [11], тетра-н-бутиламмоній бромід [12] тощо, що суттєво покращує екологічність процесу та відповідає принципам зеленої хімії.

Каталізатор 10% Pd/C, у протилежність від фосфінових комплексів Pd(II) та Pd(0), є достатньо стійким до умов реакційного середовища, що дозволяє не використовувати прийнятну для реакції Судзукі-Міяури атмосферу аргону, втім отримувати продукти з кращими виходами та чистотою.

Подальший синтез сполук **2a-e** було проведено за умов, наведених в п.11 таблиці 1, визначеними нами як оптимальних (таблиця 2).

Таблиця 2. Результати синтезу моно- та дифеніл заміщених саліцилових альдегідів **2a-e**.**Table 2.** Results of the synthesis of mono- and diphenyl substituted salicylic aldehydes **2a-e**.

№	Вихідна сполука	Продукт	Вихід, %	Чистота, %
2a			92	100

2b			93	98
2c			83	92
2d			79	90
2e			89	97

Умови реакції: 10% Pd/C (1 мг), K₂CO₃ (**2a-d**: 1.5 екв., **2e**: 3 екв.), PhB(OH)₂ (**2a-d**: 1.2 екв., **2e**: 2.4 екв.), H₂O, 120°C, 10 хв.

Висновок

Нами розроблено ефективний метод синтезу моно- та дифеніл заміщених саліцилових альдегідів **2a-e** за реакцією крос-сполучення Судзукі-Міяури бромопохідних саліцилового альдегіду **1a-e** з фенолбороновою кислотою з використанням реактора Anton Paar Monowave 50. Умови проведення реакції було оптимізовано з урахуванням принципів “зеленої хімії”, з використанням відносно екологічно безпечного каталізатора 10% Pd/C, основи K₂CO₃ та дистильованої води як розчинника. Зазначені сполуки можуть бути використані для синтезу різноманітних функціональних речовин, включно флуоресцентних барвників та барвників-сенсibilізаторів для DSSC з кумариною ланкою.

Експериментальна частина

Загальна методика синтезу сполук **2a-e**.

Бромсаліциловий альдегід **1a-e** (500 мг), K₂CO₃ (**2a-d**: 1.5 екв., **2e**: 3 екв.), PhB(OH)₂ (**2a-d**: 1.2 екв., **2e**: 2.4 екв.) та 4 мл дистильованої H₂O внесли у віалу на 10 мл G10, спорядили комплектним магнітним PTFE якорем, додали каталізатор 10% Pd/C (1 мг), та закрили силіконовою кришкою з PTFE септою. Віалу поставили у нагрівальний пристрій реактора. Реакційну суміш повільно нагріли до 120°C протягом 5 хв та утримували при цій температурі протягом 10 хв. По закінченню реакції провели контрольний аналіз методом ТШХ. Реакційну суміш перенесли в ділільну лійку, нейтралізували оцтовою кислотою до нейтрального середовища та екстрагували три рази по 10 мл хлороформу. Органічний шар поєднали, екстрагували один раз 50 мл дистильованою водою та 30 мл насиченим розчином хлориду натрію. Органічний шар посушили над безводним сульфатом натрію та концентрували на роторному випаровувачі. Очистку проводили методом флеш-хроматографії, елюент — градієнт н-гексан-хлороформ.

3-Гідрокси-[1,1'-біфеніл]-4-карбальдегід (2a). Вихід 92%, білі гольчаті кристали, T_{пл} = 71-72°C. ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.87 (s, 1H), 10.27 (s, 1H), 7.75 (d, J = 8.1 Hz, 1H), 7.72

– 7.64 (m, 2H), 7.55 – 7.46 (m, 2H), 7.46 – 7.40 (m, 1H), 7.31 – 7.22 (m, 2H). ^{13}C NMR (100 MHz, DMSO- d_6) δ 193.07, 162.05, 143.98, 139.20, 131.71, 128.81, 127.73, 126.79, 121.64, 120.64, 114.66. $[\text{MH}]^+ m/z = 199.1$.

4-Гідрокси-[1,1'-біфеніл]-3-карбальдегід (2b). Вихід 93%, жовтий порошок, $T_{\text{пл}} = 72-73^\circ\text{C}$. ^1H NMR (400 MHz, DMSO- d_6) δ 10.70 (s, 1H), 10.32 (s, 1H), 7.92 (d, $J = 2.5$ Hz, 1H), 7.84 (dd, $J = 8.6, 2.5$ Hz, 1H), 7.69 – 7.50 (m, 2H), 7.45 (t, $J = 7.2$ Hz, 2H), 7.34 (t, $J = 6.8$ Hz, 1H), 7.09 (d, $J = 8.4$ Hz, 1H). ^{13}C NMR (100 MHz, DMSO- d_6) δ 194.37, 162.41, 142.58, 132.94, 131.59, 131.44, 128.81, 127.73, 127.54, 120.63, 118.61. $[\text{MH}]^+ m/z = 199.0$.

3-Гідрокси-[1,1'-біфеніл]-2-карбальдегід (2c). Вихід 83%, світло-жовтий смолистий осад, $T_{\text{пл}} = 35-36^\circ\text{C}$. ^1H NMR (400 MHz, DMSO- d_6) δ 11.65 (d, $J = 1.9$ Hz, 1H), 9.81 (s, 1H), 7.68 – 7.59 (m, 1H), 7.54 – 7.41 (m, 5H), 7.02 (d, $J = 8.4$ Hz, 1H), 6.93 (dd, $J = 7.5, 1.1$ Hz, 1H). ^{13}C NMR (100 MHz, DMSO- d_6) δ 192.77, 162.33, 139.68, 137.38, 133.50, 128.97, 128.75, 128.00, 122.26, 118.97, 117.28. $[\text{MH}]^+ m/z = 199.0$.

2-Гідрокси-[1,1'-біфеніл]-3-карбальдегід (2d). Вихід 79%, жовтий смолистий осад, $T_{\text{пл}} = 37-38^\circ\text{C}$. ^1H NMR (400 MHz, DMSO- d_6) δ 11.36 (s, 1H), 10.09 (s, 1H), 7.81 (dd, $J = 7.7, 1.7$ Hz, 1H), 7.68 (dd, $J = 7.5, 1.8$ Hz, 1H), 7.60 – 7.53 (m, 2H), 7.49 – 7.41 (m, 2H), 7.40 – 7.34 (m, 1H), 7.18 (t, $J = 7.6$ Hz, 1H). ^{13}C NMR (100 MHz, DMSO- d_6) δ 193.23, 157.68, 139.37, 132.80, 132.25, 129.51, 128.07, 127.95, 127.75, 121.38, 121.02. $[\text{MH}]^+ m/z = 199.0$.

4'-Гідрокси-[1,1':3',1''-терфеніл]-5'-карбальдегід (2e). Вихід 89%, жовтий порошок, $T_{\text{пл}} = 101-102^\circ\text{C}$. ^1H NMR (400 MHz, DMSO- d_6) δ 11.33 (s, 1H), 10.17 (s, 1H), 8.14 (d, $J = 2.4$ Hz, 1H), 7.93 (d, $J = 2.4$ Hz, 1H), 7.78 – 7.73 (m, 2H), 7.70 – 7.64 (m, 2H), 7.48 (td, $J = 7.7, 2.3$ Hz, 4H), 7.44 – 7.34 (m, 2H). ^{13}C NMR (100 MHz, DMSO- d_6) δ 195.27, 158.09, 141.33, 139.44, 134.57, 133.63, 131.70, 129.37, 129.18, 128.79, 128.06, 128.04, 127.86, 127.11, 120.73. $[\text{MH}]^+ m/z = 275.0$.

Подяка

Роботу виконано за фінансової підтримки НДР МОН України, державний реєстраційний номер 0122U001485.

References

1. Saumya Jos, N.R. Suja. Chiral Schiff base ligands of salicylaldehyde: A versatile tool for medical applications and organic synthesis-A review. *Inorg. Chim. Acta* **2023**, 547, 121323. <https://doi.org/10.1016/j.ica.2022.121323>
2. Manoj Kumar, Atresh Kumar Singh, Alok Kumar Singh, Rajesh K. Yadav, Satyam Singh, Atul P. Singh, Amit Chauhan. Recent advances in 3d-block metal complexes with bi, tri, and tetradentate Schiff base ligands derived from salicylaldehyde and its derivatives: Synthesis, characterization and applications. *Coord. Chem. Rev.* **2023**, 488, 215176. <https://doi.org/10.1016/j.ccr.2023.215176>
3. H. F. Xie, C. Yu, Y. L. Huang, H. Xu, Q. Zhang, X. Sun, X. Feng, C. Redshaw. A turn-off fluorescent probe for the detection of Cu^{2+} based on a tetraphenylethylene-functionalized salicylaldehyde Schiff-base. *Mater. Chem. Front.* **2020**, 4, 1500-1506. <https://doi.org/10.1039/C9QM00759H>
4. Srivastava, Devyani, Kociok-Köhn, Gabriele, Gosavi, Suresh W., Sharma, Vinod K. Copper(ii) and cobalt(iii) Schiff base complexes with hydroxy anchors as sensitizers in dye-sensitized solar cells (DSSCs). *RSC Adv.* **2023**, 13, 9046-9054 <https://doi.org/10.1039/D3RA00344B>
5. Yong Tian, Kaixuan Wang, Houpeng Zhang, Xianming Wu, Chaofan Zhong. Novel polymeric metal complexes of salicylaldehyde schiff base derivative being used for dye sensitizer. *Tetrahedron* **2022**, 113, 132756 <https://doi.org/10.1016/j.tet.2022.132756>
6. Lončarić, Gašo-Sokač, Jokić, Molnar. Recent Advances in the Synthesis of Coumarin Derivatives from Different Starting Materials. *Biomolecules* **2020**, 10 (1), 151. <https://doi.org/10.3390/biom10010151>
7. He, J., Liu, Y., Gao, J., Han, L. New D–D– π -A triphenylamine–coumarin sensitizers for dye-sensitized solar cells. *Photochem. Photobiol. Sci.* **2017**, 16 (7), 1049–1056. <https://doi.org/10.1039/C6PP00410E>

8. Das, S. G., Doshi, J. M., Tian, D., Addo, S. N., Srinivasan, B., Hermanson, D. L., Xing, C. Structure–Activity Relationship and Molecular Mechanisms of Ethyl 2-Amino-4-(2-ethoxy-2-oxoethyl)-6-phenyl-4H-chromene-3-carboxylate (sHA 14-1) and Its Analogues. *J. Med. Chem.* **2009**, 52 (19), 5937–5949. <https://doi.org/10.1021/jm9005059>
9. Minutolo, F., Bellini, R., Bertini, S., Carboni, I., Lapucci, A., Pistolesi, L., Macchia, M. Monoaryl-Substituted Salicylaldoximes as Ligands for Estrogen Receptor β . *J. Med. Chem.* **2008**, 51 (5), 1344–1351. <https://doi.org/10.1021/jm701396g>
10. Aktoudianakis, E., Chan, E., Edward, A. R., Jarosz, I., Lee, V., Mui, L., Dicks, A. P. “Greening Up” the Suzuki Reaction. *J. Chem. Educ.* **2008**, 85 (4), 555. <https://doi.org/10.1021/ed085p555>
11. Alshammari, H. M., Aldosari, O. F., Alotaibi, M. H., Alotaibi, R. L., Alhumaimess, M. S., Morad, M. H., Adil, S. F., Shaik, M. R., Islam, M. S., Khan, M., Alwarthan, A. Facile Synthesis and Characterization of Palladium@Carbon Catalyst for the Suzuki-Miyaura and Mizoroki-Heck Coupling Reactions. *Applied Sciences* **2021**, 11 (11), 4822. <https://doi.org/10.3390/app11114822>
12. Ang, J. W. J. Integrating Green Chemistry into Teaching Laboratories: Aqueous Suzuki–Miyaura Cross-Coupling Reaction Using a Recyclable Fluorous Precatalyst. *J. Chem. Educ.* **2021**, 98 (1), 203–207. <https://doi.org/10.1021/acs.jchemed.0c00072>
13. Vargass, B. P., Rosa, C. H., Rosa, D. da S., Rosa, G. R. “Green” Suzuki-Miyaura cross-coupling: An exciting mini-project for chemistry undergraduate students. *Educ. Quim.* **2016**, 27 (2), 139–142. <https://doi.org/10.1016/j.eq.2015.11.006>
14. D. Obermayer, D. Znidar, G. Glotz, A. Stadler, D. Dallinger, C. O. Kappe. Design and Performance Validation of a Conductively Heated Sealed-Vessel Reactor for Organic Synthesis. *J. Org. Chem.* **2016**, 81, 11788–11801. <https://doi.org/10.1021/acs.joc.6b02242>

Надіслано до редакції 06.11.2024

Прийнято до друку 19.12.2024

P. V. Trostianko*, V. V. Pashko†, S. M. Kovalenko*. Synthesis of Mono- and Diphenyl Substituted Salicylic Aldehydes, Important Building Blocks for the Synthesis of Fluorescent Dyes and Dyes-Sensitizers for DSSC

*V.N. Karazin Kharkiv National University, School of Chemistry, 4 Svobody sq., 61022, Kharkiv, Ukraine

†Donau Lab Ukraine LLC, Stratehichne Hwy 16, Kyiv, 03028, Ukraine

Salicylic aldehydes are of interest for the synthesis of many biologically active compounds, ligands, fluorescent probes and functional dyes-sensitizers for DSSCs, etc.

In this work, we have developed a method for the synthesis of mono- and diphenyl substituted salicylic aldehydes by the Suzuki-Miyaura cross-coupling reaction of the corresponding bromosalicylaldehydes with phenylboronic acid, in order to study the possibilities of modifying the structure of functional dyes. A series of experiments were performed on the variation of synthesis conditions, replacement of solvents and catalysts both under conditions of convection heating at atmospheric pressure and using a conductive heating reactor with a sealed vessel Anton Paar Monowave 50.

In the process of searching for optimal reaction conditions, great attention was paid to the modern requirements of "green chemistry". The best result was obtained by carrying out the Suzuki-Miyaura cross-coupling reaction of bromosalicylaldehydes with phenylboronic acid in water using the $\text{Pd}(\text{PPh}_3)_4$ catalyst. This method was improved by using 10% Pd/C as a catalyst, which has a number of advantages, namely: greater availability, ease of storage and use, lower cost and environmental safety. Also, the 10% Pd/C catalyst is more resistant to environmental influences, which allowed the reactions to be carried out without the use of an inert atmosphere in a conductive heating reactor with a sealed tank Anton Paar Monowave 50.

Keywords: organic synthesis, salicylaldehydes, cross-coupling, Suzuki-Miyaura reaction, green chemistry, eco-friendly synthesis, palladium catalyst, Anton Paar Monowave 50 reactor, fluorescent dyes, DSSC.

Received 06.11.2024

Accepted 19.12.2024

Kharkiv University Bulletin. Chemical Series. Issue 43 (66), 2024

ЕТИЧНІ НОРМИ ПУБЛІКАЦІЇ НАУКОВИХ РЕЗУЛЬТАТІВ ТА ЇХ ПОРУШЕННЯ.

Редакційна колегія робить все можливе для дотримання етичних норм, прийнятих міжнародним науковим товариством, і для запобігання будь-яких порушень цих норм. Така політика є важливою умовою плідної участі журналу в розвитку цілісної системи знань в галузі хімії та суміжних галузях. Діяльність редакційної колегії значною мірою спирається на рекомендації Комітету з етики наукових публікацій (Committee of Publication Ethics), а також на цінний досвід міжнародних журналів та видавництв. Подання статті на розгляд означає, що вона містить отримані авторами нові нетривіальні наукові результати, які раніше не були опубліковані. Кожну статтю рецензують щонайменше два експерти, які мають усі можливості вільно висловити мотивовані критичні зауваження щодо рівня та ясності представлення матеріалу, його відповідності профілю журналу, новизни та достовірності результатів. Рекомендації рецензентів є основою для прийняття остаточного рішення щодо публікації статті. Якщо статтю прийнято, вона розміщується у відкритому доступі; авторські права зберігаються за авторами. За наявності будь-яких конфліктів інтересів (фінансових, академічних, персональних та інших), учасники процесу рецензування мають сповістити редакційну колегію про це. Всі питання, пов'язані з можливим плагіатом або фальсифікацією результатів ретельно обговорюються редакційною колегією, рівно як спори щодо авторства та доцільність роздроблення результатів на невеличкі статті. Доведені плагіат чи фальсифікація результатів є підставами для безумовного відхилення статті.

STATEMENT ON THE PUBLICATION ETHICS AND MALPRACTICE. The Editorial Board has been doing its best to keep the ethical standards adopted by the world scientific community and to prevent the publication malpractice of any kind. This policy is considered to be an imperative condition for the fruitful contribution of the journal in the development of the modern network of knowledge in chemistry and boundary fields. The activity of the Editorial Board in this respect is based, in particular, on the recommendations of the Committee of Publication Ethics and valuable practice of world-leading journals and publishers. The submission of a manuscript implies that it contains new significant scientific results obtained by authors that were never published before. Each paper is peer reviewed by at least two independent experts who are completely free to express their motivated critical comments on the level of the research, its novelty, reliability, readability and relevance to the journal scope. These comments are the background for the final decision about the paper. Once the manuscript is accepted, it becomes the open-access paper, and the copyright remains with authors. All participants of the review process are strongly asked to disclose conflicts of interest of any kind (financial, academic, personal, etc.). Any indication of plagiarism or fraudulent research receives extremely serious attention from the side of the Editorial Board, as well as authorship disputes and groundless subdivision of the results into several small papers. Confirmed plagiarism or fraudulent research entail the categorical rejection of the manuscript.

ІНФОРМАЦІЯ ДЛЯ АВТОРІВ. Журнал публікує статті російською, англійською та українською мовами. До публікації приймаються: огляди (за погодженням з редколегією); оригінальні статті, обсяг 6-10 журнальних сторінок; короткі повідомлення, обсяг до 3 журнальних сторінок. Крім звичайного списку літератури, в статті обов'язково повинен бути другий список, всі посилання якого дані латиницею. Правила підготовки цього списку наведені в розділі «Транслітерація» на сайті журналу. Обидва списки повинні бути повністю ідентичні. При рецензуванні статей один з критеріїв - наявність посилань на публікації останніх років. Стаття обов'язково повинна містити резюме російською, українською та англійською мовами. У всіх трьох необхідно вказати назву статті, прізвища авторів і ключові слова. Орієнтовний обсяг резюме - 1800 знаків (без урахування заголовку і ключових слів). Редакція приймає електронний (MS Word) і два роздрукованих (для харків'ян) тексту рукопису. Адреси вказані в розділі «Контакти» на сайті журналу. Супровідний лист до статті, виправленої відповідно до зауважень рецензента, повинен містити відповіді на всі зауваження. Подається електронний і один роздрукований (для харків'ян) варіант. Рукописи, які пройшли рецензування, прийняті до публікації і оформлені відповідно до правил для авторів, приймаються у форматі doc (не docx) електронною поштою (chembull@karazin.ua). Роздрукований варіант не потрібен. Докладніша інформація розміщена на сайті журналу <http://chembull.univer.kharkov.ua>.

INFORMATION FOR AUTHORS. Papers in Ukrainian, Russian and English are published. These may be invited papers; review papers (require preliminary agreement with Editors); regular papers; brief communications. In preparing the manuscript it is mandatory to keep the statement on the publication ethics and malpractice, which can be found on the web-site and in each issue. The article should contain summaries in English, Russian, and Ukrainian. In all three it is necessary to indicate the title of the article, the names of the authors and the keywords. The approximate volume of summary is 1800 characters (excluding the title and key words). The help in translation is provided by request for foreign authors.. Any style of references is acceptable, but all references within the paper must be given in the same style. In addition, the second, transliterated, list of references is required if at least one original reference is given in Cyrillic. See section "Transliteration" of the web-site for details. Please use papers of previous issues as samples when prepare the manuscript. The MS Word format is used. Standard fonts (Times New Roman, Arial, Symbol) are preferable. Figures and diagrams are required in vector formats. Figure captions are given separately. All figures, tables and equations are numbered. Please use MS Equation Editor or MathType to prepare mathematical equations and ISIS Draw to prepare chemical formulas and equations. The decimal point (not coma) is accepted in the journal. Please avoid any kind of formatting when prepare the manuscript. Manuscripts may be submitted to the Editor-in-Chief via e-mail chembull@karazin.ua. For more detailed information see the journal web-site <http://chembull.univer.kharkov.ua>.

Наукове видання

Вісник
Харківського національного університету
імені В.Н. Каразіна

Серія «Хімія»
Вип. 43 (66)
Збірник наукових праць
Українською та англійською мовами.

Технічний редактор:
Д.О. Анохін

Підписано до друку «27» грудня 2024. Формат 60х84/8.
Зам. № 54/24. Ум.-друк. Арк. 3,83 Обл.-вид. арк. 4,79.
Тираж 100 пр. Ціна договірна.

61022, Харків, майдан Свободи, 4
Харківський національний університет імені В.Н. Каразіна,
Видавництво Харківського національного університету імені В.Н. Каразіна

Надруковано: ХНУ імені В.Н. Каразіна
61022, м. Харків, майдан Свободи, 4.
Тел.: 705-24-32

Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.09