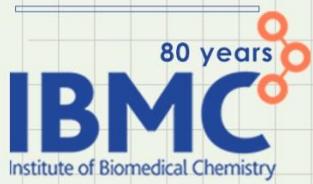




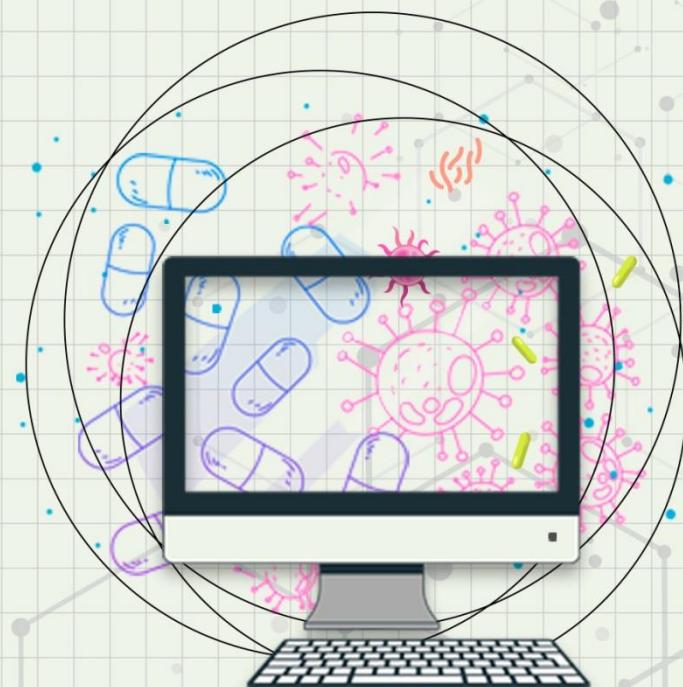
Russian Academy
of Sciences



XXX Symposium on Bioinformatics

and Computer-Aided Drug Discovery

PROCEEDINGS BOOK



Institute of Biomedical Chemistry
Moscow, Russia (Virtual), September 16-18 2024

Russian Academy of Sciences
Ministry of Science and Higher Education of Russian Federation
Institute of Biomedical Chemistry
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
of the Russian Academy of Sciences

Committee

Organization committee:

Vladimir Poroikov (Chairman)
Institute of Biomedical Chemistry, Russia

Roman Efremov (Chairman)
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Russia

Scientific committee:

James Devillers
CTIS, Rillieux La Pape, France

Frank Eisenhaber
A*STAR Bioinformatics Institute, Singapore

Rajesh Kumar Goel
Punjabi University - Patiala, India

Athina Geronikaki
Aristotle University of Thessaloniki, Greece

Alexander Kel
geneXplain GmbH, Germany

Alexey Lagunin
Pirogov Russian National Research Medical University, Russia

Uko Maran
University of Tartu, Estonia

José L. Medina-Franco
National Autonomous University of Mexico, Mexico

Kyoung Tai No
Yonsei University, Republic of Korea

Tudor Oprea
University of New Mexico, USA

Dmitry Osolodkin
Chumakov Federal Scientific Center for Research and Development of Immune-and-Biological Products of Russian Academy of Sciences

Vladimir Palyulin
Faculty of Chemistry - Lomonosov Moscow State University, Russia

Mikhail Panteleev

Center for Theoretical Problems of Physico-Chemical Pharmacology of the Russian Academy of Sciences,
Russia

Kunal Roy

Jadavpur University, India

Garikapati Narahari Sastry

CSIR-North East Institute of Science and Technology, India

Marcus Scotti

Universidade Federal da Paraíba, Brasil

Hanoch Senderowitz

Bar-Ilan University, Israel

Dmitry Shulga

Faculty of Chemistry - Lomonosov Moscow State University, Russia

Vyacheslav Tribulovich

Saint-Petersburg State Institute of Technology, Russia

Anna Tsantili-Kakoulidou

National & Kapodistrian University of Athens, Greece

Pavel Vassiliev

Volgograd State Medical University, Russia

Weiliang Zhu

Shanghai Institute of Materia Medica, Chinese Academy of Sciences, China

Roman Zubarev

Karolinska Institutet, Sweden

Local Organizing Committee:**Tatyana Pleshakova** (Chair)

Institute of Biomedical Chemistry

Polina Savosina (Secretary)

Institute of Biomedical Chemistry

Dmitry Druzhilovskiy

Institute of Biomedical Chemistry

Anastasia Rudik

Institute of Biomedical Chemistry

Vladlen Skvortsov

Institute of Biomedical Chemistry

Kira Stefanovich

Institute of Biomedical Chemistry

DOI: 10.18097/BCADD2024

ISBN: 978-5-900760-20-9

PROCEEDINGS BOOK OF THE XXX SYMPOSIUM
“BIOINFORMATICS AND COMPUTER-AIDED DRUG
DISCOVERY” – Moscow: Institute of
Biomedical Chemistry, 2024

The materials of the XXX International Symposium “Bioinformatics and Computer-Aided Drug Discovery” (Virtual, 16-18 September 2024) are presented. This Symposium is dedicated to the emerging challenges and opportunities for *in silico* drug discovery. Contemporary fields of biomedical science devoted to the analysis of normal and pathological states of the organism and revealing the pathological processes at the cellular and molecular levels are discussed.

The main topics include: development and practical application of computational methods for finding and validation of new pharmacological targets, *in silico* design of potent and safe pharmaceutical agents, optimization of the structure and properties of drug-like compounds, rational approaches to the utilization of pharmacotherapeutic remedies in medical practice.

This information will be useful for researchers whose investigations are dedicated to creating computational methods and their application to drug research and development using bio- and chemoinformatics methods based on post-genomic technologies. It can also be useful for undergraduate, graduate, and postgraduate students specializing in the relevant fields.

Responsible editors: Corr. Member of Rus. Acad. Sci. V.V. Poroikov, Prof. R.G. Efremov



Dear Colleagues,

On behalf of the Organizing Committee and the Administration of the Institute of Biomedical Chemistry, I am very happy to welcome the participants of the XXX Symposium “Bioinformatics and Computer-Aided Drug Discovery”.

Being initiated by my teacher, the Full Member of the Russian Academy of Sciences, Professor Alexander Archakov, the first Symposium hold in the framework of the II Russian National Congress “Man and Drugs” in 1995. Since then, this meeting has been held annually, organized by the Corresponding Member of Russian Academy of Sciences, Professor Vladimir Poroikov, leading Russian scientist in the field of medical bioinformatics and computer-aided drug discovery, jointly with the Full Member of Russian Academy of Sciences, Professor Nikolay Zefirov, and since 2018 - jointly with Professor Roman Efremov.

In 2024 we are celebrating the 300 years of the Russian Academy of Sciences and 80 years of the Institute of Biomedical Chemistry. The XXX Symposium is dedicated to these significant events, which highlights the continuity and importance of the progress of biomedical science in the discovery of new medicines. Analysis of big biomedical data, development of artificial intelligence & machine learning, extension and expansion of new bio- and chemoinformatics methods provide the pre-requisites for better understanding of pathological processes in the organism, and identification of the promising biomarkers and pharmacological targets. Wide application of in silico methods not only saves human efforts, time and money spent on finding of new more safety and potent drugs, but provides the strong basis for integration of information and extraction of new knowledge in this multidisciplinary area.

In 2023, previous XXIX Symposium attracts more than 400 participants including 46 speakers from Belarus, Brazil, Georgia, Germany, India, Indonesia, Ireland, Israel, Japan, Mexico, The Netherlands, Pakistan, Philippines, Russia, Sweden and USA. Almost seventy young researchers took part in the Young Scientists Contest, and the best reports were awarded by the special Diplomas.

This year more than 450 researchers from 50 countries have been registered to take part in the Symposium. To extend the communication between the participants, the Organizing Committee arranged a distributed in time E-Poster Session that is going during not traditional two-three hours but for 18 days, which helps to overcome the limitations of stirring life of the most people working in science and the time difference for the participants from different continents.

The Symposium allows researchers from many countries to exchange by novel ideas and to discuss the main challenges and opportunities in the field of bioinformatics, chemoinformatics, medicinal chemistry and pharmacology, as well as to find new partners for future collaborative projects.

The main topics of the Symposium are especially important due to the active involvement of our Institute in the project “Digital Biodesign and Personalized Health Care”. This is a global project on the digitalization of health monitoring and healthcare management. As part of this project, Institute of Biomedical Chemistry

developing a digital information platform designed to optimize treatment using modern pharmacotherapy and taking into account the individual peculiarities of the patient.

Supporting and developing traditions of this Symposium, we will make the better future together. I would like to thank all the participants of the Symposium and wish you to get a new knowledge due to the nice lectures, to have very fruitful discussions and to find new friends and partners in collaborative projects for a mutual benefit!

Director of the Institute of Biomedical Chemistry,
Doctor of Biological Sciences



Elena Ponomarenko



Dear Colleagues!

We are pleased to welcome you as participants of the Jubilee, XXX Symposium “Bioinformatics and Computer-Aided Drug Discovery” (BCADD-2024).

Our Symposia are dedicated to the emerging challenges and opportunities in computer-aided drug discovery. This series of annual Symposia started in 1995 in the framework of the Second Russian National Congress “Man and Drugs”. Originally, it was initiated by the Full Member of the Russian Academy of Sciences Alexander Archakov and co-chaired by Prof. Vladimir Poroikov. An essential contribution to the organization of the first Symposia was made by Professor Alexis Ivanov. In 1996-2017 the Symposia were co-chaired by the Full Member of the Russian Academy of Sciences Nikolay Zefirov and Prof. Vladimir Poroikov. Significant impact on the next Symposia have been provided by Professor Oleg Raevsky, who has initiated the organization of the Russian Section of the International QSAR Society.

At the BCADD-2024 Symposium two invited lectures will be presented, to remind great achievements of Full Member of the Russian Academy of Sciences Nikolay Zefirov and Professor Oleg Raevsky.

Since 2018 the mutual efforts to organize and perform the Symposia are applied by Prof. Vladimir Poroikov and Prof. Roman Efremov.

Many world-wide famous researchers presented their lectures at the past symposia including Per Artursson (Uppsala University, Sweden), Igor Baskin (Lomonosov Moscow State University, Russia), Artem Cherkasov (University of British Columbia, Canada), Alexey Egorov (Lomonosov Moscow State University, Russia), Frank Eisenhaber (A*STAR Bioinformatics Institute, Singapore), Alexey Finkelstein (Institute of Protein Research, Russia), Viktor Finn (VINITI, Russia), Alexander Gabibov (Institute of Bioorganic Chemistry, Russia), Mikhail Gelfand (Institute for Information Transmission Problems, Russia), Jerome Golebiowski (CNRS GDR “Odorant Odor Olfaction”, France), Viktor Kuzmin (Bogatsky Physico-Chemical Institute, Ukraine), José Medina-Franco (National Autonomous University of Mexico, Mexico), Alexander Nemukhin (Lomonosov Moscow State University, Russia), Kyoung Tai No (Yonsei University, Republic of Korea), Oleg Raevsky (Institute of Physiologically Active Compounds, Russia), Narahari G. Sastry (CSIR-North East Institute of Science and Technology, India), Hanoch Senderowitz (Bar-Ilan University, Israel), Oliver Steck and Andreas Vitte (Tripos, Germany), Igor Tetko (Institute of Structural Biology, Helmholtz Zentrum München, Germany), Vladimir Tumanyan (Institute of Molecular Biology, Russia), Alexandre Varnek (University of Strasbourg, France), Gennady Verkhivker (Chapman University, Irvine, USA), Erik Weber (Environmental Protection Agency, USA), and others.

At the upcoming, XXX Symposium, plenary/keynote lectures and oral talks will be presented by the experienced as well as younger scientists from many countries including Armenia, Australia, China, Germany, Hungary, India, Israel, Mexico, Peru, Russia, Sweden Taiwan, United States and others. Their lectures cover the wide topics dedicated to the development and application of in silico methods for drug

discovery & development. Besides the COVID-19 pandemics is over now, some lessons from *in silico* studies on SARS-CoV-2/COVID-19 will be presented in several talks, which will be useful to combat new biogenic threats in the future.

It is necessary to emphasize that the traditional Young Scientists Contest (YSC) aroused great interest: 41 abstracts by undergraduates and graduates, as well as researchers without scientific degrees under the age of 30 were submitted for participation in the competition. The YSC abstracts were evaluated by seventeen Members of the International Scientific Committee (ISC) including distinguished scientists from Brazil, China, Germany, Greece, India, Israel, Mexico, Russia. Based on the voting of the ISC members and taking into account the geographical diversity of the participants, 18 abstracts have been selected for the flash presentations. The best presentations will be awarded by the Diploma of the First, Second and Third Degrees.

To extend the communication between the participants, the Organizing Committee arranged a distributed in time E-Poster Session that is going during not traditional two-three hours but for 18 days, which helps to overcome the limitations of stirring life of the most people working in science and the time difference for the participants from different continents. The best posters will be awarded by special Diploma.

The Symposium on Bioinformatics and Computer-Aided Drug Discovery are arranged by scientists for scientists; neither commercial entity is involved in preparing the meeting nor registration fee is requested.

Let us use the Symposium discussion platform to exchange original scientific ideas, attractive methodological solutions, and breakthrough multidisciplinary technologies. This is especially important in connection with the current situation in the world, which complicate international scientific and educational relationships, efficient exchange of information and data. These factors have always been at the heart of scientific creativity, especially in the field of biomedical research.

We believe that holding our Symposium in the current conditions, involving the participation of scientists from many countries, will help to develop scientific diplomacy, preserve and increase professional and human relations of colleagues, establish new creative connections, and, as a result, increase the efficiency of computer technologies for the discovery of new medicines. We hope that our Symposium will also contribute to reducing tension in the world. The online format provides unique opportunities for this, including talks given by our authoritative colleagues from all over the world.

Welcome to the sessions of the XXX Symposium on Bioinformatics and Computer-Aided Drug Discovery. We wish you very exciting lectures, fruitful communication and valuable discussions!



Vladimir Poroikov

Corresponding Member of the Russian
Academy of Sciences, Prof. Dr.



Roman Efremov

Prof. Dr.

Scientific Program

Monday September 16, 2024

Chairpersons: Vladimir Poroikov and Roman Efremov

8:30	Opening of the Symposium
------	--------------------------

Plenary lecture

9:00	COMPUTATIONAL DESIGN AND REPURPOSING OF DRUGS FOR CORONAVIRUSES AND DRUG RESISTANT PATHOGENS  David Winkler Monash Institute of Pharmaceutical Sciences, Melbourne, Australia
------	---

Oral presentations

10:00	SUCCESSFUL APPLICATION OF COMPUTING METHODS TO DEVELOPMENT SARS-COV-2 INHIBITORS  Vladimir Sulimov Dimonta Ltd., Research Computing Center, Lomonosov Moscow State University, Moscow, Russia
10:20	PITFALLS OF SARS-COV2 MAIN PROTEASE COVALENT INHIBITION MODELING WITH THE COMBINED QUANTUM AND MOLECULAR MECHANICS APPROACHES  Igor Polyakov Lomonosov Moscow State University, Moscow, Russia
10:40	MODELLING LETHALITY AND TERATOGENICITY OF ZEBRAFISH (DANIO RERIO) DUE TO β -LACTAM ANTIBIOTICS EMPLOYING THE QSTR APPROACH  Aniket Nath Jadavpur University, Kolkata, India

Keynote lectures

11:00	SIXTY YEARS IN SCIENCE OF PROFESSOR NIKOLAY S. ZEFIROV: FROM ORGANIC SYNTHESIS, CONFORMATIONAL ANALYSIS AND REACTION DESIGN - TO MEDICINAL CHEMISTRY AND CADD  Vladimir Palyulin Faculty of Chemistry, Lomonosov Moscow State University, Moscow, Russia
11:30	EXTENDED SIMILARITY INDICES: BENEFITS OF COMPARING MORE THAN TWO OBJECTS SIMULTANEOUSLY. THEORY, SPEED, CONSISTENCY, AND DIVERSITY SELECTION  Károly Héberger HUN-REN Research Centre for Natural Sciences, Budapest, Hungary

Oral presentations

12:00	MOLECULAR MODELING OF BACTERIAL RESISTANCE. THE ROLE OF DYNAMIC BEHAVIOR OF PROTEIN COMPLEXES WITH SUBSTRATES OR INHIBITORS  Maria G. Khrenova Lomonosov Moscow State University, Moscow, Russia
-------	--

12:20	DECIPHERING CAP1 AS A SIGNIFICANT DRUGGABLE TARGET IN COMBATING CANDIDA ALBICANS PATHOGENESIS AND MULTIDRUG RESISTANCE Neha Jaiswal National Institute of Technology Raipur, Raipur, India
12:40	MODELING, SYNTHESIS AND IN VITRO TESTING OF PEPTIDES BASED ON UNUSUAL AMINO ACIDS AS POTENTIAL ANTIBACTERIAL AGENTS Armen Sargsyan Scientific and Production Center «Armbiotechnology» NAS RA, Yerevan, Armenia

13:00-15:00

Lunch break

Chairpersons: Károly Héberger and Vladimir Palyulin

Keynote lectures

15:00	DEVELOPMENT AND APPLICATION OF A WEB-BASED INTEGRATED PLATFORM D3CARP FOR TARGET PREDICTION AND VIRTUAL SCREENING Weiliang Zhu Shanghai Institute of Materia Medica Chinese Academy of Sciences, Shanghai, China
15:30	COMBATING CYSTIC FIBROSIS: COMPUTATIONAL STUDIES ON CFTR Hanoch Senderowitz Bar-Ilan University Faculty of Exact Sciences, Ramat-Gan, Israel

Oral presentations

16:00	ESTIMATION OF RETENTION TIME OF ORGANIC PESTICIDES IN HUMAN MILK USING QSPR AND READ-ACROSS METHODS AN ALTERNATIVE APPROACH TO EXPERIMENTAL HAZARDS ASSESSMENT Ankur Kumar Jadavpur University, Kolkata, India
16:20	MULTI-TARGET NEURAL NETWORK MODEL OF ANXIOLYTIC ACTIVITY OF CHEMICAL COMPOUNDS BASED ON CORRELATION CONVOLUTION OF ENERGY SPECTRA OF MULTIPLE DOCKING Pavel Vassiliev Volgograd State Medical University, Volgograd, Russia
16:40	NETWORK PHARMACOLOGY REVEALED THE POTENTIAL OF BITTER HONEY IN SUPPRESSION OF CEREBRAL MALARIA-INDUCED INFLAMMASOME Michael Daniyan Obafemi Awolowo University, Ile-Ife, Nigeria

Keynote lectures

17:00	STRUCTURAL PHARMACOLOGY OF TRANSIENT RECEPTOR POTENTIAL CHANNELS Arthur Neuberger Columbia University, New York, USA
17:30	FINE MAPPING OF BIOMOLECULAR SURFACES USING THE NEW MOLECULAR SURFACE TOPOGRAPHY (MST) WEB TOOL Yury Trofimov Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia

Oral presentations

18:00	UNIFIED LATIN AMERICAN NATURAL PRODUCT DATABASE LANAPDB Alejandro Gomez Garcia National Autonomous University of Mexico, Mexico City, Mexico
18:20	IS CHEMICAL SPACE OF RUSSIAN VENDOR CATALOGUES SUFFICIENT FOR LEAD DISCOVERY Dmitry Osolodkin Chumakov FSC RnD IBP Russian Academy of Sciences (Institute of Poliomyelitis), Moscow, Russia
18:40	LATIN AMERICAN PHYTOCHEMICAL DERIVATIVES AS PROMISING CANDIDATES FOR GALLSTONE DISEASE THERAPY INSIGHTS FROM MOLECULAR SCREENING, MOLECULAR DOCKING, DENSITY FUNCTIONAL THEORY, AND MOLECULAR DYNAMICS STUDIES Jaime Tamayo Universidad Nacional Mayor de San Marcos, Lima, Peru

Tuesday September 17, 2024

Chairpersons: Kunal Roy and Vladimir Sulimov

Keynote lectures

10:00	CHRONOBIOTICSDB - WORLD FIRST DATABASE OF CIRCADIAN RHYTHM'S PHARMACOLOGICAL MODULATORS Ilya Solovev Pitirim Sorokin Syktyvkar State University, Syktyvkar, Russia
10:30	COCONUT 2.0 DATABASE AND AUTOMATED LITERATURE MINING USING DECIMER.AI Kohulan Rajan Friedrich Schiller University, Jena, Germany

Oral presentations

11:00	MACHINE LEARNING ESTIMATION OF THE SMALL MOLECULE SELECTIVITY INDEX VALUE FOR INFLUENZA VIRUS STRAIN AH1N1 Alexey Egorov National Research Nuclear University MEPhI, Moscow, Russia
11:20	MACHINE LEARNING PREDICTIONS OF COCRYSTAL FORMATION TO ENHANCE ACTIVE PHARMACEUTICAL INGREDIENTS PROPERTIES FOR CANCER PREVENTION Nguyen Quoc Khanh Le Taipei Medical University, Taipei, Taiwan
11:40	CHEMOMETRICS GUIDED LEAD IDENTIFICATION AND DESIGN OF NOVEL ANALOGS WITH TRYPANOTHIONE REDUCTASE ACTIVITY BASED ON 2-AMINOBENZIMIDAZOLE SCAFFOLDS AND MOLECULAR SIMULATIONS FOR ADDRESSING LEISHMANIASIS Arpita Biswas Brainware University, Kolkata, India

Keynote lectures

<p>12:00</p> <p>ON A SIMPLE FRAMEWORK OF DIMENSIONALITY REDUCTION FOR CLASSIFICATION MODELING OF SPARSE ENVIRONMENTAL TOXICITY DATA</p> <p> Kunal Roy</p> <p>Jadavpur University, Kolkata, India</p>
<p>12:30</p> <p>ACTIVATION OF GLYCOGENOLYSIS WITHOUT AN ACTIVATION OF ATP CONSUMPTION CAN CAUSE CELL DEATH</p> <p> Victor Vitvitsky</p> <p>Center for Theoretical Problems of Physico-Chemical Pharmacology Russian Academy of Sciences, Moscow, Russia</p>

13:00-15:00

Lunch break

Chairpersons: Athina Geronikaki and Alexey Lagunin

Young Scientists flash presentations

<p>15:00</p> <p>MOLECULAR DOCKING OF SECONDARY METABOLITE COMPOUND OF KAWISTA (<i>Limonia acidissima</i>) AS HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR-2 (HER-2) INHIBITORS</p> <p> Azka Khoirunnisa</p> <p>Department of Pharmacy, Faculty of Health Science, University of Muhammadiyah Malang, Indonesia</p>
<p>15:10</p> <p>PREDICTION OF PROTEIN SECONDARY STRUCTURES BASED ON SUBSTRUCTURAL MNA DESCRIPTORS OF MOLECULAR FRAGMENTS</p> <p> Oleg Zakharov</p> <p>Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia</p>
<p>15:20</p> <p>STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF HYPOTHETICAL PROTEINS OF LUMPY SKIN DISEASE VIRUS TOWARD IDENTIFICATION OF VACCINE TARGETS</p> <p> Bhaavikka Agarwaal</p> <p>Mahindra University, Hyderabad, India</p>
<p>15:30</p> <p>COMPUTATIONAL EXPLORATION OF NATURAL NOVEL HYBRID MOLECULES AS RAF-1 KINASE ANTAGONISTS FOR BREAST CANCER THERAPEUTICS</p> <p> Navya Aggarwal</p> <p>Immunoncology and Molecular Theragnostics Lab, Centre for Medical Biotechnology, Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida, India</p>
<p>15:40</p> <p>FOOD CHRONOBIOATICS: KEY COMPOUNDS DATABASE FOR CHRONONUTRITION</p> <p> Denis Golubev</p> <p>Pitirim Sorokin Syktyvkar State University, Syktyvkar, Russia</p>
<p>15:50</p> <p>IN-SILICO DESIGN AND STUDY OF NOVEL MANNICH BASE DERIVATIVES AGAINST THE SELECTED TARGETS FOR ITS ANTIMICROBIAL ACTIVITY</p> <p> Guru Aribam Mansi Devi</p> <p>Harsha college of pharmacy, Bengaluru, Karnataka, India</p>
<p>16:00</p> <p>QUANTITATIVE PREDICTION OF HUMAN IMMUNODEFICIENCY VIRUS DRUG RESISTANCE</p> <p> Ekaterina Stolbova</p> <p>National Research University Higher School of Economics, Moscow, Russia</p>
<p>16:10</p> <p>4'-FLUORO-5,7-DIHYDROXYFLAVONE – PIPERAZINE HYBRIDS AS VEGFR-2 INHIBITORS: DESIGN, IN-SILICO STUDY, SYNTHESIS, AND ANTICANCER ACTIVITY</p> <p> Kalyani Thombre</p> <p>Department of Pharmaceutical Science, Smt. Kishoritai Bhoyar College of Pharmacy, Kamptee, Nagpur (MS), India</p>

16:20	DRUG METABOLISM PREDICTION USING GRAPH NEURAL NETWORKS  Nikita Polomoshnov Moscow State University, Moscow, Russia
16:30	MOLECULAR DYNAMIC SIMULATION OF SOME N-N-DISUBSTITUTED PIPERAZINE DERIVATIVES EXHIBITING ANTICHOLINESTERASE ACTIVITY  Viktor Ghamaryan Russian-Armenian University, Institute of Biomedicine & pharmacy, laboratory of structural bioinformatics Yerevan, Armenia
16:40	CREATION OF SAR MODELS FOR PREDICTION OF T-CELL EPITOPEs WITH HUMAN LEUKOCYTE ANTIGENS BASED ON PROTEIN STRUCTURAL FORMULAS  Anton Smirnov Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia
16:50	VIRTUAL SCREENING OF NEW POTENTIAL INSECT EPOXIDASE CYP15A1 INHIBITORS  Polina Yakovets Research Institute for Physical Chemical Problems, Belarusian State University, Minsk, Belarus
17:00	LIGAND-PROTEIN BINDING SITE ANNOTATION USING GRAPH NEURAL NETWORKS  Alexey Ereshchenko The Federal State Unitary Enterprise Dukhov Automatics Research Institute, Moscow, Russia
17:10	CASTOR OIL AND LAXATIVE ACTIVITY: VIRTUAL SCREENING OF RICINOLEIC ACID ANALOGUES  Bayoudh Sirine Pharmaceutical Sciences Department B, College of Pharmacy of Monastir, University of Monastir, Monastir, 5000, Tunisia
17:20	WORLD WIDE APPROVED DRUGS: FROM BIG BIOMEDICAL DATA TO GLOBAL SMALL MOLECULE DRUG DATABASE  Polina Savosina Institute of Biomedical Chemistry, Moscow, Russia
17:30	A GENERAL PROTOCOL FOR THE CONSTRUCTION OF STRUCTURE-ACTIVITY LANDSCAPES OF NON-CANONICAL PEPTIDES  Edgar López-López DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico
17:40	POTENTIAL NEW INHIBITOR MOLECULES FOR SARS-COV-2 PLPRO: AN IN SILICO AND SYNTHETIC APPROACH  Gianfranco Sabadini Organic Chemistry Laboratory, Institute of Chemistry and Biochemistry, Faculty of Sciences, University of Valparaíso. Gran Bretaña 1111, Playa Ancha, Valparaíso, Chile
17:50	AUTOMATED ANALYSIS OF STRUCTURE-MULTIPLE PROPERTY RELATIONSHIPS: IMPACT ON SMARTS  Jesús Israel Espinoza Castañeda DIFAQUIM research group, Department of Pharmacy, School of Chemistry, National Autonomous University of Mexico

Keynote lectures

18:00	AUTOMATING THE DESIGN OF GLYCOMIMETIC AGENTS  Robert J. Woods University of Georgia, Athens, USA
-------	--

18:30	SQUARE ANTIPRISM IS A KEY DETERMINANT FOR POTASSIUM ION SELECTIVITY Anton Chugunov Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry Russian Academy of Sciences, Moscow, Russia
-------	---

Wednesday September 18, 2024

Chairpersons: Alexander Kel and Olga Tarasova

Keynote lectures

9:00	COMPREHENSIVE COMPUTATIONAL SYSTEMS BIOLOGY MODEL OF BLOOD PLATELET SIGNALLING: A TOOL FOR BASIC RESEARCH, DIAGNOSTICS AND PHARMACOLOGY Anastasia Sveshnikova Dmitriy Rogachev National Medical Research. Center of Pediatric Hematology, Oncology, Immunology Ministry of Healthcare of Russian Federation, Moscow, Russia
9:30	ANALYSIS OF CHEMICALS-VIRUS-HOST INTERACTIONS BASED ON LARGE-SCALE BIOMEDICAL TEXT AND DATA MINING Olga Tarasova Institute of Biomedical Chemistry, Moscow, Russia

Oral presentations

10:00	EVALUATION OF NATURAL FLAVONOID COMPOUNDS N AMYLOID- (AB) INHIBITION FOR ALZHEIMER'S DISEASE TREATMENT Arli Aditya Parikesit Indonesia International Institute for Life Sciences, Jakarta, Indonesia
10:20	QSAR MODELS FOR PREDICTION OF CYTOTOXIC IC50 AND GI50 VALUES OF SUBSTANCES IN RELATION TO NON-TUMOR CELL LINES Elena Lisitsa Pirogov Russian National Research Medical University, Moscow, Russia
10:40	DIGEP-PRED 2.0 A WEB-SERVICE FOR PREDICTING DRUG-INDUCED CELL SIGNALING AND GENE EXPRESSION CHANGES Sergey Ivanov Institute of Biomedical Chemistry, Moscow, Russia

Keynote lectures

11:00	CHEMICAL PROTEOMICS FOR OVERCOMING DRUG RESISTANCE Roman Zubarev Karolinska Institutet, Stockholm, Sweden
11:30	HOW MANY DRUG TARGETS DO WE NEED? Alexander Kel geneXplain GmbH, Wolfenbuettel, Germany

Oral presentations

12:00	MOLECULAR MODELING OF HUMAN LINE-1 ORF2 PROTEIN Anna Kulakova Lomonosov Moscow State University, Moscow, Russia
-------	--

12:20	<p>THE QSAR STUDY OF THE HYDROLYSIS OF DINITROSYL IRON-SULFUR COMPLEXES ✉ Victor Luzhkov Federal Research Center of Problem of Chemical Physics and Medicinal Chemistry Russian Academy of Sciences, Moscow, Russia</p>
12:40	<p>IN SILICO EVALUATION OF THE MUTAGENICITY, GENOTOXICITY, AND CARCINOGENICITY OF LEVETRACETAM, A NEW-GENERATION ANTIEPILEPTIC ✉ Sezen Yilmaz Sarialtin Ankara University, Ankara, Turkey</p>

13:00-16:00

Lunch break

Chairpersons: Maria Khrenova and Dmitry Shulga

Oral presentations

16:00	<p>CONSTRUCTING BAYSIAN NETWORKS TO DETERMINE THE RISKS OF DRUG INTERACTIONS BASED ON INSTRUCTIONS ✉ Yurii Titov Plekhanov Russian University Of Economics, Moscow, Russia</p>
16:20	<p>SMALL-MOLECULE ACTIVATORS FOR GLUCOSE-6-PHOSPHATE DEHYDROGENASE (G6PD) USING MACHINE LEARNING APPROACHES ✉ Madhu Sudhana Saddala University of California, Irvine, USA</p>
16:40	<p>OLEG A. RAEVSKY - SCIENTIST, TEACHER, PERSON ✉ Oleg Tinkov Shevchenko Transnistria State University, Tiraspol, Moldova</p>

Keynote lectures

17:00	<p>ALPHAFOLD-LIKE MODELS FOR SMALL MOLECULE STRUCTURE PREDICTION ✉ Petr Popov Constructor University, Bremen, Germany</p>
17:30	<p>PROTEIN 3D STRUCTURE IDENTIFICATION BY ALPHAFOLD A PHYSICS-BASED PREDICTION OR RECOGNITION USING HUGE DATABASES? ✉ Alexei Finkelstein Institute of Protein Research Russian Academy of Sciences, Pushchino, Russia</p>

Plenary lectures

18:00	<p>EXPLORING THE CHEMICAL SPACE AND MULTIVERSE OF FOOD CHEMICALS AND NATURAL PRODUCTS ✉ Jose Medina-Franco Universidad Nacional Autonoma de Mexico, Mexico City, Mexico</p>
18:40	<p>Closure of the XXX Symposium on Bioinformatics and Computer-Aided Drug Discovery</p>

PLENARY LECTURES

COMPUTATIONAL DESIGN, AND REPURPOSING OF DRUGS FOR CORONAVIRUSES AND DRUG RESISTANT PATHOGENS

D.A. Winkler^{1,2,3}

¹School of Biochemistry and Chemistry, La Trobe Institute for Molecular Science, La Trobe University, Australia

²Department of Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Australia

³School of Pharmacy, University of Nottingham, UK

In the last 20 years, the world has been threatened with three different coronaviruses (CoV): Severe Acute Respiratory Syndrome coronavirus (SARS-CoV); Middle East Respiratory Syndrome coronavirus (MERS-CoV); and SARS-CoV-2. These viruses pose serious global pandemic threats, with estimated case fatality rates of 15% for SARS, 34% for MERS and 1-3% for SARS-CoV-2. Sadly, the SARS-CoV-2 pandemic has infected >500M people and killed an estimated 15M so far and severely disrupted world economies. The pandemic created an urgent need to understand where the virus came from and which domestic and farm animals may harbour the virus, and to find new drugs to treat COVID-19, the disease caused by this coronavirus. This will not be the last coronavirus to threaten humanity, hence we need better tools to track virus origin, and to identify drugs active against future coronavirus threats. Neglected tropical diseases also continue to create high levels of morbidity and mortality in a sizeable fraction of the world's population, despite ongoing research into new treatments. Some of the most important technological developments that have accelerated drug discovery for diseases of affluent countries have not flowed down to neglected tropical disease drug discovery.

Computational methods are increasingly accessible to all scientists, in stark contrast to many other advanced technologies. They are beginning to make significant inroads into discovery of drugs for neglected tropical diseases due to the increasing availability of large databases that can be used to train ML models, increasing accuracy of these methods, lower entry barrier for researchers, and widespread availability of public domain machine learning codes.

I will discuss in silico modelling and screening approaches used to estimate the SARS-CoV-2 susceptibility of humans and other important animal species. Surprisingly, we found humans to have the highest susceptibility. I will also illustrate how state-of-the-art computational methods can rapidly identify drugs from existing drug libraries that may be repurposed to treat COVID-19 patients. Individual drug protection against CoVs may be short-lived, given their rapid mutation rates and the development of drug resistance. Thus, CoV drugs should hit multiple targets within viruses to minimize resistance. We describe how this computational screening pipeline can be expanded in the future to identify drugs with broad-spectrum activity against a wide diversity of coronaviruses. I will describe the current state of development of in silico CoV drug screening, the challenges and pitfalls of these approaches, and how such methods may be used to develop drugs for future CoV pandemics and neglected tropical diseases generally.

This paper also discusses the use of computational methods such as molecular docking and molecular dynamics simulations and, increasingly, machine learning to discovery of new drugs and repurposing of existing drugs for pandemics, resistant pathogens, and neglected tropical diseases. I also provide a perspective on development of machine learning methods and the use of other artificial intelligence methods for drug discovery. The current roadblocks to, and likely impacts of, synergistic new technological developments in ML methods for drug discovery in the future are also discussed.

1. Piplani, S. et al. (2022) in *Frontiers of COVID-19: Scientific and Clinical Perspectives of the Novel SARS-CoV-2*, Adibi, S., Griffin, P., Sanicas, M., Rashidi, M., & Lanfranchi, F. (eds.), Springer Nature.
2. Piplani, S., et al. (2022) *Int. J. Mol. Sci.*, 23, 7704 .
3. Piplani, S., et al. (2022) *Front. Mol. Biosci.*, 9, 781039.
4. Piplani, S., et al. (2021) *Mol. Biomed.*, 2, 28.
5. Piplani, S., et al. (2021) *Sci. Rep.*, 11, 13063.
6. Winkler, D.A. (2024) *J. Math. Chem.*. DOI: 10.1007/s10910-023-01568-3
7. Winkler, D.A. (2022) *Expert. Opin. Drug Dis.*, 17(9), 925-927.
8. Winkler, D.A. (2022) *Nat. Mach. Intell.*, 4(2), 102-103 .
9. Winkler, D.A. (2021) *Front. Chem.*, 9, 39.
10. Ung, P., Winkler, D.A. (2011) *J. Med. Chem.*, 54(5), 1111-1125.
11. Wijffels, G., et al. (2011) *J. Med. Chem.*, 54 (13): 4831-4838.

EXPLORING THE CHEMICAL SPACE AND MULTIVERSE OF FOOD CHEMICALS AND NATURAL PRODUCTS

J. F. Avellaneda-Tamayo, J.L. Medina-Franco

*DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry,
Universidad Nacional Autónoma de México, Mexico City, Mexico*

Natural products correspond to secondary metabolites of natural origin, whether plant, animal, or other sources, which are synthesized especially in low concentrations by the living being in question. They are known as complex chemical compounds thanks to millions of years of evolution that have favored their complexity and specificity, both at a structural level and in their participation in biochemical processes, which determines human interest in their study.

The cheminformatics study of natural chemical libraries allows us to accurately and systematically describe the physicochemical, constitutional and structural characteristics of the molecules that belong to the database, and thus know in advance the possibility of finding suitable candidates for the advancement of the processes in the development of molecules with different functionalities. This criterion is applied both in the field of drug design and in the field of new materials, food additives, agrochemicals, among others.

This work is an exhaustive compilation of the cheminformatics analyzes carried out on two different libraries of natural origin, such as FooDB [1], and its potential approximation to the chemical space of compounds from food, and NAPROC-13, a database for dereplication of natural products using ^{13}C -NMR [2].

We thank the support of DGAPA, UNAM, Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), grant no. IV200121. We also thank the Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC), UNAM, for the computational resources to use Mitzli supercomputer at UNAM under project LANCAD-UNAM-DGTIC-335.

1. Avellaneda-Tamayo, J.F., Chávez-Hernández, A.L., Prado-Romero, D.L., Medina-Franco, J.L. (2024) *J. Chem. Inf. Model.*, 64(4), 1229-1244.
2. Avellaneda-Tamayo, J.F., Agudo-Muñoz, N.A., Sánchez-Galán, J.E., López-Pérez, J.L., Medina-Franco, J.L. (2024) *ChemRxiv*, DOI: 10.26434/chemrxiv-2024-spksf

KEYNOTE LECTURES

SIXTY YEARS IN SCIENCE OF PROFESSOR NIKOLAY S. ZEFIROV: FROM ORGANIC SYNTHESIS, CONFORMATIONAL ANALYSIS AND REACTION DESIGN - TO MEDICINAL CHEMISTRY AND CADD

V.A. Palyulin

Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

Nikolay S. Zefirov was born in Yaroslavl, Russia, on 13th September, 1935. He obtained a M.S. degree in Chemistry from Moscow State University in 1958, then PhD in 1961 and a Dr. Sci. degree in 1966. From his first appointment as a research chemist, Nikolay Zefirov has held various teaching and research appointments in the Department of Chemistry. In 1971 he was appointed as Head of the Heterocyclic Compounds Laboratory. At that stage, the work of Zefirov's laboratory enjoyed recognition for research excellence both inside and outside Russia. He was elected a Corresponding Member of the Russian Academy of Sciences in 1981 and a Full Member (Academician) in 1987. In 1989 professor Zefirov was appointed as Director of the Institute of Physiologically Active Compounds, retaining at the same time his university positions. In addition, in 1993 he headed the Organic Chemistry Division at Moscow State University and in 2014 had founded the Division of Medicinal Chemistry and Advanced Organic Synthesis. He carried out a colossal research, pedagogical, and organizational job in his three affiliations: Institute of Physiologically Active Compounds, Moscow State University, and the Zelinsky Institute of Organic Chemistry where he headed the Laboratory of Mathematical Chemistry and Computer-Aided Organic Synthesis. Professor Zefirov was also a founding member of the International Academy of Mathematical Chemistry.

The high academic ranks reflect the outstanding contribution Professor Zefirov has made to science. The record of his research output includes *inter alia* synthesis of various polycyclic and cage hydrocarbons, which were either unknown before or hard to synthesize. A new class of spiro-condensed cyclopropanes – triangulanes – was discovered and synthetic methods were devised for chain, branched, and cyclosubstituted triangulanes. An unprecedented phenomenon of competitive covalent binding of nucleofugal anions in carbocationic processes was opened up, which resulted in dozens of new reactions, the introduction of many new reagents such as μ -oxodiphenyliodosotriflate (Zefirov reagent), and the synthesis of new and hard to obtain covalent perchlorates and fluorosulfates. His studies in conformational analysis and dynamic stereochemistry are well known and became textbook material. He discovered new conformational effects such as the “hockey sticks” effect, the effect of coordination stabilization of unstable conformations and the effect of boat-like conformations in diheterobicyclononanes. A general equation relating the product ratio with the rate constants for conformationally mobile systems was derived, the well-known cases of Curtin-Hammett and conformational-equilibria control being the limiting cases. Algebraic chirality criteria for point 3D configurations and their superpositions were deduced. The general approach to the quantitative description of cyclic molecule conformations based on the functions of torsional angles was suggested and successfully used in numerous stereochemical studies.

The research interests of Nikolay Zefirov included such rapidly growing branches as medicinal chemistry and computer-aided molecular design. In fact, Professor Zefirov was a founder of systematic *in silico* medicinal chemistry studies in Russia. He also organized the official approval of the specialty “medicinal chemistry” and its teaching in Russia both for students and PhD students at universities. It is necessary to mention the development under the guidance of Professor Zefirov of advanced methodologies and computer programs for the design of new organic reactions and organic compounds with desirable properties; development of new molecular descriptors for QSAR, application of artificial neural networks in QSAR/QSPR, development of the molecular field topology analysis method and the solution of inverse problem in QSAR for various descriptors. A number of potent anticancer agents were discovered based on “structure–activity” analysis, molecular modelling, and unique laboratory cancer models. The construction of the first 3D molecular models of all glutamate, GABA_A and GABA_C, adenosine, melatonin and other receptors, detailed analysis of binding sites and molecular dynamics simulations combined with *de novo* design resulted in the development of highly active

neuroprotective agents and cognition enhances. Molecular modelling was also very instrumental in repurposing of antihistamine drug Dimebon as a promising candidate for the cure of neurodegenerative diseases.

The current series of symposia was organized thirty years ago by Professor Zefirov jointly with Professor Poroikov and they co-chaired them together for more than two decades.

Professor Zefirov passed away in 2017, but hundreds of researchers and professors all over the world who were working with him continue to develop his scientific ideas.

EXTENDED SIMILARITY INDICES: BENEFITS OF COMPARING MORE THAN TWO OBJECTS SIMULTANEOUSLY. THEORY, SPEED, CONSISTENCY, AND DIVERSITY SELECTION

R.A. Miranda-Quintana¹, D. Bajusz², A. Rácz³, K. Héberger³

¹Department of Chemistry, University of Florida, Gainesville, USA

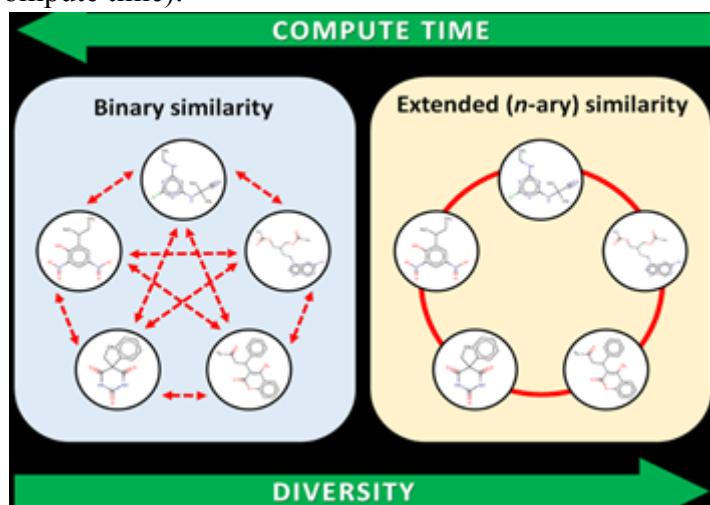
²Medicinal Chemistry Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary

³Plasma Chemistry Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary

Determining the similarity of molecules is essential in drug design. The coding of structures with 0- and 1-bit sequences is called fingerprints and the similarities are determined by pairwise comparisons. The most commonly used measure of similarity is the asymmetric Tanimoto coefficient.

We introduced multiple, n -fold fingerprint comparisons instead of the usual pairwise (binary) comparison. As parallel to dichotomous (binary) coefficients (such as e.g., the Tanimoto coefficient), extended indices were created (defined). Moreover, we explored their properties and optimized the number of comparisons. We have shown the dependence of the similarity coefficients on the length of the fingerprints encoding the structure of the molecules and found that it is worth using weightings. We have determined which coefficients should be used optimally (extended Baroni-Urbani-Buser and extended Faith with 1-similarity counters) [1].

We have demonstrated that comparisons based on new types of extended similarity measures allow the calculation of similarities significantly faster and that it can also calculate the diversity of databases much more efficiently than traditional methods. Internal and external consistency analysis shows whether the n -fold and binary indices rank molecules in the same way. The method of t-distributed stochastic neighbor embedding (t-SNE) has shown that the compactness of different compounds can also be better specified with the new indices than when using traditional binary coefficients. Our indices can also be applied to agglomerative hierarchical clustering [2]. A comparison of binary- and the extended similarity indices can be found below (larger diversity and less compute time):



This study was supported by the National Research, Development and Innovation Office of Hungary (OTKA, contracts K134260 and PD134416) and by the University of Florida, start-up grant.

1. Miranda-Quintana, R.A., Bajusz, D., Rácz, A., Héberger, K. (2021) *Journal of Cheminformatics*, 13, 33.
2. Miranda-Quintana, R.A., Bajusz, D., Rácz, A., Héberger, K. (2021) *Journal of Cheminformatics*, 13, 32.

DEVELOPMENT AND APPLICATION OF A WEB-BASED INTEGRATED PLATFORM D3CARP FOR TARGET PREDICTION AND VIRTUAL SCREENING

W. Zhu

Shanghai Institute of Materia Medica, Shanghai, China

Resource- and time-consuming biological experiments are unavoidable in traditional drug discovery, which have directly driven the evolution of various computational algorithms and tools for drug-target interaction (DTI) prediction. For improving the prediction reliability, we integrated the multiple-conformation based docking, 2D/3D ligand similarity search and deep learning approaches to construct a comprehensive webserver, namely D3CARP, for target prediction and virtual screening. Specifically, 9,352 conformations with positive control of 1970 targets were used for molecular docking, and approximately 2 million target-ligand pairs were used for 2D/3D ligand similarity search and deep learning. Besides, the positive compounds were added as references, and related diseases of therapeutic targets were annotated for disease-based DTI study. The accuracies of the molecular docking and deep learning approaches were 0.44 and 0.89, respectively. And the average accuracy of five ligand similarity searches was 0.94. The strengths of D3CARP encompass the support for multiple computational methods, ensemble docking, utilization of positive controls as references, cross-validation of predicted outcomes, diverse disease types, and broad applicability in drug discovery. The D3CARP is freely accessible at <https://www.d3pharma.com/D3CARP/index.php>.

With this website, we successfully identified the target proteins for a few natural compounds, which were validated by bioassays.

1. Shi, Y., Zhang, X., Yang, Y., et al. (2023) *Computers in Biology and Medicine*, 164, 107283.

COMBATING CYSTIC FIBROSIS: COMPUTATIONAL STUDIES ON CFTR

N. Khazanov, L. Simchaev, M. Zhenin, M.S. Bahia, L. Lublin, A. Xue, H. Senderowitz

Department of Chemistry, Bar Ilan University, Ramat Gan, Israel

Mutations in CFTR, a chloride ion channel of the ABC transporters superfamily, is the cause of the genetic disease Cystic Fibrosis (CF). CFTR is composed of five domains (MSD1-NBD1-R-MSD2-NBD2) all bearing CF-causing mutations. Multiple cryo-EM structures of the full-length protein from different organisms, some in complex with FDA-approved CF therapies, have been published. Yet, atomic-scale structures and dynamics of WT and in particular mutant CFTR remain largely unknown for two main reasons: (1) Experimental structures represent only snapshots of the highly dynamic and allosteric protein. However, the protein's dynamics is critical for understanding the mechanism of action of CF-causing mutations and for developing mutation-specific therapies. (2) Out of the many CF-causing mutations, structures were only solved for the WT construct and for the construct bearing the most prevalent mutation, F508del. Still, these structures provide excellent starting points for constructing reliable models for multiple CFTR constructs and for performing molecular dynamic simulations to study their structures, energetics, and dynamics under near physiological conditions. These structures could also generate testable hypotheses regarding putative binding sites for small compounds modulators. Here we describe a series of such simulations.

We begin by presenting replica exchange MD simulations for a small series of mutated NBD1 constructs. NBD1 is considered a hotspot for CF-causing mutations, many of which compromise the thermal stability of the domain. Our simulations yield a favorable correlation between experimentally determined thermal stability data and fluctuation profiles. Next, we extended our calculations to multiple mutated NBD1 and NBD2 constructs. Using FoldX and Rosetta, we obtained, once more, a favorable correlation between calculated $\Delta\Delta G$ values and thermal stability data and were able to pinpoint the structural determinants which lead to compromised or increased stability. Finally, using lengthy MD simulations at elevated temperatures we were able to shed light into the mechanism of action (MOA) of two CF-causing mutations, L467P and A559T.

Next, we turned our attention to the full-length protein, performing MD simulations, computational electrophysiology simulations and molecular docking on several constructs of the apo protein (WT and mutants) and on the protein in complex with FDA-approved CFTR modulators bound to putative binding sites. By analyzing the computational results, we could shed light on the MOA of several CF-causing mutations (e.g., Q359K/T360K, P67L, I1234V), suggest alternative binding sites for CFTR modulators, and follow the pathway of chloride ions through the CFTR channel. Many of our results were found to be in agreement with experimental data.

Taken together this work demonstrates that computational studies may be able to predict the effect and MOA of CF-causing mutations, provide mechanistic insights into the effect of modulators, and therefore potentially pave the way towards the development of mutations-specific therapies.

1. Sabusap, C.M. et al (2021) *Journal of Biological Chemistry*, 296, 100598.
2. Bahia, S., et al. (2021) *Journal of Chemical Information and Modeling*, 61, 1762-1777
3. Aluma, B., et al. (2020) *Respiratory Medicine*, 170, 106027.
4. Mei-Zahav, M., et al. (2018) *J. Cyst. Fibros.*, 17, e41-e45.
5. Wang, C, et al. (2018) *J. Biol Chem.*, 293, 17685-17704.
6. Simhaev, L., et al. (2017) *J. Chem. Inf. Model.*, 2017, 57, 1932-1946.

STRUCTURAL PHARMACOLOGY OF TRANSIENT RECEPTOR POTENTIAL CHANNELS

A. Neuberger

Columbia University, New York, USA

Transient Receptor Potential (TRP) channels are ion-conducting membrane proteins that are subdivided into 7 subfamilies. The various members can roughly be categorized as either ion homeostasis master regulators/gate keepers or sensors of environmental stimuli of chemical (drugs, natural compounds, toxins etc.) and/or physical nature (e.g., temperature, mechanical stress). Given their central role in human sensing and ion regulation, it is not surprising that TRP channels are involved in various human diseases including many forms of human cancer. Moreover, as pain sensors, TRP channels also represent some of the most sought-after membrane proteins for the development of new, addiction-free pain therapy approaches. We utilize cryo-electron microscopy (cryo-EM) to resolve high-resolution, near-atomic structures of various important TRP channels in different conformations, fore most open and closed states, in the presence of different agonists or antagonists to inform novel drug design approaches. In particular, we show 3 study examples that are representative for three approaches of how we use cryo-EM as a tool to foster new drug development approaches: (1) “inspired by nature” – synthesis of a novel small-molecule drug inspired by the natural regulation of a human calcium master regulator and oncochannel, (2) “old drugs, new target(s)” – exploration and exploitation of how already approved drugs can work on new (TRP channel) targets, therewith fostering the urgently needed drug repurposing, and (3) natural compounds – resolving the molecular mechanism of natural (ant-)agonists’ action on TRP channels. Our structural pharmacology of how we can use the natural inactivation mechanism of human calcium master regulator and oncochannel TRPV6 by intracellular Calmodulin for the design of a novel hTRPV6 inhibitor that also features the highest known affinity to date (1) as well as of the molecular mechanism of hTRPV6’s inhibition by the FDA-approved drug econazole (2) and by natural compounds such as genistein (soy) and tetrahydrocannabivarin (a phytocannabinoid) (3) demonstrate the power of cryo-EM, especially of combined with complimentary computational methods such as advanced molecular dynamic simulation and functional recordings.

FINE MAPPING OF BIOMOLECULAR SURFACES USING THE NEW MOLECULAR SURFACE TOPOGRAPHY (MST) WEB TOOL

Yu.A. Trofimov^{1,2}, N.A. Krylov¹, I.I. Veretenenko^{1,3}, A.O. Chugunov^{1,3}, R.G. Efremov^{1,3,4}

¹*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, Russia*

²*National Research Nuclear University MEPhI, Moscow, Russia*

³*Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Russia*

⁴*National Research University Higher School of Economics, Moscow, Russia*

Structural biology provides an ever-growing number of atomic resolution structural data, giving valuable information about isolated intermediates of biomolecules and their complexes, as well as their conformational ensembles. At the same time, *in silico* methods like molecular dynamics (MD) simulations are often used to explore the small-scale flexibility of molecules, their properties and the energetics of inter- / intra-molecular interactions. A fruitful approach in structural studies is to inspect the characteristics of molecular surfaces. The physico-chemical properties of surfaces (e.g., landscape, hydrophobicity and charge distributions), on the one hand, are determined by the molecular structure itself, and, on the other hand, they substantially determine the molecular function. Such properties distributed on a complex 3D surface are often reduced to 2D maps for ease of visual and computational analyses. Compared to the 3D representation, this approach provides a way to catch the fine details of the molecular surface on flat maps, compare global and local variations in surface features in different conformations of molecules, and evaluate the complementarity of interactions in molecular complexes.

Here we present a new Molecular Surface Topography (MST) web tool, which was developed as an evolution and generalization of the Protein Surface Topography (PST) method [1]. MST provides the capability to build 2D projections of molecular surfaces and to represent as a map a number of surface features, including surface landscape, molecular hydrophobicity potential (MHP) and electrostatic potential (ESP). Three types of projection are available for visualization: planar, cylindrical and spherical. MST is freely available at <https://model.nmr.ru/cell>.

Three application cases of the tool with an appropriate projection geometry are presented. (i) Mapping the membrane surface using a planar projection. The relationship between the clustering of lipids in mixed DOPC/DOPS bilayers and the mosaicity of MHP distribution on the membrane-water interface is shown based on MD simulation data [2]. (ii) Cylindrical maps provide a detailed representation of the hydrophobic organization of the conductive pores of TRPV ion channels based on the cryo-EM structures. It has been shown that the conformation-dependent hydrophobicity of the pore determines the conductive or non-conductive states of the channels [3]. (iii) The spherical projections of a small peptide, conotoxin, demonstrate the complementarity of MHP and ESP distributions between the molecules in the toxin-receptor complex obtained using X-ray crystallography. The maps suggest that the charge distribution on the toxin surface provides the required orientation for its binding to the protein site [1]. Such “molecular portraits” obtained with the MST tool help to identify patterns of physico-chemical properties of the surface, which can be useful in the functional characterization of small compounds, macromolecules and their complexes, and can also provide a new insight in the future engineering of molecules with predefined properties.

The work was supported by RSF grant 23-14-00313.

1. Koromyslova, A.D., et al. (2014) *J. Chem. Inf. Model.*, 54, 1189.
2. Veretenenko, I.I., et al. Nanoscale lipid domains determine the dynamic molecular portraits of mixed DOPC/DOPS bilayers in a fluid phase: a computational insight. In press.
3. Trofimov Y.A. et al. (2024) *Commun. Chem.* 7, 119 (2024).

ICHRONOBIOTICSDB - WORLD FIRST DATABASE OF CIRCADIAN RHYTHM'S PHARMACOLOGICAL MODULATORS

I. Solovev, D. Golubev, N. Kotelina, A. Yagovkina

*Pitirim Sorokin Syktyvkar State University, Medical Institute,
Laboratory of Translational Bioinformatics and Systems Biology, Syktyvkar, Russia*

Chronobiotics are drugs [1], both experimental and used in medical practice, constituting a rather heterogeneous, from a chemical point of view, group of substances that can modify the parameters of the circadian rhythm of fluctuations in various physiological and biochemical parameters, such as the expression of the “clock” genes themselves in organisms-models and cell cultures or the expression of clock-controlled genes. The class of chronobiotic drugs has been known for more than 50 years, since the properties of the hormone melatonin were discovered and described in detail in the clinic. Despite attempts to systematize chronobiotics, there is not yet a unified classification of these pharmacological agents (natural chrononutrients, synthetic targeted circadian rhythm modulators, hypnotics, chronobiotic hormones are identified), and there is no single source of knowledge about chronobiotics. Creating the world's first curated and updated database of chronobiotic drugs (circadian rhythm modulators) and organizing access to it through the Internet information and telecommunications network is an extremely urgent fundamental task of chronobiology, chronomedicine and pharmacoinformatics/bioinformatics.

The purpose of the study is to create a relational database of chronobiotics “ChronobioticsDB” (cb-db.ru) using the php programming language [2] and MySQL as a database management system. Access to the database will be provided through the website. This database is filled using PubMed data on chronobiotics, which are manually extracted from articles and annotated. Each chronobiotic compound card contains links to primary data sources, an image of the molecule, the chronobiotic formula in machine-readable SMILES format, and a name according to IUPAC nomenclature. To enrich the card for each compound, it is planned to synchronize the database with ChemSpider, DrugBank, Chemb, ChEBI, Engage, UniProt, PubChem in order to increase the relevance of information about each chronobiotic. The biological and pharmacological value of the database is given by synchronization with the databases of the FDA, CLINICALTRIALS.GOV, Selleckchem, KEGG, MSDS, TOXNET, etc (in case of overlap, the cards will contain information about the unique properties of the chronobiotic). The database will be initially utilised as a training-set for AI searching for novel chronobiotics.

This work was funded by the Russian Science Foundation Grant “Design of the world's first pharmacological database of circadian rhythm modulators (ChronobioticsDB) and organisation of the access to it” no. 24-75-00108

1. Solovev, I.A., Shaposhnikov, M.V., Moskalev, A.A. (2021) *Clocks & Sleep*, 3, 429-441.
2. Achour, M., Betz, F., Dovgal, A., et al. PHP Manual. 2023.

COCONUT 2.0 DATABASE AND AUTOMATED LITERATURE MINING USING DECIMER.AI

K. Rajan¹, V. Chandrasekhar¹, S. Kanakam¹, N. Sharma¹, J. Schaub¹, A. Zielesny², C. Steinbeck¹

¹Institute for Inorganic and Analytical Chemistry, Friedrich Schiller University Jena, Jena, Germany

²Institute for Bioinformatics and Chemoinformatics, Westphalian University of Applied Sciences,
Recklinghausen, Germany

Natural products (NPs) represent a class of small molecules biosynthesised by living organisms that exhibit significant pharmacological and industrial potential due to their bioactive properties. Despite the existence of several databases, there is no comprehensive digital repository that consolidates open information about natural product structures. The COllecTion of Open Natural prodUcTs (COCONUT) [1] stands as one of the most extensive online databases, aggregating both elucidated and predicted NPs from open-access sources. The upcoming COCONUT V2.0 COCONUT V2.0 will incorporate data from approximately 63 databases and datasets, integrating structural information, computed molecular properties, and, where available, annotated data on nomenclature, source organisms, and bibliographic references. The ongoing standardisation and curation process uses the newly developed cheminformatics microservices [2]. A modern web interface facilitates database queries based on structure, substructure, similarity, and calculated properties. Additionally, COCONUT V2.0 will feature a submission platform with reporting capabilities. Still, the printed literature contains a wealth of information, and there is a continually expanding corpus of publications describing novel chemical structures. However, this information is not machine-readable and is primarily intended for human readers. This emphasises the necessity for automated methods to extract and convert such data into machine-readable formats for inclusion in public databases. However, the majority of existing tools for literature mining and optical chemical structure recognition (OCSR) are proprietary or subject to restrictive licensing agreements [3]. DECIMER.ai [4] is an open-source platform that addresses this challenge by leveraging recent advances in deep learning, natural language processing, and computer vision. This platform demonstrates the capability to automatically segment, classify, and translate chemical structure images into computer-readable formats. The web application facilitates end-to-end extraction and curation of chemical data from scientific literature. A process is underway to fine-tune Large Language Models (LLMs) to extract organism-specific information and relevant metadata from the printed literature. All models, codes, and datasets are publicly released to reduce manual effort in systematically mining the chemical knowledge from publications, with extracted data from natural products slated for inclusion in COCONUT. With the integration of DECIMER's automated literature mining capabilities, COCONUT 2.0 will become an even more comprehensive, curated, and open-source repository of natural product research.

1. Sorokina, M., Merseburger, P., Rajan, K., Yirik, M.A., Steinbeck, C. (2021) *J. Cheminform.*, 13, 2
2. Chandrasekhar, V., Sharma, N., Schaub, J., Steinbeck, C., Rajan, K. (2023) *J. Cheminform.*, 15, 98
3. Rajan, K., Brinkhaus, H.O., Zielesny, A., Steinbeck, C. (2020) *J. Cheminform.*, 12, 60
4. Rajan, K., Brinkhaus, H.O., Agea, M.I., Zielesny, A., Steinbeck, C. (2023) *Nat. Commun.*, 14, 5045

ON A SIMPLE FRAMEWORK OF DIMENSIONALITY REDUCTION FOR CLASSIFICATION MODELING OF SPARSE ENVIRONMENTAL TOXICITY DATA

A. Banerjee, K. Roy

Drug Theoretics and Cheminformatics Laboratory, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India

Experimental toxicity data of chemicals against several environmental endpoints are often limited. One of the most commonly used *in silico* approaches for toxicity prediction is Quantitative Structure-Activity Relationship (QSAR), which generates predictions for the query compounds. However, the reliability of the predictions from QSARs derived from small datasets is often questionable from a statistical point of view. This is due to the presence of a larger number of descriptors as compared to the number of training compounds, which reduces the degree of freedom of the developed model. To reduce the overall prediction error for a particular QSAR model, we have proposed here the computation of the novel Arithmetic Residuals in K-groups Analysis (ARKA) descriptors. We have reduced the number of modeling descriptors in a supervised manner by partitioning them into K classes ($K = 2$ here) depending on the higher mean normalized values of the descriptors to a particular response class, thus preventing the loss of chemical information. A scatter plot of the data points using the values of two ARKA descriptors (ARKA_1 vs. ARKA_2) can potentially identify activity cliffs, less confident data points, and less modelable data points. We have used here five representative environmentally relevant endpoints (skin sensitization, earthworm toxicity, milk/plasma partitioning, algal toxicity, and rodent carcinogenicity of hazardous chemicals) with graded responses to which the ARKA framework was applied for classification modeling. On comparing the performance of the models generated using conventional QSAR descriptors and the ARKA descriptors, the prediction quality of the models derived from ARKA descriptors was found, based on multiple graded-data validation metrics-derived decision criteria, much better than the models derived from QSAR descriptors signifying the potential of ARKA descriptors in ecotoxicological classification modeling of small data sets. Additionally, this holds true for the Read-Across approach as well, since the Read-Across predictions using ARKA descriptors supersede the predictions generated from QSAR descriptors. For the ease of users, a Java-based expert system (<https://sites.google.com/jadavpuruniversity.in/dtc-lab-software/arithmetic-residuals-in-k-groups-analysis-arka>) has been developed that computes the ARKA descriptors from the input of QSAR descriptors.

ACTIVATION OF GLYCOGENOLYSIS WITHOUT AN ACTIVATION OF ATP CONSUMPTION CAN CAUSE CELL DEATH

M.V. Martinov¹, S.I. Sudarkina², F.I. Ataullakhanov^{1,3}, V.M. Vitvitsky¹

¹*Center for Theoretical Problems of Physicochemical Pharmacology, Russian Academy of Sciences, Moscow, Russia*

²*National Research University Highest School of Economics, Moscow, Russia*

³*Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia*

Glycolysis and the associated glycogen metabolism are important components of the energy metabolism in animal cells. Understanding the mechanisms of interaction of these metabolic systems is of great importance both for understanding the general principles of regulation of cellular metabolism, and for solving a big number of actual applied problems in the field of medicine, pharmacology, and biotechnology, such as prevention and treatment of diabetes, treatment of glycogen storage diseases, development of new drugs, optimization of antitumor therapy, etc. The basic principles of regulation of these metabolic systems were formulated many years ago. At the same time, the regulation of glycolysis and glycogen metabolism in most specific cells and tissues has not been studied well enough. For example, in contracting skeletal muscles an activation of glycogen phosphorylase (an enzyme that catalyzes the conversion of glycogen to glucose-1-phosphate) leads to activation of glycolytic flux and ATP production due to activation of glycogen consumption. However, in resting muscles, activation of glycogen phosphorylase does not lead to either activation of glycolysis or glycogen consumption. The mechanisms of this regulation remain unclear. To understand the mechanisms of this regulation we analyzed the interaction between glycogenolysis and glycolysis using a mathematical model of muscle energy metabolism in mammalian white skeletal muscles. System analysis of the model demonstrates that under steady-state conditions activation of glycogen phosphorylase provides a significant increase in the rate of ATP production and enhances energy charge stabilization in contracting muscles. The model reveals a significant role of phosphate (a substrate of glycogen phosphorylase) in regulation of glycogenolysis and in stabilization of [ATP] and energy charge in muscles. In addition, phosphocreatine has been shown to serve not only as a source for ATP production, but also as a phosphate depot, which is released during ATP depletion and accelerates glycogen phosphorylase reaction. It was found, however, that activation of glycogen phosphorylase in a resting muscle cannot provide stable glycogen consumption due to depletion of phosphate. Moreover, the system analysis reveals that the forced acceleration of glycogen phosphorylase reaction in a resting muscle does not increase the rate of ATP production but stimulates accumulation of phosphorylated glycolytic intermediates instead which can cause osmotic damage to the muscle cells.

This study was supported by the Russian Science Foundation, grant no. 23-24-00178 (for M.V.M. and V.M.V.).

AUTOMATING THE DESIGN OF GLYCOMIMETIC AGENTS

R.J. Woods, Y. Xiao

Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA

Introduction. Specific interactions between carbohydrates and proteins underlie the initiation or progression of many diseases. Carbohydrate-binding proteins (human, bacterial or viral lectins and adhesins) and carbohydrate-processing enzymes (glycosyltransferases and glycosidases) are therefore important targets for therapeutic intervention, however the creation of drug-like molecules that can competitively inhibit carbohydrate-binding sites is uniquely challenging. The optimization of a glycomimetic inhibitor involves the synthesis and screening of chemical analogs in an attempt to increase the inhibitory potential and biological activity. Despite the synthetic challenges, the benefit of employing the native carbohydrate as a scaffold is that it intrinsically confers the desired specificity. The fundamental challenge in the creation of a glycomimetic is that of divining which modifications will lead to enhanced affinity without compromising specificity.

Methodology. Here we present our work towards the development and validation of a computational strategy that leverages the benefits of computational modeling and structural biology. Specifically, we are creating and automated computational approach that uses carbohydrate-protein co-crystal (or NMR) structures as the basis for lead optimization by modifying the bound oligosaccharide *in situ*.

Results & Discussion. We have successfully constructed a library of 545 drug-like aliphatic and aromatic moieties, extracted from Sigma-Aldrich, ensuring that they are available for synthetic strategies. Each moiety may be conjugated at any available hydroxyl or amino position in a carbohydrate, which for a single monosaccharide with five available positions for derivatization gives rise to 545^5 (4.8×10^{13}) potential glycomimetic compounds. Here we will discuss the automated screening of this moiety library with regard to several protein receptors, including influenza hemagglutinin. In the course of this work, we have noted weaknesses in current computational energy functions, including the lack of CH- π interactions, which we have recently corrected[1] in the popular docking software AutoDock VINA-Carb [2].

Conclusions. The successful completion of this project will facilitate the application of sophisticated modeling techniques by users who are either not experts in modeling, or not experts in carbohydrate chemistry.

We thank the National Institutes of Health for funding (R01 GM-135473)

1. Xiao, Y., Woods, R.J. (2023) *J. Chem. Theory Comput.*, 19, 5503-5515.
2. Nivedha, A.K., Thieker, D. F., Hu, H., Woods, R.J. (2015) *J. Chem. Theory Comput.*, 12, 892-901.

SQUARE ANTIPRISM IS A KEY DETERMINANT FOR POTASSIUM ION SELECTIVITY

K. Scherbakov¹, A. Vassilevskii², A. Chugunov²

¹ Institute of Biomedical Chemistry, Moscow, Russia

² Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia

The mechanism of ion channel selectivity is a key and yet unresolved question in biochemistry. It is well established that potassium channel selectivity is provided by the selectivity filter (SF) structure that replicates the geometry of the hydrated K⁺ ion, which is surrounded by eight water molecules forming a square antiprism. The SF is constituted of five layers of carbonyl/threonine side chain oxygen atom quads, which form four K⁺-binding sites (S1–S4), each built up of eight oxygen atoms, perfectly arranged in a square antiprism. Such architecture allows dehydration with no energy penalty for K⁺, but not Na⁺, Ca²⁺, or most other cations. Whether this principle is universal for all known K⁺-channels and if the K⁺-recognising sites are unique for their SF is unknown. To answer these questions, we set up a nearly protein universe-wide geometric scanning for square antiprismatic sites, using not only all known K⁺-channel structures, but also all other membrane proteins, and a representative subset of all other proteins. We found that:

1. the SF is a unique structure persisting in K⁺-channels and some related proteins in a highly conserved form, but it is absent from other proteins;
2. conductive and non-conductive SF states may be clearly delineated by a 1.25 Å RMSD threshold at sites S1–S3, providing a precise and easy criterion for geometrical assessment of the SF state;
3. single K⁺-binding sites are found not only in SFs, but also in other channel domains;
4. among other membrane proteins, there are prominent cases of K⁺-permeable proteins including K⁺-transporters and NaK, CNG and HCN channels, which also feature K⁺-binding sites;
5. in Na⁺,K⁺-ATPase the “quality” of the found K⁺-binding sites correlates with the protein state: in a K⁺-selective E2 state RMSD from a solvated K⁺ template is lower than in Na⁺-selective E1 state;
6. in other membrane as well as non-membrane proteins the antiprismatic sites often coincide with the binding sites for other cations.

Our work provides a way for determination of K⁺ binding sites in proteins as well as explanation of high K⁺ selectivity of the potassium channels.

COMPREHENSIVE COMPUTATIONAL SYSTEMS BIOLOGY MODEL OF BLOOD PLATELET SIGNALLING: A TOOL FOR BASIC RESEARCH, DIAGNOSTICS AND PHARMACOLOGY

A.N. Sveshnikova^{1,2}, M.A. Panteleev^{1,2}

¹*Center for Theoretical Problems of Physico-Chemical Pharmacology, Russian Academy of Sciences, Moscow, Russia*

²*National Medical Research Center of Pediatric Hematology, Oncology and Immunology named after Dmitry Rogachev, Moscow, Russia*

Anuclear blood cells, platelets, are the basis for the formation of blood clots in human vessels. Dysfunction of platelets is observed in various diseases, both hereditary and acquired. Anti-platelet therapy is mostly used after ischemic stroke; however, action of therapy could be patient-specific. Previously we developed a series of computational models for the description of platelet intracellular signaling [1, 2] and a system of experimental methods that allows us to characterize patient's platelets [3]. In order to construct a personalized model of platelet signaling, we also performed platelet proteomics studies [4]. Next, we developed an algorithm for the personalization of platelet computational model. First, the protein copy numbers for the signal transduction enzymes were taken from the patient's proteomics data. Second, the model was divided into functional modules. The central module described the platelet calcium signaling, as calcium is the main driver of platelet activation. Personal cytosolic calcium concentration dynamics in response to various agonists was measured by flow cytometry [1], and characteristic distributions of calcium oscillations were taken from single-platelet microscopy [2]. Together, these data were used for the adjustment of model parameters and module validation. The other modules were platelet receptor activation (for G-protein coupled receptors and tyrosine kinase receptors), platelet integrin activation and platelet mitochondria-dependent necrosis. The unknown parameters of these modules were determined based on flow cytometry data for the investigated patient by means of parameter estimation techniques. The validation of the whole model was conducted against patient functional testing, including aggregometry and end-point flow cytometry data.

Thus, we have developed an experimental-theoretical approach to computer-aided study of patients' platelets and *in silico* personalized prediction of the patient responses to therapy. These data will help in the future to select personalized therapy for these patients.

This study was supported by Russian Science Foundation grant 23-74-00057.

1. Sveshnikova, A.N., Balatskiy, A.V., Demianova, A.S., et al. (2016) *Journal of Thrombosis and Haemostasis*, 14, 2045–2057
2. Martyanov, A.A., Balabin, F.A., Dunster, J.L., et al. (2020) *Biophysical Journal*, S0006349520303465
3. Martyanov, A. A., Tesakov, I. P., Khachatryan, L. A. et al. (2023) *Blood Advances* 7, 4936–4949.
4. Garzon Dasgupta, A.K., Martyanov, A.A. , Ignatova, A.A., et al. (2024) *Pediatric Research*, 95, 966–973

ANALYSIS OF CHEMICALS-VIRUS-HOST INTERACTIONS BASED ON LARGE-SCALE BIOMEDICAL TEXT AND DATA MINING

O.A. Tarasova¹, N.Yu. Biziukova¹, R.R. Taktashov¹, D.A. Karasev¹, B.N. Sobolev¹, N.S. Ionov¹, S.M. Ivanov^{1,2}, A.V. Dmitriev¹, A.V. Rudik¹, D.A. Filimonov¹, V.V. Poroikov¹

¹Institute of Biomedical Chemistry, Moscow, Russia

²Pirogov Russian National Research Medical University, Moscow, Russia

The development of safe and effective antivirals is of great importance due to the new challenges associated with the unexpected rapid spread of viral infections, such as SARS-CoV-2, leading to a global pandemic. For the treatment of viral infections caused by hepatitis C, herpesviruses, human immunodeficiency virus (HIV), predicting the efficacy and safety of therapeutic strategies is important due to the problems of latency and drug resistance. The development of novel drugs is rather time-consuming (taking up to dozen years) and costly process [1]. At the same time, drug development and health-related biomedical studies generate a huge amount of information that is published and deposited in databases. This information can be used to gain new knowledge about structure-property relationships by applying AI/ML methods.

Our study aims to develop an in silico approach for the extracting knowledge about viruses and the host (the human body), and potential antiviral agents based on the mining of massive amounts of scientific publications.

Based on the results obtained earlier, we developed an integrated method for the automated extraction of comprehensive information about the interactions between virus, host and chemical compounds. Using this information, it is possible to acquire new knowledge about antiviral agents acting on the both viral and host targets, identify both active and inactive compounds, and to assess their possible side effects and toxicities. Moreover, our approach provides the information on the co-morbidities associated with the particular viral infection.

The developed approach is based on the automated selection of relevant publications using deep neural network models followed by the recognition of named entities and the extraction of the associations between them. The conditional random fields [2] and naïve Bayesian [3] approaches were used for automated recognition of chemical and biological named entities. The associations between chemical and biological objects (viral and host proteins, genes, miRNAs) are extracted using the set of template phrases and the naïve Bayesian approach. Belonging of a particular biological object to the specific organism is identified using the automated queries to biological databases.

The developed approach was used to analyze data for a number of viruses, including hepatitis B and C viruses, SARS-CoV-2, influenza A and B, and herpes simplex virus. We collected data on the interactions of these viruses with (1) the host (human body), (2) potential antiviral agents and also extracted information on interactions between potential antiviral agents and host proteins is also extracted. Based on the data analysis performed, we have created a freely available knowledgebase on the interaction of chemical compounds with viral proteins and their host targets. The developed approach allows the identification of both well-studied and recently discovered novel interactions between a virus, a host and a chemical compound. Information on virus-host interactions is useful for a comprehensive understanding of viral infection mechanisms and the development of new therapeutic strategies.

This study was supported by the Program for Basic Research in the Russian Federation for a long-term period 2021–2030 (project No. 124050800018-9).

1. Catacutan, D.B., Alexander, J., Arnold, A., Stokes, J. (2024) *Nature Chemical Biology*, online ahead of print. doi: 10.1038/s41589-024-01679-1
2. Biziukova, N.Y., Ivanov, S.M., Tarasova, O.A. (2024) *Big Data Mining and Analytics*, 7, 107.
3. Tarasova, O.A., Rudik, A.V., Biziukova, N.Y. et al. (2022) *Journal of Cheminformatics*, 14, 55.

CHEMICAL PROTEOMICS FOR OVERCOMING DRUG RESISTANCE

H. Lyu¹, A. Kel,² R.A. Zubarev^{1,3,4}

¹Karolinska Institutet, Stockholm, Sweden

²geneXplain GmbH, Wolfenbüttel, Germany

³RUDN University, Moscow, Russia

⁴The National Medical Research Center for Endocrinology, Moscow, Russia

FDA approval currently has a staggering 97% failure rate at clinical trials for oncology; typically, due to issues with drug efficacy, toxicity or resistance to the drug. A promising approach to overcoming drug resistance is finding a suitable complementary molecule (agent). The agent by itself doesn't have to be toxic to cancer cells, and may be neutral or even promote cancer cell proliferation, making the main therapy more effective. A classic example is the combination of Tegafur [1-(2-tetrahydrofuryl)-5-fluorouracil] and uracil in a 1:4 ratio, with Tegafur being the main anticancer drug, and uracil - the "innocent" agent [1]. Addition of this agent significantly potentiates the toxicity of Tegafur to cancer cells. The difficulty of scaling up this approach to a wide variety of anticancer agents is in finding the proper agent. As the agent action must be specific to the action mechanism of the main anticancer drug, finding the proper agent is a nontrivial task that has so far required either high-throughput screening or pure luck. Recently, a third alternative emerged because of the progress in chemical proteomics and the development of effective data analysis tools. This alternative is the intelligent analysis of comparative proteomics data on resistant and sensitive cells for the agent based on the knowledge (or predictive modeling) of the cellular mechanics. This approach is still under development, but first results are encouraging.

First, we grew lung cancer cells resistant to auranofin (at least a 2-fold increase in LC50) from the original population of lung cancer cells sensitive to this drug. Then we performed comprehensive comparative proteomics analysis of the resistant cells versus sensitive cells using novel approach PISA-REX proteomics [2] that combines solubility, expression and redox analysis in a single experiment. The PISA-REX data were then analyzed with the latest geneXplan data processing tool Genome Enhancer [3]. Genome Enhancer uses upstream search, an integrated promoter and pathway analysis, and identifies potential drug targets responsible for the observed proteome changes. In the first step of analysis, the transcription factors that regulate differential proteins are identified with the use of the TRANSFAC® database of transcription factors binding sites [4]. On the second step, the tool searches for common master-regulators of the identified transcription factors by building a data-specific signal transduction network using the TRANSPATH® database of mammalian signal transduction and metabolic pathways [5]. The identified master regulators are prospective drug target candidates. The HumanPSD™ database is employed to identify targeting these candidates drugs that have been tested in clinical trials. In parallel, the cheminformatic tool PASS [6] predicts small molecules that can affect the identified targets.

In our case, Genome Enhancer correctly identified the thioredoxin pathway as the main difference between the sensitive and resistant cells with thioredoxin as the main target, but other pathways were also listed. Based on the set of revealed potential targets, Genome Enhancer predicted the molecules that can affect them. We are currently testing how these molecules affect LC50 of auranofin.

1. Maehara, Y., Kakeji, Y., Ohno, S., Baba, H., Sugimachi, K. (1997) *Oncology*, 11, 14.
2. Saei, A. A., Lundin, A., Lyu, H., ..., Zubarev, R. A. (2024) *Adv. Sci.*, in press.
3. <https://genexplain.com/genome-enhancer/>
4. <https://genexplain.com/transfac/>
5. <https://genexplain.com/transpath/>
6. <https://genexplain.com/pass/>

HOW MANY DRUG TARGETS DO WE NEED?

A. Kel^{1,2}, R. Zubarev²

¹geneXplain GmbH, Wolfenbüttel, Germany

²Karolinska Institutet, Stockholm, Sweden

In oncology, the question of how many drug targets we need is pivotal in determining the future direction of cancer treatment. With the increasing complexity of cancer biology, it is unclear whether focusing on a few highly effective targets or identifying a broader range of targets will yield better therapeutic outcomes. This talk will explore the balance between quantity and quality of drug targets in the quest to optimize cancer treatment strategies. To answer this question, we analyzed one of the most comprehensive datasets on cancer drug treatment proteomics that includes 287 A549 adenocarcinoma proteomes affected by 56 cytotoxic drugs [1].

Deconvolution of this dataset into a space of protein expression dimensions revealed a set of 13 “orthogonal cell death pathways” that characterize different mechanisms of cell death. To understand the details of the molecular mechanisms of regulation of protein expression in each of these cell death pathways we applied the AI driven technology Genome Enhancer [2]. Genome Enhancer analyses multi-omics data and reconstructs molecular network governing the gene and protein expression in a given pathological state of the tumor cells. It identifies master-regulators in the network and predicts the optimal drug targets. Genome Enhancer applies *Upstream Analysis* algorithm [2, 3] for the identification of key master-regulators responsible for pathologic gene regulation in cancer cells of the studied states. The analysis comprised of three main steps: 1) AI-based algorithm to scan promoters of the genes that encode differentially expressed proteins using TRANSFAC® database [4] and to identify complexes of transcription factors (TF), which regulate these genes; 2) a graph search algorithm using TRANSPATH® database [5] to identify common regulators of the TFs selected on the previous step; 3) drug targets are then selected on the basis of HumanPSD database, containing the information about the approved drugs and their targets, as well as the therapies under development, and their targets. It also reveals new potential drug targets using chemoinformatic approach of PASS tool [6].

The performed analysis revealed a set of networks regulating activity of all 13 “orthogonal cell death pathways” that we think cover the majority of possible cancer cell death destines. These networks are characterized by fairly different although partially overlapping sets of master regulators – that can be considered as a comprehensive set of anti-cancer drug targets. The multi-target-oriented approach towards treatment prescription accelerates the off-label drug usage and enables effective treatment selection for clinically complicated cases with no obvious treatment options.

1. Saei AA, Beusch CM, Chernobrovkin A, et al., (2019) *Nat Commun.*, 10(1), 5715.
2. Kel, A., Boyarskikh, U., Stegmaier, P. et al. (2019) *BMC Bioinformatics*, 20, 119.
3. <https://genexplain.com/genome-enhancer/>
4. <https://genexplain.com/transfac/>
5. <https://genexplain.com/transpath/>
6. <https://genexplain.com/pass/>

ALPHAFOLD-LIKE MODELS FOR SMALL MOLECULE STRUCTURE PREDICTION

G. Derevyanko¹, P. Popov²

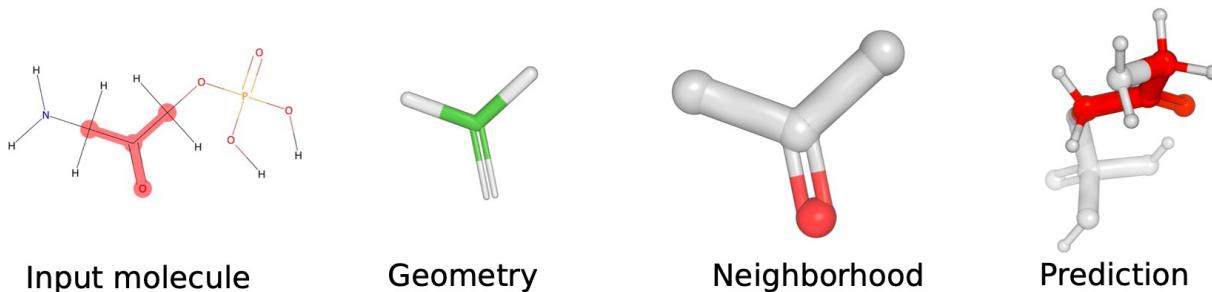
¹*Skolkovo Institute of Science and Technology, Moscow, Russia*

²*Constructor University, Bremen, Germany*

Recent success of AlphaFold2 [1] in predicting structures of proteins from multiple sequence alignments (MSA) raises the question: can we generalize this approach to other important types of molecules? The positive answer to this question opens a door to overcoming the lack of structural data needed to train the model for predicting structures of RNA, proteins with non-standard amino-acids, and small molecules. One of the key components of AlphaFold2 is a protein representation, where each amino acid residue is described as a rigid body, while the inner degrees of freedom are predicted separately. This coarse-graining method allowed extracting the orientation of amino acid residues from MSA, which plays a big role during the co-evolution of the neighboring residues. In this work, we propose a novel representation of a small molecule as a set of rigid bodies with additional constraints, which is suitable for small molecular structure modeling. We further adjusted the Evoformer block architecture from AlphaFold2, such that it incorporates explicit positions of rigid bodies eliminating the need for a separate structural module. We have tested our method on the task of predicting ground-state molecular structures using the curated dataset Molecule3d extracted from PubChemQC of about 4 millions molecules. We show that the model outperforms RDKit and previously published baselines. Particularly, our model achieves the average root mean square deviation (RMSD) of 0.83 over

~700,000 molecules in the testing subset of the Molecule3d dataset. The mean evarage error of the predicted distance matrix between the heavy atoms of the molecules is 0.30 versus 0.53 for the RDKit ETKDG algorithm [2] and 0.66 for the DeeperGCN-DAGNN + Distance [3].

This study was supported by Russian Scientific Foundation project No. 22-74-10098.



1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021) *Nature*, 596(7873), 583–589.
2. Riniker, S., Landrum, G.A. (2015) *Journal of chemical information and modeling*, 55(12), 2562– 2574.
3. Xu, Z., Luo, Y., Zhang, X., Xu, X., Xie, Y., Liu, M., Dickerson, K., Deng, C., Nakata, M., Ji, S. (2021) *arXiv preprint arXiv*, 2110.01717.

PROTEIN 3D STRUCTURE IDENTIFICATION BY ALPHAFOLD: A PHYSICS-BASED PREDICTION OR RECOGNITION USING HUGE DATABASES?

A.V. Finkelstein,^{1,2} D.N. Ivankov³

¹*Laboratory of Protein Physics, Institute of Protein Research of the Russian Academy of Sciences,
Pushchino, Moscow Region, Russia*

²*Biology Department, the Lomonosov Moscow State University, Moscow, Russia*

³*Center of Molecular and Cellular Biology, Skolkovo Institute of Science and Technology, Moscow, Russia*

The great success of AlphaFold programs poses the questions: (i) What is the main reason for this success? (ii) What AlphaFolds does: physics-based prediction of the spatial structure of a protein from its amino acid sequence or recognition of this structure from similarity of the target sequence to some parts of sequences with already known spatial structures? The answers given here are: (i) the main reason for the AlphaFold's success is the usage of huge databases which already cover virtually all protein superfamilies existing in Nature; (ii) using these databases, multiple sequence alignments, and coevolutionary information – like correlations of amino acid residues of the contacting chain regions – AlphaFold recognizes a spatial structure by similarity of the target sequence (or its parts) to related sequence(s) with already known spatial structures. We emphasize that this does not diminish the merit and utility of AlphaFold but only explains the basis of its success.

ORAL PRESENTATIONS

SUCCESSFUL APPLICATION OF COMPUTING METHODS TO DEVELOPMENT SARS-COV-2 INHIBITORS

A.V. Sulimov^{1,2}, I.S. Ilin², A.S. Tashchilova¹, O.A. Kondakova², D.C. Kutov^{1,2}, V.B. Sulimov^{1,2}

¹Dimonta Ltd., Moscow, Russia

²Research Computing Center, Lomonosov Moscow State University, Moscow, Russia

The COVID-19 pandemic has triggered a wave of publications on the application of computational methods to the search for inhibitors of target proteins of the SARS-CoV-2 coronavirus [1]. Against the backdrop of a huge number of publications, we found it interesting to review those studies and the computer methods used in them that turned out to be successful: the inhibitor candidates found in them were confirmed in experimental testing.

In developing anti-viral drugs against SARS-CoV-2, the structure-based methodology was preferred over ligand-based methods. Between 2020 and early 2024, inhibitors were developed for eight therapeutic targets - SARS-CoV-2 coronavirus proteins for which 3D structures were known: Mpro, PLpro, nsp14, nsp15, 2'-O-methyltransferase (nsp16), viral helicase (nsp13), RNA-dependent RNA polymerase (RdRp) and spike protein. The most widely used computational methods were molecular modeling methods, and docking [2] certainly ranks first in the list of molecular modeling methods. Docking is the main modelling tool of hit discovery at the initial stage of drug development. Docking plays a secondary, although important, role in the hit-to-lead optimization. Docking programs find a best position of a ligand in the active site of the target protein and estimate the free energy of the protein-ligand binding, the so-called docking score. The literature reviewed shows that it is necessary to improve the accuracy of docking programs to improve the efficiency of early-stage drug development. Supercomputer resources have become increasingly used for docking. To increase the efficiency of docking, it should be supplemented by post-docking processing of the top scored compounds. For this purpose, quantum-chemical calculations of the enthalpy of protein-ligand binding by semiempirical methods, MM/GBSA or MM/PBSA methods and molecular dynamics study of stability of the protein-ligand complexes along long trajectories are used. The larger the size of the screened database, the more important it is to use post-processing to reduce the number of top-scored molecules to a level that can be processed experimentally. Pharmacophore models reflecting key protein-ligand interactions also played an important role. When selecting the best compounds for experimental testing, programs for clustering chemically similar compounds are used to increase the diversity of candidates, and an analysis of the interactions of candidates with the target protein is carried out, including the identification of hydrogen bonds. The database content used for virtual screening greatly influences the success of inhibitor discovery. Screening of very large databases has not yet led to high hit detection efficiency. For screening very large ligand databases containing up to tens of billions of molecules, neural network methods, machine learning algorithms, in general, artificial intelligence (AI) methods, in combination with docking are used. AI helps to select compounds with supposedly good docking scores and avoid docking of ligands with obviously bad scores. Screening of databases of available ligands of ready-made and/or synthesizable compounds is very widely used for the initial discovery of hits. This approach significantly reduces material and time costs at the initial stage of searching for hits.

The widespread use of computer modelling methods to develop inhibitors targeting SARS-CoV-2 proteins has led to a significant, but not yet sufficient, acceleration in the development of new effective drugs. Success of drug development is determined not only by high accuracy of computational methods but also by the availability of methods for experimental verification of modelling results.

This research was financially supported by the Russian Science Foundation, Agreement no. 21–71–20031.

1. Gurung, A.B., Ali, M.A., Lee, J. et al. (2021) *Biomed Res. Int.*, 2021, 8853056.
2. Sulimov, V.B., Kutov, D.C., Taschilova, et al. (2021) *Curr. Top. Med. Chem.*, 21, 6.

PITFALLS OF SARS-COV2 MAIN PROTEASE COVALENT INHIBITION MODELING WITH THE COMBINED QUANTUM AND MOLECULAR MECHANICS APPROACHES

I.V. Polyakov^{1,2}, M.G. Khrenova^{1,2}

¹Lomonosov Moscow State University, Moscow, Russia

²Emanuel Institute of Biochemical Physics RAS, Moscow, Russia

The protein-ligand interactions of target enzyme and potential drug molecules can be assessed with different models including the quantum mechanics-based approaches which are invaluable in studying covalent inhibitors [1]. Combined quantum and molecular mechanics (QM/MM) methods are routinely used to study mechanisms of chemical reactions in proteins and protein complexes [2]. The SARS-CoV-2 main protease (M^{pro}) interaction with the drug molecules was recently simulated with different QM/MM protocols [3, 4]. M^{pro} inhibition mechanism by PF-07321332 drug nirmatrelvir was uncovered; first, the ion pair intermediate is formed by Cys145 side chain deprotonated by the His41 side chain, second, the covalent C-S bond is formed, third, the resulting intermediate undergoes a proton transfer from the protonated but the resulting free energy profiles were in a striking disagreement of up to 11 kcal/mol in the reaction barrier height which translates in enormous discrepancies of the catalytic rate constant.

In this work we report the results of QM/MM modeling of the nirmatrelvir reaction with the M^{pro} . The potential energy QM/MM surface (PES) reaction profiles were obtained with the combination of ChemShell [5] and Turbomole. The PBE0-D3/6-31G**//CHARMM36 level of theory was employed, total atom count was over 10.5 thousand and the quantum part comprised 132 atoms. Two different starting points from the relaxed QM/MM molecular dynamics trajectories were chosen such as to mimic the enzyme-substrate complex (ES) and the reaction intermediate (INT). The resulting energy profiles were found to be in striking disagreement. One of the profiles was characterized by several low energy barriers (2-7 kcal/mol), while the other contained an 18 kcal/mol energy barrier.

The source of these disagreements was found to be in the structure of the active site. The INT model enzyme-substrate structure was characterized by closer S-C attack distance and better solvation of the nitrogen atom of the nitrile group of nirmatrelvir. Basically, the static PES profiles imply that there is a non-obvious reaction coordinate associated with the substrate position on the surface of the M^{pro} active site which is not probed during the conventional PES scans but can be grasped by the QM/MM molecular dynamics simulation if the sampling coordinate and trajectory length were chosen correctly. Thus, one needs to be extremely cautious doing computational design of covalent inhibitors if the active site is on the surface of the protein.

This research was funded by the Russian Science Foundation (grant number 19-73-20032). The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

1. Sadybekov, A.V., Katritch, V. (2023) *Nature* 616, 673–685.
2. Khrenova, M.G., Mulashkina, T.I., Kulakova, A.M., et al. (2024) *Moscow Univ. Chem. Bull.* 79, 86–92.
3. Grigorenko, B.L., Polyakov I.V., Khrenova, M.G., et al. (2023) *J. Am. Chem. Soc.* 145(24), 13204-13214.
4. Ramos-Guzmán, C.A., Ruiz-Pernia, J.J., Tunon, I. (2021) *Chem. Commun.* 57, 9096.
5. Lu, Y., Sen, K., Yong, C., et al. (2023) *Phys. Chem. Chem. Phys.*, **25**, 21816-21835.

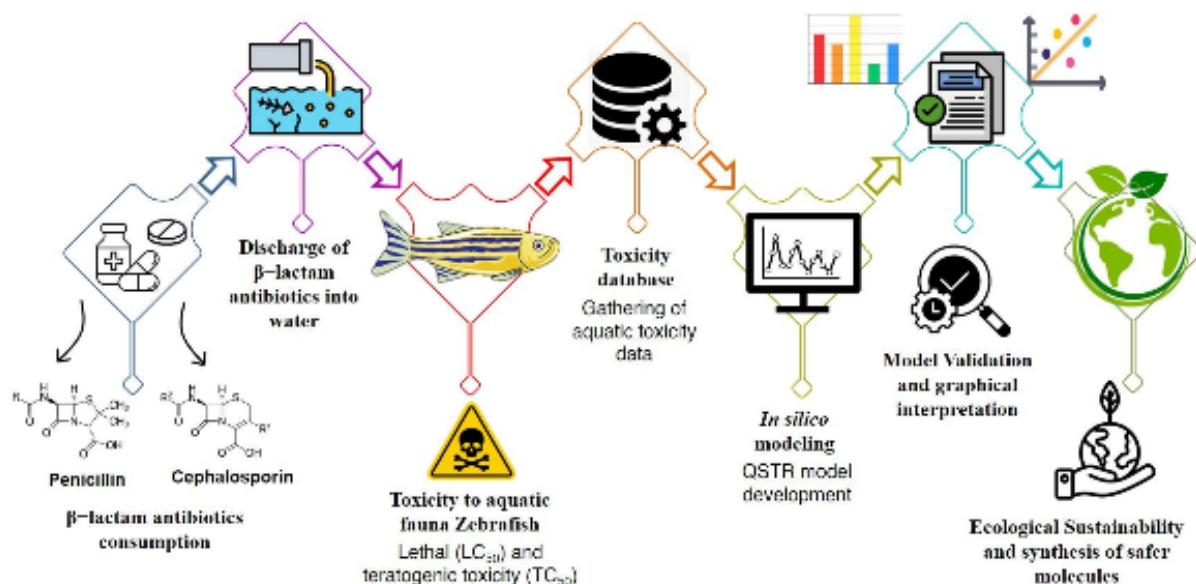
MODELLING LETHALITY AND TERATOGENICITY OF ZEBRAFISH (*DANIO RERIO*) DUE TO β -LACTAM ANTIBIOTICS EMPLOYING THE QSTR APPROACH

A. Nath¹, P.K. Ojha¹, K. Roy²

¹*Drug Discovery and Development Laboratory, Department of Pharmaceutical Technology,
Jadavpur University, Kolkata, India*

²*Drug Theoretics and Cheminformatics Laboratory, Department of Pharmaceutical Technology,
Jadavpur University, Kolkata, India*

Nowadays, β -lactam antibiotics are one of the most consumed OTC (over-the-counter) medicines in the world. Its frequent use against several infectious diseases leads to the development of antibiotic resistance. Another unavoidable risk factor of β -lactam antibiotics is environmental toxicity. Numerous terrestrial as well as aquatic species have suffered due to the excessive use of these pharmaceuticals. The eco-toxicity data for β -lactam antibiotics are quite limited, so the unavailable toxicity data for β -lactams are required to be further explored. To investigate such toxicities, in this present study, we have performed a toxicity assessment employing a novel *in silico* technique like quantitative structure-toxicity relationships (QSTRs) to explore toxicity against zebrafish (*Danio rerio*). We have developed single as well as inter-endpoint QSTR models for the β -lactam compounds to explore important structural attributes responsible for their toxicity, employing median lethal (LC_{50}) and median teratogenic concentration (TC_{50}) as the endpoints. We have shown how an inter-endpoint model can extrapolate unavailable endpoint values with the help of other available endpoint values. To verify the models' robustness, predictivity, and goodness-of-fit, several universally popular metrics for both internal and external validation were extensively employed in model validation (single endpoint models: $R^2 = 0.631\text{--}0.750$, $Q^2_{FI} = 0.607\text{--}0.684$; inter-endpoint models: $R^2 = 0.768\text{--}0.840$, $Q^2_{FI} = 0.678\text{--}0.760$). Again, these models were engaged in the prediction of these two responses for a true external set of β -lactam molecules without response values to prove the reproducibility of these models. By performing this study, we can say that the effort and resources dedicated to experimental research on extrapolating toxicities of new β -lactam compounds to humans and other animals could be reduced, which will ultimately help to achieve the ideas behind green chemistry and the 3Rs (Refinement, Replacement, and Reduction) by minimization of testing of animals.



MOLECULAR MODELING OF BACTERIAL RESISTANCE: THE ROLE OF DYNAMIC BEHAVIOR OF PROTEIN COMPLEXES WITH SUBSTRATES OR INHIBITORS

M.G. Khrenova^{1,2}, A.V. Krivitskaya², Ivan V. Smirnov^{1,3}, A.M. Egorov¹

¹Lomonosov Moscow State University, Moscow, Russia

²Fundamental Research Centre of Biotechnology of the Russian Academy of Sciences, Moscow, Russia

³Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia

Among the mechanisms of antibiotic resistance, the important role plays those associated with the interaction of bacterial enzymes with beta-lactam antibiotics, since it is this group of compounds that occupies approximately two thirds of the market for antibacterial therapeutics. These proteins include antibiotic targets – penicillin-binding proteins, as well as beta-lactamases, which are responsible for inactivation of antibiotics. From the point of view of enzymology, all these enzymes belong to the class of hydrolases. Therefore, for a comprehensive study of the mechanisms of reactions, including their comparison and the search for ways to control enzymatic activity, it is necessary to develop common approaches. A large amount of experimental material has been accumulated in the literature on the substrate specificity of proteases of various types, including serine and zinc-dependent ones.

Herein, we present the results of molecular modeling, which includes calculations of molecular dynamic trajectories with classical and combined QM/MM potentials, analysis of geometric parameters and characteristics of electron density, as well as data processing using artificial intelligence methods to determine differences in stationary catalytic parameters observed in the experiment. Also, we demonstrate the importance of dynamic behaviour of complexes and its relation to the binding efficiency. The following examples will be considered: (1) mechanisms of interaction of metallo-beta-lactamases L1 and NDM-1 with cephalosporins, carbapenems and inhibitors, organic boric acids and unithiol; (2)interpretation of the effect of amino acid substitutions in penicillin-binding protein PBP-2 on interaction with ceftriaxone.

The work was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement no. 075-15-2024-536. The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

DECIPHERING CAP1 AS A SIGNIFICANT DRUGGABLE TARGET IN COMBATING *CANDIDA ALBICANS* PATHOGENESIS AND MULTIDRUG RESISTANCE

N. Jaiswal, A. Kumar

Department of Biotechnology, National Institute of Technology, Raipur (CG), India

Candida albicans is a significant opportunistic fungal pathogen known for its virulence and capacity to develop multidrug resistance (MDR), complicating treatment efforts. Its pathogenicity is driven by factors such as adhesion to host tissues, morphological switching between yeast and filamentous forms, biofilm formation, and the secretion of hydrolytic enzymes. These mechanisms allow *C. albicans* to evade the host immune response, persist on medical devices, and resist available antifungal treatments. In our study, we investigated the CAP1 protein as a potent therapeutic target due to its critical role in these processes. Further we identified its localization and it was found that CAP1 is located in the cytoplasm that further make it as a viable drug target. The gene ontology analysis reveals that CAP1 is involved in crucial cellular functions, including metabolism and regulation, suggesting that inhibiting CAP1 could disrupt essential processes. Our findings reveal that CAP1 is expressed in both planktonic, hyphal and biofilm stages of *C. albicans*, playing a pivotal role in the transition from planktonic to hyphal and biofilm states. The interaction analysis via string database and cytoscape highlights the extensive protein-protein interaction network centered around CAP1. This network includes key proteins such as ALS3, HWP1, TUP1, SAP4 and others involved in MDR like MRR1, MDR2, CDR1 and biofilm formation. CAP1's interactions with these proteins suggest its crucial role in phase switching, regulating virulence and pathogenicity. The identification of CAP1 as a central hub protein within this network underscores its significance in the regulation of MDR and biofilm formation of *C. albicans* which highlights its potential as a promising futuristic target for the development of effective antifungal agents.

MODELING, SYNTHESIS AND *IN VITRO* TESTING OF PEPTIDES BASED ON UNUSUAL AMINO ACIDS AS POTENTIAL ANTIBACTERIAL AGENTS

A.S. Sargsyan¹, L.T. Karapetyan¹, A.V. Mkhitarian, L.A. Stepanyan, T.H. Sargsyan^{1,2}, Yu.M. Danghyan¹, A.V. Sargsyan, G.G. Oganezova¹, N.A. Hovhannisyan^{1,2}

¹Scientific and Production Center "Armbiotechnology" NAS RA, , Yerevan, Armenia

²Yerevan State University, Yerevan, Armenia

Many proteases are of great interest for pharmacology due to their key role in various pathologies. Bacterial collagenase (EC 3.4.24.3) is a quite attractive target for drug development as the inhibitors of bacterial collagenolytic protease may stop disease propagation caused by infections [1].

A search for peptides with the ability to inhibit bacterial collagenases was conducted. For this purpose, multiple peptides with protective groups were modeled. The structures of compounds capable of interacting with bacterial collagenase were predicted based on docking analysis using AutoDock Vina software. The crystallographic structure of collagenase G was taken from the RCSB database (PDB-ID: 2Y50). After *in silico* screening, the selected compounds were synthesized using the activated esters method. The influences of the synthesized compounds on collagenase activity were studied *in vitro* by the known method [2]. Using various concentrations of the peptides the values of IC₅₀ were determined.

According to the obtained results, the most suitable molecules were:

Name	ΔG (kcal/mol)	IC ₅₀ (mM)
N-t-BOC-(S)-b-[4-allyl-3-(pyridine-4-yl)-5-thioxo-1,2,4-triazol-1yl]-a-alanine	-8.8	1.45
N-t-BOC-(S)-b-[4-allyl-3-(pyridine-4-yl)-5-thioxo-1,2,4-triazol-1yl]-a-alanyl-glycine	-8.5	1.35
N-t-BOC-(S)-b-[4-allyl-3-(pyridine-3-yl)-5-thioxo-1,2,4-triazol-1yl]-a-alanine	-9.7	1.56
N-t-BOC-(S)-b-[4-allyl-3-(pyridine-3-yl)-5-thioxo-1,2,4-triazol-1yl]-a-alanyl-glycine	-8.0	1.75

Furthermore, the antibacterial activity of this compounds was tested. The multidrug-resistant strains *P. aeruginosa* 5249, *P.aeruginosa* 9059, *P. aeruginosa* 80, *K. pneumoniae* 63, *E.coli* ESBL 64, *S. maltophilia* 9288 and *Staphylococcus aureus* MDC 5233 were selected as test cultures. The lowest concentration of the compound that prevented visible bacterial growth after incubation was 0.05mM.

This study was funded by the High Education and Science Committee of RA, in the frames of the research projects № 21T-2I235 and № 23T/AA-006.

1. Schönauer E., et. al. (2017) *J. Am. Chem. Soc.*, 139, 12696–12703.
2. Sargsyan A.S., et. al. (2023) *JCTM*, 58(3), 615-620.

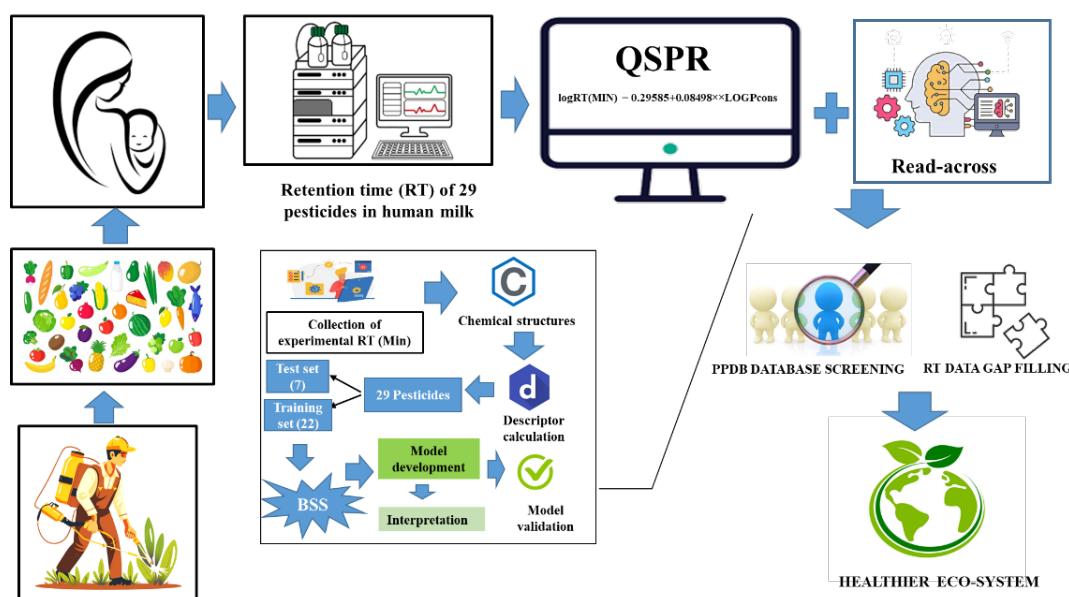
ESTIMATION OF RETENTION TIME OF ORGANIC PESTICIDES IN HUMAN MILK USING QSPR AND READ-ACROSS METHODS: AN ALTERNATIVE APPROACH TO EXPERIMENTAL HAZARDS ASSESSMENT

A. Kumar¹, P. K. Ojha¹, K. Roy²

¹Drug Discovery and Development Laboratory, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India

²Drug Theoretics and Cheminformatics (DTC) Laboratory, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India

Uncontrolled and excessive utilization of pesticides results in serious health effects in humans due to their accumulation in breast milk, blood, etc. Health effects associated with direct or indirect (via food, water, air, etc.) consumption of excessive pesticides include neural effects, carcinogenic effects, liver and kidney disease, abnormal behavior, genotoxicity, infertility, dull behavior in new-born babies, autism spectrum disorders, etc. So, it is necessary to identify the pesticides, harmful and carcinogenic food hazards, and related compounds in human milk. Retention time (RT) prediction is one of the analytical techniques to identify the chemical hazards in milk, it requires complex processes, sophisticated equipment, and highly skilled labor, but it requires complex processes, sophisticated equipment, highly skilled labor, high cost, etc. Therefore, we have developed a QSPR univariate model and read-across method using RT of 29 pesticides in human milk (Ultrahigh Performance Liquid Chromatography Coupled to Tandem Mass Spectrometry) for the estimation of the retention time of pesticides in human milk to overcome the analytical problems and experimental costs. Statistical results ($R^2=0.909$, $Q^2_{LOO}=0.891$, $Q^2_{F1}=0.945$, $Q^2_{F2}=0.942$, Mean Absolute Error (MAE; 95% data)=0.026, Model Quality based on MAE-based criteria='GOOD', Average $rm^2_{(LOO)}=0.852$, Delta $rm^2_{(LOO)}=0.055$, Average $rm^2_{(test)}=0.917$, Delta $rm^2_{(test)}=0.037$; read-across: $Q^2_{F1}=0.960$, $Q^2_{F2}=0.960$, MAE= 0.020) show that models are highly validated, robust, accurate and easily interpretable. From the study, it was inferred that lipophilicity (LOGPcons descriptor) is responsible for the high retention time, which will be helpful in the design of pesticides and chemicals with desired RT. PPDB database screening shows the model's real-world application. This study will be useful for data-gap filling (prediction of RT of untested and new compounds in milk) and identification of food hazards in human milk, which will maintain some healthier eco-systems and also serve as a vital application for food informatics.



- Pedersen, T.L., Smilowitz, J.T., Winter, C.K., Emami, S., Schmidt, R.J., Bennett, D.H., Hertz-Pannier, I., Taha, A.Y. (2021) *J. Agric. Food Chem.*, 69(23), 6676-6689.

MULTI-TARGET NEURAL NETWORK MODEL OF ANXIOLYTIC ACTIVITY OF CHEMICAL COMPOUNDS BASED ON CORRELATION CONVOLUTION OF ENERGY SPECTRA OF MULTIPLE DOCKING

P.M. Vassiliev, M.A. Perfilev, A.V. Golubeva, A.N. Kochetkov, D.V. Maltsev

Volgograd State Medical University, Volgograd, Russia

According to WHO data for 2023, almost 1 billion people worldwide suffer from mental disorders [1]. Most of them are anxiety disorders, which leads to an intensification of the search for new anxiolytic drugs. The development of new approaches to the search for anxiolytic substances using artificial intelligence methods is one of the current areas of modern bioinformatics and pharmacology.

The aim of this work is to construct a multi-target classification neural network model of the dependence of the anxiolytic activity of chemical compounds on the parameters of the correlation convolution of the energy spectra of their multiple docking.

The training set was formed on the basis of the original verified database [2] of the structures of 537 known experimentally studied substances, of which 273 compounds had pronounced anxiolytic activity, and 264 compounds were low active or inactive. Optimized 3D models of these compounds were constructed sequentially using the molecular mechanics method in MarvinSketch 15.6.15 and then by the semi-empirical quantum chemical method PM7 in MOPAC2012. By means of the Open Targets system [3] and the original Microcosm BioS v23.11.2 system [4] using 273 active compounds from the database [2], 22 biotargets relevant to anxiolytic activity were identified. For these biotargets, 22 valid experimental X-ray 3D models were found in the UniProt, PDBe and RCSB PDB databases using the method [5]. Using the original MSite v21.04.22 program according to the method [6], 27 spaces for multiple docking were formed for each target protein. Ensemble docking was performed using AutoDock Vina 1.1.1 according to the method [5], binding energy spectra were obtained for 537 compounds for the entire volume of analyzed proteins (a total of ~1.6 million values). Correlation convolution of the obtained spectra was performed for each target protein using the original CorrConv v24.1.24 program, resulting in 22 convolutional variables. Training of neural networks with the architecture of a two-layer bottleneck perceptron was carried out in the Statistica program [7] using 7 sampling options. In total, ~30 000 neural networks were trained.

The accuracy indicators of the best found model MLP 22-14-2 (tanh, softmax) on the combined training set were $Acc=90.7\%$, $Sens=90.9\%$, $Spec=90.5\%$, $AUC_{ROC}=91.5\%$.

The resulting model is used to search for new substances with anxiolytic activity.

This study was supported in the framework of the state assignment of the Ministry of Health of the Russian Federation No. 2302240009-9 "Development of a methodology for computer searching multi-target pharmacologically active compounds based on multiple docking and technology of the convolutional neural networks of various architectures".

1. WHO (2024) *Mental Health, Brain Health and Substance Use*, <https://www.who.int/teams/mental-health-and-substance-use/overview>
2. Vassiliev, P.M., Maltsev, D.V., Perfilev, M.A. et al. (2022) *Certificate of state registration of the database RU2022621744*.
3. Open Targets (2023), <https://www.opentargets.org/>
4. Vasilyev, P.M., Luzina, O.A., Babkov D.A., et al. (2019) *Journal of Structural Chemistry*, 60(11), 1827-1832.
5. Vassiliev, P.M., Maltsev, D.V., Spasov, A.A., et al. (2023) *Pharmaceuticals*, 16(5), 731.
6. Vassiliev, P.M., Kochetkov, A.N., Spasov, A.A., et al. (2021) *Volgograd scientific and medical journal*, 3, 57-61.
7. Hilbe, J.M. (2007) *The American Statistician*, 61(1), 91-94.

NETWORK PHARMACOLOGY REVEALED THE POTENTIAL OF BITTER HONEY IN SUPPRESSION OF CEREBRAL MALARIA-INDUCED INFLAMMASOME

M.O. Daniyan¹, O.B. Adeoye², E. Osirim³, I.D. Asiyambola¹, A. Shonhai⁴

¹Department of Pharmacology, Faculty of Pharmacy, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria

²Department of Biochemistry, Benjamin Carson (Snr.) School of Basic Medical Sciences, Babcock University, Ilishan-Remo, Ogun State, Nigeria

³Department of Physiology, Redeemer's University, Ede, Osun State, Nigeria

⁴Department of Biochemistry and Microbiology, Faculty of Science, Engineering and Agriculture, University of Venda, Thohoyandou, South Africa

Cerebral malaria (CM) is a fatal complication of *Plasmodium falciparum* infection. The biological and physiological links between CM, inflammation, and inflammasome, testifies to the complexity of its pathology. Resistance to available and affordable drugs, worsening economic crisis, and urgent need for integration of orthodox with traditional/alternative medicine, have made the search for sustainable pharmacotherapy a necessity. With increasing dependence on natural products for daily medical needs in many parts of the world, and the renewed interest in the integration of traditional with orthodox medicine, the need for a better understanding of the pharmacology of natural products is inevitable. Honey is one of the mostly used natural products for varieties of beneficial health purposes. Previous work from our teams on the medicinal properties of bitter honey has established botanical and bioactive markers, inhibitory effects on pancreatic alpha-amylase activity, and anti-dyslipidaemia, cardio-protective, and ameliorative effects on hepatorenal damage in streptozotocin-induced diabetic rats. Network pharmacology, an interdisciplinary field of research that aims to understand the multifaceted interactions between biological systems and drugs at a molecular level, has found usefulness in identifying complex interactions between drugs and biological systems, identifying new drug targets, repurposing existing drugs, and improving the safety and efficacy of drugs. In this report, we have identified bitter honey (BH) derived phytochemicals using gas chromatography coupled with mass spectrometry (GC-MS), and 9 targets from genes associated with CM, inflammation, inflammasome, and BH phytochemicals using Venn analysis. Network analysis revealed significant functional and physical interactions among these targets and NOD-, LRR-, and pyrin domain-containing protein 3 (NLRP3). Molecular docking of bitter honey-derived phytochemicals against these targets identified 3 most promising phytochemical candidates for further experimental validation. Based on these results, we predict that bitter honey may aid in the suppression of CM-mediated inflammasome cell death via its interactions with these targets.

UNIFIED LATIN AMERICAN NATURAL PRODUCT DATABASE: LANaPDB

A. Gómez-García¹, A. Kathrin-Prinz², D.A. Acuña-Jiménez³, W.J. Zamora³, H.L. Barazorda-Ccahuana⁴, M.Á. Chávez-Fumagalli⁴, M. Valli⁵, A.D. Andricopulo⁵, V.S. Bolzani⁶, D.A. Olmedo⁷, P.N. Solís⁷, M.J. Núñez⁸, J.R. Rodríguez-Pérez⁹, H.A. Valencia-Sánchez⁹, H.F. Cortés-Hernández⁹, O.M. Mosquera-Martinez¹⁰, O. Koch², J.L. Medina-Franco¹

¹National Autonomous University of Mexico, Mexico City, Mexico

²Institute of Pharmaceutical and Medicinal Chemistry, Münster, Germany

³University of Costa Rica, San Pedro, Costa Rica

⁴Universidad Católica de Santa María, Arequipa, Peru

⁵University of São Paulo, São Carlos, Brazil

⁶São Paulo State University, Araraquara, Brazil

⁷University of Panama, Panama, Panama

⁸University of El Salvador, San Salvador, El Salvador

⁹Universidad Tecnológica de Pereira, Pereira, Colombia

¹⁰Universidad Nacional Abierta y a Distancia, Dosquebradas, Colombia

Historically, natural products (NPs) have been the largest source of bioactive compounds for medicinal chemistry. From 1946 to 1980, of all small molecules approved as drugs worldwide, 53% were NPs and NP-derivatives and from 1981 to 2019, 64% were NPs and synthetic drugs with a NP pharmacophore [1]. NPs have a unique and favorable property profile that is particularly useful in drug discovery. Compound databases of NPs play a crucial role in drug discovery and have implications in other areas, such as food chemical research, ecology and metabolomics. In this context, Latin America is one of the largest biodiverse regions in the world, it contains at least a third of the global biodiversity [2]. Hence, several Latin American countries have been developing NP collections and making them public. From Latin American NP databases, since 2017, more than ninety compounds have been identified with therapeutical effects in the treatment of diverse diseases or symptoms, such as chagas disease, leishmaniasis, schistosomiasis, age-related diseases and pain [3].

The long-term goal of the project is to collect, unify, and standardize the Latin American NP collections available in the public domain into one public database. The first version of the Latin American Natural Product database (LANaPDB) was recently published [4] and is freely available. The current version contains 13,578 compounds, coming from ten different databases of seven different countries. As part of the project, we have analyzed the chemical contents of LANaPDB in terms of the coverage in the chemical space, distribution of physicochemical properties of pharmaceutical interest, structural content and diversity as well as the commercial availability of the NPs in this collection. LANaPDB is a collective effort of researchers from seven Latin American countries to ensemble a public and representative library of natural products in a geographical region with a large biodiversity.

1. Newman, D.J., Cragg, G.M. (2020) *J. Nat. Prod.*, 83, 770–803.
2. Raven, P.H., Gereau, R.E., Phillipson P.B., et al (2020) *Sci Adv*, 6, 6228.
3. Gómez-García A., Jiménez, D.A.A., Zamora, W.J. ,et al (2023) *Pharmaceuticals*, 16, 1388.
4. Gómez-García, A., Medina-Franco, J.L. (2022) *Biomolecules*, 12, 1202.

IS CHEMICAL SPACE OF RUSSIAN VENDOR CATALOGUES SUFFICIENT FOR LEAD DISCOVERY?

D.I. Osolodkin^{1,2}, V.I. Uvarova¹, A.D. Fomina^{1,3}

¹FSASI “Chumakov FSC R&D IBP RAS” (Institute of Poliomyelitis), Moscow, Russia

²Sechenov First Moscow State Medical University, Moscow, Russia

³Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

Virtual screening of vendor catalogues followed by experimental assessment of hits is a commonly accepted practice of lead discovery in academic medicinal chemistry. While ZINC database comprises a large chemical space covered by numerous vendors, a majority of these compounds is not available for a large-scale procurement from Russian Federation. Consequently, only a few vendors are easily available for Russian scientists eagering to discover new relevant chemotypes. We have performed several virtual screens for potential antivirals in the catalogues of different Russia-based vendors, and the hits were further evaluated in phenotypic and target-based experimental assays. Hit rates in these assays were strikingly different, from ~30% for anti-flavivirus activity to ~1% for protease inhibition [1-3]. Since the methodology and the retrospective validation of the virtual screens gave a hint for their relevance, we supposed that variation of the hit rates could be inferred from the catalogue composition, because no good hits could be found in a catalogue that does not contain relevant molecules.

In the current study, we present the property-based and chemical space-based analysis of the catalogues of the Russian vendors and compare them with our training sets as well as with the other databases, such as ChEMBL and databases of the approved drugs. Based on this analysis, we make conclusions on the applicability of these vendor catalogues for wide scale lead identification.

This study was supported by the State Research Funding for FSASI “Chumakov FSC R&D IBP RAS” (Institute of Poliomyelitis) from the Ministry of Science and Higher Education of the Russian Federation, project no. FNZG-2024-0005.

1. Fomina, A.D., Uvarova, V.I., Kozlovskaya, L.I., et al. (2024) *Molecular Informatics*, 43, e202300279.
2. Zakharova, M.Y., Kuznetsova, A.A., Uvarova, V.I., et al. (2021) *Frontiers in Pharmacology*, 12, 773198.
3. Orlov, A.A., Khvatov, E.V., Koruchekov, A.A., et al. (2019) *Molecular Informatics*, 38, 1800166.

LATIN AMERICAN PHYTOCHEMICAL DERIVATIVES AS PROMISING CANDIDATES FOR GALLSTONE DISEASE THERAPY: INSIGHTS FROM MOLECULAR SCREENING, MOLECULAR DOCKING, DENSITY FUNCTIONAL THEORY, AND MOLECULAR DYNAMICS STUDIES

J. Tamayo, A. Latorre-Palomino, V. García

¹Academic Department of Physical Chemistry, Universidad Nacional Mayor de San Marcos, UNMSM, Lima, Peru

Cholelithiasis, commonly referred to as gallstones, signifies the presence of concretions, or solid masses, within the biliary ducts or the gallbladder. This pathological condition exhibits a global prevalence ranging from 10% to 20% in the adult population, with an elevated incidence of up to 20% observed in industrialized nations. Gallstones formation is associated with an imbalance in cholesterol levels and an elevated concentration of unconjugated bilirubin. Within this context, the nuclear receptor FXR (Farnesoid X receptor) functions as a pivotal sensor for bile acids, orchestrating enterohepatic circulation by regulating gene expression within the intricate network of cholesterol and triglyceride metabolism. Conversely, the enzyme beta-glucuronidase plays a critical role in the deconjugation of glucuronidamines from bilirubin, culminating in the precipitation of calcium bilirubinate. Consequently, the central objective of this research endeavor is the identification of phytochemical compounds sourced from the Latin American database to design novel nature-derivatives with the potential to act as modulators of FXR and inhibitors of beta-glucuronidase. The research methodology was structured as follows: Initially, all herbal compounds were subjected to a bioactivity probability filter by PASS online and Molinspiration tool, followed by a comprehensive docking procedure involving the shortlisted candidates through AutodockVina software. Subsequently, those phytochemicals demonstrating favorable energy interactions were subjected to a *in silico* ADMET assessment using online platforms such as SwissADME, PreADMET, pkCSM and Datawarrior to design novel derivatives through rational isosterism. Subsequent investigations encompassed Density Functional Theory (DFT) calculations, Non-Covalent Interactions (NCI) and Molecular Dynamics simulations, allowing for an in-depth analysis of the reactivity, nature interactions and temporal stability, respectively. This systematic approach enabled the elucidation of 2 potential candidates derived from Latin American cultural ancient phytochemicals with therapeutic promise in the management of cholelithiasis.

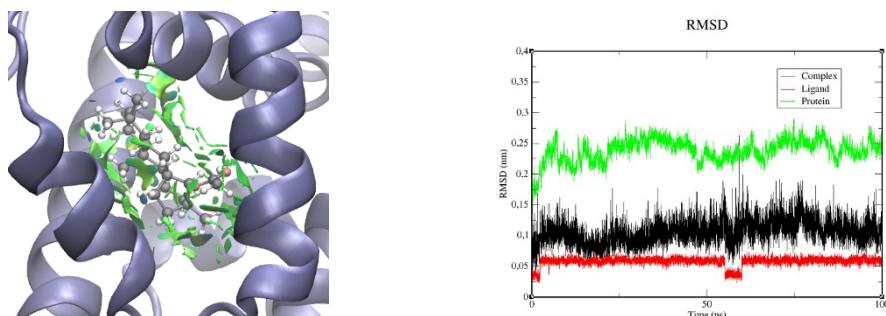


Fig. 1. Isosurface interaction by NCIPILOT4 and RMSD from molecular dynamics of the best natural derivative and FXR.

1. Jonker, J. W., Liddle, C., Downes, M. (2012). *The Journal of Steroid Biochemistry and Molecular Biology*, 130(3-5), 147-158.

MACHINE LEARNING ESTIMATION OF THE SMALL MOLECULE SELECTIVITY INDEX VALUE FOR INFLUENZA VIRUS STRAIN A/H1N1

A.D. Egorov¹, Ya.V. Gorohov², M.M. Kuznetsov¹, S.S. Borisevich^{1,2}

¹National Research Nuclear University “MEPhI”, Institute of Intelligent Cybernetic Systems, Moscow, Russia

²Ufa Federal Research Center of the Russian Academy of Sciences, Ufa Institute of Chemistry, Ufa, Russia

An important task in the field of medicinal chemistry is the creation of predictive models that allow assessing the value of a given biological activity of compounds, anticipating the stages of synthesis and in vitro (in vivo) experiments [1]. The most modern method for estimating the values of various quantities, if it is impossible to derive an analytical formula, is machine learning. In this work, machine learning models (both classical and neural network) were used to predict (the classification problem and separately the regression problem were solved) the antiviral activity against the influenza virus strain A/H1N1 of small organic molecules. The compound database was compiled by analyzing data presented in the ChEMBL database and in some new publications from 2022 to 2024. The collected database contains about 2000 compounds, for each of which the values of IC_{50} (inhibitory activity of compounds against influenza virus strain A/H1N1 - MTT test), CC_{50} (cytotoxicity against the MDCK cell line) and selectivity index (SI - ratio of CC_{50} to IC_{50}). Using classical machine learning methods (random forest, gradient boosting, logistic regression) and using modern neural network models (BERT-based transformers), the problem of predicting the exact value of the selectivity index and the problem of binary classification (assigning an active or inactive compound to a class) were solved. The accuracy of solving the classification problem was up to 0.95, depending on various parameters. The prediction results are in agreement with the experimental data.

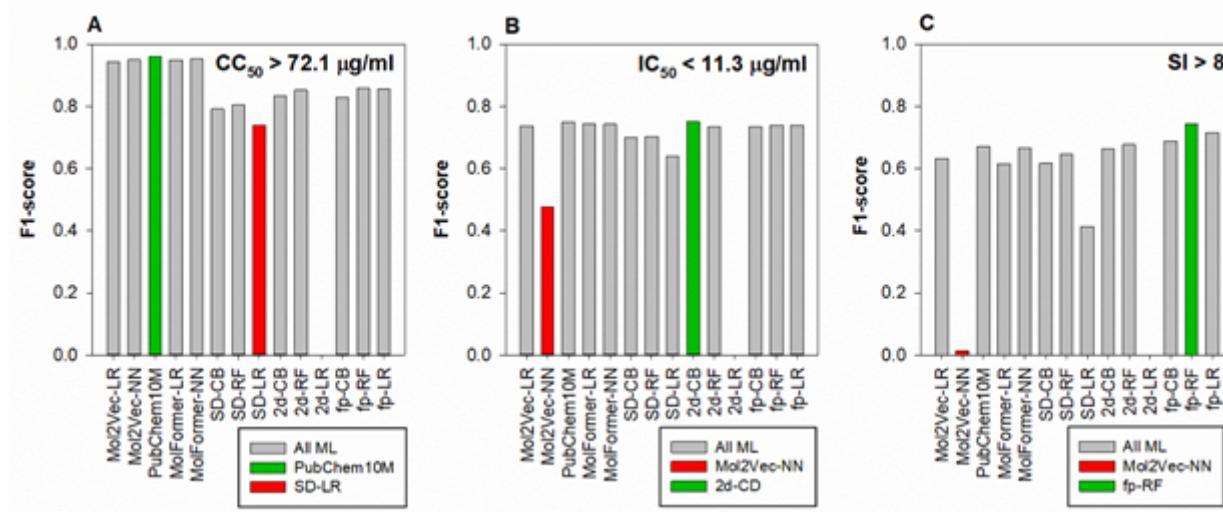


Fig. 1. Biological parameters prediction f_2 -score in binary classification task.

1. Zhavoronkov, A., Ivanenkov, Y.A., Aliper A., Veselov, M.S., Aldinskikh, V.A., Aldinskaya, A.V., Terentiev, V.D., Polykovskiy, D.A., Kuznetsov, M.D., Asadulaev, A., Volkov ,Y., Zhulus, A., Shaykhetmetov, R.R., Zherbak, A., Minaeva, L.I., Zagribelnyy, B.A., Lee, L.H., Soll, R., Madge, D., Xing, L., Guo ,T., Aspuru-Guzik, A. D(2019) *Nat. Biotechnol.*, 37(9), 1038–1040.

MACHINE LEARNING PREDICTIONS OF COCRYSTAL FORMATION TO ENHANCE ACTIVE PHARMACEUTICAL INGREDIENTS PROPERTIES FOR CANCER PREVENTION

N.Q.K. Le^{1,2}

¹AIBioMed Research Group, Taipei Medical University, Taipei, Taiwan;

²Professional Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

The objective of this study is to leverage machine learning (ML) techniques to predict the formation of cocrystals with enhanced physicochemical properties of active pharmaceutical ingredients (APIs), particularly anticancer drugs, to improve their oral bioavailability and therapeutic efficacy.

Traditional methods for developing and optimizing cocrystals are often time-consuming and resource-intensive. This study advances previous work by integrating machine learning models, specifically LightGBM (LGBM) and support vector machine (SVM), to predict cocrystal formation, thereby significantly reducing the time and cost associated with experimental trials. Furthermore, this study explores the use of deep learning encoders and feature selection techniques to enhance the prediction accuracy and generalization capabilities of the models.

The study utilized various ML models and feature extraction techniques, such as UniRep, BERT, and SSA, to encode amino acid sequences of anticancer peptides (ACPs) and non-ACPs. Key findings include:

- The UniRep encoder combined with SVM achieved the highest performance with an accuracy of 77.2%.
- Feature selection techniques like LGBM and ET significantly improved model performance by selecting the most relevant features from the dataset.
- The implementation of LGBM-related algorithms and non-tree-based algorithms (e.g., TabNet) provided a robust framework for cocrystal formation prediction.

The predictive models demonstrated improved accuracy and efficiency in identifying potential cocrystal formations, enhancing the physicochemical properties of APIs, and thereby improving their oral bioavailability and therapeutic efficacy. These findings are significant as they pave the way for the development of more effective anticancer drugs with reduced side effects and increased patient compliance.

In conclusion, this study successfully demonstrates the application of machine learning techniques in predicting the formation of cocrystals with enhanced physicochemical properties, offering a promising approach to improve the bioavailability and efficacy of anticancer drugs. Future research will focus on conducting more extensive experiments on feature combinations and selections, as well as applying domain-specific knowledge to further refine the predictive models.

This study was supported by the National Science and Technology Council, Taiwan [grant number MOST111-2628-E-038-002-MY3].

1. Alley, E.C., Khimulya, G., Biswas, S., AlQuraishi, M., Church, G.M. (2019) *Nature Methods*, 16(12), 1315–1322.
2. Rao, B., Zhou, C., Zhang, G., Su, R., Wei, L. (2020) *Briefings in Bioinformatics*, 21(5), 1846–1855.
3. Lv, Z., Cui, F., Zou, Q., Zhang, L., Xu, L. (2021) *Briefings in Bioinformatics*, 22(5), bbab008.
4. Yuan, Q., Chen, K., Yu, Y., Le, N.Q.K., Chua, M.C.H. (2023) *Briefings in Bioinformatics*, 24(1), bbac630.

CHEMOMETRICS GUIDED LEAD IDENTIFICATION AND DESIGN OF NOVEL ANALOGS WITH TRYPANOOTHIONE REDUCTASE ACTIVITY BASED ON 2-AMINOBENZIMIDAZOLE SCAFFOLDS AND MOLECULAR SIMULATIONS FOR ADDRESSING LEISHMANIASIS

A. Biswas, A. Bhattacharjee, S. Kar, P.K. Ojha

*Drug Discovery and Development Laboratory (DDD Lab), Department of Pharmaceutical Technology,
Jadavpur University, Kolkata, India*

We have developed a robust quantitative structure–activity relationship (QSAR) model against Leishmaniasis, a major Neglected Tropical Disease (NTD), affects millions globally. Current treatments are hampered by infection relapse, high toxicity and lengthy regimens. A contemporary study investigated 2-aminobenzimidazole scaffold for inhibition of Trypanothione reductase (TR), a critical mediator of Leishmaniasis etiogenesis. However, the obtained activity (IC_{50}) was moderately potent however they showed admissible toxicity *in vivo*. This inspired us to innovate a QSAR model of TR inhibitory activity is built on the available *in vivo* experimental data. Simple 2D molecular descriptors with ease of interpretability was employed so that the elucidated structural features closely associated with the inhibitory activity can be utilized to design new molecules. 2D QSAR was leveraged to select the final model strictly catering to OECD guidelines and Intelligent Consensus Prediction (ICP) was deployed to further enforce the model prediction. Hydrophobicity, aromatic ring, hydrogen bond acceptor/donor, as well as heteroatoms (nitrogen, fluorine, etc.) enhance the restrictive activity. The developed model was deployed to anticipate the TR suppressing activity of the DrugBank compounds. QSAR-guided favourable structural modifications were undertaken to generate prospective analogs of the top leads. The molecular interactions and bonding affinity of the analogs with respect to co-crystallized experimental TR inhibitor (TRL190) at the binding site was assessed using molecular docking. A comprehensive *in-silico* ADMET evaluation was conducted and resilience of protein-ligand interactions of the top candidates (DB03231-A6 and DB12269-A4) with respect to the apo-TR was studied using 300ns molecular dynamics simulation. Upon cumulative retrospection DB12269-A4 emerges as the most promising lead analog justifying additionally *in vitro* and *in vivo* validation for prospective application in Leishmaniasis.

EVALUATION OF NATURAL FLAVONOİD COMPOUNDS İN AMYLOİD-B (AB) INHİBİTİON FOR ALZHEIMER'S DİSEASE TREATMENT

A.M. G.Surya¹, C.B. Trifena², I.K. Cahyadi², J. Shemuel¹, A.A. Parikesit¹

¹Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia

²Department of Biotechnology, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that is marked by the gradual buildup of amyloid plaques and neurofibrillary tangles within the brain. This accumulation leads to significant cognitive decline and memory loss. To date, only a couple of pharmaceutical drugs have received approval for the treatment of Alzheimer's, yet these medications do not directly address the underlying amyloid β (A β) plaques and neurofibrillary tangles (NFT) that are characteristic of the disease. Consequently, there has been a growing interest in the potential of plant-based herbal remedies, which are believed to enhance cognitive abilities and mitigate the symptoms associated with AD. In light of this, the current research is focused on investigating the efficacy of natural flavonoid compounds in inhibiting the formation of amyloid β , thereby offering a novel therapeutic approach for AD. Utilizing a computational, structure-based methodology, this study seeks to identify and evaluate flavonoids that could serve as effective inhibitors of amyloid β [1]. The three-dimensional structures of the amyloid β protein target, along with those of various flavonoid compounds, have been sourced from the Protein Data Bank (PDB) and the PubChem database, respectively. The initial phase of the study involves a thorough screening of the flavonoids to assess their drug-like properties, employing ADME (Absorption, Distribution, Metabolism, and Excretion) profiling and toxicity predictions. This is accomplished using the SWISS-ADME and ProTox-II web servers. Subsequently, a Quantitative Structure-Activity Relationship (QSAR) analysis is conducted to identify compounds that may act as inhibitors of amyloid beta aggregation or antagonists to the beta-amyloid protein. Following the screening process, molecular docking studies are undertaken to determine the binding affinities of the flavonoid ligands to the amyloid β protein target. This is achieved through the use of the CB-Dock2 tool. Visualization techniques are then applied to further analyze the interactions between the ligands and the protein target. The final stage of the research involves conducting molecular dynamics simulations. These simulations are performed using the CABSFLEX2 web-server, which allows for the prediction of protein movements and fluctuations. The simulations focus on the root mean square fluctuations (RMSF) of the most promising flavonoid compound identified. Through the combined use of drug-likeness assessments, toxicity evaluations, and QSAR analysis, epicatechin has emerged as the flavonoid compound with the highest potential to act as an amyloid β inhibitor. Notably, two distinct poses of epicatechin have been identified, each exhibiting the same binding energy but occupying different pocket sizes within the protein structure. The comprehensive findings of this study underscore the promise held by natural flavonoid compounds as alternative therapeutic agents in the fight against Alzheimer's Disease, particularly in targeting the activity of amyloid β plaques. The insights gained from this research could pave the way for the development of more effective treatments for AD, harnessing the power of natural compounds to combat this debilitating condition.

This study was supported in the framework of internal research grant of the Department of Research and Community Service of Indonesia International Institute for Life Sciences (LPPM i3L) at the year 2024.

1. Pratama, R.A., Astina, J., Parikesit, A.A. (2023) *J. Trop. Biodivers Biotechnol.*, 8, 83124.

QSAR MODELS FOR PREDICTION OF CYTOTOXIC IC₅₀ AND GI₅₀ VALUES OF SUBSTANCES IN RELATION TO NON-TUMOR CELL LINES

E.A. Lisitsa¹, A.A. Lagunin^{1,2}

¹*Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia*

²*Department of Bioinformatics, Institute of Biomedical Chemistry, Moscow, Russia*

A cytotoxicity test is conducted to evaluate the safety of new drugs in preclinical studies of potential substances. In actual experimental studies, it can be challenging to organize such testing with dozens of cell lines due to the complexity and labor-intensive nature of the process. We propose to make a prediction of cytotoxicity prior to conducting a real-world experiment, based solely on the structural formulas of potential substances. We have early developed classification SAR structure-activity relationships models for prediction of cell-line cytotoxicity including action on non-tumor (normal) cell lines [1]. The goal of this study is to develop quantitative structure-activity relationships (QSAR) models in order to predict the cytotoxic IC₅₀ and GI₅₀ values of substances in relation to non-tumor cell lines.

The creation of a dataset with information on the structural formulas, cytotoxic IC₅₀ and GI₅₀ values of experimentally tested compounds was based on the freely available ChEMBL database. The DataWarrior program and the KNIME analytical platform were used to process ChEMBL data. Quantitative structure-activity relationship models were created using the GUSAR software [3-5]. From 33 non-tumor human cell lines analyzing in the ChEMBL database (v. 33) 17 ones were selected based on the availability of data with IC₅₀ or GI₅₀ values for more than 100 compounds (BEAS-2B, BJ, CCD-18Co, GES1, HaCaT, HEK-293T, HEK293, HFF, HFL1, HMEC-1, HUVEC, MCF-10A, MRC5, NHDF, PBMC (homo sapiens), TERT-RPE1, and WI-38). For each of these cell lines, a training set with unique organic compound structures was created. Only two cell lines, HUVEC and MRC5, had more than 100 GI₅₀ records in the ChEMBL database, so the training set related with GI₅₀ values were created for these two lines. The IC₅₀ and GI₅₀ values were converted into pIC₅₀ and pGI₅₀ values. Then these data were used to create QSAR models for predicting the cytotoxicity of molecules. The accuracy of the predictions was assessed using R², Q², and standard deviation (SD). For pIC₅₀, the R² values varied from 0.694 to 0.910, Q² from 0.579 to 0.839, and SD from 0.250 to 0.669. The mean values of R², Q² and SD for QSAR models predicted pGI₅₀ values were 0.792, 0.699 and 0.654, respectively. The created QSAR models show sufficiently accuracy for using in the assessment of the cytotoxicity of new compounds. The created models have been implemented to the freely available CLC-Pred web service for cytotoxicity prediction [2]: <https://www.way2drug.com/Cell-line/>.

This work was supported by the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (No. 122030100170-5).

1. Lagunin, A.A., Dubovskaja, V.I., Rudik, A.V., et al. (2018) *PLoS One*, 13(1), e0191838.
2. Lagunin, A.A., Rudik, A.V., Pogodin, P.V., et al. (2023) *Int. J. Mol. Sci.*, 24(2), 1689.
3. Filimonov, D.A., Zakharov, A.V., Lagunin, A.A., Poroikov, V.V. (2009) *SAR QSAR Environ Res.*, 20(7-8), 679.
4. Lagunin, A.A., Sezganova ,A.S., Muraviova, E.S., et al. (2024) *SAR QSAR Environ Res.*, 35(1), 1.
5. Lagunin, A.A., Romanova, M.A., Zadorozhny, A.D., et al. (2018) *Front Pharmacol.*, 9, 1136.

DIGEP-PRED 2.0: A WEB-SERVICE FOR PREDICTING DRUG-INDUCED CELL SIGNALING AND GENE EXPRESSION CHANGES

S.M. Ivanov^{1,2}, A.V. Rudik¹, A.A. Lagunin^{1,2}, D.A. Filimonov¹, V.V. Poroikov¹

¹*Department of Bioinformatics, Institute of Biomedical Chemistry, Moscow, Russia*

²*Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia*

The analysis of drug-induced gene expression profiles (DIGEPs) measured in cell lines is widely used to estimate the potential therapeutic and adverse drug effects, and also molecular mechanisms of drug action. However, the corresponding experimental data is absent for many existing drugs and drug-like compounds. To solve this problem, we created the DIGER-Pred 2.0 web-service [1] which allows predicting DIGEPs and direct drug targets by structural formula, and connect them by paths in signaling and regulatory networks. It is based on the combined use of structure-activity relationships (SARs) and network analysis.

To perform prediction of DIGEP, we created the SAR models based on the corresponding data from Comparative Toxicogenomics Database (CTD) and Connectivity Map (CMap). To perform prediction of direct drug targets and associated molecular mechanisms of action (MoAs), we created SAR models based on the data from ChEMBL and PubChem. Both types of SAR models were built by PASS (Prediction of Activity Spectra for Substances) software. The mean accuracy of prediction calculated by leave-one-out cross validation was 86.5 % for 13377 genes and 94.8 % for 2932 proteins (CTD data), and it was 97.9 % for 2170 MoAs corresponding to 1940 individual human proteins. SAR models (mean accuracy - 87.5 %) were also created for CMap data given on MCF7, PC3, and HL60 cell lines with different threshold values for the logarithm of fold changes: 0.5, 0.7, 1, 1.5, and 2.

Using only the structural formula of a compound, the user can obtain information on potential gene expression changes in several cell lines and direct drug targets, and, then, perform two types of analysis. First, the user can perform enrichment analysis to identify KEGG and Reactome pathways, Gene Ontology biological processes and diseases from DisGeNET associated with predicted DIGEP. Second, the network analysis can be performed. To do it, the transcription factor (TF) enrichment analysis of predicted up- and down- regulated genes is performed. Next, the upstream analysis is executed to link the estimated TFs with the direct targets of a compound predicted by SAR models through a signaling network. As a result, the user can obtain information on potential drug targets, which are probable master regulators responsible for drug-induced gene expression changes in several cell lines under different conditions. All the above-mentioned results can be obtained based only on the structural formula of a query compound and can be used to estimate the therapeutic, side and toxic effects of drug-like compounds, along with the underlined molecular mechanisms.

DIGER-Pred 2.0 web-service is freely available at <https://www.way2drug.com/digep-pred>.

This work has been supported within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) (No. 122030100170–5).

1. Ivanov, S.M., Rudik, A.V., Lagunin, A.A., Filimonov, D.A., Poroikov, V.V. (2024) *Mol. Inform.*, 43, e202400032.

THE QSAR STUDY OF THE HYDROLYSIS OF DINITROSYL IRON-SULFUR COMPLEXES

V.B. Luzhkov^{1,2}, V.B. Krapivin¹, N.A. Sanina¹

¹*Federal Research Center of Problems of Chemical Physics and Medicinal Chemistry RAS,
Chernogolovka, Moscow region, Russia*

²*Department of Fundamental Physico-Chemical Engineering, Lomonosov Moscow State University, Moscow, Russia*

The dinitrosyl iron-sulfur complexes $[\text{Fe}^+(\text{NO})_2\text{L}_2]$ (DNICs) are important inorganic compounds with many interesting chemical, biological and medicinal applications. In this presentation we report a computational study of the hydrolysis of thio ligands (L) in DNICs in water and the related quantitative structure–activity relationships (QSARs). The quantum chemical calculations of molecular structures and electronic properties of reaction products and intermediates have been carried out in the framework of density functional theory and implicit water model. In particular, the TPSSH and M06 functionals and def2-TZVP basis set was used in calculations.

The thio ligand hydrolysis in a prototypic DNIC $[\text{Fe}^+(\text{NO})_2(\text{SCH}_2)_2]$ was found to be an exothermic process with small activation energy, whereas exchange of NO for H_2O is thermodynamically unfavorable [1]. The calculations have predicted lower activation barrier for the associative mechanism with concerted replacement of SCH_2 by H_2O than for dissociative mechanism with homolytic bond cleavage of the Fe–S coordination bond in water. The mechanism of hydrolysis that involves participation of OH^- was found to be less probable at pH 7. The coordination numbers of the central iron in the studied reactions varied from two to five and the optimized geometries of reactants in solution displayed tetrahedral, trigonal bipyramidal, trigonal pyramid, and triangular shapes.

The energies of the Fe–S coordination bonds $D(\text{Fe–S})$ of 26 thiocarbonyl and thiolate DNICs have been determined from the DFT calculations of thio ligands abstraction and exchange for H_2O molecule [2,3]. The bond energies of the first ligand have the values of $D(\text{Fe–S})$ in the range of 16.6–65.4 kJ mol⁻¹ for neutral thiocarbonyl and 87.0–133.8 kJ mol⁻¹ for charged thiolate ligands. Overall, the computed energies of the Fe–S bonds are significantly lower than the known bond energies of sulfur in organic compounds. The calculated energies of hydrolysis acquire negative values for some of thiocarbonyl ligands that points to an unstable character of these compounds in water. The DFT calculations can also provide useful data on the Gibbs free energies of hydrolysis ΔG and the equilibrium constants K in these reactions. The estimated equilibrium constant for the DNIC with the SCH_2 ligand is $\ln(K)=8.15$.

Quantitative structure–activity relationships have been examined between $D(\text{Fe–S})$ and the reactivity indices (RI) of free L [2, 3]. Linear correlations have been found between the calculated $D(\text{Fe–S})$ and partial electronic charge on atom S and the energies of frontier MOs of the free thio ligands in water. These correlations include also linear combinations of the frontier MO energies (electron negativity χ and chemical hardness η). Low values of $D(\text{Fe–S})$ obtained for DNICs can indicate that perturbation theory might be applicable to describe energetics of the Fe–S coordination bonds on a qualitative level. The found correlations with the reactivity indices that are connected to perturbation theory framework support this suggestion. Two-parameter linear regression models based on RI of only a free L were also built to quantify stability of DNICs in hydrolysis of L or substitution for other thio ligands. These two-parameter models can be used to predict thermodynamic stability of a new DNICs in the reactions of hydrolysis. The predicted in this way Fe–S bond energy of glutathione agrees with the experimental data on thio ligand substitution by this thiol.

This work was financially supported by the FRC PCP MC RAS state task No. 124020500019-2.

1. Luzhkov, V.B., Krapivin, V.B., Sanina, N.A., Aldoshin, S.M., (2023) *Russ. J. Phys. Chem. A*, **9**, 813.
2. Luzhkov, V.B., Krapivin, V.B., Sanina, N.A. Aldoshin, S.M. (2024) *Polyhedron*, **257**, 117013.
3. Krapivin, V.B., Luzhkov, V.B., Sanina, N.A., (2024) *Russ. Chem. Bull.*, **73**, inpress.

MOLECULAR MODELING OF HUMAN LINE-1 ORF2 PROTEIN

A.M. Kulakova^{1,2}, T.I. Mulashkina¹, M.G. Khrenova^{1,3}

¹*Lomonosov Moscow State University, Moscow, Russia*

²*Institute of Biomedical Chemistry, Moscow, Russia*

³*Bach Institute of Biochemistry, Federal Research Centre “Fundamentals of Biotechnology” of the Russian Academy of Sciences, Moscow, Russia*

The long interspersed element-1 (LINE-1) is an ancient genetic parasite that has the ability to move to different genome locations through a ‘copy and paste’ mechanism. LINE-1 retrotransposition can lead to a variety of genetic disorders and may play a role in the development of cancer, autoimmune diseases and the aging process. Human LINE-1 encodes two proteins: open reading frame 1 protein and open reading frame 2 protein (ORF2p). ORF2p consists of several domains, two of them being a reverse transcriptase (RT) domain and endonuclease domain. Inhibiting their activities could be a promising approach for therapy.

Herein, we present the results of ORF2p full-atom molecular modelling. Firstly, we study the dynamic behaviour of different form of ORF2p: apo-form, complex with RNA template, hybrid complex with RNA template and DNA primer, and complex with RNA template, DNA primer and deoxythymidine triphosphate. All studied system has an ‘open’ or ‘thumb up’ conformation of ORF2p reverse transcriptase core, which corresponds to the active form of ORF2p. All these calculations were the molecular dynamics simulation in NPT (p=1 atm., T=300 K) ensemble with classical CHARMM force field.

Also, we demonstrate the dynamic behavior of enzyme-substrate complex of ORF2p with thymidine triphosphate which was calculate using molecular dynamic simulation with combined quantum mechanics/molecular mechanics (QM/MM) potentials. The quantum part was calculated using Kohn-Sham method of the density functional theory with the PBE0 functional with D3 dispersion correction and the 6-31G** basis. The molecular mechanical part was described using the CHARMM force field.

The work was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement no. 075-15-2024-643. The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

IN SILICO EVALUATION OF THE MUTAGENICITY, GENOTOXICITY, AND CARCINOGENICITY OF LEVETRACETAM, A NEW-GENERATION ANTIEPILEPTIC

S. Yılmaz-Sarıaltın¹, C.Ö. Yalçın²

¹Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Toxicology, Ankara, Türkiye

²Blacksea Technical University, Faculty of Pharmacy, Department of Pharmaceutical Toxicology, Trabzon, Türkiye

Epilepsy is a chronic neurological disease that affects approximately 50 million people globally, including pregnant women [1]. The choice of antiepileptic drugs during pregnancy continues to be debated, although research has been conducted for a long time. According to the latest treatment protocols, some new-generation antiepileptic drugs, such as lamotrigine, levetiracetam, and oxcarbazepine, can be administered during pregnancy with minimal risk [2]. Although the majority of women with epilepsy have uneventful pregnancies, 90% of children born to women with epilepsy are reported to be healthy. Adverse effects include increased risk of stillbirth, preterm delivery, small for gestational age at birth, low Apgar score at 5 minutes, neonatal hypoglycemia, neonatal infection, respiratory distress syndrome, major congenital malformations, and cognitive and behavioral effects [3]. It is not only the treatment of epilepsy but also seizures during pregnancy that cause harmful effects on the mother and newborn [4, 5]. Exposure to a mutagenic, genotoxic, carcinogenic chemical agent during pregnancy may result in teratogenicity. In our study, the mutagenic, genotoxic, and carcinogenic effects of levetiracetam, one of the most widely prescribed new-generation antiepileptic drugs, and its metabolites were evaluated using computational methods. Levetiracetam is thought to be effective in controlling seizures during pregnancy and relatively free of teratogenic side effects. However, the FDA has classified levetiracetam as pregnancy category C [6]. In this category, which belongs to the previous FDA classification, animal reproduction studies conducted for drugs have shown an adverse effect on the fetus, and/or there are no adequate and well-controlled studies in humans but the potential benefits may warrant the use of the drug in pregnant women despite the possible risks [7].

The metabolite analysis of levetiracetam was conducted using MetaTox (v.2.0), a computer technology in silico on the PassOnline platform (<http://way2drug.com/passonline/predict.php>). Levetiracetam and its known and predicted metabolites were analyzed by VEGA (v.1.2.3), EPA TEST (v.4.2.1 and 5.1.2), and Danish Q(SAR) for their potential mutagenic, genotoxic, and carcinogenic behavior. Molecular structural alerts for mutagenicity were carried out using an expert rule-based QSAR prediction tool, OECD QSAR Toolbox (v.4.7). The organ-specific carcinogenicity potentials were predicted by ROSC-Pred, which is a web service for rodent organ-specific carcinogenicity prediction. The developmental toxicity potentials were evaluated using VEGA and EPA TEST models.

Based on the available data in the literature, levetiracetam has three known metabolites: (2S)-2-(2-oxopyrrolidin-1-yl)butanoic acid (UCB-L 057) (**M1**), (2S)-2-(4-hydroxy-2-oxopyrrolidin-1-yl)butanamide (**M2**), and 4-{[(1S)-1-carboxypropyl]amino}-3-hydroxybutanoic acid (**M3**) [6]. By the MetaTox analysis, **M2** was also predicted as a metabolite of levetiracetam. The predicted metabolites to be formed by aliphatic hydroxylation are 2-(4-hydroxy-2-oxopyrrolidin-1-yl)butanamide (**M2**), 2-(3-hydroxy-2-oxopyrrolidin-1-yl)butanamide (**M4**), 3-hydroxy-2-(2-oxopyrrolidin-1-yl)butanamide (**M5**), 2-(2-hydroxy-5-oxopyrrolidin-1-yl)butanamide (**M6**), 4-hydroxy-2-(2-oxopyrrolidin-1-yl)butanamide (**M7**), and 2-hydroxy-2-(2-oxopyrrolidin-1-yl)butanamide (**M8**). Our further in silico assessments were conducted on levetiracetam and its eight known/predicted metabolites. VEGA KNN-read-across, EPA TEST consensus models predicted **M7** as mutagenic with a prediction value higher than 0.50. SarPy-IRFMN and KNN-read-across models predicted **M6** and **M8** as mutagenic with low and moderate reliability, respectively. Other metabolites were evaluated as non-mutagenic in tested models. All tested substances were predicted as carcinogenic in the IRFMN-ISSCAN-CGX carcinogenicity model with mostly good reliability. The organs where levetiracetam showed the highest carcinogenicity potential were estimated to be the oral cavity, nasal cavity, and liver in male rats ($Pa=0.596$, 0.540 , and 0.550 , respectively), while the uterus and nasal cavity in female rats ($Pa=0.652$ and 0.456 , respectively). Levetiracetam, **M1**, **M2**, **M4**, **M5**, **M6**, and **M7** were predicted carcinogenic in the hematopoietic system in female mice with a Pa value of 0.419 , 0.421 , 0.498 , 0.530 , 0.495 , 0.411 , and 0.494 , respectively. Levetiracetam, **M1**, **M5**, **M7**, and **M8** were predicted to have carcinogenic potential in oral and nasal cavities in male and female rats. Neither levetiracetam nor its metabolites were predicted to be

active in the chromosomal aberration model. However, **M2**, **M4**, and **M5** were predicted as genotoxic *in vivo* micronucleus (MN) activity model. **M5** was estimated to be active in both *in vitro* and *in vivo* MN models with moderate reliability. H-acceptor-path3-H-acceptor structural alert for mutagenicity has been identified for all substances in the OECD QSAR Toolbox evaluation. The EPA TEST developmental toxicity consensus models estimated levetiracetam and all its metabolites (**M1-M8**) as developmental toxicants (prediction value \geq 0.65).

In conclusion, the mutagenicity predictions for **M2**, **M4**, **M5**, **M6**, **M7**, and **M8** are inconclusive, and there is a potential for mutagenic effects that cannot be excluded. We should also consider the potential for developmental toxicity and carcinogenicity of levetiracetam and all its metabolites (**M1-M8**). It's important to note that these *in silico* data should be confirmed with *in vivo* experiments.

1. WHO (2024) <https://www.who.int/news-room/fact-sheets/detail/epilepsy>
2. Moores, G., et al. (2021) *CMAJ*, 193(32), E1253.
3. Li, Y., Meador, K.J. (2022) *Continuum (Minneapolis Minn)*, 28(1), 34-54.
4. Lavu, A., et al. (2023) *Br. J. Clin. Pharmacol.*, DOI: 10.1111/bcp.15752
5. Tomson, T., Battino, D. (2012) *Lancet Neurol.*, 11, 803-813.
6. KEPPRA® (levetiracetam). https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/021035s078s080%2C021505s021s024lbl.pdf
7. Leek, J.C., Arif, H. (2024) <https://www.ncbi.nlm.nih.gov/books/NBK507858/>

CONSTRUCTING BAYSIAN NETWORKS TO DETERMINE THE RISKS OF DRUG INTERACTIONS BASED ON INSTRUCTIONS

Yu.P. Titov, D.D. Kybrakov, N.V. Kilmishkin

Plekhanov Russian University Of Economics, Moscow, Russia

With the development of modern computing systems, opportunities arise for automatic information analysis, including the use of artificial intelligence systems. Within the framework of such systems, the task of searching for combinations of drugs that are undesirable, from the point of view of the likelihood of negative interactions, is considered. Solving this problem statistically is impossible due to the huge number of possible combinations of 3-5 drugs taken simultaneously, and the development of chemical, biological and other models is complicated by the accuracy of the data obtained, the cost and limitations of the models themselves, and difficulties in creating an explanatory and legal component.

To determine the risks of side effects when using several medicines, it is proposed to use legally accurate instructions for medicines. The sections in the instructions with drug interactions describe only some pairwise drug interactions. When using 3-5 drugs, their pairwise interaction may be insignificant, but complex interaction can lead to a critical effect of the drugs. In this work, based on the analysis (using artificial intelligence) of instruction texts, a Bayesian network (Bayesian graph model) is built that describes all texts of instructions for medicinal products. The top-level vertices define information about the patient, his anthropometric data, known diseases and medications taken. Also at this level are all possible medicines, instructions for which are in the State Register of Medicines. Each vertex is connected by arcs to vertices-mechanisms that determine the effect of the drug and other properties on individual components of the human body. The set of vertices-mechanisms and the direct effect of the drug on the mechanism is determined from the text of the instructions, using the sections of pharmacological action, pharmacokinetics, pharmacodynamics, indications, contraindications, side effects and drug interactions. The interaction of drugs in such a system is also considered through their influence on the same vertex mechanisms. The mechanism vertices are connected by arcs to each other and to the side action vertices. The vertices indicate conditional probabilities (vertices-mechanisms and vertices-side effects) and a priori probabilities (vertices associated with personal characteristics), which are recalculated using formulas from Bayesian networks as a result of noting the use of medications. The final probabilities of side effects take into account all medications, mechanisms and personal characteristics of a person. When a threshold level of certain side effects is reached, the use of appropriate combinations of drugs for a particular patient can be judged to be undesirable or prohibited.

Determining probabilities with sufficient accuracy is an important and difficult task. At this stage of development, it is proposed to “train” since the structure of the graph is close to the structure of a neural network on various medical texts. The values obtained should not contradict the instructions and other medical texts and recommendations. In this case, there may be more than one layer of mechanisms and the probabilities can be recurrent. For such systems, convolutional and recurrent neural networks are used. It is proposed to calculate the initial probability values based on the location of words in the instructions, the words of “markers” and the frequency of occurrence of words in the instructions. In this case, information about pairwise drug interactions and the effect of a drug on a side effect will show a high influence on the corresponding mechanisms representing intermediate nodes in the Bayesian network. It should also be noted that the Bayes network representation allows for the creation of an explanatory component of the interaction, specifying mechanisms, side effects, and probabilities.

The study was supported by the Russian Science Foundation No. 23-75-30012

SMALL-MOLECULE ACTIVATORS FOR GLUCOSE-6-PHOSPHATE DEHYDROGENASE (G6PD) USING MACHINE LEARNING APPROACHES

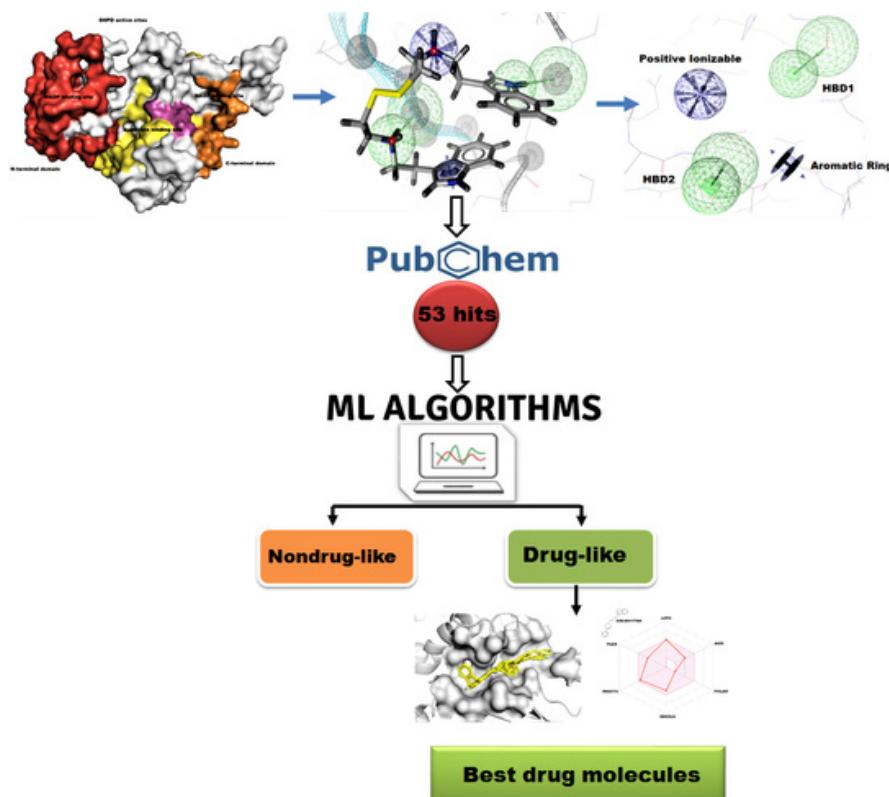
M.S. Saddala¹, N. Patya²

¹*Department of Bioinformatics, UCI-MIND, University of California, Irvine, USA*

²*Department of Computational Biology and Bioinformatics, University of Himachal Pradesh, Himachal Pradesh, India*

Glucose-6-Phosphate Dehydrogenase (G6PD) is a crucial and widely distributed enzyme in the cytoplasm that catalyzes the conversion of glucose-6-phosphate into 6-phosphogluconate, a key step in the pentose phosphate pathway (PPP). This enzymatic activity is pivotal for the production of Nicotinamide Adenosine Dinucleotide Phosphate (NADPH), essential for maintaining the reductive environment of cells by regenerating glutathione. G6PD deficiency, which impairs NADPH production, results in oxidative stress, leading to potential damage to photoreceptors, retinal cells, and the integrity of the blood-retinal barrier. To address the consequences of G6PD deficiency, our study aimed to identify potential G6PD activators through a combination of pharmacophore modeling and machine learning techniques. We initially constructed pharmacophore-based models using the complex of G6PD and the known activator, compound AG1. Subsequent virtual screening processes identified 53 hit molecules that aligned with the core pharmacophore features. To further analyze these hit molecules, we performed comprehensive molecular descriptor calculations, clustering, and Principal Component Analysis (PCA). These steps allowed for a refined classification of the hits. Additionally, multiple statistical machine learning methods were employed to assess the drug-likeness of the identified compounds. This robust analysis categorized the 53 hits into 18 drug-like and 35 nondrug-like compounds.

For the drug-like compounds, we utilized our established cheminformatics pipeline for further evaluation. This included molecular docking studies to assess binding affinities and *in silico* ADMET (absorption, distribution, metabolism, excretion, and toxicity) analysis to predict the pharmacokinetic and safety profiles of these compounds. From these evaluations.



OLEG A. RAEVSKY - SCIENTIST, TEACHER, PERSON

V.Y. Grigorev¹, O.V. Tinkov², V.P. Kazachenko¹

¹Institute of Physiologically Active Compounds at Federal Research Center of Problems of Chemical Physics and Medicinal Chemistry, Russian Academy of Sciences, Chernogolovka, Russia

²Shevchenko Transnistria State University, Tiraspol, Moldova

O.A. Raevsky was born on May 10, 1940 in the city of Rostov-on-Don. In 1957, he graduated from high school and entered the Chemical Faculty of Rostov State University, where he earned a degree in physical chemistry. After graduation in 1962, he qualified as a research chemist. In the period from 1962 to 1979, he worked at the A.E. Arbuzov Institute of Organic and Physical Chemistry, Kazan branch, USSR Academy of Sciences and worked his way up from a laboratory assistant to a senior researcher. In that period, O.A. Raevsky defended his dissertations for the degree of Candidate of Chemical Sciences (1966) and Doctor of Chemical Sciences (1978). He became one of the first chairmen of the Council of Young Scientists in the country. Since 1979, he began working at the Institute of Physiologically Active Compounds of the Russian Academy of Sciences.

Oleg Alekseevich's research interests were diverse and included physical chemistry, spectroscopy, molecular modeling, studying structure-property relationships, designing and computer searching for new physiologically active compounds with specific properties (drugs), developing computer programs, and creating databases. He proposed innovative computational methods, such as assessing the quantitative contribution of hydrogen bonding to the properties and biological activity of chemical compounds, determining the similarity of chemical structures, and using discrete regression models to find structure-activity relationships. Together with his team, O.A. Raevsky developed a series of computer programs, which were registered with Rospatent and were successfully used in molecular design research at various domestic and international universities and companies.

Oleg Alekseevich Raevsky was author of more than 400 scientific publications, including two textbooks for students of chemical universities and two monographs "Properties of Chemical Compounds and Drugs as Functions of Their Structure" (2013) and "Modeling Structure-Property Relationships" (2015). In 1983-1986, Oleg Alekseevich taught an elective course "Designing Biologically Active Compounds" at Mendeleev University of Chemical Technology. In 1988, he was awarded the title of Professor in Physical Chemistry. Under his guidance, 15 doctoral and candidate's dissertations have been successfully defended.

O.A. Raevsky was one of the organizers of the Russian Section of the International Scientific Society QSAR and Molecular Modeling Society, the chairman of which he was elected twice for a five-year term from 1996 to 2006. Since 1994, he was a member of the Board of this international society. From 1993 to 2005, O.A. Raevsky was a member of the editorial board of the international scientific journals QSAR & Combinatorial Science and SAR and QSAR in Environmental Research. He was an active participant in the symposia "Bioinformatics and Computer-Based Drug Design," which were held within the framework of the Russian National Congresses "Man and Medicine." In 2013, the Russian Academy of Natural History awarded O.A. Raevsky the honorary title "Founder of the Scientific School: Computer Molecular Design of Physiologically Active Compounds".

POSTERS

IN SILICO SCREENING OF SOME PYRIDO[2,3-D]PYRIMIDIN AS A NEW INHIBITOR OF TELOMERASE ACTIVITY

B. Addoum^{1*}, M. Chbaib^{2*}, R. Achagar³, M. El Mzibri¹, M. Attaleb¹, I. Chaoui¹, A. Elmakssoudi³, L. Benbace¹

¹*National Center for Energy, Nuclear Sciences and Techniques (CNESTEN), Biology and Medical Research Unit - Life Sciences Division, Rabat, Morocco*

²*Higher Institute of Nursing Professions and Health Techniques, Annex Guelmim, Morocco*

³*Laboratory of Organic Synthesis, Extraction, and Valorization, Department of Chemistry, Faculty of Sciences Ain Chock, Hassan II University, Casablanca, Morocco*

*Equal contribution

Telomerase, an enzyme responsible for maintaining telomere length, plays a crucial role in cellular aging and cancer proliferation. Inhibiting telomerase activity has emerged as a promising strategy for anticancer therapy. In this study, we performed in silico screening of a series of pyrido[2,3-d]pyrimidin derivatives to identify potential telomerase inhibitors. We evaluated the binding affinity and stability of these compounds within the telomerase active site using molecular docking, virtual screening, and molecular dynamics simulations. The results revealed that all the telomerase inhibitor drugs tested got docked with negative binding energy onto the target protein. The 5-phenyl-7-(p-tolyl)pyrido[2,3-d]pyrimidin-4-amine docked into 5CQG active site appeared to have the best inhibitor because it gives energy complex about -6.3 Kcal/mol. Moreover, the molecular interaction studies showed that telomerase structure had multiple active site residues for all studied compounds including Met 769, PRO 770, LYS 721, VAL 70 2, and GLN 767.

Further studies are needed to testify the medicinal use of pyranopyrazoles as a new tool to fight against coronavirus infection. Our findings suggest that pyrido[2,3-d]pyrimidin derivatives could serve as potent telomerase inhibitors, offering a new avenue for developing targeted anticancer therapies. Further experimental validation and optimization of these compounds are warranted to advance their potential as therapeutic agents.

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF HYPOTHETICAL PROTEINS OF LUMPY SKIN DISEASE VIRUS TOWARD IDENTIFICATION OF VACCINE TARGETS

B. Agarwal¹, B. Ganguly²

¹*Mahindra University, Hyderabad, India*

²*Indian Immunologicals Limited, Hyderabad, India*

Lumpy Skin Disease (LSD) is a major scourge of livestock. The disease is caused by a member of *Poxviridae*, having a large genome typical of pox viruses, encoding more than 150 genes of which several protein coding genes are hypothetical in nature without any known function¹. The current work aimed at structural and functional characterization of all hypothetical proteins (HPs) of the LSD virus toward the identification of probable vaccine targets. Briefly, the HP sequences were mined from UniProt² and clustered into representative sequences using CD-HIT³. InterProScan⁴ was used to confirm true HPs based on absence of any known function. Thereafter, the structure was predicted using AlphaFold⁵ and the predicted structure was used for determination of function using VAST⁶ and DeepFRI⁷. Homology to essential host proteins was determined with DEG⁸ BLAST and probable vaccine targets were identified with VaxiJen⁹.

UniProt search yielded 900 HPs of LSD viral origin, which clustered into 110 representative sequences, of which 68 were found to be true HPs. The structural predictions were successfully obtained for all HPs of the LSD virus with reasonable quality. Functional inferences based on the predicted structure allowed further characterization of the HPs. Finally, 05 probable vaccine targets were identified among these HPs. It is expected that our structo-functional characterization of the HPs of the LSD virus will shed greater insights into the pathophysiology of the disease and encourage vaccinological investigations.

1. Yadav, P., Kumar, A., Nath, S.S. et al. (2024) *BMC Genomics*, 25, 196.
2. UniProt Consortium (2008) *Nucleic Acids Research*, 36(Suppl 1), D190.
3. Huang, Y., Niu, B., Gao, Y. et al. (2010) *Bioinformatics*, 26, 680.
4. Jones, P., Binns, D., Chang, H.Y. et al. (2014) *Bioinformatics*, 30, 1236.
5. Jumper, J., Evans, R., Pritzel, A. et al. (2021) *Nature*, 596, 583.
6. Gibrat, J.F., Madej, T., Bryant, S.H. (1996) *Current Opinion in Structural Biology*, 6, 377.
7. Gligorijević, V., Renfrew, P.D., Kosciolak, T. et al. (2021) *Nature Communications*, 12, 3168.
8. Luo, H., Lin, Y., Liu, T. et al. (2021) *Nucleic Acids Research*, 49(D1), D677.
9. Doytchinova, I.A., Flower, D.R. (2007). *BMC Bioinformatics*, 8, 1.

COMPUTATIONAL EXPLORATION OF NATURAL NOVEL HYBRID MOLECULES AS RAF-1 KINASE ANTAGONISTS FOR BREAST CANCER THERAPEUTICS

N. Aggarwal¹, M. Chourasia², B. Bondhopadhyay¹

¹*Immunoncology and Molecular Theragnostics Lab, Centre for Medical Biotechnology, Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida, India*

²*Drug Discovery Lab, Centre for Computational Biology and Bioinformatics, Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida, India*

Raf-1 is one of the three isoforms belonging to the Raf family of serine/threonine kinases. This kinase protein has been found to be highly implicated in breast cancer. Currently, no natural or specific inhibitors exist against Raf-1. To address the shortage of specific inhibitors for Raf-1 in breast cancer, our objective is to develop safe and selective anticancer therapeutics by integrating natural products with existing pan-Raf drugs.

We compiled 31 pan-Raf drugs from various databases and sourced 449,008 molecules from the SuperNatural 3.0 database. Utilizing the Schrodinger Suite 2023-1, we performed homology modeling using the BRAF structure (8C7X) and the Raf-1 (P04049) FASTA sequence. The drug ligands underwent standard precision (SP) docking, while high throughput virtual screening (HTVS) was applied to the natural product library, resulting in 3,701 compounds. Fragmentation of these compounds produced 8,786 natural product fragments and 1,311 drug fragments. This process generated 54,279 unique hybrids, of which 25,000 passed HTVS screening. The top 1,000 hybrids were then subjected to SP docking, followed by ADME predictions. Four hybrids that met drug-like criteria underwent further evaluation using multiparameter optimization (MPO), and the top three hybrids were selected after extra precision (XP) docking.

SP docking of the top 1,000 hybrids produced a best score of -16.375 kcal/mol. Four hybrids – NHP-C01, NHP-C02, NHP-C03, and NHP-C04—met the selection criteria, with MPO scores of 0.68, 0.64, 0.62, and 0.66 out of 1.00, respectively, and XP docking scores ranging from -3.362 to -12.042 kcal/mol. NHP-C01, NHP-C02, and NHP-C03 were chosen for further analysis. In this study, semi-synthetic hybrid drugs, NHP-C01, NHP-C02, and NHP-C03, were identified by combining natural products with pan-Raf drugs. These hybrids demonstrate promising potential for Raf-1 inhibition, based on their interactions, binding energy scores, and predicted ADME profiles. The potential binding of these lead compounds, emphasizes on the need for their further evaluation.

The authors acknowledge Amity University Uttar Pradesh, Noida for providing the facilities for the conduct of the present study. The authors also acknowledge ICMR-DHR (R.12013/21/2023-HR/ E-Office: 8206474).

THE STUDY OF A CLUSTER OF COMPOUNDS WITH ANTI-INFLAMMATORY ACTIVITY IN A SERIES OF SUBSTITUTED AMIDES OF ANTHRANILIC ACID BY MOLECULAR DOCKING ON INTERLEUKIN 2

K.V. Andryukov

Perm state pharmaceutical academy, Perm, Russia

The study of interaction with biological targets responsible for the mechanism of anti-inflammatory activity makes it possible to search for a cluster in which the analyzed compound binds to the enzyme. The biological targets of anti-inflammatory activity - enzymes include protein molecules, for example, cyclooxygenase [1], lipoxygenase and interleukin. The interaction by molecular docking is characterized by the formation of intermolecular hydrogen bonds with amino acid residues of the "active" site of the biological target.

The aim of the work is to carry out molecular modeling of a cluster of compounds with anti-inflammatory activity (PVA) in a series of substituted amides of anthranilic acid by the method of molecular docking according to interleukin 2 (IL-2). The cluster modeling is based on multiple linear regression analysis and spatial (Cartesian) three-dimensional coordinates. By changing the coordinates, a significant "active" section of IL-2 was searched.

Ligand-receptor interactions were modeled using the AutoDock 4.0 program as part of the MGL Tools 1.5.6 software package. During molecular docking, a three-dimensional model of the IL-2 molecule was used, information about which was obtained from the RCSB Protein Data Bank database: PDB ID code: 1M48 [2]. The search for the binding site was carried out by a combination of correlation analysis with regression analysis, the Statistica 10 program.

When carrying out molecular docking of 29 substituted anthranilic acid amides, the results of a study on a biologically active substance (BAS) – a leader with anti-inflammatory activity - acetyl salicylic acid (ASA) were used as a starting point. The structure of the ASK– IL-2 complex with a residue of the amino acid arginine (ARG 38A) allows us to explore several points of interaction within the "active" anti-inflammatory cluster. Four interaction sites in the form of dots were obtained. One of the found points, the most significant according to the correlation model, was selected with the coordinates of the oxygen atom of the carboxyl group interacting with the amino acid residue ARG 38A of the active site: $x = -0,243$; $y = 12,381$; $z = -11,943$; with a binding energy equal to -5,22 kcal / mol. The data of a significant point recorded in a multiple regression model for salicylic acid derivatives [3] were used to select PVA conformations for substituted anthranilic acid amides, according to the minimum deviation from the predicted binding energy. The evaluation of the selection results was carried out using the correlation coefficient of binding energy with experimental values of biological activity (PVA exp.) obtained using the carrageenan edema model [1]. The correlation equation based on the results of the selection turned out to be the following: $PVA\ exp. = -24.85 - 11.69 \times Be$ correlated with PVA ($R = 0,730$; $F = 30,89$; $p < 0,00001$; $N = 29$).

The results of the study make it possible to identify the PVA descriptor for constructing forecasting and search models: "structure – anti-inflammatory activity" and to build a cluster of the "active" IL-2 site. The use of the obtained correlation and regression model of the cluster makes it possible to ensure the constancy of the modeling conditions, the interaction of the analyzed compounds with IL-2 during molecular docking.

1. Andryukov, K., Korkodinova, L. (2018) *Pharm. Chem. J.*, 52(5), 411-414.
2. Arkin, M.A., Randal, M., DeLano, W.L., et al. (2003) *Proc. Natl. Acad. Sci. USA*, 100, 1603-1608.
3. Andryukov, K. (2024) *Cardiovascular Therapy and Prevention*, 23 (S6), 155.

IN SILICO DESIGN OF NATURAL HYBRID DRUGS AGAINST PI3K-ALPHA FOR BREAST CANCER THERAPY

N. Aggarwal, S. Sen, B. Bondhopadhyay

Immunoncology and Molecular Theragnostic Lab, Centre for Medical Biotechnology, Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida, India

PI3K-alpha is one of the most influential isoforms of PI3K protein family which promotes cell proliferation and metastasis in almost half of all breast cancer cases. Despite its critical role, there is a scarcity of synthetic and natural therapeutic options targeting its inhibition. Thus, the study herein, aims to address this gap by developing natural hybrid antagonists targeting PI3K-alpha through innovative fragment-based and combinatorial *in silico* approaches.

Our methodology involved sourcing 25 pan-PI3K and PI3K-alpha targeting drugs, along with natural compounds from the COCONUT database. The Schrodinger Suite 2023-1 was used to perform all *in silico* experiments. We performed high throughput virtual screening (HTVS), followed by standard (SP) and extra precision (XP) docking of the parent dataset. We fragmented these drugs and natural compounds into Murcko scaffolds and heterogenous fragments, subsequently hybridizing them into two categories: natural fragments with drug Murcko scaffolds (Category 1) and drugs fragments with natural compound Murcko scaffolds (Category 2) using the BREED algorithm. These hybrids underwent further HTVS, SP, and XP docking, induced fit docking, and ADME prediction.

The results were promising, with Category 1 and Category 2 hybrids achieving top docking scores of -13.354 kcal/mol and -12.670 kcal/mol, respectively. These natural hybrids outperformed their parent drugs and natural compounds in binding affinity, pharmacological profiles, and adherence to Lipinski's rule of five. Remarkably, hybrids NH-01 and NH-06 exhibited superior binding and pharmacological behavior compared to their parent molecules.

Our findings suggest that natural-drug hybrid antagonists offer a novel and potent therapeutic avenue for targeting PI3K-alpha in breast cancer treatment. The exceptional performance of these hybrids underscores their potential, necessitating further testing in *in-vitro* and *in-vivo* systems to validate their efficacy as breast cancer therapeutics.

The authors acknowledge Amity University Uttar Pradesh, Noida for providing the facilities for the conduct of the present study. The authors also acknowledge ICMR-DHR (R.12013/21/2023-HR/ E-Office: 8206474).

CHARACTERIZING CONFORMATIONAL STATES IN GPCR STRUCTURES USING MACHINE LEARNING

I. Buyanov, P. Popov

iMolecule, Skolkovo Institute of Science and Technology, Moscow, 121205, Russia

The rapid breakthroughs in machine learning and structural biology resulted in a novel phase of structural analysis of G protein-coupled receptors (GPCRs), which are crucial in signal transduction and major targets in drug discovery [1]. GPCR signaling pathways are closely related with GPCR conformational states [2]. In this study, we introduced a novel machine learning approach (STAGS) [3] to annotate the conformational states of GPCRs, which are represented as high-dimensional feature vectors, capturing the structure-based information about amino acid residue pairs involved in the activation pathway. Utilizing molecular dynamics simulations, we generated a dataset of GPCR conformations in both inactive and active states. The machine learning model, trained on this dataset, distinguishes between inactive-like and active-like conformations with close to perfect accuracy. The interpretability of our model facilitates large-scale analysis of molecular dynamics trajectories, providing deeper insights into GPCR conformational dynamics. The STAGS workflow (Fig. 1) includes 1) dataset collection, 2) extracting pair-wise distance features from GPCR structures to create high-dimensional vectors, 3) training machine learning model to classify conformational states, and 4) applying the trained model to analyze conformational ensembles obtained, for example, with molecular dynamics simulations. The developed method can be used in analysis of GPCR conformations as well as in structure-based drug discovery targeting GPCRs.

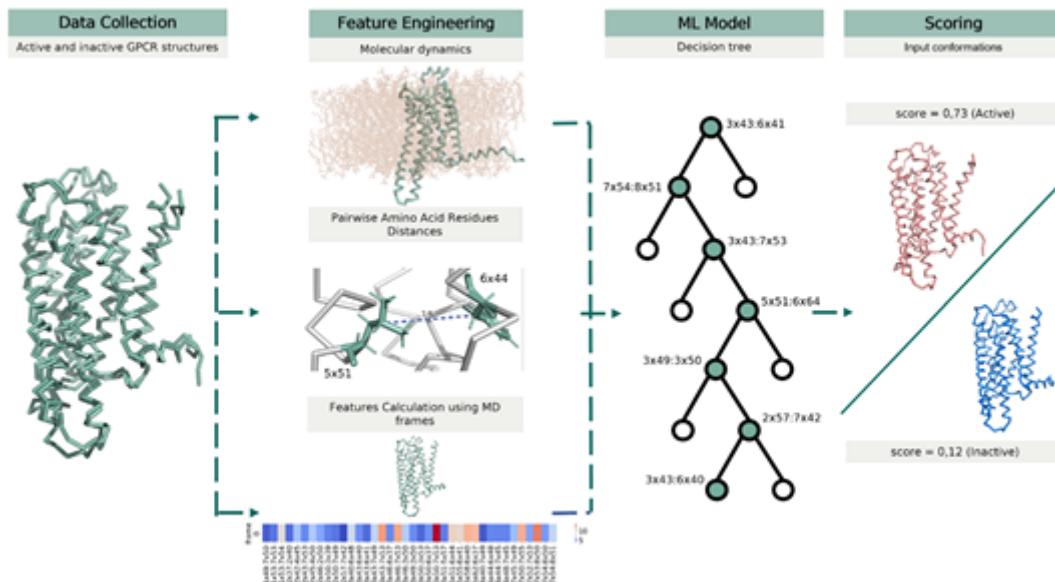


Fig. 1. Overview of the STAGS workflow comprising the dataset collection, feature engineering, machine learning model training and application. Data Collection represents a composition of a dataset of active and inactive conformations of class A GPCRs. Feature Engineering illustrates a calculation of the descriptor corresponding to the pair-wise distances between amino acid residues from the molecular dynamics simulations. ML Model illustrates a decision tree from the random forest. Scoring demonstrates the calculated scores for active- and inactive-like conformations of a GPCR.

This study was supported by Russian Scientific Foundation project no. 22-74-10098.

1. Insel, P. A. et al. (2019) *Trends Pharmacol. Sci.*, 40, 378–387.
2. Sriram, K., Insel, P. A. (2018) *Mol. Pharmacol.*, 93, 251–258.
3. Buyanov, I., Popov, P. (2024) *Sci. Rep.*, 14, 1098.

MOLECULAR DYNAMICS DESIGN AND SYNTHESIS OF BENZOIC ACID DERIVATIVES AS TRPC6 CHANNEL AGONISTS: POTENTIAL DRUGS FOR THE TREATMENT OF AUTISM SPECTRUM DISORDERS

D. Cabezas¹, G. Sabadini¹, J. Romero², M. Mellado³, A. Palacios¹, I. García¹, J. Mella¹

¹Facultad de Ciencias y Facultad de Odontología, Universidad de Valparaíso, Chile

²Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Chile

³Facultad de Medicina y Ciencias de la Salud, Universidad Central de Chile, Chile

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by persistent challenges in social interaction, communication, and repetitive patterns of behavior. Around 75 million people in the world have autism spectrum disorder, that is 1% of the world's population. The current rate of children with autism is 1 in 36, and there is no pharmacological cure. Most therapy is aimed at treating symptoms such as aggression, hyperactivity, and sleep disorders associated with autism. Some potential targets associated with autism are TRPC6, mGluR5, PTPRD, 5-HT and GABA-R. TRPC6 is a member of the transient receptor potential (TRP) channel family, which regulates calcium influx in various cells and tissues. Aberrant TRPC6 channel activity has been associated with neurological disorders, including ASD. Therefore, modulating TRPC6 channel function presents a potential avenue for therapeutic intervention in ASD. In recent years, several agonists and modulators of the TRPC6 channel have been identified, demonstrating the potential to regulate its activity and restore normal calcium signaling. In addition, a recent study of our group has shown that hyperforin, a TRPC6 agonist, is capable of reversing autistic behavior in fly models of this pathology [1]. However, there is still a need for the development of novel compounds with improved efficacy and selectivity.

In order to find new and better molecules (with affinity in the nM range), we carried out a virtual screening of 202 molecules from our research group. We found that a benzimidazole compound exhibited a good fit in the agonist groove of TRPC6 and was able to establish a hydrogen bond with Trp680, a key amino acid in agonism, but was unable to establish relevant interactions with other important amino acids such as Lys676 and Asn702 (XPGScore = -9.246 kcal/mol). To this end, we measured the distances to these residues and determined that the insertion of an OH and a COOH group were promising. After this change, the scoring value improved substantially. Additionally, we found that switching from benzimidazole to benzothiophene further improved the scoring (Hit compound BT11, XPGScore= -11.452 kcal/mol). In this study, we focus on the compound BT11 and its derivatives as potential agonists of the TRPC6 channel. BT11 has shown promising results in molecular dynamics simulations (100 ns), exhibiting favorable interactions with key residues in the agonist binding site (Trp680, Phe675, Phe679, Lys676 and Lys698). The compound BT11 demonstrates pore expansion in the channel similar to the reference ligand GSK1702934A (XPGScore = -9.235 kcal/mol), with superior and statistically significant results ($R^2= 0.8322$). We aim to synthesize and evaluate a series of BT11 derivatives to assess their agonistic activity on the TRPC6 channel. Additionally, we will investigate the effects of BT11 and selected derivatives in animal models of ASD, specifically using fruit flies (*Drosophila melanogaster*). Figure 1 represents a summary of the present work.

This study was supported by project Fondecyt Grant N°=1240289 from Anid, Chile.

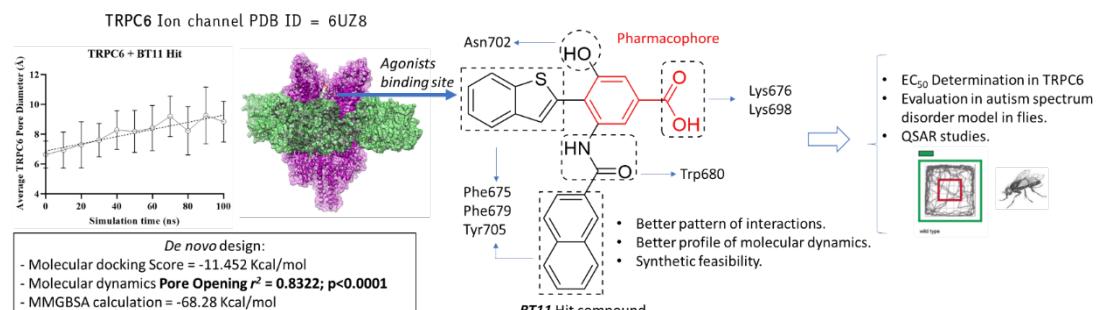


Figure 1. Graphical abstract.

1. Palacios-Muñoz, A. et al. (2022) *Molecular Psychiatry*, 27, 3328–3342.

PREDICTING FLUORESCENCE/SINGLET OXYGEN GENERATION QUANTUM YIELD RATIO FOR BODIPYs USING QSPR AND MACHINE LEARNING

P.P. Chebotaev¹, A.A. Buglak¹, A. Sheehan², M.A. Filatov²

¹*Faculty of Physics, Saint-Petersburg State University, Saint-Petersburg, Russia*

²*School of Chemical and Biopharmaceutical Sciences, Technological University Dublin, Dublin, Ireland*

Photodynamic regulation of biological processes was first described in 1900 and since then this phenomenon has been actively investigated. Predicting the ratio of fluorescence to singlet oxygen generation for functional dyes is crucial for emerging 21st century theranostics methods that combine fluorescence imaging and photodynamic therapy (PDT). The development of new functional dyes for theranostics is often costly and time-consuming due to chemical modification and post-synthetic screening of large libraries of compounds. Our work aims to adapt machine learning algorithms for the simultaneous prediction of fluorescence and photosensitizing ability of heavy-atom-free BODIPY compounds. To date, there have been no attempts to construct QSPR models for predicting the quantum yield of singlet oxygen generation except for the works of Buglak et al. [1]. We analysed the ratio between fluorescence quantum yield (Φ_{Fl}) and singlet oxygen generation quantum yield (Φ_{Δ}) for over 70 BODIPY structures in polar (acetonitrile) and non-polar (toluene) solvents, which respectively mimic hydrophilic and hydrophobic cell environments. QSPR models were developed from more than 5000 calculated molecular descriptors, including quantum chemical descriptors. We applied multiple linear regression (MLR), support vector regression (SVR), and random forest regression (RFR) methods for model building and optimization. The analysis revealed the significance of various descriptors, with those related to the electronic structure, polarizability, ionization potential, and topological features playing crucial role. Notably, descriptors related to 2D atom pairs, particularly the arrangement of carbon and nitrogen atoms, emerged as highly influential for compounds in acetonitrile. High statistical parameters of the models demonstrated their accuracy in predicting the $\Phi_{\text{Fl}}/\Phi_{\Delta}$ ratio, with the RFR model performing best for compounds in toluene and the SVR model for compounds in acetonitrile. Also, all models show sufficient statistical parameters ($R^2 = 0.73 - 0.91$) for both media and they could be used for predicting $\Phi_{\text{Fl}}/\Phi_{\Delta}$.

Our findings demonstrate the applicability of the QSPR methodology for studying the $\Phi_{\text{Fl}}/\Phi_{\Delta}$ ratio, providing a valuable tool for pre-synthetic screening of new functional dyes. These predictive models offer a simple and effective approach to the search for novel photosensitizers, replacing the need for random synthesis of new molecular libraries. Furthermore, they can guide the synthesis of dyes with a desired $\Phi_{\text{Fl}}/\Phi_{\Delta}$ ratio in specific environments, such as solvents of varying polarity. Such predictive models not only hold promise for photodynamic therapy but also have implications in fine chemical synthesis and the development of new theranostic drugs. Also, we proved that QSPR machine learning methods can be useful for predicting key photochemical parameters for BODIPY photosensitizers.

1. Buglak, A.A., Charisiadis, A., Sheehan, A., et al. (2021) *Chem. – A Eur. J.* 27, 9934–9947.

SYNTHESIS AND ACTIVITY OF 3,4-DIHYDROPYRIMIDIN-2-ONES(THIONES) USING VARIOUS ORGANIC CATALYSTS

N. Chiobanu, F. Macaev

ICh USM, Chisinau, Moldova

Recently, the number of publications on the chemistry of 3,4-dihydropyrimidin-2-ones (thiones) obtained by three-component condensation according to the Biginelli reaction has increased significantly. This is due not only to the preparative availability of 3,4-dihydropyrimidin-2-ones (thiones), but also to their manifestation of a wide range of pharmacological activity: analgesic, antibacterial, antihypertensive, etc., which makes further searches among them very promising [1]. One of the important products of the three-component reaction under the Biginelli reaction conditions is Monastrol (ethyl-6-methyl-4-(3-hydroxyphenyl)2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxyl)

Monastrol plays an important role in human activity and is a bioactive dihydropyrimidine derivative of primary importance [2]. With the development of biochemistry, when the chemical components of organic structures began to be deciphered, it became clear how huge a role pyrimidines play in human life, because The pyrimidine fragment is part of many vital organic compounds, which has demonstrated a completely new mechanism of anti-cancer action due to its specific effect on cell division (mitosis).

The biological role of Monastrol and its derivatives has led to significant interest in its synthesis and is a 3-component one-pot synthesis based on the interaction of acetoacetic ester, thiourea and 3-hydroxybenzaldehyde, which avoids waste from multi-step purification and the formation of residues. In addition, adducts include almost all the atoms from the reactants in their structures, the so-called atom efficiency. Water is a common by-product. The synthesis is catalyzed by various inorganic acids, organic compounds or under microwave irradiation. The yield and purity of the product are achieved by balancing the appropriate conditions: temperature, choice of catalyst and solvent.

Catalysis plays a fundamental role in the synthesis of 3,4-dihydropyrimidin-2-ones(thiones). To increase the yield, reaction time, selectivity of environmental pollution, environmental safety, waste and costs in synthesis, we previously synthesized a series of imidazole-based ionic liquids with various inorganic anions [1]. From the point of view of “green chemistry”, which is as waste-free and environmentally friendly as possible, it is of interest to use β -cyclodextrin and low-esterified pectin as catalysts. This allows us to fully comply with modern environmental requirements while obtaining a sufficiently high yield of the final product.

Obtaining Oxymonastrol (ethyl-6-methyl-4-(3-hydroxyphenyl)-2-oxo-1,2,3,4-tetrahydropyrimidine-5-carboxyl) under identical conditions is more efficient: synthesis proceeds faster in time and yield the final product is higher [2].

The advantages of the proposed method are: the availability of the reagents used, the simplicity of the synthesis method and the method of purifying the final product by crystallization, compliance with the linear relationship of the theoretical principles of green chemistry with practical experience in application, features that allow you to get as close as possible to eco-friendly catalytic conditions.

The structures of the compounds are confirmed by elemental analysis data, IR and NMR spectral data.

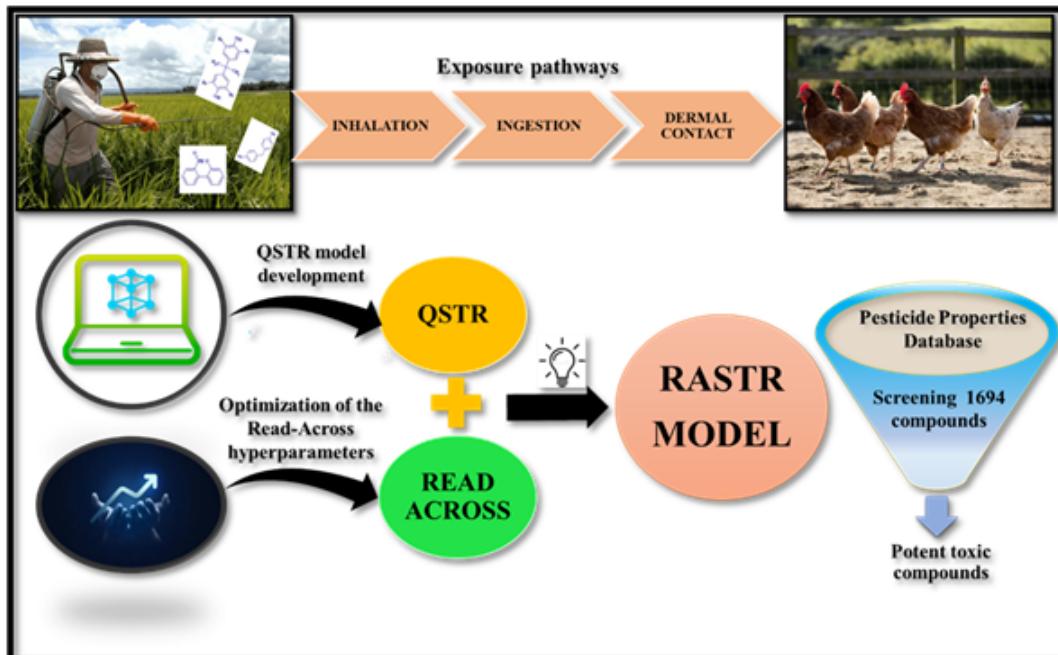
1. Kurbanova, M. (2010) *Org. Chem.*, 46, 599.
2. Macaev, F. (2013) *SPRINGER Science&Business media*, 313-331.

CHEMOMETRICS-DRIVEN PREDICTION AND PRIORITIZATION OF DIVERSE PESTICIDES ON CHICKENS FOR ADDRESSING HAZARDOUS EFFECTS ON PUBLIC HEALTH

S. Das, P.K. Ojha

*Drug Discovery and Development Laboratory, Department of Pharmaceutical Technology,
Jadavpur University, Kolkata, India*

The extensive use of various pesticides in the agriculture field badly affects both chickens and humans, primarily through residues in food products and environmental exposure. This study offers the first quantitative structure-toxicity relationship (QSTR) and quantitative read-across-structure toxicity relationship (q-RASTR) models encompassing the LOEL and NOEL endpoints for acute toxicity in chicken, a widely consumed protein. The study's significance lies in the direct link between chemical toxicity in chicken, human intake, and environmental damage. Both the QSTR and the similarity-based read-across algorithms are applied concurrently to improve the predictability of the models. The q-RASTR model was generated by combining read-across derived similarity and error-based parameters, alongside structural and physicochemical descriptors. Machine Learning approaches (SVM and RR) were also employed with the optimization of relevant hyperparameters based on the cross-validation approach, and the final test set prediction results were compared. The PLS q-RASTR models for NOEL and LOEL endpoints showed good statistical performance, as traced from the external validation metrics Q^2_{F1} : 0.762-0.844; Q^2_{F2} : 0.759-0.831 and MAEtest: 0.195-0.214. From this study, it was found that lipophilicity and electronegativity are responsible for the toxicity of pesticides towards chickens. On the other hand, polarity, hydrophilicity, and large numerical value of SE (LK) & SD similarity (GK) descriptors will reduce the toxicity of pesticides towards chickens. The developed models were further used to screen the Pesticide Properties DataBase (PPDB) for potential toxicants in chickens. Thus, established models can address eco-toxicological data gaps and development of novel and safe eco-friendly pesticides.



STUDYING THE PREDICTIVE AND IN VITRO ACTIVITY OF 5-ARYLIDENTIOBARBITURATES AND THEIR DERIVATIVES

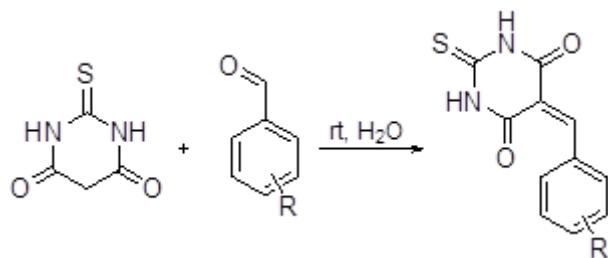
A.S. Skrylkova², D.M. Egorov^{1,2}, I.L. Kuzikova^{1,2}, Z.A. Zhakovskaya^{1,2}

¹St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia

²St. Petersburg State Institute of Technology, St. Petersburg, Russia

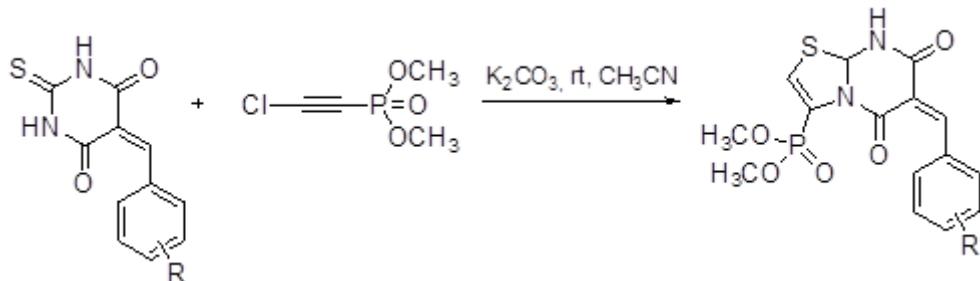
Derivatives of barbituric acid are promising objects for their modification in order to obtain new compounds with biological activity. The activity prediction of potential compounds was performed using the PASS Online resource. Based on the prediction, a number of compounds were selected for synthesis and study of biological activity *in vitro*.

The synthesis of target structures was carried out by the Knoevenagel condensation using a modified method.



The antimicrobial properties of the resulting 5-arylidene-thiobarbiturates were studied. The forecast predicted moderate activity of the target structures against a wide range of bacteria and fungi, but research on *Staphylococcus aureus* 209P, *Staphylococcus aureus* ATCC6538, *Candida utilis* LIA-01, *Pseudomonas aeruginosa* 0387 in most cases showed a negative result.

Of particular interest is the study of the possibility of predicting the biological activity of complex molecules using the PASS Online resource. For this purpose, the previously obtained 5-arylidene-thiobarbiturates were modified.



Despite the fact that the forecast did not show pronounced antifungal and antibacterial activities, several derivatives showed *in vitro* antibacterial activity against the bacteria *Staphylococcus aureus* 209P, *Staphylococcus aureus* ATCC6538.

This work was supported by the Russian Science Foundation (RSF), Project no. 23-13-00224.

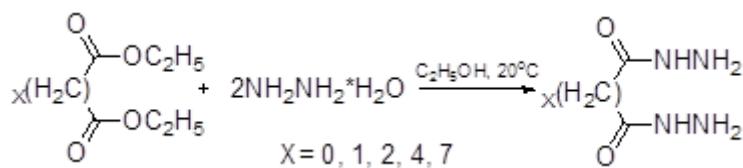
PREDICTION AND STUDY OF ANTIBACTERIAL ACTIVITY OF BISHYDRAZIDES

E.A. Gerasimova^{1,2}, D.M. Egorov^{1,2}, I.L. Kuzikova¹, Z.A. Zhakovskaya¹

¹St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia

²St. Petersburg State Institute of Technology, St. Petersburg, Russia

Organic acids play an important role in the body of living organisms, including humans. It is of interest to study the properties of derivatives of carboxylic acids, in the case of this work - hydrazides of dibasic acids. Prediction of the biological activity of the studied substances using the PASS Online service showed high potential activity as an antibacterial drug. The forecast shows moderate probabilities of activity Pa reaching values of 0.591, with Pi 0.114.



However, in vitro studies of antifungal (*Candida utilis* LIA-01) and antibacterial activities (*Staphylococcus aureus* 209P, *Staphylococcus aureus* ATCC6538, *Pseudomonas aeruginosa* 0387) showed low efficiency of the studied objects; only in the case of nonanedicarboxylic acid hydrazide was activity confirmed against *Staphylococcus aureus* 209P, *Staphylococcus aureus* ATCC6538.

This work was supported by the Russian Science Foundation (RSF), Project no. 23-13-00224.

LIGAND-PROTEIN BINDING SITE ANNOTATION USING GRAPH NEURAL NETWORKS

A. Ereshchenko, S. Evteev, A. Vyacheslavov, D. Adjugim, A. Malyshev, Ya. Ivanenkov

The Federal State Unitary Enterprise Dukhov Automatics Research Institute, Moscow, Russia

The use of machine learning (ML) algorithms has been increasingly prevalent in the field of drug discovery. One application of ML-based methods is structure-based drug design. In particular, multiple ML-based models have been developed for identifying ligand binding sites, but the number of ML solutions for binding site annotation remains limited.

Here we present a graph neural network (GNN) based approach for annotating ligand-protein binding sites using 3D protein structural information. The aim of the annotation is to define which atom types (H-bond donors, H-bond acceptors, hydrophobic) are more suitable in a specified macromolecular environment. Our model represents data as a graph containing a point in a given space and the surrounding protein heavy atoms within a fixed radius. The nodes, representing protein atoms, are characterized by chemical features such as atom name, residue name, residue type, etc. The graph edges are labeled with interatomic distances. The graph features pass through a multilayer GNN using an attention-based transformer operator [1], a global pooling layer, and a sequence of linear layers that provide the final prediction in the range [0,1]. A separate neural network was trained for each atom type using a dataset of more than 8000 ligand-protein complexes from RCSB PDB. To annotate a point in space (which can be an arbitrary point or an atom of a molecule), all scores given by the models are balanced using exponential coefficients to ensure more accurate classification.

To evaluate the performance of our models, we compared it with existing solutions using the Astex Diverse Set [2]. Our method showed comparable or superior accuracy in predicting hydrogen bond acceptor, hydrogen bond donor and hydrophobic atoms compared with energy-based methods [3] and [4]. Furthermore, utilizing the scores calculated by our approach as features allowed us to train high-precision models for the purposes of binding site classification and small molecule affinity prediction, showcasing a wide range of applicability of the developed technology in the field of drug discovery.

1. Shi, Y., et al. (2020) *arXiv*, 2009.03509
2. Hartshorn, M.J., et al. (2007) *J. Med. Chem.*, 50, 726–741.
3. Ravindranath, P.A., et al. (2016) *Bioinformatics*, 32(20), 3142–3149.
4. Harris, R., et al. (2008) *Proteins*, 70, 1506–1517

APPLICATION OF SAR METHODOLOGY TO IDENTIFY FRESHWATER ALLELOCHEMICALS WITH ANTI-CYANOBACTERIAL EFFECTS AGAINST PLANKTONIC CYANOBACTERIA

E.A. Kurashov¹, A.G. Rusanov², E.V. Fedorova³

¹ Papanin Institute for Biology of Inland Waters of the Russian Academy of Sciences, Borok, Yaroslavl oblast, Russia

² St. Petersburg Research Center RAS, Institute of Limnology of the Russian Academy of Sciences, Petersburg, Russia

³ Quantori, Global Headquarters Cambridge, Massachusetts, USA

Controlling harmful cyanobacterial “blooms” through developing a new generation of algaecides based on allelochemical substances is a challenge facing modern aquatic ecology and biotechnology. The present article is devoted to the use of the SAR (Structure-Activity-Relationship) methodology to identify allelochemicals from aquatic macrophytes (floating-leaved *Nuphar lutea* (L.) Sm. and several species of submerged macrophytes: *Ceratophyllum demersum* L., *Myriophyllum spicatum* L., *Elodea canadensis* Michx, and species of the genus *Potamogeton*) effective against planktonic cyanobacteria. Identification and detection of compounds were performed using gas chromatography-mass spectrometry. In recent years, the use of SAR to identify biological activities for natural compounds has become increasingly widespread. One example is the identification of natural antioxidants from marine algae that could be developed for further industrial applications [1]. We applied the PASS (Prediction of Activity Spectra for Substances) computer program to predict biological activity spectra of the major components of macrophyte metabolomes and discover their ecological potential against cyanobacteria [2]. A study of the biological activities of major low-molecular weight organic compounds showed that monocarboxylic acids, gallic acid, cis6-octadecenoic acid, cis9-octadecenoic acid, palmitoleic acid, linolenic acid, and 9-cis12-cis-linoleic acid are the most promising compounds for the experimental verification and creation of nature-like algaecides of a new generation [3]. PASS predictions were successfully compared to the available information on the biological activity of those compounds and confirmed experimentally.

The present study shows that some organic acids significantly inhibit the growth of *Synechocystis aquatilis* Sauvageau and *Aphanizomenon flos-aquae* Ralfs ex Bornet and Flahault and can be used as algaecides for suppression of cyanobacteria. The inhibitory effect of the combined mixture of these allelochemicals is stronger than the effect of each individual component, suggesting that there are can be various mechanisms of cyanobacterial growth inhibition.

The study was funded by Grant of the Russian Science Foundation no. 22–24–00658, <https://rscf.ru/project/22-24-00658>. We also express gratitude to the Resource Center «Cultivation of Microorganisms» at St. Petersburg State University for providing the culture of *Synechocystis aquatilis* for experiments.

1. Kelman, D., Posner, E.K., McDermid, K.J., Tabandera, N.K., Wright, P.R., Wright, A.D. (2012) *Marine Drugs*, 10(2), 403–416
2. Filimonov, D.A., Druzhilovskiy, D.S., Lagunin, A.A., Gloriozova, T.A., Rudik, A.V., Dmitriev, A.V., Pogodin, P.V., Poroikov, V.V. (2018) *Biomedical Chemistry: Research and Methods*, 1(1), e00004
3. Kurashov, E.A., Fedorova, E.V., Krylova, J.V., Kapustina, L.L., Mitrukova, G.G., Protopopova, E.V. *J. Sib. Fed. Univ. Biol.*, 2023, 16(2), 232–251.

ANALYSIS OF ENSEMBLE DOCKING APPLICABILITY TO DIVERSE HOMOLOGY MODELS OF ORTHOFLAVIVIRUS NS1 PROTEIN

A.D. Fomina^{1,2}, V.A. Palyulin², D.I. Osolodkin¹

¹FSASI “Chumakov FSC R&D IBP RAS”, Moscow, Russia

²Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

The genus *Orthoflavivirus* includes enveloped viruses transmitted by ticks and mosquitoes. These viruses cause dangerous diseases such as dengue fever, Zika fever, West Nile fever, and tick-borne encephalitis. By 2024, there are no approved direct-acting antiviral drugs against orthoflavivirus infections. Vaccines have been developed to prevent diseases caused by some orthoflaviruses, but they cannot be used to treat diseases after onset, do not provide complete protection against viral infection, and their development and widespread use are complicated by antibody-dependent enhancement of infection, which is observed for some flaviviruses (dengue virus, Zika virus).

Non-structural protein NS1 plays an important role in the reproduction cycle of orthoflaviruses by stabilising the replicative complex. Moreover, in the acute form of the disease it causes hyperpermeability of blood vessels due to interaction with endothelial cells and destruction of the glycocalyx. Small molecules that interact with the NS1 protein may be used as lead compounds for the discovery of anti-orthoflavivirus drugs. However, complexes of the orthoflavivirus NS1 protein with small molecules have not been described, complicating the application of structure-based molecule design. Nevertheless, the structure of the apo form of the NS1 protein from mosquito-borne orthoflaviruses was studied using X-ray crystallography and cryo-electron microscopy, resulting in seven full-length structures. These structures can be used as templates for homology modelling for other orthoflavivirus NS1 proteins, thus producing enough structural data to employ ensemble docking.

Recently we described a method of systematic ensemble docking [1], which uses all available structural data for docking-based virtual screening of small molecules. Though the improvement in accuracy due to the usage of multiple structures of the same protein has been shown multiple times, the question whether using homologous proteins from organisms of different species may help to find potential broad-spectrum inhibitors remains open. We have performed homology modelling based on the sequences of epidemiologically significant orthoflaviruses. Since there are no available crystal structures of NS1 proteins from tick-borne orthoflaviruses, the available structures deriving from mosquito-borne orthoflaviruses were used, though the amino acid sequences of these groups have inter-group similarity between 34 and 41 percent. The systematic selection process by the diversity of the atomic coordinates was perplexed by the randomness of coordinate generation in the flexible parts of NS1 protein structure. That is why we decided to determine whether the ensemble could be selected based on docking results.

The models and experimentally determined structures were used for docking of 5000 diverse drug-like molecules from ZINC15. The results were ranked and the sum of ranks distributions and rank-order correlation were analysed to check discrimination ability between virtual screening results and randomly generated data. Nevertheless, the applicability of homology modelling of diverse closely related proteins as the structures source for ensemble docking remains limited.

1. Fomina, A.D., Uvarova, V.I., Kozlovskaya, L.I., et al. (2024). *Molecular Informatics*, 43, e202300279.

**IN-SILICO SCREENING OF PHYTOCHEMICAL CONSTITUENTS FROM
GLOBIMETULA OREOPHILA FOR DRUG-LIKE COMPOUNDS USING CRYSTAL
STRUCTURE OF FALCIPAIN I AND II:
HOPE FOR NEW ANTIMALARIAL DRUGS**

**G. Dauda^{1,2}, H.A. Bila², B. Bashar³, M. Abdullahi², A.N. Hamza², M.S. Yahaya², G.M. Muhammad⁴,
I.A. Musa¹, M.M. Aliyu², H.S. Hassan Halimatu²**

¹*Department of Pharmaceutical and Medicinal Chemistry, University of Abuja, Nigeria*

²*Department of Pharmaceutical and Medicinal Chemistry, Ahmadu Bello University, Zaria, Nigeria*

³*National Agency for Food and Drug Administration and Control, Gusau, Nigeria*

⁴*Department of Pharmacology and Therapeutics, Ahmadu Bello University, Zaria, Nigeria*

Malaria continues to be one of the most devastating global health problems due to the high morbidity and mortality it causes in endemic regions. The search for new antimalarial targets is vital because of the increasing prevalence of drug resistance in malaria parasites. Malarial proteases constitute promising therapeutic targets as they play important roles in the parasite life cycle. Falcipain-2 and falcipain-3, which belong to the cysteine proteases, mediate massive degradation of host hemoglobin to release amino acids for parasite nutrition. This study aims to evaluate the inhibitory activity of phytochemical constituents from *Globimetula oreophila* leaves extract as malaria-associated protein using an *in silico* docking study. The prospects of inhibition of *Plasmodium falciparum* vital enzyme; a cysteine protease (falcipain-2 and 3) of quercetin (DG1), dihydrostilbene (DG2), 4'-methoxy-isoliquiritigenin (DG3), stigmasterol (DG4) and quercetin (DG5) isolated from the leave of *Globimetula oreophila* were investigated using *in silico* studies.

The result obtained from the molecular docking analysis demonstrated that the isolated compounds were found to interact with the residues at the active site and other sub-unites regulating the specificity of the proteases. The compounds have a better binding affinity within the binding pockets of falcipain-2 and falcipain-3, which is evident that the binding process was principally favored by hydrogen bonds, van der Waals, and other hydrophobic interactions. Regarding the binding interactions, the most contributing features of the ligands for receptor interactions are the carbonyl group, hydroxyl group, methoxy group, and the pentacyclic structure. Compound DG1 had a lower binding energy of -7.8 Kcal/mol and thus, better affinity within the active site with falcipain-2. Compounds DG1 and DG5 had the lowest binding energy of -8.2 Kcal/mol each with falcipain-3. Conclusion: The inhibition of these enzymes by the isolated compounds could offer therapeutic benefits against *Plasmodium falciparum* and as a good candidate for new sources in the development of antimalarial drugs.

MOLECULAR DYNAMIC SIMULATION OF SOME N-N-DISUBSTITUTED PIPERAZINE DERIVATIVES EXHIBITING ANTICHOLINESTERASE ACTIVITY

V.S. Ghamaryan, L.S. Hunanyan

*Russian-Armenian University, Institute of Biomedicine & Pharmacy, Laboratory of Structural Bioinformatics,
Yerevan, Armenia*

Acetylcholinesterase (AChE) inhibitors dominate the treatment of Alzheimer's disease (AD). AChE prevails in the healthy brain, while it is believed that butyrylcholinesterase (BuChE) plays an insignificant role in the regulation of acetylcholine (ACh) levels in the brain. However, BuChE activity progressively increases in patients with AD while AChE activity remains unchanged or decreases. Thus, both enzymes are legitimate therapeutic targets for alleviating the cholinergic deficit that is considered responsible for the decline in cognitive, behavioral, and global functions characteristic of AD.[1,2] Studies have shown that some piperazine derivatives can play an important role in inhibiting AChE and BuChE through the regulation of ACh [3]. In this work we demonstrate the effect of some piperazine derivatives selected earlier [4] on their possible action as cholinesterase inhibitors using molecular dynamics (MD) methods. Ligands for molecular dynamics were selected based on virtual screening, molecular docking, and dynamics with the TRPC6 structure, which is one of key protein in the calcium hypothesis of AD pathogenesis. MD simulations were conducted on AChE and BuChE based on crystal structure details obtained from the <https://www.rcsb.org/> (PDB codes: 4pqe & 7q1m) refined with a resolution of 2.90 Å and 2.79 Å, respectively. Simulations were performed on 28 processors of a Xeon 2.60 GHz cluster running Linux using GROMACS V.2023.1 software.

The MD results indicate that all the studied ligands have an affinity for the investigated proteins and show stable complex formation throughout the simulation, except for Str47, which does not interact with the AChE but interacts with the BuChE, making it of interest as a highly conservative BuChE inhibitor, and Str204 and Str210, which dissociate from the binding site and exit the active site pocket of AChE during the simulation. The molecular mechanism of interaction with the studied proteins was also investigated, where during complex formation, Str68 and the control compound have the same interaction mechanism, binding with the amino acid residues of the peripheral anionic site and forming hydrogen bonds with Tyr 72, Gly 342, and Trp 286, leading to the occlusion of the active site pocket neck of AChE. Str5 slightly differs in its binding site, simultaneously binding with the C-terminal section to the Ω-loop and Gly 342, leading to the contraction of the pocket neck and closing of the active site from the native ligand. All our ligands bind to the active site pocket of BuChE and close the pocket by their physical presence, except for Str204, which unfolds the Ω-loop, leading to the opening of the entrance to the active site. The rest of the ligands during the simulation, being in the choline binding site, make the native ligand impermeable to the BuChE active site. As a result of the study, we identified structures that may exhibit multitarget inhibition properties and verified a structure that has high specific affinity for BuChE. Thus, the results we obtained indicate that the piperazine derivatives we selected can exhibit anticholinesterase activity and are considered as therapeutic agents for AD.

1. Hampel, H., Mesulam, M.-M., Cuello, Farlow, M.R., Giacobini, E., Grossberg, G.T., Khachaturian, A.S., Vergallo, A., Cavedo, C., Snyder, P.J., Khachaturian, Z.S. (2018) *Brain*, 141(7), 1917–1933.
2. Giacobini, E. (2001) *Drugs Aging* 18, 891–898.
3. Karaytuğ, M.O., Balçı, N., Türkan, F., Gürbüz, M., Demirkol ,M. E., Namlı ,Z., Tamam, L., Gülçin, İ. (2023) *Journal of Biochemical and Molecular Toxicology*, 37(2), e23259
4. Hamaryan, V.S., Hunanyan, L.S., Popugaeva, E.A. (2021) *Collection of scientific papers of the XXVII symposium "Bioinformatics and computer-aided drug design"* DOI: 10.18097/BCADD2021.

UNVEILING THE EFFICACY OF CEFIDEROCOL IN OVERCOMING CARBAPENEM-RESISTANCE IN *PSEUDOMONAS AERUGINOSA*: INSIGHTS FROM *IN VITRO* AND *IN SILICO* ANALYSES

S. Ghosh¹, S. Basu^{1,2}, G. Ashok¹, S. Ramaiah¹, B. Veeraraghavan³, A. Anbarasu¹

¹*Medical and Biological Computing Laboratory, School of Biosciences and Technology (SBST), Vellore Institute of Technology (VIT), Vellore, India*

²*Department of Biotechnology, NIST University, Berhampur, India*

³*Department of Clinical Microbiology, Christian Medical College (CMC), Vellore, India*

The global health crisis posed by ‘high-risk’ hypermutable clones of *Pseudomonas aeruginosa*, which are extensively drug-resistant (XDR), is exacerbating mortality rates, particularly among immunocompromised individuals. Penicillin binding protein-3 (PBP3) mutations T91A and F533L in carbapenemase-positive *P. aeruginosa* rendered β-lactam antibiotics ineffective, thereby challenging β-lactam-β-lactamase inhibitor (BL-BLI) combinations. This study focused on structural insights into mutations in PBP3, with F533L and T91A mutations found in carbapenemase-positive *P. aeruginosa* isolates (n=6) identified via whole genome sequencing. Despite broad resistance to β-lactam antibiotics, these isolates demonstrated susceptibility to cefiderocol (MIC ≤ 4µg/ml). Mutations in PBP3 reduced local intra-chain interactions, slightly increased structural flexibility (~1%), and affected antibiotic interactions. Molecular dynamics simulations supported the stability of PBP3 mutants through root-mean-square deviations, radius of gyration, solvent accessibility, and density curves, favoring their persistence. Docking studies revealed that these mutations induced unfavorable stereochemical clashes with conventional antibiotics, increasing their inhibition constants (IC) up to ~50-fold. Cefiderocol, however, maintained its efficacy due to its higher binding affinity ($\Delta G: -8.2 \pm 0.4$ kcal/mol) for multiple PBP targets and lower binding affinity ($\Delta G: -6.7 \pm 0.7$ kcal/mol) to β-lactamases compared to other β-lactam antibiotics. Further molecular dynamics simulations and Molecular Mechanics Poisson Boltzmann Surface Area calculations indicated the energetically favorable binding of cefiderocol to PBP3 proteins. This study provides a structural perspective on emerging non-polar amino acid substitutions in PBP3 that contribute to XDR and advocates prioritizing existing antibiotics based on multi-target affinities to counteract the challenges posed by target-protein mutations.

Hence, antibiotics like cefiderocol with multi-target specificity can be used to overcome the challenge imposed by mutations in conventional protein targets.

This study was supported by the Indian Council of Medical Research (ICMR), Govt. of India (Research grant IRIS-2021-10630).

FOOD CHRONOBIOtics: KEY COMPOUNDS DATABASE FOR CHRONONUTRITION

D. Golubev, I. Solovev

Pitirim Sorokin Syktyvkar State University, Syktyvkar, Russia

Chronobiotics are agents that influence circadian rhythms through specific molecular mechanisms [1]. These substances interact with the body's internal clock by modulating the expression and activity of core clock genes and proteins. The concept of chronobiotic drugs has been recognized for over 50 years, beginning with the detailed clinical discovery and characterization of melatonin's properties. Despite various efforts to organize and categorize chronobiotics, there still isn't a standardized classification for these pharmacological agents. Food chronobiotics its various plant-based compounds such as anthocyanins that interact with circadian rhythm of organism and regulating proteins and gene which effect on metabolic process and health [2]. The aim of work is showed intermediate results of creating the world's first comprehensive and continually updated database of chronobiotic drugs (circadian rhythm modulators) and making it accessible via the internet. This is a crucial and timely task in the fields of chronobiology, chronomedicine, and pharmacoinformatics/bioinformatics. At present, there is no database of pharmacological modulators of circadian rhythms, including the class of chronobiotic drugs that has not yet been identified in classical pharmacological databases.

Currently, we have conducted an analysis across major scientific databases and identified over 500 unique compounds that qualify as chronobiotics, targeting molecular components such as CRY, PER, BMAL1 etc. In the future will include links to primary sources, a molecular image, a SMILES-format formula, and an IUPAC name. To enhance the database, it will be synchronized with resources like ChemSpider, DrugBank, ChEMBL, ChEBI, UniProt, and PubChem. Additionally, biological and pharmacological relevance will be increased through integration with FDA, ClinicalTrials.gov, KEGG, and other related databases.

This work was funded by the Russian Science Foundation Grant "Design of the world's first pharmacological database of circadian rhythm modulators (ChronobioticsDB) and organisation of the access to it" no. 24-75-00108

- 1.Solovev, I.A., Shaposhnikov, M.V., Moskalev, A.A. (2021) *Clocks Sleep*, 3(3), 429-441.
2. Huang, J.Q. , Lu ,M., Ho. C.T. (2021) *Food Funct.*, 12(14), 6136-6156.

RELEVANT BIOTARGETS OF STAPHYLOCOCCUS AUREUS FOR THE DEVELOPMENT OF A CONVOLUTIONAL NEURAL NETWORK MODEL FOR PREDICTING THE ANTIMICROBIAL ACTIVITY OF CHEMICAL COMPOUNDS BASED ON MULTIPLE DOCKING

A.V. Golubeva, P.M. Vassiliev, M.A. Perfilev, A.N. Kochetkov, D.A. Dushkin

Volgograd State Medical University, Volgograd, Russia

The World Health Organization has compiled a list of key pathogens for which active research into new treatments is needed due to increasing bacterial resistance to antibiotics. *Staphylococcus aureus* is included in this list as a group of microorganisms with a high level of risk. The aim of this work was to identify relevant biotargets for antimicrobial activity against *S. aureus* and further build convolutional neural networks based on multiple docking.

Using the original Microcosm BioS v23.11.2 system [1] we analyzed the generated verified database [2] of chemical compounds with high antimicrobial activity. Among 43 biotargets with activity against *S. aureus*, 11 most affine targets were found with the value of the average integral index of the level of targeted activity IndMean >0 . These targets are proteins encoded by the genes *blaZ*, *srtA*, *def*, *ligA*, *tdk*, *sspA*, *norA*, *fabI*, *spsB*, as well as DNA gyrase and topoisomerase IV. In relation to these biological targets, the indicated substances are predicted to show activity from average to significantly above average. The search for experimental 3D models for these targets was conducted in UniProtKB, PDBe and RCSB PDB databases. At the current stage of the study, 5 biotargets encoded by the *tdk*, *sspA*, *norA*, *fabI* and *spsB* genes were subjected to the validation procedure [3], during which 9 valid 3D models were identified. Optimized 3D molecular models were generated for 3,709 chemical compounds from the database, which were tested for *S. aureus* antimicrobial activity. The models were first constructed using molecular mechanics methods in MarvinSketch 15.6.15 and then the semi-empirical quantum chemical PM7 method in MOPAC2012. In accordance with the methodology [4], using the MSite v21.04.22 program, multiple docking spaces were generated, resulting in binding energy spectra being obtained for 3,709 chemical compounds across the entire volume of the analyzed proteins (in total ~4.5 million values). These spectra were subjected to a correlation-based convolution procedure. Based on the results, the classification neural network model was developed to predict the antimicrobial activity of chemical compounds against *Staphylococcus aureus*, the predictive accuracy of which exceeds 90%. The application of the multiple docking and subsequent neural network modeling procedure has shown that the identified 5 biological targets are relevant to antimicrobial activity against *Staphylococcus aureus*. The developed methodology will be applied to assess the relevance of the remaining 6 biotargets.

The work was carried out within the framework of the state assignment of the Ministry of Health of the Russian Federation № 23022400009-9 «Development of a methodology for computer search of multitargeted pharmacologically active substances based on multiple docking and convolutional neural network technology of various architectures».

1. Vasilyev, P.M., Luzina, O.A., Babkov D.A. et al. (2019) *Journal of Structural Chemistry*, 60(11), 1827-1832.
2. Golubeva, A.V., Vasilyev, P.M., Stepanenko, I.S. et al. (2024) *MedChem2024: book of abstracts*, 207.
3. Vassiliev, P.M., Maltsev, D.V., Spasov, A.A. et al. (2023) *Pharmaceuticals*, 16(5), 731.
4. Vassiliev, P.M., Kochetkov, A.N., Spasov, A.A. et al. (2021) *Volgograd Journal of Medical Research*, 3, 57-61.

APPLICATION OF THE REDUCED FACTORIAL DESIGN FOR THE PHARMACOLOGICAL OPTIMIZATION OF NEW MOLECULAR ENTITIES, THE CASE OF ANTI-INFLAMMATORIES

M.E. Campos-Aldrete¹, M.A. Meraz-Ríos², M. Guerrero-González¹

¹Departamento de Química Orgánica, Escuela Nacional de Ciencias Biológicas-Instituto Politécnico Nacional, México City, México

²Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Cinvestav-Zacatenco, México City, México

Imidazo[1,2-a]azines with the acid group decreased the inflammatory process in murine models and the effect was attributed to the inhibition of cyclooxygenase (COX) enzymes, well known as prostaglandin-endoperoxide synthase (PTGS). Due to this observation, imidazo[1,2-a]azines become an interesting alternative for the development of non-steroidal anti-inflammatory drugs that can be used in chronic degenerative diseases.

In the present work, imidazo[1,2-a]azines with different acid groups were designed and synthesized. These molecules were evaluated in *in vitro* models using both COX isoforms. The results obtained from the evaluation were used for a reduced factor analysis to select the structural elements that could improve the inhibitory activity.

The inhibitory effect of five acid derivatives of imidazo[1,2-a]azines and standards ibuprofen and indomethacin were evaluated *in vitro* on the enzymes COX isoforms, it was observed that the different substituents provide different inhibition profiles, highlighting that the group of imidazo[1,2-a]pyridines is more active than the bioisosteric imidazo[1,2-a]pyrimidines, these results were analyzed using the reduced factorial design 2^{3-1} also was performed *in silico* Docking to recognize the structural elements necessary for the inhibition of the targets. The IC₅₀ to COX1 (μM) was obtained where the compounds showed 4a (2.72), 4b (3.94), 5a (7.29), 5b (63.26), 6a (12.93) indomethacin (0.13) ibuprofen (0.2) and IC₅₀ in COX 2 (μM) 4a (1.89), 4b (2.39), 5a (8.08), 5b (41.15), 6a (5.86) indomethacin (0.09) ibuprofen (0.125).

Using factorial design, was possible optimize the inhibitory response on therapeutic targets obtaining molecule 4a as a result of factorial analysis.

Reduced factorial design can be an economical alternative to traditional QSAR or HTS methods in molecular optimization.

1. Guerrero-Gonzalez, M., Campos-Aldrete, M.E., Meraz-Ríos, M.A. (2022) *Future Medicinal Chemistry*, 14(11), 771-784.

EXAMINATION OF ADME PROPERTIES OF SOME TRIAZOLE DERIVATIVE COMPOUNDS THAT CAN BE USED AS PHARMACEUTICAL ACTIVE INGREDIENTS

E. Hacıfazlıoğlu

Recep Tayyip Erdoğan University, Technical Sciences Vocational School, Environmental Protection and Control Department, Rize, Turkey

In this study, some important ADME parameters such as physicochemical properties, lipophilicity, water solubility, pharmacokinetics, medicinal chemistry and drug-likeness properties of ten triazole derivative compounds, which may be drug active ingredients, were investigated on the SwissADME a web tool worked online. In order to evaluate drug similarity, a separate bioavailability radar was drawn for each molecule. Bioavailability radar was plotted for each molecule separately to assess drug similarity. BOILED-Egg plot was drawn for each molecule to assess passive gastrointestinal absorption (HIA) and brain penetration (BBB) based on the location of the molecules. SwissTargetPrediction, an online web tool, was used to predict the most likely protein targets of the molecules. Bioavailability Radar is seen that molecule 2 fits the description and the entire graphic line is in the pink area. We can consider this molecule as a drug candidate. Molecules 1, 4 and 7 are close to being considered as drug candidates, but there is a slight deviation from the red area at the point of insaturation feature.

1. Patel, A.A., Mehta, A.G.(2010). *Journal of Saudi Chemical Society*, 14(2), 203-208.
2. Beresford, A.P., Selick, H.E., Tarbit, M.H. (2002) *Drug Disc. Today*, 7(2), 109-116
3. Wildman, S.A., Crippen, G.M. (1999) *J. Chem. Inf. Comput. Sci.*, 39(5) 868-873.

ANTIBACTERIAL, ANTIFUNGAL AND ANTOXIDANT ACTIVITIES OF AMINE AND ALDEHYDE OF MELDRUM'S ACID: IN VITRO, MOLECULAR DOCKING AND ADMET STUDIES

N.H. Pungot^{1,2}, W.U.A. Wan Azli¹, N.A.S. Muhamad Ruzi¹, N.N. Firdaus¹, M. Zulkifli¹, M.H.M. Idris^{2,3}

¹Organic Synthesis Laboratory, Institute of Science (IOS), Universiti Teknologi MARA Selangor, Bandar Puncak Alam, Selangor, Malaysia

²Faculty of Applied Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

³Drug Discovery Center (DDC), Integrative Pharmacogenomics Institute (iPROMISE), Universiti Teknologi MARA Selangor, Bandar Puncak Alam, Selangor, Malaysia

The development of new therapeutic agents is necessary to address the major healthcare challenges posed by antimicrobial resistance and oxidative stress-related diseases. Meldrum's acid and its derivatives have demonstrated potential in various biological applications. Thus, this study investigates the antibacterial, antifungal, and antioxidant properties of amine and aldehyde derivatives of Meldrum's acid using *in vitro* assays. Furthermore, molecular docking and ADMET studies were performed to understand the mechanisms of action and evaluate the pharmacokinetic and safety profiles of these compounds. Firstly, antibacterial studies and molecular docking were conducted to observe the efficacy of treatments as antibacterial agents using disc diffusion assays, as well as to understand their mechanism by elucidating the interactions between the ligands and target protein through molecular docking. The study demonstrated that M2A (3,4-dimethoxybenzaldehyde) and M3A (4-methoxybenzaldehyde) have significant antibacterial activity against *Escherichia coli*, *Bacillus cereus*, *Streptococcus aureus*, and *Pseudomonas aeruginosa*, with inhibitory zones comparable to gentamicin. Molecular docking revealed that the M2A and M3A had good binding modes with DNA gyrase. Next, this study determined the antifungal activity, binding relationships with target proteins and ADMET properties of Meldrum's acid derivatives with amine and aldehyde substitutions. The research found that M3A (4-methoxybenzaldehyde) showed significant antifungal activity against *Candida albicans*, with an inhibition zone comparable to ketoconazole. Molecular docking revealed that M3A binds strongly to the sterol 14-alpha demethylase enzyme. Additionally, M3A demonstrated favorable ADMET properties, including good absorption and low toxicity. This advancement offers a potential new antifungal agent, addressing current drug limitations and paving the way for alternative therapeutic options. Ferric ion reducing antioxidant power (FRAP) and hydrogen peroxide scavenging assays were also conducted to assess their antioxidant properties. Based on FRAP assay, out of all compounds, only M5 (4-aminophenol) shows significant ferric ions reducing power although it is still incomparable to the effectiveness of positive control, Trolox. For hydrogen peroxide scavenging assay, M8 (2-naphthylamine) and M1A (4-hydroxybenzaldehyde) are the compounds that exhibit the highest hydrogen peroxide percentage of inhibition. Comparing how each compound reacts in different assays was also conducted. Based on the result, the relationship of FRAP value and percentage of inhibition are not directly proportional as the compounds can have high FRAP value but low percentage of inhibition. FRAP assay and hydrogen peroxide scavenging assay do not correlate to one another as both assays possess different principles targeting different radicals in determining antioxidant activity. In conclusion, Meldrum's acid derivatives could be further explored as new antibacterial and antifungal agents. Contrarily, none of these derivatives possess remarkable ferric ions reducing power. However, some of Meldrum's acid derivatives exhibit notable hydrogen peroxide scavenging activity that surpasses Trolox's scavenging ability, allowing them to act as potential antioxidants.

This study was supported by DUCS-CoE Grant (Grant no.: 600-UiTMSel (PI. 5/4) (029/2022)).

IMPRINTED ALBUMINS AS A PLATFORM FOR DRUG DELIVERY SYSTEMS

P. Ilicheva, D. Tsyupka, P. Pidenko, N. Burmistrova

Institute of Chemistry, Saratov, Russia

Serum albumins are the most abundant plasma proteins. They are highly soluble, very stable and have extraordinarily long circulatory half-life as a direct result of their size. In contrast, many therapeutic molecules are smaller than the renal filtration threshold and are rapidly lost from the circulation thereby limiting their therapeutic potential. Association, conjugation or fusion of therapeutic drugs to albumins is a well-accepted and established half-life extension technology [1]. These properties make albumins attractive building blocks for synthesis of imprinted proteins (IPs). IPs are potentially useful for the development of long-acting drug delivery systems.

The technique of bioimprinting is based on reversible changes in the three-dimensional structure of a protein induced by noncovalent interactions with a ligand [2]. IPs synthesis includes five general stages: protonation of amino acid residues of protein (i), formation of protein-template complex (ii), deprotonation (iii), fixation of the new matrix conformation via cross-linker (iv) and purification of IP binding sites (v). Development of drug delivery systems based on IPs involves time-consuming and expensive experiment. In this sense, approaches that employ computational/theoretical chemistry have the great potential at design IPs based systems [3].

Here, we report the computational approaches (molecular dynamics simulations and molecular docking) for prediction of drug loading into model drug delivery systems (albumin-based IPs). Bovine serum albumin (BSA) was used as a model protein matrix. We have adopted a comprehensive and fundamental approach to the imprinting of BSA, using a combination of experimental and theoretical results that yield a more exhaustive characterization of structural changes in the protein matrix. In this work, 4-hydroxycoumarin (4-HC) was used as template molecule. Coumarins exhibit a wide range of pharmacological activities, which includes anti-diabetic, anti-viral, anti-microbial, anticancer and anti-inflammatory activities. Among these properties, the present work compiles the research findings of 4-HC as anti-cancer agent. We simulated several stages of 4-HC-imprinted BSA synthesis to demonstrate how IPs are precisely formed and to predict the location and number of binding sites of 4-HC on the protein surface. The obtained theoretical results were used to select the appropriate ratio of protein carrier and drug for the experimental production of IP. We revealed that IPs loaded with 4-HC inhibit mouse malignant tumor cell lines in vitro. In vitro results also showed high biocompatibility of IPs without 4-HC on normal cells.

This work was supported by the Russian Science Foundation (project 22–16–00102).

1. Sleep, D., Cameron, J., Evans, L.R. (2013) *Biochim. Biophys. Acta (BBA)-General Subjects*, 12, 5526-5534.
2. Piletsky, S. (2019) Molecular Imprinting of Polymers, *Taylor & Francis Limited (Sales)*.
3. Ilicheva, P., Fedotova, E., Presnyakov, K., et al. (2024) *Molecular Systems Design & Engineering*, 9, 456-463.

INVESTIGATION OF SOME BENZIMIDAZOLE DERIVATIVE COMPOUNDS FOR THE TREATMENT OF MULTIPLE SCLEROSIS (MS) BY MOLECULAR DOCKING METHOD

F. İslamoğlu

Recep Tayyip Erdoğan University, Department of Chemistry, Rize, Turkey

Multiple Sclerosis is a progressive, neurological, and chronic disease that affects the white and gray matter (brain and spinal cord) of the central nervous system (CNS) [1]. Illness is characterized by damage to the axons, which transmit messages, and the myelin sheaths that surround and accelerate the transmission of messages [2]. In this study, the usability of twelve different benzimidazole derivatives as active pharmaceutical ingredients in the treatment of multiple sclerosis was investigated. For each molecule, a docking study was carried out with the target protein using the Auto Dock Vina program, and the highest docking score of -13.3 (Kcal/mol) was obtained in number twelve. However, when calculated for all molecules in bioactivity score, physicochemical properties, lipophilicity, water solubility, pharmacokinetic properties, drug-likeness, medicinal chemistry properties, toxicity, total energy, and dipole moments , it was determined that the most ideal molecule was (1,3-Bis((5-(ethylamino)-1,3,4-thiadiazol-2-yl)methyl)-1,3-dihydro-2H-benzimidazol-2-one (molecule 1). As a result of all these studies, it has been concluded that this molecule can be used as an effective drug in the treatment of multiple sclerosis.

1. Noh, J.M., Seon-Yeong, K., Seo, H.S., Seo, J.H., Kim, B.G, Lee, Y.S. (2009) *Bioorganic & Medicinal Chemistry Letters*, 19, 5586-5589.
2. Tiwari, A.D., Panda, S.S., Asiri, A. M., Hall, C.D. (2014) *Synthesis*, 46, 2430-2435.

AUTOMATED ANALYSIS OF STRUCTURE-MULTIPLE PROPERTY RELATIONSHIPS: IMPACT ON SMARTS

J.I. Espinoza Castañeda, J.L. Medina-Franco

*DIFAQUIM research group, Department of Pharmacy, School of Chemistry,
National Autonomous University of Mexico, Mexico*

Structure property-(activity) relationships -SP(A)Rs- is a fundamental concept in medicinal chemistry that systematically explores the association between the chemical structure of a compound and its properties (including biological activity in the case of SARs) [1]. Systematic SP(A)R analysis helps to obtain information to identify the structural features that influence biological and pharmacological effects, making it available to develop further computational techniques, such as predictive models.

Complex and multifactorial diseases such as Alzheimer's, cancer and diabetes are the consequence of the systemic breakdown of physiological networks [2], where multiple proteins and pathways are involved in the onset and development of the disease. To address complex diseases is frequently used the concept of polypharmacology, where a molecule can interact with two or more targets simultaneously generating synergistic effects and resulting in better therapeutic outcomes. Indeed, the development of multitarget drugs is a promising approach and a general trend in drug discovery [3]. The design of multi-target compounds offers advantages as compared to the conventional single-targeting molecules; a multi-targeting drug is more effective for lower acquired resistance and simplified therapy formulation [4].

To aid the design of multi-target compounds and explore systematically structure multiple-activity relationships - SMARTs - we are developing a free automatic tool to provide a global visualization of the chemical multiverse of compound data sets annotated with multiple biological activity. Of note, the chemical multiverse of a compound data set is a group of chemical spaces, each generated with complementary representations [5].

The free tool, built in Python, contributes to the quick analysis of SMARTs by performing three main processes: 1) dataset curation, 2) generation of chemical spaces based on the similarity calculation obtained with three groups of descriptors, and 3) dimensionality reduction to generate 2D visual representations of the chemical space. In this project, we are using MACCS keys (166 bits) and ECFP-4 fingerprints, and six molecular descriptors with pharmacology relevance to describe the chemical spaces. The methods to perform the dimensionality reduction are PCA and t-SNE, both integrating an activity index. The tool can use datasets in various formats with information on compounds with activity against one or more targets. In support of open science, the code is freely available on Google Colab at: Automated Multi-Target Multiverse

1. Wawer, M., Lounkine, E., Wassermann, M., Bajorath, J. (2010) *Drug Discovery Today*, 15(15), 630-639
2. Csermely, P., Ágoston, V., Pongor, S. (2005) *TRENDS in Pharmacological Sciences*, 26(4), 178-182
3. López-López, J.I., Medina-Franco, J.L. (2024) *Drug Discovery Today*, 29(7), 104046
4. Antolin, A., Workman, P., Mestres, J., Al-Lazikani, B. (2016) *Curr. Pharm. Des.*, 22(46), 6935-6945.
5. Medina-Franco, Chavez-Hernández, López-López, Saldivar-Gonzalez (2022) *Mol. Inf.*, 41, 2200116

MOLECULAR DOCKING OF SECONDARY METABOLITE COMPOUND OF KAWISTA (*LIMONIA ACIDISSIMA*) AS HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR-2 (HER-2) INHIBITORS

A. Khoirunnisa, M.A Muchlisin, A.A. Rosali, E.J. Astuti, A. Rafikayanti

Department of Pharmacy, Faculty of Health Science, University of Muhammadiyah Malang, Malang, Indonesia

The kawista plant (*Limonia acidissima*) is empirically known as a breast cancer growth inhibitor. This study aims to investigate the active components of *L. acidissima* and compare them with other anticancer drugs using in silico methods. Secondary metabolite compounds of *L. acidissima* were obtained from the Dr. Duke Phytochemical and Ethnobotanical database. Prediction of compound activity was carried out using molecular docking method using Pyrx autodock software. 31 secondary metabolite compounds of *L. acidissima* were obtained from Dr. Duke's database. From the molecular docking results, four secondary metabolite compounds were obtained: limodissimin A (-10.16 Kcal/Mol; 35.69 nM), dihydroxyacidissiminol (-10.08 Kcal/Mol; 40.97 nM), acidissiminol (-10.12 Kcal/Mol; 38.47 nM), and acidissiminol epoxide (-10.42 Kcal/Mol; 23.07 nM) had better binding energy and pKi than the positive control SYR127063 (-9.84 Kcal/Mol; 60.87 nM). Limodissimin A is predicted as the most potential secondary metabolite compound and can be further developed as an anti-cancer HER-2 positive therapy because it has the lowest predicted binding energy and Ki.

ON THE EFFICIENCY OF CARBAPENEM ACTIVATION BY BETA-LACTAMSES

D.D. Kholmanskikh^{1,2}, M.E. Kuznetsov^{1,2}, D.A. Tatarnikova¹, A.V. Krivitskaya^{2,3}, M.G. Khrenova^{1,2}

¹*Lomonosov Moscow State University, Moscow, Russia*

²*Emanuel Institute of Biochemical Physics of the Russian Academy of Sciences, Moscow, Russia*

³*Fundamental Research Centre of Biotechnology of the Russian Academy of Sciences, Moscow, Russia*

Antibiotic resistance is a constant problem of bacterial infection treatment. Beta-lactams represent the main group of antibiotics that disrupt cell wall building due to the inhibition of penicillin-binding proteins. Beta-lactamases are enzymes that appeared evolutionary to inactivate beta-lactam compounds. Now, four classes of beta-lactamases are known. Three of them, A, C and D, are serine hydrolases and B class are zinc-dependent enzymes, called metallo-beta-lactamases. Beta-lactam compounds are also different. Among them, the most known are penicillins, cephalosporins, monobactams and carbapenems. Many experimental studies are conducted to determine the rate steady state kinetic parameters of antibiotic inactivation by beta-lactamases. Theoretical studies can serve rational background of variations in catalytic parameters.

The binding efficiency might be related to the conformational diversity as well as flexibility of the entire protein. This can be studied using classical molecular dynamic (MD) simulations with the subsequent analysis of trajectories. Further hydrolysis rate depends on the interactions of the enzyme with the substrate in the active site. For hydrolases there are two most important participants of the reaction. Those are a nucleophilic particle and a so-called oxyanion hole. The nucleophile can be either an oxygen atom of the catalytic serine residue or an oxygen atom of the hydroxide anion for serine and metallo-beta-lactamases, respectively. The oxyanion hole is formed by either NH groups or zinc cations depending on the particular type of beta-lactamase. The proper way to deepen the understanding of these interaction is molecular dynamics with combined quantum mechanics / molecular mechanics (QM/MM) potential. Herein, we demonstrate results on QM/MM MD simulations of several carbapenems with a metallo-beta-lactamase NDM-1 to explain differences in their observed kinetic parameters.

The authors acknowledge financial support from the Russian Science Foundation, project 19-73-20032. The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University.

MATHEMATICAL MODELING OF EPILEPSY DRUG THERAPY

P. Kondrakin, F. Kolpakov

Sirius University of Science and Technology, Sirius, Russia

Epilepsy is a chronic neurological disorder characterized by a predisposition to spontaneous seizures. Despite the availability of more than 30 antiepileptic drugs, about one-third of epilepsy patients are drug-resistant. Using computer modeling to study the mechanisms of neuronal response to epilepsy drug therapy can enhance general understanding of this complex disease, improve treatment efficacy, and reduce the risk of side effects. Computational neuroscience methods have developed very rapidly in recent years, and modern successes in seizure modeling have enabled direct correspondences between experimentally measurable quantities and simulated biophysical processes, such as changes in ion concentrations and membrane potential fluctuations. This progress allows for biologically meaningful modeling of the biophysical pathways targeted by antiepileptic drugs.

This study aims to develop a mathematical model of neuronal activity that accurately describes the functioning of nerve cells, taking into account the dynamics of the primary targets of antiepileptic drugs. To achieve this goal, we reproduced results from several existing synaptic transmission models on the BioUML platform (biouml.org), focusing on key cellular processes important in the pathology of epilepsy.

Based on reproduced models, we constructed a modular model (see Figure 1). The core of the model was based on [1] and was further extended to include the regulation of GABA and glutamate dynamics by astrocytes as detailed in [2]. Additionally, the model incorporated GABAA, GABAB, and mGlu receptors [3], as well as the inhibitory effects of interneurons [4]. The model includes 247 variables, 42 differential equations, 189 parameters, and 234 assignment operations.

As a result, we have developed the modular computer model of neuronal activity. It captures the dynamics of the main molecular targets of antiepileptic drugs and makes it possible to simulate their effects. The model considers the principal molecular mechanisms involved in controlling excitation and inhibition within the neuronal synaptic network. Special attention is given to the dynamics of ion currents, receptors, and neurotransmitters. Comparative analysis of the numerical results from the constructed model with published clinical trial data revealed a qualitative reproduction of the primary effects observed during first-line drug therapy for epilepsy. The model is well-suited for use in the study of epilepsy drug therapy. In the future, we plan to integrate the modular model into the multilevel model we previously developed [5], with the aim of modeling drug therapy at various levels of brain organization.

The study was supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-10-2021-093).

1. Toman, M., Wade, J., Verkhratsky, A. et al. (2023) *Scientific Reports*, 13, 1.
2. Flanagan, B., McDaid, L., Wade, J. et al. (2021) *Frontiers in Cellular Neuroscience*, 15.
3. Li, L., Zhou, J., Sun, H., Liu, J. et al. (2020) *Comput and Math Methods in Medicine*, 2020.
4. Borjkhani, M., Bahrami, F., & Janahmadi, M. (2018) *PLOS ONE*, 13, 3.

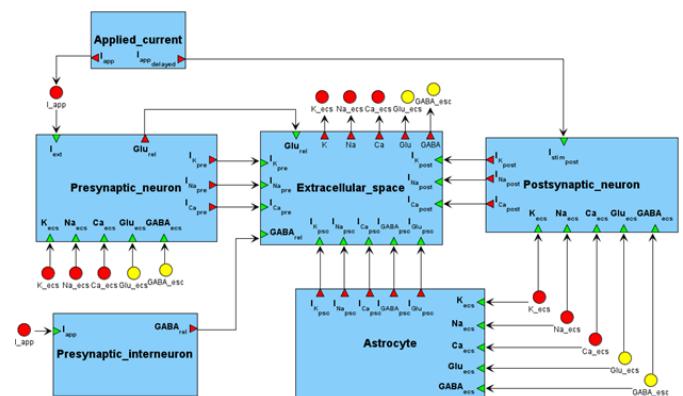


Figure 1. Modular structure of the model.

SEARCH FOR PROTEIN TARGETS FOR CATIONIC PEPTIDES USING COMPUTER MODELING

N.A. Koroleva, O.G. Kovtun, A.A. Lushnikova

Federal State Budgetary Institution «N.N. Blokhin National Medical Research Center of Oncology» of the Ministry of Health of the Russian Federation, Moscow, Russia

Search for tumor-specific overexpressed molecular targets for new anticancer compounds and evaluating of the mechanisms of their interactions is important step in development of effective drugs. Amphiphilic arginine and lysine- rich cationic peptides (CPs) with a branched molecular structure are considered as a perspective antitumor compounds, enriched in arginine and lysine, of amphiphilic nature.

Objective is to study proposed targets for some CPs using binding simulation of ligands (CP) and key proteins expressed in tumor cells.

Results. The first step in the search for relevant protein targets is the systematization of the data obtained from corresponding experimental resources. We analyzed two cell lines originated from gastrointestinal stromal tumor (GIST) – GIST naïve and its drug-resistant derivate – GIST/T1 IMR. The following proteins were found to be overexpressed and involved in high cell proliferation in these cell lines: chaperone multifunctional proteins nucleolin (NCL), nucleophosmin-core (NPM), tyrosine kinase Kit (KIT), carrying marker extended amino acid deletion: C-KIT11 ex.del. V560-Y578; then, platelet-derived growth factor receptor alpha (PDGFR), fibroblast growth factor receptor 2 (FGFR2), vascular endothelial growth factor receptor 2 (VEGFR2), heat shock protein HSP 90-beta (HSP 90), polycomb complex protein BMI-1 (BMI1), mitogen-activated protein kinase 1 (ERK1).

The optimal X-ray structures were found in Protein Database – PDB for these molecular targets: NCL (2KRR), NPM (2P1B), KIT (6KLA), PDGFR (5GRN), FGFR2 (6LVK), VEGFR2 (1Y6A), HSP 90 (5UCJ), BMI1 (7ND1), ERK1 (4QTB). Next step - pairwise molecular modeling was carried out using selected receptors in order to identify ones with high affinity for PCs. Molecular modeling was performed in Maestro (Schrödinger) software using the Glide protocol. Protein preparation includes optimal pH = 7.0 and OPLS3 force field, modeling parameters scaling factor 1, partial charge cut-off= 0.25.

The following results were obtained by pair molecular docking - the average interaction scores (glide scores) are indicated between CP (AM2, NC783) and active centers of the proteins: NCL (-7.83), NPM (-8.53), KIT (-7.23), PDGFR (-7.52), FGFR2 (-9.6), VEGFR2 (-7.23), HSP 90 (-9.37), BMI1 (-5.93), ERK1 (-8.17).

Thus, we've proven, that the most promising molecular targets in GIST sensitive and drug-resistant cells are proteins overexpressed as receptors on tumor cell surface or in cell cytoplasm: NCL, NPM, PDGFR, FGFR2, HSP 90, ERK1. These data are compatible with the results obtained by in vitro experiments. In particular, significant overexpression of the chaperone proteins NCL and NPM was found in tumor cells, which, when interacting with PC, induce apoptosis with subsequent induction of apoptosis in large proportion of tumor cells (up to 75-80%).

Conclusions. Computer modelling using molecular binding simulation by docking is a relevant approach to select PCs with high cytotoxicity for both sensitive and drug-resistant GIST cells due to inactivation by PCs of several key proteins important for cell signaling. So, cationic peptides under study are considered as perspective ones for development new anticancer drugs with selective activity.

1. Unk, M., Jezeršek Novaković, B., Novaković, S. (2023) *Cancers (Basel)*, 15(5), 1498.
2. Floris, G., Sciot, R., Wozniak, A., Van Looy, T., Wellens, J., Faa, G., Normant, E., Debiec-Rychter, M., Schöffski, P. (2011) *Clin. Cancer Res.*, 7(17), 5604-14.
3. Lushnikova, A.A., Andreev, S.M., Kozhikhova, K.V., et al. (2018) *Allergy*, 1315, 234-235.

ACTIVATION GATE OF TRPV1 ION CHANNEL IN DIFFERENT FUNCTIONAL STATES OF ITS CONDUCTING PORE

I.V. Lazarev^{1,2}, Y. N. Trofimov^{1,3}, R.G. Efremov^{1,4,5}

¹*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, Russia*

²*First Moscow State Medical University MSMU, Moscow, Russia*

³*National Research Nuclear University MEPhI, Moscow, Russia*

⁴*National Research University Higher School of Economics, Moscow, Russia*

⁵*Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Russia*

TRPV1 is a homotetrameric cation channel capable of activation when exposed to high temperatures ($>43^{\circ}\text{C}$), low pH values (<5.9), or some chemicals; it also plays a critical role in thermal sensation, and nociception in the human body. Given that existing drugs designed to combat chronic pain and neurologic pathologies are limited by a lack of efficacy and the presence of side effects, TRPV1 is one of the most promising targets for analgesic development.

Among the models of TRPV1 detected by cryo-EM methods, three main states can be distinguished for the channel pore: α -closed, π -closed, and π -open [1]. The π -states are so named because of the presence of the π -helix segment in the central part of the pore-forming helices S6. The α -closed state is characterized by a fully α -helical conformation of S6, which leads to rotation of the C-terminal part of the helices by about 100° . The structure of the activation gates of the channel — the pore region responsible for modulation of its conductance, also differs significantly for the α and π -states. Although numerous models of TRPV1 in basic conformations, as well as in many intermediate states, have been obtained experimentally, a better understanding of the dynamics and energetics of channel functioning requires a detailed analysis of the functional characteristics of these states *in silico*.

In this work, we studied in detail the gate region of rat TRPV1. To do this, we performed molecular dynamics simulations for the three basic states of TRPV1 and used a molecular modeling-based approach called “dynamic molecular portrait” (DMP) [1] to characterize the structural intermediates of the protein. The pore forming fragments of helices S6 (residues 663:690) and neighboring regions of helices S5 (residues 560-600) were considered to analyze inter-helical interactions (h-bonds, unfavorable hydrophobic-hydrophilic and favorable hydrophobic-hydrophobic contacts) and the protein-water complementary (hydration of hydrophobic and hydrophilic surface areas of the helices). The analysis showed that the transitions between the TRPV1 states are accompanied by the oppositely directed changes in the complementarity degree of the interactions, which compensate for unfavorable contributions of each other. Thus, the transition from the α -closed to the π -states leads to the breaking of some h-bonds in S6 due to the formation of the π -bugle, on the one hand, and the appearance of the T685-D576 h-bond between the helices S6 and S5, on the other. The α -closed state is characterized by the most favorable packing of the gate lining residues, which are hydrophobic and dehydrated in this state. However, the transition to the π -states leads to the decreasing of the gate packing quality and simultaneously to the increasing of S6-S5 contacts complementarity. Variations in the hydration of hydrophobic areas of the helices additionally balance the overall interactions complementarity for the different channel states. In this way the DMP approach revealed the fine balance between the various types of intermolecular interactions, which presumably provide transitions between the functional states of TRPV1 pore without significant energy barriers. Such an approach can also be useful for structural and functional analysis of the other types of ion channels.

This study was supported by RSF grant 23-14-00313.

1. Trofimov, Y.A., et al. (2024) *Commun. Chem.* **7**, 119.

MODELING OF THE PHOSPHORYL TRANSFER MECHANISM IN THE ACTIVE SITE OF PROTEIN KINASE A

M.S. Leonova¹, A.V. Krivitskaya², T.I. Mulashkina^{1,2}

¹*Moscow State University, Moscow, Russia*

²*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia*

Protein phosphorylation is one of the most important mechanisms in cell signaling and gene regulation. Consequently, this reaction is treated as a therapeutic target for cancerous, immune, inflammatory and neurodegenerative diseases. Enzymes called protein kinases catalyze the reaction of the γ -phosphate transfer from the ATP to specific residues such as serine, threonine or tyrosine. Among the superfamily of protein kinases, cAMP-dependent protein kinase (protein kinase A, PKA) was the first to be characterized. Moreover, PKA has been analyzed most thoroughly ever since becoming a model of all kinases because of the highly conserved core of the enzyme.

Although, PKA is one of the best studied kinases, there are still questions about nature of the phosphoryl transfer mechanism. Phosphorylation reactions are followed by the cleavage of the P-O bond. Considering transition state, the nucleophile and leaving group may be bonded to phosphorus to varying degrees, so associative and dissociative mechanisms may be recognized. In the case of associative mechanism, the bond formation with the nucleophile starts when the ADP and the γ -phosphate group are still rather close. Conversely, dissociative mechanism describes transition state nature where the new bond is yet to be formed properly but the bond between the ADP and the leaving group is already almost broken.

Thus, the aim of this study was to determine the type of mechanism in the reaction of serine phosphorylation of the substrate SP20 in the active site of PKA. The geometric and electron-density criteria of structures corresponding to possible conformations of the enzyme-substrate complex were analyzed. In addition, the Gibbs energy profile of the serine phosphorylation reaction in the active site of PKA was predicted using molecular modeling methods.

In this study, molecular dynamic trajectories were clustered, and three stable conformations of the enzyme-substrate complex were found, differing by the mutual arrangement of the substrate's serine and the phosphate tail of ATP. Representative frames were selected from these clusters, and molecular dynamics calculations were performed with the potentials of the combined quantum mechanics/molecular mechanics (QM/MM) method: for analyzing the state of the enzyme-substrate complex — without adding a bias potential, and for plotting the Gibbs energy profile — with the addition of a bias potential using the umbrella sampling method. The quantum subsystem included side chains of residues Lys72, Asp166, Lys168, serine of substrate SP20 and 7 molecules of water, as well as two magnesium cations and their coordination spheres represented by side chains of residues Asn171 and Asp184, phosphate groups of the ATP and 3 molecules of water. This subsystem was described at the DFT level with the PBE0 hybrid functional and D3 dispersion correction, using the 6-31G** basis set. The CHARMM force field was used to describe the MM subsystem. The difference between the distances of the breaking bond ($P-O_{ATP}$) and nucleophilic attack ($P-O_{Ser}$) was chosen as the reaction coordinate: $\Delta = d(P-O_{ATP}) - d(P-O_{Ser})$.

The mean value of the breaking bond length and the value of the Laplacian of electron density along the line of the breaking P-O bond can be used as criteria for determining the type of mechanism. Applying these criteria to the cleaved $P-O_{ATP}$ bond showed that, regardless of the conformation of the enzyme-substrate complex, the phosphorylation reaction in the active site of PKA occurs by a dissociative mechanism. Analysis of the Gibbs energy profiles computed for the phosphorylation reaction of serine in the active site of PKA also showed that the reaction proceeds by a dissociative mechanism. In addition, the Gibbs energy profile for the conformational transition of the detected conformations was obtained, and it was determined which of the conformations is the most reactive.

This work was supported by the Russian Science Foundation (project no. 23-13-00011). The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University.

LARGE-SCALE PREDICTION OF BIOLOGICAL ACTIVITIES WITH ACTIVE-IT SYSTEM

V.L. Almeida^{1,2}, O.D.H. dos Santos³, J.C.D. Lopes¹

¹Chemoinformatics Group (NEQUIM), UFMG, Belo Horizonte, MG, Brasil

²Serviço de Fitoquímica e Prospecção Farmacêutica, FUNED, Belo Horizonte, MG, Brasil

³Departamento de Farmácia, Escola de Farmácia, UFOP, Ouro Preto, MG, Brazil

Traditional testing methods in pharmaceutical development can be time-consuming and costly, but *in silico* evaluation tools like the Active-IT system can offer a solution. This system uses multi-conformational pharmacophore fingerprints to predict small organic molecules' biological and pharmacological activities. It utilizes recursive partition for multiple model construction and validation, resulting in more diverse and robust predictions. The system has four modules: 3D-Pharma for generating descriptors, ExCVBA for modeling through machine learning methods, an extensive database of activity models, and Active-IT for making predictions. Data is gathered from the PubChem BioAssay database and modeled using SVM and Naïve Bayes machine learning methods. Models are constructed using a recursive stratified partition method and undergo an activity randomization (Y-random) process to prevent overfitting. More than 3500 bioactivity models were built, each comprising 30 SVM and 30 Naïve Bayes models and 120 randomized models (Fig. 1).

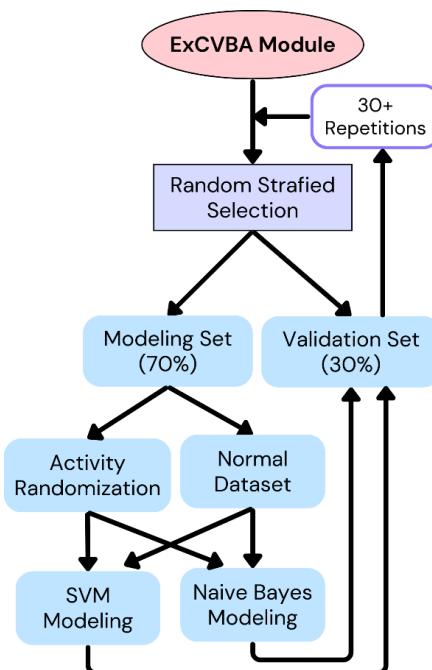


Fig.1. Flowchart of model building with ExCVBA module of Active-IT system.

The Active-IT module uses the raw score distributions of active and inactive molecules among multiple models to calculate the probability of the compound being active (P_a) or inactive (P_i). The P_a - P_i value estimates the compound's activity potential. This study evaluated five bioactive compounds of ayahuasca tea using the Active-IT system. Results showed that the system accurately predicted the activity already described by these compounds and proposed new biological activities and targets. The predictions were then validated both internally and externally. The external validation used known targets described in several public databases, such as DrugBank, IUPHAR/BPS, Therapeutic Target Database (TTD), and Comparative Toxicogenomics Database (CTD). The external validation results are remarkable, with 43% to 48% of 33 known targets correctly predicted. This level of accuracy in large-scale virtual screening 3D methods has not been mentioned in the literature before.

This study was supported by the Brazilian Science Without Border program (CNPq fellowships 202407/2014-4 to JCDL and 249299/2013-5 to VLA) and FAPEMIG fellowship BIP-00213-24 to VLA.

CONFORMATIONAL DYNAMIC PHARMACOPHORE FINGERPRINTS WITH 3D-PHARMA

B.F. Domingues¹, F.M. Santos¹, A.M. José¹, V.L. Almeida^{1,2}, J.C.D. Lopes¹

¹Chemoinformatics Group (NEQUIM), UFMG, Belo Horizonte, MG, Brasil

²Serviço de Fitoquímica e Prospecção Farmacêutica, FUNED, Belo Horizonte, MG, Brasil

3D-Pharma is a Ligand-based methodology that utilizes 3D potential pharmacophores to build molecular multidimensional descriptors for small molecules. The methodology involves the calculation of dominant tautomers and protonation states of molecular structures, conformational sampling, partial charge calculations, and assigning pharmacophore types. The pharmacophore fingerprints are built with all 3-point potential pharmacophores (3-PPP) of each molecular representation. The uniconformational vectors are converted to a unique (modal) multiconformational fingerprint representing all chemical and conformational molecular dynamics (Fig. 1).

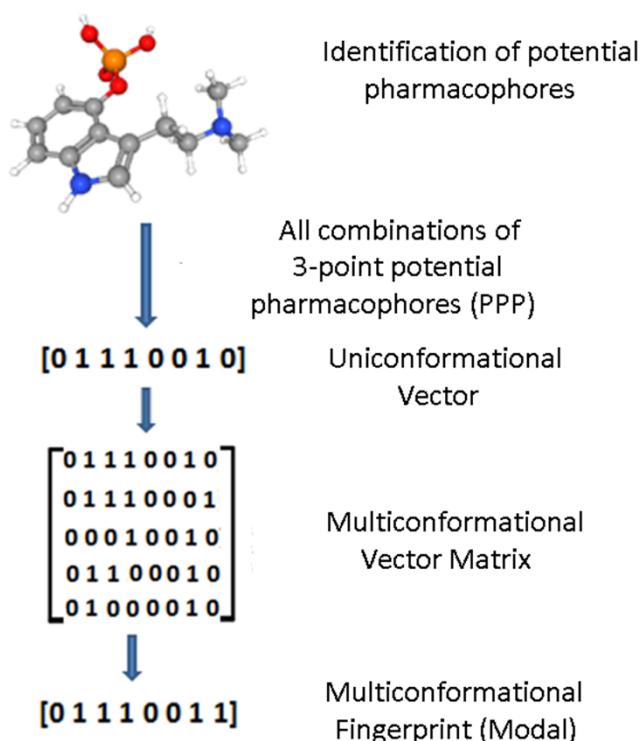


Fig. 1. Construction of unique (modal) multi-conformational pharmacophore fingerprint.

The study examined ten protein targets of the DUD datasets and found that 3D-Pharma consistently produced high-quality models. Using an ensemble template approach, the query is built from a set of known active molecules, where the model is a unified fingerprint based on the frequency of specific pharmacophore triplets. The study created 30 models for ten targets from the DUD database using three external datasets (Drugs, PDB Ligands, and WOMBAT Ligands), consistently producing high-quality models. The 3D-Pharma approach enables it to perform calculations at a speed comparable with standard 2D methods. The dataset size used for modeling significantly impacted the model quality, with most datasets producing very good models. The analysis showed that 3D-Pharma outperforms other LBVS tools regarding global accuracy, scaffold hopping, and early recovery capacities. Overall, the data suggests that 3D-Pharma is a reliable and predictive approach for drug discovery.

This study was supported by the Brazilian Science Without Border program (CNPq fellowships 202407/2014-4 to JCDL and 249299/2013-5 to VLA) and FAPEMIG fellowship BIP-00213-24 to VLA.

POWER METRIC (PM) ROBUSTNESS COMPARED WITH OTHER ENRICHMENT METRICS

V.L. Almeida^{1,2}, J.C.D. Lopes¹

¹Chemoinformatics Group (NEQUIM), UFMG, Belo Horizonte, MG, Brasil

²Serviço de Fitoquímica e Prospecção Farmacêutica, FUNED, Belo Horizonte, MG, Brasil

The Power Metric (PM) is a new metric that belongs to enrichment-type metrics, like Enrichment Factor (EF) and ROC enrichment (ROCE) (Lopes et al., 2017). It can be expressed using the formula

$$PM = \frac{TPR}{TPR + FPR}$$

TPR is the true positive rate, and FPR is the false positive rate. Power Metric expresses the probability of correctly classifying an active compound (De Winter and Lopes, 2018). This work aims to analyze the statistical robustness of the Power Metric compared to other enrichment metrics. The methodology involves extensive testing, including generating ranks of active compounds based on the exponential distribution (Truchon and Bayly, 2007). Several simulations were performed with varying numbers of total and active compounds and the model quality. Each simulation was repeated 1 million times, and the results were analyzed through the variations of mean, standard deviations (STD), and relative error (STD/mean) of the metrics. For the Power Metric, the values of the above parameters are maintained independently of the dataset composition, highlighting its reliability.

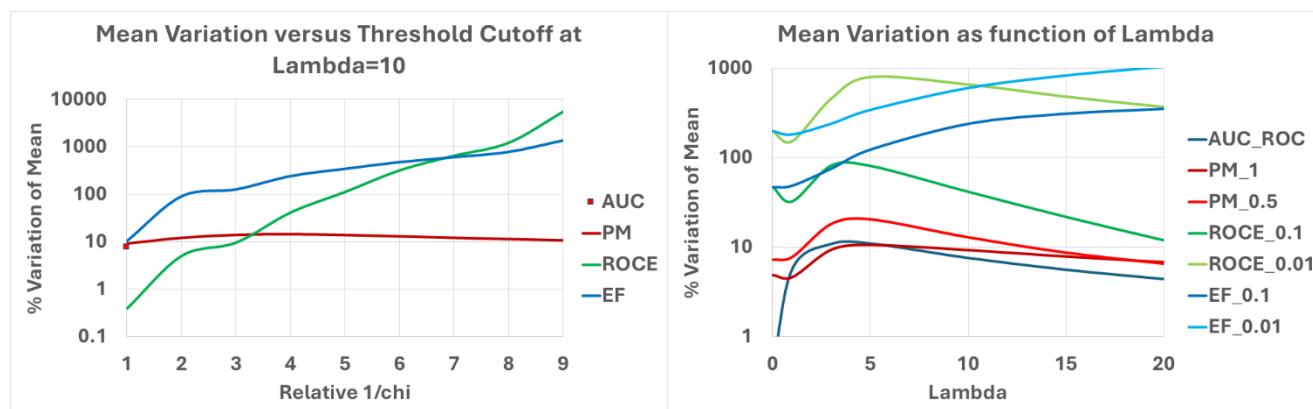


Fig. 1. Percentage variation of the mean against cutoff (A) and model quality (lambda) (B).

The Power Metric's reliability and robustness make it an invaluable tool for meta-analysis studies, particularly those involving early recovery capabilities. Thus, it can replace AUC with advantages. Even at smaller threshold values, the Power Metric maintains its robustness, which sets it apart from other less robust metrics. These practical advantages make Power Metric an appealing choice for virtual screening studies, where its robustness can significantly enhance the screening process's accuracy.

This study was supported by the Brazilian Science Without Border program (CNPq fellowships 202407/2014-4 to JCDL and 249299/2013-5 to VLA) and FAPEMIG fellowship BIP-00213-24 to VLA.

1. Lopes, J.C.D., et al. (2017) *J. Cheminform*, 9, 7
2. De Winter, H., Lopes, J.C.D. (2018) *J. Cheminform*, 10, 13
3. Nicholls A (2008) *J. Comput. Aided Mol. Des.*, 22, 239–255
4. Truchon, J.-F., Bayly, C. I. (2007) *J. Chem. Inf. Model.*, 47(2), 488–508

A GENERAL PROTOCOL FOR THE CONSTRUCTION OF STRUCTURE-ACTIVITY LANDSCAPES OF NON-CANONICAL PEPTIDES

E. López-López^{1,2}, *O. Robles*³, *F. Plisson*⁴, *J.L. Medina-Franco*¹

¹DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico

²Department of Pharmacology, Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico

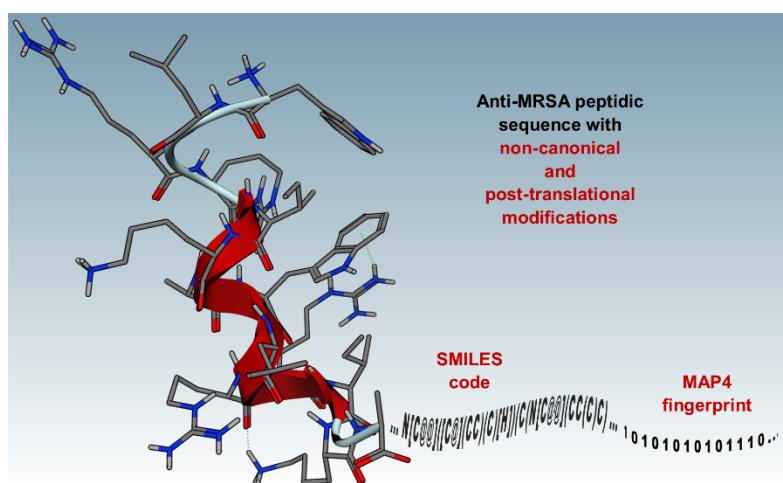
³Medicinal Chemistry and Chemogenomics Laboratory, Faculty of Bioanalysis-Veracruz, Universidad Veracruzana, Mexico

⁴Department of Biotechnology and Biochemistry, Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico

Peptide structure-activity/property relationship (P-SA/PR) studies focus on understanding how the structural variations of peptides influence their biological activities and other functional properties. This knowledge accelerates the rational design and optimization of peptide-based drugs, biomaterials, or diagnostic agents. These studies examine peptide structures from their primary sequences, essentially encoded from the 20 amino acids. Current approaches often exclude peptide libraries with post-translational and synthetic modifications, for this reason, it is imperative the development of a novel protocol capable of decoding SPR on non-canonical peptides [1].

In this work, we introduce a novel methodology based on the use of MAP4 fingerprint [2] to construct the first structure-activity landscape to study non-canonical peptides. We constructed and studied a dataset of 223 antimicrobial peptides with reported activity against methicillin-resistant *Staphylococcus aureus* (MRSA) strains, which allows us to identify peptidic activity cliff(s). Namely, the present protocol allows us the systematic study of the peptide activity from sequence, structure, and tridimensional perspectives, which allows us to identify structural motifs that play a crucial role in the peptidic anti-MRSA activity.

This is the first computational study to systematically explore the activity landscape of peptides with non-canonical residues, emphasizing the quantification of structural similarity, which opens new perspectives on the rational design of drugs, biomaterials, antibodies, and food chemicals.



1. López-López, E., Robles, O., Plisson, F., Medina-Franco, J. L. (2023) *Digital Discovery*, 2(5), 1494-1505.
 2. Capecchi, A., Probst, D., Reymond, J. L. (2020). *Journal of Cheminformatics*, 12, 1-15.

SOME INSIGHTS FROM THE ANALYSIS OF DATASET OF DRUGS, THEIR MOLECULAR SCAFFOLDS AND MEDICAL INDICATIONS

G. Malakhov^{1,2}, P. Pogodin¹

¹*Institute of Biomedical Chemistry, Moscow, Russia*

²*Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia*

Bemis-Murcko scaffolding [1] is a powerful tool for compound clustering and subsequent analysis. Using ChEMBL database [2] and RDKit library [3], we have compiled the dataset of known small molecule drugs, their molecular scaffolds and associated medical indications augmented with the interactive interface (<https://rsf-23-73-01058.github.io/scaffolds-of-the-known-drugs/>). These data can help to understand both pharmacological promiscuity of the drugs associated with a particular scaffold and diversity of structures of drugs associated with a particular indication. The data can help to perceive the diversity of known drugs in general [4].

Using this dataset: 1) we assessed the relationship between drugs' scaffolds and medical indications. For this purpose, we built scaffold-based classifiers, one for each medical indication, and evaluated them in Leave-One-Out Cross-Validation. Average Balanced Accuracy values were about 0.7 and 0.6 after micro and macro averaging, respectively. These findings testify that the concept of scaffolds should be of use for finding novel indications for drugs and (Q)SAR in general. 2) We compared the diversity of scaffolds of drugs belonging to the different classes defined using USAN stems and major protein families assigned to them according to ChEMBL (Nuclear receptors, Phosphodiesterases, G protein-coupled receptors, Proteases, Ion channels, Kinases), the diversity of kinase inhibitors' scaffolds was the lowest according to our analysis, which corresponds well to the fact that the most of kinase inhibitors interact with the conserved ATP binding site of kinases. Although, we assessed the silhouette values for scaffold grouping corresponding to the drugs' classes and these values were near zero for each class indicating the decent baseline of the scaffold diversity for drugs belonging to distinct classes. 3) We calculated the molecular complexity scores for drug scaffolds and compared the complexity of scaffolds describing only one and more than one drug. Scaffolds describing more than one drug were statistically simpler, thus, application of molecular complexity scores to the scaffolds support the hypothesis that promiscuous entities in general have lower complexity.

Here we present the novel dataset and its application in studies on drug scaffolds' relations to the medical indications, their diversity and complexity.

The study was supported by the Russian Science Foundation, grant № 23-73-01058, (<https://rsrf.ru/en/project/23-73-01058/>).

1. Bemis, G.W., Murcko, M.A. (1996) *J. Med. Chem.*, 39(15), 2887–2893
2. Mendez, D., Gaulton, A., Bento, A.P., et al. (2019) *Nucleic Acids Research*, 47(D1), D930-D940
3. rdkit/rdkit: 2023_09_1 (Q3 2023) Release Beta.
4. Malakhov, G., Pogodin, P. (2024) *Data in Brief*, 54, 110417.

EXPLORING INTERACTION OF AMPHOTERICIN B AND PLANT FLAVONS IN MEMBRANE ENVIRONMENT BY MOLECULAR DYNAMICS

A. Malykhina, S. Efimova, O. Ostroumova

Laboratory of Membrane and Ion Channel Modeling, Institute of Cytology of Russian Academy of Sciences, Saint Petersburg 194064, Russia

Amphotericin B (AmB) due to its wide spectrum of action, low resistance of pathogenic strains and high clinical efficacy, remains the gold standard in the treatment of serious systemic fungal infections. The main mechanism of action of AmB is considered to be the formation of pores by complexes of AmB and ergosterol in the plasma membrane of fungi which leads to leakage of electrolytes and other molecules from the intracellular medium and subsequent cell death. However, AmB use is limited with nephrotoxicity, cardiotoxicity and hemolysis. The reason for the high toxicity of this antifungal agent is the possibility of the formation of complexes with cholesterol and the subsequent formation of pores in the membranes of mammalian cells. Combination of AmB with plant flavones seems to be a good strategy for finding less toxic and more effective formulation. The effects of 26 plant small molecules characterizing by disordering membrane action on AmB pore-forming ability were tested by electrophysiological assays with model lipid membranes. 13 molecules were shown to increase AmB activity, 1 agent had no effect and 12 compounds were able to decrease it. The latter fact contradicted previous ideas about the influence of small molecules on AmB pores via membrane curvature stress, which should cause an increase in AmB-induced current. We noticed that “decreasing” flavones have an identical structural motif composed of a carbonyl group, which is linked to a fragment containing several conjugated double bonds.

Molecular dynamics simulation was employed to investigate the interaction of AmB and plant flavones (chrysarin, wogonin, baicalein, apigenin, scutellarein, luteolin, morin and fisetin) in membrane. GROMACS 2023.2 was used to perform simulations with CHARMM36m all-atom force field. Tested molecules were parametrized based on known 3D structure (NMR data for AmB [1]) using CGenFF. Model membranes were assembled in CHARMM-GUI Membrane Builder, contained 100 POPC, 50 ERG molecules and ionized with 2M KCl. One molecule of AmB was placed in the membrane in such a way that mycosamine sugar was located in the headgroup region while the polyene ring was inside the hydrophobic part of the bilayer. One molecule of flavone was placed in the center of the membrane at the distance of 12 Å away from AmB. The size of the simulation box was approximately $6.7 \times 6.7 \times 9.0$ nm. Energy minimization was followed by a six-step equilibration process with gradually turning off the position restraints on lipid molecules. Production simulations were performed for 100 ns at 25°C and 1 bar using the Nosé -Hoover and semi-isotropic pressure coupling approach with C-rescale barostat.

Possessing several polar groups, it was energetically unfavorable for the flavone molecules to stay in hydrophobic lipid core, thus, they tended to move up to the polar region of membrane. Two of the tested flavones, morin that increased AmB activity and chrysarin that decreased it, managed to bind to antibiotic under such conditions. The difference between them was quite remarkable: chrysarin occupied the polyene chain of AmB while morin stayed in connection with hydroxyl chain. The polyene chain of AmB is crucial for interaction with ergosterol and pore formation (and we observed replacement of ergosterol by chrysarin). Thus, chrysarin could prevent pore assembling and lead subsequent decrease in AmB-induced current. On the other hand, morin had almost no interaction with the polyene chain. Unexpected interaction of morin and polar chain of AmB in the bilayer might provide a field for speculation about the alternative mechanism of potentiating effect: morin might stabilize the pore by binding to the polar part of the AmB molecule that is not involved in its interaction with sterol. The obtained results shed some light to AmB pore functioning and could empower future drug development.

This study was supported by the Russian Foundation of Science No. 22-74-10023.

1. Umegawa, Y., Yamamoto, T., Dixit, M., et al. (2022) *Science Advances*, 8(24), eab02658.

IN-SILICO DESIGN AND STUDY OF NOVEL MANNICH BASE DERIVATIVES AGAINST THE SELECTED TARGETS FOR ITS ANTIMICROBIAL ACTIVITY

G.A. Mansi Devi, A. Surega

Harsha college of pharmacy, Bengaluru, Karnataka, India

During the past few decades, the treatment of infectious diseases is increasingly challenging and crucial due to the accumulation of many problems related to pathogens developing resistance against antimicrobial drugs. So, to overcome this there is a strong push to discover and develop the new antimicrobial drugs. The objective of the study focuses on designing novel compounds of Mannich base to combat antimicrobial resistance. This study describes detailed information regarding the targets available for antimicrobial activity. From that number of proteins, 3 proteins have been selected namely Aminoglycoside Phosphotransferase (3TYK), InhA - Enoyl ACP reductase (6SQ5), MurE UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2,6-diaminopimelate ligase, (8G6P) based on different mechanism which are considered to be a new target for the treatment of bacterial infections. With the hope that focusing on this target will enhance the effectiveness of the drug in overcoming resistance. Mannich bases are a class of compounds with well-known antimicrobial properties, So, this work attempts to identify and design a novel molecule for bacterial infections. A detailed literature survey was done on different pharmacophores which are reported as inhibitors of the selected targets that are considered for these studies. Pharmacophores are important and they define the key features necessary for the compounds to interact and bind with its biological target. Based on those pharmacophores, 10 compounds have been designed for mannich base. The best compounds were selected based on molecular docking studies and ADMET (Absorption, Distribution, Metabolism, Elimination and Toxicity) analysis by using Simulation Plus software to predict the interaction and safety profile of the compounds. The selected compounds were synthesized, and they were characterized by using IR (Infrared Spectroscopy), NMR (Nuclear Magnetic Resonance) and Mass Spectrometry and the synthesized compounds are evaluated for antimicrobial activity.

The results of our study would include the identification of promising Mannich base compound with potential antimicrobial activity characterized by favourable interaction profiles and ADMET properties by improving their chances of successful introduction and development efforts are required to bring these promising novel compound to the market. The combination of advanced computational methods, synthesis, and rigorous testing provides a strong foundation for developing new and effective antimicrobial agents. Ongoing research and development is essential to translate these findings into viable therapeutic options.

NEW CATALYSTS FOR THE PRODUCTION OF 3,4-DIHYDROPYRIMIDINE-2(1H)-THIONES BASED ON SALICYL ALDEHYDES AND *IN SILICO* ANALYSIS OF BIOACTIVITY OF OBTAINED COMPOUNDS

I.V. Mineyeva^{1,2}

¹*Faculty of Chemistry, Belarusian State University, Minsk, Belarus*

²*Research Institute for Physical Chemical Problems, Belarusian State University, Minsk, Belarus*

Multifunctional 3,4-dihydropyrimidine-2(1H)-thiones have a diverse spectrum of action and are actively used medically as antibacterial, antiviral, antitumor, analgesic and anti-inflammatory, antiaggregation, and antihypertensive agents.

Using catalysts based on salts of rare earth metals, approaches have been developed to involve labile salicylic aldehydes **1a-d** in the Biginelli multicomponent reaction (Fig. 1).

The analysis of pharmacological activity of **4a-d** was carried out using computer technologies *in silico* on the PassOnline platform (<http://way2drug.com/passonline/predict.php>). Calculations of acute rat toxicity using GUSAR, rodent organ-specific carcinogenicity using ROSC-Pred, *in silico* prediction of adverse drug effects ADVERPred, prediction of interaction with the undesirable targets (antitarget prediction), prediction of drug-induced changes of gene expression profile using DIGEP-Pred, prediction of interaction of pharmacological substances with human kinome using KinScreen, prediction of toxicity taking into account the metabolism of drug using MetaTox, prediction of interaction with molecular targets, prediction of substrate/metabolite specificity were performed. Possible effective interactions with molecular targets are calculated and predictions of various types of activity are made. A prediction of interaction with tumor and non-tumor cell lines has been made.

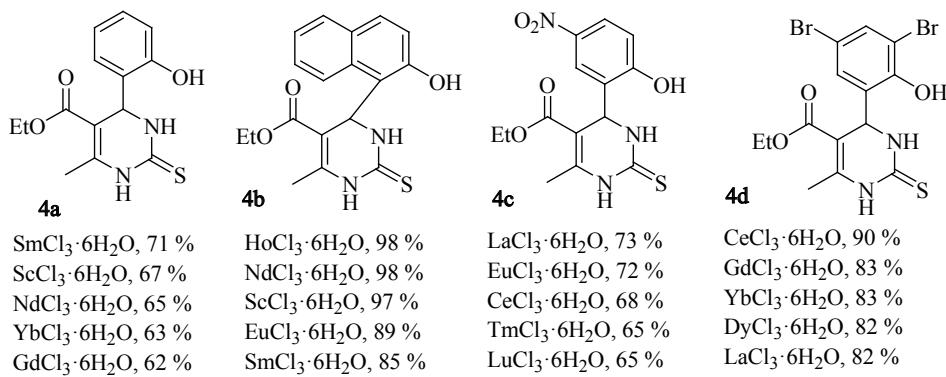
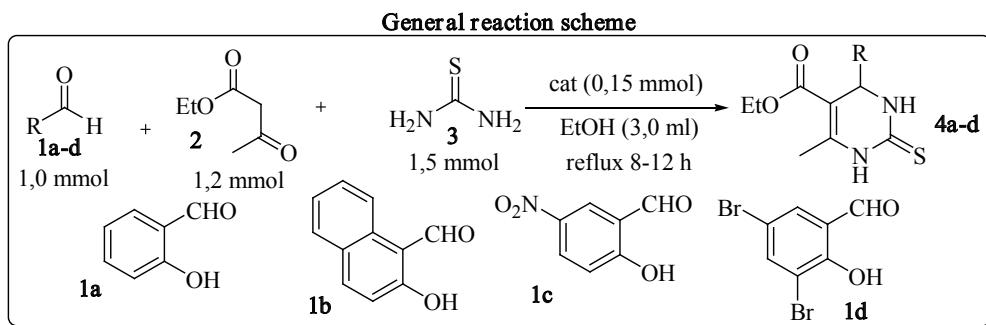


Fig. 1. Structures and yield of all new compounds **4a-d**

The results obtained indicate the possibility of obtaining new heterocyclic for the purpose of obtaining new molecular tools and drug prototypes.

This work was supported by the SPSR (Belarus) no. 20240340.

RARE EARTH CHLORIDES AS NEW CATALYSTS FOR THE MULTICOMPONENT BIGINELLI REACTION WITH PARTICIPATION 5-AMINOTETRAZOLE AND *IN SILICO* ANALYSIS OF BIOACTIVITY OF OBTAINED COMPOUNDS

I.V. Mineyeva^{1,2}

¹Faculty of Chemistry, Belarusian State University, Minsk, Belarus

²Research Institute for Physical Chemical Problems, Belarusian State University, Minsk, Belarus

Pyrimidines belong to a privileged class of heterocyclic scaffolds that are known to have a wide range of biological and pharmacological activities. In addition, tetrazoles are an important subunit of many natural and synthetic compounds exhibiting biological activity. Tetrazoles fused with pyrimidines are known to have a wide range of biological properties, including antimicrobial, antidepressant and antituberculosis activity. New derivatives of 3,4-dihydropyrimidine **1-6** were obtained based on the multicomponent Biginelli reaction with the participation of 5-aminotetrazole (**7**), aldehydes **8**, **9**, 1,3-dicarbonyl compounds **10-12** and salts of rare earth metals as catalysts for this transformation upon boiling in ethanol (Fig. 1).

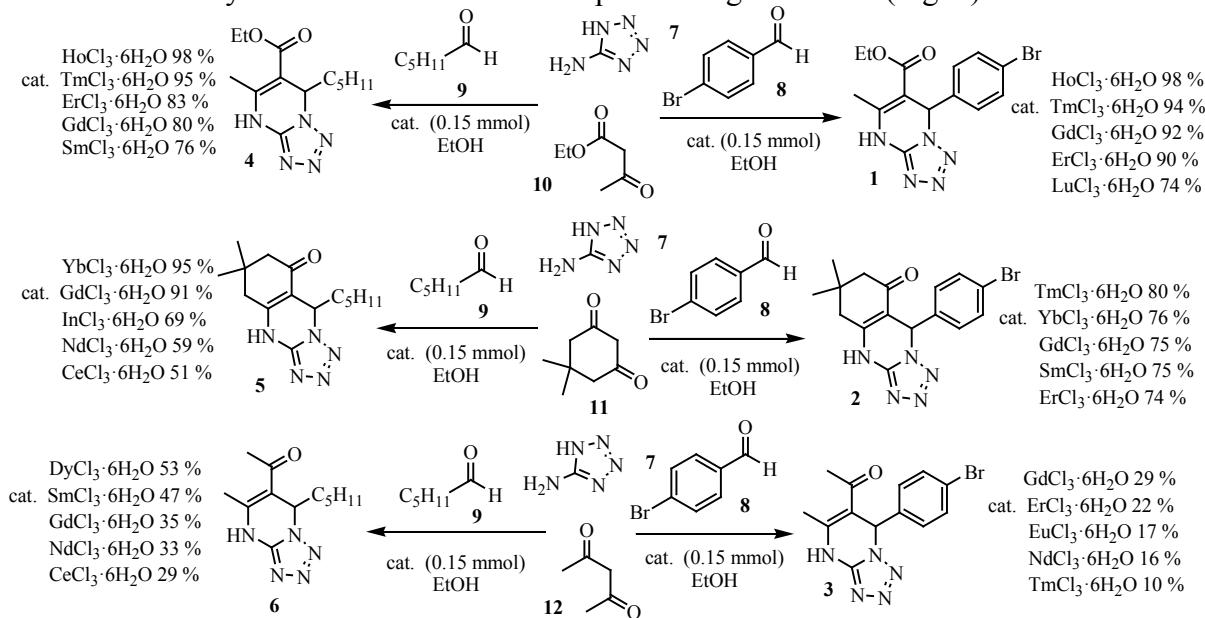


Fig. 1. Structures and yield of all new compounds **1-6**

Analysis of the pharmacological activity of the obtained compounds was carried out using *in silico* computer technologies on the PassOnline, SwissADME, ProTox-II platform and the PerMM service.

Compounds **1-6** may be in demand as new ligands for the preparation of transition metal complexes; their modification at the tetrazole fragment is possible, which makes it a potentially promising building block for pharmaceutical and biomedical research.

This work was supported by the SPSR (Belarus) no. 20240340.

CHLORIDES AND TRIFLATES OF METAL IN ETHANOL AS NEW CATALYSTS OF THE MULTICOMPONENT BIGINELLI REACTION WITH THE PARTICIPATION OF FURFURAL

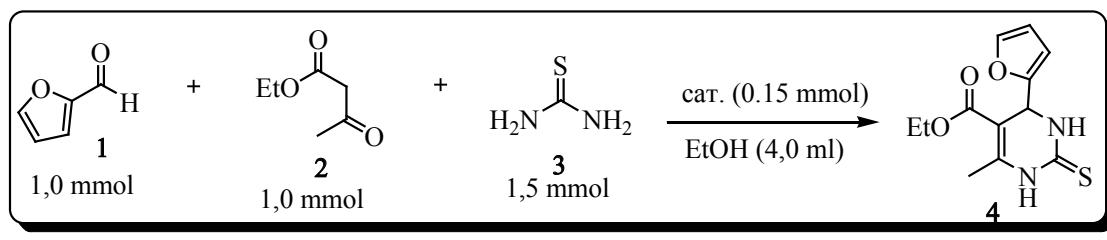
I.V. Mineyeva^{1,2}

¹Faculty of Chemistry, Belarusian State University, Belarus

²Research Institute for Physical Chemical Problems, Belarusian State University, Minsk, Belarus

New data have been obtained on the use of catalysts based on rare earth metals and some metal triflates in the multicomponent Biginelli reaction with the participation of labile furfural aldehyde (**1**), ethyl acetoacetate (**2**), thiourea (**3**) when boiled in ethanol (Figure 1).

The analysis of pharmacological activity of **4a** was carried out using computer technologies *in silico* on the PassOnline platform (<http://way2drug.com/passonline/predict.php>).



InCl ₃ ·6H ₂ O	12,6 %
YbCl ₃ ·6H ₂ O	34,0 %
ScCl ₃ ·6H ₂ O	36,7 %
ErCl ₃ ·6H ₂ O	37,1 %
CeCl ₃ ·6H ₂ O	39,1 %
YCl ₃ ·6H ₂ O	42,9 %
EuCl ₃ ·6H ₂ O	43,2 %

DyCl ₃ ·6H ₂ O	44,4 %
NdCl ₃ ·6H ₂ O	50,1 %
LaCl ₃ ·6H ₂ O	52,3 %
SmCl ₃ ·6H ₂ O	55,1 %
GdCl ₃ ·6H ₂ O	61,1 %
HoCl ₃ ·6H ₂ O	90,8 %
TmCl ₃ ·6H ₂ O	93,4 %

List of proteins predicted as possible direct targets

Polyadenylate-binding protein 1	0,7806
Eukaryotic translation initiation factor 4H	0,6923
Cytochrome P450 2J2	0,5822
Protein tyrosine kinase 2 beta	0,5204
Nuclear receptor subfamily 4 group A member 1	0,4884
Plectin	0,4491
Tyrosyl-DNA phosphodiesterase 1	0,4221
Microtubule-associated protein tau	0,4218
Serine/threonine-protein kinase NEK6	0,4204

Fig. 1. Synthesis of product **4** and analysis of its potential biological properties

Compound **4** may be in demand as a new ligand for the preparation of transition metal complexes; modification at reaction centers is also possible, which makes it a potentially promising building block for pharmaceutical and biomedical research.

This work was supported by the SPSR (Belarus) no. 20240340.

DEVELOPING NOVEL GLIPTIN ANALOGS WITH A PIPERAZINE CORE FOR IMPROVED DPP-4 INHIBITION IN THE TREATMENT OF TYPE 2 DIABETES MELLITUS: A COMPUTATIONAL STUDY

T.K.S. Muhammed¹, C.S. Chandran¹, B. Valliyot², K.M. Ammu¹, O. Archana¹, K.P. Adithya¹, A. Aravind¹

¹College of Pharmaceutical Sciences, Government Medical College, Kannur, Kerala, India

²Department of Medicine, Government Medical College, Kannur, Kerala, India

Diabetes Mellitus Type 2 (T2DM) is an increasingly prevalent metabolic disease characterized by decreased insulin secretion and resistance. The FDA has authorized five types of drugs to control it: thiazolidinediones, biguanides, meglitinides, sulfonylureas, α -glucosidase inhibitors, and dipeptidyl peptidase-4 (DPP-4) inhibitors. Owing to the side effects of current medications, researchers are investigating novel DPP-4 inhibitors (gliptins) for their improved safety profile. This study aimed to develop gliptin analogs with a piperazine core as DPP-4 inhibitors utilizing the Schrödinger Suite 2022-1. This suite integrates field-based QSAR, molecular docking, and molecular dynamics simulations. The study used forty-five piperazine derivatives with IC₅₀ values between 0.43 nM and 455 nM to determine spatial characteristics through Gaussian field-based QSAR, yielding promising statistical results. Contour maps revealed both advantageous and disadvantageous interactions around the piperazine core, leading to the development of 10 new compounds with enhanced predicted IC₅₀ values, strong binding affinities, and docking scores. Among these, Compound G86 showed superior binding affinity and inhibitory activity against DPP-4 compared with sitagliptin, suggesting its potential for further investigation. The QikProp analysis predicted favorable pharmacokinetic and physicochemical properties and a better safety profile for G86. Additionally, molecular dynamics simulations indicated that the G86-1X70 complex was more stable than the sitagliptin-1X70 complex (Fig. 1). This study highlights the effectiveness of computational methods in accelerating the discovery of anti-diabetic drugs and sets the stage for future research and therapeutic advancements in managing T2DM.

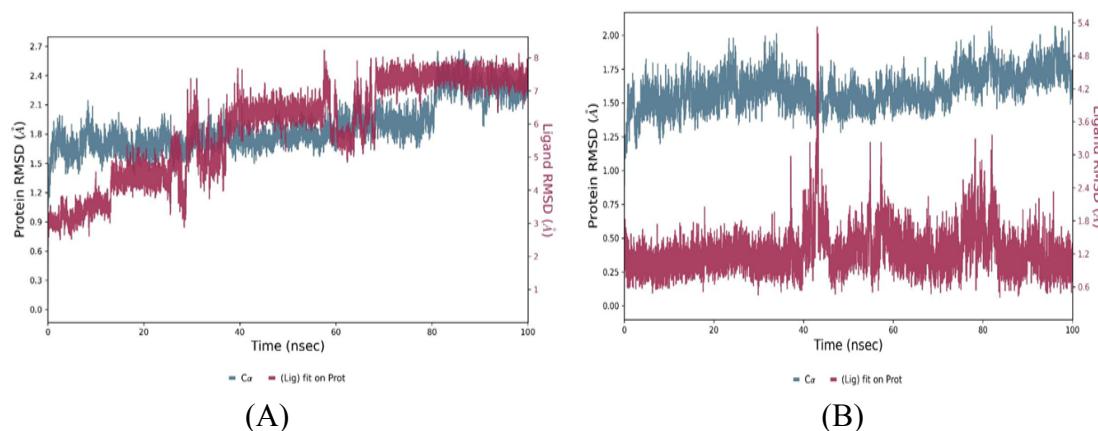


Fig. 1. Protein –Ligand RMSD of (A) G86-1X70 complex, (B) A44 (Reference)-1X70 complex.

DETERMINATION OF SUBSTRATE ACTIVATION IN ACTIVE SITES OF HYDROLASES USING NEURAL NETWORKS

T.I. Mulashkina^{1,2}, M.A. Krinitksiy^{1,3,4}, M.G. Khrenova^{1,2,5}

¹*Moscow State University, Moscow, Russia*

²*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia*

³*Moscow Institute of Physics and Technology, Moscow, Russia*

⁴*Shirshov Institute of Oceanology of the Russian Academy of Sciences, Moscow, Russia*

⁵*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia*

For enzyme-catalyzed chemical reactions to occur, the substrate must be activated at the enzymes active site. This can be determined by analyzing Laplacian electron density maps in the reaction region. These maps show deconcentration of electrons near the electrophile in the direction of nucleophile attack for reactive molecules and concentration of electrons for nonreactive molecules. By studying molecular dynamics trajectories calculated using quantum mechanics and molecular mechanics methods (QM/MM), the efficiency of substrate activation can be estimated quantitatively from the ratio between reactive and nonreactive states. A typical QM/MM trajectory is tens of picoseconds long corresponds to tens of thousands of frames in which Laplacian maps. Manually processing such a volume of information is difficult. So, in this work, we developed a convolutional neural network based on the ResNet50 architecture. This network is capable of distinguishing between reactive and non-reactive enzyme-substrate complexes.

The neural network was trained on images of Laplacian electron density maps calculated for enzyme-substrate complexes of the main protease of SARS-CoV-2 with three different substrates. During training and validation, artificial data augmentation was used to increase invariance to rotations and translations of the neural network. To evaluate its performance, sets of Laplace electron density maps were randomly selected from QM/MM MD trajectories of bacterial metallo-beta-lactamase NDM-1 complexes and antibiotic imipenem as well as capralactam-lipase CALB complexes. The training and test data were balanced by the target feature, the reaction state or its absence.

Based on the training results, the proposed neural network distinguishes between reactive and non-reactive states in all data sets with a correct answer rate exceeding 97%. The difference between the training and test samples does not exceed one percentage point.

The attention field of the neural network was also analyzed. This analysis showed that the key factor in determining the reactivity of enzyme-substrate complexes is indeed deconcentration of electron density near electrophilic atoms.

This work was supported by the Scientific and Educational School of Moscow State University (project no. 23-Sh03-04).

EXPLORING INTERACTIONS OF ACETYLCHOLINESTERASE WITH STEROIDAL PREGNANES THROUGH MACHINE LEARNING AND ATOMISTIC SIMULATION

O.M. Ogunyemi¹, O.S Macaulay¹, G.A Gyebi², M.M. Ogunyemi³, M.O Agunloye⁴, C.O. Olaiya¹

¹Biomolecular Structure Group, Nutritional and Industrial Biochemistry Research Unit, Department of Biochemistry, College of Medicine, University of Ibadan, Ibadan, Nigeria

²Department of Biochemistry, Faculty of Science and Technology Bingham University, Karu, Nasarawa, Nigeria

³Africa Centre of Excellence for Mycotoxin and Food Safety, Federal University of Technology Minna, Minna, Nigeria

⁴Department of Physiology, Faculty of Biomedical Sciences, Kampala International University, Kansanga, Uganda

Acetylcholinesterase (AChE), an enzyme critical for the regulation of neurotransmission, is one of the most important drug targets in Alzheimer's disease (AD). The current FDA-approved drugs including donepezil, galantamine and rivastigmine are able to reduce symptoms and delay progression to certain extents in early-stage AD, however, they show little results in advanced AD cases. Besides, these therapeutic agents are associated undesirable effects including non-target-specificity, poor bioavailability and high hepatotoxicity tendency. Thus, screening for more efficient AChE inhibitors is an overarching goal in AD drug discovery. Previously, we reported the *in vitro* AChE inhibitory potential of isolated pregnanes from plant extracts. Further exploration of this class of compounds may help to expand the chemical space coverage of AChE in search for more efficient inhibitors. A trained and validated Random Forest Regressor (RFR)-based machine learning algorithm predicted 843 pregnanes with $\text{pIC}_{50} \geq 5$ from the initial 1,583 pregnane compounds. A dense distribution of the compounds was observed within the space of molecular weight around 500 Da and within the space of LogP around 5. Inhibition constant prediction then suggested 67 compounds with $\text{pKi} \geq 7$. Structure-based virtual screening revealed several AChE binders including 21-[(3-Hydroxy-2-naphthyl)oxy]pregnane-2-one which made strong molecular contacts with the catalytic active site, peripheral anionic site (PAS), oxyanion hole and anionic sub-site of the AChE binding gorge through multiple hydrogen bonds and hydrophobic interactions. The resulting AChE-pregnane complexes exhibited structural stability and conformational flexibility as indicated by the dynamic parameters obtained from 100 ns full atomistic molecular dynamic simulation. Overall, the identified pregnanes are promising AChE inhibitors subject to experimental validation.

IN SILICO PREDICTION OF ANTI VIRAL ACTIVITY OF 1-((1R, 3R, 5R, 7R) ADAMANTAN-2-YL)-3,5-DIMETHYL-1H-PYRAZOL-4-YL)-1H-TETRAZOLE AND (1S, 3S, 5S, 7S-1,3-DI(1H-TETRAZOL-1-YL)ADAMANTAN

M. A. Gureev¹, I. S. Ershov^{2,3}, E.D. Fatuev², V.A. Ostrovskii^{2,3}

¹HSE University, Saint Petersburg, Russia

² Saint Petersburg State Institute of Technology (Technical University), St. Petersburg, Russia

³ Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg, Russia

Tetrazolyladamantanes are among the promising drug candidates for influenza virus therapy [1]. Recently, we synthesized two new derivatives of the tetrazolyladamantanes series: 1-((1R, 3R, 5R, 7R)adamantan-2-yl)-3,5-dimethyl-1H-pyrazol-4-yl)-1H-tetrazole (HED001) and (1S, 3S, 5S, 7S-1,3-di(1H-tetrazol-1-yl) adamantan (HED006). In this work, we evaluated the potential antiviral activity of compounds HED001 and HED006.

The most popular proteins related to the influenza A/H1N1 virus were used for calculations: hemagglutinin (8VQN), nucleoprotein (3TG6), N-terminal domain of the acidic protein polymerase (5FDG), basic protein polymerase 2 (4P1U), neuraminidase (3TI6) and M2 channel (6BMZ). To assess the binding of small molecules, the molecular docking method based on the GlideScore scoring function was used. The docking area is oriented to the coordinates of the centroid of the reference ligand (if it is present in the PDB file), which is also a source of information on the pharmacophoric characteristics of the active substances. In the absence of a control compound, literature data are used. For each compound, 15 folds in the active cavity of the protein were generated. The following is estimated: the value of the evaluation function (the lower the value, the better the result), the cluster ability of docking solutions, and their coincidence with the arrangement of active substances associated with the protein under study, specified in the PDB file. The calculation results are given in the table.

Table. Values of the evaluation function with the clustering parameter of docking solutions (** - high, ** - medium, * - low)

GlideScore@Protein (Kcal/mol)	HMG	NP	PA	PB2	NA	M2
HED001	-4.92(*)	-6.01(*)	-3.28 (*)	-5.61 (**)	-3.42	-8.34 (***)
HED006	-4.41 (*)	-6.09 (***)	-4.34 (*)	-5.21 (***)	-4.60 (*)	-7.41 (***)

The most interesting target among all for the compounds under consideration is the M2 channel. The ligands bind identically to the control compound (spiroadamantane) and completely reproduce its pharmacophoric characteristics. The adamantyl core provides lipophilic contacts in the center of the active cavity of the channel, interacting with Ala30, Val27 in chains A-D. Compound HED001, due to a higher density of lipophilic contacts, has better affinity, which is largely provided by the methyl groups of the pyrazole ring. Polar contacts of tetrazole with histidines (His37 in chains A-D) also enhance affinity for the M2 channel. Compound HED006 binds somewhat worse, however, in general, it also reproduces the lipophilic contacts of the reference compound, and with a lower density of contacts than HED001.

This research was funded by the Russian Science Foundation, grant number 23-13-00224.

1. Kiselev, O.I. (2012) Chemo Drugs and Influenza Chemotherapy, Rostok: St. Petersburg, Russia,

VIRTUAL AND EXPERIMENTAL SCREENING OF SUBSTANCES WITH ANXIOLYTIC ACTIVITY USING A CONVOLUTIONAL NEURAL NETWORK MODEL BASED ON MULTIPLE DOCKING

M.A. Perfilev¹, P.M. Vassiliev¹, A.V. Golubeva¹, A.N. Kochetkov¹, D.V. Maltsev¹, A.S. Tarasov¹, K.R. Magomedova¹, M.E. Blohin², O.A. Luzina²

¹*Volgograd State Medical University, Volgograd, Russia*

²*Novosibirsk Institute of Organic Chemistry named after N.N. Vorozhtsov, Novosibirsk, Russia*

Previously, according to the developed multiple docking methodology [1], we created a classification neural network model for predicting the anxiolytic activity of chemical compounds based on the correlation convolution of the energy spectra of their multitarget docking in 27 spaces of each of the 22 relevant biotargets. To train the neural networks, we used information from a verified database on the chemical structure and level of anxiolytic activity of known compounds [2]. The neural networks were trained in the Statistica program [3]. The best neural network model has the MLP 22-14-2 architecture (tanh, softmax), the accuracy rates of the model $Acc = 90.7\%$, $Sens = 90.9\%$, $Spec = 90.5\%$, $AUC_{ROC} = 91.5\%$. The aim of this work is to demonstrate the applicability of the developed methodology for searching for new multitarget pharmacologically active compounds using convolutional neural network models based on multiple docking using the example of screening anxiolytic substances. First, using the found model, we performed a virtual screening of five new derivatives of natural compounds. According to the obtained results, three compounds out of five probably have a pronounced anxiolytic effect. Then, as a primary stage of the experimental study, we conducted a basic behavioral test “light/dark transition” on white mature male mice weighing 20-30 g. We recorded the number of transitions between the chambers of the installation, as well as the time the animal spent in the light chamber. The studied substances were administered intragastrically with an atraumatic probe 30 minutes before the experiment at a dose equimolar to 1 mg/kg of the comparison drug (diazepam). We found that two compounds out of three predicted as having pronounced anxiolytic activity statistically significantly differ from the control (according to the nonparametric Kruskall-Wallis criterion with Dunn’s posttest, $p < 0.05$). Thus, the results of the computer forecast and experimental testing show good agreement, the search accuracy was 66.7%. A more detailed experimental study of the active substances found is recommended, as well as further development of the methodology for constructing convolutional neural network models based on multiple docking.

The work was carried out within the framework of the state assignment of the Ministry of Health of the Russian Federation No. 23022400009-9 “Development of a methodology for computer search for multitarget pharmacologically active compounds based on multiple docking and convolutional neural network technology of various architectures.”

1. Vassiliev, P.M., Kochetkov, A.N., Spasov, A.A., Perfilev, M.A. (2021) *Volgograd Scientific and Medical Journal*, 3, 57-61.
2. Vassiliev, P.M., Maltsev, D.V., Perfilev, M.A. et al. (2022) *Certificate of State Registration of the Database RU2022621744*.
3. Hilbe, J.M. (2007) *The American Statistician*, 61(1), 91-94.

COMPUTER SCREENING OF BIOLOGICAL ACTIVITY OF AZOLE AND ANILINE DERIVATIVES

E.A. Petukhova, E.A. Danilova

Ivanovo State University of Chemistry and Technology, Ivanovo, Russia

Mortality from malignant neoplasms is gradually increasing worldwide. For example, in 2020, 18.1 million cases of cancer were reported. Nowadays, the general methods of treating cancer are chemotherapy and radiation therapy, as well as radical surgery, which affects not only the tumor tissue, but also the healthy cells adjacent to the tumor. An important task is to develop the affordable and less toxic drugs in comparison with those used to treat cancer patients. The development of new compounds showing antitumor effect is difficult, expensive and time-consuming process. In this regard, the use of modern computer programs, which make it possible, based on the structural formula of a compound, to predict the probability of a molecule exhibiting certain properties, becomes more relevant.

The purpose of this work is to investigate the range of biological activity (including cytotoxicity) *in silico* compounds potentially exhibiting antitumor activity (Fig. 1) using PASS and CLC-Pred computer programs.

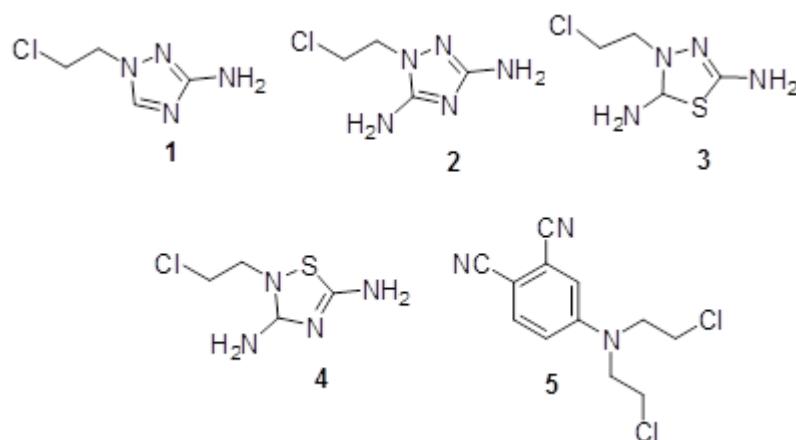


Fig.1. Structural formulas of the studied molecules

The prediction of the spectrum of biological activity showed that the molecules studied (**1-5**) can exhibit antitumor properties with a probability of more than 90 %.

The cytotoxicity prediction results showed that molecules **1** and **2** had a 75 % and 86 % probability of being potentially cytotoxic to colon carcinoma. Compounds **3**, **4** and **5** with a probability of 53 %, 54 %, and 65 %, respectively, are hypothetically capable of exhibiting cytotoxic properties against skin melanoma, brain glioblastoma, and prostate carcinoma, respectively.

Therefore, computer prediction of biological properties allowed us to assume that the synthesis of such molecules would be appropriate, and the resulting compounds could potentially be used as active pharmaceutical substances in the synthesis of drugs for the treatment of oncological diseases.

DRUG METABOLISM PREDICTION USING GRAPH NEURAL NETWORKS

N. Polomoshnov^{1,2}, A. Rudik¹

¹*Institute of Biomedical Chemistry, Moscow, Russia*

²*Moscow State University, Moscow, Russia*

In the current drug development process, an integral part of testing candidate molecules is the prediction of their metabolism, since biotransformations in the human body can significantly affect the toxicity and efficacy of a compound. Currently, there are several metabolism prediction programs such as MetaTox[1], BioTransformer[2], XenoSite[3], etc., but many of them either predict only possible sites of metabolism, or predict an excessive number of metabolites, or have reduced learnability due to complicated initial structure - for example, they split metabolism into two successive steps ("oxidation" and "conjugation"), thus covering not all possible reactions, or, like SyGMA[4], have no learnable part at all, working only with predefined algorithms.

In addition, most models rely on the representation of molecules as descriptors (SMILES [8], MNA [6] and others) or fingerprints such as Morgan's fingerprint [5]. Descriptors reflect the two-dimensional structure of the molecular graph only implicitly, which complicates the model training process, and fingerprints are often not trained for a specific problem, which risks potentially under-representing the information about the compound needed for the problem.

We present a two-part model architecture for predicting the metabolism of drug compounds. The first part is based on biotransformation rules selected from open sources. The second part is a classifier based on graph attention networks (GATs [7]), forming embeddings for all inferred substrate-metabolite pairs and selecting the inferred correct ones.

Xenobiotic biotransformation data combined from DrugBank, ChEMBL, MetXBioDB and Accelrys Metabolite databases and transformation rules from MetXBioDB, SyGMA and GLORY [10] databases were used to train the model. Substrate-metabolite pairs were represented as molecular graphs with vector representations of nodes and edges. The hyperparameters for the neural network were selected using the Optuna [9] framework implementing Tree-structured Parzen Estimator [11]. As a result, the model showed good results on test metabolic maps, which was evaluated in different ways: coverage of individual metabolic maps at the generation stage, as well as using the MCC metric to evaluate the quality of the classifier - this metric was used because of its robustness to class imbalance [12]. It can be stated that the model is applicable to real medical chemistry problems.

1. Rudik, A.V., Bezhentsev, V.M., Dmitriev, A.V., Druzhilovskiy, D.S., Lagunin, A.A., Filimonov, D.A., Poroikov, V.V. (2017) *Journal of Chemical Information and Modeling*, 4, 57.
2. Djoumbou-Feunang, Y., Fiamoncini, J., Gil-de-la-Fuente, A., Greiner, R., Manach, C., Wishart ,D.S. (2019) *Journal of Cheminformatics*, 2, 11.
3. Zaretzki, J., Matlock, M., Swamidass, S.J. (2013) *Journal of Chemical Information and Modeling*, 12, 53.
4. Ridder, L., Wagener, M.(2008) *ChemMedChem*, 3.
5. Morgan, H.L. (1965) *Journal of Chemical Information and Modeling*, 2, 5.
6. Filimonov, D., Poroikov, V., Borodina, Y., Gloriozova, T. (1999) *J. Chem. Inf. Comput. Sci.*, 39.
7. Velickovic, P., et. al. (2017) *arXiv:1710.10903*.
8. Weininger, D. (1988) *J. Chem. Inf. Comput. Sci.*, 1, 28.
9. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M. (2019) *KDD*
10. de Bruyn Kops, C., Šícho, M., Mazzolari, A., Kirchmair, J. (2021) *Chemical Research in Toxicology*, 2, 34.
11. Shuhei Watanabe (2023) *arXiv:2304.11127*
12. Chicco, D., Jurman, G (2023) *BioData Mining*, 4, 16

MATHEMATICAL MODELING OF CAR-T IMMUNOTHERAPY IN MULTIPLE MYELOMA

A. Punko¹, A. Volkova²

¹*Higher School of Economics, Moscow, Russia*

²*Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow, Russia*

Despite the high efficacy of CAR-T immunotherapy in the treatment of patients with multiple myeloma, the duration of response remains a therapeutic challenge. Efforts are underway to increase CAR-T cell persistence to improve treatment outcomes. We found clinical trial results with 15 CAR-T cell designs that differed in affinity and avidity binding. Mathematical modelling of CAR-T immunotherapy in multiple myeloma was performed to compare the efficacy of the developed CAR designs. A systematic review of mathematical models was conducted to select a base model, and in vitro data from 17 studies and clinical data from 26 studies were obtained for subsequent modelling. This study integrates key drug-specific and system-specific parameters responsible for CAR-T cell functionality in preclinical and clinical studies. Using a mathematical model, we were able to characterize (i) the destruction of multiple myeloma tumor cells by CAR-T cells in 13 cell lines with effector to target cell ratios ranging from 1:256 to 80:1, (ii) the multiphase profile of CAR-T cell pharmacokinetics including distribution, expansion, contraction and persistence phases, and (iii) the dynamics of clinical biomarkers in multiple myeloma patients. In addition, we compared the efficacy of CAR-T designs with different binding affinity and avidity. It is shown that the lower the value of the dissociation constant, the higher the expansion and persistence of CAR-T cells in the blood. BCMA and M-protein biomarker value decreases by 100% of the initial level at a value of 14 nanomoles or less. It is also shown that constructs that bind two epitopes of BCMA antigen have longer persistence. But with the observed values of dissociation constant and other parameters on biomarker profiles, no effect of changing the number of epitopes was found.

ANTIVIRAL ACTIVITY OF SOME NEW OXOCIN DERIVATIVES: *IN SILICO* STUDY AND *IN VITRO* RESULTS

**I.A. Pustolaikina¹, S.Sh. Normatov², R.A. Muzaparov¹, A.F. Kurmanova¹, A.S. Rakhimzhanova¹,
A.L. Stalinskaya³, I.V. Kulakov³**

¹Karaganda Buketov University, Karaganda, Kazakhstan;

²Infochemistry Scientific Center, ITMO University, Saint Petersburg, Russia;

³University of Tyumen, Tyumen, Russia

A set of new oxocins was synthesized at the University of Tyumen by the research group of Professor I.V. Kulakov [1-2]. Oxocins are compounds containing the oxocine ring, which is an eight-membered unsaturated aromatic ring with one oxygen and seven carbon atoms. These compounds are structural analogues of natural integrastatins, which have a wide range of biological activity, including antiviral properties. According to the World health statistics 2024 report such infectious diseases as human immunodeficiency (HIV) and coronavirus COVID-19 remain among the most dangerous, so it seemed interesting to test new synthetic compounds for their ability to prevent those using modern methods of *in silico* studies.

In this work, *in silico* study of the antiviral activity of 12 new oxocin derivatives against the HIV-1 (PDB ID: 3V81) and SARS-CoV-2 (PDB ID: 7AAP) viral proteins was carried out using molecular docking and molecular dynamics simulation techniques. Nevirapine and Molnupiravir were used as the reference drugs to determine the antiviral activity of HIV-1 and SARS-CoV-2, respectively.

To evaluate the antiviral activity of new oxocine derivatives, a molecular docking was performed using the Lamarckian genetic algorithm (LGA) with the help of AutoDock 4.2.6, MGLTools 1.5.7 and Discovery Studio 3 programs. It was shown, that ligand **6a** have the best affinity with HIV-1 (-12,44 kcal/mol) and SARS-CoV-2 (-8.51 kcal/mol) viral proteins compared to other oxocin derivatives and also compared to Nevirapine (-8.48 kcal/mol) and Molnupiravir (-5.79 kcal/mol) reference drugs. Ligand **6a** have also demonstrated the best anti-COVID activity in an *in vitro* study [2]. Interaction analysis for ligand **6a** with HIV-1 and SARS-CoV-2 viral proteins complexes showed the formation of more active bonds with receptors compared to other compounds.

The stability of ligand **6a** - HIV-1 (Figure, a) / SARS-CoV-2 (Figure, b) viral proteins complexes was confirmed by molecular dynamics simulation using VMD, NAMD, Grace software and the CHARMM-GUI website.

Performed *in silico* study of the antiviral activity of some new oxocin derivatives indicates that ligand **6a** (*Z*-3-(2-(2,5-dimethyl-11,12-dihydro-5H-5,11-epoxybenzo[7,8]oxocino[4,3-b]pyridin-3-yl)-2-oxoethylidene)-3,4-dihydroquinoxalin-2(1H)-one) having the best binding energies, inhibition constants, demonstrating active interaction and good RMSF, RMSD, SASA values. Compound **6a** is a potential agent against HIV-1 and SARS-CoV-2 viruses, so it can be recommended for further *in vivo* clinical trials.

This study was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP23488790).

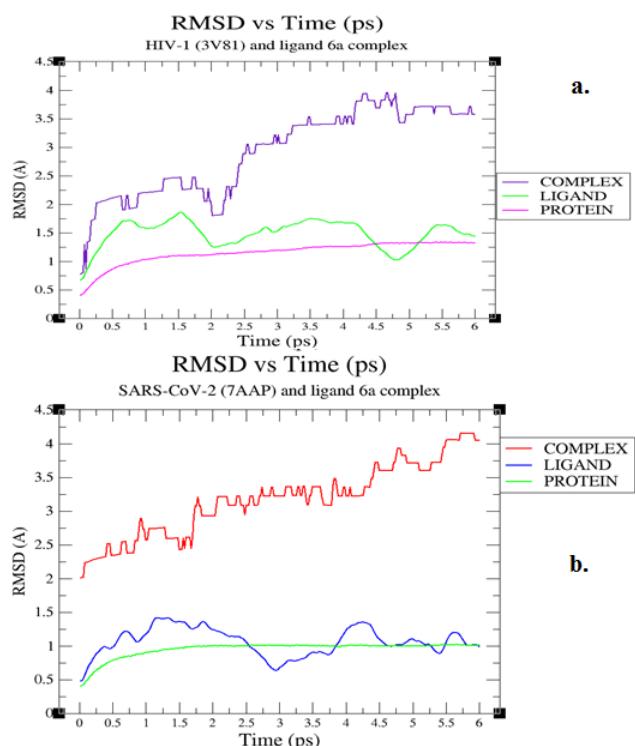


Fig. RMSD plot for ligand **6a** complexes

1. Kulakov, I.V., Stalinskaya, A.L., Chikunov, S.Y., & Gatilov, Y.V. (2021) *New Journal of Chemistry*, 45(7), 3559.
2. Stalinskaya, A.L., Martynenko, N.V., Shulgau, Z.T., Shustov, A.V., Keyer, V.V., Kulakov, I.V. (2022) *Molecules*, 27(12), 3701.

IDENTIFICATION OF PROTEOFORMS IN EXPERIMENTAL ISCHEMIC STROKE IN MICE: COMPARISON OF DATA FROM 2D ELECTROPHORESIS AND AN INDEPENDENT EXPERIMENT WITH MASS SPECTROMETRIC IDENTIFICATION

A.V. Rybina

Institute of Biomedical Chemistry, Moscow, Russia

Experimental data from two papers by different research groups, obtained in mice using techniques of autohypoxia-induced hypoxic preconditioning (HPC) and middle cerebral artery occlusion (MCAO)-induced cerebral ischemia, were analyzed. In the first study [1], 2D electrophoretic maps were obtained and protein identification was performed using the MALDI-TOF MS method. In the second, biological samples from different groups of mice were analyzed by LC-MS [2], with raw data available in the ProteomeXchange database (accession code PXD032141). Previously, an algorithm was proposed for the preliminary identification of protein proteoforms associated with post-translational modifications (PTMs) based on 2D electrophoresis data [3]. In this paper an attempt is made to prove its performance. The authors of the first paper identified 8 groups of spots on the 2D electrophoretic maps corresponding to 8 proteins with at least 2 proteoforms (which were not identified). The identification of peptides with PTMs was repeated using the original data from the second paper. Of the 8 target proteins, 7 were stably identified. For 3 of them, 6 possible modifications were found. 4 variants match the spots on the gel.

The work was performed within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (№ 122030100170-5).

1. Bu, X. et al. (2011) *J. Neurochem.*, 117(2), 346-356.
2. Ramiro, L. et al. (2023) *BMC Neurosci.*, 24(1), 7.
3. Rybina, A. (2023) *Biomedical Chemistry: Research and Methods*, 6(1), e00191.

POTENTIAL NEW INHIBITOR MOLECULES FOR SARS-COV-2 PLPRO: AN *IN SILICO* AND SYNTHETIC APPROACH

G. Sabadini¹, D. Cabezas¹, M. Mellao², J. Mella¹

¹*Organic Chemistry Laboratory, Institute of Chemistry and Biochemistry, Faculty of Sciences,
University of Valparaíso, Valparaíso, Chile*

²*Institute of Research and Postgraduate Studies, Faculty of Medicine and Health Sciences,
Central University of Chile, Santiago, Chile*

At the beginning of 2020, a new betacoronavirus was reported in the city of Wuhan, China, named 2019-nCoV. By mid-2023, approximately 800 million confirmed cases and nearly 7 million deaths related to this new coronavirus had been reported worldwide [1]. At the end of 2020, the first vaccine against SARS-CoV-2 was released, capable of inducing the formation of antibodies against the virus, making it a good strategy to prevent new infections. However, other short-term strategies are needed to treat patients with severe symptoms, necessitating the use of a specific drug.

The PLpro protease is a potential therapeutic target, and new strategies are proposed to develop inhibitors. PLpro is a non-structural protein of the virus that plays an important role in evading antiviral immune response mechanisms [2].

In this study, our objective was to obtain potential PLpro inhibitors using computational techniques such as QSAR, molecular docking, and molecular dynamics. We formulated a combined QSAR equation that used physicochemical and Free-Wilson descriptors. The r^2 , q^2 , and r^2 test values were 0.833, 0.770, and 0.721, respectively [3].

From the equation, we found that the presence of an aromatic and basic system is crucial for obtaining a good ligand. We identified two aminopyrimidine molecules with activities of 7.03 for Py3 and 8.08 for Py5 in pIC_{50} values. These molecules were synthesized in two steps: first, a chalcone was synthesized through the Claisen-Schmidt condensation reaction. The second step involved converting the chalcone to an aminopyrimidine structure via a cyclocondensation with guanidine in saturated KOH in ethanol.

QSAR have successfully identified promising PLpro inhibitors. The aminopyrimidine molecules synthesized through Claisen-Schmidt condensation and cyclocondensation with guanidine highlight the potential for therapeutic strategies to treat severe COVID-19 cases.

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. <https://covid19.who.int/?mapFilter=cases>.
2. Shin, D., Mukherjee, R., Grewe, D. et al. (2020) *Nature*, 587, 657–662 (2020).
3. Sabadini, G., Mellado, M., Morales, C., Mella, J. (2024) *Pharmaceuticals (Basel, Switzerland)*, 17(5), 606.

APPLICATION OF MDM2-RECRUITING PROTAC TO OVERCOME TRANSPORTER-MEDIATED CHEMORESISTANCE OF TUMOR CELLS

A. Sagaidak, T. Grigoreva, D. Novikova, V. Tribulovich

*Laboratory of Molecular Pharmacology, St. Petersburg State Institute of Technology
(Technical University), St. Petersburg, Russia*

Currently the approach of targeted protein degradation in proteasomes is gaining popularity as a way to modulate the stability of proteins that play a crucial role in the development of malignant diseases. The advantage of such an approach is the ability to suppress undruggable proteins. A chimeric molecule that directs the target protein to proteolysis, PROTAC (PROteolysis TArgeting Chimera), is created within this approach. PROTAC is a molecule that combines a ligand for the protein of interest and a ligand for E3 ubiquitin ligase connected by a linker. Forced ubiquitylation of the target by E3-ubiquitin ligase provides its proteasomal degradation [1].

It is proposed to use an approach similar to PROTAC to affect transmembrane proteins, in particular P-glycoprotein. The activity of ABC transporters, and primarily P-glycoprotein, in healthy cells provides the release of substances undesirable for cells; however, during therapeutic exposure it causes cell insensitivity to the treatment, in particular, leads to chemoresistance of tumor cells.

To date, the main way to overcome ABC transporter-mediated tumor resistance is the combined use of chemotherapeutics with small molecule compounds capable of inhibiting transporters. All three generations of P-glycoprotein inhibitors that have been proposed target the inner cavity of the transport protein, the site for substrate binding. At the same time, as it turned out, such compounds can both reduce and enhance activity of the transporters depending on the concentrations used. Therefore, it can be expected that the interaction with PROTAC will definitely suppress the functioning of the transporters, since it will not just inhibit the transport protein, but provoke its exit from the membrane and subsequent degradation.

The aim of this work was to test the applicability of the PROTAC concept to overcome ABC-mediated chemoresistance of tumor cells by the example of P-glycoprotein.

The interaction of PROTAC with P-glycoprotein was evaluated using models reflecting two different transporter configurations, as well as a model of E3-ubiquitin ligase MDM2. Using a combination of *in silico* methods, a number of sites for potential binding to small molecule ligands, as well as to the ubiquitin ligase, was identified within the P-glycoprotein. In particular, the site optimal for the PROTAC development and a new hit compound capable of binding to the transport protein and suppress its activity were revealed. In the case of the E3 ubiquitin ligase, it was shown that varying substituents in the ligase-binding fragment of a potential PROTAC can significantly increase its affinity to the ubiquitinylating enzyme.

This study was supported by the Russian Science Foundation (project no. 23-13-00344).

1. Zhang, D., Ma, B., Liu, D. et al. (2024) *EBioMedicine*, 105, 105212.

WORLD WIDE APPROVED DRUGS: FROM BIG BIOMEDICAL DATA TO GLOBAL SMALL MOLECULE DRUG DATABASE

P. Savosina, D. Druzhilovskiy, D. Filimonov, V. Poroikov

Institute of Biomedical Chemistry, Moscow, Russia

Despite the increase in the number of approved drug databases (DB) accessible via the World Wide Web, the information presented in those DB reflects only the knowledge about well-known pharmaceutical substances approved in a few countries. The majority of such DB contains the information on drugs approved by the FDA- and EMA, since the data from these medicines regulatory agencies are well standardized. Information on pharmaceuticals developed and used in other countries is lacking or unreliable. However, such knowledge is necessary for the development and application of computer-aided drug design methods because it not only enhance pharmacotherapeutic knowledge widening the covered chemical space, but also allow revealing previously unknown “structure-activity” relationships.

National drug registries can provide information on “locally” approved drugs. However, they are difficult to find because there is no single resource that provides access to them. Moreover, taking into account all the difficulties of processing, the combined data from national drug registries is an example of “big data”. There is no uniform processing algorithm to retrieving, collecting, and analyzing this data.

Recognizing the need for a single source of knowledge on approved medicines, regardless of their origin or country of use, we created the World Wide Approved Drugs (WWAD) database. The pharmacotherapeutic information on organic small molecule drugs has been extracted from the 71 national medicine registers and aggregated in WWAD. This makes it possible to identify the drugs used in one or two countries and significantly expand our knowledge of existing medicines. The official documents published by medicines regulatory authorities after the drug approval process provide the high accuracy of the aggregated information. An additional search of information about the collected drug was performed using scientific articles to expand the knowledge about the compound’s interactions with molecular targets and their pharmacological activities. Verification and standardization of chemical structures, unification of biological activity and therapeutic indication terms allow the information from WWAD to be used in computer-aided drug design without preprocessing.

The current version of WWAD contains information on over 4400 unique low molecular weight active pharmaceutical substances. Each compound is described in terms of its chemical structure, pharmacotherapeutic application, molecular targets, therapeutic indications and biological activities with links to data sources. The data on the national medicine register that granted the first approval for the therapeutic application and the corresponding date of the first approval are also presented. In addition, the information on the association of therapeutic agents with known three-dimensional complexes of their targets is included in WWAD. For a drug repurposing estimation, the results of prediction of biological activity spectra using PASS program are added to the database.

Our WWAD web resource is freely available for academic non-commercial utilization at <https://www.way2drug.com/wwad>. Information from WWAD database is presented in an interactive table with the ability to search globally or filter the data presented by a specific value. This web resource allows the study and analysis of data about small molecule pharmaceuticals regardless of their origin or approval country without the need to search in each national medicine register.

The WWAD database copyrights have been registered by Rospatent: No. 2024621221, 21.03.2024.

The study was performed in the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) (no. 122030100170-5).

INTERMOLECULAR INTERACTIONS OF 17.1 PEPTIDE WITH TNFR1 AND MTS-1 PROTEINS

K. Scherbakov^{1,2,3}, D. Yurkina³, E. Romanova³, R. Zigashin⁴, D. Yashin³, L. Sashchenko³

¹*Higher School of Economics, Saint-Petersburg, Russia*

²*Institute of Biomedical Chemistry, Moscow, Russia*

³*Institute of Gene Biology, Moscow, Russia*

⁴*Shemyakin & Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia*

Protein-protein (PPi) and protein-peptide interactions are the prospective and yet poorly discovered space of potential drug targets. Until recent times it was thought that PPi interfaces are undruggable due to their relatively flat surfaces. But nowadays there are cases of frontier studies where peptidomimetics with the desired biological activity are discovered.

The common example of PPis is the immune response that is preferentially regulated by receptors of cytokines, such as TNFR1. We have recently found that the fragment of Tag protein, designated by us 17.1, can interact both with the TNFR1 and Mts1 proteins. The MTS1-17.1 complex showed cytotoxicity against tumor cells that contained TNFR1 on their surfaces. To reveal intermolecular interactions that are responsible for the formation of the 17.1-Mts1 and 17.1-TNFR complexes with atomistic resolution, which may be useful for the development of the new anti-cancer peptidomimetics, we have performed molecular docking and molecular dynamics (MD) studies.

Docking data shows that 17.1 binds on the surface of the TNFR1 in an extended conformation. The obtained complex served as a starting point for a 100-ns MD run. The 17.1 peptide shows a high mobility on the surface of TNFR1 with RMSD values of backbone atoms up to 12 Å. The complex of 17.1 and TNFR1 adopts 4 major representative conformations. Despite the different conformations of the 17.1, we established some key residues of the peptide and the TNFR1 that are responsible for the complex formation in every representative conformation. These residues are Y144, R150, V152, R154, T155 in the case of the 17.1, and C56, S72, K75 R177, E79 for the TNFR1. So we conclude that despite the different conformations that peptide adopts, some key interactions preserve in every conformation considered.

According to docking data, two Mts1 subunits bind one 17.1 molecule in the long groove on the protein surface. MD calculations were also conducted for 17.1:Mts1 complex with a docking structure as a starting point. MD simulations show that a described groove enlarges during the MD run which allows it to fully accommodate the 17.1 peptide. The 17.1 peptide here forms the only dominant representative conformation. So only this conformation was further considered. The 17.1-Mts1 complex forms many strong and highly specific interactions such as H-bonds and salt bridges. The residues involved in these interactions are D151, V152, Q153, R154, T155 in the 17.1 and S44, F45, K57, A54 and R49 in the Mts-1.

So we revealed the key intermolecular interactions and residues responsible for the formation of TNFR-17.1 and Mts1-17.1 complexes. We believe that obtained data can serve as a starting point for a computer-aided drug design of the new class of anti-cancer peptidomimetics.

IDENTIFYING LEAD COMPOUNDS FOR RAF-1 INHIBITION : INTEGRATING AI WITH MOLECULAR DOCKING AND ADME PROFILING

N. Aggarwal, S. Sen, B. Bondhopadhyay

Immunoncology and Molecular Theragnostic Lab, Centre for Medical Biotechnology, Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida, India

The serine/threonine kinase C-Raf, also commonly referred to as Raf-1, forms one of the three isoforms of the Raf protein family. In general, the activity of Raf-1 has been linked to multiple physiological functions such as cell differentiation, proliferation and survival. It forms an important part of various cell signalling pathways, especially the mitogen-activated protein kinase (MAPK) pathway. Thus, any dysregulation in the expression of this protein is associated with many diseases, such as cancer. Specifically, in breast cancer, the anomalous activation of the Raf family of proteins, including B-Raf and Raf-1 have frequently been found to increase metastasis, angiogenesis and aid the uncontrolled proliferation of cancer cells, contributing to the growth and progression of the tumor. Hence, targeting Raf-proteins provides significant potential for the development of successful therapeutic interventions in order to inhibit the growth of the tumor, metastasis and oncogenic signalling by modulating the MAPK/ERK pathway.

At present, there is a severe lack of studies on the use of natural inhibitors of Raf-1, and at the same time, there are no drugs that specifically target the said protein. In general, Sorafenib and Regorafenib are the two drugs commonly used for inhibiting Raf-1, however, these were originally meant for the inhibition of the B-Raf protein, and hence may induce adverse effects due to off target effects. This lack of specificity and safety prompted the search for natural analogues of Sorafenib, that could provide better prospects in these areas. This was achieved by leveraging artificial intelligence algorithms like NNScore 2.0, that runs on neural networks and Deepchem which uses random forest classifiers, along with relevant *in silico* screening software like Glide in Schrodinger Suite 2023-1, to build a lead optimization and safety profile for the drug development of these natural analogues. This was done by evaluating their binding scores, by performing MMGBSA and by looking into their ADME/Tox profile. The reference ligand, Sorafenib, has a greater free binding energy (binding energy = -56.17 kcal/mol) but a lower docking score than the natural molecules in all the three tools utilized. Among the top four ligands in all the programs, ZINC000002148905 (binding energy = -56.66 kcal/mol) was the best natural analogue. In addition, ZINC1092713 and ZINC000002148905, with binding energies of -43.32 and -38.55 kcal/mol, respectively, were deemed to be the most promising candidates for the development of new therapies against breast cancer. Hence, these natural analogues provide a promising avenue for the further development of in-vitro and in-vivo analyses, such that successful therapeutic drugs can be introduced for the mitigation of breast cancer.

GENETIC INFORMATION COMPRESSION PROGRAM

V. Shylau¹, V. Litvyak²

¹*International Sakharov Environmental Institute of Belarussian State University,
Republic of Belarus*

²*Institute of Starch and Starch-containing Raw Materials Processing –
Branch of Russian Potato Research Centre, Russia*

We have developed a computer program for editing genetic information based on the principle of fractal compression of self-similar structures by translating the letters of the genetic alphabet into a binary digital code with subsequent reversible transformation. The program ensures the maximum degree of genome compression (up to 99%) while preserving the original information and protein-coding functions. Compression of genetic and amino acid sequences is a pressing problem for theoretical and applied areas of bioinformatics, molecular biology and medicine. It was found that classical text compression algorithms cannot lead to the desired result when working with genetic material. The computer program we have developed uses the principles of information recoding and sequential fractal compression [1-4]. The program algorithm is based on the principle of translating the letters of the genetic alphabet into a digital code with subsequent reversible transformation of self-similar structures. During the implementation of the genome editing program algorithm, the DNA text is converted into a binary digital code and broken down into large blocks of numbers. The program's task is to find all the large blocks and replace them with similar small ones, which are then replaced with letters of the genetic alphabet. If these actions are repeated many times, it is possible to significantly reduce the size of DNA while preserving the original information and protein-coding functions. The program's algorithm provides a high degree of DNA compression of up to 99% by detecting and encoding iteration factors. For example, the human insulin gene, originally containing 4044 nucleotides, was transformed by a computer program into a sequence consisting of only 32 nucleotides. (AACGTTAACTGGTTAATCTTCTCTAAGAGAA). The second example of the implementation of a genetic information compression program is the mRNA of the delta strain of severe acute respiratory syndrome coronavirus (SARS-CoV-2), in which the number of nucleotides, originally equal to 29821, is compressed to a size of 10 nucleotides (UCUUAAGAGA). The principle of genome editing resembles the method of writing an abstract on the text of a book, i.e. Miniaturized biosystems ("compressed" genes) retain the recursiveness inherent in the original source (genomic DNA), so a living organism can read the original genetic text and execute all the commands encoded in it. Genes "compressed" using the developed computer program can be used as a template for the synthesis of short peptides - informational analogues of hormones, enzymes, neurotransmitters and other biologically active molecules, as well as for genome editing using CRISPR technology. "Compressed genomic RNA" can become the basis for the synthesis of antigens from short peptides or mRNA in order to create vaccines against SARS-CoV-2, HIV, etc.

It should be emphasized that the developed computer program, based on one of the variants of the scientific hypothesis of fractal editing of genetic information, requires experimental verification of the biological effectiveness of the molecules synthesized with its help, which is expected to be carried out in the near future.

1. Barnsley, M.F. (2012) Fractals Everywhere: New Edition. Dover Publications.
2. Peitgen, H.-O., Èurgens, J., Saupe, D. (2006) Chaos and Fractals: New Frontiers of Science. Springer Science & Business Media.
3. Fisher, S. (2012) Fractal image compression: theory and application. Springer Science & Business Media.
4. Tkachuk, V. at al. (2018) GA-based Fractal and Neuroevolution Compressions Conference Paper.
<https://www.researchgate.net/publication/328007415>.

THE EFFECT OF POINT MUTATIONS ON THE STRUCTURE AND DYNAMICS OF THE TRANSMEMBRANE DOMAIN OF THE SARS-COV-2 SPIKE PROTEIN (ACCORDING TO THE COMPUTATIONAL EXPERIMENT)

P.A. Shilova^{1,2}, E.T. Aliper¹, R.G. Efremov^{1,3,4}

¹*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia*

²*Lomonosov Moscow State University, Moscow, Russia*

³*National Research University Higher School of Economics, Moscow, Russia*

⁴*Moscow Institute of Physics and Technology (State University), Dolgoprudny, Russia*

The SARS-CoV-2 spike (S) protein, a homotrimeric, membrane-anchored fusion protein, plays a critical role in viral infectivity by mediating host cell recognition and membrane fusion. Data suggest that the S protein's transmembrane domain (S-TMD), which remains incompletely understood, is functionally involved in viral fusion and virus viability. Therefore, elucidating the trimerization mechanism and possible S-TMD conformations within the viral envelope is essential for comprehending the molecular mechanism of S protein-mediated fusion.

S-TMD is formed by three α -helices arranged in a coiled-coil structure. The composition analysis of the side chains packed at the helix-helix interface, which mirrors interactions between S-TMD chains, is crucial to understanding the trimer's structural properties. By introducing point mutations into the trimer, we can assess the contribution of specific amino acid residues to the interface formation in each helix, gaining insight into the mechanisms of the S-TMD trimerization.

In the current study, we investigate the role of 10 aminoacidic residues in the stability of the helix-helix interfaces based on the model of S-TMD (S_OPT) constructed by Aliper et al. [1]. The results of the analysis were used to predict amino acid positions which are more sensitive to S-TMD's destabilization. The involvement of each studied amino acid in the trimerization process can be applied in future studies of the transmembrane stabilization mechanisms in the full-length Spike protein, as well as to gain a better understanding of transmembrane helix packing in general.

The stability of each S_OPT model with amino acid substitutions inserted into the POPC bilayer with the solvent, was evaluated via molecular dynamics simulations. The original computational framework, proposed and applied by Aliper et al. [1], was used during the analysis, with key parameters including the number of the hydrophobic protein-protein and protein-lipid contacts, free volume (FV) in the lumen of the trimer, root-mean-square deviation (RMSD) of atomic positions in the α -helix backbone. The presence of an amino acid residue at the α -helix interface at a certain point in the trajectory was determined by the values of the accessible surface area (ASA). The impact of mutations on the stability of the model's structure was assessed according to two main criteria: 1) deviation from the original protein model structure during the simulation due to disruption of the α -helix interaction by lipid molecules; 2) changes in the amino acid residues' contribution to helix-helix interface formation compared to the key residues identified for the interfaces of the wild-type S_OPT model.

Examination of S_OPT trajectories revealed that residues G1219, G1223 and M1233 are crucial for the stability of the S-TMD. In contrast, mutagenesis of the residues W1212, F1220, I1227, and M1229 led to protein interface altercations without structural disruption involving lipid molecules. Models made by substituting I1216, I1225 and V1230 residues with alanine showed neither interface nor structural disruption. Overall, our results demonstrate that certain residues of the S_OPT model play a key role in the stabilization process. With this information, the packing mechanism of full-length SCoV-2 S protein can be further investigated, as well as possible ways to modulate stability of the S protein and other fusion proteins. Such insights into the structural and functional characteristics of the trimer are of great importance, as S-TMD has the potential to become a new therapeutic target [2].

The work was supported by the Russian Science Foundation (23-14-00313).

1. Aliper, E.T., Krylov, N.A., Nolde, D.E., Polyansky, A.A., Efremov, R.G. (2022) *Int. J. Mol. Sci.*, 23, 9221.
2. Westerfield, J.M., Barrera, F.N. (2020) *The Journal of Biological Chemistry*, 295(7), 1792-1814.

DEVELOPMENT OF PREDICTIVE MODELS FOR GLP-1R AGONISTS USING PATENT-DERIVED ACTIVE COMPOUNDS AND DEEP LEARNING TECHNIQUES

D.O. Shkil, K.V. Balakin

Moscow Institute of Physics and Technology, Moscow, Russia

In drug discovery, predicting target-specific biological activity presents significant challenges, particularly in high-throughput screening (HTS) programs, where only a small fraction of screened compounds exhibit activity, while the majority remain inactive. This discrepancy complicates the development of reliable predictive models, as the limited number of active compounds often leads to insufficient training data and a higher likelihood of false positives [1, 2]. One notable example of this challenge is the GLP-1R (Glucagon-Like Peptide-1 Receptor) HTS program [3].

GLP-1R is a critical therapeutic target, particularly in the treatment of type 2 diabetes and obesity [4,]. The search for potent GLP-1R agonists is therefore of high interest. However, in one of the GLP-1R HTS program, out of approximately 370,000 screened compounds, only 23 were identified as active. Subsequent analysis revealed that these active compounds were unsuitable for model development due to poor drug-likeness, as determined through visual inspection of functional groups and descriptor calculations.

Given these limitations, we hypothesized that active compounds from various patents could serve as a more representative dataset for training predictive models. We compiled a dataset of active molecules from a wide range of patents and employed deep learning techniques to develop predictive models. In this study, we developed a binary classification model for molecular activity prediction using a Graph Attention Network (GAT) architecture, implemented with PyTorch Geometric. The model processes molecular structures represented as graphs, where atoms and bonds are treated as nodes and edges, respectively. The GAT model comprises three GATConv layers with 8 attention heads each, followed by fully connected layers to produce the final classification output. The model was trained using the Adam optimizer, and class imbalance was addressed by applying a weighted binary cross-entropy loss.

Results on the test dataset: Accuracy: 0.9946, Precision: 0.4818, Recall: 1.0, F1-score: 0.6503, ROC AUC: 0.9997, PR AUC: 0.9528.

Evaluation: The model achieved high overall accuracy and ROC AUC, indicating a strong ability to distinguish between active and inactive molecules. The recall of 1.0 reflects the model's effectiveness in identifying all active molecules, making it particularly valuable in applications where false negatives are critical to avoid. However, the relatively low precision suggests a higher rate of false positives, which could be problematic in scenarios where the cost of false positives is high. Despite this, the model's high PR AUC demonstrates its robustness in handling class imbalance. Future work should focus on improving precision while maintaining high recall, potentially through further tuning of classification thresholds or model refinement.

1. Soufan, O., et al. (2015) *PLoS One*, 10(12), e0144426.
2. Butkiewicz, M., et al. (2017) *Chemical Informatics*, 3(1), 1.
3. <https://pubchem.ncbi.nlm.nih.gov/bioassay/624172>
4. Sztanek, F. et al. (2024) *Biomedicines*, ,12(6), 1320.
5. Drucker, D.J. (2024) *Diabetes Care*, dci240003

GENE CO-EXPRESSION NETWORK OF FEW SELECTED PUTATIVE BEHAVIOURAL GENES FROM SOCIAL AND NON-SOCIAL ORGANISMS USING *IN SILICO* APPROACH

A. Singh, K.V. Ramesh

Department of Biotechnology & Genetics, School of Sciences, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

Behaviour that gets influenced by the characteristics of related organisms is defined as social behaviour. Social organisms such as mammals (mice, apes, and humans) and insects (honey bees, wasps, ants, and mites) depict a higher level of social interactions. Despite differences at the ecological and evolutionary level, social animals exhibit common behavioural traits such as foraging, defensive and mite grooming. Several studies have reported that candidate genes and molecular pathways of social behaviour are identified from behaviour and physiology of organisms. Therefore, the interaction of these genes via constructing co-expression network will explore the conserved and diverse domains in behaviour associated molecular pathways of social and non-social animals. The present work focuses on gene interaction network of five important behavioural genes such as *ast* (allatostatin), *egr-1* (early growth response 1), *hr-38* (hormone receptor 38), *ace* (acetylcholinesterase) and *nrxn* (neurexin) from social and non-social organisms. For the analysis of the role of selected genes in social behaviour, KEGG pathway analysis was used to identify pathways for differentially expressed genes (DEGs). Among the several pathways, these genes interact in insulin-like growth factor signaling (insulin signaling), dopamine and ecdysteroid signaling, NPF (neuropeptide F) signaling, behavioural and neural plasticity. To begin with, the co-expression and text mining data of these genes were obtained using STRING database. In the second step, ‘.csv’ file, containing the name of interacting genes and combined score was used as an input by Cytoscape software to generate co-expression gene network. The topology of these network consists of interacting genes from insulin-like growth factor signaling, dopamine and ecdysteroid signaling, NPF (neuropeptide F) signaling, and behavioural and neural plasticity pathways. From the connections, it is evident that protein kinases such as *egr-1* form the maximum connection with other genes and therefore, protein kinases are the “hub gene”. Other functional connections in the sub-group of the network include *hr-38*, *nrxn*, *ast* and *ace*. Edges of the network suggest that behaviour-related genes are differentially expressed or physically bound in various areas of the brain and ovaries. Further, among various pathways, insulin-like growth factor signaling pathway is the central one for the regulation of sucrose responsiveness, that will further regulate the behavioural modulation in social and non-social organisms.

CASTOR OIL AND LAXATIVE ACTIVITY: VIRTUAL SCREENING OF RICINOLEIC ACID ANALOGUES

S. Gabsi¹, N. Ben Fatma^{1,2}, F. Hfaeidh¹, S. Bayoudh¹, M. El Kaibi^{1,2}

¹Pharmaceutical Sciences Department B, College of Pharmacy of Monastir, University of Monastir, Monastir, Tunisia

²Drug Development Laboratory LR12ES09, College of Pharmacy of Monastir, University of Monastir, Monastir, Tunisia

Introduction. Castor oil, derived from the seeds of *Ricinus communis* L., has been utilized for centuries for its diverse therapeutic and industrial applications [1]. The objective of this study was to identify potential candidates for future drug development.

Materials and methods. The molecules related to ricinoleic acid (RA) that we investigated were obtained from the “Related compounds” section in PubChem®, which includes all parent compounds, mixtures, salt forms, as well as similar compounds and conformers with structural similarities to the compound of interest in terms of 2D or 3D chemical structure. We employed the Python library Dockstring to conduct virtual screening on 100 randomly selected molecules. Their SMILES codes were loaded from a CSV file, along with the target file and the configuration file containing grid box coordinates. Subsequently, 3D conformations were generated, followed by molecular docking using Autodock Vina®. The resulting data was then recorded for analysis.

Results and discussion. The virtual screening of the 100 molecules related to RA showed that some molecules had better affinity score than RA. Among these molecules, certain ones exhibited more negative interaction scores, suggesting a potential to establish stronger, more stable, and more specific interactions with the target molecule. The top 10 molecules with the highest scores were selected. The difference in the affinity scores between the RA-related molecules and RA can be attributed to additional bonds formed by related molecules, such as salt bridges, Pi-Sigma type hydrophobic bonds, and the number of connections established. This may explain the improvement in affinity. Further analysis of these molecules highlighted four candidates (26, 28, 38, and 4) with favorable physicochemical, pharmacokinetic, and toxicological properties, suggesting their potential for therapeutic development.

Conclusion. Through molecular modeling and virtual screening, we identified several promising compounds with potential enhanced efficacy compared to ricinoleic acid. However, to translate these findings into clinical practice, rigorous experimental validation, including molecular dynamics studies and *in vitro* and *in vivo* experiments, is essential.

1.Dawson, W. R. (1929) *Aegyptus*, 10(1), 47-72.

CREATION OF SAR MODELS FOR PREDICTION OF T-CELL EPITOPES WITH HUMAN LEUKOCYTE ANTIGENS BASED ON PROTEIN STRUCTURAL FORMULAS

A.S. Smirnov¹, D.A. Filimonov², A.A. Lagunin^{1,2}

¹*Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia*

²*Institute of Biomedical Chemistry, Moscow, Russia*

One of the primary goals of immunoinformatics is the prediction of interactions between T-cell receptors (TCR) and their ligands, known as epitopes. An epitope for T-lymphocytes is a linear protein fragment interacted with a major histocompatibility complex (MHC) molecule. Human MHC molecules often is called human leukocyte antigen (HLA). HLA class one (HLA I) is a membrane-bound protein consisting of two chains: a highly variative in population alpha chain and constant beta-2-microglobulin. T-lymphocytes can only recognize epitopes that are bound to HLA. An epitope binds with an alpha chain. Neoepitopes and viral epitopes typically appear on HLA I. Prediction of binding epitopes with different HLA I alleles can help in vaccine and immunotherapeutic agent development against virus infections and oncology diseases. This theme has become a hot topic in bioinformatics, and many authors suggest their own decisions. The latest approaches are based on artificial neural networks with fixed-length input. The quality metrics are good only for epitopes with length at 9 residues, because datasets are biased towards nine. Current software work with protein sequences as for letter sequences for prediction or converts them into vectors of numbers without using chemical structure information.

In this work, binding epitopes with HLA proteins are analyzed using the structure–activity relationships (SAR) approach when epitopes are represented by their structural formulae instead letter sequences. The SAR models were created and validated by the modified version of Prediction of Activity Spectra for Substances (MultiPASS) software which allows using different levels of multilevel neighborhoods of atoms (MNA) descriptors to describe the peptide's structural formula. The MultiPASS uses modified naïve Bayes algorithm for revealing relationships between structural formulae and interaction with HLA alleles. We successfully used similar approach for prediction of interaction between epitopes and CDR3 regions of T-cell receptors [1]. The dataset from NetMHCpan-3.0 was used for training and validating models [2]. This dataset has provided a strict-5-fold cross-validation split. No identical 8-mer segment was shared between partition as possible [3]. The dataset contains two types of assays: binding affinity measurements and mass-spectrometry assays. The binding affinity data were transformed to binary by threshold $0.426 (1 - \log_{10}(500 \text{ nM}) / \log_{10}(50000))$. It consists of 126404 non-active and 40814 active peptides for 104 HLA I alleles. The mass-spectrometry data consists of 80220 active and 3309176 non-active peptides for 51 HLA I alleles. We also used split by assay types for training and validating. The areas under the receiver operating curve and under the precision-recall curve (AUROC and AUPRC, respectively) were used as model quality metrics.

As result, SAR models for 84 and 51 HLA I alleles were created using binding affinity data and mass-spectrometry data, respectively. The best models for binding affinity data have median AUROC 0.85 (0.85 [0.80;0.88], first and third quartile in brackets) and AUPRC 0.63 [0.48; 0.73]. The best models for mass-spectrometry data have median AUROC 0.98 (0.98 [0.94;0.99], first and third quartile in brackets) and AUPRC 0.75 [0.57; 0.81]. Although models based on binding affinity data have lower quality metrics, than analogues, the models based on mass spectrometry data works significantly better than models based on binding affinity data.

This work was supported by the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (no. 122030100170-5).

1. Smirnov, A., Rudik, A.V., Filimonov, D.A., Lagunin, A.A. (2023) *Immunology*, 169, 447.
2. Nielsen, M., Andreatta, M. (2016) *Genome Med.*, 8, 33.
3. Qu, W., You, R., Mamitsuka ,H., et al. (2023) *Bioinformatics*, 39, btad551.

QUANTITATIVE PREDICTION OF HUMAN IMMUNODEFICIENCY VIRUS DRUG RESISTANCE

E. Stolbova^{1,2}, L. Stolbov¹, D. Filimonov¹, V. Poroikov¹, O. Tarasova¹

¹*Institute of Biomedical Chemistry, Moscow, Russia*

²*National Research University Higher School of Economics, Moscow, Russia*

Human immunodeficiency virus (HIV) drug resistance remains a persistent global problem. Although highly active antiretroviral therapy (HAART) can be used to effectively treat HIV infection, some individuals may experience poor therapeutic outcomes, in particular due to HIV drug resistance. This phenomenon represents a major challenge in the management of HIV, as it can lead to increased viral load, disease progression and higher rates of HIV transmission. Data containing viral sequences from patients with insufficient treatment results and characteristics of viral drug resistance can be used to build quantitative and qualitative models for predicting drug resistance.

Numerous studies on the qualitative prediction of HIV drug resistance were performed using various machine learning methods such as Naïve Bayesian approach, random forests, decision trees, support vector machines, and neural networks [1,2]. In our previous study, we developed regression models predicting HIV drug resistance to eight protease inhibitors using random forest regression (RFR), support vector regression (SVR), and self-consistent regression (SCR) [3].

This study aims to build quantitative models of HIV drug resistance to a broader range of antiretroviral drugs, including non-nucleoside reverse transcriptase inhibitors (NNRTIs), nucleoside reverse transcriptase inhibitors (NRTIs), and integrase inhibitors (INIs). We considered drugs with sufficient observations to build robust models, thereby excluding those with insufficient data. Specifically, we built models to predict resistance to two NNRTIs – efavirenz (EFV) and nevirapine (NVP), and six NRTIs – lamivudine (3TC), abacavir (ABC), zidovudine (AZT), stavudine (D4T), didanosine (DDI), tenofovir disoproxil fumarate (TDF) and two INIs including raltegravir (RAL) and elvitegravir (EVG). We built RFR and SVR models based on genotype-phenotype data. The models predict resistance fold ratios, using binary vectors with values of 0 or 1 to indicate the presence of specific pentapeptides in each amino acid sequence as independent variables. The performance of the models was evaluated using the coefficient of determination (R^2) and root mean square error (RMSE) in five-fold cross validation.

Overall, our models are characterized by reasonable performance for predicting resistance to HIV enzymes. The mean performance metrics were $R^2 = 0.79$, RMSE = 0.448 for NNRTIs; $R^2 = 0.717$, RMSE = 0.245 for NRTIs; $R^2 = 0.716$, RMSE = 0.315 for INIs. Differences in the model performance for different drugs may be associated with the quality of the original experimental data. Despite these limitations, our approach allows for quantitative prediction of resistance levels based on structure-resistance relationships.

Therefore, the quantitative models developed in our study may be useful for analyzing HIV resistance to drugs used in HAART regimens, predicting the efficacy of the specific drugs and drug combinations for a particular individual, and optimizing treatment strategies.

This study is supported by the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) (No. 124050800018-9)

1. Steiner, M., Gibson, K., Crandall, K. (2020) *Viruses*, 12(5), 560.
2. Tarasova, O., Filimonov, D., Poroikov, V. (2017) *J. Bioinform. Comput. Biol.*, 15, 1650040.
3. Stolbova, E.A., Stolbov, L.A., Filimonov, D.A., Poroikov, V.V., Tarasova, O.A. (2024) *Viruses*, 16, 1132.

DRUG-DRUG INTERACTION MECHANISM PREDICTION USING BAYESIAN METHODS IN MACHINE LEARNING

R.R. Taktashov, A.V. Dmitriev, D.A. Filimonov

Institute of Biomedical Chemistry, Moscow, Russia

The combined use of drugs (polypharmacy), which has become considerably more prevalent in recent years, can have a versatile effect on the human body: it can cause beneficial effects (pharmacokinetic or pharmacodynamic parameter modulation to achieve the required result in the context of treating or preventing specific diseases) or lead to undesirable, possibly life-threatening Drug-Drug Interactions (DDIs). It is especially important to develop and improve computational methods, which are, in the realm of DDI research, often the only available method for researchers, to predict and/or classify potential DDIs. A clinically significant result of DDI prediction is not just the fact of DDI occurrence, but also an indication of the DDI mechanism, taking into account the severity level of individual DDIs.

We created classification models for pharmacokinetic and pharmacodynamic DDI mechanism prediction for pairs of substances with severity level indication while utilizing PASS algorithms and PoSMNA molecular descriptors. To create a training set, we used the MecDDI database, which is predicated on a hierarchical classification of DDI mechanisms. The platform is unique in that it provides detailed descriptions and graphical illustrations for each of the more than 178000 DDIs of 1922 FDA-approved drugs. Assigned severity levels of each DDI are indicated according to the ORCA DDI classification system, which is focused on clinical management. An Invariant Accuracy Prediction (IAP) criterion, which is similar to AUC (Area Under ROC Curve), was calculated in Leave-Pair-Out Cross Validation (LPO-CV) procedure to assess the predictive power of PASS algorithms. IAP values closest to 1 correspond to a higher predictive capability of the selected model.

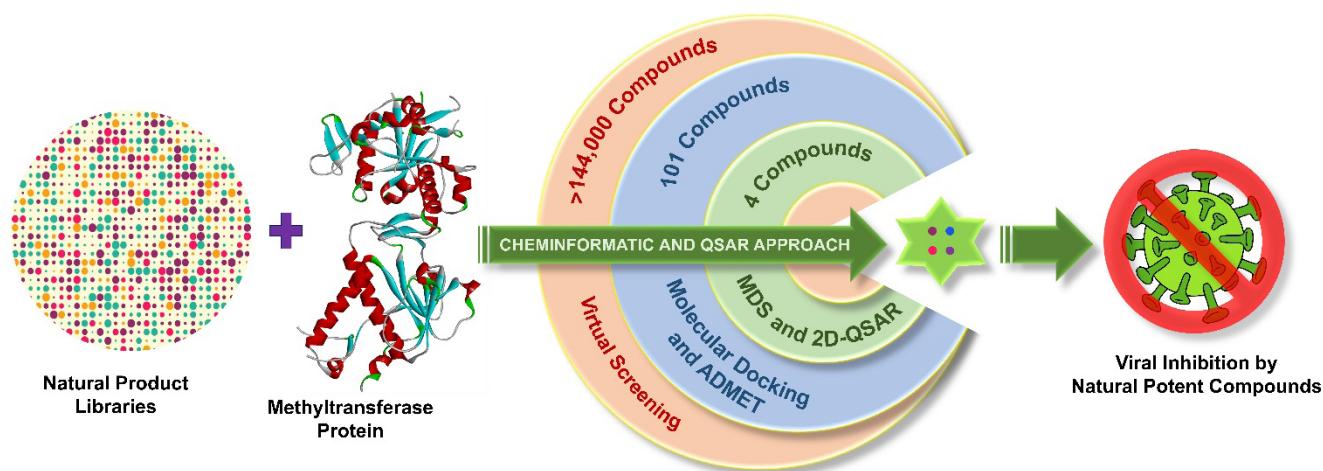
As a result of the training and validation procedures, prediction models were obtained for 96 mechanisms, classes and severity levels of MLV, with an average IAP value of 0.8 for all models. The model we developed is freely available on the Way2Drug web portal (<https://www.way2drug.com/ddi/>), where the user can indicate pairs of chemical structures of interest in MOL files, SMILES strings or generic names. Chemical structures can also be depicted by the user using the drawing tool in MarvinJS web add-on. A spectrum of mechanisms of pharmacokinetic and pharmacodynamic DDIs with severity level indication can also be obtained for substances that have not yet been synthesized, which is important in the early stages of drug development.

CHEMINFORMATICS AND QSAR-BASED IDENTIFICATION OF NATURAL BIOACTIVE COMPOUNDS AS POTENT INHIBITORS OF SARS-COV-2 N-7 METHYLTRANSFERASES

J. Thomas, J. Satija

Centre for Nanobiotechnology (CNBT), Vellore Institute of Technology, Vellore, Tamil Nadu, India

Since the emergence of the COVID-19 pandemic, to date, no specific and efficient therapeutic cure is available for the SARS-CoV-2 viral infection. Few reports have been observed with the late cardiovascular or neurodegenerative disorders that are associated with altered immunity. The nsp14 protein of SARS-CoV-2 is a key target for antiviral drug development as it controls replication of virus and evasion of host immune system. In this study, a consensus of cheminformatics approaches involving virtual screening, molecular docking, ADMET profiling, molecular dynamics simulations, free-energy landscape, MM-GBSA, DFT and 2D-QSAR analysis is employed to discover the inhibitor of nsp14 protein. Our study discovers potent novel natural compounds against the nsp14 protein that can hinder the replication of SARS-CoV-2 by inhibiting N-7 methyltransferases. Overall, this investigation reveals four novel potent natural bioactive compounds, namely ZINC2132169, ZINC8791872, ZINC8952607 and ZINC6624334 with better stability with nsp14 protein and non-toxic nature. Interestingly, all four hit compounds formed a stable complex with the binding pocket residues of nsp14 protein in a similar way as of control SAH. Further, the MSA analysis revealed that hit compounds do not affect the normal human methyltransferases i.e. the Glycine N-methyltransferase and thus are safe to use in humans. Further, 2D-QSAR studies supported the cheminformatics findings and predicted the pIC₅₀ values in the range of 5.30 - 6.71 nM which matches with the pIC₅₀ values of the established compound. This strongly suggests that the action of these hit compounds would inhibit the function of nsp14 and eventually lead to a decrease in viral load in the host. Overall, the present study has provided immense insights toward a deeper understanding of the inhibition of nsp14 protein at the atomic level and identification of the novel natural bioactive compounds. These findings need to be established further through *in vitro* and *in vivo* studies in the near future for their efficacy to fight against SARS-CoV-2 infection.



Graphical Abstract. Computational screening of natural product library against nsp14 pf SARS-CoV-2 to identify a potent bioactive compound.

1. Thomas, J., Satija, J. (2024) *ChemistrySelect*, 9, e202304932.

4'-FLUORO-5,7-DIHYDROXYFLAVONE – PIPERAZINE HYBRIDS AS VEGFR-2 INHIBITORS: DESIGN, IN-SILICO STUDY, SYNTHESIS, AND ANTICANCER ACTIVITY

K.R. Thombre, K.R. Gupta, M.J. Umekar

Department of Pharmaceutical Science, Smt. Kishoritai Bhoyar College of Pharmacy, Kamptee, Nagpur (MS), India

4',5,7-trihydroxyflavone and its analogs have distinct structural properties, they are valuable in medicinal chemistry and are frequently used to create novel bioactive compounds with a wide range of targeted pharmacological activities and minimal side effects. To get more potential candidates, we designed a novel series of 7-((substituted piperazine) alkoxy) 4'-fluoro,5,7-dihydroxyflavone derivatives and docked with Vascular endothelial growth factor receptor 2 (VEGFR-2) with PDB ids-4ASD. VEGFR-2 is an angiogenic mediator that promotes rapid development and early metastasis. Some designed derivatives were considered for synthesis based on their good affinity with relevant receptors, binding energy, toxicity profile, and Lipinski rule of five. Molecular docking validation were performed by calculating RMSD value. Synthesis of these derivatives was performed in two steps. In the first step, Fluoro apigenin and 1,4-dibromo butane were incorporated with potassium carbonate as a base and anhydrous DMF as a solvent to create 4'-fluoro-5,7-dihydroxy flavones (**S1–S10**), which were then further processed with different substituted piperazines in the second step. All the synthesized derivatives were characterized by elemental and spectral analysis, confirming the proposed structure of the derivative. The newly synthesized 4'-fluoro-5,7-dihydroxyflavone – piperazine hybrids were screened for anticancer activity. The anticancer activities of some synthesized derivatives based on their physical, spectral, and elemental analysis were performed on the MCF cell line followed by an MTT assay. The synthesized hybrids **S1, S5, S7, and S10** show good anticancer activity as compared to other derivatives.

SELECTIVITY OF LIPID BINDING TO THE VANILLOID SITE OF TRPV ION CHANNELS

I.I. Veretenenko^{1,2}, Y.A. Trofimov^{1,3}, R.G. Efremov^{1,2,4}

¹*Shemyakin–Ovchinnikov Institute of bioorganic chemistry RAS, Moscow, Russia*

²*Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Russia*

³*National Research Nuclear University MEPhI, Moscow, Russia*

⁴*National Research University “Higher School of Economics”, Moscow, Russia*

Thermo-TRPVs (Transient Receptor Potential Vanilloids) are non-selective ion channels, that are expressed, among others, in the ends of sensory neurons and function as sensors of various external stimuli, such as temperature and some chemical compounds (e.g., capsaicin - vanilloid, the active component of hot chili peppers, which acts as an agonist of TRPV1). These channels are involved in nociception, for instance, by their activation in response to high temperature. Despite several proposed mechanisms of TRPVs thermo-induced activation [1,2], the nature of TRPV thermosensitivity remains enigmatic. In the context of this problem, the “vanilloid binding pocket” (VBP) site draws attention due to the presence of lipid in the closed state of TRPV3 and its absence in the heat activated one. It was assumed that this lipid can function as a “fuse”, stabilizing the channel in the closed state at low temperature and initiating a wave of conformational rearrangements in the course of leaving VBP as the temperature rises [2]. One member of the TRPV subfamily, TRPV1, is known to contain phosphatidylinositol (PI) in its VBP [3,4], whereas the same site of TRPV3 is likely to be non-selective for lipid type [2].

The aim of this study is to identify the mechanism of lipid selectivity in VBP. For this purpose, molecular dynamics (MD) simulations were performed for TRPV1 and TRPV3 with five most common types of phospholipids inserted in VBP: PI, phosphatidylserine (PS), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG). Analysis of MD data showed that the PI selectivity of TRPV1 is determined by the structural features of the site: the characteristic “bottleneck”, which holds the lipid phosphogroup, and the site extension, where the large polar head of PI is most favorably located. Free energy calculations confirmed the preferable binding of PI. At the same time, in the VBP of TRPV3 there are two substitutions among the “bottleneck”-forming residues: Ser→Phe and Glu→Gln. This leads to the destruction of the stable “bottleneck” structure and prevents lipid positioning similar to TRPV1, which causes loss of the site selectivity.

A comprehensive understanding of the mechanism of lipid binding in the VBP site opens new possibilities for the design of pharmacological interventions, as this site also competitively binds a number of TRPVs agonists [5]. Moreover, mutations of the “bottleneck”-forming residues can affect the interaction of VBP with lipids, opening the prospect of engineering TRPVs with altered properties: in particular, channels with modified temperature activation threshold, which can be used in thermogenetics [6] – a method of activating cells (e.g., neurons) by locally increasing the temperature with laser or ultrasound radiation.

This study was supported by RSF grant no. 23-14-00313.

1. Kwon, D.H., Zhang, F., Suo, Y., et al. (2021) *Nat. Struct. Mol. Biol.*, 28 (7), 554–563.
2. Nadezhdin, K.D., Neuberger, A., Trofimov, Y.A., et al. (2021) *Nat. Struct. Mol. Biol.*, 28(7), 564–572.
3. Nadezhdin, K.D., Neuberger, A., Nikolaev, Y.A., et al. (2021) *Nat. Comm.*, 12(1), 2154.
4. Zhang, K., Julius, D., Cheng, Y. (2021) *Cell*, 184(20), 5138-5150
5. Yelshanskaya, M.V., Sobolevsky, A.I. (2022) *Front. Pharmacol.*, 13, 900623.
6. Ermakova, Y. G., Lanin, A.A., Fedotov, I.V. et al. (2017) *Nat. Comm.*, 8(1), 15362.

A NEW VERSION OF RETENTION TIME PREDICTOR PROGRAM

A.I. Voronina, A.V. Rybina, V.S. Skvortsov

Institute of Biomedical Chemistry, Moscow, Russia

This paper presents a new version of the Retention Time Predictor (RTP) program and web service for predicting the retention time of peptides on a chromatographic column in mass spectrometry experiments. The previous version of RTP implemented a calculation algorithm that is a modification of the well-known program SSRCalc version 3 [1]. The only differences were the addition of retention factor values for modified amino acid residues and the addition of an algorithm to calculate the isoelectric point (pI) value from the pIPredict program [2]. Modifications used: Tandem Mass Tag and Isobaric Tags for relative and absolute quantification tags, phosphorylation, acetylation, formylation, methylation, carbamidomethylation, oxidation and double oxidation of a number of residues. In addition to routine improvements to the user interface, the new version adds a new calculation method that still uses an additive scheme with correction based on peptide length and mass and pI value. In contrast to the previous variant, the values of the retention coefficients for all residues were recalculated and the value of the hydrophobicity index (% acetonitrile) is predicted as the target value, the same as used in the Chronologer program [3]. The RTP program and web service are freely available at <http://lpcit.ibmc.msk.ru/RTP>.

The work was performed within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (№ 122030100170-5)

1. Krokhin, O.V. (2006) *Analytical Chemistry*, 78(22), 7785-7795.
2. Skvortsov, V.S., Voronina, A.I., et al. (2021) *Biomedical Chemistry: Research and Methods*, 4(4), e00161.
3. Wilburn, D.B., Shannon, A.E., et al. (2023) *bioRxiv*, 5(30), 542978.

VIRTUAL SCREENING OF NEW POTENTIAL INSECT EPOXIDASE CYP15A1 INHIBITORS

P. Yakovets^{1,2}, Y. Faletrov^{1,2}

¹Research Institute for Physical Chemical Problems, Belarusian State University, Minsk, Belarus

²Faculty of Chemistry, Belarusian State University, Minsk, Belarus

Long-term use of synthetic or natural insect growth regulators often leads to the development of resistance to them in insects [1]. One of the approaches at the initial stages of searching for new effective pest control agents is the design of compound structures using molecular docking. CYP15a1 epoxidase is an interesting target for the development of new regulators of insect populations, and compounds containing an alkyne fragment can act as potential inhibitors of its activity [2]. In our work, 40 structures of compounds from plants, including medicinal and edible ones, containing an alkyne or alkene fragment were selected as ligands for virtual screening using Autodock Vina together with the auxiliary tool FYTdock [3], and 23 model structures of insect cytochrome CYP15a1, created using the AlphaFold 3 service were selected as the target protein.

We found that 3'-Geranylchalconaringenin (Pubchem: CID10070028) from *Humulus lupulus* binds to the structures of CYP15a1 (UniProt: A0A2J7PVT0, A0A836ELH5) from the termites *Cryptotermes secundus* and ants *Pseudoattta argentina* with localization of the phenolic fragment near the heme with binding energies (E_{bind}) of -11.4 and -8.3 kcal/mol, respectively. Xanthoangelol (Pubchem: CID643007) from *Artocarpus altilis*, *Angelica keiskei* and other plants binds to the structures of CYP15a1 (UniProt: A0A2J7PVT0, A0A836ELH5) from *Cryptotermes secundus* and *Pseudoattta argentina* with E_{bind} -10.9 and -8.4 kcal/mol, respectively, and Xanthoangelol B (Pubchem: CID10409180) from *Angelica keiskei* with the structure of CYP15a1 (UniProt: A0A2J7PVT0) from *Cryptotermes secundus* with E_{bind} -10.6 kcal/mol and the location of the phenolic fragment near the heme. Ostruthin (Pubchem: CID5281420) from *Halosciastrum melanotilingia* and other plants binds to the structure of CYP15a1 (UniProt: A0A2J7PVT0) from *Cryptotermes secundus* with E_{bind} -10.6 kcal/mol and orientation of the coumarin moiety near the heme. 9-Angeloyloxy-7-methoxy-10,11-epoxy-6,7,10,11-tetrahydro-5,6-dehydro-alpha-farnesene (Pubchem: CID129829779) from *Anisotome pilifera* binds to the structure of CYP15a1 (UniProt: A0A2J7PVT0) from *Cryptotermes secundus* with E_{bind} -8.9 kcal/mol and the orientation of the terminal alkene near the heme, N-Isobutyl-2,4,8,10,12-tetradecapentaenamide (Pubchem: CID5318518) from *Zanthoxylum piperitum* binds to this structure with E_{bind} -8.4 kcal/mol and the orientation of the double bond at position 12 near the heme. (2Z,4E)-N-(2-methylpropyl)undeca-2,4-dien-8,10-diynamide (Pubchem: CID15609885) from *Echinacea angustifolia*, *Spilanthes* and other plants binds to the structure of CYP15a1 (UniProt: A0A2J7PVT0) from *Cryptotermes secundus* with E_{bind} -8.0 kcal/mol and the orientation of the terminal alkyne near the heme, and Neopellitorine A (Pubchem: CID636555) from *Artemisia dracunculus* binds to this structure with the orientation of the triple bond at position 9 near the heme with E_{bind} -7.9 kcal/mol.

Based on *in silico* evaluation, a number of natural compounds, including those from medicinal or edible plants, were identified as potential covalent inhibitors or substrate of CYP15a1 epoxidase from the termite *Cryptotermes Secundus*, which are pests for buildings, trees, agricultural lands, and the ant *Pseudoattta argentina*. Thus, the data obtained allow us to substantiate the importance of *in vitro* experimental studies of these compounds as potential regulators of insect populations.

The research was supported by the grant of the State Scientific Research Institute (Belarus) No. 20210560.

1. Hammock, B.D., Mumby, S.M., Lee, P.W., et al. (1977) *Pesticide Biochemistry and Physiology*, 7(3), 261–272.
2. Helvig, C., Koener, J.F., Unnithan, G.C., et al. (2004) *Proceedings of the National Academy of Sciences*, 12(101), 4024–4029.
3. Faletrov, Y.V., Staravoitava, V.A., Dudko, A.R., et al. (2022) *Research Square*, DOI:10.21203/rs.3.rs-1456627/v1. (preprint).

PREDICTION OF PROTEIN SECONDARY STRUCTURES BASED ON SUBSTRUCTURAL MNA DESCRIPTORS OF MOLECULAR FRAGMENTS

O. Zakharov¹, A. Rudik², D. Filimonov², A. Lagunin^{1,2}

¹*Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia*

²*Department of Bioinformatics, Institute of Biomedical Chemistry, Moscow, Russia*

The study aims to enhance the accuracy of protein secondary structure (PSS) prediction by introducing a novel approach based on structural formulae of peptides represented as a list of Multilevel Neighbourhoods of Atoms (MNA) descriptors [1]. This method captures the molecular fragment characteristics of proteins to predict their secondary structures. Traditional methods for PSS prediction, such as the GOR method [2] and modern deep learning techniques like SSREDNs [3] and CRRNN [4], primarily rely on sequence-based data and evolutionary information. These approaches, while effective, often fail to accurately capture the complexity and variability of protein folding. The proposed method leverages MNA descriptors, which offer a nuanced representation of molecular structure without relying on specific bond configurations, thereby providing a more detailed analysis of protein secondary structures. Using a dataset of over 335,000 secondary structure annotations from 37,000 proteins from PDB records with resolutions 2 Å and less, the study developed classification structure-property relationships models utilizing MNA descriptors and a Bayesian algorithm implemented in MultiPASS software [5]. The average prediction accuracy (AUC) for eight types of secondary structures was 0.902, as determined by leave-one-out cross-validation. When tested on independent datasets (ASTRAL and CB513), the models achieved AUCs of 0.860 and 0.889, respectively. The study also introduced a web application, MNA-PSS-Pred, for the prediction of protein secondary structures for peptide and protein sequences, which is freely available at <https://www.way2drug.com/MNA2DFinder/>. This tool provides high accuracy and is accessible for broad scientific use, demonstrating its significance in advancing PSS prediction methodologies.

This study has been supported within the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021-2030) (No. 122030100170-5).

1. Fomenko, A.E., Sobolev, B.N., Filimonov, D.A., Poroikov, V.V. (2003) *Biofizika*, 48, 595-605.
2. Garnier, J., Gibrat, J. F., Robson, B. (1996) *Methods Enzymol.*, 266, 540-553.
3. Wang, Y., Mao, H., Yi, Z. (2017) *Knowledge-Based Systems*, 118, 115-123.
4. Zhang, B., Li, J., Lü, Q. (2018) *BMC Bioinformatics*, 19, 293.
5. Smirnov, A. S., Rudik, A. V., Filimonov, D. A., Lagunin, A.A. (2023) *Immunology*, 169, 447-453.

AUTHOR INDEX OF ABSTRACTS

- Addoum B. 65
Agarwal B. 66
Aggarwal N. 67
Andryukov K.V. 68
Bayoudh S. 126
Biswas A. 53
Bondhopadhyay B. 69
Buyanov I. 70
Chebotaev P.P. 72
Chiobanu N. 73
Chugunov A. 31
Daniyan M.O. 47
Das S. 74
Dauda G. 80
Egorov A.D. 51
Egorov D.M. 75
Ereshchenko A. 77
Espinoza Castañeda J.I. 90
Fedorova E.V. 78
Finkelstein A.V. 37
Fomina A.D. 79
Gerasimova E.A. 76
Ghamaryan V.S. 81
Ghosh S. 82
Golubev D. 83
Golubeva A.V. 84
Gómez-García A. 48
Guerrero-González M. 85
Hacıfazlıoğlu E. 86
Héberger K. 21
Idris M.H.M. 87
Ilicheva P. 88
İslamoğlu F. 89
Ivanov S.M. 56
Jaiswal N. 43
Kel A. 35
Khoirunnisa A. 91
Kholmanskikh D.D. 92
Khrenova M.G. 42
Kondrakhin P. 93
Koroleva N.A. 94
Kulakova A.M. 58
Kumar A. 45
Lazarev I.V. 95
Le N.Q.K. 52
Leonova M.S. 96
Lisitsa E.A. 55
Lopes J.C.D. 97, 98, 99
López-López E. 100
Luzhkov V.B. 57
Malakhov G. 101
Malykhina A. 102
Mansi Devi G.A. 103
Medina-Franco J.L. 17
Mella J. 71
Mineyeva I.V. 104, 105, 106
Muhammed T.K.S. 107
Mulashkina T.I. 108
Nath A. 41
Neuberger A. 24
Ogunyemi O.M. 109
Osolodkin D.I. 49
Ostrovskii V.A. 110
Palyulin V.A. 19
Parikesit A.A. 54
Perfilev M.A. 111
Petukhova E.A. 112
Polomoshnov N. 113
Polyakov I.V. 40
Popov P. 36
Punko A. 114
Pustolaikina I.A. 115
Rajan K. 27
Roy K. 28
Rybina A.V. 116
Sabadini G. 117
Saddala M.S. 62
Sagaidak A. 118
Sargsyan A.S. 44
Savosina P. 119
Scherbakov K. 120
Sen S. 121
Senderowitz H. 23
Shilova P.A. 123
Shkil D.O. 124
Shylau V. 122
Singh A. 125
Smirnov A.S. 127
Solovev I. 26
Stolbova E. 128
Sulimov V.B. 39
Sveshnikova A.N. 32
Taktashov R.R. 129
Tamayo J. 50
Tarasova O.A. 33
Thomas J. 130
Thombre K.R. 131
Tinkov O.V. 63
Titov Yu.P. 61
Trofimov Yu.A. 25
Vassiliev P.M. 46
Veretenenko I.I. 132
Vitvitsky V.M. 29
Voronina A.I. 133
Winkler D.A. 16
Woods R.J. 30
Yakovets P. 134
Yılmaz-Sarıaltın S. 59
Zakharov, O. 135
Zhu W. 22
Zubarev R.A. 34

