

## **Integration of Computer Bioinformatics Methods in Research (B1, B2, B4, G7.2, G7.7) for the Development of Therapeutic Agents**

Modern pharmaceutical science increasingly relies on bioinformatic and computational approaches to accelerate the discovery and optimization of novel therapeutic agents. Publications (B1, B2, B4, G7.2, G7.7) present the application of *in silico* methods in the design of galantamine derivatives for Alzheimer's disease treatment, neuropeptide ligands for Parkinson's disease, antiviral molecules against SARS-CoV-2, and bibliometric analysis of scientific activity around the nutraceutical lycopene. Emphasis is placed on molecular modeling, docking analysis, quantum-chemical calculations, and bibliometric tools as key components of modern multidisciplinary research.

The development of new bioactive molecules increasingly relies on computer-aided drug discovery (CADD), which offers an efficient and cost-effective pathway from molecular design to the selection of candidate therapeutics. Studies (B1, B2, B4, G7.2, G7.7) demonstrate the central role of *in silico* methods in predicting structural-biological interactions, pharmacokinetic properties, and multi-target activity, particularly in complex diseases such as Alzheimer's, Parkinson's, and COVID-19.

Alongside molecular modeling, bibliometric analysis serves to map scientific progress, especially in dynamic fields such as nutraceuticals. This review consolidates specific examples illustrating how the integration of theoretical and experimental approaches accelerates the discovery of novel therapeutic agents.

### **Research Objectives**

The primary goal of the reviewed publications is the development of new therapeutic agents addressing significant medical challenges. Specifically, efforts are directed toward discovering drugs with high efficacy and low toxicity, suitable for Alzheimer's and Parkinson's therapy. Additional focus is placed on identifying potential antiviral agents against COVID-19, including the repurposing of existing drugs for new therapeutic

targets. A detailed bibliometric analysis of nutraceutical products, particularly lycopene, is also conducted, examining its proven and potential benefits for human health.

### ***Scientific Contributions***

The reviewed publications significantly contribute to the field of biopharmaceutical sciences, particularly in the following areas:

- Development of novel hybrid peptide-norgalantamine compounds with demonstrated low toxicity and enhanced inhibitory activity against acetylcholinesterase (AChE), as well as antioxidant effects, potentially applicable in Alzheimer's therapy.
- Synthesis of neurotensin(8–13) analogs showing improved memory, motor function, and reduced dopaminergic neuron damage in a murine Parkinson's disease model.
- Investigation of favipiravir and deferiprone highlights the latter's potential as an effective antiviral agent against SARS-CoV-2, with higher affinity for target proteins Mpro and ACE2 and low neurotoxicity risk.
- Bibliometric analysis underscores the broad application and scientific significance of lycopene, identifying key research directions and leading institutions in the field.

### ***Methodologies***

The publications employ a comprehensive set of methodologies combining computational, experimental, and analytical approaches. Special attention is given to computational methods such as molecular docking, virtual screening, and quantum-chemical calculations (DFT). Molecular docking is used to study interactions between developed compounds and their biological targets (e.g., acetylcholinesterase, viral Mpro enzyme, ACE2 receptor). Specialized software tools (Molegro Virtual Docker, SeeSAR) enable precise assessment of binding affinity and complex stability. Quantum-chemical calculations provide key insights into molecular tautomeric forms and stability under physiological conditions. *In silico* ADME-Tox analyses are also performed to evaluate pharmacokinetic properties.

Experimental methods include chemical synthesis of novel molecules, *in vitro* cytotoxicity assays, membrane permeability studies, and metal-chelating ability

tests. *In vivo* studies on animal models assess behavioral and biochemical effects. Analytical techniques encompass statistical evaluation of experimental results, bibliometric analysis using Web of Science and VOSviewer, and specialized software (DATAN V.5, PAMPA Explorer) for data processing.

## **Conclusion**

Modern drug discovery is characterized by the growing integration of computational and bioinformatic methods in the design and preclinical evaluation of novel therapeutics. This review summarizes the contributions of *in silico* approaches in several research areas: galantamine derivatives for Alzheimer's, neuropeptide ligands for Parkinson's, comparative analysis of antiviral molecules against SARS-CoV-2, and bibliometric mapping of lycopene-related research. Key roles of molecular modeling, structural docking, quantum-chemical calculations, virtual screening, and bibliometric analyses are highlighted as driving forces in biopharmaceutical innovation. The presented data demonstrate the effectiveness of multidisciplinary integration between theoretical and experimental methodologies. Computational methods enable rapid, cost-efficient identification and optimization of candidate molecules, accelerating the transition from molecular design to biological activity.

These results underscore the importance of *in silico* tools as an integral part of modern drug and nutraceutical development.

## **Publications G7.4 and G8.3: Analysis of Codon Usage and Codon Context in Prokaryotic and Mitochondrial Genomes—Development of Specialized Bioinformatics Tools and Their Applications**

Publications G7.4 and G8.3 present a comparative analysis of approaches for studying codon usage and codon context in prokaryotic and mitochondrial genomes. The focus is on mechanisms controlling gene expression and translation at the molecular level, employing newly developed software tools and empirical/statistical methods for genomic sequence analysis.

### ***Key Objectives:***

1. **Development of Gene Triplet Analysis (GTA) software**—A specialized tool for quantitative analysis of codon frequency and context in sequenced genomes. GTA processes GenBank (gbk) files, enabling analysis of protein-coding, regulatory, and non-coding regions. The software evaluates translational dynamics in bacterial and mitochondrial genomes, assessing start/stop codon frequency and contextual dependencies.
2. **Large-scale codon usage analysis in 158 prokaryotic genomes**—Identification of codon usage patterns in highly expressed genes, ribosomal proteins, and weakly expressed genes. Comparative analysis examines variations within bacterial strains and across taxonomic groups, including missing/atypical initiator codons. Correlation analysis tests the alignment between formal bacterial taxonomy and genomic codon usage patterns.
3. **Codon pair analysis and gene expression efficiency**—Statistical evaluation of codon pair distribution (sense:sense, sense:stop) in *E. coli*, testing hypotheses on translation termination efficiency. Experimental validation uses the chloramphenicol acetyltransferase (*cat*) gene, demonstrating that 3'-terminal codon pairs significantly influence protein expression levels.

### ***Methodologies:***

- Custom bioinformatics tools and scripts (VBS, VBA, Perl) automate genomic data analysis, including local database creation (NCBI, DDBJ, EMBL, KEGG), gene classification by expression level, and codon/codon pair frequency analysis.
- Visualization tools generate interpretable graphs and color-coded tables.

### ***Contributions:***

- Public release of GTA and other specialized tools.
- Insights into evolutionary and functional regulation of gene expression via codon usage patterns.

- Experimental evidence linking codon context to translational efficiency, guiding gene engineering strategies.
- Open-access data and tools for synthetic biology and biotechnology applications.

### **Publication B.3: An Algorithm and Application for Efficient DNA Fiber Data Analysis**

This study presents a novel algorithm and Java-based software ("DNA Size Finder") for automated analysis of DNA fiber labeling images, overcoming labor-intensive manual analysis in DNA replication dynamics studies.

#### ***Features:***

- Image preprocessing, Otsu thresholding, modified skeletonization (Zhang–Suen algorithm), fiber tracing with bend-angle control, and post-processing for measurement accuracy.
- Fully automated ("factory mode") and manual correction modes, configurable user profiles, platform independence, and no reliance on commercial tools.
- Validation on hydroxyurea-treated cells shows results comparable to manual analysis (for fibers  $<10\text{ }\mu\text{m}$ ).

**GitHub:** [DNA Size Finder](#)

---

### **Publication G7.1: Deglycation Activity of *E. coli* Phosphoglucose Isomerase**

This study reveals a novel deglycation function of the glycolytic enzyme phosphoglucose isomerase (Pgi) in *E. coli*, catalyzing removal of glucose-6-phosphate (G6P)-modified Amadori products from proteins.

### ***Key Findings:***

- *In vitro* and *in vivo* evidence of Pgi's deglycation activity, with Pgi-deficient cells accumulating more Amadori products.
  - Mechanistic model linking deglycation to Pgi's isomerization activity.
  - Structural homology between Pgi and bacterial Amadoriase (FrIB), suggesting evolutionarily conserved secondary functions.
- 

### **Publication G7.3: Segregational Instability in *E. coli* of Plasmids Carrying Human Interferon Gamma Gene Variants**

This work examines how 3'-end truncations of human interferon-gamma (hIFN $\gamma$ ) affect plasmid stability in *E. coli*.

### ***Results:***

- A 9-nucleotide truncation (hIFN $\gamma\Delta$ 1) causes severe instability (plasmid loss by 50 hours).
  - Linear correlation between mRNA levels and plasmid loss rate ( $\theta$ ).
  - Mathematical modeling (Stewart-Levin equations) quantifies instability parameters ( $\Delta$ ,  $\Theta$ ,  $\theta$ ).
- 

### **Conceptual Model for Breast Cancer Diagnostics and Precision Therapy**

A proposed *in silico* knowledge discovery (KDD) framework integrates:

- Medical imaging (CAD, deep learning: CNN, VGG-19).
- Biopsy/histopathology (AI algorithms: LYNA).
- Molecular biology (NGS, biomarkers: BRCA1/2, CHEK2).
- Clinical data analysis (KNIME, Weka; machine learning: SVM, ANN, PCA).

### Contributions:

- Unified multimodal data platform for personalized therapy.
  - AI-driven decision support for risk assessment and treatment efficacy prediction.
- 

### Publications G7.1, G7.5, G7.6, G8.1, G8.2: Glycation and Related Biochemical Processes in *E. coli* and *In Vitro* Systems

These studies explore glycation mechanisms, biological effects, and protective strategies, highlighting:

- **Pgi's deglycation activity**—A newly identified function in *E. coli*.
  - **Purine compounds as glycation modulators**—Theophylline and xanthine inhibit glycation; uric acid exacerbates it.
  - **Deoxyribonuclease activity of allene acids**—Potential for DNA-targeted agent development.
  - **Pro-glycation effect of L-lysine**—Contradicts its presumed inhibitory role.
  - *E. coli* as a model for glycation-induced aging—Linked to DNA/protein damage and asymmetric division.
- 

### G7.9.1: Chemistry and Environmental Protection Educational Kit

A hands-on learning kit for Bulgarian schools, including:

- Chemical reagents and lab equipment.
- Manuals for basic and advanced experiments.
- Designed to enhance STEM education and environmental awareness.

## **Интеграция на компютърни биоинформатични методи в изследвания (B1, B2, B4, G7.2, G7.7) за разработването на терапевтични агенти**

Съвременната фармацевтична наука все по-често разчита на биоинформатични и компютърни подходи за ускорено откриване и оптимизация на нови терапевтични агенти. В публикации (B1, B2, B4, G7.2, G7.7) се представят приложението на *in silico* методи при дизайн на галантаминови производни за лечение на болестта на Алцхаймер, неuropeптидни лиганди за Паркинсонова болест, антивирусни молекули срещу SARS-CoV-2, както и при библиометричен анализ на научната активност около нутрицевтика ликопен. Подчертава се значението на молекулното моделиране, докинг анализа, квантово-химичните изчисления и библиометричните инструменти като ключови компоненти в съвременните мултидисциплинарни изследвания.

Разработването на нови биоактивни молекули все по-често се основава на компютърно подпомагано лекарствено откриване (CADD), което предлага ефективен и рентабилен път от молекулния дизайн до селекцията на кандидат-терапевтици. В изследвания (B1, B2, B4, G7.2, G7.7) се демонстрира централната роля на *in silico* методи за предсказване на структурно-биологични взаимодействия, фармакокинетични свойства и мултитаргетна активност, особено при сложни заболявания като Алцхаймер, Паркинсон и COVID-19.

Наред с молекулярното моделиране, библиометричният анализ служи за картографиране на научния напредък, особено в динамични области като нутрицевтиката. Това ревю обединява конкретни примери, илюстриращи как интеграцията между теоретични и експериментални подходи ускорява процеса на откриване на нови терапевтични агенти.

### **Цели на изследванията**

Основната цел на разглежданите публикации е разработването на нови терапевтични агенти, които да адресират значими медицински предизвикателства. Специфично, усилията са насочени към откриването на лекарства с висока

ефективност и ниска токсичност, подходящи за терапията на болестта на Алцхаймер и Паркинсон. Допълнително е акцентирано върху идентифицирането на потенциални антивирусни средства срещу COVID-19, като е разгледана и възможността за репозициониране на съществуващи лекарства за нови терапевтични цели. Паралелно с това се извършва подробен библиометричен анализ на нутрицевтични продукти с акцент върху ликопена, разглеждайки неговите доказани и потенциални ползи за човешкото здраве.

## **Научни приноси**

Разглежданите публикации значително допринасят за областта на биофармацевтичните науки, особено в следните направления: Разработени са нови хибридни пептидно-норгалантаминови съединения с доказана ниска токсичност и повишена инхибиторна активност спрямо ацетилхолинестераза (AChE), както и с демонстриран антиоксидантен ефект, потенциално приложими в терапията на болестта на Алцхаймер. Синтезирани са аналози на невротензин(8–13), които демонстрират подобрене в паметта, двигателната функция и намалено увреждане на допаминергичните неврони в миши модел на болестта на Паркинсон. Изследването на фавипиравир и деферипрон откроява потенциала на последния като ефективен антивирусен агент срещу SARS-CoV-2, с доказан по-висок афинитет към целевите протеини Mpro и ACE2 и нисък риск от невротоксичност. Библиометричният анализ подчертава широкото приложение и научната значимост на ликопена, идентифицирайки ключови изследователски направления и водещи институции в областта.

Разглежданите публикации използват комплексен набор от методологии, комбиниращи компютърни, експериментални и аналитични подходи. Особено внимание е обърнато на компютърните методи като молекулярен докинг, виртуален скрининг и квантово-химични изчисления (DFT). Молекулярният докинг е използван за детайлно изследване на взаимодействията между разработените съединения и биологичните им цели като ацетилхолинестераза, вирусния ензим Mpro и рецептора ACE2. Използвани са специализирани софтуерни инструменти като

Molegro Virtual Docker и SeeSAR, които позволяват прецизна оценка на афинитета и стабилността на формираните комплекси. Квантово-химичните изчисления предоставят ключова информация за тавтомерните форми на молекулите и тяхната стабилност при физиологични условия и различни разтворители. Също така са извършени и *In silico* анализи на ADME-Tox параметри, които са важни за оценката на фармакокинетичните характеристики на изследваните съединения.

Експерименталните методи включват химичен синтез на нови молекули, последван от задълбочени *in vitro* анализи за цитотоксичност, пропускливост през биологични мембрани и способност за хелиране на метали. Проведени са и *in vivo* проучвания върху животински модели, включително тестове за поведенчески и биохимични ефекти. Аналитичните методики, използвани в тези изследвания, включват статистически анализи за оценка на експериментални резултати, библиометричен анализ с използване на базите данни Web of Science и VOSviewer, както и специализирани софтуерни инструменти като DATAN V.5 и PAMPA Explorer за обработка на експериментални данни.

## **Заключение**

Развитието на съвременната лекарствена наука се характеризира с нарастваща интеграция на компютърни и биоинформатични методи в процесите на откриване, дизайн и предклинична оценка на нови терапевтични агенти. Настоящото ревю обобщава приноса на *in silico* подходи в няколко изследователски направления: разработването на галантаминови производни за лечение на болестта на Алцхаймер, проектирането на неuropeптидни лиганди с потенциална активност срещу Паркинсонова болест, сравнителния анализ на антивирусни молекули срещу SARS-CoV-2 и библиометричното картографиране на научната продукция, свързана с нутрицевтика ликопен. Подчертава се ключовата роля на молекулното моделиране, структурния докинг, квантово-химичните изчисления, виртуалния скрининг и библиометричните анализи като движещи сили на иновациите в биофармацевтичните изследвания. Представените данни и подходи демонстрират ефективността на мултидисциплинарната интеграция между

теоретични и експериментални методологии. Компютърните методи позволяват бързо и икономически ефективно идентифициране и оптимизация на кандидат-молекули, което ускорява прехода от молекулен дизайн към биологична активност.

Тези резултати подчертават значението на *in silico* инструментариума като неразделна част от съвременния научен арсенал в разработването на нови лекарства и нутрицевтични продукти.

**В публикации Г7.4 и Г8.3 е направен, анализ на кодоновата употреба и кодоновия контекст в прокариотни и митохондриални геноми: разработка на специализирани биоинформатични инструменти и техните приложения**

В публикации Г7.4 и Г8.3 се представя сравнителен анализ на различни подходи за изучаване на употребата на кодони и кодоновия контекст в геномите на прокариоти и митохондрии. Основният акцент е поставен върху механизмите, които контролират процесите на генната експресия и транслация на молекулярно ниво, като за целта са разработени нови софтуерни инструменти, както и приложени емпирични и статистически методи за анализ на геномни последователности. Изследванията имат няколко основни цели, като първата от тях е свързана с разработването на специализиран софтуерен инструмент, наречен Gene Triplet Analysis (GTA). Този софтуер е предназначен за количествен анализ на честотата и контекста на кодоните в секвенирани геноми. Програмата GTA е създадена да работи директно с файлове във формат GenBank (gbk), което позволява анализ не само на протеин-кодиращите области, но и на регулаторните и не кодиращи области, намиращи се непосредствено преди старт кодоните и след стоп кодоните. Софтуерът е приложим за изследване на динамиката и регулацията на транслационните процеси в бактериални и митохондриални геноми, като оценява честотата на срещане на старт и стоп кодоните, както и тяхната контекстуална зависимост.

Вторият аспект на изследванията включва мащабен анализ на кодоновото използване в 158 прокариотни генома, като се поставя акцент върху

идентифицирането на специфични модели и закономерности при употребата на кодони в различни групи гени. Изследването разглежда три основни генни подгрупи: силно експресирани гени, рибозомни протеини и слабо експресирани гени. Извършен е детайлен сравнителен анализ на вариациите в използването на кодони както между различни щамове в рамките на един бактериален вид, така и между различни бактериални видове от различни таксономични групи. Важно място в този анализ заема идентифицирането на липсващи или атипични инициаторни кодони. За по-дълбоко разбиране на връзката между бактериалната таксономия и геномните характеристики, свързани с кодоновото използване, е приложен корелационен анализ. Този анализ има за цел да провери доколко формалната бактериална таксономия съответства на Геномната теория, която се основава на моделите на кодонова употреба.

Третата основна цел на изследванията е свързана с комплексния анализ на кодоновите двойки и връзката им с ефективността на генната експресия. В този контекст, вниманието е насочено специално към генома на *Escherichia coli*. Статистически е оценено разпределението на всички възможни кодонови двойки (sense:sense и sense:stop) в бактериалния геном и е изследвано влиянието на тези двойки върху транслацията. Въз основа на тези данни е формулирана и експериментално тествана хипотезата, че определени статистически предпочитани кодонови двойки могат да оказват влияние върху ефективността на терминирането на транслацията. Като експериментален модел е използван генът за хлорамфеникол ацетилтрансфераза (cat). Експерименталните резултати показват, че промените в 3'-терминалните кодонови двойки могат значително да променят количеството на експресирания протеин, потвърждавайки биоинформатично постулираната връзка между кодоновия контекст и ефективността на транслацията.

За реализирането на тези научни цели е разработен широк набор от биоинформатични инструменти и персонализирани скриптове на езици като Visual Basic Studio (VBS), Visual Basic for Applications (VBA) и Perl. Тези софтуерни решения са използвани за автоматизиране на всички ключови етапи от геномния анализ, включително създаването и управлението на локални геномни бази данни

от публични източници като NCBI, DDBJ, EMBL и KEGG. Допълнително е реализирано автоматично класифициране на гените според нивото им на експресия, както и преброяване и статистически анализ на честотата на кодоните и кодоновите двойки в големи масиви от геномни данни. Скриптовете предоставят възможност и за идентифициране на нестандартни или липсващи кодони, което има значение за разбирането на алтернативните механизми на транслация. Разработен е и софтуер за визуализация на резултатите в лесно интерпретируеми графики и цветно кодирани таблици.

Основните приноси на представените изследвания включват разработването и публичното предоставяне на специализирани софтуерни инструменти, сред които ключово място заема програмата Gene Triplet Analysis (GTA). В допълнение, мащабният анализ на кодоновото използване е осигурил ценни данни за различията и приликите в моделите на използване на кодони сред различните бактериални групи и типове гени, което има важно значение за разбирането на еволюционните процеси и функционалната регулация на генната експресия. Експерименталното доказателство за влиянието на кодоновите двойки върху транслационната ефективност предоставя нови практически насоки за оптимизация на генно-инженерни стратегии и подобрява разбирането на фундаменталните процеси на генната експресия. Всички генерирани резултати, както и използваните софтуерни инструменти и изходни кодове, са свободно достъпни за научната общност, което позволява по-нататъшни изследвания и потенциални приложения в областта на синтетичната биология и биотехнологиите. Тези приноси значително задълбочават разбирането за сложността на генетичния код и ролята на кодоновия контекст в регулацията на генната експресия и еволюцията на геномите.

**В публикацията B.3 An algorithm and application to efficiently analyse DNA fibre data** се представя разработването на нов алгоритъм и софтуерно приложение за автоматизиран анализ на изображения от ДНК фибърно маркиране. Основната цел на изследването е да се преодолее проблемът с трудоемкия и субективен

ръчен анализ, който е характерен за тази утвърдена техника за изследване на динамиката на ДНК репликацията на единична молекулна резолюция.

Създаденото приложение, наречено **“DNA size finder”**, е разработено на Java и интегрира цялостен алгоритмичен работен поток за обработка на изображения. Включени са етапи като: предобработка на изображения, автоматична сегментация чрез Otsu праг, модифицирана скелетизация (на базата на Zhang–Suen алгоритъм), нов метод за трасиране на фибри с контрол на ъгъла на завой, постпроцесинг за оптимизиране на точността на измерванията, както и автоматичен анализ на оцветяването на фибрите.

Приложението позволява както напълно автоматичен („factory mode“), така и ръчно-коригиран режим на работа. Всяка сесия се базира на конфигурируеми потребителски профили, което осигурява възпроизводимост и адаптивност към различни експериментални условия. Софтуерът е напълно безплатен, платформено независим и не изисква използване на външни комерсиални инструменти.

Тестването на „DNA size finder“ при контролни и третирани с хидроксиурея клетки показва, че резултатите от автоматичния анализ са съпоставими с тези от ръчен анализ, особено при стандартни фибри с дължини под 10  $\mu\text{m}$ . Приносът на авторите е както в създаването на функционален инструмент за високопроизводителен анализ, така и в алгоритмичните подобрения, които разширяват възможностите на метода за оценка на репликационната динамика.

Разработеният софтуер е достъпен с отворен код на адрес:

👉 <https://github.com/Teodor03/DNA-size-finder>

**В публикацията Г.7.1.: „Deglycation activity of the *Escherichia coli* glycolytic enzyme phosphoglucose isomerase“** се разглежда нова и досега недостатъчно проучена функция на добре охарактеризирания гликолитичен ензим фосфоглюкозоизомераза (Pgi) от *Escherichia coli*. Основната цел на изследването е да се установи дали Pgi притежава способност за катализиране на дегликация – процес на премахване на неензимно присъединени захарни остатъци (Амадори продукти) от протеини, модифицирани с глюкоза-6-фосфат (G6P). Тази хипотеза е мотивирана от биохимични и структурни сходства между Pgi и вече известни бактериални амадориази.

Изследването използва комбинация от биохимични, клетъчни и структурни подходи:

- **Ин витро дегликационни тестове**, при които гликирани с G6P модели (лизозим и BSA) се третират с пречистен Pgi и се анализират с NADPH-зависима ензимна реакция;
- **Кинетичен анализ**, чрез който се определят параметри като  $K_m$ ,  $V_{max}$  и  $k_{cat}$ ;
- **Клетъчна експериментална система**, сравняваща Pgi-дефицитни и Pgi-профициентни щамове на *E. coli*, с последващо измерване на нивата на Амадори продукти (APs), крайни продукти на гликацията (AGEs) и G6P;
- **Флуориметрични и имунологични методи** (вкл. ELISA) за детекция на AGEs и CML;
- **Биоинформатичен структурен анализ**, чрез който се сравняват активните сайтове на Pgi и амадориазата FrlB с помощта на jFATCAT и структурни данни от PDB.

Принос на изследването

1. **Идентифицира се нова биохимична функция на Pgi** – ензимът проявява дегликационна активност спрямо G6P-свързани Амадори продукти, което не е описано досега.

2. **Демонстрира се физиологична значимост на тази активност**, чрез повишено натрупване на гликирани продукти в Pgi-дефицитни клетки, независимо от по-ниското съдържание на G6P.
3. **Предлага се механистичен модел**, при който Pgi катализира обратима изомеризация, водеща до освобождаване на протеина от захарната модификация.
4. **Разкрива се структурна и функционална хомология между Pgi и бактериалната амадораза FrlB**, подкрепяща възможността за еволюционно съхранена допълнителна активност на гликолитични ензими.
5. **Поставя се основа за по-широки изследвания** на аналогични механизми при други организми, включително еукариоти, предвид високата консервативност на Pgi.

**В публикацията Г.7.3. Segregational instability in E. coli of expression plasmids carrying human interferon gamma gene and its 3'-end truncated variants.**

Има за цел да изясни как структурни промени в рекомбинантния ген на човешки интерферон-гама (hIFN $\gamma$ ), по-специално системни **скъсявания на 3'-края**, влияят върху **сегрегационната стабилност** на експресионни плазмиди в *Escherichia coli*. Изследва се връзката между структурата на експресирания ген, нивото на продукция на mRNA и протеин, и степента на загуба на плазмиди по време на клетъчно делене в не-селективни условия. За целта са конструирани осем плазмиди на база pBR322, носещи пълната ORF на hIFN $\gamma$  или различни по дължина **3'-крайно скъсени** варианти (от hIFN $\gamma$  $\Delta$ 1 до hIFN $\gamma$  $\Delta$ 7). Те са експресирани в *E. coli* LE392 под контрола на силен конститутивен промотор. Използвани са следните подходи:

- **Създаване на последователни култури** чрез периодични инокулации в LB среда без селекционен натиск, поддържайки клетките в експоненциална фаза.

- **Оценка на плазмидната стабилност** чрез броене на колонии, устойчиви на ампицилин, спрямо общия брой колонии (250), след серийни разреждания.
- **Квантитативно определяне на продукцията** на рекомбинантен hIFN $\gamma$  чрез ELISA и на mRNA чрез хибридизация със специфичен радиоактивно белязан олигонуклеотид.
- **Математическо моделиране на популационната динамика** с помощта на уравненията на Стюарт и Левин, като са изчислени параметрите:  $\Delta$  – разлика в специфичния растежен темп между плазмид-позитивни и -негативни клетки,  $\Theta$  – специфична скорост на загуба на плазмид,  $\theta$  – относителна честота на загуба на плазмид (индикатор за нестабилност).

#### Приноси

- **Доказано е, че 3'-крайните скъсявания на гена за hIFN $\gamma$  силно влияят върху стабилността на експресионните плазмиди.**
  - Най-нестабилен е вариантът със скъсяване от 9 нуклеотида (hIFN $\gamma\Delta$ 1), при който плазмид-позитивната популация изчезва до 50-ия час от култивирането.
- **Открита е линейна корелация между нивата на hIFN $\gamma$  mRNA и степента на плазмидна загуба ( $\theta$ )** – по-висока транскрипционна активност води до по-висока нестабилност.
- **Липсва директна корелация между степента на скъсяване и загубата на плазмид**, което показва, че не дължината сама по себе си, а функционалният израз на гена играе ключова роля.
- **Предложен е нов количествен подход за оценка на стабилността на плазмидите**, съчетаващ експериментални данни с математическо моделиране, което може да се използва при проектиране на по-стабилни експресионни системи.

В публикацията ***“Conceptual model of integrated approach for in silico knowledge data discovery for breast cancer diagnostics and precision therapy”***, се предлага концептуален модел за **интегриран подход** към диагностика и прецизна

терапия при рак на гърдата, чрез комбиниране и анализ на данни от всички етапи на заболяването – от скрининг до лечение и проследяване. Моделът е базиран на парадигмата за **in silico откриване на знание (Knowledge Data Discovery, KDD)** и използва съвременни методи от областта на изкуствения интелект, машинното обучение и биоинформатиката. Основната идея е създаването на **интелигентна научна аналитична платформа**, която да подпомага клиничното вземане на решения и персонализираната медицина. В статията е представена **концептуална архитектура**, която интегрира:

- **Медицински изображения** – мамография, ултразвук, PET/CT, MRI, обработени с CAD (Computer-Aided Detection) и deep learning методи (напр. CNN, VGG-19).
- **Биопсия и хистопатология** – анализ на HE оцветяване, структурни и текстурни характеристики на тъканите, включително автоматизирана оценка с AI алгоритми (напр. LYNA).
- **Молекулярна биология** – експресионни профили на гени, молекулярни подтипове на рак на гърдата (Luminal A/B, HER2, Basal-like), и данни от NGS и биомаркери (напр. BRCA1/2, CHEK2).
- **Интегрирани данни от различни етапи** – включително пациентски досиета, данни от скрининг, генетика и клинични фактори.
- **Инструменти за анализ** – използвани са софтуерни платформи като **KNIME** и **Weka** за анализ на големи обеми от данни и прилагане на машинно обучение (SVM, ANN, PCA и др.).

Моделът обединява всички тези аспекти в единен **научен аналитичен работен поток (scientific workflow)**, който осигурява възпроизводимост, проследимост и автоматизация на анализа.

- Предложен е **интегриран in silico модел**, който обединява мултимодални медицински и биологични данни за рак на гърдата.
- Създаден е **компютърен сценарий** за проследяване на всички етапи от лечението, с възможност за последваща имплементация в клинична среда.

- Разработката включва **механизми за интелигентно вземане на решения**, базирани на анализ на големи масиви данни и алгоритми за машинно обучение.
- Моделът предоставя **рамка за прецизна медицина**, насочена към персонализирани терапевтични стратегии, включително оценка на риска, прогнозиране на ефективността на лечението и избор на терапия според молекулния профил.
- Статията допринася за развитието на **дигитална онкология**, като демонстрира как съвременни ИТ и AI подходи могат да се прилагат за повишаване на ефективността на диагностиката и лечението.

**В публикации G7.1, G7.5, G7.6, G8.1 и G8.2 са насочени за изследванията върху гликация и свързани биохимични процеси в *Escherichia coli* и in vitro системи**

Представените научни трудове разглеждат разнообразни аспекти на гликацията – неензимен процес на модификация на биомолекули със захари – и нейните биологични последици в прокариотни и in vitro модели. Основната изследователска линия е насочена към механизмите на възникване, биологичните ефекти и потенциалните защитни/инхибиторни стратегии спрямо гликационно увреждане, с фокус върху *Escherichia coli*, модели на стареене и структурно променени протеини. Цел на тези изследвания е да се разкрият **механизмите на гликационно увреждане и възможностите за тяхното ограничаване или неутрализиране**, както в клетъчни системи (*E. coli*), така и в биомолекулни in vitro модели. Изследванията подчертават **значението на гликацията не само като метаболитен страничен ефект, но и като активен фактор за загуба на функция, имуногенност, стареене и генотоксичност**. Работата има значение както за фундаменталната микробиология и биохимия, така и за приложни области като **биотехнологии, анти-ейджинг стратегии и фармакология на гликацията**.

Изследванията използват разнообразни експериментални подходи за анализ на гликация, дегликация и нуклеазна активност. Гликирането на протеини (лизозим,

BSA и *E. coli* протеини) се осъществява чрез инкубация с глюкозо-6-фосфат или D-рибоза, последвано от пречистване чрез гел-филтрация и ултрафилтрация. Количественото определяне на гликационни продукти включва спектрофотометрични и флуорометрични методи за измерване на Амадори продукти (APs), флуоресцентни AGEs и CML чрез ELISA. Агрегацията на протеини се оценява чрез SDS-PAGE. Дегликационната активност на фосфоглюкозоизомеразата (Pgi) се изследва с NADPH-зависим ензимен тест, използващ гликирани протеини като субстрати. Определени са и кинетични параметри ( $K_m$ ,  $V_{max}$ ,  $k_{cat}$ ), както и концентрации на вътреклетъчен G6P чрез ензимна колориметрия. Нуклеазната активност на аленови киселини се оценява чрез инкубация със суперспирализирана ДНК и електрофоретичен анализ, както и чрез тест с T4 ДНК лигаза за определяне на типа терминали след разцепване. Общите лабораторни процедури включват култивиране на *E. coli*, изолиране на хистон H1, определяне на протеинова концентрация и статистически анализ с t-тест и Mann–Whitney тест за значимост.

Основни приноси и открития от изследванията върху гликация, ензимна активност и стареене в *Escherichia coli*

Нова дегликационна функция на фосфоглюкозоизомеразата (Pgi) в *E. Coli*. Установено е, че гликолитичният ензим Pgi проявява дегликационна активност спрямо Амадори продукти (APs), модифицирани с глюкозо-6-фосфат. Ензимът катализира изомеризация, която води до възстановяване на нативния протеин. Активността е потвърдена както *in vitro*, така и *in vivo*, като Pgi-дефицитни клетки натрупват значително повече APs. Предложен е механизъм, свързващ дегликационната функция с класическата изомеразна активност на ензима.

Антигликационно действие на пуринови съединения. Шест пуринови съединения са тествани *in vitro* за ефект върху гликацията на BSA. Теофилин и ксантин показват най-силен инхибиращ ефект, докато пикочната киселина засилва гликацията. Получените резултати разкриват потенциала на определени пурины като

модулатори на гликационните процеси, което е от значение за бъдещи терапевтични разработки.

Доказателства за дезоксирибонуклеазна активност на аленови киселини. Три аленови киселини демонстрират специфична активност към ДНК, водеща до разкъсване на суперспирализирана плазмида. Активността е инхибируема от фосфатен буфер, а резултатите от лигазния анализ сочат възможно генериране на нестандартни терминали или участие на окислителен механизъм. Това подчертава потенциала им за развитие на нови ДНК-насочени агенти.

Усилващ ефект на L-лизин върху гликацията. L-лизин показва прокаталитичен ефект върху гликацията на хистон H1 и BSA, като стимулира образуването на AGEs. Ефектът е устойчив дори при съвместно третиране с антиоксиданти като ресвератрол, което поставя под въпрос употребата на L-лизин като инхибитор на гликацията.

*Escherichia coli* като модел за изследване на гликационно индуцирано стареене. Разгледан е *E. coli* като релевантен модел за проучване на молекулярни механизми на стареене, акцентирайки върху ролята на гликацията на ДНК и протеини. Наблюденията за асиметрично делене, мутагенеза, както и липсата на известни защитни ензими като фруктозамин 3-киназа, подкрепят хипотезата за AGEs като фактори, ускоряващи стареенето.

Изследванията представят иновативни приноси в областта на клетъчната биохимия и микробиологията, разкривайки нови ензимни функции, идентифицирайки биохимични регулатори и утвърждавайки *E. coli* като модел за стареене, свързано с гликацията. Тези резултати разширяват разбирането за молекулярните основи на протеиновото увреждане, клетъчната хомеостаза и потенциални терапевтични мишени.

Г7.9.1. и Г7.9.1 се отнася до създаване на набор за практическо обучение по химия и опазване на околната среда, за практическото обучение по химия и опазване на околната среда в началното и гимназиално образование на българските

училища, който да облекчи учебния процес, да подобри качеството на обучението и да събуди интереса на учениците към тази важна за техническия прогрес учебна дисциплина. Наборът включва комплект от химически реактиви, комплект от лабораторни принадлежности и ръководство за работа и е разработен във варианти, предназначени за базови експерименти, които са задължителни за всички ученици, както и за ученици с подчертан интерес към химия и опазване на околната среда