International Journal of Basic and Applied Sciences, 14 (SI-1) (2025) 423-427

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/cehxkd95
Research paper

Synthetic Biology for Drug Discovery: Unlocking New Therapeutic Pathways

Dr. L. Kaviarasan [1] *, Vibhor Mahajan [2], Sourav Rampal [3], H. D. Raghavendra Prasad [4], Dr. Naresh Kaushik [5], Dr. Snigdha Pattnaik [6], Shitij Goyal [7]

[1] *Assistant Professor, Department of Pharmacy, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India*
[2] *Center of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India.*
[3] *Chitkara Center for Research and Development, Chitkara University, Himachal Pradesh, India*
[4] *Assistant Professor, Department of Civil Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University)، Ramnagar District, Karnataka, India*
[5] *Assistant Professor, uGDX, ATLAS SkillTech University, Mumbai, India*
[6] *Professor, Department of Pharmaceutics, School of Pharmaceutical Sciences, Siksha 'O' Anusandhan (Deemed to be University)، Bhubaneswar, Odisha, India*
[7] *Quantum University Research Center, Quantum University, Roorkee, Uttarakhand, India*
**Corresponding author E-mail: kaviarasan.pharmacy@sathyabama.ac.in*

Received: May 2, 2025, Accepted: May 31, 2025, Published: October 31, 2025

Abstract

The goal of this study, which employs a Data Mining Algorithm, is to pinpoint the functional groups and their connectivity chains in phytochemicals that are used for diabetes treatment. We'll also compare these with Acarbose, a synthetic drug currently prescribed for diabetes management. Plants produce chemical substances to defend themselves against predators and parasites, and these protective compounds are known as phytochemicals. They can be quite effective as drugs. However, figuring out the similarities between different molecules poses a significant challenge in both drug discovery and molecular biology. This challenge is crucial because the biochemical properties of a molecule are closely tied to its structure. The molecular structure of organic compounds primarily consists of carbon, hydrogen, nitrogen, and oxygen atoms. Among these, carbon is particularly important as it forms the backbone of these compounds. The arrangement of carbon atoms is key to naming the compound and influences its physical and chemical properties. Together with hydrogen and oxygen, these atoms create normal functional groups that link together to form various functional group chains. Chains with the same functional groups tend to undergo similar chemical reactions. Therefore, identifying these functional group chains in phytochemicals is quite a complex task.

Keywords: *Pharmaceutical; Mechanisms; Discovery; Communicate.*

1. Introduction

Synthetic biology is also fast becoming a revolutionary domain in drug discovery, especially in the creation of new therapies to treat metabolic diseases, such as diabetes. Manipulation of biological systems at the molecular scale has a high potential in the discovery and optimization of small molecules as therapeutic agents. In this paper, the author will concentrate on how synthetic biology and AI-driven methodology can be integrated into the process of drug discovery (in particular, by taking advantage of phytochemicals and utilizing them to develop diabetes treatment). This research will involve the integration of computational techniques with biological data to gain further ground in the discovery of new therapeutic pathways that can have a tremendous influence on the management of diabetes.
Some terminologies form the core of the analysis in this paper. The functional groups are groups of atoms within a molecule that dictate chemical reactions and properties of the molecule. The Chameleon Clustering Algorithm is a computational technique that is utilized to cluster the structure of molecular structures based on their chemical properties. The AI/ML models (Artificial Intelligence/Machine Learning models) are systems based on data-driven approaches to detect patterns, which are used in this case to analyze phytochemicals and synthetic drugs. Phytochemicals are non-synthetic, naturally occurring compounds in plants that have therapeutic effects (anti-oxidant or anti-inflammatory, etc.), which we compare to synthetic drugs as possible therapeutic uses.
So far, treatments have tapped into three main pathways for pulmonary vasodilation: the nitric oxide (NO) pathway, the endothelin pathway, and the prostacyclin pathway. In the early days of understanding this disease, researchers suggested that, much like Knudsen's two-hit hypothesis in cancer, the development of pulmonary arterial hypertension (PAH) was influenced by a mix of risk factors [9]. These range from underlying health issues and genetic mutations to infections and toxic exposures, all leading to endothelial dysfunction and the onset

of PAH [1]. Since the inaugural Aspen Lung Conference focused on pulmonary circulation back in 1962, we've seen remarkable advancements in the field of pulmonary hypertension (PH) [11]. The rise of PAH, particularly due to the widespread use of appetite suppressants, sparked an urgent need for effective treatments [7]. The first treatment for PAH emerged in 1981, coinciding with the first heart-lung transplant, and in 1990, we saw the introduction of the first medical therapy that improved survival rates [3]. Over the past three decades, scientific breakthroughs have broadened the range of medications available for this condition, with several now approved by the Food and Drug Administration (FDA) [2] [5]. More recently, treatment strategies have shifted to address patients with milder forms of the disease, utilizing combinations of multiple agents [4].

2. Need for The Study

These include drugs like primaquine, mefloquine, halofantrine, and lumefantrine. However, reports of resistance to these newer options have also surfaced in various regions, prompting the urgent need to develop new, fast-acting medications that differ in both their chemical structure and how they work [15] [17].

In addition to this older development in synthetic biology, there are also papers on the use of synthetic biology tools in drug discovery involving bacteria. These studies are concerned with the fact that synthetic biology can be used to produce bioactive compounds and optimally metabolic pathways. Compared to our methodology, which involves the use of AI-based clustering algorithms to analyze functional groups of phytochemicals, this study highlights the importance of engineered microbial systems to produce therapeutic compounds in large quantities as an alternative approach to drug development.

Advantages

- First off, small-molecule drugs that are closely based on compounds already found in healthy humans tend to be better tolerated and safer than those developed from more common library screenings, including virtual ones and optimization methods.
- Next, our Chemilogics TM discovery engine is rooted in human clinical data, focusing on enriching only those molecules that are likely to promote health. This gives us confidence that we're targeting the most relevant and effective drug pathways [12].
- Lastly, our engine creates models that compare genetic and clinical data from healthy individuals with those suffering from diseases like inflammatory bowel disease, rheumatoid arthritis, cancer, metabolic disorders, and neurological issues. This ensures that our drug candidates are not only relevant but also applicable to real clinical situations. We see this as a winning formula for drug discovery – finding the right molecule for the right target to truly help people facing specific health challenges.

3. Materials and Methods

Synthetic biology is an exciting and diverse field that merges engineering with biology. Its main goal is to create and build new biological devices, parts, or systems that don't exist in nature, or to rework existing natural systems for practical uses [6].

Drug discovery is often a challenging endeavor, and it can easily go off track if we choose the wrong targets, develop drug candidates that aren't safe enough, or select inappropriate clinical indications for those candidates. By delving into biological pathways and genetic data, we can tackle these challenges head-on. At Empress, we've developed a unique approach to harness human clinical genetic data through AI and causal machine learning (ML). This allows us to identify safe and effective small molecule modulators that target disease-relevant human pathways, which can be quickly refined into drug candidates. We focus on clusters of bacterial genes that generate small molecules capable of positively influencing human health. Our process begins with synthetic biology to identify and characterize these small molecules, followed by chemical synthesis and optimization to create viable drug candidates. We believe that these will turn into safe and effective medicines, specifically targeting the most relevant pathways and clinical indications.

To explain the methodology, we give a step-by-step description of the application of the Chameleon Clustering Algorithm and AI/ML models. To begin with, molecular data of phytochemicals and Acarbose were pre-treated to obtain major features, namely, the functional groups within a given compound. Then, the Chameleon Clustering Algorithm was used to cluster similar functional groups. The initial step in the clustering process is to find and construct the initial clusters, during which the algorithm recognizes and clusters functional groups based on the atomic connectivity. These clusters are further refined in later stages by having them merged down based on the strength of their interconnections. The Euclidean distance was used to compute bond lengths of the functional groups that were connected within the molecular structure. Lastly, it was also determined that the phytochemicals resemble acarbose by comparing the sequences of functional groups connectivity and bond lengths, quantifying the level of similarity by the rule-based neural network.

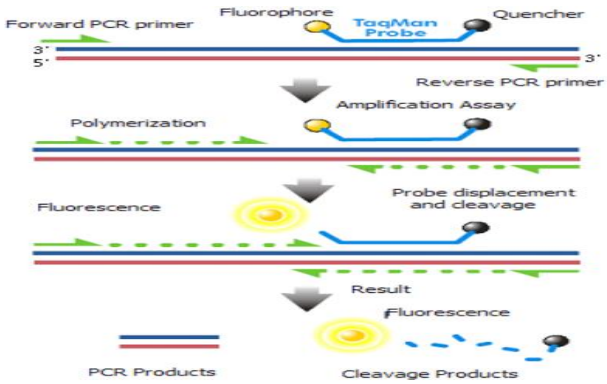

Fig. 1: TaqMan Probe Mechanism.

Figure 1 depicts the Taqman Probe Mechanism that is applied in the determination of specific DNA sequences in phytochemicals. The probe mechanism is made up of a fluorophore and a quencher, which collaborate to trace the amplification of target DNA sequences. The method is necessary in determining the functional groups of the phytochemicals by monitoring DNA-level interactions. The figure has been created using experimental results of the capability of the probe to provide accurate binding and recognition of precise molecular targets, which justify our study of phytochemical structures.

We're not exactly charting new territory in human biology, but we are stumbling upon some fresh discoveries in microbial synthetic pathways. What's exciting is how our AI and machine learning models guide us to human targets and pathways, helping us make surprising connections to diseases and clinical indications that we might have overlooked before. There are those "aha" moments when we uncover targets and pathways that we thought we understood, only to find they're doing unexpected and significant things related to specific diseases. It's thrilling to think about how we can translate these findings into clinical applications [13]. A great example of this is when we identified a therapeutic mechanism that requires engaging two targets with overlapping, compensatory biological effects. While both industry and academia have previously recognized the importance of each target on its own, translating that knowledge into real clinical success has been a challenge. By following the initial genetic clues from the coevolution of microbial chemists and human biological pathways, we've come up with a unique therapeutic hypothesis. Current AI and ML techniques are fantastic at spotting patterns in data and making connections and predictions that might otherwise slip through the cracks. However, these methods do have their limitations, particularly when it comes to the quality and quantity of the data they rely on. In drug discovery, having more and better-quality data—especially human clinical data—can make a big difference. Our AI and ML models improve even further as we integrate various data types. This includes multitopic profiling of samples from humans as hosts, combined with samples from our resident microbes, and layering in clinical features from an ever-growing number of individuals, which reflects a broader diversity of biology [8].

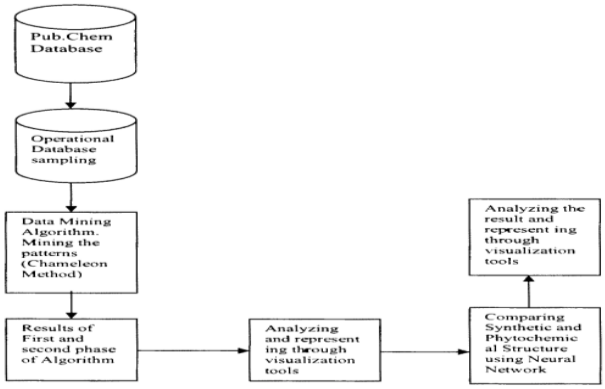

Fig. 2: Research Methodology.

We take our ethical responsibility seriously when it comes to discovering safe and effective new medicines for those battling diseases, and we see AI as a valuable tool in this mission. In the realm of drug discovery, especially when it involves patient samples, safeguarding privacy is crucial. Thankfully, there are plenty of ways to anonymize data, which enables companies to link genomic and sample analysis with clinical tests and health outcomes. We ensure that the human clinical data we use in our AI and machine learning models complies with the Health Insurance Portability and Accountability Act, along with other relevant regulations. Additionally, we rigorously assess the safety and effectiveness of our drug candidates, adhering strictly to FDA guidelines and other regulatory standards [14].

4. Results and Discussion

This merging process, guided by a dynamic model, helps uncover natural and cohesive clusters. Initially, the Chameleon algorithm creates some starting clusters and leverages the connectivity properties of each atom with its related atoms to build functional groups. Then, it merges these clusters based on the connections between the functional groups, ultimately forming final clusters that showcase the connectivity chains of functional groups in the synthetic drug Acarbose, which is used to manage diabetes [10]. In the second phase of the Chameleon Clustering Algorithm, the first-level clusters are merged to create even more clusters based on the interconnections of the functional groups, allowing the second level of Chameleon to identify those linked groups of atoms.

Fig. 3: Serotype Distribution.

Here's the text we're looking at: Acarbose, a synthetic drug, features various functional group connectivity chains. Each of these chains is treated as a rule, with potential input and output values illustrated in a truth table. The Chameleon Algorithm and Rule-Based Neural Network were crafted in 'C language and have been successfully applied to various data related to molecular structures.

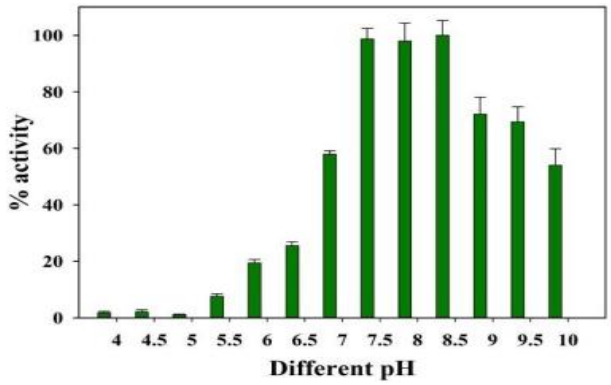

Fig. 4: Effect of pH.

The simulator system of the Neural Network looks at how different phytochemical-linked groups of atoms stack up against the functional group chains of the synthetic drug Acarbose. The results from the neural network will show how similar phytochemicals are to synthetic drugs. At the third level, we calculate the length of each cluster or functional group chain by measuring the bond lengths between the connected functional groups.

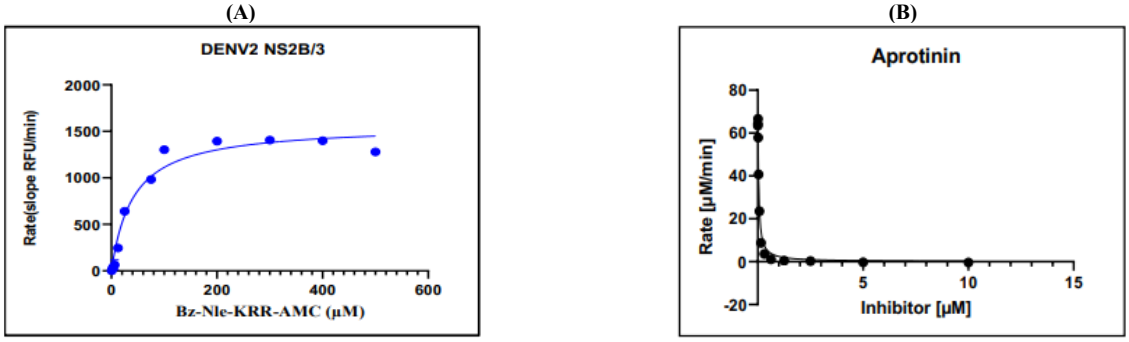

Fig. 5: Morrison Plot.

The Morrison Plot presented in Figure 5 is applied to study the bond lengths between functional groups of the overall molecular structures of the compounds. The spatial distance between atoms in phytochemicals and those in Acarbose can be compared with the help of this plot. The plot data were obtained based on the calculated bond distances, whereby the Euclidean distance techniques were used on the molecular coordinates. The figure shows that the changes in the bond length among various compounds can be associated with the therapeutic potential of the compounds, thereby proving our hypothesis on the correlation of functional groups in drug discovery.

The space between the nuclei of atoms that are bonded together is referred to as the bond distance or bond length. These lengths can change based on a variety of factors. Because bond lengths tend to be stable, the bond energies for similar types of bonds also remain consistent. In this stage, we take the X and Y coordinates of the atoms in the molecular structure and calculate the distances between them using the Euclidean distance method.

Although the application of AI/ML models and the Chameleon Algorithm presents useful information on the functional group interconnectivity of phytochemicals, these methods have intrinsic weaknesses. The quality and completeness of the data used to train the AI models are one of the greatest limitations. Noisy or incomplete information may result in inaccurate predictions and compromise the accuracy of the clustering results. Also, the Chameleon Algorithm, although effective in clustering functional groups, might not capture all the strong interactions that occur in a real biological system, which might restrict its usefulness in drug discovery in the real world. Also, a good comparison between microbial synthetic pathways and mammalian systems is difficult because microbial models are not necessarily able to mimic the complexity of mammalian biology. The variability in metabolic processes, cellular contexts, and control systems may influence the transfer of findings between microbial and mammalian systems, and this gap needs to be filled by more research and development of these models.

Besides diabetes, our strategy can be applicable to other metabolic diseases like obesity, hypertension, and hyperlipidemia, where artificial biology and AI-based drug discovery may be central to discovering new treatment regimens. As an example, the synthetic gene networks can be designed to control insulin secretion in type 2 diabetes patients, which provides a promising future for gene therapy. Besides, synthetic gene networks, when integrated in biosensing applications, may be the way to provide real-time biomarker tracking related to disease progression, a more personalized way of treatment. Synthetic biology in stem cell engineering. Synthetic biology would also be useful in designing gene circuits that regulate stem cell differentiation to create regenerative therapy for tissue repair and replacement. Such applications illustrate how synthetic biology can transform treatment approaches in various therapeutic fields to promote the science of precision medicine.

5. Conclusion

The future of translational synthetic biology really depends on how well we can connect smaller functional circuits to create larger networks that behave in predictable ways. In a previous article, we discussed four research initiatives focused on enhancing and speeding up the overall design process, making it easier to integrate biological circuitry. But beyond just improving the design cycle, applied synthetic biology could really benefit from revisiting the original inspiration behind biocomputing. Imagine being able to program higher-level decision-making into synthetic networks; this would lead to more resilient and dynamic organisms capable of multitasking. Plus, since adaptive and predictive behaviors are naturally found in all living organisms, including microbes, synthetic learning networks made from genetic and biological components could give engineered organisms a boost in automation for biosensing and similar applications. Right

now, most synthetic biology work is done in microbes, but many of the biggest challenges, especially those related to human health, are tied to mammalian systems. So, putting more effort into advancing mammalian synthetic biology is essential for developing next-generation therapies, like engineering synthetic gene networks for stem cell generation and differentiation. By tackling these challenges, our limitations won't be about the technical aspects of building or the strength of synthetic gene networks, but rather the creativity of researchers and the range of societal issues that synthetic biology can help solve.

References

[1] Usman, M. (2024). The impact of synthetic biology on genetic research and biotechnology. *Frontiers in Biotechnology and Genetics, 1*(2), 107–125.

[2] Fryszkowska, A., & Devine, P. N. (2020). Biocatalysis in drug discovery and development. *Current Opinion in Chemical Biology, 55*, 151–160. https://doi.org/10.1016/j.cbpa.2020.01.012.

[3] Fairfax, J., & Sørensen, A. (2024). Integrating telemedicine and pharmacists in chronic gastrointestinal diseases: A critical role during the COVID-19 pandemic. *Global Journal of Medical Terminology Research and Informatics, 1*(1), 23–29.

[4] Rooms, L. D., Duke, P. W., Stach, J. E. M., & Race, P. R. (2020). Unlocking the therapeutic potential of antimicrobial natural products with synthetic biology. *International Biopharmaceutical Industry*.

[5] Weth, F. R., Hoggarth, G. B., Weth, A. F., Paterson, E., White, M. P., Tan, S. T., ... & Gray, C. (2024). Unlocking hidden potential: advancements, approaches, and obstacles in repurposing drugs for cancer therapy. *British journal of cancer*, *130*(5), 703-715. https://doi.org/10.1038/s41416-023-02502-9.

[6] Cubillos-Ruiz, A., Guo, T., Sokolovska, A., Miller, P. F., Collins, J. J., Lu, T. K., & Lora, J. M. (2021). Engineering living therapeutics with synthetic biology. *Nature Reviews Drug Discovery, 20*(12), 941–960. https://doi.org/10.1038/s41573-021-00285-3.

[7] Dewangan, H., &Dewangan, T. (2024). Sophisticated design and integrative modeling of sustainable environmental practices in contemporary pharmacy and pharmaceutical industries. *Natural and Engineering Sciences, 9*(2), 395–406. https://doi.org/10.28978/nesciences.1575486.

[8] Pujol, A., Mosca, R., Farrés, J., & Aloy, P. (2010). Unveiling the role of network and systems biology in drug discovery. *Trends in Pharmacological Sciences, 31*(3), 115–123. https://doi.org/10.1016/j.tips.2009.11.006.

[9] Kala, T., &Jayabal, R. (2019). User perception about the library collections and services in Cherraan's College of Pharmacy, Coimbatore, Tamil Nadu. *Indian Journal of Information Sources and Services, 9*(1), 101–105. https://doi.org/10.51983/ijiss.2019.9.1.585.

[10] Ezeako, E. C., Solomon, A. Y., &Itam, Y. B. (2025). Prospects of synthetic biology in revolutionizing microbial synthesis and drug discovery. *Life Research, 8*(1), 6. https://doi.org/10.53388/LR20250006.

[11] Wang, X., Zhou, N., & Wang, B. (2023). Bacterial synthetic biology: Tools for novel drug discovery. *Expert Opinion on Drug Discovery, 18*(10), 1087–1097. https://doi.org/10.1080/17460441.2023.2239704.

[12] Prabhakar, C. P., & Rahman, F. (2025). Robotics and intelligent systems for autonomous industrial operations: Architectures, algorithms, and real-world applications. National Journal of Electrical Electronics and Automation Technologies, 1(3), 44–52.

[13] Rudevdagva, R., & Kingdone, G. C. (2025). Lip-reading-guided speech enhancement via self-aligning cross-attention networks. National Journal of Speech and Audio Processing, 1(4), 17–25.

[14] Abbas, M., & Gravino, D. (2025). Physical layer key generation using channel reciprocity for secure wireless body area networks. National Journal of RF Circuits and Wireless Systems, 3(1), 24–33.

[15] Müller, M. A., Schmidt, J. C., & Fischer, C. M. (2025). Sustainable VLSI design: Green electronics for energy-conscious systems. Journal of Integrated VLSI, Embedded and Computing Technologies, 2(2), 44–51.

[16] Varshavardhini, S., & Rajesh, A. (2023). An efficient feature subset selection with fuzzy wavelet neural network for data mining in big data environment. *Journal of Internet Services and Information Security, 13*(2), 233–248. https://doi.org/10.58346/JISIS.2023.I2.015.

[17] Prajapati, R. N., Bhushan, B., Singh, K., Chopra, H., Kumar, S., Agrawal, M., Pathak, D., Chanchal, D. K., & Laxmikant. (2024). Recent advances in pharmaceutical design: Unleashing the potential of novel therapeutics. *Current Pharmaceutical Biotechnology, 25*(16), 2060–2077. https://doi.org/10.2174/0113892010275850240102105033.

[18] Trosset, J.-Y., & Carbonell, P. (2015). Synthetic biology for pharmaceutical drug discovery. *Drug Design, Development and Therapy*, *9*, 6285–6302. https://doi.org/10.2147/DDDT.S58049.

[19] Weth, F. R., Hoggarth, G. B., Weth, A. F., Paterson, E., White, M. P. J., Tan, S. T., Peng, L., & Gray, C. (2024). Unlocking hidden potential: Advancements, approaches, and obstacles in repurposing drugs for cancer therapy. *British Journal of Cancer, 130*(5), 703–715. https://doi.org/10.1038/s41416-023-02502-9.

[20] Rahim, R. (2025). Mathematical model-based optimization of thermal performance in heat exchangers using PDE-constrained methods. Journal of Applied Mathematical Models in Engineering, 1(1), 17–25.

[21] Ezeako, E. C., Solomon, A. Y., Itam, Y. B., EZIKE, T., OGBONNA, C., AMUZIE, N., ... & OZOUGWU, V. (2025). Prospects of synthetic biology in revolutionizing microbial synthesis and drug discovery. *Life Res*, *8*(1), 6. https://doi.org/10.53388/LR20250006.

[22] Khan, M., & Shah, F. (2025). Evidence Based Quality Improvement for Most Businesses Sustainability. National Journal of Quality, Innovation, and Business Excellence, 2(1), 11-22.

[23] MAKLASSA, D., & NURBAYA, S. (2025). The Effect of Knowledge Management, Competence, And Work Culture Through Achievement and Digitalized Technology on The Progress of Government Organizations. Quality-Access to Success, 26(206). https://doi.org/10.47750/QAS/26.206.18.