chemistry proceedings

MDPI

Proceeding Paper

Random Forest Model Resolves the Challenges against Multi-Drug Resistant Bacteria by AgNPs †

Sona Ayadi Hassan * and Parinaz Ghadam

Department of Biotechnology, Faculty of Biological Sciences, Alzahra University, Tehran, Iran; email1@email.com

* Correspondence: s.ayadihassan@alzahra.ac.ir

† Presented at the 27th International Electronic Conference on Synthetic Organic Chemistry (ECSOC-27), 15–30 November 2023; Available online: https://ecsoc-27.sciforum.net/.

Abstract: Among the hopeless tries for finding new solutions against Multi-Drug Resistant (MDR) bacteria, Nanoparticles (NPs) especially silver nanoparticles (AgNPs) have been among the most outstanding antimicrobial and anti-MDRs because of their magnificent physical and chemical properties. Due to high costs and numerous tries and errors, there should be some strategies for predicting the anti-MDR activities of NPs. In the present study, a Machine Learning (ML) based model; Random Forest (RF) was applied to predict the anti-MDR activities of AgNPs. Once, the literature was provided, the desired information regarding the physical and chemical information besides the taxonomical information of the MDR bacteria was retrieved. Then, the preprocessing strategies were applied. Subsequently, the model was predicted with a high accuracy (R^2 = 0.73). The analysis of significant attributes revealed that Dose, DLS_size, MDR bacteria species are the most significant factors in the anti-MDR activities of silver nanoparticles. The findings proved this tool can help scientists to have reasonable assumptions toward anti-MDR activities of AgNPs before any experiments, cutting the high costs of numerous experiments.

Keywords: machine learning; Anti-MDR; AgNPs; Random Forest

1. Introduction

Antibiotics' overconsumption causes prolonged hospitalization and extensive death cases. Practices for prohibition of antibiotic resistance in microbes consist lowering antibiotics consumption, optimized drug release, modification in antibiotic targets, introduction of antibiotics for destroying and modifying the microbial enzymes etc. Unfortunately, antimicrobial resistance has been occurred against some antibiotics, which are widely administered against some fetal and pathogenic microorganisms [1]. It is so disappointing that no effective kinds of antibiotics have been produced newly to overcome this problem. On the other hand, the design and production of a new type of antibiotics is expensive and time-consuming including consumption of new compounds and numerous clinical trials [2]. The reoccurrence of resistance against newly developed antibiotics leads this process to the development of substitutes to overcome these challenges [3]. NPs are one of the most frequently produced nanomaterials, which has extended antimicrobial properties due to their special physical and chemical properties (p-chem). Since they present remarkable antimicrobial properties [4], they have been suggested as promising substitutes for antibiotics [5]. Based on the literature, metal-based NPs are vigorous agents in the attempt for the control and elimination of the microorganisms specially drug-resistant strains [6,7]. For example, a lot of investigations have reported the antimicrobial activities of AgNPs against a spectrum of multi-resistant bacteria; including MDR *Pseudomonas aeruginosa*, Methicillin-resistant *Staphylococcus aureus* (MRSA), *Klebsiella aerogenes* and etc. [8]. It is supposed that antibacterial properties of NPs are determined by three factors: (1)

Citation: Hassan, S.A.; Ghadam, P. Random Forest Model Resolves the Challenges against Multi-Drug Resistant Bacteria by AgNPs. *Chem. Proc.* **2023**, *14*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s): Name

Published: 15 November 2023

Physical and chemical characterization of a NP, (2) The taxonomical features of a bacteria and (3) Experimental conditions [9]. Since not all NPs have the capacity for being applied in the combat for eradication of antibiotic resistant bacteria, thus the discovery of such an ideal antibiotic in this case requires hard work and too much time. It is needed to have some predictions about the desired NPs with suitable antibiotic properties for the control and eradication of MDR bacteria. There has been computational methods to help the researchers to predict the physical and chemical properties of a NP before its production [10].

Machine learning (ML) answers the question of how to design and construct computers that are taught automatically by experiences and practices. The benefits of ML over practical procedures are being economic and independency of these methods to physical and chemical reactions. Something, which is so vital for medical trials. Up to now, ML has been employed to predict successful delivery of anticancer drugs [11] or direct delivery of anticancer NPs into cancer cell lines [12]. In another study, antioxidant properties of nanomaterials were predicted by ML [13]. Saadat et al. [14] suggested a ML tool for the prediction of antibacterial effects of biogenic silver NPs against gram negative and gram positive bacteria. But none of these investigations have focused on the prediction of antimicrobial effects of NPs against MDR bacteria. The bases of these predictions tools are p-chem properties, experimental conditions and microbes and cell lines as inputs. All in vitro information were accumulated from research articles and were collected in a comprehensive database. This research screens silver NPs, predicts their antibacterial effects for the control and eradication of MDR bacteria to save time and decreasing the expenses by reducing the frequency of clinical trials. Suggestion of such a ML based tool can control and inhibit the growth and biofilm formation of MDR bacteria that are of a great importance for public health.

2. Material and Methods

The path followed to implement the model is presented in Figure 1. In particular, researches regarding the anti-MDR effects of AgNPs were accumulated and data reterival was carried out in respect of physical and chemical identifications of NPs, exposure conditions and toxonomical features of the bacteria. All the original data sets have been assessed as complete. There was then the preprocessing of information, including standardisation, transversion of categorical features to numerical ones and data categorization.

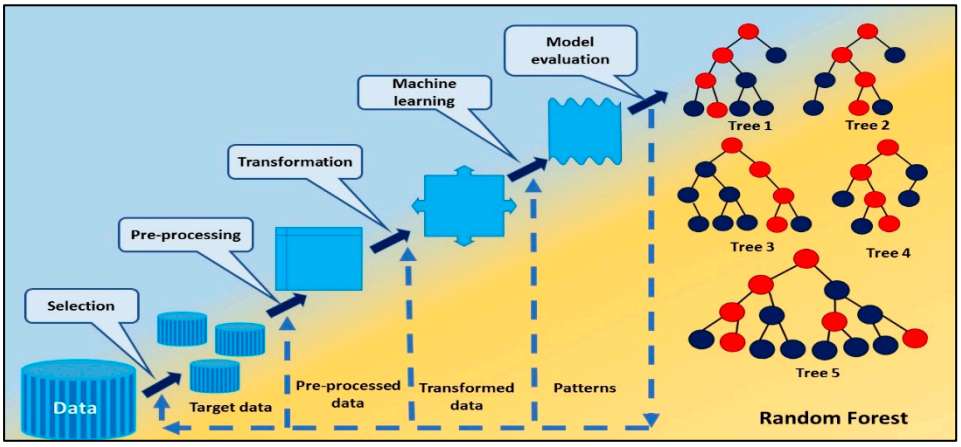

Figure 1. Random Forest model development prcess for the prediction of anti-MDR activities of AgNPs.

We set up several regression models and verified their accuracy so we could find a model with good predictability. In order to identify the attributes that have a major impact on predictions of results, an analysis of their relevance was carried out.

2.1. Data Selection

A literature search for studies looking at the effects of AgNPs on eradication or inhibition of MDR was carried out in 10.23.2022. Between 2012 and 2022, the articles collected included various keywords, such as "bactericidal", "antibacterial, "microbicide" and "antimicrobials". There were approximately 1000 articles found. As previously mentioned, the reason for the selection AgNPs is that it is the most studied case in this regard. The studies have included chemical, physical and green synthesis of AgNPs. English language, original studies with a view to antibacterial properties of NPs for only MDR bacteria published over the last 10 years and in vitro investigations are discharge factor. The discharge factor apply to case reports, reviews, studies showing binary results and studies which demonstrate their findings by means of illustrations. It appears that, in total, around 517 documents are relevant to this examination, but it is only the information concerning 70 articles which has been obtained because all of them do not meet the desired factor. In vitro studies were used solely because they replaced, reduced and enhanced animal tests with no consideration of the impact of pharmacokinetics in order to perform unexplanatory testing [15].

2.2. Retrieving Data

2.2.1. The Retrieval of Input Data

Each research was evaluated with a concentration on (i) the chemial composition of NPs (AgNPs); (ii) the physical and chemical characterization parameters and (iii) the study design experimental parameters. For having accurate predictions on the antibacterial activities of AgNPs, these variables have been obtained as an input parameters [9].

2.2.2. The Retrieval of Output Data

Several assays and methods have been reported in studies for the evaluation of antimicrobial efficacy. Some outcomes based on antibacterial measurements were filed, such as optical density (OD), colony-forming unit (CFU), minimum bactericidal concentration (MBC), minimum inhibitory concentration (MIC) and zone of inhibition (ZOI). In the output values, there were various metrics and expressions which highlighted the need for a standardised method to report scientific data.

2.3. Data Transformation

Missing data has been reported in some of the numerical attributes from the original Raw Dataset I. We have created the final dataset, which is known as "Dataset II" following selection of outcomes. There were only a handful of missing values from the inputs in our final data. We filled out the missing values by an average of the data provided in one column, because regression models do not work well with null data [16]. Categorical attributes must be transformed into numerical attributes in regression models. Multiple conversion methods are used to make the new columns appear in the main dataset; for each specified column, we created dummy variables which have been converted into a single data set. For the purpose of indicating whether or not original attributes are present, a value 0 or 1 was used [17].

For the purpose of ensuring model accuracy, numeric data should be normalized. Z Score, minmax scaler, standard scaler and median absolute strategy can be used for this process. The minmax scaler strategy has been applied in this study. This estimator scales and translates each individual attribute so that it falls within a certain range on the training set, for example, between 0 and 1 [18]. A training set has already been given to the supervised computer algorithm that can be used to approximate the output of an unidentified target function. The data was randomly divided into two series, one seri to train the model (training serie) with 70% data, and the other serie (30%) to evaluate the model efficiency (test serie).

2.4. Model Development, Evaluation and Attribute Importance Analysis

Regression techniques build models that can calculate new numerical values out of input variables. Regression modeling involves associating a function that maps inputs to continuous outputs [9]. The applied ML algorithm functions from the chemical and physical characterization of NPs and experimental situation to the ability to inhibit MDR-bacteria and helps predict the anti-MDR ability of NPs. In this research, Random Forest (RF) was applied to predict the anti-MDR activities of AgNPs. The model was developed in Python 3.8, Scikit-learn version 1.3.0. It consists of a set of decision trees and integrates interactions and feature selection naturally into the learning process. It is nonparametric, explainable, efficient, and provides high prediction accuracy for a wide range of data [19]. The model was assessed using mean absolute error (MAE), root mean square error (MSE), root mean square error (RMSE), and coefficient of determination or R squared (R^2) [20]. The significance of attributes is a monitored event where the most important characteristics are differentiated and ranked according to their importance for predicting outcomes.

3. Results

3.1. Data Transformation

In the dataset I, rows and 18 columns were derived from 70 studies which tested the antiNMPDR properties of 850 AgNPs. Data sets II were created by excluding attributes with over 20% missing values. The final input data consisted of duration (h), exposure dose (μg/mL), Core_size (nm), Zeta_potential (mV) and DLS_size (nm) and reported in numeric values. Variables with categorical values included bacteria taxonomical information, shape and coating (Figure 2). Changing the coating column to a binary format: Coated and Uncoated type was necessary in order to prevent models being overfitted. A total of 260 rows and 11 columns are contained in the final dataset. Zeta_potential, DLS_size, Dose and Core_Size have the highest percentage of missing values in this new dataset.

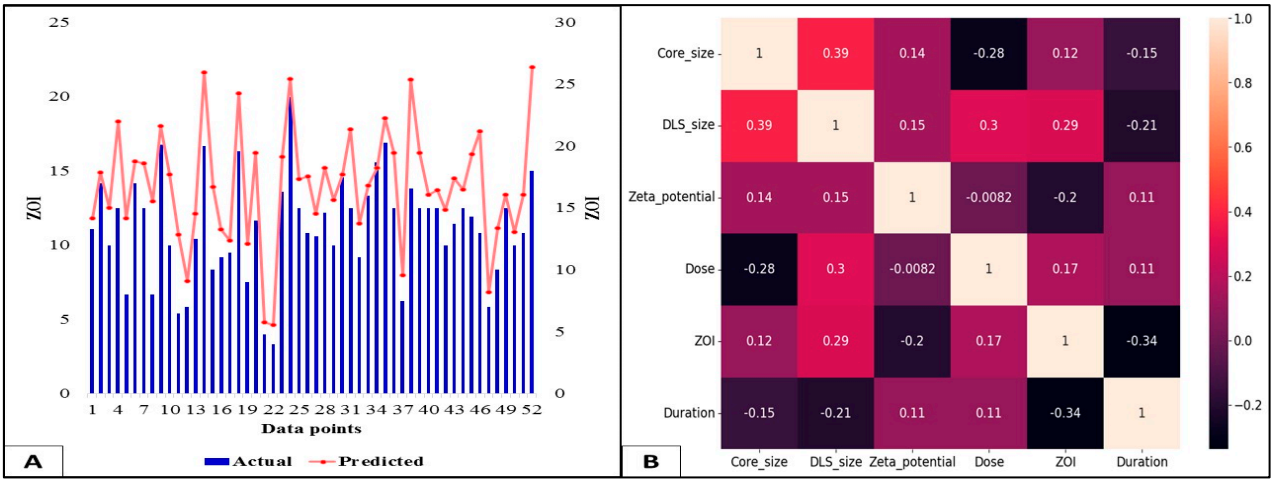

Figure 2. (**A**) Scatter plots of predicted vs. actual data points, (**B**) Heatmap of attribute significance.

3.2. Model Validation and Attribute Importance Analysis

The results of modeling reported that the developed model exhibits low MAE and MSE (4.5 and 1.6) and R^2 equal to 0.73 (Figure 3A). The predicted data points trend is so smilar to that of actual ones. Figure 3B presents that ZOI is exteremely dependent on the DLS-size, Core-size and exposure dose with the ratios of 0.29, 0.12 and 0.17 respectively. But other factors like Zeta_potential and Duration did have any impact on the anti-MDR effects of AgNPs. Besides to the impact of physical and chemical features of AgNPs, species information of the MDR bacteria is of a great importance too.

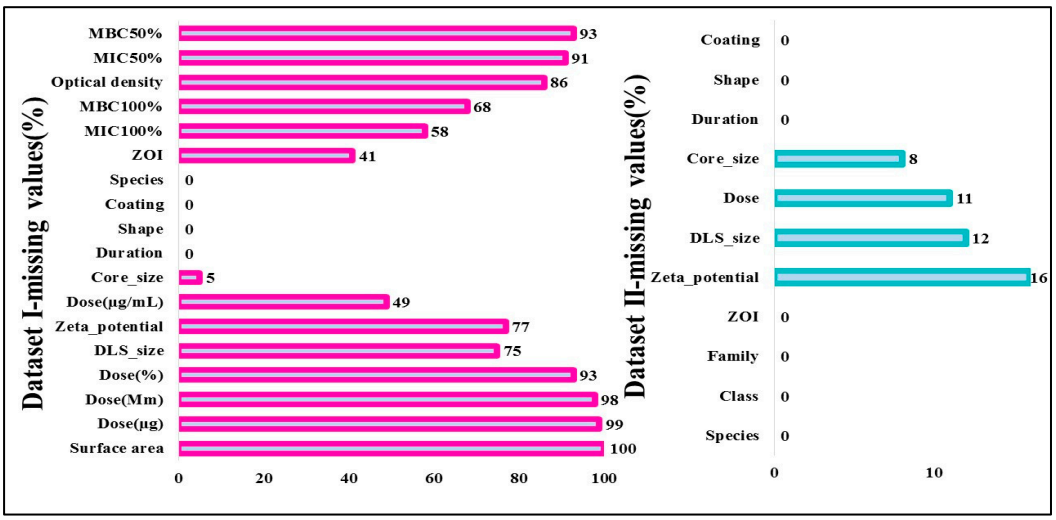

Figure 3. The level of missing values in dataset I (**Left**) and dataset II (**Right**).

4. Discussion

A machine learning model was developed in the present investigation. The model was subjected to validation to predict the anti-MDR activities of AgNPs correctly. Sucessfully, the model is in agreement with OECD rules [21] which suggests (1) a reasonable target (ZOI: Zone of Inhibition as an indicator for inhibition of a bacterial growth by AgNPs); (2) a developed algorithm (RF model, provided in supplementary file as S1); (3) an accurate framework adressing data ranges, nomical and categorical columns and high accurate validation scores. Recent investigations have presented that physical and chemical features like core-size [22], morphology [23], surface [24], surface charge [25] and hydrodynamic diameter [26] are important factors in determing the antibacterial activities of nanoparticles. Exposure dose and duration are among other determing factors [9]. Thus, it was needed to collect data about the physical and chemical properties plus exposure conditions. Consequently, numerous missing data occurred almost in all area. It is well-known that the regression models can not handle the missing data. The physical and chemical properties of a great importance, they should be kept. Another issue is the encodig of coating information into uncoated and coated, inibiting the over-fit of model. One of the challenges in determing chemical and physical properties is that there is no standard workflow for obtaining homogenous data which is vital for prediction. The reviewed articles gathered information about the anti-MDR activities of AgNPs following various strategies. In order to have a homogenous and comprehensive dataset, the characterization and antibacterial experiments should follow a standard protocol. To have consistant data, only one anti-MDR protocol was selected to gather single formed data. This choice confirms the requisite for having a standard anti-MDR method to have a reproducible and reliable model development. Random Forest is an effective model for having prediction in microbiology and genetics due to its strength for working with multi-dimensional data [13]. The results of our study presented a reliable and suitable MSE, MAE and R^2 scores which are comparable with the results of another RF model for the prediction of antibacterial activities of nanoparticles by Mirzaei et al. [9]. There are some strategies for preprocessing stage that help us to normalize data, to transform categorical features to numerical like one-hot encoding or handling missing data. There is not ay machine learning study specifically on the anti-MDR activities of AgNPs. Up now, antibacterial activities of nanoparticles have been studied with different models [9]. Due to the similar important attributes in determining antibacterial activities of nanoparticles and anti-MDR activities of them, the strategy of the previous studies were chosen. Various studies have reported that the DLS-size, dose, Core_size shape and surface area are significant factors [27,28]. Based on the researches, the scientists believe the small-sized nanoparticles are more vigorous

to penerate the cell wall of bacteria and fight against MDR barriers [29]. Additionaly, it is stated in some literatures this is the hydrodynamic diameter of a nanoparticle which is so important in revealing antibacterial effects not the core size [26]. Other investigations have shown that the antibactrial activities are tightly dependent on the species of the bacteria due to the difference in the cell wall structure of a gram-positive and gram-negative species [24]. Furthermore, the antibacterial effects are closely dose-dependent. This means: high exposure dose leads to better anti-MDR activity [30].

5. Conclusions

The microbial resistance against antibiotics has turned to be a big challenge in the world. Nnanoparticles exhibit promising antibacterial effects with numerous antibacterial mechanisms. At the present investigation, the a machine learning-based model; Random Forest was introduced to predict the anti-MDR activities of AgNPs. This is the first step in addressing a tool to help the scientists for choosing the best anti-MDR AgNPs and maniupulting the strategies for gaining the best anti-MDR AgNPs along cutting the expenses. Due to the presence of numerous gaps in database, it is emphasized to update and set the standard measurements.

Aknowledgements: This research was funded by Alzahra University, Tehran, Iran under a post-doctoral scholar.

References

1. Teixeira, M.C.; Sanchez-Lopez, E.; Espina, M.; Calpena, A.C.; Silva, A.M.; Veiga, F.J.; Garcia, M.L.; Souto, E.B. Advances in antibiotic nanotherapy: Overcoming antimicrobial resistance. *Emerg. Nanotechnol. Immunol.* **2018**, 233–259. https://doi.org/10.1016/B978-0-323-40016-9.00009-9.
2. Bartlett, J.G.; Gilbert, D.N.; Spellberg, B. Seven ways to preserve the miracle of antibiotics. *Clin. Infect. Dis.* **2013**, *56*, 1445–1450.
3. Adeniji, F. Global analysis of strategies to tackle antimicrobial resistance. *Int. J. Pharm. Pract.* **2018**, *26*, 85–89.
4. Gharpure, S.; Akash, A.; Ankamwar, B. A review on antimicrobial properties of metal nanoparticles. *J. Nanosci. Nanotechnol.* **2020**, *20*, 3303–3339.
5. Basavegowda, N.; Baek, K.-H. Multimetallic nanoparticles as alternative antimicrobial agents: Challenges and perspectives. *Molecules* **2021**, *26*, 912.
6. Allahverdiyev, A.M.; Abamor, E.S.; Bagirova, M.; Rafailovich, M. Antimicrobial effects of TiO_2 and Ag_2O nanoparticles against drug-resistant bacteria and leishmania parasites. *Future Microbiol.* **2011**, *6*, 933–940.
7. Kotrange, H.; Najda, A.; Bains, A.; Gruszecki, R.; Chawla, P.; Tosif, M.M. Metal and metal oxide nanoparticle as a novel antibiotic carrier for the direct delivery of antibiotics. *Int. J. Mol. Sci.* **2021**, *22*, 9596.
8. Dove, A.S.; Dzurny, D.I.; Dees, W.R.; Qin, N.; Nunez Rodriguez, C.C.; Alt, L.A.; Ellward, G.L.; Best, J.A.; Rudawski, N.G.; Fujii, K. Silver nanoparticles enhance the efficacy of aminoglycosides against antibiotic-resistant bacteria. *Front. Microbiol.* **2023**, *13*, 1064095.
9. Mirzaei, M.; Furxhi, I.; Murphy, F. A Machine Learning Tool to Predict the Antibacterial Capacity of Nanoparticles. *Nanomaterials* **2021**, *11*, 1774.
10. Goldberg, E.; Scheringer, M.; Bucheli, T.D.; Hungerbühler, K. Prediction of nanoparticle transport behavior from physicochemical properties: Machine learning provides insights to guide the next generation of transport models. *Environ. Sci. Nano* **2015**, *2*, 352–360.
11. Das, K.P.; Chandra, J. Nanoparticles and convergence of artificial intelligence for targeted drug delivery for cancer therapy: Current progress and challenges. *Front. Med. Technol.* **2022**, *4*, 1067144.
12. Lin, Z.; Chou, W.-C.; Cheng, Y.-H.; He, C.; Monteiro-Riviere, N.A.; Riviere, J.E. Predicting nanoparticle delivery to tumors using machine learning and artificial intelligence approaches. *Int. J. Nanomed.* **2022**, *17*, 1365–1379.
13. Mirzaei, M.; Furxhi, I.; Murphy, F.; Mullins, M. A supervised machine-learning prediction of textile's antimicrobial capacity coated with nanomaterials. *Coatings* **2021**, *11*, 1532.
14. Saadat, A.; Dehghani Varniab, A.; Madani, S.M. Prediction of the Antibacterial Activity of the Green Synthesized Silver Nanoparticles against Gram Negative and Positive Bacteria by using Machine Learning Algorithms. *J. Nanomater.* **2022**, *2022*, 4986826.
15. Polli, J.E. In vitro studies are sometimes better than conventional human pharmacokinetic in vivo studies in assessing bioequivalence of immediate-release solid oral dosage forms. *AAPS J.* **2008**, *10*, 289–299.
16. Beale, E.M.L.; Little, R.J.A. Missing values in multivariate analysis. *J. R. Stat. Soc. Ser. B* **1975**, *37*, 129–145.
17. Potdar, K.; Pardawala, T.S.; Pai, C.D. A comparative study of categorical variable encoding techniques for neural network classifiers. *Int. J. Comput. Appl.* **2017**, *175*, 7–9.

18. Ali, P.J.M.; Faraj, R.H.; Koya, E.; Ali, P.J.M.; Faraj, R.H. Data normalization and standardization: A technical report. *Mach. Learn. Tech. Rep.* **2014**, *1*, 1–6.
19. Qi, Y. Random forest for bioinformatics. In *Ensemble Machine Learning*; Springer: New York, NY, USA, 2012; pp. 307–323.
20. Mirzaei, M.; Furxhi, I.; Murphy, F.; Mullins, M. Employing Supervised Algorithms for the Prediction of Nanomaterial's Antioxidant Efficiency. *Int. J. Mol. Sci.* **2023**, *24*, 2792.
21. OECD, O. Environment health and safety publications series on testing and assessment No. 69, Guidance document on the validation of (quantitative) structure-activity relationships [(Q) SAR] models 2007.
22. Saeed, B.A.; Lim, V.; Yusof, N.A.; Khor, K.Z.; Rahman, H.S.; Abdul Samad, N. Antiangiogenic properties of nanoparticles: A systematic review. *Int. J. Nanomed.* **2019**, *14*, 5135–5146.
23. Cheon, J.Y.; Kim, S.J.; Rhee, Y.H.; Kwon, O.H.; Park, W.H. Shape-dependent antimicrobial activities of silver nanoparticles. *Int. J. Nanomed.* **2019**, *14*, 2773–2780.
24. Wang, L.; Hu, C.; Shao, L. The antimicrobial activity of nanoparticles: Present situation and prospects for the future. *Int. J. Nanomed.* **2017**, *12*, 1227–1249. https://doi.org/10.2147/IJN.S121956.
25. Maillard, A.P.V.F.; Espeche, J.C.; Maturana, P.; Cutro, A.C.; Hollmann, A. Zeta potential beyond materials science: Applications to bacterial systems and to the development of novel antimicrobials. *Biochim. Biophys. Acta (BBA)-Biomembranes* **2021**, *1863*, 183597.
26. Khurana, C.; Vala, A.K.; Andhariya, N.; Pandey, O.P.; Chudasama, B. Antibacterial activity of silver: The role of hydrodynamic particle size at nanoscale. *J. Biomed. Mater. Res. Part A* **2014**, *102*, 3361–3368.
27. Huq, M.A. Green Synthesis of Silver Nanoparticles Using Pseudoduganella eburnea MAHUQ-39 and Their Antimicrobial Mechanisms Investigation against Drug Resistant Human Pathogens. *Int. J. Mol. Sci.* **2020**, *21*, 1510. https://doi.org/10.3390/ijms21041510.
28. Krishna, P.G.; Chandra Mishra, P.; Naika, M.M.; Gadewar, M.; Ananthaswamy, P.P.; Rao, S.; Boselin Prabhu, S.R.; Yatish, K.V.; Nagendra, H.G.; Moustafa, M.; et al. Photocatalytic Activity Induced by Metal Nanoparticles Synthesized by Sustainable Approaches: A Comprehensive Review. *Front. Chem.* **2022**, *10*, 917831.
29. Menichetti, A.; Mavridi-Printezi, A.; Mordini, D.; Montalti, M. Effect of Size, Shape and Surface Functionalization on the Antibacterial Activity of Silver Nanoparticles. *J. Funct. Biomater.* **2023**, *14*, 244.
30. Pazos-Ortiz, E.; Roque-Ruiz, J.H.; Hinojos-Márquez, E.A.; López-Esparza, J.; Donohué-Cornejo, A.; Cuevas-González, J.C.; Espinosa-Cristóbal, L.F.; Reyes-López, S.Y. Dose-dependent antimicrobial activity of silver nanoparticles on polycaprolactone fibers against gram-positive and gram-negative bacteria. *J. Nanomater.* **2017**, *2017*, 4752314.