Benchmarking the physicochemical properties of 500 compounds for absorption, distribution, metabolic, excretion, and toxicity (ADMET) property prediction in *Caenorhabditis elegans*

[1]Regina K. L. Pinlac, [1]Lloyd E. T. Comia, [1]Gian N. T. Epino, [1]Rogie M. Fernandez, [1]Hannah S. Madrid, [1]Aaron S. R. Salvacion, [1]Cathleen J. M. Taloza, [1,2]John S. B. Nas

[1] Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, Manila, Philippines; [2] Department of Biology, College of Arts and Sciences, University of the Philippines Manila, Manila, Philippines. Corresponding author: J. S. B. Nas, jbnas@up.edu.ph

Abstract. *Caenorhabditis elegans* (*C. elegans*) is commonly used to screen plant extracts and purified compounds for their biological activity. The bioactivity of these compounds can be better elucidated if their bioavailability in the nematode is well understood. However, there are no established reference values for evaluating the Absorption, Distribution, Metabolic, Excretion, and Toxicity (ADMET) properties in *C. elegans*. Hence, this study evaluated nine physicochemical properties to establish a reference value for potential lead compounds in *C. elegans*. A total of 500 compounds previously used in *C. elegans*, which exhibited anticancer, antimicrobial, antifungal, and anti-neurodegenerative properties were investigated. The collected data were pooled to get the mean ±SE. The predicted reference values were compared with the existing parameters in evaluating the bioavailability of lead compounds in humans, namely Lipinski's Rule of Five (RO5), Ghose Filter Rule, Veber's Rule, and Rule of Three (RO3). The predicted ranges for the various physicochemical properties are the following: molecular weight (58.08 to 781 g mol^{-1}), MlogP (-4.47 to 6.92), H bond donor (0 to 8), H bond acceptor (0 to 17), WlogP (-3.54 to 8.02), number of atoms (4 to 53), rotatable bonds (0 to 12), PSA (0 to 270.86 Å). Overall, the predicted ranges of the physicochemical properties of the bioactive compounds used in *C. elegans* exceeded the existing reference values of the different parameters used in humans. However, the 500 compounds, on average, follow most of these rules, except RO3. These findings suggest that some compounds, which may have high bioavailability in *C. elegans*, may not be demonstrated in humans. Thus, these findings warrant further investigations.

Key Words: ADMET, anticancer, antifungal, antimicrobial, *Caenorhabditis elegans*, neurodegenerative diseases.

Introduction. For decades, *Caenorhabditis elegans* have been used as a model organism to decipher evolutionarily conserved pathways and processes associated with human diseases. *C. elegans*, as an inexpensive and straightforward disease model, was revealed to have 83% of the protein-coding genes homologous in humans (Burns et al 2010). Its complete genome sequence was also identified, and gene functions have been determined using numerous mutant phenotypes (Burns et al 2010). These instances add to the versatility of this organism, evident by the increasing involvement in various molecular studies on neurodegeneration, anticancer, antifungal, and anti-parasitic.

Approaches in drug discovery have become widespread due to the impact of increasing drug screening tools. *C. elegans* as a model in the initial screening of bioactive compounds better elucidates metabolic pathways and toxicity due to its mutant phenotypes (Carretero et al 2017). Additionally, these compounds necessitate absorption in the gut and diffusion to their target tissues for them to demonstrate a therapeutic effect (Kaletta & Hengartner 2006). There are characteristics of *C. elegans* that hinder the absorption and distribution of the compounds due to its physical and enzymatic

xenobiotic defenses (Burns et al 2010). Its four-layered cuticle acts as a selective barrier for water-soluble compounds during uptake in the intestinal lumen (Carretero et al 2017).

In drug design, the bioavailability of the compounds remains a challenge in predicting their pharmacological effect. It must conform to the pharmacokinetic parameters characterized by absorption, distribution, metabolism, excretion, and toxicity (ADMET) (Pressman et al 2017). Bioavailability remains an essential part of the pharmacokinetic parameter, because it predicts the amount of substance absorbed in the gastrointestinal tract and to the systemic circulation before reaching the target site to exert its biological effect (Rein et al 2013). Its physicochemical properties predict the bioavailability of the compounds used in humans.

The nine physicochemical properties that are used in evaluating the bioavailability of compounds are molecular weight, octanol-water partition coefficient (MlogP), hydrogen bond donor (H-bond donor), hydrogen bond acceptor (H-bond acceptor), atom-based calculation of partition coefficient (WlogP), molar refractivity, number of atoms, rotatable bonds, and polar surface area (PSA) (Nas 2020a). These physicochemical properties serve as a guide in predicting the drug-likeness of a compound.

Lipinski's Rule of Five (RO5) is the most well-adapted parameter to assess permeability and solubility (Ghose et al 1999). According to RO5, lead compounds should have the following desired values: molecular weight (≤500 g mol^{-1}), MlogP (≤4.15), H-bond donor (≤5), and H-bond acceptor (≤10) (Nas 2020a). Compounds that failed to satisfy the two desired values may have low absorption and permeation, indicating poor oral bioavailability (Pollastri 2010). Moreover, permeability pertains to the ability of a drug to cross intestinal epithelium, whereas solubility pertains to the drug's entry into the systemic circulation (Chandrasekaran et al 2018).

The Ghose Filter rule is also a suitable parameter in considering if a compound shows drug-likeness. Ghose suggested a lower range of molecular weight (160-480 g mol^{-1}), WlogP (-0.4 - 5.6), molar refractivity (40-130), and the number of atoms (20-70) (Desalermos et al 2011). The molar refractivity affects the volume of the compound, which in effect influences the London dispersive forces during its binding on a particular protein residue (Ghose et al 1999).

Veber's rule questioned the usual parameter set in molecular weight (≤500 g mol^{-1}) because it is not relevant enough to isolate the compounds with poor bioavailability. Hence, Veber's rule suggests that the number of rotatable bonds (≤10) and PSA (≤140) are more determinant in predicting the oral bioavailability of compounds (Veber et al 2002). Rotatable bonds are demonstrated by the molecule's flexibility (Veber et al 2002). This molecular flexibility is an essential quality because it is involved in facilitating molecules to easily transverse a membrane (Pajouhesh & Lenz 2005). On the other hand, PSA is the total H-bond count of a molecule. It affects the permeation rate in the membrane; thus, with a low PSA, permeation rate increases, which is also considered a desirable parameter for the bioavailability of the compounds.

The RO3 was recently established because of the innovative approach in constructing the drug-size compound libraries. This novel approach is called fragment-based discovery, wherein the hit compounds must adhere to the parameters set by the RO3, such as MlogP (≤3), molecular weight (<300), H-bond donor (≤3), H-bond acceptor (≤3), and rotatable bonds (≤3). Moreover, this rule helps create fragment libraries to discover efficient lead compounds (Congreve et al 2003). Lead compounds have a small molecular weight and less complex properties (Lipinski et al 1997). Hence, reduced complexity of compounds is a typical basis for discovering drugs (Holden-Dye & Walker 2014).

Zebrafish (*Danio rerio*) is a well-established model organism because of their evolutionary genetic similarities to humans (Desalermos et al 2011). A previous study raised questions about the physiological and genetic differences between fish and mammals. About 700 compounds were assessed and revealed to have an average molecular weight, number of H bond donor, H bond acceptor, number of rotatable bonds, and PSA lower than two reference drug sets (Desalermos et al 2011). Conversely, the average logP values of the compounds used in *D. rerio* is significantly higher than the two

drug groups, denoting a more lipophilic nature (Desalermos et al 2011). Overall, it shows that zebrafish absorbed molecules likely to be more lipophilic. The physicochemical properties of this species fall within the narrow range of values compared to the known drugs (Lagorce et al 2017). Zebrafish can be evaluated using multiple parameters, such as the biological effects, permeability, and toxicity. The use of the standardized ADMET properties in zebrafish reveals the oral bioavailability of other freshwater and marine organisms (Nas 2020b).

The ease of handling and low-cost maintenance of *C. elegans* makes it more advantageous than the zebrafish. It also has a high brooding rate, short reproductive span, and short lifespan, which helps understand the molecular mechanisms of lengthening fertility and longevity of compounds. This nematode is also an effective model for this study because of its short life span. The genetic tractability in this nematode makes it easier to conduct forward and reverse genetic screenings to study various molecular pathways, such as aging, cancer, and neurodegenerative diseases (Nas et al 2019).

This study evaluates the physicochemical properties of 500 compounds used in *C. elegans* and benchmarks their ADMET properties to establish a bioavailability prediction parameter in this model organism.

Material and Method. We retrieved 500 compounds from different studies using *C. elegans*. The keywords that we used in searching research articles include the following: "ANTICANCER" and "CAENORHABDITIS ELEGANS" or "ANTIFUNGAL" and "CAENORHABDITIS ELEGANS", or "ANTIMICROBIAL" and "CAENORHABDITIS ELEGANS" or "NEURODEGENERATION" and "CAENORHABDITIS ELEGANS" or "NEUROPROTECTION" and "CAENORHABDITIS ELEGANS" or "PARKINSON'S" and "CAENORHABDITIS ELEGANS" or "ALZHEIMER'S" and "CAENORHABDITIS ELEGANS". We used different indexing platforms, namely Google Scholar, Mendeley, Research Gate, Science Direct, and PubMed to gather these articles. After compiling the compounds, only those with available 3D structures in PubChem were used.

The compounds used in this study were assessed based on the physicochemical properties, namely molecular weight, MlogP, H-bond donor, H-bond acceptor, WlogP, molar refractivity, number of atoms, rotatable bonds, and PSA. We retrieved the values of each physicochemical property from SwissADME (http://www.swissadme.ch/). Under the RO5, the physicochemical properties are molecular weight, MlogP, H-bond donor, and acceptor. In comparison, the physicochemical properties evaluated in the Ghose filter are WlogP, molar refractivity, molar weight, and the number of atoms. Moreover, in Veber's rule, rotatable bonds and PSA are physicochemical properties. Furthermore, the RO3 parameters are MLlogP, molecular weight, rotatable bonds, H-bond donor, and H-bond acceptor.

The compounds were evaluated by classifying them based on different categories such as anticancer, antifungal, antimicrobial, and neurodegenerative. Under these classifications, we also computed their respective mean values for each physicochemical property. Also, we determined the outliers in each category through box and whisker plots. We transformed the data set to have a homogenous and normally distributed adjusted mean value per classification and overall compounds. We tested the adjusted data set for homogeneity and normality using Levene's test and Shapiro-Wilk test, respectively. Data found to be both homogenous and normally distributed were further evaluated with a one-way analysis of variance (ANOVA) using Tukey's test for post-hoc analysis. The level of significance was set at $p<0.05$. The processing of data was carried out using Statistical Package for the Social Sciences (SPSS) v. 17.0 and Microsoft Excel 365.

Results and Discussion

General overview of the physicochemical properties of the 500 compounds. Figure 1 shows the general classification of compounds used in *C. elegans*. Out of the 500 compounds examined, neurodegenerative-associated drugs have the highest count

with 198 compounds, followed by antimicrobial with 150 compounds identified. The anticancer and antifungal have only 96 and 56 counts, respectively. Overall, the figure indicates that several compounds screened in *C. elegans* were used against neurodegenerative diseases, compared to anticancer, antimicrobial, and antifungal studies.

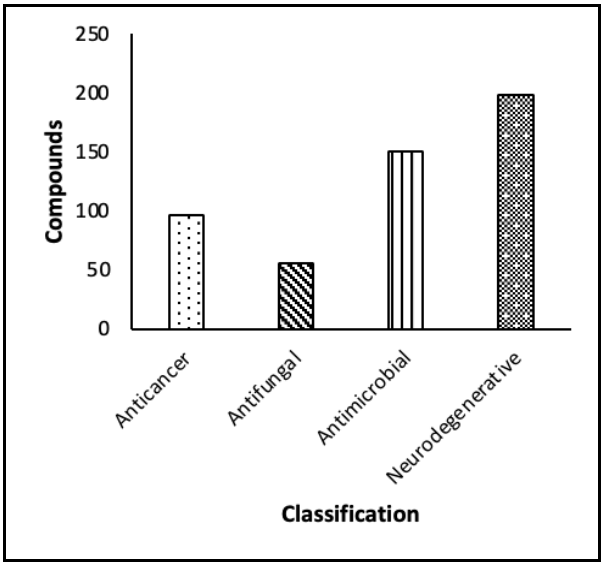

Figure 1. General classification of the compounds used in *Caenorhabditis elegans*.

The mean molecular weight, H-bond donor, H-bond acceptor, molar refractivity, number of atoms, rotatable bonds, and PSA of anticancer, antimicrobial, antifungal, neurodegenerative, and overall compounds are presented in Table 1. The mean of these compounds was higher than their respective median, as shown in Figure 2a-i.

Figure 2a shows the molecular weight of all compounds and in each classification (anticancer, antimicrobial, antifungal, neurodegenerative). There are 40 outliers in the overall compounds. Antifungal compounds have no outliers in each category, whereas neurodegenerative have nine outliers that generate the most outliers within the classifications.

The MlogP in each classification (anticancer, antimicrobial, antifungal, neurodegenerative) and overall compounds were shown in Figure 2b. There are 36 outliers within all compounds. Within the four classifications, the neurodegenerative compounds displayed the most number of outliers, with 9. On the other hand, antifungal compounds have two outliers, making them the least numbered of outliers within the classifications.

Figure 2c shows the H-bond donor of all compounds, as well as the four classifications (anticancer, antimicrobial, antifungal, neurodegenerative). All compounds generate 36 outliers. While under the four classifications, the highest number of outliers is neurodegenerative, which has nine counts. In contrast, anticancer has the least number of outliers with three.

Figure 2d shows the H-bond acceptor of the overall compounds and per classification (anticancer, antimicrobial, antifungal, neurodegenerative). There were a total of 23 outliers in the 500 compounds. Within the four classifications, neurodegenerative contributes with the highest number of outliers (12), followed by antimicrobial (6), antifungal (2), and anticancer (1).

The WlogP of all compounds and in each classification (anticancer, antimicrobial, antifungal, neurodegenerative) was shown in Figure 2e. There were 27 overall outliers. Within each category, neurodegenerative contributes the highest value of outliers which has 13, followed by antimicrobial (8), antifungal (2), and only one outlier in anticancer.

Figure 2f shows the molar refractivity of all compounds and per classification. The overall compounds have 40 outliers, while the antimicrobial compounds have ten,

followed by neurodegenerative (9) and anticancer (6). Moreover, antifungal compounds have no outliers at all.

Table 1

Mean ±SE of the physicochemical properties of the 500 compounds used in *Caenorhabditis elegans*

	n	*Physicochemical property*	*Mean ±SE*
Overall	500	Molecular Weight	412.41±13.47
		MLOGP	1.03±0.13
		H-bond Donor	3.12±0.20
		H-bond Acceptor	6.212±0.26
		WLOGP	2.07±0.14
		Molar refractivity	108.60±3.46
		Number of Atoms	28.02±0.82
		Rotatable bonds	5.73±0.48
		PSA	115.02±5.64
Anticancer	96	Molecular Weight	419.21±21.45
		MLOGP	0.52±0.25
		H-bond Donor	3.79±0.35
		H-bond Acceptor	6.97±0.47
		WLOGP	2.12±0.28
		Molar refractivity	112.19±5.71
		Number of Atoms	29.85±1.54
		Rotatable bonds	5.84±0.63
		PSA	131.92±9.35
Antimicrobial	150	Molecular Weight	431.45±23.26
		MLOGP	0.84±0.25
		H-bond Donor	3.21±0.38
		H-bond Acceptor	6.57±0.44
		WLOGP	1.87±0.27
		Molar refractivity	109.23±6.12
		Number of Atoms	29.17±1.63
		Rotatable bonds	6.25±0.93
		PSA	129.66±10.49
Antifungal	56	Molecular Weight	522.64±46.57
		MLOGP	0.74±0.38
		H-bond Donor	4.04±0.65
		H-bond Acceptor	8.09±0.97
		WLOGP	1.78±0.41
		Molar refractivity	139.52±12.27
		Number of Atoms	36.46±3.31
		Rotatable bonds	6.71±0.38
		PSA	142.48±17.64
Neurodegenerative	198	Molecular Weight	363.47±23.26
		MLOGP	1.51±0.19
		H-bond Donor	2.48±0.34
		H-bond Acceptor	5.05±0.43
		WLOGP	2.27±0.19
		Molar refractivity	97.64±5.8
		Number of Atoms	23.87±1.05
		Rotatable bonds	5.01±0.90
		PSA	87.96±9.44

Figure 2g represents the outliers of the number of atoms of the overall compounds and per classification. There were 40 outliers in overall compounds. In each classification, antimicrobial and neurodegenerative have the most numbered outliers with a count of 9, followed by anticancer (7), and antifungal has no outliers at all.

Figure 2h shows the rotatable bonds in all compounds and in per classification. Overall, there were 37 outliers. Further, neurodegenerative exhibited the most numbered

outliers in each classification with a count of 11. On the other hand, we found only one outlier in antifungal, 8 in antimicrobial, and 6 in anticancer.

The PSA of the overall compounds and in each classification was shown in Figure 2i. We found 33 outliers in the overall compounds. At the same time, in each classification, neurodegenerative has the most numbered outliers with a count of 12, followed by antimicrobial (11), antifungal (4), and anticancer (2).

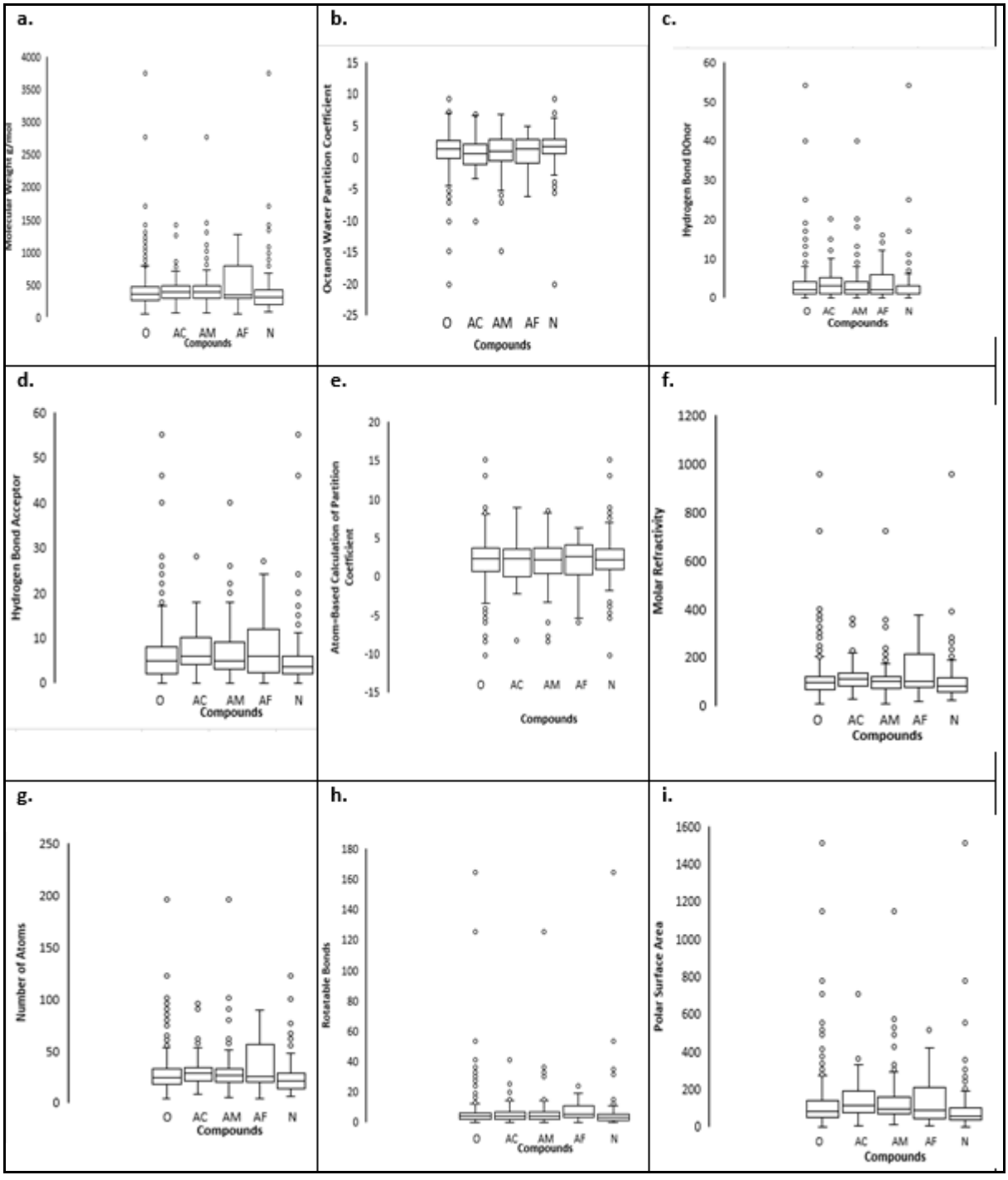

Figure 2. Box and whisker plot - physicochemical properties of compounds used in *Caenorhabditis elegans* in each classification; a - molecular weight; b - MlogP; c - H-bond donor; d - H-bond acceptor; e - WlogP; f - molar refractivity; g - number of atoms; h - rotatable bonds; i - PSA; O - overall; AC - anticancer; AM - antimicrobial; AF - antifungal; N - neurodegenerative.

Transformed mean of the physicochemical properties of the 500 compounds. We showed the adjusted mean ±SE of the different physicochemical properties in Table 2 and Figure 3, after transforming the dataset.

Table 2

Range of the adjusted mean ±SE of the 500 compounds' physicochemical properties

Physicochemical Properties	*Mean ±SE*	*Range (Min-Max)*
Molecular Weight		
Anticancer	375.74±13.85	70.62 - 705.63
Antifungal	522.64±46.57	58.08 - 1270.27
Antimicrobial	378.44±11.37	70.05 - 733.93
Neurodegenerative	307.44±9.47	79.1 - 679.78
Overall	346.79±6.47	58.08 - 781
Octanol-water Partition Coefficient (MLOGP)		
Anticancer	0.50±0.21	-3.4 - 6.54
Antifungal	1.00±0.34	-5.33 - 4.89
Antimicrobial	1.26±0.20	-5.34 - 6.73
Neurodegenerative	1.75±0.13	-2.77 - 6.14
Overall	1.31±0.10	-4.47 - 6.92
H-bond Donor		
Anticancer	3.41±0.28	0 - 10
Antifungal	2.13±0.32	0 - 8
Antimicrobial	2.13±0.15	0 - 8
Neurodegenerative	1.60±0.11	0 - 6
Overall	2.19±0.09	0 - 8
H-bond Acceptor		
Anticancer	6.75±0.42	0 - 18
Antifungal	7.39±0.87	0 - 24
Antimicrobial	5.83±0.31	0 - 17
Neurodegenerative	3.90±0.19	0 - 11
Overall	5.32±0.17	0 - 17
Atom-based Calculation of Partition Coefficient (WLOGP)		
Anticancer	2.16±0.26	-2.31 - 8.48
Antifungal	1.92±0.39	-5.43 - 6.29
Antimicrobial	2.29±0.21	-3.42 - 8.21
Neurodegenerative	2.24±0.13	-1.79 -7
Overall	2.27±0.10	-3.54 - 8.02
Molar refractivity		
Anticancer	101.98±3.99	27.64 - 205.25
Antifungal	139.52±12.27	16.38 - 375.47
Antimicrobial	94.52±2.85	6.79 - 174.23
Neurodegenerative	85.54±2.66	24.24 - 188.23
Overall	91.88±1.70	6.79 - 194.53
Number of Atoms		
Anticancer	26.74±1.02	8 - 53
Antifungal	36.46±3.31	4 - 89
Antimicrobial	25.43±0.76	5 - 51
Neurodegenerative	21.51±0.67	6 - 48
Overall	23.86±0.45	4 - 53
Rotatable bonds		
Anticancer	4.59±0.34	0 - 14
Antifungal	6.4±0.61	0 - 19
Antimicrobial	4.46±0.28	0 - 14
Neurodegenerative	3.25±0.20	0 -11
Overall	3.98±0.14	0 - 12
Polar Surface Area		
Anticancer	123.36±6.81	3.24 - 331.14
Antifungal	119.31±14.53	3.24 - 377.42
Antimicrobial	102.79±5.04	8.81 - 282.61
Neurodegenerative	65.31±3.13	0 – 189.53

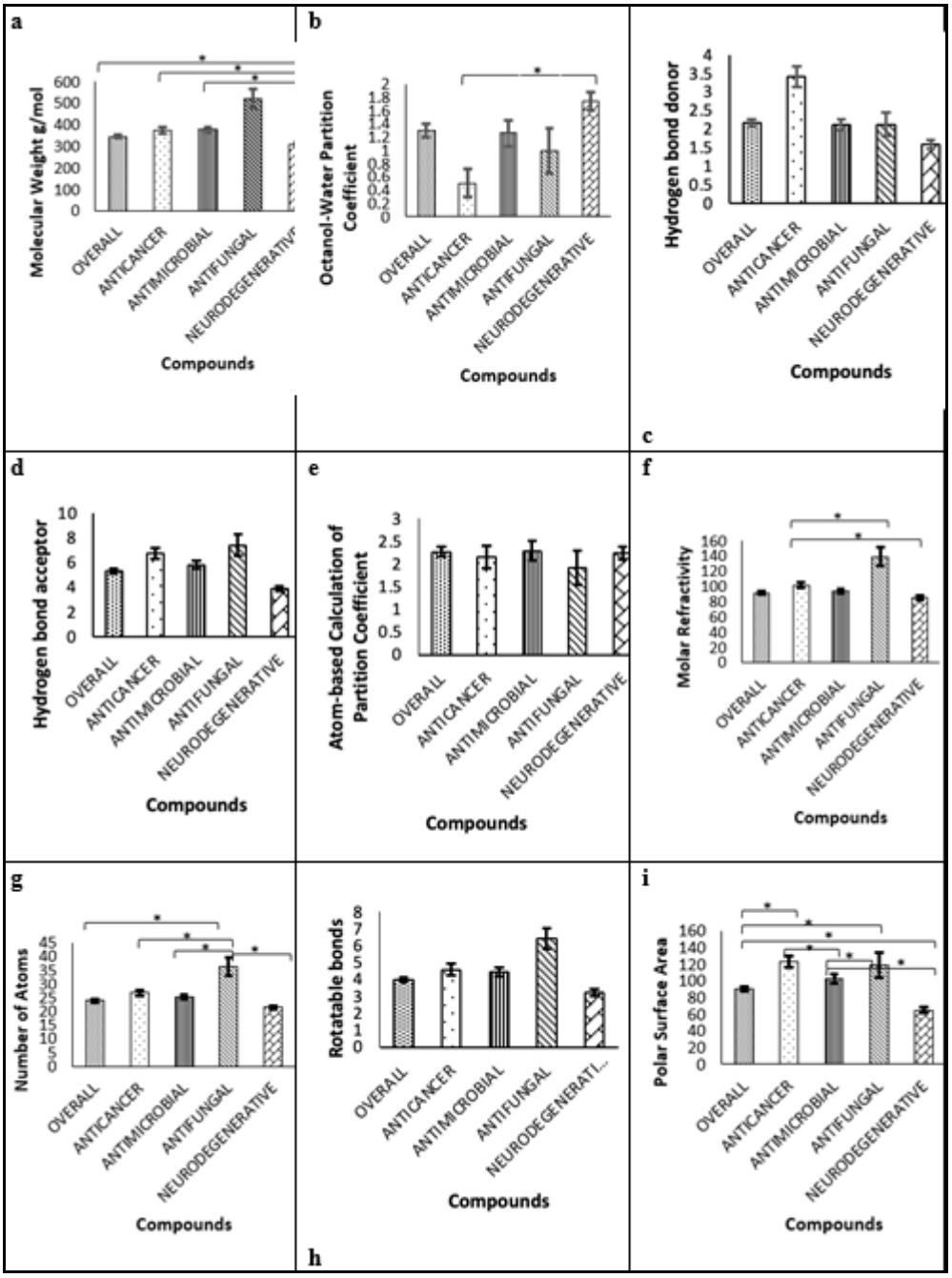

Figure 3. Adjusted mean ±SE of each physicochemical properties of the 500 compounds. a - molecular weight; b - MlogP; c - H-bond donor; d - H-bond acceptor; e - WlogP; f - molar refractivity; g - number of atoms; h - rotatable bonds; i - PSA.

The overall molecular weight of the 500 compounds evaluated was 346.79±6.47 g mol^{-1}. This mean is significantly higher than the mean of the neurodegenerative compounds by 11% ($p<0.05$), as shown in Table 3. Additionally, the anticancer and antimicrobial compounds are 22% and 23% higher ($p<0.05$) than neurodegenerative drugs.

The MlogP of the 500 compounds has an average of 1.31±0.10, as shown in Table 2. This value is comparable with the mean ±SE of the other compound, as shown in Figure 3b. Interestingly, the MlogP of the neurodegenerative compounds is 71% higher than those from anticancer ($p<0.05$), as shown in Table 3.

Table 3

Comparison of the mean± SE of the physicochemical properties of the 500 compounds

Physicochemical properties	Comparison	p-value	
Molecular Weight	Overall vs Anticancer	0.6032	ns
	Overall vs Antimicrobial	0.449	ns
	Overall vs Antifungal	>0.9999	ns
	Overall vs Neurodegenerative	0.0175	s
	Anticancer vs Antimicrobial	0.4448	ns
	Anticancer vs Antifungal	>0.9999	ns
	Anticancer vs Neurodegenerative	0.0002	s
	Antimicrobial vs Antifungal	>0.9999	ns
	Antimicrobial vs Neurodegenerative	0.0002	s
	Antifungal vs Neurodegenerative	>0.9999	ns
MLOGP	Overall vs Anticancer	0.2422	ns
	Overall vs Antimicrobial	0.4856	ns
	Overall vs Antifungal	0.3931	ns
	Overall vs Neurodegenerative	0.3501	ns
	Anticancer vs Antimicrobial	0.1438	ns
	Anticancer vs Antifungal	0.2462	ns
	Anticancer vs Neurodegenerative	0.0412	s
	Antimicrobial vs Antifungal	0.4057	ns
	Antimicrobial vs Neurodegenerative	0.3343	ns
	Antifungal vs Neurodegenerative	0.2264	ns
Hydrogen Bond Donor	Overall vs Anticancer	0.2042	ns
	Overall vs Antimicrobial	0.4852	ns
	Overall vs Antifungal	0.4844	ns
	Overall vs Neurodegenerative	0.346	ns
	Anticancer vs Antimicrobial	0.2447	ns
	Anticancer vs Antifungal	0.2444	ns
	Anticancer vs Neurodegenerative	0.1638	ns
	Antimicrobial vs Antifungal	0.4992	ns
	Antimicrobial vs Neurodegenerative	0.3577	ns
	Antifungal vs Neurodegenerative	0.3583	ns
Hydrogen Bond Acceptor	Overall vs Anticancer	0.2676	ns
	Overall vs Antimicrobial	0.4125	ns
	Overall vs Antifungal	0.1846	ns
	Overall vs Neurodegenerative	0.2701	ns
	Anticancer vs Antimicrobial	0.3616	ns
	Anticancer vs Antifungal	0.4026	ns
	Anticancer vs Neurodegenerative	0.1372	ns
	Antimicrobial vs Antifungal	0.2591	ns
	Antimicrobial vs Neurodegenerative	0.2131	ns
	Antifungal vs Neurodegenerative	0.1004	ns
WLOGP	Overall vs Anticancer	0.47	ns
	Overall vs Antimicrobial	0.4934	ns
	Overall vs Antifungal	0.4081	ns
	Overall vs Neurodegenerative	0.4922	ns
	Anticancer vs Antimicrobial	0.4626	ns
	Anticancer vs Antifungal	0.436	ns
	Anticancer vs Neurodegenerative	0.4773	ns
	Antimicrobial vs Antifungal	0.4023	ns
	Antimicrobial vs Neurodegenerative	0.4857	ns
	Antifungal vs Neurodegenerative	0.4085	ns

Note: s - significantly different at $p<0.05$; ns - not significantly different.

Table 3

Comparison of the mean± SE of the physicochemical properties of the 500 compounds (continuation)

Physicochemical properties	*Comparison*	*p-value*	
Molar Refractivity	Overall vs Anticancer	0.1461	ns
	Overall vs Antimicrobial	0.3914	ns
	Overall vs Antifungal	>0.9999	ns
	Overall vs Neurodegenerative	0.2543	ns
	Anticancer vs Antimicrobial	0.2303	ns
	Anticancer vs Antifungal	0.0001	s
	Anticancer vs Neurodegenerative	0.0523	s
	Antimicrobial vs Antifungal	0.9999	ns
	Antimicrobial vs Neurodegenerative	0.1781	ns
	Antifungal vs Neurodegenerative	0.9999	ns
Number of Atoms	Overall vs Anticancer	0.2778	ns
	Overall vs Antimicrobial	0.3744	ns
	Overall vs Antifungal	0.0051	s
	Overall vs Neurodegenerative	0.3155	ns
	Anticancer vs Antimicrobial	0.3997	ns
	Anticancer vs Antifungal	0.031	s
	Anticancer vs Neurodegenerative	0.1564	ns
	Antimicrobial vs Antifungal	0.0149	s
	Antimicrobial vs Neurodegenerative	0.2192	ns
	Antifungal vs Neurodegenerative	0.007	s
Rotatable Bonds	Overall vs Anticancer	0.3812	ns
	Overall vs Antimicrobial	0.4064	ns
	Overall vs Antifungal	0.1134	ns
	Overall vs Neurodegenerative	0.3557	ns
	Anticancer vs Antimicrobial	0.4756	ns
	Anticancer vs Antifungal	0.1996	ns
	Anticancer vs Neurodegenerative	0.2656	ns
	Antimicrobial vs Antifungal	0.1794	ns
	Antimicrobial vs Neurodegenerative	0.2832	ns
	Antifungal vs Neurodegenerative	0.1069	ns
Polar Surface Area	Overall vs Anticancer	0.0003	s
	Overall vs Antimicrobial	0.1045	ns
	Overall vs Antifungal	0.0015	s
	Overall vs Neurodegenerative	0.0038	s
	Anticancer vs Antimicrobial	0.0326	s
	Anticancer vs Antifungal	0.3581	ns
	Anticancer vs Neurodegenerative	>0.9999	ns
	Antimicrobial vs Antifungal	0.0524	s
	Antimicrobial vs Neurodegenerative	0.0001	s
	Antifungal vs Neurodegenerative	>0.9999	ns

Note: s - significantly different at $p<0.05$; ns - not significantly different.

The molar refractivity of the 500 compounds evaluated in antifungal was 139.52±12.27, as shown in Table 2. This is significantly higher than the mean of the anticancer compounds by 26% ($p<0.05$). However, anticancer compounds are significantly higher than the mean of neurodegenerative compounds by 16% ($p<0.05$) based on Table 3 and Figure 3f.

The computed adjusted mean of the number of atoms in antifungal was 36.46±3.31, as shown in Table 2. This is significantly higher than the mean of overall compounds, anticancer, antimicrobial, and neurodegenerative compounds by 34%, 26%, 30%, and 41% ($p<0.05$) as shown in Table 3 and Figure 3g, respectively.

The overall adjusted mean of the polar surface area evaluated was 90.81±2.74, significantly higher than the mean of neurodegenerative compounds by 28%($p<0.05$), as shown in Table 2. This value is comparable with the mean ±SE of the other compound, as shown in Figure 3i. Meanwhile, the mean of anticancer and antifungal compounds is

higher than the overall mean by 26% and 23% ($p<0.05$), respectively. The mean of anticancer compounds, 123.36±6.81, is higher than that of antimicrobial compounds by 16% ($p<0.05$). However, the mean of antimicrobial compounds (102.79±5.04) is higher than the mean neurodegenerative compounds by 36%, but lower than antifungal compounds by 16% ($p<0.05$) (Table 3).

Comparison of the predicted values of each physicochemical property with the established parameters in humans. The predicted values and the reference values for molecular weight, MlogP, H-bond donor, H-bond acceptor, WlogP, molar refractivity, number of atoms, rotatable bonds, and PSA are presented in Table 4.

Table 4

Comparison between the Reference and Computed Range of each Physicochemical Properties of the 500 compounds

Physicochemical Properties		Range	Reference Value RO5	Ghose Filter	Veber's	RO3
Molecular Weight	Anticancer	70.62 - 705.63	≤500	160 to 480		<300
	Antifungal	58.08 - 1270.27				
	Antimicrobial	70.05 - 733.93				
	Neurodegenerative	79.1 - 679.78				
	Overall	58.08 - 781				
MlogP	Anticancer	-3.4 - 6.54	≤5			≤3
	Antifungal	-5.33 - 4.89				
	Antimicrobial	-5.34 - 6.73				
	Neurodegenerative	-2.77 - 6.14				
	Overall	-4.47 - 6.92				
H-bond Donor	Anticancer	0 - 10	≤5			≤3
	Antifungal	0 - 8				
	Antimicrobial	0 - 8				
	Neurodegenerative	0 - 6				
	Overall	0 - 8				
H-bond Acceptor	Anticancer	0 - 18	≤10			≤3
	Antifungal	0 - 24				
	Antimicrobial	0 - 17				
	Neurodegenerative	0 - 11				
	Overall	0 - 17				
WlogP	Anticancer	-2.31 - 8.48		-0.4 to 5.6		
	Antifungal	-5.43 - 6.29				
	Antimicrobial	-3.42 - 8.21				
	Neurodegenerative	-1.79 -7				
	Overall	-3.54 - 8.02				
Molar Refractivity	Anticancer	27.64 - 205.25		40 to 130		
	Antifungal	16.38 - 375.47				
	Antimicrobial	6.79 - 174.23				
	Neurodegenerative	24.24 - 188.23				
	Overall	6.79 - 194.53				
Number of Atoms	Anticancer	8 - 53		20 to 70		
	Antifungal	4 - 89				
	Antimicrobial	5 - 51				
	Neurodegenerative	6 - 48				
	Overall	4 - 53				
Rotatable Bonds	Anticancer	0 - 14			≤10	≤3
	Antifungal	0 - 19				
	Antimicrobial	0 - 14				
	Neurodegenerative	0 -11				
	Overall	0 - 12				
PSA	Anticancer	3.24 - 331.14			≤140	
	Antifungal	3.24 - 377.42				
	Antimicrobial	8.81 - 282.61				
	Neurodegenerative	0 – 189.53				
	Overall	0 – 270.86				

The predicted values were obtained from the adjusted range when we transformed the outliers. Based on the generated adjusted range, the predicted values exceed the established parameters, namely, RO5, Ghose Filter Rule, Veber's Rule, and RO3, except for the number of atoms whose predicted value is much lower. These parameters are used for evaluating the human oral bioavailability of compounds.

Despite the range exceeding the reference values, the average value of the compounds (in Table 2) falls within the reference value. In addition, the mean molecular weight of the antifungal compounds violated the molecular weight of RO5, Ghose filter, and RO3, whereas the mean of the other compounds violated only the RO3. Only the mean H bond donor value of the anticancer compounds violated RO3. The mean values of the compounds exceeded the H bond acceptor value of RO3. Additionally, the mean molar refractivity of the antifungal compounds surpassed the desired range in the Ghose Filter rule. Despite the compounds following the desired value of the number of rotatable bonds in Veber's rule, their mean value exceeded the desired value in RO3.

The 500 compounds gathered from different studies were proven to have their efficacy on *C. elegans*. Furthermore, numerous *C. elegans* strains employed in the various articles evaluated represented some human diseases, such as Parkinson's disease, Duchenne muscular disorder, pathogenic infection, and cancer (Nas et al 2020). Hence, *C. elegans* has been utilized as a typical animal screening model for anti-neurodegenerative, antimicrobial, anticancer, and antifungal research.

Moreover, using *C. elegans* as a model has various drug discovery benefits, including a high brooding rate, simple anatomy, short lifespan, genetic similarity in humans, and ease of genetic screening (O'Reilly et al 2014; Chen et al 2020; Markaki & Tavernarakis 2020). Several proof-of-concept studies have already been presented, such as the adaptability of *C. elegans* in compound screening, drug target identification, and decoding mechanisms of drug action (Pajouhesh & Lenz 2005). To simplify, this nematode model could be easily tested for the efficacy of novel drugs on complex processes involved in human diseases (Holden-Dye & Walker 2014). The physicochemical properties significantly impact its efficiency in animal models as it predicts its absorption, distribution, metabolism, clearance, and toxicity (Hou et al 2007).

We evaluated the physicochemical properties of the compounds to determine the bioavailability in terms of ADMET properties. The behavior of the compounds depends on the physicochemical properties, which have biological effects once orally administered (Chandrasekaran et al 2018). After the outliers were determined per classification, we gathered the range of the outliers. The compounds considered outliers in all physicochemical properties are bacitracin, bleomycin, caspofungin, cecropin A, colistin, lysozyme, and micafungin. Cecropin A, colistin, and lysozyme were antimicrobial compounds, while caspofungin and micafungin were antifungal. Additionally, bacitracin and bleomycin are classified as neurodegenerative and anticancer compounds, respectively. Even though these compounds were tested against *C. elegans*, they did not show any drug-likeness. It fails to satisfy the existing parameters established for human bioavailability.

Based on Table 2, the adjusted mean of the molecular weight, H-bond donor, H-bond acceptor, molar refractivity, number of atoms, rotatable bonds, and PSA of the overall and the four classifications are much lower than the value of their unadjusted mean. This implies that some compounds used in *C. elegans* may have low bioavailability in its body. The permeation, diffusion, absorption, metabolism, or excretion may be impeded in the body (Chandrasekaran et al 2018). Conversely, the adjusted mean of the MlogP and the WlogP of the overall compounds and the four classifications shows a higher value within their adjusted mean compared to their unadjusted mean. The outliers have lower values that cause the unadjusted mean to be lower. When transformed, the adjusted mean increased the threshold. Both properties determine the lipophilicity of the compound, which may indicate that some of the compounds below the adjusted range are more soluble (Lipinski et al 1997; Ghose et al 1999). This instance may result in the compound having low absorption or high clearance (Bhal 2007).

We compared the adjusted range of each physicochemical property per classification based on the different rules mentioned in the study. In terms of molecular

weight of overall compounds and four categories, the adjusted range was much higher. It did not satisfy the RO5 (≤500 g mol-1), Ghose Filter (160-480 g mol^{-1}), and RO3 (<300). Studies have shown that compounds with high molecular weight are poorly absorbed and quickly excreted (Lipinski et al 1997). These findings indicate that some compounds predicted to have high bioavailability in *C. elegans* may differ in humans.

The MlogP of the overall compounds and the other four classifications have an adjusted range exceeding the reference value for RO5 (≤4.15) and RO3 (≤3). Higher logP indicates a highly lipophilic compound, resulting in poor permeation or absorption because of low aqueous solubility (Bhal 2007).

Moreover, in RO5, H-bond donors should be ≤5, and for H-bond acceptors, the values should be ≤10. The physicochemical properties of some compounds may not satisfy this rule since their adjusted range for the overall compounds and for the other four classifications were relatively higher. However, the RO3 states that both physicochemical properties should be ≤3. Some compounds' H-bond donor and acceptor did not satisfy the rule as it may be comparably higher. H-bond donors significantly affect the passive diffusion of the cell membrane, which is an essential factor in drug absorption and distribution (Coimbra et al 2021). Compounds with a higher number of H-bond donors have a negative effect on the permeability and partition of the drug's membrane (Rafi et al 2012). It will not contribute during ligand binding and can detract from it. It can also decrease the affinity, affecting the hydrophobic membrane region, increasing water desolation when penetrating the blood. H-bond donor guides the solubility of the drugs, both in water and lipids, and the affinity with their targets (Rafi et al 2012).

The adjusted range of WlogP and molar refractivity did not satisfy the reference values used in the Ghose Filter Rule. This higher logP indicates poor permeation or absorption because they have decreased aqueous solubility (Bhal 2007). The molar refractivity of ions in a molecule represents the arrangements of their electron shells and provides information about their electronic polarization. Most of the biochemical process happens in aqueous media. Consequently, molar refraction and polarizability of aqueous drug compounds deliver notable details that may be essential in pharmaceutical and medicinal chemistry (Sawale et al 2016). Moreover, the adjusted molar refractivity value indicates that there is also an increase in molecular polarizability. In the Lorentz-Lorenz formula, this relationship also affects the volume and molecular weight of the compounds resulting from the overall compounds with greater values.

In addition, only the adjusted range of the number of atoms in antifungal compounds did not satisfy the Ghose Filter Rule (20-70) because it was much lower. Other classifications, including the overall compounds, satisfy the rule as their values are within. Compounds with a higher number of atoms have reduced absorptivity (Brenner & Stevens 2018). This property may also be affected by the polarity of the compounds (Rein et al 2013). Charged compounds tend to have lower cellular permeability, leading to a high excretion rate (Aldred 2008). In contrast, uncharged compounds are reabsorbed through cell membranes (Aldred 2008). It only means that these compounds may have poor bioavailability due to their poor absorption and high clearance.

The adjusted range was higher for the number of rotatable bonds of overall compounds and four classifications. Some compounds may not have satisfied Veber's Rule (≤10) and RO3 (≤3). The number of rotatable bonds indicates the measure of molecular flexibility of the drugs, which has an inverse effect on the bioavailability of a compound (Jia 2020). Compounds with higher rotatable bonds have decreased oral bioavailability (Lagorce et al 2017). Ideally, the rotatable bonds should be lower to have a higher oral bioavailability (Veber et al 2002).

Lastly, for the PSA of overall compounds and four classifications, the adjusted range was higher, and some compounds may satisfy Veber's Rule (≤140) and RO3 (≤3). If the polar surface area is higher, it may indicate a poor permeation in the cell membranes (Nas et al 2020). Compounds with a lower polar surface area may have an increased permeation rate than lipophilicity (Veber et al 2002).

Conclusions. In the 500 compounds evaluated in this study, 39.6% were studied for neurodegenerative diseases, 19.2% for cancer, 11.2% for antifungal, and 30% for

antimicrobial. The predicted ranges for the various physicochemical properties are the following: molecular weight (58.08 to 781 g mol^{-1}), MlogP (-4.47 to 6.92), HB donor (0 to 8), HB acceptor (0 to 17), WlogP (-3.54 to 8.02), number of atoms (4 to 53), rotatable bonds (0 to 12), PSA (0 to 270.86 Å). There is a significant difference in the mean values of the molecular weight, MlogP, molar refractivity, number of atoms, and PSA. We hypothesize that the strains of *C. elegans* may influence the differences in the values of the different physicochemical properties in the overall anticancer, neurodegenerative, antimicrobial, and antifungal. Consequently, the predicted ranges of the physicochemical properties exceed the existing parameters in evaluating the oral bioavailability of lead compounds in humans, namely RO5, Ghose Filter Rule, Veber's Rule, and RO3. Moreover, the higher values of physicochemical properties in *C. elegans* suggest that some compounds with high bioavailability in *C. elegans* may not be the same in humans. Conversely, the average value of the different physicochemical properties of the compounds falls within the expected ranges of most of the rules, except in RO3. Altogether, the present evaluation of the physicochemical properties of the 500 compounds may establish the ideal values for screening lead compounds in *C. elegans*. Also, the bioavailability of the compounds in the different strains of *C. elegans* may show an interesting result. Hence, further investigation is suggested.

Conflict of Interest. The authors declare that there is no conflict of interest.

References

Aldred E. M., 2008 Pharmacology: A handbook for complementary healthcare professionals. Elsevier Health Sciences, Philadelphia, 376 p.

Bhal S. K., 2007 LogP—Making sense of the value. Advanced Chemistry Development, Application Note, Toronto, 4 p.

Brenner G. M., Stevens C. W., 2018 Pharmacology. Elsevier, 552 p.

Burns A. R., Wallace I. M., Wildenhain J., Tyers M., Giaever G., Bader G. D., Nislow C., Cutler S. R., Roy P. J., 2010 A predictive model for drug bioaccumulation and bioactivity in *Caenorhabditis elegans*. Nature Chemical Biology 6(7):549-557.

Carretero M., Solis G. M., Petrascheck M., 2017 *C. elegans* as model for drug discovery. Current Topics in Medicinal Chemistry 17(18):2067-2076.

Chandrasekaran B., Abed S. N., Al-Attraqchi O., Kuche K., Tekade R. K., 2018 Computer-aided prediction of pharmacokinetic (ADMET) properties. In: Dosage Form Design Parameters, Academic Press, pp. 731-755.

Chen X., Li H., Tian L., Li Q., Luo J., Zhang Y., 2020 Analysis of the physicochemical properties of acaricides based on Lipinski's Rule of Five. Journal of Computational Biology 27(9):1397-1406.

Coimbra J. T., Feghali R., Ribeiro R. P., Ramos M. J., Fernandes P. A., 2021 The importance of intramolecular hydrogen bonds on the translocation of the small drug piracetam through a lipid bilayer. RSC Advances 11(2):899-908.

Congreve M., Carr R., Murray C., Jhoti H., 2003 A 'rule of three' for fragment-based lead discovery? Drug Discovery Today 8(19):876-877.

Desalermos A., Muhammed M., Glavis-Bloom J., Mylonakis E., 2011 Using *Caenorhabditis elegans* for antimicrobial drug discovery. Expert Opinion on Drug Discovery 6(6):645-652.

Ghose A. K., Viswanadhan V. N., Wendoloski J. J., 1999 A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. Journal of Combinatorial Chemistry 1(1):55-68.

Holden-Dye L., Walker R., 2014 Anthelmintic drugs and nematocides: studies in *Caenorhabditis elegans*. WormBook, The C. elegans Research Community, pp. 1-29.

Hou T., Wang J., Zhang W., Xu X., 2007 ADME evaluation in drug discovery. 6. Can oral bioavailability in humans be effectively predicted by simple molecular property-based rules? Journal of Chemical Information and Modeling 47(2):460-463.

Jia C. Y., Li J. Y., Hao G. F., Yang G. F., 2020 A drug-likeness toolbox facilitates ADMET study in drug discovery. Drug Discovery Today 25(1):248-258.
Kaletta T., Hengartner M. O., 2006 Finding function in novel targets: *C. elegans* as a model organism. Nature Reviews. Drug Discovery 5(5):387-399.
Lagorce D., Douguet D., Miteva M. A., Villoutreix B. O., 2017 Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. Scientific Reports 7(1):46277, 15 p.
Lipinski C. A., Lombardo F., Dominy B. W., Feeney P. J., 1997 Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews 46(1-3):3-26.
Markaki M., Tavernarakis N., 2020 *Caenorhabditis elegans* as a model system for human diseases. Current Opinion in Biotechnology 63:118-125.
Nas J., Dangeros S., Chen P., Dimapilis R., Gonzales D., Hamja F., Ramos C., Villanueva A., 2020 Evaluation of anticancer potential of *Eleusine indica* methanolic leaf extract through Ras-and Wnt-related pathways using transgenic *Caenorhabditis elegans* strains. Journal of Pharmaceutical Negative Results 11(1):42-46.
Nas J. S. B., 2020a Exploring the binding affinity and non-covalent interactions of anthocyanins with aging-related enzymes through molecular docking. Philippine Journal of Health Research and Development 24(3):9-19.
Nas J. S. B., 2020b Screening of flavonoids from *Muntingia calabura* aqueous leaf extract and its potential influence on different metabolic enzymes in *Danio rerio*. AACL Bioflux 13(5):3046-3055.
Nas J. S. B., Roxas C. K. F., Acero R. R. G., Gamit A. L. P., Kim J. P., Rentutar J. A., Ching A. C., Saludares A. Q., 2019 *Solanum melongena* (eggplant) crude anthocyanin extract and delphinidin-3-glucoside protects *Caenorhabditis elegans* against *Staphylococcus aureus* and *Klebsiella pneumoniae*. Philippine Journal of Health Research and Development 23(4):17-24.
O'Reilly L. P., Luke C. J., Perlmutter D. H., Silverman G. A., Pak S. C., 2014 *C. elegans* in high-throughput drug discovery. Advanced Drug Delivery Reviews 69-70:247-253.
Pajouhesh H., Lenz G. R., 2005 Medicinal chemical properties of successful central nervous system drugs. NeuroRx 2(4):541-553.
Pollastri M. P., 2010 Overview on the Rule of Five. Current Protocols in Pharmacology 49(1):9-12.
Pressman P., Clemens R. A., Hayes A. W., 2017 Bioavailability of micronutrients obtained from supplements and food: A survey and case study of the polyphenols. Toxicology Research and Application 1:1-7.
Rafi S. B., Hearn B. R., Vedantham P., Jacobson M. P., Renslo A. R., 2012 Predicting and improving the membrane permeability of peptidic small molecules. Journal of Medicinal Chemistry 55(7):3163-3169.
Rein M. J., Renouf M., Cruz-Hernandez C., Actis-Goretta L., Thakkar S. K., da Silva Pinto, M., 2013 Bioavailability of bioactive food compounds: a challenging journey to bioefficacy. British journal of clinical pharmacology 75(3):588-602.
Sawale R. T., Kalyankar T. M., George R., Deosarkar S. D., 2016 Molar refraction and polarizability of antiemetic drug 4-amino-5-chloro-N-(2-(diethylamino) ethyl)-2 methoxybenzamide hydrochloride monohydrate in {aqueous-sodium or lithium chloride} solutions at 30^{o}C. Journal of Applied Pharmaceutical Science 6(3):120-124.
Veber D. F., Johnson S. R., Cheng H. Y., Smith B. R., Ward K. W., Kopple K. D., 2002 Molecular properties that influence the oral bioavailability of drug candidates. Journal of Medicinal Chemistry 45(12):2615-2623.
*** http://www.swissadme.ch/

Received: 07 August 2021. Accepted: 15 September 2021. Published online: 20 November 2021.
Authors:
Regina Kamil Leal Pinlac, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018041181@feu.edu.ph
Lloyd Emer Tandoy Comia, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, Manila, 1000 Philippines, e-mail: 2018040391@feu.edu.ph
Gian Niño Trinchera Epino, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018036431@feu.edu.ph
Rogie Mariano Fernandez, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018029391@feu.edu.ph
Hannah Santiago Madrid, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018032251@feu.edu.ph
Aaron Shane Rumbaoa Salvacion, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018037071@feu.edu.ph
Cathleen Joyce Milabao Taloza, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, 1000 Manila, Philippines, e-mail: 2018039241@feu.edu.ph
John Sylvester Brusola Nas, Department of Medical Technology, Institute of Arts and Sciences, Far Eastern University, Manila, Philippines 1000, Department of Biology, College of Arts and Sciences, University of the Philippines Manila, 1000 Manila, Philippines, e-mail: jbnas@up.edu.ph

How to cite this article:
Pinlac R. K. L., Comia L. E. T., Epino G. N. T., Fernandez R. M., Madrid H. S., Salvacion A. S. R., Taloza C. J. M., Nas J. S. B., 2021 Benchmarking the physicochemical properties of 500 compounds for absorption, distribution, metabolic, excretion, and toxicity (ADMET) property prediction in *Caenorhabditis elegans*. ABAH Bioflux 13(2):58-73.