Chemistry Extended Essay

Investigation on the Effect of Molecular Polarizability on the Lipophilicity

(Indicated by Partition Coefficients) of Various Carcinogenic and

Chemotherapeutic Alkaloids.

Research Question:

"To what extent do the polarizability values' effect on the LogP values of alkaloids used as FDA-approved chemotherapeutic drugs and among the carcinogenic substances present in betel nut and other tobacco-based products be evaluated?"

`

Word count: 4041

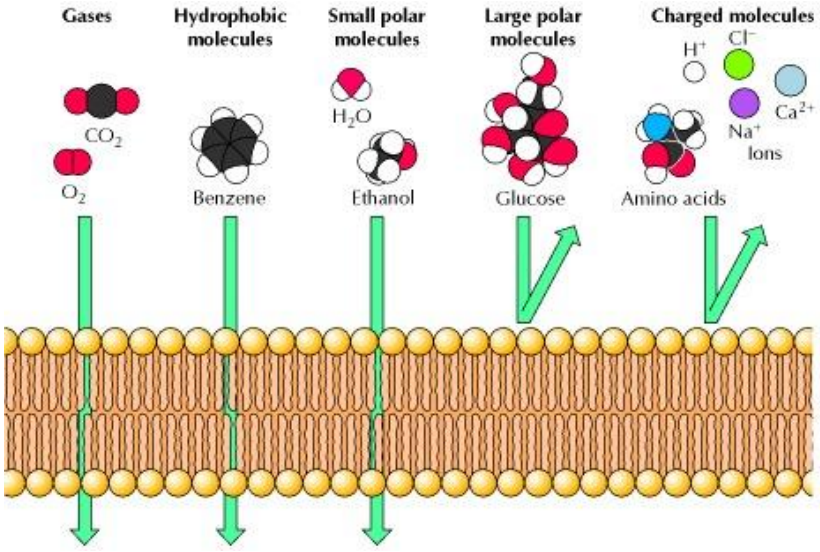
Table Of Contents

1. Introduction

The chemical structure of alkaloids (nitrogenous base compounds) used such as chemotherapeutic drugs determines their distribution across the plasma cell membrane in the human body. The agents must enter the cells through the cell membrane which is selectively permeable toward smaller polar molecules and hydrophobic molecules than large polar molecules as shown in figure 1. The drug compounds enter the bloodstream through passive diffusion across the membrane.

Cancerous cells in particular propagate by taking control of the cell cycle of healthy cells in the human body. Drugs used in chemotherapy target different aspects (phases) of the cell cycle.[1]

Figure 1: A diagrammatic representation of the plasma membrane[2]



[1] *The cell cycle and cancer | PNAS* (no date). Available at: https://www.pnas.org/doi/10.1073/pnas.94.7.2776 (Accessed Dec 11 2022).

[2] *Cell membranes - the cell - NCBI bookshelf* (no date). Available at: https://www.ncbi.nlm.nih.gov/books/NBK9928/ (Accessed: January 4, 2023).

Many factors decide the distribution of the drug compounds across the human body such as blood flow, plasma protein binding, lipid solubility, and crossing the blood-brain and placental barriers[3]. Of these factors, I chose to analyze the lipid solubility[4] of selected chemotherapeutic and cancer-causing compounds using LogP and polarizability values.

For the carcinogenic alkaloids, Areca nut compounds were analyzed. It is the seed of a popular palm tree and is chewed by over 600 million people across the world and serves as one of the leading causes of cancer[5]. Cigarettes and other most commonly used forms of similar chemical compositions are taken into consideration as well.

This study helps to classify the most widely used compounds taking into consideration the variance of lipophilicity in terms of their structural components using the partition coefficient and polarizability values associated with each compound based on the intermolecular forces of attraction, functional groups, overall stability, and strength of the molecules to predict the extent of their diffusion across the membrane.

This sets the focus of my investigation:

"To what extent do the polarizability values' effect on the LogP values of alkaloids used as FDA-approved chemotherapeutic drugs and among the carcinogenic substances present in betel nut and other tobacco-based products be evaluated?"

[3] Thomas, G. (2007). "Phases of drug action" *Fundamentals of Medicinal Chemistry*. Wiley.

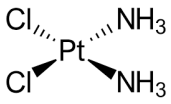
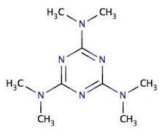
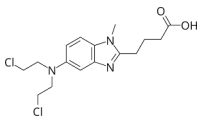
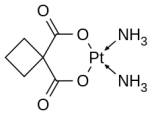
[4] Thomas, G. (2007). "3.3.1 Fundamentals of water solubility" *Fundamentals of Medicinal Chemistry*. Wiley.

[5] Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine : eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 5, 2023).

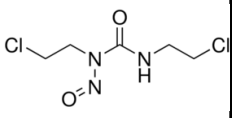
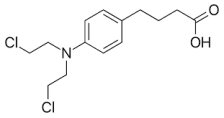
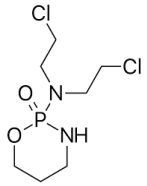
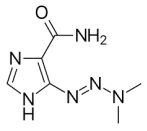
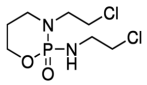
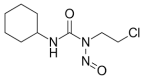
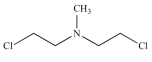
2. Background Information

2.1 Selection of chemotherapeutic and carcinogenic compounds

Table 1: List of selected chemotherapeutic compounds[6]

Drugs (FDA-approved)	IUPAC name	Molecular Formula	Molecular Structure	Molecular Mass (Mr) Mr-range: (150-800g/mol)
Cisplatin	dichloro platinum diamine	$Cl_2H_6N_2Pt$		301.10
Altretamine	N2, N2, N4, N4, N6,N6-hexamethyl-1,3,5-triazine-2,4,6-triamine	$C_9H_{18}N_6$		210.28
Bendamustine	4-{5-[bis(2-chloroethyl)amino]-1-methyl-1H-1,3-benzodiazol-2-yl} butanoic acid	C16H21Cl2N3O2		358.26
Carboplatin	7,7-diamino-6,8-dioxa-7-platinaspiro[3.5]nonane-5,9-dione	C6H12N2O4Pt		371.25

[6]Cancer, C.C. (no date) *Types of chemotherapy*, *Chemocare*. Available at: https://chemocare.com/chemotherapy/what-is-chemotherapy/types-of-chemotherapy.aspx#:~:text=Plant%20alkaloids%20are%20chemotherapy%20treatments,also%20known%20as%20antimicrotubule%20agents. (Accessed: December 20, 2022).

Carmustine	1,3-bis(2-chloroethyl)-3-nitrosourea	C5H9Cl2N3O2		214.05
Chlorambucil	4-{4-[bis(2-chloroethyl)amino]phenyl}butanoic acid	C14H19Cl2NO2		304.21
Cyclophosphamide	2-[bis(2-chloroethyl)amino]-1,3,2 lambda 5-oxazaphosphorine-2-one	C7H15Cl2N2O2P		261.09
Dacarbazine	5-(dimethyl triaz-1-en-1-yl)-1H-imidazole-4-carboxamide	C6H10N6O		182.18
Ifosfamide	3-(2-chloroethyl)-2-[(2-chloroethyl)amino]-1,3,2 lambda 5-oxazaphosphorine-2-one	C7H15Cl2N2O2P		261.10
Lomustine	3-(2-chloroethyl)-1-cyclohexyl-3-nitrosourea	C9H16ClN3O2		233.70
Mechlorethamine	bis(2-chloroethyl)(methyl)amine	C5H11Cl2N		156.05

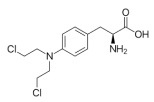
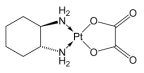
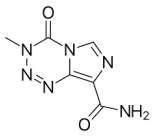
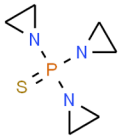
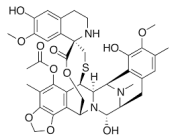
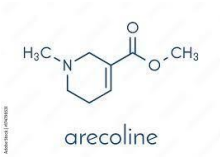
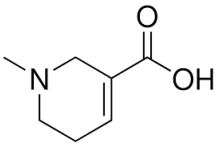
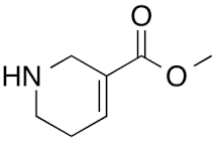
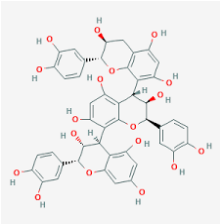
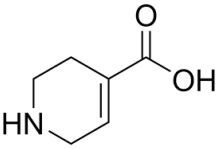
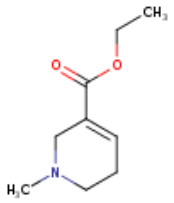
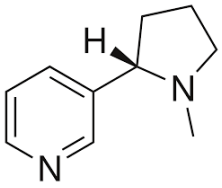
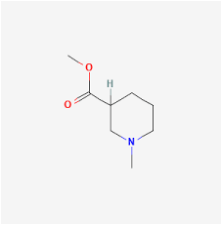
Melphalan	(2S)-2-amino-3-{4-[bis(2-chloroethyl)amino]phenyl}propanoic acid	C13H18Cl2N2O2		305.20
Oxaliplatin	(3aR,7aR)-octahydro-2',5'-dioxaspiro[cyclohexa[d]1,3-diaza-2-platinacyclopentane-2,1'-cyclopentane]-3',4'-dione	C8H12N2O4Pt		397.30
Temozolomide	3-methyl-4-oxo-3H,4H-imidazo[4,3-d][1,2,3,5]tetrazine-8-carboxamide	C6H6N6O2		194.15
Thiotepa	tris(aziridin-1-yl)-λ⁵-phosphanethione	C6H12N3PS		189.23
Trabectedin	(1R,2R,3R,11S,12S,14R,26R)-5,6',12-trihydroxy-6,7'-dimethoxy-7,21,30-trimethyl-27-oxo-3',4'-dihydro-2'H-17,19,28-trioxa-24-thia-13,30-diazaspiro[heptacyclo[12.9.6.1³,¹¹.0²,¹³.0⁴,⁹.0¹⁵,²³.0¹⁶,²⁰]triacontane-26,1'-isoquinoline]-4,6,8,15,20,22-hexaen-22-yl acetate	C39H43N3O11S		761.84

Table 2: List of Areca nut alkaloids and nicotine product alkaloids[7]

Compounds	Iupac name	Molecular Formula	Molecular Structure	Molecular Mass (Mr) Mr-range: (120-900 g/mol)
Arecoline	methyl 1-methyl-1,2,5,6-tetrahydropyridine-3-carboxylate	C8H13NO2		155.20
Arecaidine	1-methyl-1,2,5,6-tetrahydropyridine-3-carboxylic acid	C7H11NO2		141.17
Guvacoline	Methyl 1,2,5,6-tetrahydropyridine-3-carboxylate	C7H11NO2		141.17
arecatannin A1	2-(3,4-dihydroxyphenyl)-8-[2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2H-1-benzopyran-4-yl]-4-[2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2H-1-benzopyran-8-yl]-3,4-dihydro-2H-1-benzopyran-3,5,7-triol	C45H38O18		866.77
isoguvacine	1,2,3,6-tetrahydropyridine-4-carboxylic acid	C6H9NO2		127.14

[7] Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine: eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 8, 2023).

homoarecoline	ethyl 1-methyl-1,2,5,6-tetrahydropyridine-3-carboxylate	C9H15NO2		169.22
nicotine imine	1-methyl-2-(pyridin-3-yl)-3,4-dihydro-2H-pyrrol-1-ium	$C_{10}H_{14}N_2$		162.23
Methyl N-methylpiperidine-3-carboxylate	-	C8H15NO2		157.21

2.2 Polarizability and Di-pole moments[8]

An electric di-pole has two partial (+/-) charges (Q) separated by a certain distance 'L.' The product of (QL) is known as the electric di-pole moment (μ)[9]. Larger molecules have larger di-pole moments due to their separation. A polar molecule has a permanent di-pole with partial charges as opposed to the temporary induced dipoles in nonpolar molecules.

μ/D = ΔX (change in electronegativities of the atoms)

[8]Junmei Wang, Xiang-Qun Xie, Tingjun Hou, and Xiaojie Xu (2007) *Fast Approaches for Molecular Polarizability Calculations*, *pubs.acs.org*. ACS publications. Available at: https://pubs.acs.org/doi/10.1021/jp068423w (Accessed: January 6, 2023).

[9] Dipole-moment representation: An arrow from the partial negative to the partial positive charge and with units Cm or one Debye (D).

Temporary induced di-pole moments (μ^*) present in non-polar molecules are influenced by the presence of an external electric field (E)[10].

$\mu^* \propto$ E which can be represented with the constant of proportionality [11]:

$\mu^* = \boldsymbol{(\alpha)}$ E

The greater the magnitude of **(α),** the greater the distortion in the molecule's electron cloud due to the external electric field. In molecules with fewer electrons, there is a stronger influence of nuclear charges which results in a low polarizability value. Larger atoms have a relatively smaller influence on their outer electrons due to their nuclei hence their outer electrons are more susceptible to being displaced more. The units for Polarizability are measured in cubic angstroms: $1\ \text{Å}^3 = 10^{-24}\ cm^3$.

2.3 Solubility of Compounds and Molecular polarity

The solubility of compounds determines the degree of dissolving in the human bloodstream (bioavailability) through the plasma (90% water) membrane[12]. The more polar a given drug compound is, the fewer doses of it are needed to achieve the targeted pharmacological results[13]. The intermolecular forces between the solutes and the solvents within themselves must be overcome in order for the solutes to completely dissolve with the solvents.

---

[10] *University of Wisconsin–Eau Claire* (no date). Available at: https://www.chem.uwec.edu/Chem406_F06/Pages/lecture_notes/lect05/Atkins-Ch16_small.pdf (Accessed: January 22, 2023).

[11] (α): Constant of proportionality (denotes polarizability)

[12] *Physiology, blood plasma - statpearls - NCBI bookshelf* (no date). Available at: https://www.ncbi.nlm.nih.gov/books/NBK531504/ (Accessed: January 24, 2023).

[13] Savjani, K.T., Gajjar, A.K. and Savjani, J.K. (2012) *Drug solubility: Importance and enhancement techniques*, *ISRN pharmaceutics*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3399483/ (Accessed: January 16, 2023).

The more soluble a given compound is in water as a polar solvent, the more polar the given compounds are.

Table 3: Solubilities of the chemotherapeutic compounds in water at 25 C

Compounds	Predicted Solubility (mg/ml) from ALOGPS[14]
Cisplatin	69.6 mg/mL
Altretamine	3.1 mg/mL
Bendamustine	0.0618 mg/mL
Carboplatin	79.8 mg/mL
Carmustine	1.53 mg/mL
Chlorambucil	0.0773 mg/mL

[14]*Cisplatin* (no date) *Uses, Interactions, Mechanism of Action | DrugBank Online*. Available at: https://go.drugbank.com (Accessed: January 15, 2023).

Cyclophosphamide	15.1 mg/mL
Dacarbazine	1.36 mg/mL
Ifosfamide	15.0 mg/mL
Lomustine	0.755 mg/mL
Mechlorethamine	33.4 mg/mL
Melphalan	-
Oxaliplatin	27.5 mg/mL
Temozolomide	-

Thiotepa	9.27 mg/mL
Trabectedin	0.328 mg/mL

Table 4: Solubilities of the carcinogenic compounds in water at 25 C

Compounds	Predicted Solubility (g/L) from ALOGPS[15]
Arecoline	446 g/L
Arecaidine	510 g/L
Guvacoline	-
Arecatannin A1	-
Isoguvacine	-
Homoarecoline	297 g/L
Nicotine Imine	0.21 g/L
Methyl N-methyl piperidine-3-carboxylate	-

[15] *Browsing metabolites* (no date) *Human Metabolome Database: Browsing metabolites*. Available at: https://hmdb.ca/metabolites (Accessed: January 25, 2023).

The overall polarity of molecules determines the net dipole moments in the compounds. Symmetric compounds with even charge distributions primarily exhibit nonpolar characters with no net dipole moment values while more asymmetric charge distributions within a structure have net dipole moments[16]. The compounds should also at least have one polar covalent bond in order for them to exhibit some degree of polarity.

2.4 Partition coefficient (LogP) values[17]

The lipophilicity is a drug's tendency to be facilitated through the membrane and is given by the partition coefficient which is the ratio of the concentration of a given drug in Octan-1-ol and of that in water in a system that is in equilibrium as given in equation (1). The movement of the drug molecules from the hydrophilic side of the cell membrane to the hydrophobic side is known as partitioning. The partition coefficient (P) depends on the equilibrium concentrations, temperature, and polarity of solvents affecting the amount of solute in each medium.

The following equation gives the partition coefficient (P):

$$logP = \frac{[Drug]\ in\ octan\ 1\ ol}{[Drug]\ in\ water}$$

[16]Baker, M. *et al.* (no date) *7.3 molecular polarity and dipole moments*, *Chemistry Fundamentals*. UCF Pressbooks. Available at: https://pressbooks.online.ucf.edu/chemistryfundamentals/chapter/7-3-molecular-polarity-and-dipole-moments/ (Accessed: January 28, 2023).

[17] *Partition coefficient* (no date) *Partition Coefficient - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/medicine-and-dentistry/partition-coefficient#:~:text=The%20partition%20coefficient%20of%20a,and%20the%20value%20is%20calculated. (Accessed: January 7, 2023).

The partition coefficient[18] is a factor of the lipophilicity of set compounds. Empirically, its value is determined using the shake flask method[19] by shaking the drug compound with equal parts of two immiscible solvents until equilibrium is attained. Octanol-water partitioning system is the most commonly used system to determine the values of LogP. The more polar compounds will yield negative LogP values in favor of a greater concentration of them in water as opposed to certain compounds yielding positive LogP values the more non-polar towards the concentration of octanol. Ideally, the compounds targeting the oral, intestinal, and other major organ membranes of the body should have LogP values between 1.5-1.8, and the drugs targeting the CNS[20] and blood-brain barriers should have LogP values around 2[21].

Other empirical methods of finding LogP values include the HPLC (High-performance liquid chromatography). This is a much faster and an efficient alternative to the shake flask method[22].

---

[18] *Partition coefficient* (no date) *Partition Coefficient - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/medicine-and-dentistry/partition-coefficient#:~:text=The%20partition%20coefficient%20of%20a,and%20the%20value%20is%20calculated. (Accessed: January 7, 2023).

[19] *LogP—making sense of the value - acdlabs.com* (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/making_sense.pdf (Accessed: January 17, 2023).

[20] CNS: Central Nervous System

[21] *LogP—making sense of the value - acdlabs.com* (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/making_sense.pdf (Accessed: January 16, 2023).

[22] *LogP—making sense of the value - acdlabs.com* (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/making_sense.pdf (Accessed: January 16, 2023).

2.5 Classifying the chemotherapeutic drugs[23]

Table 5: Classification of Chemotherapeutic compounds

Classification	Compound(s)	Uses
The platinum compounds (metal salts)	Cisplatin, Carboplatin, and Oxaliplatin	are used to treat lung, bladder, and testicular cancers
The Nitrogenous Mustards	Cyclophosphamide, ifosfamide, melphalan, chlorambucil, and bendamustine	used to treat blood and other lymph system cancers
Azirdine derivative	Thiotepa	is used to treat fluid effusion of cancer through tissues.
Nitrosoureas	Carmustine and Lomustine	are used to treat brain tumors
Triazenes, Hydrazines, and related compounds	Dacarbazine, Temozolomide, and Altretamine	are used to treat melanoma and skin cancers.
Extremely large compound	Trabectedin	is used to treat soft tissue cancers (sarcomas).

[23] Holland, J.F. and Colvin, M. (2003) "Alkylating Agents," in *Cancer Medicine* . BC Decker Inc.

2.6 Classifying the carcinogenic compounds[24]

Table 6: Classification of Carcinogenic compounds

Classification	Compound(s)
Most Toxic compound in Betel Nut	Arecoline
Pyridine type (nitrogen atoms in existing benzene rings) compounds	Nicotine imine, Guvacoline, Isoguvacine, Arecaidine, and Homoarecoline
Extremely large compounds and other exceptions:	Methyl-N-methyl piperidine-3-carboxylate and Arecatannin A1

3. Hypothesis

With an increase in LogP values (non-polar compounds), the polarizability values of the molecules should increase due to the increased number of London dispersion forces present in the molecules[25]. Larger molecules should have more intermolecular forces of attraction which should displace electrons relative to each other to a greater extent if the compounds have greater molecular masses.

On the other hand for more polar compounds, their LogP values should be negative so with an increase in polarity, the LogP values should decrease. With a decrease in the LogP values, there are not as many London dispersion forces of attraction as they are replaced mostly

[24] Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine: eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 8, 2023).

[25] London Dispersion Forces: Temporary induced dipoles are mostly present in non-polar characteristics of a compound.

by di-pole di-pole forces based on the overall polarities of the bonds. This should decrease the overall polarizability of the molecule with a decrease in LogP values.

4. Table of variables

Table 7: *Table of variables*

Variable	Type of variable	Effect of variable
Polarizability of the different agents	Independent	-
Class of compound	Independent	-
LogP values	Dependent	-
Solvents for partitioning	Controlled	The polarity of the solvent (water) and the non-polar organic compound (Octanol) are the two primary parameters used to determine the LogP values. The concentrations of the select compounds in either of the solvents in the immiscible mixture depend on the interaction of the compound

		with each medium
The concentration of the solvents	Uncontrolled	The concentration of either of the solvents present in the immiscible mixture can affect the logP values and the equilibrium's position.
Temperature	Controlled	As the temperature affects the equilibrium concentrations of the compounds, a constant parameter for temperature must be maintained.

5. Approach to the Research Question

5.1 Selection of databases

Tabe 8: *List of secondary source software/databases used*

	Name of source	Type of source	Data collected from source
D1	ChemAxon (DrugBank) As a part of the	Computational software	Predicted LogP values and Polarizability values

	HMDB database[26]		of both compounds
D2	Chemicalize and ACD chem labs (Chem spider)[27]	Computational software	Predicted polarizability values and logP values of both compounds
D3	FDA drugs[28]	Database software	List of chemotherapeutic alkaloids legally approved and available for treatment
D4	ALOGPS (part of HMDB database)	Computational software	Predicted logP values for carcinogenic compounds.
D5	HMDB (ChemAxon)	Computational software	Empirical logP values of chemotherapeutic compounds

[26]*Human metabolome database* (no date) *Human Metabolome Database*. Available at: https://hmdb.ca/ (Accessed: January 9, 2023).

[27] *Search ChemSpider* (no date) *ChemSpider*. Available at: http://www.chemspider.com/Default.aspx (Accessed: January 9, 2023).

[28]Center for Drug Evaluation and Research (no date) *Drug approvals and databases*, *U.S. Food and Drug Administration*. FDA. Available at: https://www.fda.gov/drugs/development-approval-process-drugs/drug-approvals-and-databases (Accessed: January 9, 2023).

HMDB is a reputed database source from "The Metabolomics Innovation Center" (TMIC)[29]. TMIC uses a collection of its databases which include HMDB (Human Metabolome Database) contains detailed information along with empirical and predicted values from reputable and widely used software like ChemAxon and the others mentioned above. It is used for application metabolomics, clinical chemistry, biomarker discovery, and general education.

5.2 Methodology

The predicted logP and polarizability values for the chemotherapeutic drugs were compiled using D1 and D2. For the carcinogenic compounds, D5 was also used to get predicted LogP values.

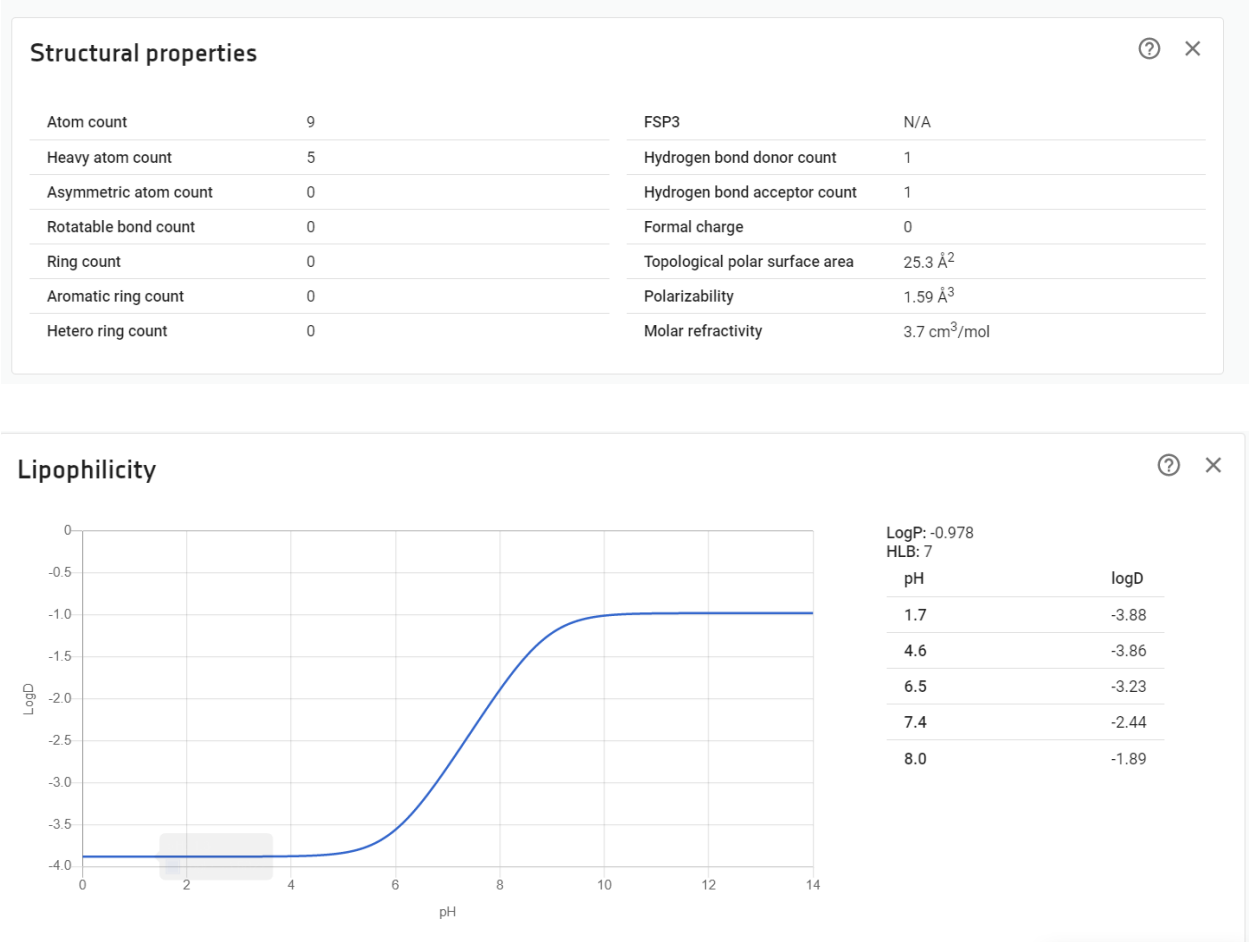
1) The predicted and empirical data sets were obtained from the drop-down option for the different variables present in the properties tab for each of the chosen compounds.

Figure 2: HMDB dropdown menu for predicted properties

Property	Value	Source
logP	2.15	ALOGPS
logP	2.49	ChemAxon
logS	-4.3	ALOGPS
pKa (Strongest Acidic)	12.21	ChemAxon
pKa (Strongest Basic)	-3	ChemAxon
Physiological Charge	0	ChemAxon
Hydrogen Acceptor Count	9	ChemAxon
Hydrogen Donor Count	5	ChemAxon
Polar Surface Area	138.07 Å²	ChemAxon
Rotatable Bond Count	3	ChemAxon
Refractivity	153.19 m³·mol⁻¹	ChemAxon
Polarizability	67.51 Å³	ChemAxon

[29] *Databases* (2022) *The Metabolomics Innovation Centre*. Available at: https://metabolomicscentre.ca/software-databases/databases/ (Accessed: January 18, 2023).

Figure 3: Chemspider menu for predicted polarizability and LogP properties



Structural properties

Atom count	9	FSP3	N/A
Heavy atom count	5	Hydrogen bond donor count	1
Asymmetric atom count	0	Hydrogen bond acceptor count	1
Rotatable bond count	0	Formal charge	0
Ring count	0	Topological polar surface area	25.3 $Å^2$
Aromatic ring count	0	Polarizability	1.59 $Å^3$
Hetero ring count	0	Molar refractivity	3.7 cm^3/mol

Lipophilicity

LogP: -0.978
HLB: 7

pH	logD
1.7	-3.88
4.6	-3.86
6.5	-3.23
7.4	-2.44
8.0	-1.89

2) With the data at hand, the classification of chemotherapeutic agents was used to get a set average with absolute uncertainties ((max-min)/ 2) for the LogP and polarizability values. The carcinogenic compounds already had a minimal data set chosen and were mostly classified under the category of betel nut alkaloids.
3) Empirical logP values were then collected for both types of compounds.
4) The predicted polarizability vs predicted and empirical LogP graphs were plotted on excel for the coverage of the chemotherapeutic compounds as well as the carcinogenic agents to analyze the trend and contrast the predicted model.

6. Raw Data Collection

6.1 Chemotherapeutic Alkaloids

6.1.1 Predicted Polarizability values and Predicted LogP values

Tabe 9: *Raw and Processed Data for all the chemotherapeutic compounds*

Compound name	Polarizability / Å^3 D1	Polarizability / Å^3 D2	LogP D1	LogP D2	Mean predicted Polarizability/ $Å^3$	Absolute uncertainty predicted polarizability	Mean logP	Absolute uncertainty predicted logP
Cisplatin	10.31	16.65	0.041	-	13.48	3.17	0.041	-
Altretamine	23.7	25.2	2.22	2.42	24.45	1.5	2.32	0.1
Bendamustine	38.19	36.8	1.66	2.69	37.5	1.39	2.175	0.5
Carboplatin	18.27	-	0.14	-	18.27	-	0.14	-
Carmustine	18.8	18.5	1.02	1.12	18.65	0.15	1.07	0.05
Chlorambucil	31.98	31.7	3.94	3.10	31.84	0.14	3.52	0.42

Cyclo- phosphamide	23.72	23.0	0.097	0.23	23.36	0.36	0.1635	0.067
Dacarbazine	17.78	18.3	-0.43	-0.28	18.04	0.26	-0.355	0.075
Ifosfamide	23.94	23.0	0.097	0.23	23.47	0.47	0.1635	0.133
Lomustine	23.61	22.9	2.16	2.76	23.255	0.71	2.46	0.3
Mechlorethamine	15.84	15.3	1.52	0.91	15.57	0.27	1.215	0.31
Melphalan	31.38	31.2	0.25	1.79	31.29	0.09	1.02	0.77
Oxaliplatin	21.29	-	1.73	-	21.29	-	1.73	-
Temozolomide	16.88	18.1	-0.28	-	17.49	0.61	-0.28	-
Thiotepa	18.23	19.5	-1	0.52	18.87	0.635	-0.24	0.76
Trabectedin	77.72	77.6	3.99	3.10	77.66	0.06	3.545	0.445

6.2 Carcinogenic Alkaloids

6.2.1 Predicted Polarizability and Predicted LogP values

Table 10: *Raw and Processed Data for all the carcinogenic compounds*

Compound name	Polarizability / Å^3 D1	Polarizability / Å^3 D2	logP D1	logP D2	logP D4	Mean predicted Polarizability/ $Å^3$	Absolute uncertainty in predicted polarizability	Mean predicted logP	Absolute uncertainty in predicted logP
Arecoline	17.1	16.6	0.65	1.28	0.55	16.85	0.25	0.83	0.37
Arecaidine	14.89	14.7	-2.5	0.77	-1.4	14.80	0.095	-1.04	1.64
Guvacoline	-	14.7	-	1.24	-	14.7	-	1.24	-
arecatannin A1	86.26	-	4.44	-	3.38	86.26	-	3.91	0.53
isoguvacine	12.74	12.7	-2.6	0.68	-2.3	12.72	0.02	-1.4	1.64
homoarecoline	19.19	18.4	1.01	1.81	0.86	18.8	0.395	1.23	0.475
Nicotine imine	18.14	-	-2.9	-	-2	18.14	-	-2.45	0.9
Methyl N-methylpiperidine-3-carboxylate	17.69	-	0.57	-	0.55	17.69	-	0.56	0.01

7. Discussion of the results

7.1 An analysis of the chemotherapeutic agents

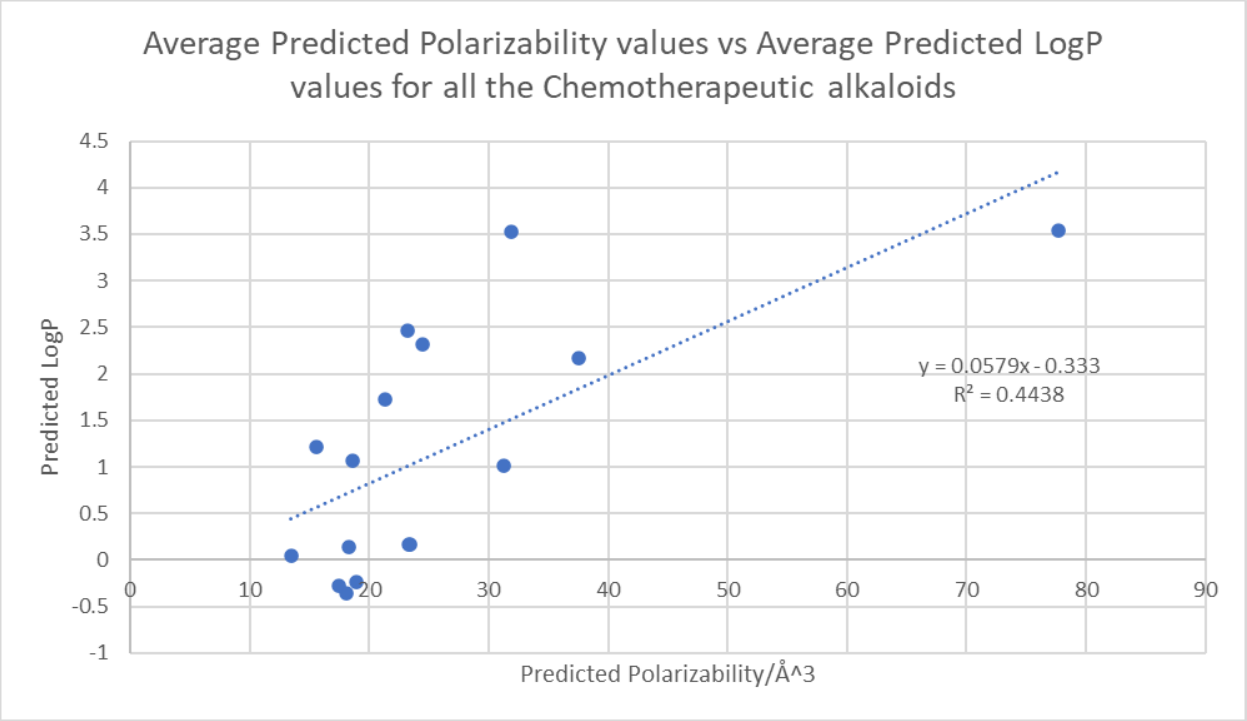
7.1.1 Average Predicted LogP vs Average Predicted Polarizability values

Table 11: *Mean Predicted Polarizability and LogP for all the chemotherapeutic compounds*

Compound name	Mean predicted Polarizability/ $Å^3$	Mean logP
Cisplatin	13.48	0.041
Altretamine	24.45	2.32
Bendamustine	37.5	2.175
Carboplatin	18.27	0.14
Carmustine	18.65	1.07
Chlorambucil	31.84	3.52
Cyclo- phosphamide	23.36	0.1635

Dacarbazine	18.04	-0.355
Ifosfamide	23.47	0.1635
Lomustine	23.255	2.46
Mechlorethamine	15.57	1.215
Melphalan	31.29	1.02
Oxaliplatin	21.29	1.73
Temozolomide	17.49	-0.28
Thiotepa	18.87	-0.24
Trabectedin	77.66	3.545

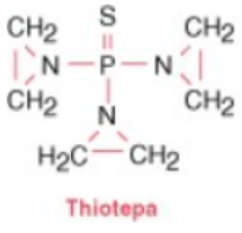
Figure 4: Graph of predicted logP values vs predicted polarizability values



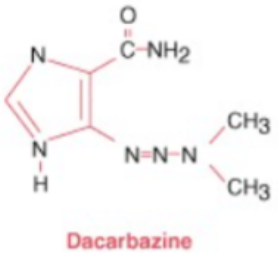
In the graph, there is a positive correlation between the predicted polarizability and the predicted LogP values. With an R^2 value of approximately 0.44, it's closer to a moderate correlation as the value is less than 0.5. The average predicted LogP values increase (slope is 0.0579) with polarizability values.

Aziridines and Epoxides:

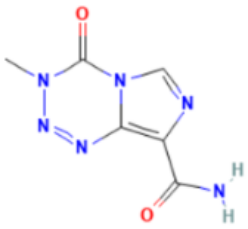
Thiotepa (Azirdine rings)



Dacarbazine



Temozolomide



The negative log P values of these compounds (exceptions to the trend) could be due to the electronegativity difference among certain bonds being relatively high, much like the P = S in

Thiotepa is polar (Log value of -0.24) due to the two lone pairs of electrons on Sulfur which disrupt the overall symmetrical structure of the molecule. The nitrogen atom in thiotepa's aziridine ring[30] gets protonated and opens up.[31] As Thiotepa branches out[32] upon interaction with aqueous media and takes part in reactions readily, its mechanism adds to its overall polar nature as a compound. This is further evidenced by the solubility of Thiotepa in water which is 9.27 mg/mL and exhibits more polar characteristics which leads to it having a lesser polarizability value due to fewer London dispersion forces of attraction.

As for Dacarbazine and Temozolomide, perhaps hydrogen bonding in the N-H bond will allow the compound to make polar bonds with the partially negative oxygen atom from the water molecule. The Double bonds across the pyridine type rings in either of the compounds could also add to the polarity of the compounds due to the delocalization of the pi-bonds on top of the existing sigma bonds. Other polar bonds such as C = O[33] could make their overall structures more polar. Dacarbazine has a LogP value of -0.355 similar to that of Temozolomide with -0.28.

---

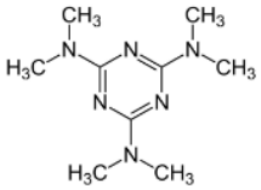
[30] Aziridine ring: Nitrogen centered ring with three other functional groups

[31] *Thiotepa* (no date) *Thiotepa - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/chemistry/thiotepa#:~:text=Mechanism%20of%20Toxicity&text=Thiotepa%20(and%20TEPA)%20form%20DNA,disruption%20of%20nucleic%20acid%20function. (Accessed: January 29, 2023).

[32] *Thiotepa* (no date) *Thiotepa - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/chemistry/thiotepa#:~:text=Mechanism%20of%20Toxicity&text=Thiotepa%20(and%20TEPA)%20form%20DNA,disruption%20of%20nucleic%20acid%20function. (Accessed: January 29, 2023).

[33] High electronegativity difference (closer to 1.8 for it to be polar)

Altretamine

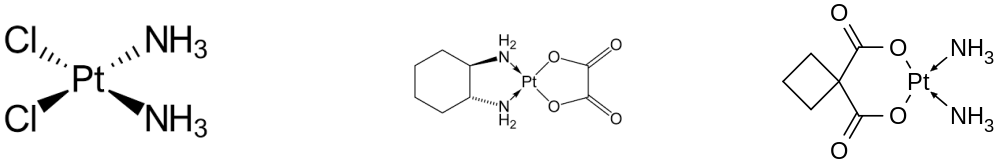


The above structure for altretamine (acts as the exception in the Hydrazine and Triazine group) has no polar bonds and has a pyridine-type structure. Perhaps the methyl groups attached to each of the nitrogen atoms surrounding the ring provide some sort of steric hindrance to keep the compound from getting protonated by the water molecule across the aqueous medium of the membrane. This allows for a plausible insight into the non-polar empirical value for LogP (2.32). Altretamine is barely soluble in water (3.1 mg/mL).

The platinum compounds have the lowest polarizability values compared to the other compounds. Their range of LogP values is from 0.041 to 1.73 with Cisplatin being the lowest. Their low polarizability values (13.48 - 21.29) could be due to the lack of temporary dipoles due to the dative coordinate bonding of ligands as lewis bases to the central transition metal (lewis acid).

The platinum compounds (metal salts):

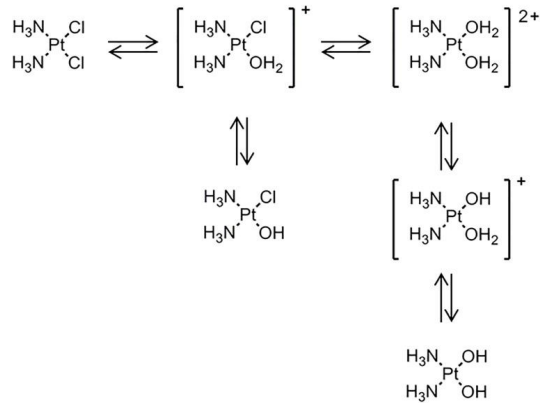
Cisplatin Oxaliplatin Carboplatin



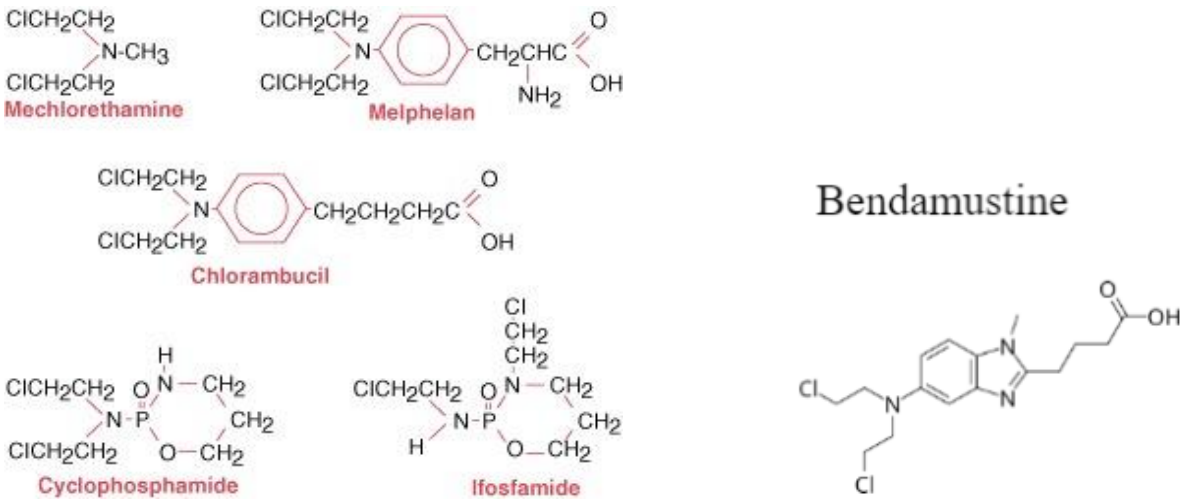
The platinum compounds have the tendency to gain a water molecule as a ligand when they diffuse through the membrane into aqueous media (Cisplatin's solubility of 69.6 mg/mL).. This mechanism is due to the dative bonds between chlorine and the central platinum atom. They

become ionized as the molecule tries to reach equilibrium with the aqueous media across the polar head of the membrane.[34]

Figure 4: Ionization of Cisplatin



The nitrogenous mustards:



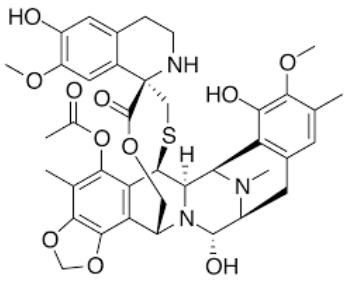
Cyclophosphamide is the most used nitrogenous mustard.[35] With its moderate solubility (15.1 mg/mL) in water, it stands out as a relatively polar molecule along with Mechlorethamine

[34] Martinho, N. (2019) *Cisplatin-Membrane Interactions and Their Influence on Platinum Complexes Activity and Toxicity*, *National Center for Biotechnology Information (NCBI)*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6336831/ (Accessed: January 8, 2023).

[35] Holland, J.F. and Colvin, M. (2003) "Alkylating Agents," in *Cancer Medicine* . BC Decker Inc.

(33.4 mg/mL). Both compounds possess asymmetric structures that allow for the distribution of a net dipole moment to be present within these molecules making their overall natures more polar. Cyclophosphamide does have the lowest LogP value in the group of 0.1635 which could primarily be due to the N-H bond that is present in its structure. This bond can form di-pole di-pole interactions with water molecules making the overall molecule more polar as opposed to the other compounds in the group. It is important to also note that Ifosfamide (structural isomer) also has an identical LogP and solubility values of 0.1635 and 15 mg/mL. Of the compounds in this group, Bendmaustine and Chlorambucil have extremely low solubility values of 0.06 and 0.07 mg/mL respectively. Despite having an OH group, their long chain of other functional groups and the benzene ring in Chlorambucil adds to their nonpolar characteristic[36] and increases their polarizability values.

Trabectedin



Tribactedin being the largest compound has the most number of bonds (more induced dipoles). With a molecular mass of 761.84g, the polarizability values increase and the dispersion forces become stronger due to its molecular mass[37]. The covalent structure of Trabectedin attracts nearby compounds more readily. More outer electrons are more easily displaced (higher

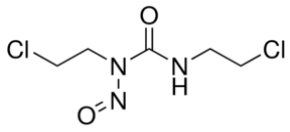
[36]More induced temporary di-poles.

[37] Petrucci, Ralph H., et al. General Chemistry: Principles and Modern Applications.Upper Saddle River, NJ: Prentice Hall, 2007

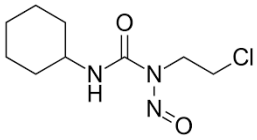
Chang, Raymond. "Chapter 13: Intermolecular Forces/ Ion-Induced Dipole and Dipole-Induced Dipole Interactions/ Dispersion, or London, Interactions." Physical Chemistry for the Biosciences. Sansalito, CA: University Science, 2005. 495-98. Print.

polarizability values). It has the highest LogP value of 3.5 and a polarizability value of 77.66 A^3. As Trbactedin has such a large structure, it has more temporary dipoles that are induced due to a much larger number of London dispersion forces of attraction that are present within the molecule. This negates the effects of certain hydrogen bonds present in the compound making its overall nature to be more non-polar (Solubility value of 0.328 mg/mL).

Carmustine



Lomustine



The nitrosoureas compounds[38] Carmustine and Lomustine have LogP values of 1.07 and 2.46 respectively. They have extremely low solubility values of 0.7 to 1.5 mg/mL which naturally leads to them having higher polarizability values of 18.65 to 23.255 A^3. These compounds despite having N-H bond and a polar double bond, their net dipoles could potentially negate vectorially due to the overlap of overall electronegativity differences between the (N-N=O) and (N-H) bonds which is evident in Carmustine. Lomstine has a more non-polar characteristic with a higher LogP[39] value. This could be due to the symmetrical hexane ring bonded due to the Nitrogen which reduces the overall electronegativity difference as opposed to the N-Cl bond in the same position in Carmustine.

[38] CNS barrier compounds administered for brain tumors.

[39] Being able to dissolve more in a concentration of Octan-1-ol than in water

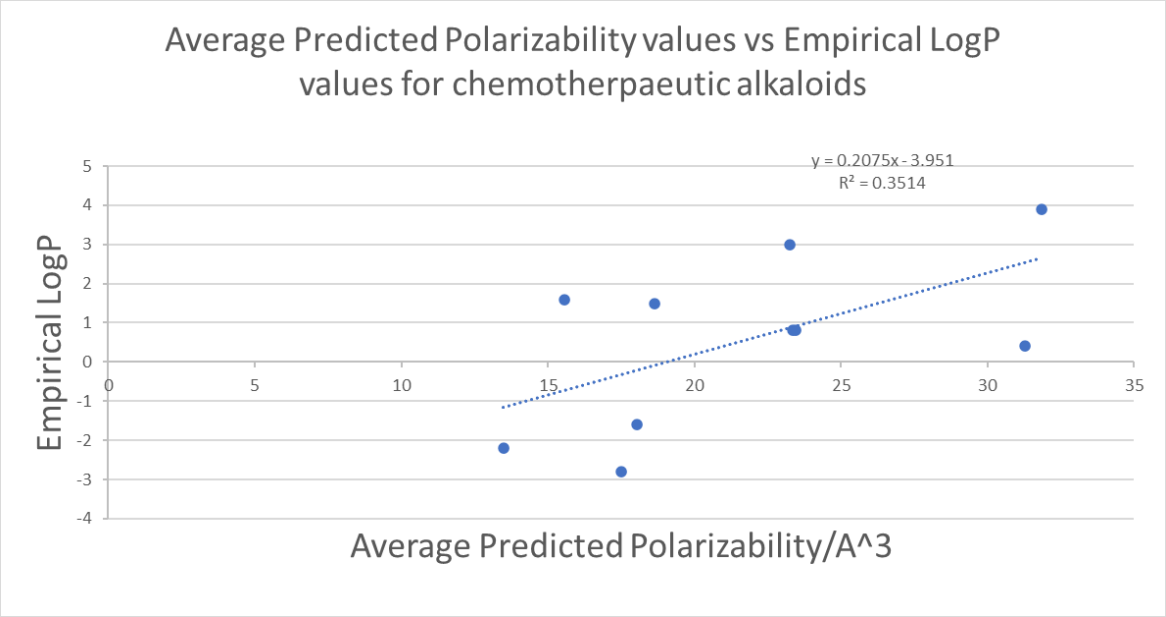
7.1.2 Empirical LogP values vs Predicted Polarizability values

Table 12: *Predicted Polarizability and empirical LogP for all the chemotherapeutic compounds*

Compound	Mean predicted Polarizability/ $Å^3$	Empirical LogP
Cisplatin	13.48	-2.19
Altretamine	24.45	1.7
Bendamustine	37.5	-
Carboplatin	18.27	-
Carmustine	18.65	1.5
Chlorambucil	31.84	3.9
Cyclo- phosphamide	23.36	0.8
Dacarbazine	18.04	-1.6

Ifosfamide	23.47	0.8
Lomustine	23.255	3
Mechlorethamine	15.57	1.6
Melphalan	31.29	0.4
Oxaliplatin	21.29	-
Temozolomide	17.49	-2.8
Thiotepa	18.87	-
Trabectedin	77.66	-

Figure 5: Graph with empirical LogP values



In the graph, there is a major variation in the data as the R^2 value of 0.3514 is relatively weak being much less than 0.5. Cisplatin has an empirical LogP value of -2.19 and a predicted value of 0.041 while Dacarbazine and Temozolomide also have significantly higher negative LogP values.. There is a large margin of difference between these two values due to the limited range of LogP values present in my data set. As my data set values only have a range from -0.34 to 3.5 (range of 3.8), the difference in Cisplatin's LogP values of 2.231 is significant much like a difference of 1.245 and 2.52 from (-0.355 to -1.6) and (-0.28 to -2.8) for Dacarbazine and Temozolomide respectively. The other compounds varied with minor differences of a maximum of 0.6 between their predicted and empirical LogP values. There is an overall shift in the graph for the empirical values to the right with lower empirical LogP values for the exceptions.

7.2 An analysis of the carcinogenic agents[40]

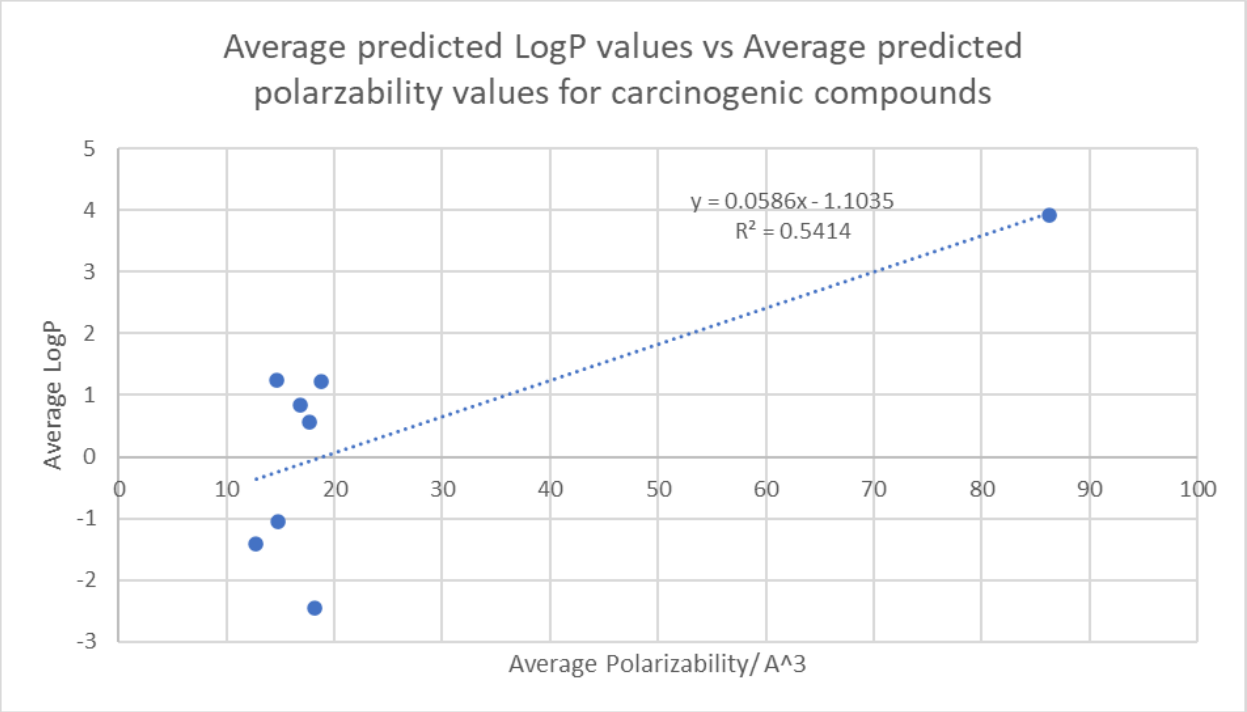
7.2.1 Predicted LogP vs Predicted Polarizability values

Table 13: *Mean Predicted Polarizability and LogP for all the carcinogenic compounds*

Compound	Mean predicted Polarizability/ $Å^3$	Mean Predicted LogP
Arecoline	16.85	0.83
Arecaidine	14.80	-1.04
Guvacoline	14.7	1.24
Arectannin A1	86.26	3.91
Isoguvacine	12.72	-1.4
Homoarecoline	18.8	1.23
Nicotine Imine	18.14	-2.45
Methyl N-methyl piperidine-3-carboxylate	17.69	0.56

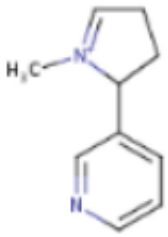
[40] Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine : eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 7, 2023).

Figure 6: Graph of predicted logP values vs Predicted polarizability values



The mean logP values for the carcinogenic compounds mostly increases with the mean polarizability values. The R^2 value is about 0.541 > 5, hence representing a moderate correlation among my data sets with a slope value of 0.0586. There are a few exceptions with negative LogP values like Nicotine Imine, Isoguvacine, and Arecaidine with values of -2.45, -1.4, and -1.04 respectively.

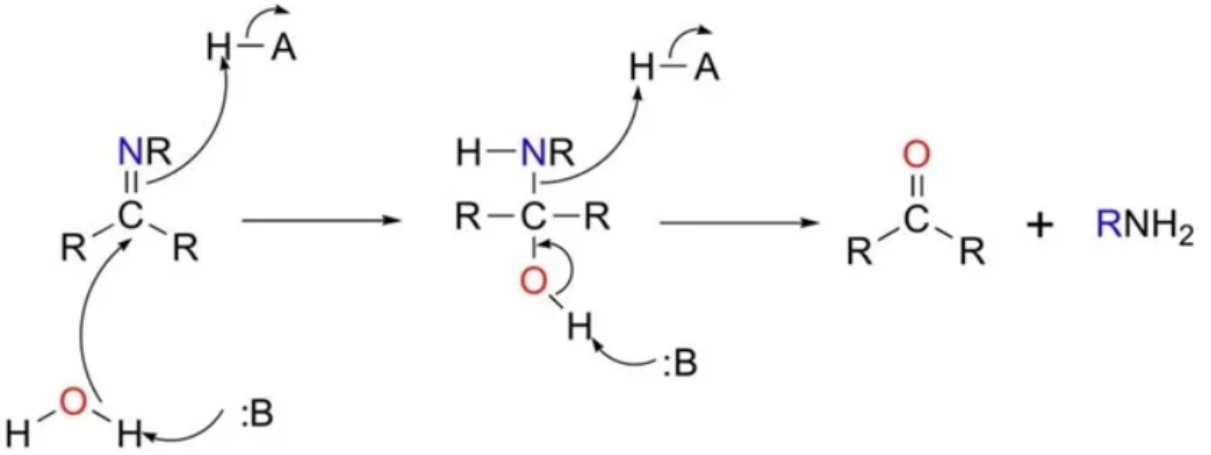
Nicotine imine



Its imine group is highly polar. It has a solubility value of 0.21 mg/mL. Despite its low

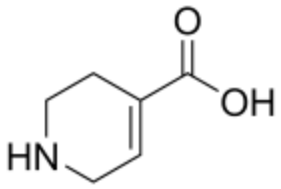
solubility value, its pi bond can still be attacked by electrophile (H+) ions from the water molecule. Due to its extent of dissociation in water, it has a negative LogP value for its more soluble in a mixture of water than a non–polar solvent with which it has no electrophilic reaction mechanism to be protonated. The nitrogen of C=N can be protonated (accepts H+ ions) easily. The electrons are separated from the attached molecule with ease. “Bond cleavage” is facilitated in this process across the aqueous medium of the membrane[41]. Nicotine imine has the lowest LogP value of -2.45 compared to all the other carcinogenic alkaloids.

Figure 6: Protonation of Imines

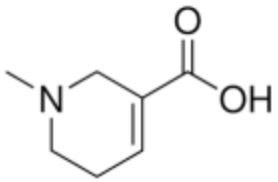


[41] Aliouche, H. (2018) *Imine hydrolysis*, *News*. Available at: https://www.news-medical.net/life-sciences/Imine-Hydrolysis.aspx (Accessed: January 21, 2023).

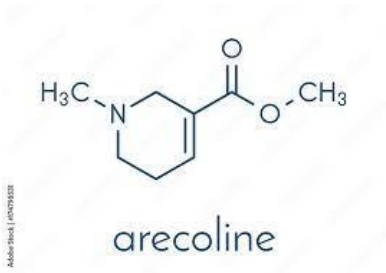
Isoguvacine



Arecaidine



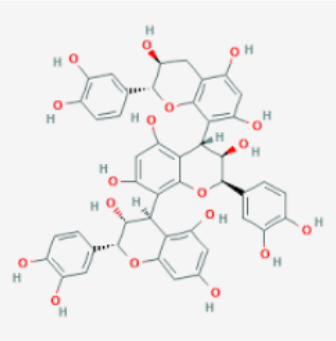
Isoguvacine and Arecaidine also have negative LogP values of -1.4 and -1.04 respectively. Both of these are pyridine-based compounds that show polar characteristics. Isoguvacine in particular has a hydrogen bond (OH) in its pyridine ring and this gives it a very polar bond, making its logP value more negative as previously established. As Isoguvacine has more hydrogen bonds than Isoguvacine, it allows it to form di-pole di-pole forces with the water molecules than Arecaidine. Arecaidine has a solubility value of 510 mg/mL which is one of the highest in the group in order to establish its polar nature.



Taking Arecoline (the most toxic betel nut compound) as an example, the only polar bond that is present is the C=O which is strongly induced by the existence of bulky methyl groups attached to the carbon and oxygen atoms to possibly have an even distribution of charges acting against the dipole moment of the oxygen in C=O. This helps the structure maintain some uniformity in the net dipole distribution. This gives a positive LogP value of 0.83 for Arecoline which is similar to the other groups of compounds that lie above the polarizability axis.

Arectannin A1 is part of a select group of Areca nut compounds that are Tannins[42]. These Tannins are a select group of compounds with polyphenolic content in the Areca nut[43].

Arectannin A1



The large molecular mass of Arectannin A1 much like Trabectedin has many bonds. For the aforementioned reasons for larger molecular mass compounds, it has higher LogP and Polarizability values adding to its non-polar nature as a polyphenol[44] with an extremely nonpolar LogP value of 3.9.

8. Conclusion

With R^2 values around 0.5 for a linear trend in both the sets of compounds, my hypothesis was mostly proven to be valid. This is due to the positive correlation between the LogP and polarizability values With higher polarizability values, there are more London

[42] Chen, X., He, Y., & Deng, Y. (2021, August 18). *Chemical composition, pharmacological, and toxicological effects of betel nut*. "2.2 Tannins," Evidence-based complementary and alternative medicine : eCAM. Retrieved January 12, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/

[43] RS;, C. Y. V. S. (n.d.). *Separation of polyphenols and arecoline from areca nut (areca catechu L.) by solvent extraction, its antioxidant activity, and identification of polyphenols*. Journal of the science of food and agriculture. Retrieved January 24, 2023, from https://pubmed.ncbi.nlm.nih.gov/23494978/

[44] Haminiuk, C. W. I., Plata-Oviedo, M. S. V., de Mattos, G., Carpes, S. T., & Branco, I. G. (2014, October). *Extraction and quantification of phenolic acids and flavonols from Eugenia pyriformis using different solvents*. Journal of food science and technology. Retrieved January 17, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4190214/#:~:text=Polyphenols%20are%20often%20most%20soluble,(Kim%20and%20Lee%202001).

dispersion forces present making the overall molecular structures nonpolar giving them more +LogP values.

As highlighted in the analysis, Dacrabazine, Temozolomide, and Cisplatin serve as primary examples to this deviation in the trend. Especially Cisplatin with an empirical LogP of -2.41 as opposed to 0.041 being its predicted LogP value. This is due to its ionizability in water and other structural parameters as discussed before. Altretamine is another example of this trend that had a slightly more polar empirical LogP value of about 1.3 and a predicted value of 2.32 much like many of the other +LogP value compounds. The rest of the compounds had marginal deviations of less than 0.5. This helped shift the empirical trend to the right from the original predicted trend.

The really large compounds from both categories of compounds such as Arectannin A1 and Tribactedin had higher polarizability values and overall nonpolar natures due to the greater number of temporary dipoles that were induced among the partial charges in their large molecular masses.

The nicotine substituent specifically showed more polar characteristics than the betel nut compounds allowing for the presence of nitrogen atoms in the pyridine rings that are capable of forming hydrogen bonds with water molecules and its protonation mechanism. These compounds have negative LogP values and have high solubility values which makes them more polar by nature.

To conclude, the overall structural natures were analyzed to determine the extent to which the chosen compounds favor the lipophilicity of diffusing across the cell membrane. The Chemotherapeutic compounds showed varying degrees of structural polarity but the compounds that were being used the most for administration much like Cyclophosphamide, Cisplatin, etc

exhibited a relatively low LogP value along with relatively high solubility values backed up their structural arrangements. The carcinogenic compounds followed a similar trend to a greater extent with most of the compounds having higher solubility values and especially nicotine based compounds having the lowest LogP values to affect the bloodstream and diffuse membrane most readily.

9. Evaluation of sources and methodology

Weakness: My data sources were limited to reliable computational software sources for the predicted values. This serves as a major limitation in data collection with the availability of only 3 database sources that could cover all the compounds that were chosen to be analyzed. Some compounds didn't have certain values from 2 out of the three database sources so their averages were taken only from one source. Of which, the chemotherapeutic agents only had two databases or Polarizability values as opposed to three for the carcinogenic compounds.

Strength: The overall availability of data across various compounds through the use of predicted models from reliable sources such as ChemAxon and ALOGPS compiled in the HMDB database for most of my compounds. As I chose compounds from specific groups, their classifications were already prevalent and were widely accepted by the scientific community. They were used to analyze the molecular polarities shared by the compounds of each group.

Strength: Despite the lack of coherence in data collection, the average uncertainties that were calculated from the table, the uncertainties in both predicted LogP values, as well as the polarizability values for most of the carcinogenic compounds, only gave up to a maximum of 1 in both variables. The carcinogens were very consistent with a very similar uncertainty range but the maximum being 1.6 just for one compound.

Weakness: The empirical values were only present for LogP for a decent amount of chemotherapeutic agents as opposed to just two from the carcinogenic agents. Another key weakness to note is the fact that the empirical LogP values had no reference source to possibly outline some systematic errors. As a result, the analysis was limited to solubility and molecular polarity.

Strength: The use of an additional predicted parameter such as the solubility values to identify further scope for analysis and to back up existing LogP values helped to analyze and verify the classifications at a greater extent in terms of molecular polarity.

Weakness: The solubility values for Temozolomide and Melphalan in the predicted set were not available and hence were not considered for their structural analysis.

Weakness: The net di-pole moment values for the compounds were not as available to be taken into consideration as an additional parameter for the molecular polarities.

10. Extensions/Alternative methods

The binding affinity of a drug to its protein receptor can be analyzed based on intermolecular forces of attraction as an extension to its structural characteristics. In order to get a clearer picture of the polarity of the compounds, the net di-pole or magnetic moments can be obtained from computational software. The acid dissociation constant values in an aqueous medium like water can also be an additional factor that can help determine the extent of ionization of compounds. The compounds can be classified using different means such as their range of molecular masses to see how the trend compares to that of just polarizability and Log values rather than using it just as a parameter to back up the existing relationship between the

LogP and Polarizability values in terms of the intermolecular forces of attraction. The LogD[45] Values of compounds can be used as a variable on their own or as an additional parameter to discuss the lipophilicity of compounds based on their extent of ionizability in a polar or nonpolar solvent. For non-ionizable compounds, LogD = LogP at all pH ranges[46].

11) Unresolved Questions

After the analysis, despite exhibiting more polar characteristics, certain compounds like nicotine imine had higher polarizability values despite it having the highest -LogP value in all the data sets combined. This can be analyzed further using specific nicotine-based compounds and other variants. Variants within a certain group of compounds such as Bendamustine exhibit varied non-polar properties. Dacarbazine's negative LogP values differed from its low solubility value; this was not explored in the analysis. The variation between empirical and predicted logP values among the chemotherapeutic alkaloids cannot be fully analyzed due to limited data sources for the empirical LogP values and the understanding of the computational way of predicting LogP values was never truly answered due to the lack of reference from the HMDB database which limited the scope to analyze and understand the systematic errors associated with the empirical methodologies.

[45] Distribution of **ionizable** compounds in the different concentrations of solvents.

[46] *Lipophilicity descriptors: Understanding when to use LogP & LogD - ACD/Labs* (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/logp_vs_logd.pdf (Accessed: January 31, 2023).

12) Bibliography

Lipophilicity descriptors: Understanding when to use LogP & LogD - ACD/Labs (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/logp_vs_logd.pdf (Accessed: January 31, 2023).

Lipophilicity descriptors: Understanding when to use LogP & LogD - ACD/Labs (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/logp_vs_logd.pdf (Accessed: January 31, 2023).

Haminiuk, C. W. I., Plata-Oviedo, M. S. V., de Mattos, G., Carpes, S. T., & Branco, I. G. (2014, October). *Extraction and quantification of phenolic acids and flavonols from Eugenia pyriformis using different solvents*. Journal of food science and technology. Retrieved January 17, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4190214/#:~:text=Polyphenols%20are%20often%20most%20soluble,(Kim%20and%20Lee%202001).

RS;, C. Y. V. S. (n.d.). *Separation of polyphenols and arecoline from areca nut (areca catechu L.) by solvent extraction, its antioxidant activity, and identification of polyphenols*. Journal of the science of food and agriculture. Retrieved January 24, 2023, from https://pubmed.ncbi.nlm.nih.gov/23494978/

Chen, X., He, Y., & Deng, Y. (2021, August 18). *Chemical composition, pharmacological, and toxicological effects of betel nut*. "2.2 Tannins," Evidence-based complementary and alternative medicine: eCAM. Retrieved January 12, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/

Aliouche, H. (2018) *Imine hydrolysis*, *News*. Available at: https://www.news-medical.net/life-sciences/Imine-Hydrolysis.aspx (Accessed: January 21, 2023).

Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine: eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 7, 2023).

Chang, Raymond. "Chapter 13: Intermolecular Forces/ Ion-Induced Dipole and Dipole-Induced Dipole Interactions/ Dispersion, or London, Interactions." Physical Chemistry for the Biosciences. Sansalito, CA: University Science, 2005. 495-98. Print.

Petrucci, Ralph H., et al. General Chemistry: Principles and Modern Applications.Upper Saddle River, NJ: Prentice Hall, 2007

Martinho, N. (2019) *Cisplatin-Membrane Interactions and Their Influence on Platinum Complexes Activity and Toxicity*, *National Center for Biotechnology Information (NCBI)*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6336831/ (Accessed: January 8, 2023).

Aziridine (no date) *Aziridine - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/nursing-and-health-professions/aziridine (Accessed: January 31, 2023).

Thiotepa (no date) *Thiotepa - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/chemistry/thiotepa#:~:text=Mechanism%20of%20Toxicity&text=Thiotepa%20(and%20TEPA)%20form%20DNA,disruption%20of%20nucleic%20acid%20function. (Accessed: January 29, 2023).

Thiotepa (no date) *Thiotepa - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/chemistry/thiotepa#:~:text=Mechanism%20of%20Toxicity&text=Thiotepa%20(and%20TEPA)%20form%20DNA,disruption%20of%20nucleic%20acid%20function. (Accessed: January 29, 2023).

Databases (2022) *The Metabolomics Innovation Centre*. Available at: https://metabolomicscentre.ca/software-databases/databases/ (Accessed: January 18, 2023).

Center for Drug Evaluation and Research (no date) *Drug approvals and databases*, *U.S. Food and Drug Administration*. FDA. Available at: https://www.fda.gov/drugs/development-approval-process-drugs/drug-approvals-and-databases (Accessed: January 9, 2023).

Search ChemSpider (no date) *ChemSpider*. Available at: http://www.chemspider.com/Default.aspx (Accessed: January 9, 2023).

Human metabolome database (no date) *Human Metabolome Database*. Available at: https://hmdb.ca/ (Accessed: January 9, 2023).

Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine: eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 8, 2023).

Holland, J.F. and Colvin, M. (2003) "Alkylating Agents," in *Cancer Medicine*. BC Decker Inc.

LogP—making sense of the value - acdlabs.com (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/making_sense.pdf (Accessed: January 16, 2023).

LogP—making sense of the value - acdlabs.com (no date). Available at: https://www.acdlabs.com/wp-content/uploads/download/app/physchem/making_sense.pdf (Accessed: January 16, 2023). .

Partition coefficient (no date) *Partition Coefficient - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/medicine-and-dentistry/partition-coefficient#:~:text=The%20partition%20coefficient%20of%20a,and%20the%20value%20is%20calculated. (Accessed: January 7, 2023).

Partition coefficient (no date) *Partition Coefficient - an overview | ScienceDirect Topics*. Available at: https://www.sciencedirect.com/topics/medicine-and-dentistry/partition-coefficient#:~:text=The%20partition%20coefficient%20of%20a,and%20the%20value%20is%20calculated. (Accessed: January 7, 2023).

University of Wisconsin–Eau Claire (no date). Available at: https://www.chem.uwec.edu/Chem406_F06/Pages/lecture_notes/lect05/Atkins-Ch16_small.pdf (Accessed: January 22, 2023).

Baker, M. *et al.* (no date) *7.3 molecular polarity and dipole moments*, *Chemistry Fundamentals*. UCF Pressbooks. Available at: https://pressbooks.online.ucf.edu/chemistryfundamentals/chapter/7-3-molecular-polarity-and-dipole-moments/ (Accessed: January 28, 2023).

Thomas, G. (2007). *Fundamentals of Medicinal Chemistry*. Wiley.

Browsing metabolites (no date) *Human Metabolome Database: Browsing metabolites*. Available at: https://hmdb.ca/metabolites (Accessed: January 25, 2023).

Cisplatin (no date) *Uses, Interactions, Mechanism of Action | DrugBank Online*. Available at: https://go.drugbank.com (Accessed: January 15, 2023).

Savjani, K.T., Gajjar, A.K. and Savjani, J.K. (2012) *Drug solubility: Importance and enhancement techniques*, *ISRN pharmaceutics*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3399483/ (Accessed: January 16, 2023).

Petrucci, Ralph H., et al. General Chemistry: Principles and Modern Applications.Upper Saddle River, NJ: Prentice Hall, 2007

Chang, Raymond. "Chapter 13: Intermolecular Forces/ Ion-Induced Dipole and Dipole-Induced Dipole Interactions/ Dispersion, or London, Interactions." *Physical Chemistry for the Biosciences*. Sansalito, CA: University Science, 2005. 495-98. Print.

Chen, X., He, Y. and Deng, Y. (2021) *Chemical composition, pharmacological, and toxicological effects of betel nut*, *Evidence-based complementary and alternative medicine: eCAM*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8387188/ (Accessed: January 5, 2023).

Thomas, G. (2007). "3.3.1 Fundamentals of water solubility" *Fundamentals of Medicinal Chemistry*. Wiley.

Thomas, G. (2007). "Phases of drug action" *Fundamentals of Medicinal Chemistry*. Wiley.

Cell membranes - the cell - NCBI bookshelf (no date). Available at: https://www.ncbi.nlm.nih.gov/books/NBK9928/ (Accessed: January 4, 2023).

The cell cycle and cancer | PNAS (no date). Available at: https://www.pnas.org/doi/10.1073/pnas.94.7.2776 (Accessed Dec 11, 2022).

Physiology, blood plasma - statpearls - NCBI bookshelf (no date). Available at:https://www.ncbi.nlm.nih.gov/books/NBK531504/ (Accessed: January 24, 2023).