Original Research Article

Potential Antineoplastic Structural Variations of Uracil Mustard (Uramustine) Retaining Cytotoxic Activity and Drug-likeness Suitable for Oral Administration

ABSTRACT

Aims: To present 12 new variants of uracil mustard having drug-like properties and cytotoxic functional group, by utilizing uracil mustard (uramustine) as a lead compound. Utilize rigorous substructure and similarity of a molecular scaffold to determine drug like variants. Physicochemical properties determined indicate the variants have favorable drug-likeness.

Study design: Conduct molecular database search utilizing features of substructure and similarity based upon uracil mustard.

Place and Duration of Study: Department of Chemistry, Medicinal Chemistry Study Section, University of Nebraska at Omaha, Omaha Nebraska between January 2015 to March 2015.

Methodology: Uracil mustard consists of the pyrimidine derivative uracil, having the bifunctional nitrogen mustard cytotoxic moiety covalently bonded onto the ring. A systematic search, utilizing substructure component and similarity, within an in-silico database search successfully determined 12 variants. Rigorous criteria for drug-likeness was implemented to screen potential candidates that included the application of the Rule of 5. In addition, maintaining the cytotoxic moiety of nitrogen mustard was crucial.

Results: A total of 12 variants of uracil mustard was identified after an extensive molecular database search using rigorous criteria. All 12 variants, and including uracil mustard, showed zero violations of the Rule of 5, thereby indicating favorable drug-likeness. Values of polar surface area for all compounds at less than 80 Angstroms^2 are suitable for central nervous system penetration. Polar surface area, number of atoms, and Log P for all compounds increased as the molecular weight increases. Structure substituents include nitrogen mustard *groups*, hydroxyl, alkyl, and carbonyl moieties. Cluster analysis discerned greatest similarity among members of this group.

Conclusion: Applying rigorous search criteria within a molecular data base, for comparison and reject, successfully identified 12 variants of uracil mustard that show favorable drug-likeness in addition to cytotoxic capability. The design of new antitumor agents is important for increasing efficacy of the clinical treatment of cancer.

Keywords: uracil mustard, uramustine, cancer, leukemia, lymphoma

1. INTRODUCTION

Uramustine or uracil mustard is an alkylating chemotherapy agent that is particularly effective for the treatment of lymphomas, Non-Hodgkin's lymphoma, chronic lymphocytic leukemia, chronic myelogenous leukemia and chronic leukemias [1,2]. It is used in lymphatic malignancies such as non-Hodgkin's lymphoma. Nitrogen mustard agents alkylate the DNA and thereby induce damage to the DNA, which in turn is a cytotoxic effect primarily to cancer cells [1]. This due to the take up of uracil due to the need to form nucleic acids during rapid

cycles of cell division. The DNA damage (cross-linking) leads to apoptosis of the affected cells. This agent is cell cycle-phase nonspecific. Chemically it is a derivative of nitrogen mustard and uracil. Notable among its advantages are: 1) small dose requirement; 2) uniform tolerability; 3) simple regimen requirements [1].

Previous studies have shown the efficacy of uracil mustard for the treatment of thrombocythemia, chronic lymphatic leukemia, and lymphoma [3]. Investigators record objective improvement in patients having solid tumor, lymphoma, and leukemia [4,5]. Clinical studies have also demonstrated the effectiveness of uracil mustard for treatment of Hodgkin disease, non-Hodgkin lymphoma, and chronic lymphatic leukemia [6, 7, 8, 9].

Uracil mustard has been shown to bring about a dramatic relief in patients having Hodgkin's disease and multiple myeloma, in addition to be well tolerated [10]. After extensive clinical trials this agent has been found effective in treatment of granulocytic leukemias [10] and childhood acute leukemia [11]. Uracil mustard is declared an effective drug for controlling thrombocytosis with minimal effects on leukocytes and erythrocytes [12]. In the case of elderly patients fighting non-Hodgkin's lymphoma, the fifth most common malignancy for male and female in the United States, the use of uracil mustard would be effective [13, 6, 7]. This lymphoma originates in the lymphatic system, causing greater than 18,000 deaths in the United States annually.

In-silico studies have been successful for characterizing the molecular activity of uracil mustard with cellular targets such as DNA [15]. Optimization of molecular scaffolding with concurrent elucidation of physicochemical properties has been successful for characterization of drug efficacy treating dermal neoplasm [16]. This study presents 12 variants of uracil mustard resulting from in-silico substructure and similarity mining with use of rigor enjoined by uracil mustard scaffolding parametric limitation.

2. MATERIAL AND METHODS

2.1 Molecular Modeling and Assembly of Molecular Variants

Molecular modeling (2-D) was accomplished utilizing ACD/ChemSketch modeling v. 10.00 (Advanced Chemistry Development, 110 Yonge Street, Toronto Ontario, M5C 1T4 Canada). In silico structure search for a substituent replacement was accomplished using chemical substructure and similarity search with Molsoft L.L.C. (Molsoft L.L.C. 11199 Sorrento Valley Road, S209 San Diego, CA 92121) and Molinspiration (Molinspiration Chemiformatics, Nova ulica 61, SK-900 26 Slovensky Grob, Slovak Republic). Various properties such as polar surface area, violations of Rule of 5, molecular volume, number of oxygens, nitrogens, amines ($-NH_n$), and hydroxyls (-OH), were determined using Molinspiration Properties Calculations module (Molinspiration Chemiformatics, Nova ulica 61, SK-900 26 Slovensky Grob, Slovak Republic). Biological activity of all compounds was determined by Molinspiration drug-likeness and bioactivity scoring (Molinspiration Chemiformatics, Nova ulica 61, SK-900 26 Slovensky Grob, Slovak Republic).

2.2 Pattern Recognition

To identify underlying associations and patterns within the descriptors multivariate numerical data matrix, then various pattern recognition techniques were implemented. Included in this analytical approach is a hierarchical cluster analysis accomplished by KyPlot v. 2.0 Beta 15 (copyright Koichi Yoshioka 1997-2001). Other pattern recognition elucidation by K-means nonhierarchical cluster analysis and discriminant analysis were performed by PAST v. 1.80 (copyright Oyvind Hammer, D.A.T. Harper 1999-2008).

2.3 Numerical Analysis of Multivariate Physicochemical Properties

Descriptive statistical analysis, Pearson r, and coefficient of determination for all numerical data where indicated was performed by Windows 7 Microsoft Office Professional Plus 2013 EXCEL (EXCEL 2013). Screening for numerical outliers was done by Grubb's Test (extreme studentized deviate) by GraphPad Software **(**2236 Avenida de la Playa, La Jolla, CA 92037 USA). Multiple regression analysis was performed by GraphPad InStat v. 3.0 for Windows 95 (HJ Motulsky, GraphPad InStat 3.0 GraphPad Software, Inc., San Diego California USA, www.graphpad.com).

3. RESULTS AND DISCUSSION

Uracil mustard (uramustine) is known to be effective in the treatment of a multitude of neoplastic diseases. Data mining for substructure replacement has undergone progress in recent years and presents an efficacious tool for drug design, particularly with the presence of effective parent constructs that are applied in seeding the mining process [15,16]. Utilizing uracil mustard as the parent compound for drug design (see compound 1, Figure 1) the outcome nitrogen mustard agents 2 through 13 were determined after an algorithmic search process generating numerous candidates, of which, included those having unfavorable bioavailability attributes (ie. high formula weight and polar surface area) to be identified and eliminated. This rate of ruination makes it necessary of having information descriptive of successful candidates, concerning the desired (drug-like) physicochemical properties. Therefore, substituent replacement is clearly an important approach in rational drug design [17].

In general, the aromatic heterocyclic organic pyrimidine ring remains in addition to the bifunctional cytotoxic nitrogen mustard group that renders the alkylating antineoplastic action. The parent compound with 12 variants are presented in Figure 1. The variants of uracil mustard are essentially derivatives having an additional substitution in place of the hydrogen (-H) located onto N-3. Substituents located onto N-3 then will introduce variations of pharmaceutical properties such as Log P, polar surface area, molecular weight, etc. It is vital to maintain a favorable drug-likeness of properties while inducing variation that would benefit the clinical efficacy of the compound. The drug-likeness of the compounds was monitored and rigorously made to adhere to the Rule of 5, which is a robust approach to identify potential drugs having good membrane permeability and orally active (easily absorbed) [18]. An orally active drug will have no more than one violation of the following criteria [18]: 1) No more than 5 hydrogen bond donors (-OH and -NH_n); 2) Not more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms); 3) A molecular mass less than 500 Daltons; 4) An octanol-water partition coefficient Log P not greater than 5.

The molecular structure of uracil mustard, with 12 variants are presented in Figure 1. Notably, each structure possesses the nitrogen mustard group responsible for the nucleic acid alkylating cytotoxic activity (located at the C-5 position). All variations of uracil mustard (#2 to #13) have substituents on the N-3 position. Beginning with compound #2, the N-3 substituent is an acyl group ($CH_3C(=O)$-), followed by compound #3 having a propanoyl group ($CH_3CH_2C(=O)$-), followed by compounds #4 to #8 having alkyl substitutions

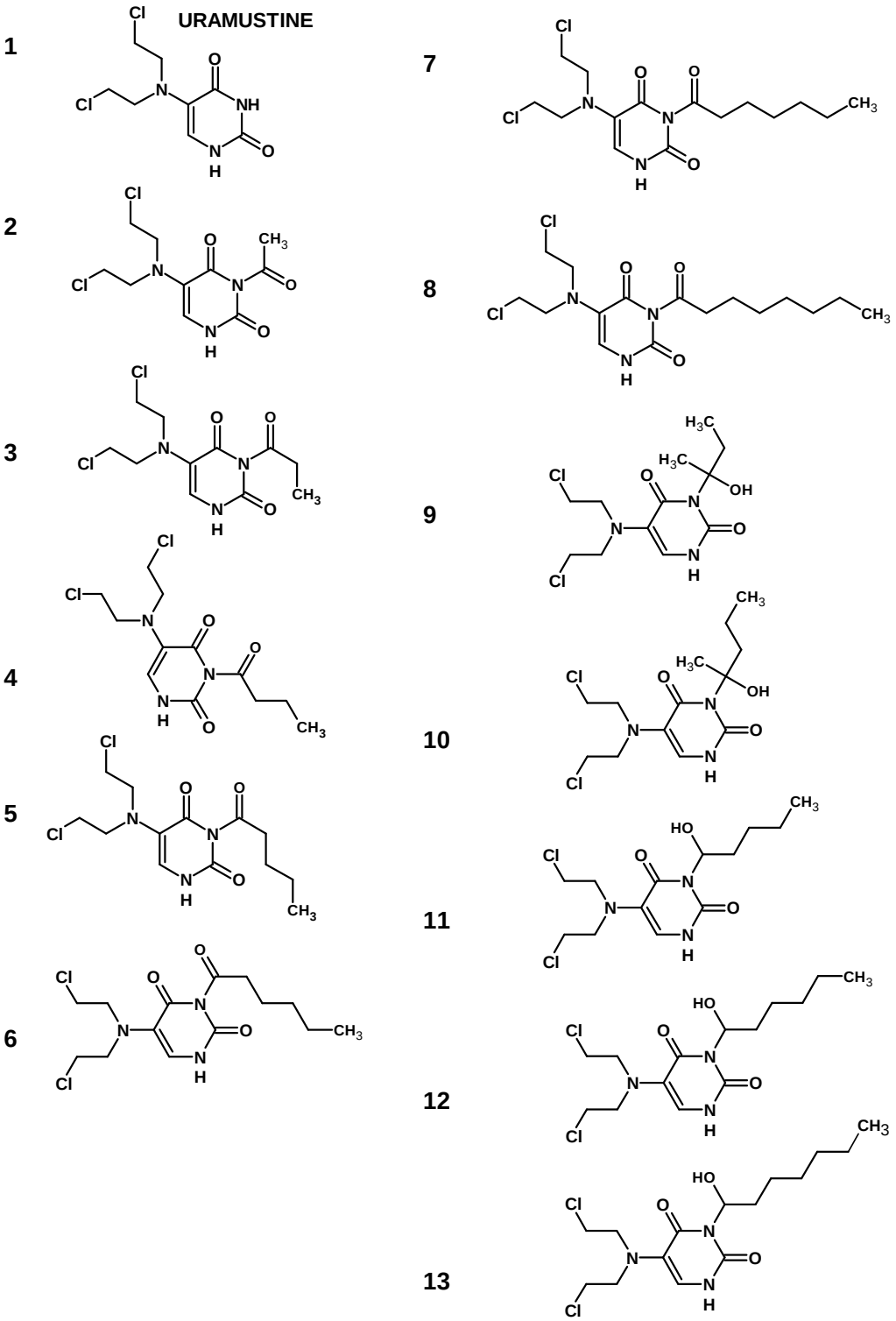

Fig. 1. Molecular structure of uracil mustard and its variants. The molecular scaffolding includes the pyrimidine derivative uracil, alkylating nitrogen mustard moiety (covalently bonded to C-5), hydroxyl, alkyl, and carbonyl moieties (collectively attached to the pyrimidine uracil on N-3).

($CH_3(CH_2)nC(=O)$-) where n = 2 to 6, respectively. Compounds #2 to #8 constitute a homologous series (differ by a constant unit, generally a (-CH2-) group) of agents which have been shown to have similar and often times increasing beneficial medicinal activity [19]. Likewise, for compounds #11 to #13 shown in Figure 1, the N-3 position has substituents ($-C(OH)(CH_2)_nCH_3$) where n = 3 to 5, respectively. Hence, two homologous series of compounds have been identified within this group of compounds.

Physicochemical properties useful for evaluating pharmaceutical potential are presented for all 13 compounds, see Table 1. Noteworthy points include the very strong positive correlation (Pearson r > 0.9000) for Log P to the number of atoms, molecular weight, rotatable bonds, and molecular volume. In addition, there is a strong positive correlation (Pearson r > 0.6000) for polar surface area to the number of atoms, molecular weight, and molecular volume. The range in Log P values for this group of compounds is broad (0.563 to 3.968), having value of 3.405. Previous studies have shown the efficacy of varying the Log P of medicaments in the case of antineoplastic drugs [16], as well as antibacterial drugs [20, 21, 22]. In these studies the medicinal activity of the agent was increased in the case of antibacterial drugs or enhanced for antineoplastic agents.

Table 1. Physicochemical Properties of Compounds

Drug	Log P	Polar Surface Area (A^2)	Number of Atoms	Molecular Weight	Number O & N	Number of –OH & $-NH_2$	Rule of 5 Violations	Rotatable Bonds	Molecular Volume (A^3)
1 uramustine	0.563	68.9	15	252.1	5	2	0	5	199.0
2	0.885	75.2	18	294.1	6	1	0	5	234.9
3	1.387	75.2	19	308.2	6	1	0	6	251.8
4	1.947	75.2	20	322.2	6	1	0	7	268.6
5	2.452	75.2	21	336.2	6	1	0	8	285.4
6	2.957	75.2	22	350.3	6	1	0	9	302.2
7	3.462	75.2	23	364.3	6	1	0	10	319.0
8	3.968	75.2	24	378.3	6	1	0	11	335.8
9	1.298	78.3	20	324.2	6	2	0	7	273.9
10	1.858	78.3	21	338.2	6	2	0	8	290.7
11	2.363	78.3	22	352.3	6	2	0	9	307.5
12	2.868	78.3	23	366.3	6	2	0	10	324.3
13	3.373	78.3	24	380.3	6	2	0	11	341.1

A^2 =Angstroms^2; A^3 =Angstroms^3

Notably, all compounds have zero violations of the Rule of 5, indicating favorable membrane permeation and absorption (i.e. drug-likeness). Values of polar surface area are kept low, having a small range of 68.9 Angstroms^2 to 78.3 Angstroms^2. These values again indicate favorable drug absorption [23]. Previous investigation as to drug-likeness of known pharmaceuticals suggest that all compounds #1 to #13 would be greater than 50% absorbed from the intestinal tract [23]. These properties support the potential of the 12 variants of uracil mustard.

Further understanding of the underlying relationships between these compounds is accomplished using pattern recognition method of hierarchical cluster analysis (a multilevel hierarchy, where clusters at one level are joined as clusters at the next level) [24]. Results of hierarchical cluster analysis utilizing a single linkage (minimum distance between elements of each cluster) and utilizing Euclidian distance (the most common distance measure, are the geometric distance in multidimensional space) is presented in a divisive vertical dendrogram (see Figure 2).

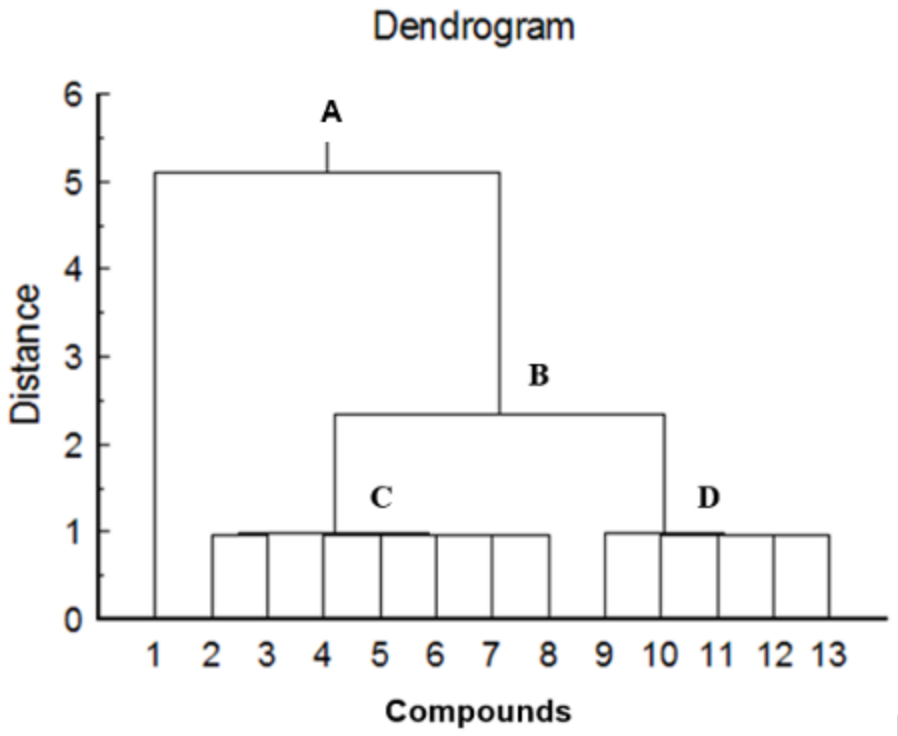

Fig. 2. Hierarchical cluster analysis of all compounds utilizing single linkage cluster conditions with Euclidean distance. Compound 1 is uracil mustard joined at node A with all remaining compounds. Node B joins compounds 2 to 13 into two sub clusters under node C (2, 3, 4, 5, 6, 7 and 8) and node D (9, 10, 11, 12, and 13).

Interestingly, the compounds #2 to #8 are clustered together to be more similar to each other and joined at node C. It follows, that compounds #9 to #13 are likewise determined to be more similar to each other and joined at node D. These two clusters are in turn joined at super node B. The super node A joins node B to the parent compound (#1). This analysis has distinguished the parent compound uracil mustard (#1) from the 12 variations of ring substituents.

Other trends in properties can be discerned by non-hierarchical cluster analysis. K-means cluster analysis is used to classify observations through a K number of clusters, and aims to find a grouping of objects which maximizes or minimizes some evaluating criterion [24]. Results of K-means cluster analysis places these compounds into four distinct clusters, as follows: cluster 1) compounds #4, #5, #9, #10; cluster 2) compounds #6, #7, #11, #12; cluster 3) compounds #8, #13; cluster 4) uracil mustard (#1), compounds #2, and #3. Therefore, a connectivity is identified among the parent compound uracil mustard and variants #2 and #3. To follow, the analogy is determined by discriminant analysis being a process used to determine which variables discriminate between two or more naturally occurring groups. Results of discriminant analysis indicate the commonality among uracil mustard (#1) to compounds #9, #10, #11, #12, and #13 (compounds having alcoholic homologous series). Distinct from these compounds are those compounds #2, #3, #4, #5, #6, #7, and #8 (compounds having alkanoyl homologous series).

Grubbs' test, also known as the maximum normed residual test or extreme studentized deviate test, is a statistical test used to detect outliers in a univariate data set assumed to

come from a normally distributed population [25]. Analysis of all 13 compounds for outliers utilizing Grubb's test, showed no outliers among the values of Log P, the number of atoms, molecular weight, number of oxygen/nitrogen atoms, and molecular volume. These findings, shows consistency with the parent drug uracil mustard.

Other discernable and highly linear trends of these compounds are the steadily increasing polar surface area (Pearson r = 0.6887) and number of atoms (Pearson r = 0.9996) as dependent variables, when compared to molecular weight (independent variable), see Figure 3.

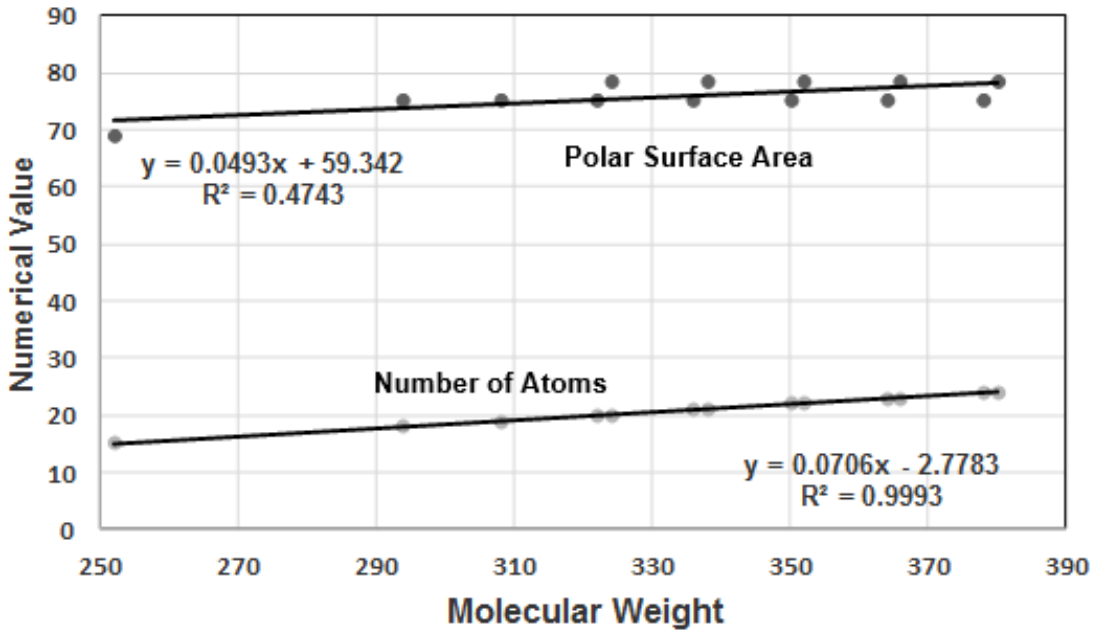

Fig. 3. Comparison of polar surface area and number of atoms (dependent variables) to compound molecular weight (independent variable). Number of atoms are highly linear with very strong positive relationship (Pearson r = 0.9996, coefficient of determination, R^2 = 0.9993) with increase of molecular weight. Polar surface area is highly linear to increase of molecular weight with a strong positive relationship (Pearson r = 0.6887, coefficient of determination R^2 = 0.4743).

A 2-way plot of Log P (dependent variable) to molecular weight (independent variable) indicated an increasing strong exponential trend (R^2 = 0.9307), see Figure 4. The increase of Log P within a homologous series of compounds has been determined in previous studies to generally enhance and improve medicinal activity [19]. The rise in Log P values corresponding increase of molecular weight is consistent with homologous series of drug agents and found in other studies [20, 21, 22].

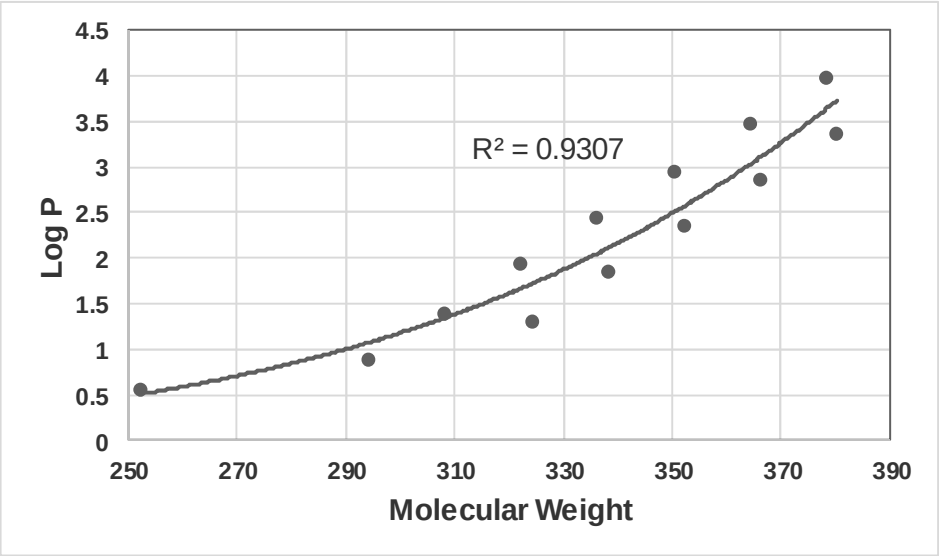

Fig. 4. Trend for Log P compared to molecular weight increase is a steady increase along (coefficient of determination, R^2 = 0.9307, correlation coefficient r = 0.9647).

The general goal of multiple regression is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. One use of multiple regression is a prediction or estimation of an unknown, dependent Y value corresponding to a set of independent X values. A second application of multiple regression is to understand the functional relationships between the dependent and independent variables, to try to see what might be causing the variation in the dependent variable [24]. Multiple regression analysis of the descriptors presented in Table 1 resulted in a mathematical description that accounts for 100% of the variance within the model (R^2 = 1.000). The properties applied include molecular weight (MW), polar surface area (PSA), number of atoms (nAtoms), number of hydroxyl and amine groups (nOHNH), number of oxygen & nitrogen atoms (nON), and molecular volume (MV). The model is expressed as follows:

MW = 36.521 + 0.74068(PSA) + 5.900(nAtoms) - 2.155(nOHNH) -1.036(nON) + 0.5455(MV)

The equation describing the model can be used to predict important molecular properties that will allow the additional design of analogous compounds.

The diversity of potential drug targets is so enormous, that it is possible to find a common denominator for all of them [26]. The strategy which leads to success is focus on particular drug classes and the development of the specific activity score for each of these classes. This is accomplished for compounds presented in this study and listed in Table 2. The distribution of activity scores for the four most important drug classes is presented in Table 2. For ion channel modulator activity and drug-likeness, the best score falls between -1.30 to 0.50 [26]. For kinase inhibitor activity and drug-likeness, the best score falls between -1.30 to 0.50 [26]. For protease inhibitor activity and drug-likeness, the best score falls between -1.10 to 0.50 [26]. For other enzyme inhibitor activity and drug-likeness, the best score falls between -1.10 to 0.50 [26]. Therefore, all compounds (#1 to #13) presented in this study show the best biological activity score for active compounds and drug-likeness. This is

further evidence that these variants of uracil mustard will be effective and useful in clinical application.

Table 2. Biological Activity of Compounds by Score

Drug	Ion Channel Modulator	Kinase Inhibitor	Protease Inhibitor	Enzyme Inhibitor
1 uramustine	-0.76	-0.50	-1.03	-0.26
2	-1.02	-0.05	-0.52	-0.14
3	-0.92	-0.08	-0.48	-0.11
4	-0.88	0.04	-0.40	0.0
5	-0.83	0.08	-0.32	0.02
6	-0.79	0.10	-0.27	0.02
7	-0.76	0.11	-0.23	0.01
8	-0.73	0.11	-0.20	0.01
9	-0.39	0.33	-0.15	0.16
10	-0.37	0.34	-0.09	0.17
11	-0.35	0.35	-0.03	0.18
12	-0.33	0.35	0.0	0.17
13	-0.32	0.34	0.02	0.17

The design of anticancer drugs is an important aspect for the clinical treatment of cancer. Improvement in the treatment and improvement of clinical outcome is enhanced by the introduction of the versatile and novel drugs initiated by rational drug design. Presented here are 12 variations of uracil mustard that have shown useful drug-likeness and possess the cytotoxic nitrogen mustard alkylating functional group. Further studies of novel drug designs would be useful for advancing the treatment of cancer.

4. CONCLUSION

Twelve structures which are variants of uracil mustard were identified utilizing rigorous physicochemical criteria of substructure and similarity to this nitrogen mustard anticancer agent. Additions to the scaffold of uracil mustard included hydroxyl groups, alkyl carbon chains, and carbonyl groups. The wide range of Log P values from 0.885 to 3.968 contributes to a diverse potential in the use of these cytotoxic variants of uracil mustard, as shown in previous studies. Hierarchical cluster analysis and discriminant analysis distinguished uracil mustard from these variants, however K-means cluster analysis identified two variants having an acyl ($CH_3C(=O)$-) group (#2) and propanoyl ($CH_3CH_2C(=O)$-) group (#3) bonded to the #3 nitrogen of uracil base to be most similar to uracil mustard. As compound molecular weight increases the polar surface area, number of atoms, and Log P increase, respectively. Multiple regression determined the equation accounting for 100% of variance modeling. Producing cytotoxic variants of uracil mustard showing zero violations of the Rule of 5 (good oral availability), this study demonstrates the efficacy of drug design following rigorous criteria for substructure and similarity. Design of novel or improved anticancer agents ultimately will benefit the patient as well as the clinical choices for

treatment of neoplastic disease. Variation of physicochemical properties can benefit the efficacy of anticancer drugs and should be further investigated for the benefit of patients.

ACKNOWLEDGEMENTS

This study was supported by the College of Arts & Sciences of the University of Nebraska, Omaha, Nebraska 68182 USA.

Competing Interests

Authors have declared that no competing interests exist.

AUTHORS' CONTRIBUTIONS

Dr. Ronald Bartzatt designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript.

CONSENT

Not applicable.

ETHICAL APPROVAL

Not applicable.

REFERENCES

1. Shanbrom E, Miller S, Haar H, Opfell R. Therapeutic spectrum of uracil-mustard, a new oral antitumor drug. JAMA. 1960;174(13):1702-05.

2. Kennedy BJ, Theologides A. Uracil mustard, a new alkylating agent for oral administration in the management of patients with leukemia and lymphoma. N Engl J Med. 1961;264:790-93.

3. Kennedy BJ, Torkelson JL, Torlakovic E. Uracil mustard revisited. Cancer. 1999;85(10):2265-72.

4. Lane M, Lipowska B, Hall TC, Colsky J. Observations on the clinical pharmacology of 5-[bis(2-chloroethyl) amino] uracil (uracil mustard; NXC-34462). Cancer Chemother Rep. 1960;9:31-6.

5. Fortuny IE, Theologides A, Kennedy BJ. Single, combined and sequentrial use of uracil mustard and prednisone. Minn Med. 1972:55:715-6.

6. Kennedy BJ, Theologides A. Effect of uracil mustard in leukemia and lymphoma. Clin Res. 1960;8:251.

7. Kennedy BJ, Theologides A. Effect of uracil mustard in chronic lymphatic leukemia and malignant lymphoma. Cancer Chemother Rep. 1961;11:77-80.

8. Lane M, Kelly MG. The antitumor activity of 5-bis(2-chlorethyl)-aminouracil (uracil mustard). Cancer Res. 1960;20:511-7.

9. Lyttle DA, Petering HB. 5-bis(2-chloroethyl)-aminouracil, a new antitumor agent. J Natl Cancer Inst. 1959;23:153-62.

10. Ellis GP, West GB. Progress in medicinal chemistry. Google E-Book: Butterworth-Heinemann; 2011.

11. Porter FS, Holowach J, Thurman WG. Uracil mustard therapy in acute leukemia in children. Cancer Chemother Rep. 1962;18:79-82.

12. Shamasunder HK, Gregory SA, Knospe WH. Uracil mustard in the treatment of thrombocytosis. JAMA. 1980;244(13):1451-5.

13. Hinan MJ, Morrison VA. Therapeutic approaches to nonHodgkin's lymphoma in the elderly patient. Hematology. 2009;2(2):173-182.

14. Kohorn E. Uracil mustard and 5-fluorouracil combination chemotherapy: a historic record. Conn Med. 2013;77(7):433-6.

15. Doweyko AM, Mattes WB. An application of 3D-QSAR to the analysis of the sequence specificity of DNA alkylation by uracil mustard. Biochemistry. 1992;31(39):9388-92.

16. Bartzatt R. Two novel alkylating structural analogues of chloramphenicol suitable for topical treatment of dermal neoplasms. J Adv. Med Pharm Sciences. 2014;1(1):11-23.

17. Nogrady T. Medicinal chemistry a biochemical approach. 2nd ed., Oxford: Oxford University Press; 1988.

18. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46(1-3):3-26.

19. Thomas G. Fundamentals of medicinal chemistry. West Sussex: John Wiley & Sons; 2003.

20. Bartzatt R, Cirillo S, Cirillo JD. Antibacterial agents inhibiting growth of ampicillin resistant *Escherichia coli*. Current Trends in Medicinal Chemistry. 2013;7:27-34.

21. Bartzatt R, Cirillo S, Cirillo JD. Design of ciprofloxacin derivatives that inhibit growth of methicillin resistant *Staphylococcus aureus* (MRSA) and methicillin susceptible *Staphylococcus aureus* (MSSA). Medicinal Chemistry. 2010;6:51-6.

22. Bartzatt R, Cirillo S, Cirillo JD. Derivatives of cephalothin that inhibit ampicillin resistant *Escherichia Coli*. Medicinal Chemistry. 2007;3(1):45-9.

23. Ertl P, Rohde B, Selzer P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. J Med Chem. 2000;43:3714-17.

24. Bow S. Pattern Recognition. New York: Marcel Dekker; 1984.

25. Grubbs, FE. Sample criteria for testing outlying observations. The Annals of Mathematical Statistics. 1950; 21(1):27-58.

26. Drug-likeness & bioactivity score. Molinspiration. Accessed 4 April 2015. Available: http://www.molinspiration.com/docu/miscreen/druglikeness.html

DEFINITIONS, ACRONYMS, ABBREVIATIONS

Term: PSA, polar surface area; A, angstroms; MW, molecular weight; nAtoms, number of atoms; nOHNH, number hydroxyl and amine groups; nON, number of oxygen and nitrogen atoms.

APPENDIX

NONE