2. Material and Methods

2.1 Hardware and Software

ACD/ChemSketch Version11 freeware, Version12 Advanced Chemistry Development Inc., Toronto, Canada is a software for drawing chemical structures that comes with other functionalities such as calculation of molecular properties, 2D and 3D structure cleaning, structure naming, and prediction of log P. The software is available in two versions: the commercial and freeware version. The freeware version does not include ACD/Dictionary, technical support, ACD/Lab extension for ChemDraw, and the function to search files by structure (Liew 2012). This program was used to design the chemical structure of the Anthrapyrazole AP compounds training sets, test sets and cross validation sets then to optimize their energy and to calculate their physicochemical properties.

Argus lab 4 software contains: An interactive 3D molecule builder that allows the user to build and manipulate complex structures; and, a rich suite of computational methods, both quantum mechanical and molecular mechanical, for calculating ground and excited states properties. The program is free for academic use.

Minitab statistical software16 for windows is an statistics package, developed at the Pennsylvania State University by researchers Barbara F. Ryan, Thomas A. Ryan, Jr., and Brian L. Joiner in 1972. Minitab is distributed by Minitab Inc., a privately owned company headquartered in State College, Pennsylvania, with subsidiaries in Coventry, England (Minitab Ltd.), Paris, France (Minitab SARL) and Sydney, Australia (Minitab Pty.). The latest version of the software is available in 7 languages: English, French, German, Japanese, Korean, Simplified Chinese, & Spanish.

All of the above mentioned programs were installed in an HP and Toshiba laptops, the operating system of which; is Microsoft Windows XP professional, Version 2007 and 2010 respectively. ChemSketch and Argus lab software were used to design the training and validating sets of anthrapyrazole compounds, optimize their geometry and calculate their total energies, heat of formation and dipole moments using AM1, PM3 and MNDO, semi-empirical quantum mechanical methods. ChemSketch software was also used to calculate the molar refractivity, molar volume and the logarithm of octanol /water partition coefficients (log P). Minitab Statistical Software16 for Windows (2009) was used for performing the regression analysis of the AP training sets and then calculating the regression equations (QSAR Models) and r-square values, while Microsoft excel 2010 was used for graphic correlation for the cross validation (internal and external validation sets).

2.2 Data Set (Biological activity data)

The training and validation sets for both the desoxy and dihydroxy Anthrapyrazole compounds were taken from the literature,(*Showalter et.al.1987*). Both of the groups having their biological activity expressed as optimum dose (O.D.) against P388 leukemia *in vivo* tumor cell lines and expressed as IC_{50} *in vitro* against murine L1210 Leukemia cell lines.

2.3 Computational Methodology

2.3.1 Desoxy AP compounds

The molecular geometries of desoxy AP compounds, both the training and cross validation sets are tabulated with their structure and the corresponding biological activity for both the P388 leukemia *in vivo* tumor cell lines and murine L1210 Leukemia cell lines as shown in table (I).

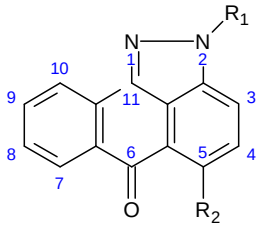

Table (1): Biological Data of desoxy anthrapyrazole against Murine L1210 leukemia *in Vitro* and P388 leukemia *in Vivo*

compound No.	**R1**	**NR2R3**	**L1210 IC_{50}**	**P388 mg/kg**
1	CH_3	$HN(CH_2)_2NH(CH_2)_2OH$	7.1E-07	100
2	$(CH_2)_2OH$	$HN(CH_2)_2NH(CH2)_2OH$	1.8E-06	25
3	$(CH_2)_2OH$	$HN(CH_2)_2NEt_2$	8.8E-07	400
4	$CH_2CH_2NH_2$	$HN(CH_2)_2NH(CH_2)_2OH$	8E-08	25
5	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2NH_2$	6.9E-08	25
6	$(CH_2)_2NHCH_2OH$	$HN(CH_2)_2NH(CH_2)_2OH$	7.4E-08	50
7	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2NMe_2$	3.2E-08	6.25
8	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2NEt_2$	6E-08	12.5
9	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NH(CH_2)_2OH$	3.2E-08	25
10	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2N_2$	4.6E-08	25
11	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NHMe$	2.7E-08	25
12	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NEt_2$	3.9E-07	50

These structures were initially optimized by using the molecular mechanics MMFF option in Arguslab-4. No electrostatic interactions were considered in these initial optimizations. The resulting optimized structures were then processed through the AM1, PM3 and MNDO semi empirical molecular orbital calculations furnished by Arguslab-4 package to obtain three sets of physiochemical parameters

representing total energies (E_t), standard heat of formation (ΔH_f) and dipole moment (μ) as shown in tables (2, 3, and 4).

Table (2): Total energies (E_t) of desoxy anthrapyrazole

compound No.	E_{tAM1}	E_{tPM3}	E_{tMNDO}
1	92372.68	92726.892	92824.19
2	94449.73	95333.13	95242.41
3	135427.3	52137.39	136182.3
4	92816.9	94557.18	93431.08
5	113436.5	113091.5	114161.6
6	126737.1	127386.3	127535.2
7	145694.8	145452	146546.2
8	179779	183497.3	180591.8
9	131864.6	135518.9	132714.5
10	106635	102304.8	103219.3
11	121195.3	120562.9	122015.6
12	170116.9	169066.6	170983.7

Table(3): Dipole moment of desoxy anthrapyrazole

compound No.	μ_{AM1}	μ_{PM3}	μ_{MNDO}
1	1643.062	1642.321	1642.423
2	1653.134	1652.696	1652.903
3	1895.878	510.0475	1895.571
4	1667.896	1665.52	1667.7
5	1648.741	1648.098	1648.485
6	1908.489	1908.456	1908.28
7	1926.916	1926.868	1926.746
8	2245.09	2245.629	2244.762
9	1834.127	1845.86	1933.956
10	1425.215	1417.362	1417.851
11	1558.33	1565.507	1557.966
12	2161.901	2161.546	2161.648

Table (4) Heat of formation of desoxy AP compounds

compound No.	H_f AM1	H_f PM3	H_f MNDO
1	188671.8	180233.6	189316.5
2	201684.4	193014.6	202727.2
3	242486.5	149973.1	243483.5
4	197798.2	189608	198637.2
5	21841.8	208142.3	219367.6
6	242654	232611.9	234732.8
7	257849.6	247388.9	258958.2
8	299107.1	292320.1	300209.7
9	251192.8	244341.8	252332.4
10	211440.8	197509.8	208241.8
11	229587.9	219210.9	230641
12	289269.6	278043.6	290418

The structures of the same set of desoxy AP compounds were also optimized using the molecular mechanics option in ChemSketch version 12. The optimization is based on modified molecular mechanics which take into account bond stretching, angle bending, and internal rotation and Van der Waals non-bonded interactions. The resulting optimized structures were then processed to calculate three sets of physiochemical parameters representing molar volume (MV), molecular refractivity (MR), polarizability and partition coefficient (log P) as shown in tables (5).

Table (5): Molar Refractivity, Molar volume, polarizability and Log P of the desoxy AP

compound No.	MR Cm3	MV Cm3	Polarizability	LogP
1	94.55	242.9	37.48	1.48
2	100.19	257.1	39.72	0.70
3	109.09	296.3	43.24	2.80
4	101.26	256.0	40.14	0.76
5	101.26	256.0	40.14	0.76
6	107.86	277.8	42.75	0.44
7	112.14	301.0	44.45	1.48
8	121.36	333.1	48.11	2.55

9	121.36	333.1	48.11	2.55
10	110.13	295.8	43.66	2.85
11	115.72	318.9	45.87	3.08
12	130.26	372.3	51.64	4.64

2.3.2 Dihydroxy AP compounds

The molecular geometries of dihydroxy AP compounds are tabulated with their structure and the corresponding biological activity for both the P388 leukemia in vivo tumor cell lines and murine L1210 Leukemia cell lines as shown in table (6).

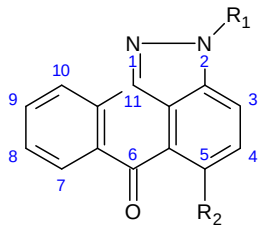

Table (6): Biological Data of dihydroxy anthrapyrazole against Murine L1210 leukemia *in Vitro* and P388 leukemia *in Vivo*

compound No.	R_1	NR_2R_3	IC_{50} M	(Optimum Dose) /kg/
13	CH_3	$HN(CH_2)_2NH(CH_2)_2OH$	1.5E-07	100
14	$(CH_2)_2OH$	$HN(CH_2)_2NH(CH_2)_2OH$	7.8E-07	200
15	$(CH_2)_2OH$	$HN(CH_2)_2N\ Et_2$	7.3E-07	400
16	$CH_2CH_2NH_2$	$HN(CH_2)_2NH(CH_2)_2OH$	5.8E-07	12.5
17	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2NH_2$	1.6E-06	50
18	$(CH_2)_2NHCH_2OH$	$HN(CH_2)_2NH(CH_2)_2OH$	7.4E-07	6.25
19	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2N\ Me_2$	2.3E-07	12.5
20	$(CH_2)_2NH(CH_2)_2OH$	$HN(CH_2)_2N\ Et_2$	5.1E-07	50
21	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NH(CH_2)_2OH$	1.3E-07	12.5
22	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NH_2$	4.6E-08	12.5
23	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2NH\ Me$	7.4E-06	25
24	$(CH_2)_2N(Et)_2$	$HN(CH_2)_2N\ Et_2$	5.5E-07	100

These structures were initially optimized by using the molecular mechanics MMFF option in Arguslab-4. No electrostatic interactions were considered in these initial optimizations. The resulting optimized structures were then processed through the AM1, PM3 and MNDO semi empirical molecular orbital calculations furnished by Arguslab-4 package to obtain three sets of physiochemical parameters representing total energies (E_t), standard heat of formation (ΔH_f) and dipole moment (μ) as shown in tables (7, 8, and 9).

Table (7) total energies (E_t) of dihydroxy AP compounds

compound No.	$E_{t\ AM1}$	$E_{t\ PM3}$	$E_{t\ MNDO}$
13	108780.5	109456.9	109452.3
14	102630.1	107180.5	103246.2
15	154150.4	153711	151843.5
16	104964.9	106024.3	105673.5
17	140736.6	14103.19	140990.5
18	159407.4	160329.9	160300.2
19	181429.5	181441.6	182300
20	224049.3	223491.3	225056
21	161587.7	105365	162395.8
22	30187.84	130926.9	30834.54
23	144595.2	144306.2	145383.7
24	148872.1	149656.2	150005.7

Table (8): Dipole moment (μ) of dihydroxy AP

compound No.	μ_{AM1}	μ_{PM3}	μ_{MNDO}
13	1840.549	1843.423	1839.883
14	1822.94	1843.139	1822.868
15	2169.615	2168.73	2147.623
16	1816.075	1813.527	1815.859
17	1898.758	1898.914	1897.517

18	2184.357	2184.348	2184.18
19	2206.234	2206.555	2205.99
20	2557.184	2557.186	2556.986
21	2078.625	1428.713	2078.375
22	488.8187	1655.452	488.915
23	1789.562	1789.121	1789.349
24	1901.099	1873.759	1900.813

Table (9): Heat of formation H_f of desoxy AP compounds

compound No.	H_f AM1	H_f PM3	H_f MNDO
13	219777.3	210427.2	220723.7
14	H_f 224562.4	218325.6	225510.1
15	275907.3	265010.3	273923.8
16	224643.8	214538.8	225658.7
17	260415.5	250617.6	260975.7
18	290021.8	279019.1	291277.9
19	308281.9	296842	309491.1
20	3580751	345777.7	359453
21	295613.5	227651.5	296792.9
22	149691.4	239595.5	150636.1
23	267685.4	256417.8	268788.2
24	282722.5	272096.8	284219.1

The structures of the same set of desoxy AP compounds were also optimized using the molecular mechanics option in ChemSketch version 12. The optimization is based on modified molecular mechanics which take into account bond stretching, angle bending, and internal rotation and Van der Waals non-bonded interactions. The resulting optimized structures were then processed to calculate three sets of physiochemical parameters representing molar volume (MV), molecular refractivity (MR), polarizability and partition coefficient (log P) as shown in tables (10).

Table (10): MR, MV, polarizability & Log P of the dihydroxy AP

compound No.	MR	MV	Polarizability	Log P
13	96.25	237.4	38.15	1.24
14	101.90	251.6	40.39	0.46
15	110.79	290.8	43.92	2.55
16	102.96	251.1	40.81	0.51
17	102.96	251.1	40.81	0.51
18	109.56	272.3	43.43	0.20
19	113.85	295.5	45.13	1.24
20	123.07	327.6	48.78	2.30
21	123.07	327.6	48.78	2.30
22	111.86	290.3	44.34	2.60
23	117.42	313.4	46.55	2.84
24	131.96	366.8	52.31	4.39

The above data were used to find out correlations between the various biological data expressed in different ways and the physiochemical quantities, namely the total energy, dipole moment, heat of formation, polarizability, molar volume, molar refractivity and log P for both, the desoxy and dihydroxy anthrapyrazole compounds under study; statistically tested for significant correlations using the formula:

$$r = \frac{n(\Sigma xy)-(\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2)-(\Sigma x)^2}\ -\ \sqrt{n(\Sigma y^2)-(\Sigma y)^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ (2\text{-}1)$$

Where r is the linear correlation coefficient which takes the values +1 when there is perfect positive correlation, -1 when there is a perfect negative correlation and zero when there is no correlation, n is the number of data points in the paired data set and x and y refer to each item in the two column of data to be correlated. Except for few trial runs, the correlation coefficients and the regression equations were calculated using Minitab16 software (2009).

2.4 Regression analysis

Regression analysis was performed using Minitab16 software (2009) as follows; start, regression, regression and then the biological activity (dose, IC_{50} or – log IC_{50}) is put in a response space and the two of the calculated physiochemical properties (total energy , molar volume, molar refractivity, log P, dipole moment, heat of formation and polarizability) were put in the predictions space. In general, a QSAR model is acceptable when it has an $\boldsymbol{r^2}$ value greater than 0.6 and r^2 (*CV*) greater than 0.5. The results of this process; the regression equations and the $\boldsymbol{r^2}$ values for the desoxy and dihydroxy AP compounds were tabulated in appendixes (A & B).

The regression equations with high $\boldsymbol{r^2}$ values (> 0.6), were subjected to further studies to choose the best QSAR model via cross validation of the training and validation sets.