VERIZONA PUBLISHER
Publishing Your Ideas

www.verizonaonlinepublishing.com
Vol: 1, Issue: 3

Journal of Pharma and Pharmaceutical Sciences

Anti-pharmacophore Identification and Structure-Activity Relationships for the Inhibition of Cellular Protein and DNA Syntheses by a Series of Thiosemicarbazones and Thiosemicarbazides

Ahmet Altun[1,2*], Gamze Fındık[2]

[1]Department of Genetics and Bioengineering, Fatih University, Büyükçekmece, Istanbul 34500, Turkey

[2]Department of Physics, Fatih University, Büyükçekmece, Istanbul 34500, Turkey

***Corresponding author:** Ahmet Altun, Department of Genetics and Bioengineering, Fatih University, Büyükçekmece, Istanbul 34500, Turkey; Tel: +90-212-8663300; Fax: +90 212 8663402; E mail: aaltun@fatih.edu.tr

Article Type: Research, **Submission Date**: 11 November 2015, **Accepted Date:** 11 January 2016, **Published Date:** 23 March 2016.

Citation: Ahmet Altun and Gamze Fındık (2016) Anti-pharmacophore Identification and Structure-Activity Relationships for the Inhibition of Cellular Protein and DNA Syntheses by a Series of Thiosemicarbazones and Thiosemicarbazides. J.Pharm Pharm Scien 1(3): 13-20. doi: https://doi.org/10.24218/vjpps.2016.13.

Abstract

Background: Structure-activity relationship (SAR) of a series of thiosemicarbazone and thiosemicarbazide derivatives (57 compounds) that are potent inhibitors of herpes simplex virus (HSV) has been investigated for their inhibitory effects on cellular protein and DNA syntheses by electron-conformational method. SAR and QSAR studies were already carried out for these compounds in terms of their anti-HSV activity and dermal toxicity. The effect of a compound on cellular growth needs also to be known before its usage in chemotherapy. Therefore, this is a complementary study in designing new nontoxic antiviral drugs with reduced side effects on protein and DNA syntheses.

Methods: Electron-conformational method, which considers both electronic and geometrical structures of compounds, was applied. Computed geometrical and electron structural parameters of each atom and bond were arranged as a matrix for each of the 57 compounds. A submatrix that is common (absent) in the inhibitory (noninhibitory) compounds within some certain tolerances was searched to obtain the anti-pharmacophore responsible for the inhibition of cellular protein and DNA syntheses. A multivariable regression analysis was carried out to obtain a classical QSAR equation that relates the inhibitory concentrations with some physicochemical characteristics of the compounds.

Results: The molecular fragment associated to the revealed submatrix is formed by three charges situated at some specific distances. It is considered responsible for the inhibition of cellular protein and DNA syntheses (anti-pharmacophore). The realization of the identified anti-pharmacophore has been discussed on several compounds. The found QSAR equation has been validated on a test series and shown that it can be used in predicting the extent of the inhibitory effect of the untested compounds.

Conclusion: This study allows one to design new anti-herpes simplex virus agents with reduced side effects on protein and DNA syntheses.

Keywords: Drug design, Electron-conformational method (ECM), Inhibitory activity, Pharmacophore, SAR.

Introduction

Thiosemicarbazones of 2-acetylpyridine, 2-acetylquinoline, 1-acetylisoquinoline and related compounds were evaluated as inhibitors of herpes simplex virus types 1 and 2 (HSV-1 and HSV-2) in vitro and in a cutaneous herpes guinea pig model by Shipman and co-workers [1]. Structure–inhibitory and structure–noninhibitory activity relationships of these compounds against HSV-1 were previously examined by means of the electron-conformational method (ECM) with the aim to guide developing derivatives of existing drugs with enhanced potency and synthesizing new antiviral agents [2]. Since all antiviral agents cannot be used in chemotherapy due to their side effects, a complementary ECM study of these compounds was performed on structure–dermal toxicity/non-toxicity relations [3]. Before using such compounds in antiviral chemotherapy, it is essential to determine their effects on cellular protein and DNA syntheses. Shipman et al. [1] performed four biochemical in vitro tests to determine the effects of the substances on cellular growth: ^{3}H-amino acid incorporation into acid-precipitable material as well as Lowry assay measurements for the inhibitory concentrations of protein biosynthesis and ^{3}H-thymidine incorporation into acid-precipitable material as well as diphenylamine (DPA) assay measurements for the inhibitory concentrations of DNA synthesis. In this study, a substructure present only in the compounds that inhibit cellular protein and DNA syntheses has been identified in terms of geometric and electronic parameters by processing available experimental data [1] with ECM. The absence of this revealed fragment in a compound is then

one of the essential criteria to have minimal side on cellular growth during chemotherapy. A quantitative structure-activity relationship (QSAR) model has also been devised to assess the amount of inhibitory effects of related compounds on protein and DNA syntheses.

Materials and methods

Training Set

The investigated compounds are described through Table 1, Table 2, and Figure 1. Their 50% inhibitory concentrations taken from Ref. [1] are also given in Table 2. The correlation between the inhibitory concentrations of isotopic (incorporation of ^{3}H-amino acids = $C_{3H\text{-}aa}$) and nonisotopic (Lowry protein assay = C_{Lowry}) protein synthesis tests is

Table 1: Main skeleton types of the compounds investigated[a]

Skeleton Type	Molecular Formula
S_1	Py-C(R_2)=NNHC(S)R_1
S_2	Py-C(CH_3)=NNHC(R_1)R_2
S_3	*o*-CH_3-Py-C(CH_3)=NNHC(S)R_1
S_4	Qu1-C(CH_3)=NNHC(S)R_1
S_5	Qu1-CH(CH_3)-NHNHC(S)R_1
S_6	Qu2-C(CH_3)=NNHC(S)R_1
S_7	Qu2-CH(CH_3)-NHNHC(S)R_1
S_8	Py-CH(CH_3)-NHNHC(S)R_1

[a] See Figure 1 for the abbreviations Py, Qu1, and Qu2

$$C_{3H\text{-}aa} = 1.083 \cdot C_{Lowry} \quad (1)$$

($R^2 = 0.99$; $S = 9.43$; $MAD = 4.10$)

where S is the standard deviation of the estimate, and MAD is the mean absolute deviation of the estimate. Since MAD is the average of the absolute values of individual errors, it does not reflect the sign of the error. Therefore, S that is derived using squares of individual errors is a better measure of the error of the estimate [4]. The smaller the S and MAD values, the better the model constructed. R^2 is the ratio of the sum of the squares of the deviation of the estimated values from the average value to that of the measured values from the average value [4]. Therefore, R^2 that lies between 0 and 1 is a measure of the error relative to the average value. The closer the R^2 gets to unity, the better the constructed model becomes predictive.

The corresponding correlation for the inhibitory concentrations of the isotopic (incorporation of ^{3}H-thymidine = $C_{3H\text{-}thy}$) and nonisotopic (DPA assay = C_{DPA}) DNA synthesis tests is

$$C_{3H\text{-}thy} = 0.935 \cdot C_{DPA} \quad (2)$$

($R^2 = 0.99$; $S = 10.06$; $MAD = 4.09$)

Besides, the correlations between the two isotopic and the two nonisotopic tests are as follows:

$$C_{3H\text{-}aa} = 0.906 \cdot C_{3H\text{-}thy} \quad (3)$$

($R^2 = 0.99$; $S = 9.06$; $MAD = 4.10$; **N48** and **N57** excluded)

$$C_{DPA} = 1.180 \cdot C_{Lowry} \quad (4)$$

($R^2 = 0.99$; $S = 10.50$; $MAD = 5.00$; **N48** excluded)

Therefore, the correlations between these four different tests are in general satisfactory. However, in DNA synthesis tests, $C_{3H\text{-}thy}$ is considerably different than C_{DPA} for **N48** to **N51** (see Table 2). This disparity arises from the inhibition of phosphorylation of ^{3}H-thymidine rather than inhibition of DNA synthesis [1]. Therefore, the compounds **N48** to **N51** are noninhibitors of both protein and DNA syntheses. **N52** is a noninhibitor of protein synthesis but it has rather small inhibitory concentrations for DNA synthesis ($C_{3H\text{-}thy}$ = 0.95 μg/ml; C_{DPA} = 4.4 μg/ml). For this compound, the smallness of $C_{3H\text{-}thy}$ than C_{DPA} should arise from the inhibition of phosphorylation of ^{3}H-thymidine. **N52** was initially considered as a noninhibitor of both protein and DNA syntheses but this supposition was later tested during the present ECM study. The rest of compounds have been classified as the inhibitors of protein (DNA) synthesis when the inhibitory concentrations are less than 10 (7) μg/ml and as the noninhibitors when the concentrations are 10 (7) μg/ml or more. With these initial settings, the class of inhibitors (**N1** to **N43**) or noninhibitors (**N44** to **N52**) of protein and DNA syntheses in the training set consists of the same molecules (see Table 2). The compounds **N53** to **N57** has been considered as a test set. **N53** to **N56** are inhibitors of both protein and DNA syntheses while **N57** is a noninhibitor of both types of cellular syntheses.

Electron-Conformational Method

The identification of pharmacophore, a group of atoms in a specific geometric arrangement that is considered responsible for a bioactivity demonstration, is one of the most important problems in molecular modelling for drug design and screening. Electron-conformational method (ECM) [5-8], known also as electron-topological method (ETM) [2,3,9-12], was developed for pharmacophore identification and pharmacophore-based bioactivity prediction. Its objects are two (active and inactive) or more bioactivity sets of a series of molecules. Its computational part is a consequence of the following steps: (1) conformational analyses; (2) quantum-chemical calculations; (3) electron-conformational matrix of contiguity (ECMC) formation from the data obtained in the first two steps; (4) ECMC processing to select pharmacophore. The first two steps are traditional ones and the others implement the core of ECM.

In ECM, each molecule in view is described by a set of parameters arranged as a matrix (ECMC), being symmetric with respect to the diagonal elements. The diagonal elements of ECMC are chosen among the atomic parameters, such as atomic charges, valence activities, atomic polarizabilities, atomic contributions to HOMO or LUMO energies, etc. The off-diagonal elements of ECMC are one of the electronic bonding parameters for chemically bonded pairs of atoms, such as Wiberg index, bond order, bond energy (total, covalent or ionic), bond polarizability, etc., and are interatomic distances for the nonbonded pairs of atoms.

After forming ECMCs of all compounds with a unique atomic and bond parameters that are deemed most important for the activity demonstration, the submatrix that is present in all active compounds but absent in all inactive ones is searched within some tolerance values specific to each type of parameters. During this search process, each submatrix of a template ECMC that belongs to one of the active compounds is compared with

Citation: Ahmet Altun and Gamze Fındık (2016) Anti-pharmacophore Identification and Structure-Activity Relationships for the Inhibition of Cellular Protein and DNA Syntheses by a Series of Thiosemicarbazones and Thiosemicarbazides. J.Pharm Pharm Scien 1(3): 13-20. doi: https://doi.org/10.24218/vjpps.2016.13.

Table 2: The investigated molecules,[a] their experimental inhibitory concentrations to the cellular protein and DNA syntheses [1], and their presently calculated physicochemical parameters

				50% inhibitory concentration (μg/ml)									
N	**Skeleton**	**R_1**	**R_2**	**Protein Synthesis**		**DNA Synthesis**		**DM**	**E_{HYD}**	**E_{LUMO}**	**A_s**	**R**	**P**
				$C_{3H\text{-}aa}$	**C_{Lowry}**	**$C_{3H\text{-}thy}$**	**C_{DPA}**	**(Debye)**	**(kcal/mol)**	**(eV)**	**(Å²)**	**(Å³)**	**(Å³)**
		Training Set (Inhibitory Molecules)											
1	S_1	NH[CH(CH$_3$)$_2$-CH$_2$-CH(CH$_3$)$_3$]	CH_3	0.12	0.29	0.23	0.23	10.8	-4.6	2.0	575	96	36.8
2	S_1	NH[CH_2-C_6H_5]	CH_3	3.20	3.20	1.50	2.90	10.9	-9.4	1.9	542	92	33.6
3	S_1	NH[*p*(CF_3)-C_6H_4]	CH_3	0.60	5.00	1.40	4.30	13.0	-9.4	1.7	556	94	33.3
4	S_1	N(CH_3)$_2$	CH_3	2.00	2.10	0.24	1.80	10.6	-5.5	1.9	432	69	25.8
5	S_1	N(CH_3)(C_6H_{11})	CH_3	0.25	0.29	0.32	0.33	10.3	-3.8	2.0	540	90	34.2
6	S_1	*m*(CH_2OH)-X_1	CH_3	3.10	2.50	2.00	4.70	11.4	-6.8	2.0	510	87	33.0
7	S_1	X_2	CH_3	0.60	0.60	0.50	1.10	9.9	-5.7	1.9	479	84	31.7
8	S_1	X_9	CH_3	0.19	0.13	0.73	0.45	9.4	-7.3	2.0	586	107	39.0
9	S_1	X_4	CH_3	1.70	1.10	0.40	0.80	10.7	-3.9	1.9	501	86	32.3
10	S_1	X_5	CH_3	1.00	0.80	0.75	0.99	10.5	-3.8	1.9	529	93	35.2
11	S_1	X_{11}	CH_3	0.05	0.13	0.12	0.06	10.6	-2.3	1.8	616	113	43.4
12	S_1	N(CH_3)$_2$	C_2H_5	1.10	1.50	0.34	1.00	10.7	-4.7	2.0	457	74	27.6
13	S_1	X_5	C_2H_5	0.38	0.39	0.06	0.32	10.6	-2.8	1.9	545	98	37.1
14	S_1	NH(C_3H_5)	C_3H_7	1.30	0.75	0.16	2.40	11.0	-8.2	1.9	528	82	31.1
15	S_1	X_9	$CH(CH_3)_2$	0.08	0.18	0.10	0.16	9.4	-5.8	2.0	634	116	42.7
16	S_3	X_9	-	0.52	2.40	0.35	0.28	11.6	-6.2	2.0	589	111	40.8
17	S_4	NH-C_2H_5	-	1.80	0.80	1.20	1.40	11.1	-7.9	1.7	514	88	32.0
18	S_4	N(*i*-C_4H_9)$_2$	-	4.70	5.80	3.20	4.40	10.7	-2.3	2.2	604	116	43.0
19	S_4	X_9	-	2.60	4.50	0.07	0.35	11.9	-7.5	1.5	634	125	45.2
20	S_4	X_4	-	1.50	0.50	1.50	1.40	10.8	-4.1	1.8	557	105	38.5
21	S_5	N(CH_3)$_2$	-	0.03	0.01	0.04	0.03	8.0	-6.1	1.7	489	89	32.4
22	S_5	*p*CH_3-X_1	-	2.70	4.90	0.13	0.44	7.2	-4.9	2.5	559	105	38.9
23	S_5	X_9	-	3.70	4.10	2.20	3.10	5.9	-8.5	2.5	628	125	45.6
24	S_5	X_5	-	0.92	1.34	0.29	0.98	8.2	-5.0	2.5	584	113	41.8
25	S_6	NH-C_3H_7(n)	-	0.76	1.10	1.00	1.60	10.8	-7.5	2.1	549	91	33.8
26	S_7	$NHCH_2$-C_6H_5	-	8.80	2.50	3.30	6.00	7.5	-10.5	2.4	597	111	40.2
27	S_7	X_3	-	5.70	6.50	1.10	0.89	7.6	-7.1	2.4	635	129	46.8
28	S_8	X_7	-	0.08	0.39	0.19	4.02	8.6	-5.8	3.1	469	77	29.1
29	S_2	NH-C_6H_5	Se	1.10	1.20	0.36	2.00	8.5	-9.0	2.5	514	79	30.3
30	S_1	NHNHC(S)N(CH_3)(C_6H_5)	-	0.55	0.81	0.34	0.37	3.7	-13.5	1.8	612	113	41.4
31	S1	NH[CH_2-*o*(CH_3)-C_6H_4]	CH_3	<1	<1	0.70	<1	10.9	-8.2	2.0	558	97	35.4
32	S1	NH(C_3H_5)	C_2H_5	<0.32	<0.32	<0.32	<0.32	11.0	-8.8	2.0	501	78	29.3
33	S1	X9	C_2H_5	<0.32	<0.32	<0.32	<0.32	9.6	-6.5	1.9	613	111	40.8
34	S1	X5	$CH(CH_3)_2$	<0.32	<0.32	<0.32	<0.32	10.7	-2.3	2.5	565	103	38.9
35	S3	NH(C_3H_5)	-	<0.32	<0.32	<0.32	<0.32	11.0	-8.0	1.7	493	79	29.3
36	S3	X5	-	0.1	<0.1	0.14	0.22	10.5	-2.6	2.1	534	99	37.1
37	S4	NH-C_3H_7(n)	-	<1	<1	<1	<1	10.9	-7.2	1.7	542	92	33.8
38	S4	N(CH_3)-C_6H_4-*p*(OCH_3)	-	<1	<1	<1	<1	11.6	-8.7	1.8	625	119	42.3
39	S5	N(CH_3)-C_6H_{11}	-	<0.1	0.36	0.09	0.61	7.0	-4.7	2.5	593	110	40.8
40	S5	X7	-	<0.1	0.04	0.06	0.17	7.4	-5.7	2.5	521	96	35.3
41	S5	X10	-	<0.1	<0.1	0.07	0.15	6.7	-6.3	2.5	563	107	39.6
42	S7	X4	-	0.0	<0.1	0.11	0.17	8.6	-4.8	2.4	562	104	39.0
43	S8	X4	-	<0.32	<0.32	<0.32	<0.32	8.0	-5.0	3.1	503	86	32.8

				50% inhibitory concentration (μg/ml)									
N	**Skeleton**	**R_1**	**R_2**	**Protein Synthesis**		**DNA Synthesis**		**DM**	**E_{HYD}**	**E_{LUMO}**	**A_s**	**R**	**P**
				$C_{3H\text{-}aa}$	**C_{Lowry}**	**$C_{3H\text{-}thy}$**	**C_{DPA}**	**(Debye)**	**(kcal/mol)**	**(eV)**	**(Å²)**	**(Å³)**	**(Å³)**
		Training Set (Noninhibitory Molecules)											
44	S1	$NH[CH_2\text{-}[CH(OH)]_4\text{-}CH_2OH]$	CH_3	28	22	7	8	9.3	-26.1	2.0	648	99	38.1
45	S2	NH_2	O	515	500	570	600	7.4	-10.3	2.2	302	51	19.0
46	Py	$C(CH_3)=NNH_2$		22	16	7.3	24	2.8	-7.7	2.8	318	44	15.7
47	S2	$NH\text{-}C_6H_5$	O	32	16	15	33	13	-8.9	1.4	497	80	28.6
48	S_1	NH_2	CH_3	348	284	2.00	58	10.9	-12.6	1.9	394	59	22.1
49	S_1	$NH(CH_3)$	CH_3	32	22	0.90	10	11.0	-9.1	1.9	421	64	23.9
50	S1	NH[CH2-*o*(C_5H_4N)]	CH_3	27	39	1.70	14	10.8	-10.4	1.9	541	89	32.9
51	S7	X6	-	0.6	0.68	1.00	19	6.9	-7.7	2.2	640	128	46.3
52	S_1	$NH(C_3H_5)$	CH_3	27	16	0.95	4.4	11	-9.7	1.9	477	73	27.4
		Test Set											
53	S_1	NH[*m*(F)-C_6H_4]	CH_3	2.00	1.30	1.40	2.10	12.5	-9.7	1.8	515	89	31.7
54	S_1	X_{12}	CH_3	0.20	0.17	0.27	0.32	10.5	-3.0	2.0	530	95	36.0
55	S_1	$NH(C_3H_5)$	$CH(CH_3)_2$	4.30	3.70	1.80	2.70	10.9	-8.0	2.1	522	82	31.1
56	S_2	X_5	Se	0.28	0.20	0.23	0.24	11.3	-2.3	1.9	535	83	33.8
57	Py	$C(CH_3)=NN(CH_3)C(S)N(CH_3)_2$	-	100	67	8	33	9	-3.0	2.7	459	74	27.6

[a] See Figure 1 and Table 1 for the abbreviations used in describing the molecules

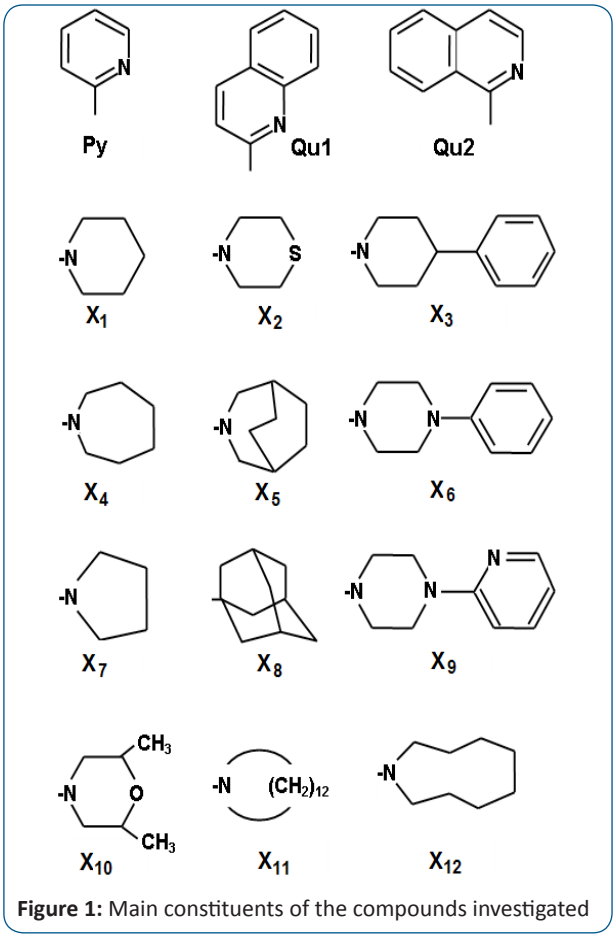


Figure 1: Main constituents of the compounds investigated

each submatrix of ECMCs that belong to the rest of compounds (see below). The submatrix revealed is considered responsible for activity demonstration, and it is called as electron-conformational submatrix of contiguity (ECSC) or as pharmacophore.

The atom-in-molecule electronic characteristics in different compounds may be analogous for different types of atoms and be different for the same atomic type. ECM can easily treat such cases since it compares parameters forming ECMC rather than the atom types. For the same reason, the compounds whose skeletons are quite different to each other can also be processed simultaneously in ECM.

ECMC Formation

Conformational analysis and quantum-chemical calculation parts of the presently considered molecules have been already published [2]. The results of these previous computations [2] were used in forming ECMCs of all investigated molecules. The diagonal elements of ECMCs were chosen as effective charges on atoms (Q_i) in electron charge unit *e* while their off-diagonal elements were taken as Wiberg indices (W_{ij}) that measure the bond order for chemically bonded atoms, and as interatomic distances (R_{ij}) in Å for chemically nonbonded atoms.

Throughout this study, the atoms entering the submatrix ECSC that is considered responsible for the bioactivity demonstration were circled on the molecule figures. Since inhibition of protein and DNA syntheses has a negative selectivity in drug design, the revealed ECSC of this study is called as anti-pharmacophore (*A-Pha*) in the remaining. To be able to separate the type of matrix elements, the charges on atoms and Wiberg indices were shown on ECMC and ECSC as italic and bold, respectively. All stable conformers of the present molecules were processed in the present ECM study. However, only the conformer containing *A-Pha* was accepted as inhibitory.

Results and Discussion

Selection and Refinement of Anti-pharmacophore

To determine the inhibitory activity feature of both protein and DNA syntheses in the form of matrix elements being common in the ECMCs of all inhibitory compounds and absent in the ECMCs of all noninhibitory ones, ECMC of each inhibitory compound was taken once as a template.

Each template ECMC was compared with all the other ECMCs within some tolerance values to obtain a representative ECSC submatrix for inhibitory compounds. With each selected template compound, the found best representative ECSC belongs to three chemically nonbonded atoms. The revealed ECSC matrix elements are enclosed by solid lines in Figure 2 on the ECMC of one of the inhibitory conformers of **N2**. Since the comparisons of ECMCs are performed without noting the atom types, the three atoms entering the ECSC are shown with the labels x, y and z rather than atom types in the following.

Only three tolerance parameters, one of which for diagonal elements (Δ_1), one for bond parameters (Δ_2), and one for interatomic distances (Δ_3), are used in the feature searching process since it is impractical to use different tolerances for all $n \cdot (n+1)/2$ matrix elements, where n is the dimension of the ECMC, or equivalently, the number of atoms in the template compound. The ECSC of **N2** given in Figure 2 represents the inhibitory activity best when Δ_1 and Δ_3 are ±0.09 and ±0.55, respectively. As the atoms entering the ECSC are chemically nonbonded, Δ_2 has no effect on the realization of the feature.

The ECSC found on **N2** enters all inhibitory agents except **N4**, **N8**, and **N28** within the given tolerances above. Any of the noninhibitory compounds, including **N52**, does not include the ECSC within the above tolerances. Noting that **N52** is noninhibitory for protein synthesis but inhibitory for DNA synthesis according to the experimental tests, we searched a representative ECSC for inhibitory compounds that also includes **N52**. However, when the chosen tolerance parameters are arranged to include **N52**, the majority of the noninhibitory compounds already includes the found ECSC. Therefore, ECM classifies **N52** uniquely as noninhibitory for both protein and DNA syntheses. As partly discussed above, appearance of **N52** as an inhibitor of DNA synthesis should arise from the technical limitations of the tests of DNA synthesis: inhibition of phosphorylation of ^{3}H-thymidine in the isotopic test, and slight underestimation of inhibitory concentration in nonisotopic colorimetric test (see below for more details about the quality of tests).

The parameters entering *A-Pha* are representative of the considered bioactivity in some certain ranges. In most cases, the atom-in-molecule characteristics in the ECSCs of template compounds are not at the middle of these ranges. In this study, each parameter of ECSC shown in Figure 2, and, at each step the accompanying tolerance were both scanned with successive increments of ±0.01 so that both midpoints of the validity ranges and flexibility tolerances can be determined more accurately and

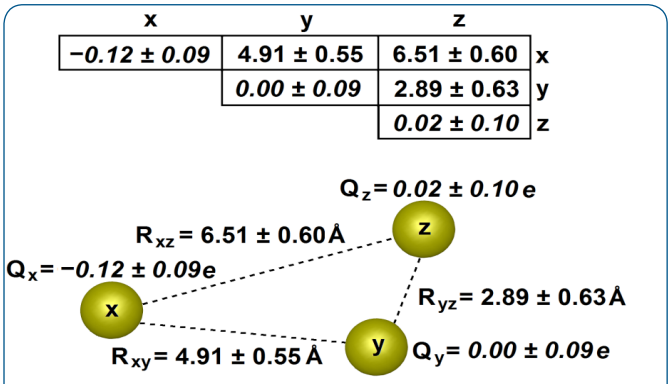
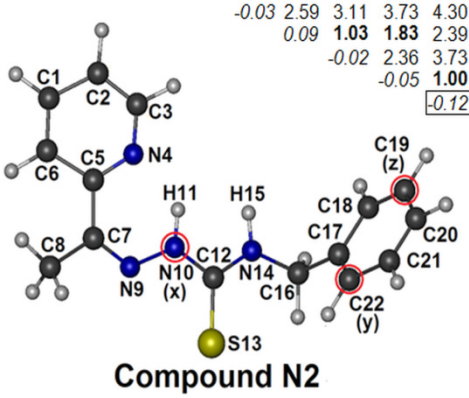


Figure 3: The anti-pharmacophore (*A-Pha*) of the protein and DNA syntheses

C1	C2	C3	N4	C5	C6	C7	C8	N9	N10 (x)	H11	C12	S13	N14	H15	C16	C17	C18	C19 (z)	C20	C21	C22 (y)	
0.04	**1.25**	2.43	2.74	2.43	**1.66**	3.79	4.42	4.83	5.12	4.55	6.43	7.66	6.87	6.31	8.30	8.90	9.24	10.02	10.55	10.28	9.43	C1
-0.03	**1.60**	2.29	2.72	2.41	4.21	5.19	4.99	4.90	4.12	6.11	7.54	6.28	5.58	7.65	8.07	8.23	8.92	9.50	9.38	8.65	C2
0.12	**1.32**	2.26	2.74	3.61	4.84	4.13	3.81	2.98	4.94	6.43	5.06	4.36	6.42	6.77	6.98	7.67	8.20	8.04	7.32	C3
-0.15	**1.52**	2.38	2.40	3.73	2.86	2.71	2.03	3.95	5.31	4.34	3.83	5.78	6.26	6.73	7.52	7.93	7.57	6.71	N4
0.11	**1.23**	**1.04**	2.62	2.43	2.89	2.56	4.24	5.32	4.94	4.61	6.40	7.06	7.70	8.56	8.91	8.42	7.45	C5
-0.03	2.59	3.11	3.73	4.30	3.93	5.63	6.67	6.30	5.90	7.75	8.45	9.00	9.86	10.28	9.86	8.91	C6
0.09	**1.03**	**1.83**	2.39	2.61	3.63	4.30	4.67	4.64	6.07	6.84	7.74	8.65	8.85	8.16	7.08	C7
-0.02	2.36	3.73	4.09	4.82	5.07	5.99	6.07	7.32	8.12	9.11	10.00	10.10	9.30	8.23	C8
-0.05	**1.00**	2.06	2.46	3.03	3.64	3.80	4.96	5.80	6.82	7.78	7.96	7.19	6.01	N9
-0.12	**0.95**	**1.27**	2.64	2.28	2.38	3.68	4.51	5.45	6.46	6.77	6.11	4.91	N10 (x)
0.10	1.98	3.51	2.39	2.09	3.85	4.55	5.25	6.23	6.67	6.19	5.09	H11
0.31	**1.18**	**1.26**	1.98	2.51	3.49	4.62	5.70	5.97	5.23	3.88	C12
-0.45	2.69	3.54	3.18	4.16	5.54	6.56	6.59	5.59	4.21	S13
-0.17	**0.95**	**0.98**	2.42	3.38	4.54	5.04	4.52	3.20	N14
0.11	2.08	2.73	3.29	4.41	5.08	4.77	3.63	H15
0.12	**1.02**	2.60	3.83	4.35	3.86	2.48	C16
0.01	**1.19**	2.46	2.84	2.46	**1.65**	C17
-0.01	**1.71**	2.45	2.82	2.43	C18
0.01	**1.21**	2.44	2.82	C19 (z)
0.00	**1.70**	2.45	C20
0.01	**1.22**	C21
0.00	C22 (y)



Compound N2

Figure 2: ECMC of the one of the inhibitory conformers of compound **N2** (due to its symmetry, only the upper triangle is shown; the matrix elements representing hydrogens bonded to carbon atoms are omitted for simplicity; the diagonal elements, which are shown as italic, are charges on atoms whereas the off-diagonal elements are interatomic distances for chemically nonbonded atoms and Wiberg indices for bonded atoms; Wiberg indices are demonstrated as bold; the ECMC elements shown with rectangular surroundings constitute the ECSC responsible for the inhibition of cellular protein and DNA syntheses; the atoms entering ECSC are circled on compound **N2**)

free from the chosen template. The submatrix refined in this way (see Figure 3) enters all inhibitory agents and is not seen any of the noninhibitors (**N52** is again classified as noninhibitory). In other words, *the feature given in Figure 3 achieves the complete separation of the inhibitory compounds from the noninhibitory agents*. As mentioned above, the inhibitory compounds **N4**, **N8**, and **N28** do not contain the feature before refinement. Slight modifications in some parameters shown in Figure 2 and in accompanying tolerances cause the inclusion of the feature by these compounds. As an example, just adjusting the charge of z being 0.01 ± 0.09 as 0.02 ± 0.10 results in the inclusion of the feature by **N16**. Further growth of any tolerance values shown in Figure 3 causes noninhibitory compounds to include the feature. If it is decreased, the number of inhibitors including the feature decreases, too.

When the ECMCs of noninhibitory compounds are taken as templates, there appears no common ECSC comprising all noninhibitors without covering inhibitory compounds. Therefore, the absence of *A-Pha* given in Figure 3 is the only ECM criterion for a drug to have no side effect on cellular growth for the present set of molecules. The revealed *A-Pha* is present in the inhibitory compounds of the test set (**N53** to **N56**) while it is absent in the noninhibitory compound of the test set (**N57**). This validates the applicability of the search of *A-Pha* in a compound to determine whether it is noninhibitory or not.

Realization of Anti-pharmacophore

Some examples are given here to realize *A-Pha*. The noninhibitory compound **N50** differs from **N2** with the presence of nitrogen atom instead of $C^{22}H$ moiety that corresponds to the y atom in ECSC (see Figure 2). Since the charge on this nitrogen is out of the limits allowable for the y atom of the *A-Pha*, **N50** is noninhibitory, emphasizing the importance of electronic structure in addition to geometrical alignment.

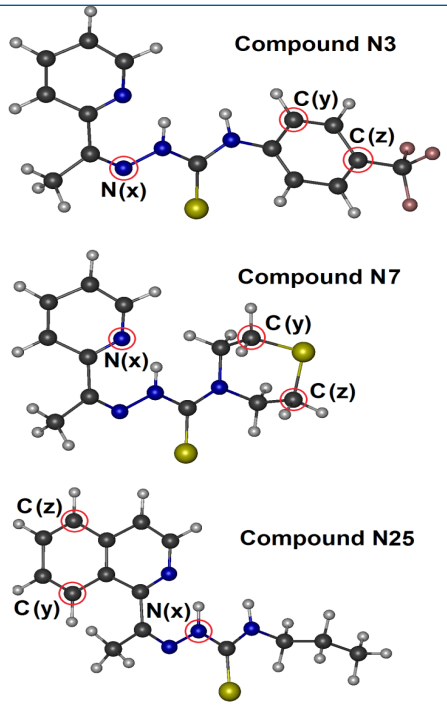


Figure 4: Demonstration of *A-Pha* on N3, N7 and N25

As the training set molecules have various skeleton types and various substituents attached to them, the appearance of *A-Pha* is expected on different geometrical fragments of the molecules [2,3-5-12], as shown in Figure 4 on **N3**, **N7**, and **N25**. Thus, the attempts to inspect the presence or absence of *A-Pha* in a compound simply from molecular formula without forming ECMC may be misleading.

If the CF_3 moiety of **N3** (see Figure 4) is substracted and its sulfur atom is replaced with oxygen, the noninhibitory compound **N47** is obtained. In **N47**, the charge on the x atom (the circled nitrogen atom of **N3** in Figure 4) falls outside the limit allowed for *A-Pha* due mainly to the replacement of S with O. Thus, **N47** becomes noninhibitory.

N52, being a noninhibitor of protein synthesis, differs from the inhibitory compounds **N32** and **N35** with the replacement of a methyl group with an ethyl group (**N32**) and addition of a methyl group at the position 6 of the pyridine (**N35**), respectively (see Figure 5). Since **N52** does not contain *A-Pha*, it can be inferred that *A-Pha* in **N32** and **N35** must be realized via these ethyl and methyl groups, as confirmed searching *A-Pha* within their ECMCs.

When a thiosemicarbazone derivative is converted to a thiosemicarbazide, in other words, when the azomethine bond is reduced, the atoms forming *A-Pha* change in some molecules (see **N9** and **N43** in Figure 6). However, *A-Pha* is realized with the same structural fragment in some others (see **N19** and **N23** in Figure 6).

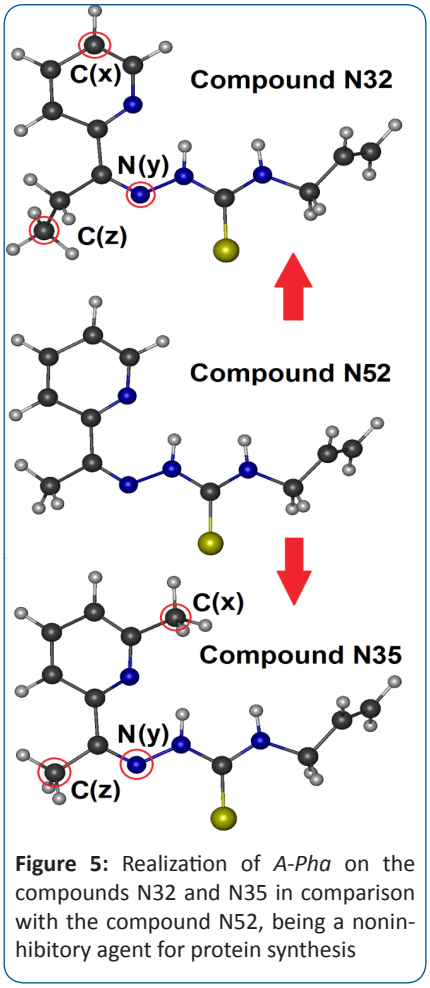


Figure 5: Realization of *A-Pha* on the compounds N32 and N35 in comparison with the compound N52, being a noninhibitory agent for protein synthesis

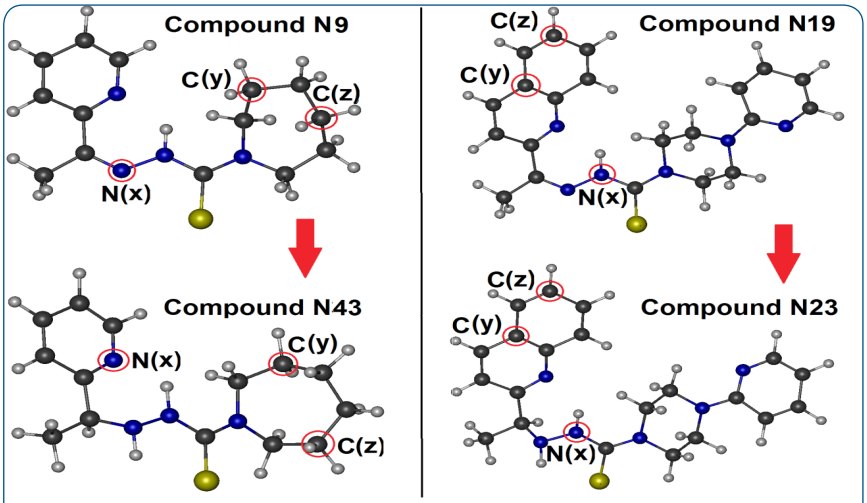


Figure 6: The effect of the reduction of azomethine bond on the realization of *A-Pha*

Quantification of the Inhibitory Effect

Multivariable regression analyses were performed for predicting the extent of inhibitory effects of thiosemicarbazones and thiosemicarbazides on cellular growth numerically. Among many calculated physicochemical parameters, only those that are statistically significant at the 99% or higher confidence level (i.e., $p \leq 0.01$) are included in the correlation equation. The parameters found significant (see Table 2) are dipole moment (*DM*), energy of lowest unoccupied molecular orbital (E_{LUMO}), hydration energy (E_{HYD}), surface area (A_S), refractivity (R), polarizability (P), and $\delta_{A\text{-}Pha}$ (taken 1 and 0 in the presence and absence of *A-Pha* in a molecule, respectively). E_{HYD} and A_S values were obtained by grid-based calculations on solvent-accessible surface and by taking the solvent probe radius as 1.4 Å. *DM* and E_{LUMO} values of the compounds were calculated as described in Ref. [2] while E_{HYD}, A_S, R, and P were calculated by using HyperChem program package [13]. The inhibitory compounds **N31** to **N43** of the training set could not be included in the data set used in deriving linear regression equation since their exact experimental inhibitory concentrations are not available (see Table 2). Therefore, the initial data set includes inhibitory compounds **N1** to **N30** and noninhibitory compounds **N44** to **N52**.

At most the R^2 values derived for the equations in estimating log of inhibitory concentrations measured with isotopic tests (log $C_{3H\text{-}aa}$ and log $C_{3H\text{-}thy}$) and with nonisotopic tests (log C_{Lowry} and log C_{DPA}) are 0.6 and 0.8, respectively. Therefore, attempts to derive correlation equations with high statistical significance for estimating log of each inhibitory concentration were not so successful. The estimates of log of the sum of inhibitory concentrations measured with isotopic tests [log ($C_{3H\text{-}aa}$ + $C_{3H\text{-}thy}$)] and with nonisotopic tests [log (C_{Lowry} + C_{DPA})] have at most R^2 of 0.62 and 0.94, respectively. It may then be concluded that the errors in measuring the inhibitory concentrations with the isotopic experimental tests are much larger and less systematic than those with the nonisotopic experimental tests. It thus seems that the nonisotopic biological tests (Lowry protein assay and DPA assay) are more accurate and the sum of the concentrations obtained in these significantly correlated tests in our data set (see above) benefit from error cancellations. Actually, systematic methodological assessment studies find that stable isotopic experimental methods measure rates of DNA synthesis far better than the ^{3}H-thymidine incorporation method since ^{3}H-thymidine perturbs the DNA synthesis and cell replication [14]. It has also been shown for protein assay techniques that methodical variations can potentially translate into differential protein expression patterns, which can be falsely taken to be biologically significant[15].

In terms of the above considerations, log (C_{Lowry} + C_{DPA}) abbreviated as log φ can be taken as a therapeutic index. Since a compound with inhibitory concentrations less than 10 µg/ml and 7 µg/ml in the tests of protein and DNA syntheses is inhibitory, the decisive value of this index to categorize the compounds as inhibitory or noninhibitory is log (10 + 7) = 1.23. If log φ of a compound is obtained equal to or more than 1.23, it is a noninhibitor of cellular protein and DNA syntheses; otherwise, it is inhibitory. The larger the log φ value, the better the compound for chemotherapy in terms of its effect on cellular growth. The best linear regression equation found for predicting log φ by

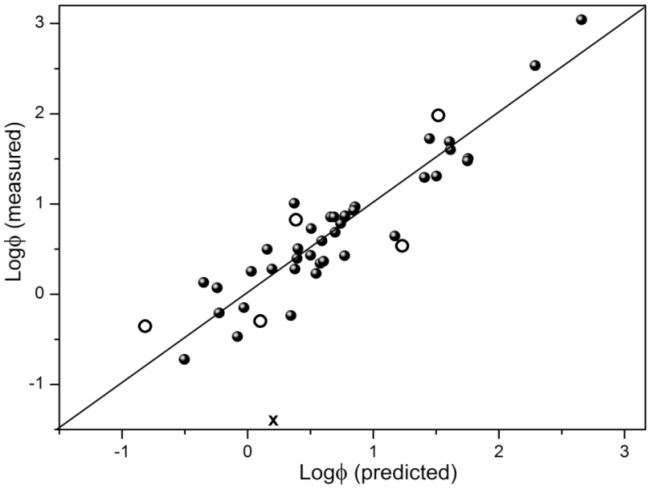


Figure 7: Plot of observed versus predicted values for the log of total inhibitory concentration ϕ. Filled and unfilled circles belong to the molecules in training and test series, respectively. Cross sign belongs to N21, which is excluded from the training set in deriving the regression equation

excluding **N21** is as follows:

$$\log \varphi = -0.847 \cdot \delta_{A\text{-}Pha} + 0.179 \cdot DM + 0.101 \cdot R - 0.140 \cdot P - 0.112 \cdot E_{HYD} - 0.014 \cdot A_S + 0.946 \cdot E_{LUMO} \quad (5)$$

$(R^2 = 0.94; S = 0.30; MAD = 0.21)$

Although the estimated log φ value of **N21** is significantly higher than the measured value, the regression equation classifies **N21** properly as inhibitory (see crossed data point in Figure 7). For all other molecules in the training set (filled circles in Figure 7), the estimated log φ values are reasonably close to the measured values. The derived regression equation estimates log φ values of the test compounds within its error bar (see unfilled circles in Figure 7). Therefore, the present regression equation is predictive in determining the extent of cellular inhibitory effects of analogous thiosemicarbazones and thiosemicarbazides.

As expected, presence of *A-Pha* in a molecule increases the inhibitory potency. Similarly, the increase in solvent-accessible surface area A_S and polarizability P increases the inhibitory potency, too. Since hydration energy E_{HYD} is a measure of the stability of a compound in a solvent environment, the stability increase of a compound causes to the lowering of its inhibitory potency. Analogously, the increase in the LUMO energy E_{LUMO} (or, equivalently the decrease in electron affinity), dipole moment DM, and refractivity R lowers the inhibitory potency.

Conclusion

Previously, a series of thiosemicarbazone and thiosemicarbazide derivatives demonstrating inhibitory activity against HSV-1 was investigated by means of ECM [2,3]. Two features responsible for inhibiting HSV-1 and three features that abolish the inhibition of HSV-1 were revealed as submatrices of ECMCs (i.e., ECSCs) of some compounds [2]. Moreover, three features of toxicity and three features of non-toxicity were determined in the form of ECSCs [3]. All of these features can be used when planning to develop and synthesize new non-toxic inhibitors of HSV-1. However, any potential antiviral drug needs to be studied in relation to its inhibitory effect on the protein and DNA syntheses. As a continuation of the previous structure-activity relationship studies on some thiosemicarbazone and thiosemicarbazide derivatives [2,3], the present ECM study finds the molecular fragment responsible for the inhibiton of protein and DNA syntheses in the form of ECSC as three charges placed at some specific distances. An agent that inhibits the protein or DNA synthesis has a negative selectivity in chemotherapy. Thus, it can be concluded that one of the criterions for a more effectual screening of new antiviral drugs is the absence of this ECM in a compound. Physicochemical parameters that affect the inhibitory concentrations of cellular protein and DNA syntheses have been determined. A QSAR equation has been devised in terms of these parameters for predicting the extent of the inhibitory concentration in a compound. The effectiveness of this QSAR equation has also been validated on the molecules outside the training set.

Acknowledgment

This work is supported by the Scientific Research Fund of Fatih University under the project number of P50011102_Y (1679).

References

1. Shipman C, Smith SH, Drach JC, Klayman DL. Thiosemicarbazones of 2-acetylpyridine, 2-acetylquinoline, 1-acetylisoquinoline, and related compounds as inhibitors of herpes simplex virus in vitro and in a cutaneous herpes guinea pig model. Antiviral Res. 1986; 6(4): 197-222.
2. Altun A, Kumru M, Dimoglo A. Study of electronic and structural features of thiosemicarbazone and thiosemicarbazide derivatives demonstrating anti-HSV-1 activity. J. Mol. Struct. (Theochem). 2001; 535(1-3): 235-246. doi:10.1016/S0166-1280(00)00582-0.
3. Altun A, Kumru M, Dimoglo A. The role of conformational and electronic parameters of thiosemicarbazone and thiosemicarbazide derivatives for their dermal toxicity. J. Mol. Struct. (Theochem). 2001; 572(1-3): 121-134. doi:10.1016/S0166-1280(01)00576-0.
4. Makridakis S, Hibon M. Evaluating accuracy (or error) measures. INSEAD Working Paper Series, 95/18/TM, Fontainebleau, France; 1995.
5. Bersuker IB, Bahceci S, Boggs JE, Pearlman RS. An electron-conformational method of identification of pharmacophore and anti-pharmacophore shielding: application to rice blast activity. J Comput Aided Mol Des. 1999; 13(4): 419-434.
6. Bersuker IB, Bahceci S, Boggs JE. Improved electron-conformational method of pharmacophore identification and bioactivity prediction. Application to angiotensin converting enzyme inhibitors. J Chem Inf Comput Sci. 2000; 40(6): 1363-1376.
7. Altun A, Golcuk K, Kumru M, Jalbout AF. Electron-conformational study for the structure-hallucinogenic activity relationships of phenylalkylamines. Bioorg Med Chem. 2003; 11(18): 3861-3868.
8. Mert S, Kasımoğulları R, İça T, Çolak F, Altun A, Ok S. Synthesis, structure-activity relationships, and in vitro antibacterial and antifungal activity evaluations of novel pyrazole carboxylic and dicarboxylic acid derivatives. Eur J Med Chem. 2014; 78: 86-96. doi: 10.1016/j.ejmech.2014.03.033.
9. Bersuker IB, Dimoglo AS. The electron-topological approach to the QSAR problem. In: Lipkowitz KB, Boyd DB, editors. Reviews in Computational Chemistry Vol. 2. New York: VCH; 1991. 423-461 pp.
10. Dimoglo AS, Chumakow YM, Dobrova BN, Saraçoğlu M. Electron-topological investigations of the structure-antitumor activity relationship of thiosemicarbazone derivatives. Arzneimittelforschung. 1997; 47(4): 415-419.
11. Shvets N, Terletskaya A, Dimoglo A, Chumakov Y. Study of the electronic and structural features characteristic of 4,5-dihydro-1-phenyl-1H-2,4-benzodiazepines demonstrating antiarrhythmic activity. J. Mol. Struct. (Theochem). 1999; 463(1-2): 105-110. doi:10.1016/S0166-1280(98)00399-6.
12. Güzel Y, Saripinar E, Yilidrim I. Electron-topological (ET) investigation of structure-antagonist activity of a series of dibenzo[a,d] cycloalkenimines. J. Mol. Struct. (Theochem). 1997; 418(1):83-91. doi:10.1016/S0166-1280(97)00069-9.
13. HyperChem Release 7, Molecular Modelling Program Package; Hypercube Inc.: Gainesville, FL, USA, 2001.
14. Hu VW, Black GE, Torres-Duarte A, Abramson FP. ^{3}H-thymidine is a defective tool with which to measure rates of DNA synthesis. FASEB J. 2002; 16(11): 1456-1457.
15. Lu TS, Yiao SY, Lim K, Jensen RV, Hsiao LL. Interpretation of biological and mechanical variations between the Lowry versus Bradford method for protein quantification. N Am J Med Sci. 2010; 2(7): 325-328. doi: 10.4297/najms.2010.2325.