QSAR and QSTR study of selenocyanate derivatives to improve their therapeutic index as anti-leishmanial agents

Article *in* Medicinal Chemistry Research · February 2014

DOI: 10.1007/s00044-013-0610-8

CITATIONS
5

READS
129

2 authors:

Maryam Iman

74 PUBLICATIONS **712** CITATIONS

SEE PROFILE

Asghar Davood

Islamic Azad University, Pharmaceutical Science Branch, Tehran, Iran

44 PUBLICATIONS **305** CITATIONS

SEE PROFILE

Med Chem Res (2014) 23:818–826
DOI 10.1007/s00044-013-0610-8

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

QSAR and QSTR study of selenocyanate derivatives to improve their therapeutic index as anti-*leishmanial* agents

Maryam Iman · Asghar Davood

Received: 27 January 2013 / Accepted: 2 May 2013 / Published online: 13 July 2013

Abstract Experimental data of selenocyanate derivatives show that they are potent inhibitors of *Leishmania* growth. In the current study, a series of selenocyanate derivatives were subjected to quantitative structure activity relationship and quantitative structure toxicity relationship analyses to identify the ideal physicochemical characteristics of potential anti-leishmanial activity with limited cytotoxic effects. Selenocyanate were built using HyperChem program, and conformational studies were performed through the semi-empirical method followed by the PM3 force field. Multi-linear regression was used as a chemo metric tool. The developed models were shown to be statistically significant according to the validation parameters. Based on our computational studies four descriptors, SPAM, P2e, B08 [N–N] and G2p can affect the activity and the cytotoxicity of selenocyanate derivatives.

Keywords Cytotoxicity · Leishmaniasis · QSAR · QSTR · Selenocyanates

M. Iman
Chemical Injuries Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

M. Iman
Center for Research and Training in Skin Diseases and Leprosy, Tehran University of Medical Sciences, Tehran, Iran

A. Davood (✉)
Department of Medicinal Chemistry, Pharmaceutical Sciences Branch, Islamic Azad University, No 99, Yakhchalave, Shariatie Street, Tehran 19419, Iran
e-mail: adavood@iaups.ac.ir; adavood2001@yahoo.com

Introduction

Leishmaniasis is a parasitic disease caused by protozoan parasites of the genus *Leishmania* with distinct manifestations (Croft and Coombs, 2003). These manifestations vary from skin disease to fatal visceral leishmaniasis (VL) which is the most serious form of the disease (Monzote, 2009; Dowlati *et al.*, 1996). It is a generalized infection of the reticuloendothelial system (RES) involving the spleen, liver, bone marrow and lymph nodes. The etiological agent in the old world including the Indian sub-continental, Sudan and other African countries is *L. donovani*, and *L. infantum* around the Mediterranean basin *L. chagasi* is the causative agent of VL in several South-American countries. VL caused by *L. donovani* is severe with very high mortality even with treatment. Available drugs are either expensive or accompanied with side effects or are not always effective (Croft and Yardley, 2002). Moreover, resistance to available drugs has become a serious problem which justifies the search for new synthetic and natural origin anti-leishmanial agents (Agarwal *et al.*, 2005; Hadighi *et al.*, 2006; Hadighi *et al.*, 2007) such as the synthetic compound selenocyanate, which has been reported to have anti-*Leishmania* activities (Plano *et al.*, 2011). In previous study, it was explored that increased concentration of selenium in plasma could have anti-*Leishmania* effect. Therefore, research and development of a potent and effective anti-leishmanial drug are essential to improve therapeutic strategies.

Quantitative structure activity relationships (QSAR) and quantitative structure toxicity relationships (QSTR) are widely used in the drug design process to improve the therapeutic index, wherever detailed structural information on ligand–receptor interactions are not experimentally available (Hansch *et al.*, 2001; Hansch and Fujita, 1964; Hansch, 1969; Gaudio *et al.*, 1994; Davood *et al.*, 2009; Iman *et al.*,

2011; Davood *et al.*, 2012a, b). QSAR and QSTR models are mathematical equations relating the chemical structure to the biological activities and toxicity, respectively (Hansch and Leo, 1995; Kubinyi, 1997a, b). The first component in the definition of a QSAR and QSTR model is the computation of the structural descriptors from the three-dimensional molecular structure, various geometrical, quantum, or molecular field descriptors that have been proposed in recent years to replace the Hansch substituent constants (Gaudio *et al.*, 1994; Hansch and Fujita, 1964; Hansch, 1969). Hence a wide range of descriptors have been used in QSAR and QSTR modeling and these descriptors have been classified into different categories such as constitutional, geometrical, topological, quantum, chemical, and so on. In this study, about 3,224 descriptors were used (Todeschini and Consonni, 2009; Todeschini *et al.*, 2007). The second component of a QSAR and QSTR model is an explicit mathematical structure activity equation to establish a statistical relationship between a dependent variable (biological activity and toxicity) and a set of independent variables (descriptors) (Gaudio *et al.*, 1994; Hansch and Fujita, 1964; Hansch, 1969). The mathematical QSAR and QSTR equations can be produced using a large number of statistical models such as multi-linear regression (MLR) and partial least squares (PLS) (Gramatica and Papa, 2003; Hansch *et al.*, 2001; Cramer *et al.*, 1988). Our QSAR and QSTR models are based on the anti-*Leishmania* activity and cytotoxicity of a set of 20, selenocyanate derivatives, which were synthesized in a previous experiment (Plano *et al.*, 2011) and many of the descriptors were calculated using the Dragon (Todeschini and Consonni, 2009; Todeschini *et al.*, 2007) and HyperChem software for all of the compounds. To select the set of descriptors most relevant to the IC_{50} of the compounds, MLR models were built and QSAR and QSTR equations with stepwise selection and elimination of variables were established using SPSS and Matlab software.

In the present research, it is tried to describe the QSAR and QSTR studies that have been done to investigate the quantitative effect of the various physicochemical parameters of selenocyanate on anti-*Leishmania* activity and cytotoxicity to define which physicochemical parameters may increase anti-*Leishmania* activity while decrease cytotoxicity.

Materials and methods

Computation of structural descriptors and QSAR equations

Molecular modeling and software

The two-dimensional structures of desired compounds 1–20 (Table 1) were built using HyperChem software (version 7, Hypercube Inc.). Conformational analyses of all compounds were performed by the semi-empirical molecular orbital calculation (PM3) method using HyperChem software. The molecular structures were optimized using the Polak-Ribiere (conjugate gradient) algorithm until the root mean square gradient was 0.01 kcal mol^{-1}. Among all energy minima conformers, the global minimum of compounds were used in QSAR and QSTR calculations and then the resulting geometry was transferred to the Dragon program, developed by Milano Chemometrics and QSAR Group (Todeschini and Consonni, 2009; Todeschini *et al.*, 2007). SPSS (version 18) and Matlab (version 7.6.0, R2008a) software were used for the MLR regression method.

Table 1 The chemical structure of selenocyanates

(a) Ar-(CH2)n-SeCN		(b) Ar-(CH2)n-Se-Se-(CH2)n-Ar
Compound	Aryl ring	*n*
1a	4-Aminophenyl	0
2a	4-(*N*,*N*-dimethylamino)phenyl	0
3a	4-Nitrophenyl	1
4a	3-Nitrophenyl	1
5a	2-Nitrophenyl	1
6a	4-Bromophenyl	1
7a	4-Trifluoromethylphenyl	1
8a	4-Methylthiophenyl	1
9a	4-Methylphenyl	1
10a	4-Cyanophenyl	1
11a	3-Cyanophenyl	1
12a	2-Cyanophenyl	1
13a	Phenyl	1
14a	Naphthyl	1
15a	4-Nitrophenyl	2
16b	4-Aminophenyl	0
17b	4-(*N*,*N*-dimethylamino)phenyl	0
18b	4-Nitrophenyl	1
19b	2-Nitrophenyl	1
20b	Phenyl	1

Data set and descriptor generation

Biological data used in this study were cytotoxicity and anti-*Leishmania* activity (IC_{50}) against *L. infantum* promastigote and amastigote of selenocyanate derivatives (Plano *et al.*, 2011), which was used for subsequent QSAR and QSTR analysis as dependent variables. *Leishmania* cells have two morphological forms, promastigote and amastigote. In mammalian hosts, *Leishmania* parasites are obligatorily intracellular, which named Amastigotes. Amastigotes adapt to living within the confines of the phagolysosomal apparatus of the host cells and initiate infection.

In the insect host, *Leishmania* parasites named promastigote. It is elongated, flagellated, extracellular and motile form of this parasite and easily grown in appropriate culture media (Handman *et al.*, 2000).

A large number of molecular descriptors were calculated using the HyperChem (Table 2) and Dragon package. Some of the chemical parameters including molecular volume (*V*), molecular surface area (SA approx), surface area (SA grid) hydrophobicity (Log *P*), hydration energy (HE), refractivity (Rf), molecular polarizability (MP) and different quantum chemical descriptors including dipole moment (DM) and the highest occupied molecular orbital (HOMO) energies were calculated using HyperChem software (Table 2). Dragon software was used to calculate different functional groups, topological, geometrical, and constitutional descriptors for each molecule. The calculated descriptors were collected in a data matrix whose number of rows and columns were the number of molecules and descriptors, respectively.

Data screening and model building

The calculated descriptors are first analyzed for the existence of a constant or near-constant variable and when detected they were removed. In addition, to decrease the redundancy which existed in the descriptor data matrix, the correlation of descriptors with each other and with the activity (pIC_{50}) of the molecules was examined and collinear descriptors (i.e. $r > 0.8$) were detected. Among the collinear descriptors, the one that showed the highest correlation with activity was retained and the others were removed from the data matrix. To select the set of descriptors, which were most relevant to the anti-*Leishmania* activity and cytotoxicity (pIC_{50}), the MLR models were built and the QSAR and QSTR equations with stepwise selection and elimination of variables were established using MLR method.

In the case of each regression problem, SPSS produced many models and ranked them based on standard error of calibration and coefficient of multiple determinations, wherein some models had a large number of input variables and thus they were over-fitted. To hinder obtaining over-fitted models, the generated QSAR models were validated by the leave-one out cross-validation procedure to check their predictability and robustness. A balance between the high cross-validation correlation coefficient and low number of descriptors were used as the criterion for model selection. The overall prediction abilities of the final models were accessed using a prediction set containing about 25 % of the original molecules. To do so, the data set of activity was randomly classified to calibrate and predict the sets. The model coefficients were

Table 2 Calculated properties of selenocyanates using the Hyperchem software

Compound	Surface area (approx)	Surface area (grid)	Volume	Hydration energy	Log *p*	Refractivity	Polarizability	Dipole moment	HOMO[a]
1	303.75	330.85	492.35	−10.45	1.04	36.37	15.89	4.7	−9.1981
2	405.65	380.41	587.27	−4.65	2.08	46.1	19.56	5.905	−8.9656
3	380.3	367.92	563.89	−10.98	1.87	43.83	23.64	3.453	−10.081
4	372.85	367.8	563.03	−10.75	1.87	43.83	18.21	6.203	−9.985
5	360.83	363.3	559.69	−10.41	1.87	43.83	18.21	6.571	−10.007
6	371.74	371.06	563.2	−5.68	2.71	44.13	19	4.098	−9.758
7	384.6	381.81	578.43	−5.43	2.8	42.48	17.93	2.96	−9.9385
8	387.74	392.17	601.16	−5.75	1.83	52.5	21.64	6.547	−9.3374
9	368.47	365.25	553.21	−4.86	2.38	41.55	18.21	5.313	−9.5889
10	383.31	367.33	557.71	−10.23	1.95	42.25	18.22	2.939	−9.9054
11	377.43	368.7	557.91	−10.13	1.95	42.25	18.22	5.328	−9.8768
12	341.56	354.69	540.71	−9.18	1.95	42.25	18.22	6.742	−9.7535
13	327.8	336.29	501.92	−6.03	1.91	36.51	16.37	5.021	−9.6296
14	324.84	389.34	611.86	−6.24	2.92	52.96	23.64	5.17573	−9.0271
15	425.83	400.46	622.39	−10.75	2.12	48.59	20.05	2.778	−10.091
16	387.63	470.22	741.03	−11.55	1.64	60.35	27.3	0.1863	−7.9605
17	584.93	557.44	935.45	0.13	3.73	79.81	34.64	0.3747	−8.2371
18	533.41	551.09	901.99	−12.04	3.3	75.27	31.95	1.884	−9.6724
19	461.44	524.87	879.05	−10.05	3.3	75.27	31.95	3.94	−8.7789
20	431.09	488.12	782.74	−2.13	3.39	60.62	28.27	2.132	−9.0219

[a] Highest occupied molecular orbital

calculated using calibration data and then they were used to calculate the anti-*Leishmania* activity of the molecules in the prediction set.

Results and discussion

Quantitative structure activity relationship (QSAR) equations

Based on the procedure explained in the "Materials and methods" section, using a stepwise multiple linear regression method, the following four-parametric equation (Eq. 1) was derived for selenocyanates 1–20. The correlation coefficient matrix for the descriptors used in the MLR equation has been provided in Table 3. In the QSAR equations, n is the number of data points, R^2 is the correlation coefficient, S is the standard deviation, F is the Fisher's F value, and q^2 is the leave-one-out (LOO) cross-validated coefficient that was obtained by a multiple linear regression.

The leishmanicidal activities of the selenocyanate derivatives were tested against both the extracellular promastigote and the intracellular amastigote forms of the parasite. We explain QSAR model for anti-leishmanial activity against promastigote and amastigote in Eqs. 1 and 2, respectively.

$$\begin{aligned} \mathrm{pIC_{50}} &= (1219 \pm 062) - (17.51 \pm 106)\ \mathrm{SPAM} - (2.87 \\ &\quad \pm 022)\ \mathrm{Mor19p} - (0.52 \\ &\quad \pm 0079)\ \mathrm{Mor25m} + (0.197 \pm 0053)\ \mathrm{EEig04d} \\ n &= 20,\ F = 137.09,\ R^2 = 0.982, \\ S &= 0.076,\ p < 0.000,\ q^2 = 087. \end{aligned} \quad (1)$$

An appropriate QSAR model is indicated by large F, small S, small P value, as well as R^2 and q^2 values close to 1. In general, the regression model is significant at P value < 0.001 using the F statistics so the Eq. (1) is significant.

Equation 1 explains 98.2 % of the variance in $\mathrm{pIC_{50}}$ data wherein the relative error prediction (REP) of the equation is shown in Table 3, which describes the effect of SPAM, Mor19p, Mor25m and EEig04d indices on anti-*Leishmania* (promastigote) activity.

SPAM is among the geometrical descriptors and corresponds to average span R and the EEig04d is among the edge adjacency indices and weighted by dipole moments. The Mor19p and Mor25m are among the 3D-MoRSE descriptors that were weighted by atomic polarizabilities and atomic masses, respectively.

Equation 1 indicates that EEig04d demonstrates positive contribution and SPAM, Mor19p and Mor25m show a negative contribution toward the anti-*Leishmania* (promastigote) activity. Comparison of the coefficient of descriptors reveals that anti-*Leishmania* (promastigote) activity might be affected mainly by SPAM descriptor. The calculated $\mathrm{pIC_{50}}$ using Eq. 1 is presented in Table 3 and the graphical representation of cross-validated calculated activity and the experimental values using Eq. 1 is presented in Fig. 1. The correlation coefficient matrix for the descriptors that were used in the MLR equation 1 is shown in Table 4.

Table 3 Anti-leishmanial activity against promastigote of selenocyanates in term of $\mathrm{pIC_{50}}$

Compound	[a]$\mathrm{pIC_{50}}$ exp.	[b]$\mathrm{pIC_{50}}$ calc.	[c] |REP| %
1	2.031984	2.039709	0.003787
2	2.554396	2.266653	0.12695
3	3.167491	2.943924	0.07594
4	3.244125	3.285832	0.012693
5	3.468521	3.452337	0.00469
6	2.099633	2.011742	0.04369
7	2.669586	2.794944	0.044852
8	2.519993	2.494632	0.01017
9	2.477556	2.545088	0.026534
10	3.259637	2.301788	0.41613
11	3.275724	3.212678	0.01962
12	2.500313	2.516296	0.006352
13	2.58838	2.53161	0.02242
14	2.429457	3.460315	0.297909
15	2.420216	2.506288	0.034342
16	3.187087	3.100385	0.02796
17	3.113509	3.057398	0.01835
18	3.420216	3.422046	0.000535
19	3.537602	3.920308	0.097621
20	3.200659	3.238404	0.011655

[a] The experimentally activity ($\mathrm{pIC_{50}}$) in *Leishmania infantum*

[b] The calculated $\mathrm{pIC_{50}}$ using multi-linear regression equation 1

[c] The absolute value of percent of the relative error of prediction

Based on this model (Eq. 1) to design new and potent ligands, substituent with high value of EEig04 and low values of SPAM, Mor19p and Mor25m should be used.

Based on the procedure explained in the experimental section, using a stepwise MLR method, the following four-parametric equation 2 was derived for anti-*Leishmania* (amastigote) activity of selenocyanates 1–20.

$$\begin{aligned} \mathrm{pIC_{50}} &= (1.663 \pm 0048) + (0.429 \\ &\quad \pm 0023)\ \mathrm{F06[N{-}X]} + (2.958 \pm 0207)\ \mathrm{P2e} \\ &\quad - (0.488 \pm 006)\ \mathrm{nArCN} - (0.254 \\ &\quad \pm 0048)\ \mathrm{RTM^{+}} \\ n &= 20,\ F = 91.38,\ R^2 = 0.973, \\ S &= 0.047,\ p < 0.000,\ q^2 = 0.77. \end{aligned} \quad (2)$$

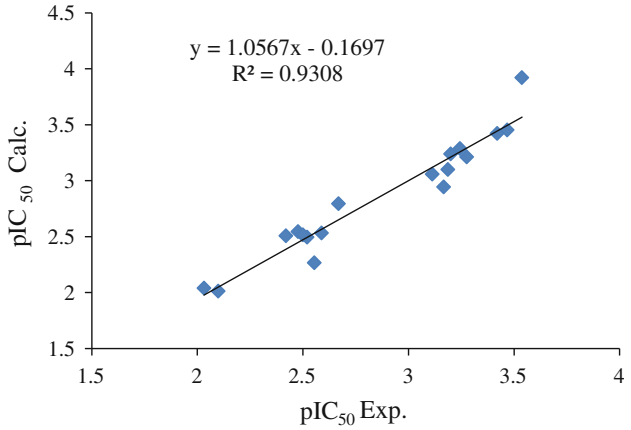

Fig. 1 Plot of cross-validated calculated activity of *L. infantum* obtained by QSAR equation 1

Equation 2 explains 97.3 % of the variance in pIC_{50} data and the REP of this equation is shown in Table 5 which describes the effect of F06 [N–X], P2e, nArCN and RTM^+ indices on anti-*Leishmania* (amastigote) activity. F06 [N–X] is among the two-dimensional frequency fingerprints and corresponds to frequency of N–X at topological distance 06 and nArCN shows the number of aromatic nitriles. P2e and RTM^+ are among the WHIM and GETWAY descriptors, which was weighted by atomic Sanderson electronegativities and atomic masses, respectively. Equation 2 indicates that F06 [N–X] and P2e demonstrate positive contribution and nArCN, and RTM^+ demonstrate negative contribution toward the activity. Comparison of the coefficient of descriptors reveals that which anti-*Leishmania* (amastigote) activity might be affected mainly by P2e.

The calculated pIC_{50} using the MLR of Eq. 2 is presented in Table 5 and the graphical representation of cross-validated of the experimental values and calculated activity using Eq. 2 have been provided in Fig. 2. The correlation coefficient matrix for the descriptors that were used in Eq. 2 is shown in Table 6.

Based on this model (Eq. 2) to design a new and potent ligands, substituent with high value of F06 [N–X] and P2e and low values of nArCN and RTM^+ should be considered.

Comparison of Eqs. 1 and 2 revealed in both of them atomic masses (Mor25m and RTM^+) have negative contribution and electronic parameter (EEig04d and P2e) have positive contribution on the leishmanicidal activity against promastigote and amastigote.

Quantitative structure toxicity relationships (QSTR)

Based on the procedure explained in the experimental section, using a stepwise multiple linear regression method, the following four-parametric equations 3 and 4 were derived for cytotoxic activity in Jurkat and THP-1 cell lines of selenocyanates 1–20, respectively.

Table 4 Pearson correlation coefficient matrix for the descriptors of selenocyanates used in the MLR activity equation 1

Correlations					
	BR EXP	SPAM	Mor19p	Mor25m	EEig04d
Pearson correlation					
BR EXP	1.000	−0.505	−0.187	0.176	0.679
SPAM	−0.505	1.000	−0.491	−0.321	−0.507
Mor19p	−0.187	−0.491	1.000	−0.050	0.226
Mor25m	0.176	−0.321	−0.050	1.000	0.256
EEig04d	0.679	−0.507	0.226	0.256	1.000
Sig. (one-tailed)					
BR EXP	–	0.012	0.215	0.229	0.000
SPAM	0.012	–	0.014	0.084	0.011
Mor19p	0.215	0.014	–	0.417	0.169
Mor25m	0.229	0.084	0.417	–	0.138
EEig04d	0.000	0.011	0.169	0.138	–
N					
BR EXP	20	20	20	20	20
SPAM	20	20	20	20	20
Mor19p	20	20	20	20	20
Mor25m	20	20	20	20	20
EEig04d	20	20	20	20	20

Table 5 Anti-leishmanial activity against amastigote of selenocyanates in term of pIC_{50}

Compound	[a]pIC_{50} exp.	[b]pIC_{50} calc.	[c]|REP| %
1	1.950782	1.995472	0.022396
2	2.034328	2.01638	0.0089
3	2.366532	2.382236	0.006592
4	2.378824	2.30718	0.03105
5	2.478862	2.496598	0.007104
6	1.935542	1.906742	0.0151
7	1.946922	1.919206	0.01444
8	1.935542	2.011142	0.037591
9	1.928118	1.952326	0.0124
10	2.383	1.488556	0.60088
11	2.217527	2.216748	0.00035
12	2.110698	1.83886	0.14783
13	2.099633	2.089902	0.00466
14	1.950782	2.629068	0.257995
15	1.924453	1.867952	0.03025
16	3.017729	2.567886	0.17518
17	2.497573	2.512076	0.005773
18	2.393619	2.368904	0.01043
19	1.90309	1.956844	0.02747
20	1.74958	1.757348	0.00442

[a] The experimentally activity (pIC_{50}) in *L. infantum*

[b] The calculated pIC_{50} using multi-linear regression equation 2

[c] The absolute value of Percent of the relative error of prediction

$$\begin{aligned} pIC_{50} = {} & (0629 \pm 0158) + (0.708 \pm 0046)\ F07[C-N] \\ & - (1113 \pm 0105)\ B08\,[N-N] \\ & + (0095 \pm 0015)\,RDF045e \\ & + (0.226 \pm 0072)\ HATS2m \\ n = 19,\ F = 85.9,\ R^2 = 0.972, & \\ S = 0.073,\ p < 0.000,\ q^2 = 094 & \qquad (3) \end{aligned}$$

Equation 3 explains 97.2 % of the variance in pIC50 data, wherein the REP of this equation is shown in Table 7. This equation describes the effect of F07[C–N], B08 [N–N], RDF045e and HATS2m indices on cytotoxicity of these compounds. F07[C–N] and B08 [N–N] are among the two-dimensional frequency and binary fingerprints and correspond to frequency of C–N and N–N at topological distance, respectively. RDF045e and HATS2m are among the RDF and GETAWAY descriptors and were weighted by atomic Sanderson electronegativities and atomic masses, respectively.

Equation 3 indicates that F07[C–N], RDF045e, and HATS2m show positive contribution and B08 [N–N] shows negative contribution toward the cytotoxicity on the Jurkat cell lines. Comparison of descriptors coefficient reveals that cytotoxicity can be affected mainly by B08 [N–N] and

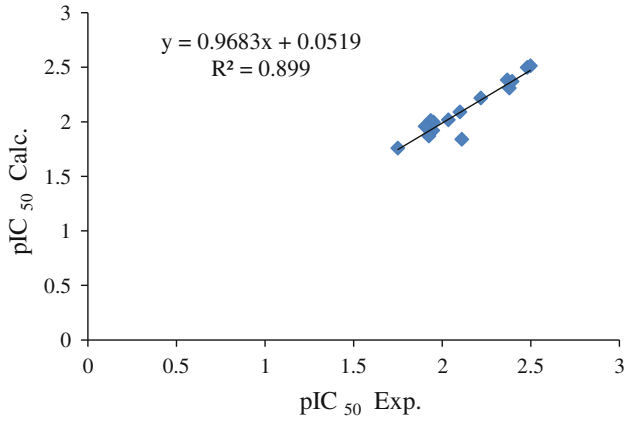

Fig. 2 Plot of cross-validated calculated activity of *L. infantum* obtained by QSAR equation 2

Table 6 Pearson correlation coefficient matrix for the descriptors of selenocyanates used in the MLR activity equation 2

Correlations					
	BR EXP	F06 [N–X]	P2e	nArCN	RTm^+
Pearson correlation					
BR EXP	1.000	0.718	−0.116	0.109	0.115
F06[N–X]	0.718	1.000	−0.501	−0.038	0.274
P2e	−0.116	−0.501	1.000	0.247	−0.102
nArCN	0.109	−0.038	0.247	1.000	−0.176
RTm^+	0.115	0.274	−0.102	−0.176	1.000
Sig. (one-tailed)					
BR EXP	–	0.000	0.313	0.323	0.315
F06[N–X]	0.000	–	0.012	0.437	0.121
P2e	0.313	0.012	–	0.147	0.334
nArCN	0.323	0.437	0.147	–	0.229
RTm^+	0.315	0.121	0.334	0.229	–
N					
BR EXP	20	20	20	20	20
F06[N–X]	20	20	20	20	20
P2e	20	20	20	20	20
nArCN	20	20	20	20	20
RTm^+	20	20	20	20	20

F07[C–N]. The calculated pIC50 using Eq. 3 is presented in Table 7 and the graphical representation of cross-validated of the experimental and the calculated activity values using Eq. 3 have been provided in Fig. 3. The correlation coefficient matrix for the descriptors that were used in Eq. 3 is shown in Table 8.

Based on the model 3, using the substituent with more values of B08 [N–N] and low values of F07[C–N], RDF045e, and HATS2m will produce new ligands with low cytotoxicity.

Table 7 Cytotoxic activity in Jurkat cell lines of selenocyanates in term of pIC_{50}

Compound	[a]pIC_{50} exp.	[b]pIC_{50} calc.	[c]|REP| %
1	1.358526	1.375925	0.012645
2	1.533132	1.551497	0.011837
3	2.163676	1.454179	0.4879
4	1.91364	1.741939	0.09857
5	1.88941	1.936454	0.024294
6	1.896196	1.890207	0.00317
7	2.136083	2.06837	0.03274
8	2.294992	2.202564	0.04196
9	1.841638	1.977925	0.068904
10	1.872895	1.87882	0.003153
11	2.059982	2.058917	0.00052
12	1.767004	1.762529	0.00254
13	1.970616	1.843622	0.06888
14	1.853872	3.356631	0.447699
15	2.619789	2.659536	0.014945
16	1.69037	2.837481	0.404271
18	1.571865	1.585339	0.008499
19	2.581699	2.596539	0.0057
20	1.583359	1.591533	0.0051

[a] The experimentally cytotoxicity pIC50 in Jurkat cell line

[b] The calculated pIC_{50} using multi-linear regression equation 3

[c] The absolute value of percent of the relative error of prediction

$$\begin{aligned} pIC_{50} &= (1.568 \pm 0064) + (0.11 \pm 0007)\ \text{DELS} \\ &\quad - (0.06 \pm 0005)\ \text{G[N..O]} - (2333 \pm 029)\ G_2P \\ &\quad - (0349 \pm 0083)\ \text{HATS4m} \\ n &= 19, F = 152.6, R^2 = 0.984, \\ S &= 0.047, p < 0.000, q^2 = 096. \end{aligned} \quad (1)$$

Equation 4 shows the cytotoxicity of compounds 1–20 on the THP-1 cell lines and explains 98.4 % of the variance in pIC50 data, wherein the REP of this equation is shown in Table 9. This equation describes the effect of DELS, G [N..O], G2P, and HATS4m indices on cytotoxicity of these compounds. DELS and G [N..O] are among the topological and geometrical descriptors and correspond to molecular electrotopological variation and distance between N..O, respectively. G2p and HATS4m are among the WHIM and GETAWAY descriptors which weighted by atomic polarizabilities and atomic masses. Equation 4 indicates that DELS shows positive contribution and G [N..O], G2P, and HATS4m show negative contribution toward the cytotoxicity on the THP-1 cell lines. Comparison of the coefficient of the descriptors reveals which cytotoxicity can be affected mainly by G2P. The calculated pIC50 using Eq. 4 is presented in Table 9 and the graphical representation of cross-validated of the experimental and calculated

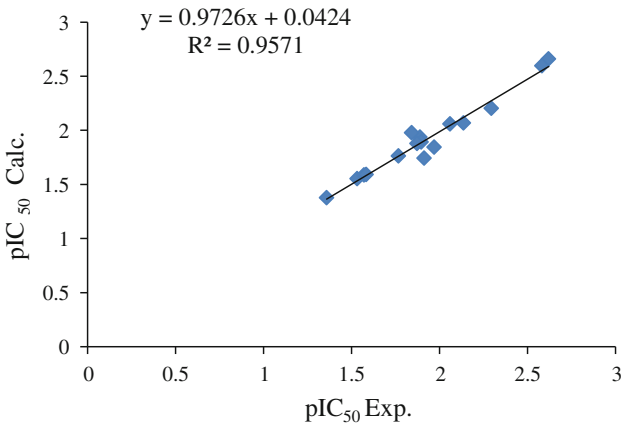

Fig. 3 Plot of cross-validated calculated cytotoxicity obtained by QSAR equation 3

Table 8 Pearson correlation coefficient matrix for the descriptors of selenocyanates were used in the MLR cytotoxicity equation 3

Correlations					
	BR EXP	F07 [C–N]	B08 [N–N]	RDF045e	HATS2m
Pearson correlation					
BR EXP	1.000	0.571	0.197	−0.209	−320
F07[C–N]	0.571	1.000	0.495	−0.357	−0.332
B08 [N–N]	0.197	0.495	1.000	−0.319	0.010
RDF045e	−0.209	−0.357	−0.319	1.000	−0.448
HATS2m	−0.320	−0.332	0.010	−0.448	1.000
Sig. (one-tailed)					
BR EXP	–	0.005	0.210	0.196	0.091
F07[C–N]	0.005	–	0.016	0.067	0.083
B08 [N–N]	0.210	0.016	–	0.092	0.484
RDF045e	0.196	0.067	0.092	–	0.027
HATS2m	0.091	0.083	0.484	0.027	–
N					
BR EXP	19	19	19	19	19
F07[C–N]	19	19	19	19	19
B08 [N–N]	19	19	19	19	19
RDF045e	19	19	19	19	19
HATS2m	19	19	19	19	19

cytotoxicity values using Eq. 4 have been provided in Fig. 4. The correlation coefficient matrix for the descriptors that were used in the MLR equation 4 is shown in Table 10.

Based on our QSAR and QSTR studies results, four descriptors SPAM, P2e, B08 [N–N] and G2p can affect the activity and the cytotoxicity and follow on therapeutic index of this series of ligand. Comparison of Eqs. 1–4 reveals increasing the distances between N–X (X = N or O) results to increase the therapeutic index due to their positive effect on the activity and negative effect on the cytotoxicity.

Table 9 Cytotoxic activity in THP-1 cell lines of selenocyanates in term of pIC_{50}

Compound	[a]pIC_{50} exp.	[b]pIC_{50} calc.	[c]|REP| %
2	1.3497	1.420442	0.049803
3	1.7747	1.81242	0.020812
4	2.0306	2.029107	0.00074
5	2.1487	2.22645	0.034921
6	1.5391	1.553239	0.009103
7	1.7033	4.233724	0.597683
8	1.7696	1.757004	0.00717
9	1.6198	1.664747	0.026999
10	2.0985	2.075251	0.0112
11	2.1192	2.094888	0.01161
12	1.767	1.721863	0.02621
13	1.6021	1.536448	0.04273
14	1.7747	1.763513	0.00634
15	1.7645	1.744764	0.01131
16	1.8153	1.803259	0.00668
17	1.6556	1.623491	0.01978
18	1.5361	0.769647	0.99585
19	2.6162	2.596031	0.00777
20	1.4921	1.501671	0.006374

[a] The experimentally cytotoxicity pIC50 in THP-1 cell line

[b] The calculated pIC_{50} using multi-linear regression equation 4

[c] The absolute value of Percent of the relative error of prediction

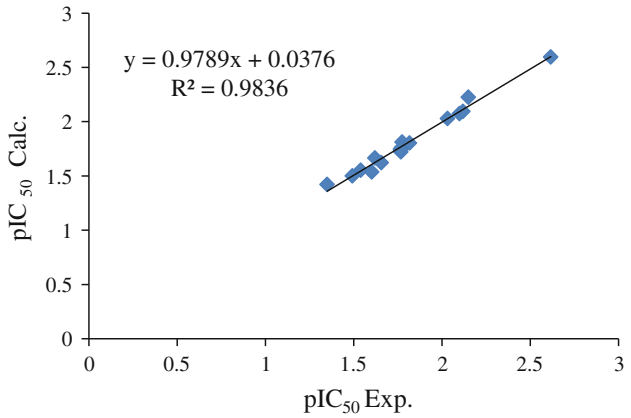

Fig. 4 Plot of cross-validated calculated cytotoxicity obtained by QSAR equation 4

Conclusions

Twenty analogs of selenocyanates with anti-leishmanial activity, using the MLR method were subjected to QSAR and QSTR studies to design a new ligand with an improved therapeutic index. Based on our present computational studies, mainly four descriptors SPAM, P2e, B08 [N–N] and G2p can affect the activity and the cytotoxicity of this series of ligands. These computational studies can offer some useful references to performing the molecular design or modification of this series of anti-leishmanial agents.

Table 10 Pearson correlation coefficient matrix for the descriptors of selenocyanates were used in the MLR cytotoxicity equation 4

Correlations					
	BR EXP	DELS	G (N..O)	G2p	HATS4m
Pearson correlation					
BR EXP	1.000	0.479	0.228	−0.286	0.297
DELS	0.479	1.000	0.749	−0.299	0.307
G (N..O)	0.228	0.749	1.000	−0.312	0.173
G2p	−0.286	−0.299	−0.312	1.000	−0.093
HATS4m	0.297	0.307	0.173	−0.093	1.000
Sig. (one-tailed)					
BR EXP	–	0.019	0.174	0.118	0.108
DELS	0.019	–	0.000	0.107	0.101
G (N..O)	0.174	0.000	–	0.096	0.239
G2p	0.118	0.107	0.096	–	0.352
HATS4m	0.108	0.101	0.239	0.352	–
N					
BR EXP	19	19	19	19	19
DELS	19	19	19	19	19
G (N..O)	19	19	19	19	19
G2p	19	19	19	19	19
HATS4m	19	19	19	19	19

These observations and experimental results provide a suitable process to explain the potent and selective inhibitory activities of these compounds. In addition, it seems that using lipophilic and electronegative moieties can improve the therapeutic index. Currently, our research group is exploring this idea for designing newer compounds with better anti-leishmanial activity.

Acknowledgments We appreciate the technical assistance of the Medicinal Chemistry Department of Azad University, Tehran Branch in performing the computational analyses.

References

Agarwal A, Ramesh Ashutosh, Goyal N, Chauhan PMS, Gupta S (2005) Dihydropyrido[2,3-*d*]pyrimidines as a new class of antileishmanial agents. Bioorg Med Chem 13(24):6678–6684

Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110(18):5959–5967

Croft SL, Coombs GH (2003) Leishmaniasis-current chemotherapy and recent advances in the search for novel drugs. Trends Parasitol 19(11):502–508

Croft SL, Yardley V (2002) Chemotherapy of leishmaniasis. Curr Pharm Des 8(4):319–342

Davood A, Nematollahi AR, Iman M, Shafiee A (2009) Synthesis and docking studies of new 1,4-dihydropyridines containing 4-(5)-Chloro-2-ethyl-5-(4)-imidazolyl substituent as novel calcium channel agonist. Arch. Pharm. Res. 32(4):481–487

Davood A, Iman M, Nematollahi A, Shafiee A (2012a) Docking and QSAR studies of new 1, 4-dihydropyridines containing 4(5)-chloro-2-methyl-5(4)-imidazolyl substituent. Med Chem Res 21(3):325–332

Davood A, Shafaroodi H, Amini M, Nematollahi A, Shirazi M, Iman M (2012b) Design, synthesis and protection against pentylenetetrazole-induced seizure of N-aryl derivatives of the phthalimide pharmacophore. Med Chem 8(5):953–963

Dowlati Y, Bahar K, Ehsasi S, Shidani B (1996) Stepwise safety trial of a killed Leishmania vaccine in Iran. Clin Dermatol 14(5): 496–502

Gaudio AC, Korolkovas A, Takahata Y (1994) Quantitative structure-activity relationships for 1,4-dihydropyridine calcium channel antagonists (nifedipine analogues): a quantum chemical/classical approach. J Pharm Sci 83(8):1110–1115

Gramatica P, Papa E (2003) QSAR Modeling of Bioconcentration Factor by theoretical molecular descriptors. QSAR Comb Sci 22(3):374–385

Hadighi R, Mohebali M, Boucher P, Hajjaran H, Khamesipour A, Ouellette M (2006) Unresponsiveness to Glucantime treatment in Iranian cutaneous leishmaniasis due to drug-resistant Leishmania tropica parasites. PLOS Med 3(5):e162

Hadighi R, Boucher P, Khamesipour A, Meamar AR, Roy G, Ouellette M, Mohebali M (2007) Glucantime-Resistant Leishmania tropica isolated from Iranian patients with cutaneous leishmaniasis are sensitive to alternative antileishmania drugs. Parasitol Res 101(5):1319–1322

Handman E, Noormohammadi AH, Curtis JM, Baldwin T, Sjolander A (2000) Therapy of murine cutaneous leishmaniasis by DNA vaccination. Vaccine 18:3011–3017

Hansch C (1969) Quantitative approach to biochemical structure-activity relationships. Accounts Chem Res 2(8):232–239

Hansch C, Fujita T (1964) A method for the correlation of biological activity and chemical structure. J Am Chem Soc 86(8):1616–1626

Hansch C, Leo A (1995) Exploring QSAR fundamentals and applications in chemistry and biology. American Chemical Society, Washington

Hansch C, Kurup A, Garg R, Gao H (2001) Chem-Bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms. Chem Rev 101(3):619–672

Iman M, Davood A, Nematollahi AR, Dehpoor AR, Shafiee A (2011) Design and synthesis of new 1,4-dihydropyridines containing 4(5)-chloro-5(4)-imidazolyl substituent as a novel calcium channel blocker. Arch Pharm Res 34(9):1417–1426

Kubinyi H (1997a) QSAR and 3D QSAR in drug design. Part 1: methodology. Drug Discov Today 2:457–467

Kubinyi H (1997b) QSAR and 3D QSAR in drug design Part 2: applications and problems. Drug Discov Today 2:538–546

Monzote L (2009) Current treatment of leishmaniasis: a review. Open Antimicrob Agents 1:9–19

Plano D, Baquedano Y, Moreno-Mateos D, Font M, Jiménez-Ruiz A, Palop JA, Sanmartín C (2011) Selenocyanates and diselenides: a new class of potent antileishmanial agents. Eur J Med Chem 46(8):3315–3323

Todeschini R, Consonni V (2009) Molecular descriptors for chemoinformatics, vol 1: Alphabetical Listing, vol 2: Appendices, References. Wiley, vol 110

Todeschini R, Consonni V, Mauri A, Ballabio D, Manganaro A (2007) Milano chemometrics and QSAR research group, vol 30. University of Milano, Milano Italy, http://michem.disat.unimib.it/chm/index.htm