ScienSage

Journal of Advanced Scientific Research

Available online through https://sciensage.info

ISSN
0976-9595
Research Article

MULTIPLE LINEAR REGRESSION ANALYSIS OF HIV DRUG OF PROTEASE INHIBITOR

M B Kalhans[1], A K Singh*[2], N B Singh[1]

[1]Department of Chemistry, Bareilly College, Bareilly, Uttar Pradesh, India

[2]Department of Physics, Bareilly College, Bareilly, Uttar Pradesh, India

*Corresponding author: draksingh@bcb.ac.in

Received: 08-02-2023; Accepted: 27-03-2023; Published: 30-04-2023

 https://doi.org/10.55218/JASR.202314403

ABSTRACT

HIV-1 protease, because of its sensitive and essential function, is an excellent target for the drug therapy. The primary chain of an aspartic protease has two different Asp-Thr-Gly sequences. The MLR approach attempts to identify and quantify the physicochemical properties of the drug and to see whether any of these properties has any effect on drug's biological activity. In the present study, 15 protease derivatives have been used as study material and on the basis of MLR equation their activity has been predicted.

Keywords: HIV-1 protease, Aspartic, MLR approach, Physicochemical.

1. INTRODUCTION

The cleavage of a large polypeptide precursor into a smaller protein function fragment is required for packaging and infectivity of budding virion needs HIV-protease. HIV- protease is a viral encoded homodimetric aspartyl protease having C2 symmetry [1, 2]. The cleavage reactions must be timed properly, allowing the immature virus to assemble properly before the polyprotein is broken. Because of its sensitive and essential function, HIV-1 protease is an excellent target for the drug therapy [3, 4]. The atomic structure of HIV-1 protease has made much of this work possible. The first structures were reported in 1989. A decade later, over 100 structures are available, including several genetic strains of this enzyme, complexes of the enzyme with many different drugs and inhibitors as well as mutant enzymes.

Since HIV-1 protease is an aspartic protease and its substrate is peptidic in nature, a number of peptide derived compounds have been identified as HIV-1 PR inhibitors [5]. The primary chain of an aspartic protease has two different Asp-Thr-Gly sequences and the apostructure of it shows these two chains are running in opposite directions with a water molecule bounded between two aspartates [6]. The substrate possesses a scissile bond which during the substrate-enzyme interaction, interacts with the corresponding binding sites on the enzyme [7]. In the present study we have focussed our attention analyzing the derivable quantitative structure activity relationship studies on peptidic and non peptidic inhibitors with the help of PM-3 and cache software.

2. METHODOLOGY

The first attempt to relate a physicochemical parameter to pharmacological effect was reported by Prichet [8-10]. The MLR approach attempts to identify and quantify the physicochemical properties of the drug and to see whether any of these properties have any effect on the drug's biological activity. If such a relation holds true, a regression equation can be drawn up which quantifies the relationship and permits a chemist to predict with source confidence that property or properties has an important role in the distribution or mechanism of action of drug. In present study there are two considerations for biological activity, namely lipophilicity and electronic factors. Lipophilicity is a measure of the journey of a drug to the site of action and electronic factors give an idea for the drug interaction with the site of action. Lipophilicity, which is a parabolic function, is expanded by Hansch and Fujita [11].

Because of the importance of the steric effects and other shape factors of molecules for receptor interaction, an Es term and a variety of other shape, size or topography term (s) have been added to the Hansch equation. These parameters are used by the application of the method of

linear multiple regression analysis [12-14].

$$log\left(\frac{1}{c}\right) = -a\pi^2 + b\pi + \rho\sigma + cE_s + dS + e$$

The 15 PR- inhibitor derivatives have been used as study material with their observed biological activity [15]. The quantum chemical and energy descriptors (Ionization potential, Electron Affinity, Atom electron density, Softness of the atom E_n and E_m, Density distribution function and Fukui Function) have been used by us for the MLR analysis of the compounds.

3. RESULTS AND DISCUSSION

The calculated values of descriptors based on atomic *i.e.* Ionization potential, Electron Affinity, Atom electron density, Softness of the atom En= and Em, Density distribution function and Fukui Function of all the fifteen compounds are given in Table 1. For MLR analysis the value of atomic softness, value of ionization potential, value of electron affinity and Fukui function of each compound have been considered.

Table 1: Values of different parameters of Cyclic Urea compounds at active sites

Compound	Ionization Potential	Electron Affinity	Atom Electron Density	Softness of The Atom En	Softness of The Atom Em	Density Distribution Function	Fukui Function
TC1	23.24764	-14.18628	1.87388	16.25876	-4.529389	12.90001	0.2342350
	14.94849	-5.887126	1.36939	10.35004	11.34036	9.427038	0.1956271
	14.48230	-5.42094	1.42612	10.20300	8.723949	9.817574	0.2037314
	17.74765	-8.686286	1.83310	13.17521	-1.023852	12.61927	0.2291375
	20.78091	-11.71955	1.81733	14.71514	-1.688521	12.51071	0.2271663
TC2	23.51813	-14.69134	1.87207	16.33458	-4.855215	12.88754	0.2340088
	14.53398	-5.707192	1.42972	10.17949	8.5131730	9.842356	0.2042457
	15.46661	-6.639817	1.39638	10.61982	10.969020	9.612839	0.1994828
	20.53987	-11.71308	1.80564	14.48187	-1.473188	12.43023	0.2257050
	20.49798	-11.67119	1.80967	14.47881	-1.577142	12.45789	0.2262088
TC3	23.47473	-14.66638	1.87381	16.31486	-4.905416	12.89952	0.2342263
	15.87499	-7.066634	1.43329	10.90784	10.102530	9.866933	0.2047557
	15.47726	-6.668906	1.39445	10.61735	11.051210	9.599553	0.1992071
	20.52984	-11.72148	1.80797	14.48117	-1.554381	12.44627	0.2259963
	20.49474	-11.68639	1.81171	14.47776	-1.651197	12.47202	0.2264638
TC4	23.60902	-14.85063	1.87289	16.36902	-5.003083	12.89319	0.2341113
	15.85212	-7.093729	1.41857	10.85347	10.630470	9.765599	0.2026529
	15.40878	-6.650385	1.37510	10.53256	11.695660	9.466346	0.1964429
	20.47511	-11.71672	1.80443	14.42551	-1.477357	12.42190	0.2255538
	20.44191	-11.68355	1.80733	14.41965	-1.551067	12.44287	0.2259163
TC5	23.47698	-14.65693	1.87370	16.31848	-4.891544	12.89877	0.2342125
	20.95419	-12.13414	1.43296	13.63347	25.944050	9.864661	0.2047086
	15.87161	-7.051556	1.43296	10.90827	10.119730	9.864661	0.2047086
	20.52933	-11.70928	1.39435	14.48399	-1.545085	9.598864	0.2260013
	20.49290	-11.67285	1.39435	14.47943	-1.638291	12.44655	0.2264550
TC6	23.56539	-17.16810	1.87163	15.75015	-7.286161	12.88452	0.2339538
	15.98108	-9.493786	1.41685	10.28082	8.962196	9.753756	0.2024071
	15.36441	-8.967121	1.37410	9.916586	9.934172	9.459462	0.1963000
	20.49061	-14.09332	1.80559	13.84809	-3.592597	12.42989	0.2256988
	20.46236	-14.06508	1.80809	13.84323	-3.651684	12.44710	0.2260113
TC7	23.47823	-14.67164	1.87361	16.31537	-4.902662	12.89815	0.2342013
	15.86840	-7.091811	1.43240	10.90203	10.127270	9.860804	0.2046286
	15.46512	-6.658529	1.39431	10.61010	11.035500	9.598589	0.1991872
	20.52999	-11.72340	1.80797	14.48081	-1.555980	12.44627	0.2259963
	20.49275	-11.68616	1.81161	14.47586	-1.649185	12.47133	0.2264513
TC8	23.60166	-14.70572	1.87374	16.40331	-4.88981	12.89904	0.2342175
	15.94486	-7.048925	1.43783	10.97678	10.08888	9.898186	0.2054043
	15.30028	-6.404348	1.38553	10.52730	11.18018	9.538147	0.1979329
	20.57374	-11.67781	1.80747	14.52425	-1.476767	12.44283	0.2259338
	20.49290	-11.59697	1.81286	14.50336	-1.606822	12.47994	0.2266075

	23.02959	-13.96034	1.87099	16.13294	-4.309332	12.88011	0.2338738
	15.84267	-6.773417	1.43530	10.95951	10.176760	9.880081	0.2050286
TC9	15.19466	-6.125407	1.46955	10.68096	8.138416	10.11655	0.2099357
	20.55361	-11.48436	1.80612	14.55150	-1.277856	12.46128	0.2257650
	20.58263	-11.51338	1.81015	14.58311	-1.405273	12.46128	0.2262288
	22.23582	-12.99115	1.86858	15.74750	-3.65301	12.86352	0.2335725
	15.06682	-5.822149	1.35701	10.43824	12.17014	9.341812	0.1938586
TC10	15.97829	-6.733625	1.45309	11.11393	9.796502	10.00324	0.2075843
	21.02762	-11.78295	1.81130	14.86671	-1.426415	12.46920	0.2264125
	20.73455	-11.48988	1.81357	14.72112	-1.399736	12.48482	0.2266963
	23.24960	-14.61167	1.87349	16.15220	-4.932262	12.89732	0.2341863
	15.37755	-6.739619	1.43136	10.59244	9.356646	9.853646	0.2044800
TC11	15.37920	-6.741274	1.39054	10.51458	10.922570	9.572637	0.1986486
	20.46066	-11.82273	1.80708	14.39842	-1.655639	12.44015	0.2258850
	20.50621	-11.86828	1.80970	14.43307	-1.746703	12.45818	0.2262125
	23.59750	-23.65279	1.87415	14.16514	-13.81416	12.90186	0.2342688
	15.97610	-16.03139	1.42706	8.733906	3.702625	9.824044	0.2038657
TC12	15.15518	-15.60715	1.37716	8.409922	5.316829	9.480527	0.1967371
	20.50152	-20.55681	1.80571	12.24119	-9.277766	12.43072	0.2257138
	20.47475	-20.53005	1.80937	12.24181	-9.935336	12.45591	0.2261713
	23.61788	-14.81608	1.87166	16.37906	-4.924171	12.88472	0.2339575
	15.68037	-6.878575	1.40661	10.74833	10.88209	9.683265	0.2009443
TC13	15.45207	-6.650273	1.37545	10.56736	11.79100	9.468756	0.1964929
	20.50063	-11.69883	1.80337	14.44560	-1.416134	12.41461	0.2254213
	20.45477	-11.65297	1.80604	14.43209	-1.479692	12.43299	0.2257550
	23.34176	-14.55670	1.87090	16.22603	-4.757365	12.87949	0.2338625
	14.63601	-5.850950	1.34798	10.07598	11.60657	9.279649	0.1925686
TC14	15.68713	-6.902074	1.38938	10.71525	11.60657	9.564651	0.1984829
	20.62248	-11.83742	1.80886	14.52790	-1.631358	12.45240	0.2261075
	20.51346	-11.82840	1.80908	14.47117	-1.601807	12.45391	0.2261350
	23.59247	-14.69091	1.86878	16.37784	-4.714065	12.86490	0.2335975
	16.19350	-7.291945	1.43972	11.11636	10.38664	9.911197	0.2056743
TC15	15.37221	-6.470648	1.37037	10.54022	11.93298	12.47601	0.1957672
	20.66429	-11.76273	1.81229	14.59301	-1.642834	12.47601	0.2265363
	20.50216	-11.60060	1.81918	14.53548	-0.785607	12.52344	0.2273975

Regression equations using the values provided in table1 are given below:

TC1=-0.0875928* En-0.0835626*Em+8.63924

rCV^2=0.749875

r^2=0.0812645

TC2=-0.0139861*En-8.81103*AED+24.283

rCV^2=0.429235

r^2=0.045991

TC3=-0.512828*En+0.423482*IP+6.09501

rCV^2=0.121087

r^2=0.1585071

TC4=-0.0474293*En-0.0676259*EA+7.49695

rCV^2=0.479485

r^2=0.325341

TC5=-0.0141025*En-1.27928*DDF+24.2768

rCV^2=0.429488

r^2=0.0459753

TC6=-0.0649327*En-48.3034*FF+19.9206

rCV^2=0.216667

r^2=0.032846

TC7=-0.151503*Em-17.9913*AED+40.4959

rCV^2=0.748959

r^2=0.219713

TC8=-0.0745367*Em+0.0733779*IP+5.60516

rCV^2=0.182449

r^2=0.0807609

TC9=-0.0202452*Em-0.0655262*EA+6.67692

rCV^2=0.105604

r^2=0.327208

TC10=-0.151543*Em-2.61403*DDF+40.5032

rCV^2=0.74857

r^2=0.219753

TC11=-0.144349*Em-125.96*FF+36.3361
rCV^2=0.567587
r^2=0.189184
TC12=-52.0334*AED+1.12734*IP+78.8736
rCV^2=0.374719
r^2=0.527176
TC13=-10.98*AED-0.0707939*EA+27.1821
rCV^2=-0.0931038
r^2=0.390185
TC14=-33064.3*AED+4801.56*DDF+25.9235
rCV^2=0.119853
r^2=0.0766715
TC15=-50.6517*AED+333.889*FF+24.2612
rCV^2=0.153395
r^2=0.0796851

4. CONCLUSIONS

MLR equations in which the rCV^2 has a negative value have no predictive value irrespective of the value of regression coefficient r^2. Regression equations in which cross- validation coefficient is greater than 0.2 and regression coefficient r^2 is less than 0.5 have low predictive power. Regression equations in which cross-validation coefficient is greater than 0.2 and the regression coefficient r^2 is greater than 0.5, has good predictive power. So the compounds discussed above have the following predictive power.

TC12> TC3> TC4

On the basis of these results we can further focus our studies on detailed pharmacokinetic research of these three compounds.

Conflict of interest
None declared

Source of funding
None declared

5. REFERENCES

1. Kohl NE et al. *Proc Natl Acad Sci USA*, 1988; **85**:4686.
2. Peng C, Ho BK, Chang TW, Chang NT. *Jour Virology*, 1989; **53**:2550.
3. Kalhans MB, Singh AK, Singh NB. *Int J Sc Tech Res*, 2019; **08**:1922.
4. Kalhans MB, Singh AK, Singh NB. *J Adv Sci Res,* 2021**; 12(1 Suppl 1):** 301-303.
5. Greuston KJ et al. *Amer Intern Med*, 1992; **11**:561.
6. Babine RE, Bender SL. *Chem Rev*, 1997; **97**:1359.
7. Wodawer A. Ericson JW. *Annu Rev Biochem*, 1993; **62**:543.
8. Ferguson J. *Proc R Soc London Ser B*, 1939; **127**:387.
9. Hansch C, Maloney PP, Fujita T, Muir RM. *Nature London*, 1962; **178:**194.
10. Richet MC, Seances CR. *Soc Biol Ses Fil*, 1893; **45**:775.
11. Hansch C, Fujita T. *J. Am Chem Soc*, 1964; **86**:1616.
12. Daniel C, Wood FS. "Fitting equation to data" Wiley New York, 1971.
13. Draper NR, Smith H. "Applied Regression Analysis" Wiley New York, 1966.
14. Suedecor GW, Cochram WG. "Statistical Helloids" IOWA State Univ Press, 1967.
15. Gupta R, Gupta SP, Hua Gao, Mekapati Suresh Babu, Debunath AK, Hansch C. *Chem Rev*, 1999; **99**:3525.