A QSAR study on some series of anti-hepatitis B virus (HBV) agents

Preet K. Arora, Vaishali M. Patil, Satya P. Gupta*

Department of Pharmaceutical Technology, Meerut Institute of Engineering and Technology, Meerut 250005, India; Satya P. Gupta Email: spgbits@gmail.com;*corresponding author.

Abstract

A quantitative structure-activity relationship (QSAR) study has been made on some series of anti-hepatitis B virus (HBV) agents, namely, a series of novel bis(L-amino acid) ester prodrugs of 9-[2-(phosphonomethoxy)ethyl]adenine, a similar series of compounds comprising of 2-amino-6-arylthio-9-[2-(phosphonoethoxy)ethyl]purine bis(2,2,2- trifluoroethyl) esters, and a series of 1-isopropylsulfonyl-2-amine benzimidazoles. In each case significant correlations are found between the anti-HBV potencies and some physicochemical and steric properties of the compounds, indicating that for the first two series the activity is controlled by the hydrophobic and the bulk properties of the molecules and, for the third series, the steric and hydrogen bonding properties of compounds are crucial for their anti-HBV potency.

Keywords: QSAR study, anti-hepatitis B virus agents, hepatitis B virus inhibitors, bis(L-amino acid) ester prodrugs of 9-[2-(phosphonomethoxy)ethyl]adenine, 2-amino-6-arylthio-9-[2-(phosphonoethoxy)ethyl]purine bis(2,2,2-trifluoroethyl) esters, 1-isopropylsulfonyl-2-amine benzimidazoles

Background:

Hepatitis B Virus (HBV) infection has now become one of the most dangerous diseases of the world. It causes both acute and chronic hepatitis [1]. Chronic HBV infection can lead to cirrhosis, liver failure and hepatocellular carcinoma [2, 3]. Presently, there are only few agents, namely α-interferon (IFN-α), i.e., lamivudine (3-Tc) (**1**), adefovir dipivoxil (**2**), and entacavir (**3**), that have been approved for clinical treatment. However, all these drugs suffer from serious side effects. Interferon-α, an immunomodulator and the first therapeutic agent developed to treat HBV infection, has very low cure rate (effective only to 30-40% patients) and produces many side effects **[4].** The three nucleoside analogs, lamivudine, adefovir dipivoxil, and entacavir, are supposed to elicit their effects mainly through the inhibition of HBV polymerase leading to decrease in viral replication **[5]**. All these three drugs suffer from high relapse rates after the cessation of the treatment **[6]**. Lamivudine has the potential of inducing drug resistant HBV after few months of therapy, and causes an associated risk of increase of viremia during the therapy in HBV patients. Although a combination therapy of lamivudine and adefovir dipivoxil may be more fruitful **[7]** since both have excellent activity and oral bioavailability and the combination has good tolerability and reduced potential to produce drug resistant HBV, this therapy is still limited **[8]**. Therefore, the medicinal chemists are paying high attention to design and develop novel classes of anti- HBV agents that may have optimal pharmacological profiles with little side effects. In this respect, the quantitative structure- activity relationship (QSAR) studies have been of great importance. They not only provide guidelines for the drug design but also throw the light on the mechanisms of drug-receptor interactions, which further aid to the rationalization of drug development. We, therefore, present here a QSAR study on some important series of HBV inhibitors.

Methodology:

We have taken three different series of HBV inhibitors that were synthesized and studied for their anti-HBV activity. Recently, a series of novel bis (L-amino acid) ester prodrugs of 9-[2-(phosphonomethoxy)ethyl]adenine (**4**) was synthesized and reported for their anti-HBV and toxic activities by Fu et al. **[9]** and a similar series of compounds comprising of 2-amino-6-arylthio-9-[2-(phosphonoethoxy)ethyl] purine bis(2, 2, 2- trifluoroethyl) esters (**5**) was earlier studied by Sekiya et al**.[10]** . However, an entirely new class of anti-HBV agents comprising of 1-isopropylsulfonyl-2-amine benzimidazoles (**6, 7**) was reported by Li et al. **[11]**. The first two series of compounds are listed I **Tables 1 and 2**, **(see supplementary material)** respectively. A combine of these two is listed in **Table 3, and Table 4 (see supplementary material) l**ists the derivatives of **6** and **7**. These tables also list the anti-HBV and toxic activities of the compounds and their physicochemical

parameters that were found to govern their potency. Among the physicochemical parameters listed in different tables, ClogP and CMR refer to calculated hydrophobicity and molar refractivity of the compounds, respectively, E_s is Taft's steric constant, and B1 is a STERIMOL parameter, a steric parameter referring to the minimum width of the substituent **[12]** .ClogP and CMR have been calculated from ChemDraw version 8.0, and values of B1 have been taken from the literature **[13].** In the activity term log ($1/EC_{50}$), EC_{50} refers to molar concentration of the compound achieving 50% inhibition of cytoplasmic HBV DNA synthesis, and in toxicity term log ($1/CC_{50}$), CC_{50} refers to the molar concentration of the compound leading to 50% extinction of HepG2.2.15 cells.

Results and Discussion:

A multiple regression analysis performed on anti-HBV and toxic data of Fu et al. on bis(L-amino acid) ester prodrugs and 9-[2-phosphonomethoxy)ethyl]adenine (**Table 1**) **(see supplementary material)** had revealed the correlations in Equations 1 and 2 **(see supplementary material)**. In these Equations, n is the number of data points, r is the correlation coefficient, r_{cv}^2 is the square of cross-validated correlation coefficient obtained by leave-one-out (LOO) jackknife procedure, F is the F-ratio between the variances of calculated and observed activities (data within parentheses refer to standard F-values at 99% level), and the figures with ± sign following the coefficients of the variables are 95% confidence intervals. Now Eq. 1 exhibits that the anti-HBV activity of this class of compounds is solely controlled by the lipophilicity of the molecules and has a simple linear dependence on ClogP. The r_{cv}^2 greater than 0.6 shows a good predictive ability of the Equation. However, the toxic data of this series of compounds was not found to be linearly dependent on ClogP, rather is shown to have a parabolic correlation (inverted parabola) with optimum ClogP value equal to zero. This means that the compound may have the lowest toxic effect when ClogP is zero. Further, Eq. 2 shows that the toxic effect of the compounds will be lower; the higher is its molar refractivity. The molar refractivity refers to the size of the molecule, so larger is the molecule, the lower would be its toxic effect. Thus we can have a good selective index of the compounds by increasing their bulkiness, maintaining their ClogP values. However, for the compounds of **Table 2 (see supplementary material)**, which are structurally very similar to those of **Table 1 (see supplementary material)**, the anti-HBV activity was found to have negative dependence on ClogP as well as on CMR (Eq. 3). If we compare the ClogP values of these compounds with ClogP values of compounds of **Table 1 (see supplementary material)**, we find that these values are much higher than those in **Table 1 (see supplementary material)**, and thus it can be concluded that higher values of ClogP may be detrimental to the activity. So there should be a limit to the value of ClogP. To find this thing, we combined both the series of compounds (**Table 3) (see supplementary material)** and obtained

the correlation as shown by Eq. 4. This Equation gives an optimum value of ClogP equal to 1.80. Thus, Equations 3 and 4 are formed **(see supplementary material).** The activity will have a positive dependence on ClogP only until its value does not reach 1.80. Beyond that, the activity will decrease as ClogP increases. In **Table 2 (see supplementary material)**, none of the compounds has its ClogP value less than 1.80, hence we got a negative correlation with ClogP. On the other hand, in **Table 1**, none of the compounds has its ClogP value greater than 1.80, hence in this case a positive correlation with ClogP was found. Since the lipophilicity of the compounds also depends on the size of the molecule, for highly hydrophobic molecules the molar refractivity has also cropped up to play a negative role in Equation 4 **(see supplementary material)** derived for the combine. For **Table 2 (see supplementary material)**, we were unable to correlate the toxic effects of the compounds with any parameters. The reason being that there was little span in the toxic data. However, when we combined them with that of compounds of **Table 1 (see supplementary material)**, a similar Equation, as found for **Table 1 (see supplementary material)** itself, was obtained (Equation 5 **in supplementary material)**.

This Equation, too, suggests that toxicity can be minimum when ClogP = 0, i.e., when there is a balance between the solubilities of the compound in lipid and aqueous phases, and it can be further controlled by controlling the size of the molecule. A control in size will also be beneficial to anti-HBV activity, but a limit of ClogP equal to zero will be slightly unsatisfactory for it. For better anti-HBV potency, ClogP should approach to 1.80 as Eq. 4 suggests. For the compounds of **Table 4 (see supplementary material)**, the anti-HBV activity was found to be correlated with Taft's steric parameter E_s, STERIMOL parameter B1 that describes the minimum width of the substituent, and an indicator parameters I_1 (Equation 6) **(see supplementary material)**. The indicator parameter I_1 has been used for regioisomerism of the compounds. For all the derivatives of **6**, $I_2 = 1$ and for all the derivatives of **7**, $I_2 = 0$. The positive coefficient of this parameter indicates that in general 6-substituted isomers **(6)** would be more active than 5- Substituted ones **(7)**. The reason of this may be that in the former the phenyl ring may be better oriented towards the active site of the receptor for interaction than in the latter. The occurrence of the E_s parameter in the Equation suggests that there would be some steric effect of the R-substituents and since E_s usually have negative values, the positive coefficient of it in the equation indicates that the bigger substituents may not be conducive to the activity. However, a positive coefficient of B1 suggests that the width of the substituent may be beneficial. The toxic effects for some of these compounds were also reported, but we could not find any good correlation for them. All the equations mentioned here represent statistically highly significant correlations, but in deriving each one, certain compounds that behaved as outliers were removed. These compounds are indicated in respective tables. However, no convincing reasons could be assigned for the aberrant behavior of these compounds in most of the cases. Outliers, in general, are a problem in QSAR, but sometimes they provide important clue to the drug design. An excellent and the only example of this here is Eq. 6, which has been derived for **Table 4 (see supplementary material)**. The outliers of this equation are compounds **6a** and **6i**, from which we can derive some important information. For compound **6a,** the observed activity is much lower than that predicted by Equation 6.

This compound has no substituent at the phenyl ring, so its low observed activity can be attributed to the absence of any substituent at the phenyl ring in this compound. And consequently it can be suggested that the substitution at the phenyl ring is essential for the good activity of the compound. Compound **6i**, which has an NO_2 group at the 3-position of the phenyl ring, also has a lower observed activity than the predicted one. The reason of this may be that nitro group may be involved in the formation of intramolecular hydrogen bond with the NH moiety in the bridge, a moiety that otherwise may presumably form a hydrogen bond with the receptor. Thus, this outlier indicates that the NH moiety may definitely be involved in the hydrogen bonding with the receptor, strengthening the drug-receptor interaction. Not only this, one can also assume that the whole carboxamide (-CONH-) or sulphonamide (-SO$_2$NH-) bridge group may be involved in the hydrogen bonding with the receptor, since CO and SO_2 are also able to do so acting as H-bond acceptors, while NH may act as H-bond donor.

Conclusion:

This QSAR analysis has shown that for certain series of anti-HBV agents, such as the series of novel bis(L-amino acid) ester prodrugs of 9-[2-(phosphonomethoxy)ethyl]adenine **(4)** and the similar series of compounds comprising of 2-amino-6-arylthio-9-[2-(phosphonoethoxy)ethyl]purine bis(2, 2, 2- trifluoroethyl) esters **(5)**, the hydrophobic property and the bulk of the molecules are very crucial. However, for the agents, such as 1-isopropylsulfonyl-2-amine benzimidazoles, this study has pointed out that the substitution at the phenyl ring and the steric properties of those substituents are highly important for their anti-HBV activity and that the hydrogen bonding ability of carboxamide or sulphonamide group, bridging the phenyl and benzimidazole rings, may be of further advantage because of their ability to form the hydrogen bonds with the receptor. Thus, the receptor of HBV inhibitors (HBV polymerase) can be assumed to possess an active site that may be capable of forming the hydrogen-bonds with the compounds.

References:

[1] EE Mast et al., *Vaccine* (1999) **17**: 1730-1733.
[2] D Lavanchy, *J Viral Hepat* (2004) **11**: 97-107.
[3] P Karayiannis, *Expert Rev Anti-Infect Ther* (2004) **2**: 745-760.
[4] R Kumar & B Aggarwal, *Curr Opin Investig Drugs* (2004) **5**: 171-178.
[5] F Zoulim, *J Hepatol (2005) 42*: 302-308.
[6] P Karayiannis, *J Antimicrob Chemother* (2003) **51**: 761 - 785.
[7] E De Clercq, *Clin Microb Rev* (2003) **16**: 569-596.
[8] K Hanazaki, *Curr Drug Targets – Inflamm Allergy* (2004), *3*, 63-70.
[9] X Fu et al., *Bioorg Med Chem Lett* (2007) **17**: 465-470.
[10] YF Li et al., *Eur J Med Chem* (2007) **42**: 1358-1364.
[11] K Sekiya et al. *J Med Chem* (2002) **45**: 3138-3142.
[12] A Verloop, The STERIMOL Approach to drug design, Marcel Dekker, New York, 1987.
[13] C Hansch and A Leo, Exploring QSAR : Fundamental and Applications in
[14] Chemistry and Biology, American Chemical Society, Washington, DC, 1995.

Supplementary material:

$$\log(1/EC_{50}) = 0.489(\pm0.143)ClogP + 6.069(\pm0.177)$$
$$n = 11, r = 0.933, r_{cv}^2 = 0.79, s = 0.26, F_{1,9} = 60.19(10.56) \tag{1}$$

$$\log(1/CC_{50}) = -0.104(\pm0.101)CMR + 0.282(\pm0.136)(ClogP)^2 + 3.391(\pm1.564)$$
$$n = 10, r = 0.935, r_{cv}^2 = 0.71, s = 0.23, F_{2,7} = 24.19(9.55), (ClogP)_{opt} = 0.0 \tag{2}$$

$$\log(1/EC_{50}) = -0.369(\pm0.347)ClogP - 0.571(\pm0.344)CMR + 15.159(\pm3.576)$$
$$n = 14, r = 0.905, r_{cv}^2 = 0.72, s = 0.24, F_{2,11} = 24.92(7.20) \tag{3}$$

$$\log(1/EC_{50}) = 0.595(\pm0.126)ClogP - 0.165(\pm0.045)(ClogP)^2 - 0.177(\pm0.091)CMR + 9.010(\pm1.347)$$
$$n = 25, r = 0.910, r_{cv}^2 = 0.65, s = 0.30, F_{3,21} = 33.69(4.87), (ClogP)_{opt} = 1.80 \tag{4}$$

$$\log(1/CC_{50}) = -0.158(\pm0.087)CMR + 0.071(\pm0.021)(ClogP)^2 + 4.664(\pm1.279)$$
$$n = 23, r = 0.923, r_{cv}^2 = 0.68, s = 0.30, F_{2,20} = 57.52(5.85), (ClogP)_{opt} = 0.0 \tag{5}$$

$$\log(1/EC_{50}) = 0.761(\pm0.403)E_s + 0.528(\pm0.423)B1 + 0.486(\pm0.357)I_1 + 5.057(\pm0.544)$$
$$n = 10, r = 0.925, s = 0.200, F_{3,6} = 11.77(9.78), r_{cv}^2 = 0.48 \tag{6}$$

Table 1: Bis (L-amino acid) ester prodrugs of 9-[2-(phosphonomethoxy) ethyl] adenine and their anti-HBV and toxic activities and physicochemical parameters **4**

Compd	R	X	n	ClogP	CMR	$\log(1/EC_{50})^a$			$\log(1/CC_{50})^a$		
						Obsd	Cald, Eq. 1	Pred, LOO	Obsd	Cald, Eq. 2	Pred, LOO
4a	Me	O	2	−1.75	11.94	4.90	5.21	5.34	2.73	3.01	3.15
4b	Isopropyl	O	2	0.10	13.79	5.88	6.12	6.14	2.59[c]	1.96	–
4c	2-Me-Pr	O	2	1.16	14.72	7.02	6.64	6.56	2.18	2.32	2.25
4d	Bz	O	2	1.08	16.96	6.51	6.60	6.61	2.59[c]	1.96	–
4e	H	S	2	−2.16	12.14	4.99	5.01	5.02	3.59	3.44	3.24
4f	Me	S	2	−1.54	13.24	5.81	5.31	5.15	2.62	2.63	2.69
4g	Isopropyl	S	2	0.31	15.10	6.12	6.22	6.23	1.54	1.82	1.92
4h	2-Me-Pr	S	2	1.37	16.02	6.67	6.74	6.75	2.47	2.32	2.20
4i	Bz	S	2	1.29	18.26	6.68	6.70	6.70	1.96	1.93	1.98
4j	Isopropyl	O	1	−0.46	12.86	6.52[b]	5.84	–	2.47	2.10	1.90
4k	2-Me-Pr	O	1	0.59	13.79	6.47	6.36	6.34	2.07	2.10	2.06
4l	Bz	O	1	0.51	16.03	6.18	6.32	6.33	1.81	1.75	1.79

[a]Taken from ref [9]. [b]Not included in the derivation of Eq. 1. [c]Not included in the derivation of Eq. 2.

Table 2: 2-Amino-6-arythio-9-[2-(phosphonomethoxy) ethyl]purine bis (2,2,2-trifluoroethyl) esters and their anti-HBV and toxic activities and physicochemical parameters.

Compd	R	ClogP	CMR	$\log(1/EC_{50})$			$\log(1/CC_{50})$
				Obsd[a]	Cald Eq. 3	Pred LOO	Obsd[a]
5a	SPh	2.95	11.45	7.30	7.54	7.62	NA[b]
5b	SPh(4-Me)	3.45	11.91	7.22	7.09	7.07	3.96
5c	SPh(3-Me)	3.45	11.91	7.04	7.09	7.10	3.82
5d	SPh(2-Me)	3.45	11.91	7.10	7.09	7.09	3.47
5e	SPh(4-Et)	3.98	12.38	6.40	6.63	6.67	3.77
5f	SPh(4-iPr)	4.38	12.84	6.01	6.22	6.29	3.77
5g	SPh(4-NO$_2$)	2.70	12.06	6.28[c]	7.28	–	3.42
5h	SPh(4-Cl)	3.67	11.94	6.92	6.99	7.00	4.21
5i	SPh(4-OMe)	2.87	12.06	7.52	7.22	7.15	NA
5j	SPh(3-OMe)	2.87	12.06	7.40	7.22	7.19	NA
5k	SPh(2-OMe)	2.47	12.06	7.10	7.37	7.56	NA
5l	SPh(4-OEt)	3.40	12.53	6.19[c]	6.75	–	NA
5m	SPh(4-O-nPr)	3.93	12.99	6.06	6.30	6.35	3.60
5n	SPh(4-O-iPr)	3.71	12.99	6.54	6.38	6.34	3.45
5o	SPh(4-O-nBu)	4.46	13.46	NA	NA	–	3.58
5p	SPh(4-O-iBu)	4.11	13.46	6.07	5.96	5.89	4.00
5q	SPh(4-O-CF$_3$)	3.98	12.11	7.22	6.78	6.66	3.09
5r	S-(2-naphthyl)	4.13	13.14	7.04[c]	6.13	–	3.70

[a]Taken from ref [10]. [b]NA: Not available. [c]Not included in the derivation of Eq. 3.

Table 3: A combine of **Tables 1** and **2**

Compd	ClogP	CMR	$\log(1/EC_{50})$			$\log(1/CC_{50})$		
			Obsd	Cald Eq. 4	Pred LOO	Obsd	Cald Eq. 5	Pred LOO

4a	−1.75	11.94	4.90	5.35	5.53	2.73	2.99	3.05
4b	0.10	13.79	5.88[a]	6.63	–	2.59	2.49	2.48
4c	1.16	14.72	7.02	6.87	6.66	2.18	2.44	2.46
4d	1.08	16.96	6.51	6.45	6.45	2.59	2.07	1.92
4e	−2.16	12.14	4.99	4.81	4.66	3.59	3.07	3.02
4f	−1.54	13.24	5.81	5.36	5.23	2.62	2.74	2.76
4g	0.31	15.10	6.12	6.50	6.54	1.54[c]	2.29	–
4h	1.37	16.02	6.67	6.68	6.68	2.47	2.27	2.24
4i	1.29	18.26	6.68	6.27	6.03	1.96	1.90	1.86
4j	−0.46	12.86	6.52	6.42	6.40	2.47	2.65	2.69
4k	0.59	13.79	6.47	6.86	6.91	2.07	2.51	2.57
4l	0.51	16.03	6.18	6.43	6.46	1.81	2.16	2.21
5a	2.95	11.45	7.30	7.30	7.31	NA	NA	NA
5b	3.45	11.91	7.22	6.99	6.98	3.96	3.63	2.85
5c	3.45	11.91	7.04	6.99	6.99	3.82	3.63	3.61
5d	3.45	11.91	7.10	6.99	6.98	3.47	3.63	3.65
5e	3.98	12.38	6.40	6.58	6.59	3.77	3.83	3.83
5f	4.38	12.84	6.01	6.18	6.25	3.77	3.99	4.05
5g	2.70	12.06	6.28[a]	7.29	–	3.42	3.28	3.26
5h	3.67	11.94	6.92	6.86	6.85	4.21	3.74	3.68
5i	2.87	12.06	7.52	7.22	7.19	NA	NA	NA
5j	2.87	12.06	7.40	7.22	7.20	NA	NA	NA
5k	2.47	12.06	7.10	7.34	7.38	NA	NA	NA
5l	3.40	12.53	6.19[a]	6.90	–	NA	NA	NA
5m	3.93	12.99	6.06	6.51	6.58	3.60	3.71	3.72
5n	3.71	12.99	6.54	6.65	6.29	3.45	3.59	3.60
5o	4.46	13.46	NA[b]	NA	NA	3.58	3.95	4.04
5p	4.11	13.46	6.07	6.29	6.36	4.00	3.74	3.71
5q	3.98	12.11	7.22	6.63	6.56	3.09	3.88[c]	–
5r	4.13	13.14	7.04[a]	6.34	–	3.70	3.80	3.82

[a]Not included in the derivation of Eq. 4. [b]NA: Not available. [c] Not included in the derivation of Eq. 5.

Table 4: 1-Isopropylsulfonyl-2-amine benzimidazoles and their anti-HBV and toxic activities and physicochemical parameters

Compd	X	R	E_s	B_1	I_1	$\log(1/EC_{50})$		
						Obsd[a]	Calcd Eq. 6	Pred LOO
6a[b]	CO	H	0.00	1.00	1.00	4.47	6.07	–
6b	CO	2-F	−0.46	1.35	1.00	6.03	5.91	5.88
6c	CO	4-F	−0.46	1.35	1.00	5.92	5.91	5.90
6d	CO	2,6-DiF	−0.92	2.70	1.00	6.15	6.27	7.53
6e	CO	2-Cl	−0.97	1.80	1.00	6.09	5.75	5.68
6f	CO	4-CH₃	−1.24	1.52	1.00	5.38	5.40	5.41
6g	SO₂	H	0.00	1.00	1.00	5.82	6.07	6.28
6h	SO₂	4-CH₃	−1.24	1.52	1.00	5.32	5.40	5.44
6i[b]	SO₂	3-NO₂	−1.01	1.70	1.00	4.98	5.67	–
7a	CO	H	0.00	1.00	0.00	5.70	5.85	5.47
7b	CO	4-F	−0.46	1.35	0.00	5.43	5.42	5.41

[a]Taken from ref [11]. [b]Not included in the derivation of Eq. 6.