Bioorganic & Medicinal Chemistry Letters, Vol 2, No.9, pp. 1165-1169, 1992
Printed in Great Britain

0960-894X/92 $5.00 + .00

QSAR OF HIV INHIBITORS

Corwin Hansch* and Litai Zhang

Department of Chemistry, Pomona College, Claremont, CA 91711

(*Received* 7 *May* 1992)

Abstract: Quantitative structure-activity relationships (QSAR) have been derived for two sets of HIV inhibitors: 1-[2-hydroxyethoxy)methyl]-6-(X-phenylthio)thymines and thiothymines and dideoxypurine ribofuranosides. It is found that the hydrophobic and especially the steric properties of substituents determine relative potencies of the derivatives.

In recent years, efforts to find effective drugs for the disease AIDS caused by the HIV infection have grown enormously. As the prospect of millions dying from the pandemic of AIDS becomes evident, a great increase in the support of research is occurring. The QSAR methodology should be of help in the efficient design of drugs. We illustrate the value of such analysis using the data in Tables I and II and compare the results with one earlier study.

From the data[1] in Table I, eq. 1 has been derived.

$$\log 1/C = 0.88\ (\pm.39)\ \Sigma\pi + 12.1\ (\pm3.8)\ L_{R}' - 1.59\ (\pm.49)\ (L_{R}')^2 + 1.17\ (\pm.86)\ B_1\text{-}3R + 1.53\ (\pm.82)\ E_s\text{-}2R - 15.3\ (\pm7.4) \quad (1$$

n = 33, r = 0.941, s = 0.500, ideal L_{R}' = 3.8 (3.6 to 3.9), all terms are significant at >0.95 level according to the F test.

In this expression, C is the molar concentration of drug affording 50% protection to MT-4 cells, $\Sigma\pi$ represents $\pi R' + \pi R$, L is the sterimol parameter[2] for the length of the substituent, B_1-3R is the sterimol parameter for the width of R in the 3-position (when R's are present in both 3 and 5 positions, only 1 B_1-3R is used, i.e. only one meta substituent appears to show the positive steric effect). E_s -2R is the Taft[3] steric constant for ortho substituents. The most important single variable is $\Sigma\pi$, however, the two terms in L_{R}' account for 75% of the variance, while $\Sigma\pi$ alone accounts for only 49%. These figures are somewhat misleading since for R', π and L are highly collinear, but $\Sigma\pi$ is not collinear with L. the important point is that the coefficient with $\Sigma\pi$ is in the range (0.5 to 1.1) normally found for inhibitors in cell culture.[4] Both R and R' appear to contribute in the same way after correction has been made for the length of R' and the width of R. Ortho substituents on the phenyl ring depress inhibitory potency as brought out by E_s-2R. The B_1-3R term is marginal since this term contributes the least to eq. 1. As eq. 1 is linear in $\Sigma\pi$ it would seem that the best way to increase activity is to make more hydrophobic derivatives, but the most hydrophobic analogues in Table I (27 and 33) have calculated log P of about 3.7. There is considerable evidence that log P of about 2 is ideal for CNS penetration[5] so that reduction of hydrophobicity in so far as it is simply related to cell penetration is in order. What is not clear from eq. 1 is, is π associated only with cell penetration or are specific hydrophobic receptor interactions possible? This could be explored by replacing a CH by N or N-O at the meta position of the C_6H_5 moiety and increasing the hydrophobicity of the other meta substituent. The N fragment would decrease log P by 1.5 and the N-O would lower it by 3.8.

Except for the $\Sigma\pi$ terms all of the other terms account for steric effects which brings out the critical fit of the ligands to their receptors.

From the data[6] in Table II, QSAR 2 has been derived.

$$\log \% = 0.78\ (\pm.22) \log P + 0.90\ (\pm.24)\ E_s\text{-}X + 2.19\ (\pm.22) \quad (2$$
$$n = 10,\ r = 0.966,\ s = 0.085$$

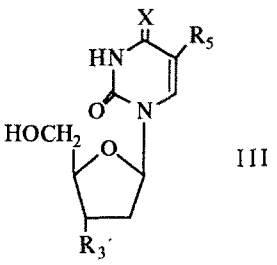
The dependent variable is the % protection of the ATH8 cells from HIV infection in the presence of 2 μM nucleoside. The two important factors brought out by eq. 2 are that the hydrophobic effect is essentially the same as for eq. 1 and that there is a negative steric effect in the 6-position. In contrast to eq. 1 it seems highly likely that one could increase activity considerably and at the same time increase the probability of the chemicals penetrating the CNS by making analogs with log P near 2. The importance of designing drugs capable of penetrating the CNS for the control of AIDS has been discussed by Lien, et al.[7] as well as others. If this could be done without introducing steric problems, an increase of about ten fold in activity is to be expected. The 2- and 6-positions are not suitable positions as the E_s-X term shows and the fact that of the three substituents in the 2 position (H, NH_2, Cl) the largest (Cl) results in a compound of very low activity. It is over ten times less active than predicated. Another point to consider is the evidence that amino and halogen substituents in the 6-position of II undergo hydrolysis which appears to be essential for their biological activity.[6] Thus, the steric term in eq. 2 may be associated with this process. The 2-NH_2, 6-SH congener in Table II is well predicated as is the 6-SH analogue. However, the former is of doubtful activity and the latter is clearly inactive. If the 2-NH_2, 6SH compound is omitted and QSAR 2 rederived essentially the same coefficients are obtained, but r = 0.921 because of the reduction in the variance. Although it is a satisfaction that these two compounds are predicted to have very low activity on the basis of the size of the sulfur atom, this lack of activity could also be due to their high toxicity which might or might not have anything to do with the size of sulfur.

The most likely position for increasing activity would be substitution on the furan ring with a more hydrophobic group. In this respect it is of interest to compare ethyl with the often used azide group.

N_3: $\pi = 0.46$, $L = 4.62$; $\sigma^* = 0.42$. C_2H_5: $\pi = 1.02$; $L = 4.11$; $\sigma^* = -0.05$.

If the inductive effect is important the N_3 might be superior, but if a shorter group would be advantageous then the more lipophilic ethyl group would be interesting.

Mahmoudian[8] has also derived a QSAR (eq. 3) for variations of III.



X = O or NH

III

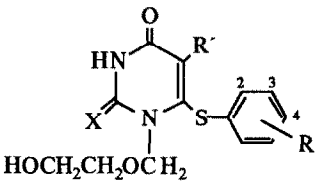
$$-\log EC_{50} = 1.02\ \pi_5 - 0.21\ L_5^2 + 1.79\ B_{3,5} + 0.49\ B_{4,3'} + 2.87 \quad (3$$

$$n = 23,\ r = 0.777,\ s = 0.78$$

In this expression π does not apply to substituents in position 3′. The effects of these substituents appear to be accounted for by steric effects alone. Although Mahmoudian shows graphically a good correlation between activity of four examples where R_5 = H and R_3' = H, NH_2, I and N_3 and B_4 we suspect more is involved especially in light of the importance of hydrophobicity terms in QSAR 1 and 2. Although the quality of correlation by eq. 3 is low, it is of interest that the π term is in agreement with the other two QSAR.

A general way in which QSAR are used to develop new lead compounds is to study the modification of those compounds which turn out to be more active than predicted (eg. #22 in Table I). It would seem that the allyl group is considerably more active than one would expect from its rather large L value. Hence, exploration of this substituent with substituents on the phenyl ring would be of interest. In Table II, there are no large positive outliers. The most obvious route to more active compounds would be increase log P of #4 by 1.5 to 2.0 units. This might be accomplished by adding substituents with positive π values to the furan ring.

Table I: Parameters used to derive equation 1.



I

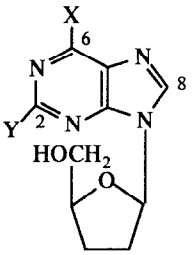
No.	Substituents			Obsd.[1] log 1/C	Calcd. log 1/C	L	B_1-3R	$\Sigma\pi$	E_s-2
	X	R	R′						
1.	O	2-Me	Me	4.15	3.90	2.87	1.00	1.12	-1.24
2.	O	2-NO_2	Me	3.85	4.08	2.87	1.00	.28	-1.01
3.	O	2-OMe	Me	4.72	4.85	2.87	1.00	.54	-0.55
4.	O	3-Me	Me	5.59	5.44	2.87	1.52	1.12	00
5.	O	3-Et	Me	5.57	5.84	2.87	1.52	1.58	00
6.	O	3-t-C_4H_9	Me	4.92	5.52	2.87	2.60	2.54	00
7.	O	3-CF_3	Me	4.35	5.22	2.87	1.99	1.40	00
8.	O	3-F	Me	5.48	5.26	2.87	1.35	.70	00
9.	O	3-Cl	Me	4.89	5.27	2.87	1.80	1.27	00
10.	O	3-Br	Me	5.24	5.24	2.87	1.95	1.42	00
11.	O	3-I	Me	5.00	5.26	2.87	2.15	1.68	00
12.	O	3-NO_2	Me	4.47	4.52	2.87	1.70	.28	00
13.	O	3-OH	Me	4.09	4.55	2.87	1.35	-.11	00
14.	O	3-OMe	Me	4.66	5.12	2.87	1.35	.54	00
15.	O	4-Me	Me[2]	3.66	6.00	2.87	1.00	1.12	00
16.	O	3,5-di-Me	Me	6.59	5.93	2.87	1.52	1.68	00
17.	O	3,5-di-Me	Me	5.89	5.88	2.87	1.80	1.96	00

Table I (cont.)

No.	Substituents			Obsd.[1] log 1/C	Calcd. log 1/C	L	B_1-3R	$\Sigma\pi$	E_s-2
	X	R	R´						
18.	S	3,5-di-Me	Me	6.66	5.93	2.87	1.52	1.68	00
19.	O	3-COOMe	Me	5.10	4.82	2.87	1.64	.55	00
20.	O	3-COMe	Me	5.14	4.39	2.87	1.60	.01	00
21.	O	3-CN	Me	5.00	4.37	2.87	1.60	-.01	00
22.	O	H	Allyl	5.60	4.71	5.11	1.00	1.10	00
23.	O	H	Et	6.96	7.18	4.11	1.00	1.02	00
24.	S	H	C_3H_7	5.00	5.83	4.92	1.00	1.55	00
25.	S	H	$CHMe_2$	7.23	7.63	4.11	1.00	1.53	00
26.	S	3,5-di-Me	Et	8.11	7.60	4.11	1.52	2.14	00
27.	S	3,5-di-Me	$CHMe_2$	8.30	8.04	4.11	1.52	2.65	00
28.	S	3,5-di-Cl	Et	7.37	7.56	4.11	1.80	2.44	00
29.	O	H	Et	6.92	7.18	4.11	1.00	1.02	00
30.	O	H	C_3H_7	5.47	5.83	4.92	1.00	1.55	00
31.	O	H	CHMe	7.20	7.62	4.11	1.00	1.53	00
32.	O	3,5-di-Me	Et	7.89	7.60	4.11	1.52	2.14	00
33.	O	3,5-di-Me	$CHME_2$	8.57	8.04	4.11	1.52	2.65	00
34.	O	3,5-di-Cl	Et	7.85	7.56	4.11	1.80	2.44	00

1. Ref. 1.
2. This compound not used to derive eq. 1.

Table II: Parameters used to derive equation 2.



II

No.	Substituent	obsd.[1] log %	Calcd. log %	log P	E_s-X
1.	2-NH_2, 6-I	1.32	1.33	0.52	-1.40
2.	6-Br	1.40	1.41	0.35	-1.16
3.	2-NH_2, 6-SH	0.78	0.76	-0.59	-1.07
4.	6-F	1.81	1.77	0.00	-0.46
5.	2,6-di-NH_2	1.23	1.28	-0.46	-0.61
6.	6-Cl	1.45	1.50	0.24	-0.97

Table II (cont.)

No.	Substituent	obsd.[1] log %	Calcd. log %	log P	E_s-X
7.	2-NH_2, 6-F	1.68	1.74	-0.05	-0.46
8.	2-NH_2, 6-Cl	1.58	1.48	0.21	-0.97
9.	6-I	1.23	1.34	0.53	-1.40
10.	2-NH_2, 6-Br	1.54	1.41	0.34	-1.16
11.	6-SH[2]	inactive	0.71	-0.66	-1.07
12.	2-Cl, 6-NH_2[2]	0	1.35	-0.37	-0.61

1. % Protection of ATH8 cells by 2 μM concentration of nucleoside. From reference 6.
2. These data points not used in the derivation of eq. 2.

Acknowledgment

This research was supported by a grant for studies in fundamental toxicology from the R.J. Reynolds Company. We are indebted to John Driscoll of the NCI for a number of helpful suggestions used to prepare this report.

References

1. Tanaka,H.; Takashima,H.; Ubasawa,M.; Sekiya,K.; Nitta,I.; Baba,M.; Shigeta,S.; Walker,R.T.; DeClercq,E.; Miyasaka,T. *J. Med. Chem.* **1992**, 35, 337.
2. Silipo,C.; Vittoria,A. *Comprehensive Medicinal Chemistry*; Ramsden, C.A., Ed.; Pergamon Press: Oxford, 1990; Vol 4, p. 153.
3. Unger,S.H.; Hansch,C. *Prog. Org. Chem.* **1976**, 12, 91.
4. Hansch,C; Kim,D.; Leo,A.; Novellino,E.; Silipo,C.; Vittoria,A. *CRC Critical Rev. Toxicol.* **1989**, 19, 185.
5. Hansch,C.; Björkroth,J.P.; Leo,A. *J. Pharm. Sci.* **1987**, 76, 663.
6. Shirasaka,T.; Murakami,K.; Ford,H.Jr.; Kelley,J.A.; Yoshioka,H.; Kojima,E.; Aoki,S.; Broder,S.; Mitsuya,H. *Proc. Natl. Acad. Sci. USA* **1990**, 87, 9426.
7. Lien,E.J.; Gao,H.; Prabhakar,H. *J. Pharm. Sci.* **1991**, 80, 517.
8. Mahmoudian,M. *Pharm. Res.* **1991**, 8, 43.