J. Enzyme Inhibition, 2001, Vol. 16, pp. 185–197
Reprints available directly from the publisher
Photocopying permitted by license only

Quantitative Structure–Activity Relationship of some HIV-1 Protease Inhibitors: A Fujita–Ban Type Analysis

S.B. MEKAPATI[a], R. SIVAKUMAR[b] and S.P. GUPTA[c,*]

[a]*Department of Chemistry, Pomona College, Claremont, CA 91711, USA;* [b]*Department of Pharmacy;* [c]*Department of Chemistry, Birla Institute of Technology and Science, Pilani – 333031, India*

(Received 7 November 2000)

A Fujita–Ban type analysis has been made on a few series of HIV-1 (human immunodeficiency virus of type 1) protease inhibitors and the activity contributions of various substituents obtained. From these activity contributions, a compound is predicted that may have better activity than ritonavir, presently prescribed for the treatment of patients suffering from HIV-1. A few other compounds are also suggested.

Keywords: QSAR, HIV-1 protease, Fujita–Ban, ritonavir

INTRODUCTION

The replicative cycle of human immunodeficiency virus of type 1 (HIV-1), which is mainly responsible for acquired immunodeficiency syndrome (AIDS), presents several viable targets which can be exploited for the development of anti-HIV drugs. Ideally, an anti-HIV drug should arrest the virulence and further infection of healthy cells without displaying toxicity towards normal cellular physiology.

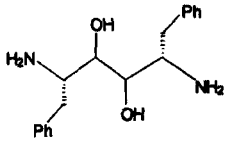
The HIV encodes an aspartyl protease (Pr), which is a homodimeric enzyme that cleaves the polyprotein products of gag and gag-pol viral genes, yielding structural proteins and enzymes that are essential to the life cycle of the virus. Inhibition of this enzyme leads to the production of non-infectious viral particles[1,2] and thus to the prevention of further propagation of the virus. Since abundant structural informations are available on this enzyme, it has become an attractive target for computer-aided drug design strategies,[3,4] and consequently a prime focus for the development of anti-HIV chemotherapy.[5]

A number of peptide-derived compounds, reflecting the structure of polyprotein substrates of HIV-Pr, have been identified as HIV-Pr inhibitors,[6] but their clinical development has been hindered by their poor pharmacokinetics, including low oral bioavailability and rapid excretion,[7] and complex and expensive synthesis.[8] Therefore, attempts have been made to develop modified peptidic inhibitors and

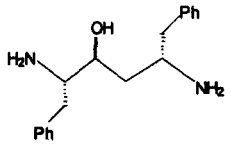
*Corresponding author.

substantial progress in the identification of agents with high oral bioavailability has been made.[9–12]

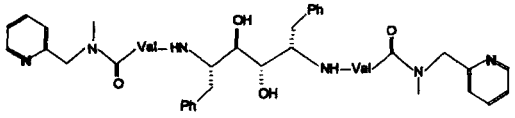
Since HIV protease exists as a C_2-symmetric homodimer, a variety of peptidic inhibitors that were developed were based on this C_2-symmetric structure. Kempf *et al.*[13,14] reported a series of inhibitors based on a symmetric core diamine (**1**) and a pseudo symmetric core diamine (**2**). Structure–activity studies on the derivatives of (**1**) and (**2**) led Kempf *et al.*[14,15] to identify two compounds (**3**) (A-77003) and (**4**) (A-80987), possessing adequate anti-HIV activity; of these two, A-80987 was found to possess better oral bioavailability in both animal models and in humans.[9] Further studies on the analogues of A-80987 (**5–7**) led to the development of ritonavir (ABT-538) (**8**) which possesses high oral bioavailability in both animals and humans and a substantially reduced rate of metabolism.[9,16] For further investigation, the analogues of ritonavir (**9**) were also studied.[16] We present here a more systematic and quantitative study on the structure–activity relationships of various analogues of A-80987 (**5–7**) and those of ritonavir (**9**) in order to investigate the possibility of designing still better analogues.



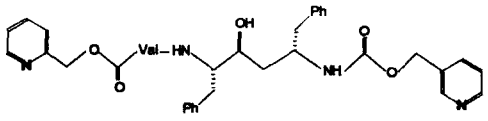
1



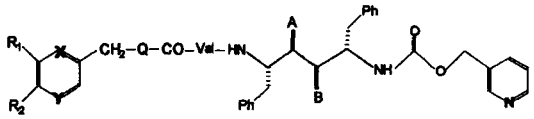
2



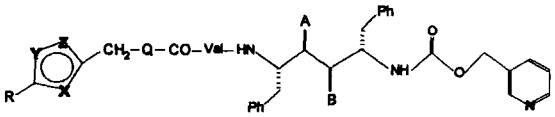
3 (A-77003)



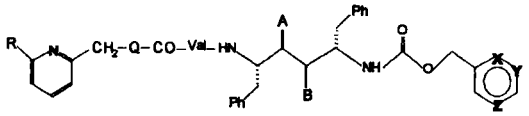
4 (A-80987)



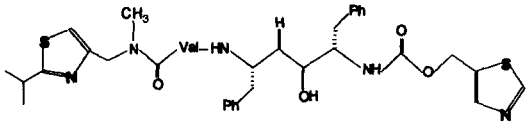
5 (A-80987)



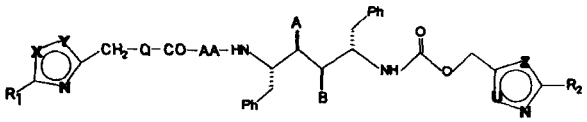
6



7



8 (ABT-538)



9

MATERIALS AND METHOD

The various series of analogues of A-80987 and a series of analogues of ritonavir that are listed in Tables I–IV have been taken from Kempf *et al.*[16] Kempf *et al.* evaluated the anti-HIV activity of these compounds in terms of the ability of the compound to block the spread of HIV-1 in the immortalized human T-cell line MT4 by measuring the cytopathic effect of the virus in those cells

by uptake of a tetrazolium dye. The EC_{50} values listed in the Tables refer to the molar concentrations of the compound producing 50% effect. Kempf *et al.*[16] had also studied the cytotoxic effects of the compounds in terms of $CCIC_{50}$, the concentration of the compound required to reduce by 50% the number of mock-infected MT4 cells.

We analyzed the structure–activity relationships of these compounds, using the Fujita–Ban method,[17] where the activity contribution of each substituent or moiety can be obtained. In this method the total activity of a compound is given by

$$\text{Activity} = \sum_i \alpha_i \chi_i + \mu \qquad (1)$$

where α_i is the activity contribution of the ith substituent relative to H or the substituent defined in the parent structure, and χ_i is a parameter which takes a value of 1 or 0 depending on the presence or absence of the ith substituent in the molecule. The constant μ is the activity of the parent structure of the molecule.

In the present case, the substitutions occurred within the rings and in the bridge groups also. Therefore, in each case the parent structure needs to be specified:

5: $R_1 = R_2 = H$; (X = N, Y = CH); Q = O; (A = OH, B = H)
6: R = H; (X = N, Y = S, Z = CH); Q = O; (A = OH, B = H)
7: R = H; Q = O; (X–Y = S, Z = N); (A = OH, B = H)
9: $R_1 = H$; (X = S, Y = CH); Q = O; (Z = S, U = CH); AA = Val; (A = OH, B = H)

Deviations from such parent structures have been parameterized. For the substituents on the rings, referred to by R, R_1, or R_2, the values of χ as defined in Equation (1) have been used. For the remainder, we have defined the values of χ as follows.

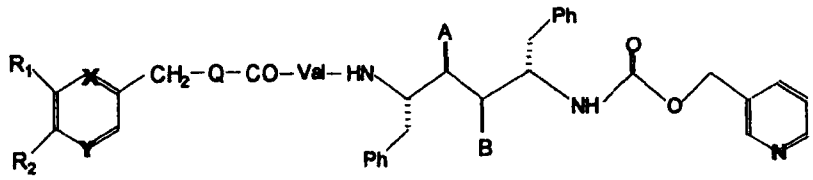
1. χ is equal to 1 in all structures for A = H, B = OH and 0 for A = OH, B = H as defined in the parent structures.
2. χ is equal to 1 in all structures if Q is NH, *N*-alkyl, CH_2, or CH_2O. It is zero for Q = O (the parent structure).
3. For the ring of **5**, $\chi = 1$ if X = CH, Y = N, or if X = Y = N. It is zero if X = N, Y = CH (parent structure).
4. For the ring of **6**, $\chi = 1$ if X = N, Y = O, Z = CH, or if X = N, Y = CH, Z = S, or if X = S, Y = N, Z = CH, or if X = CH, Y = O, Z = N, or if X = O, Y = N, Z = CH, or if X = CH, Y = N, $Z = C(CH_3)$. It will be zero for X = N, Y = S, Z = CH (parent structure).
5. For the ring of **7**, $\chi = 1$ if X–Y = CH, Z = O, or if X–Y = CH–N, Z = N. It is zero for X–Y = S, Z = N (parent structure).
6. For the rings of **9**, left hand side ring: $\chi = 1$ if X = O, Y = CH, or if X = CH, Y = S. It is zero, if X = S, Y = CH (parent structure). For right hand side ring: $\chi = 1$ if Z = O, U = CH, or if Z = CH, U = O. It is zero for Z = S, U = CH (parent structure).
7. In **9**, $\chi = 1$ for AA = Ala in the bridge group. It is zero for AA = Val (parent structure).

RESULTS AND DISCUSSION

For the compounds in Table I, the Fujita–Ban analysis revealed the contribution of substituents to anti-HIV activity of the compounds as shown in Table V. The figures within parentheses are 95% confidence intervals. The Table also reports the number of compounds (n) used in the analysis and the values obtained for the correlation coefficient (r), the standard deviation (s), and the F-statistics (F). Certain compounds as indicated in the Table (Table I) were not included in the analysis as they were exhibiting aberrant behaviour.

Table V shows that the activity contributions of certain substituents (indicated by asterisk) are statistically insignificant at 95% confidence intervals. Hence ignoring these, the activity contributions were reanalyzed and the results were those

TABLE I P3 Pyridine analogues (5) and their anti-HIV-1 potency and cytotoxic activity studied by Kempf *et al.*[16]

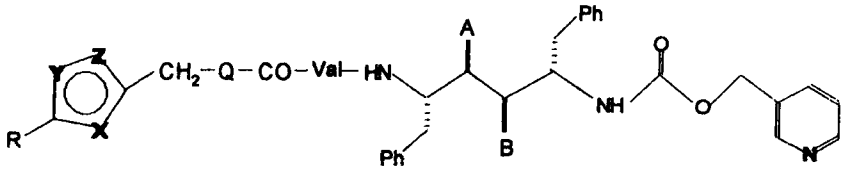


No.	R_1	R_2	X	Y	Q	A	B	log (1/EC_{50})			log (1/$CCIC_{50}$)	
								Obsd	Calcd[a]	Calcd[b]	Obsd	Calcd[c]
1	H	H	N	CH	O	OH	H	6.69	6.72	6.88	–	–
2	H	H	N	CH	NCH_3	H	OH	6.24	6.37	6.54	–	–
3	CH_3	H	N	CH	O	OH	H	7.00	7.00	7.04	–	–
4	CH_3	H	N	CH	O	H	OH	7.29	7.18	7.21	–	–
5	CH_3	H	N	CH	NCH_3	OH	H	6.48	6.47	6.53	4.09	4.16
6	CH_3	H	N	CH	NCH_3	H	OH	6.59	6.65	6.70	4.23	4.16
7	CH_3	H	N	CH	NH	OH	H	6.52	6.34	6.39	4.13	4.16
8	CH_3	H	N	CH	NH	H	OH	6.47	6.52	6.56	–	–
9	Et	H	N	CH	O	OH	H	7.72	7.46	7.45	4.09	4.16
10	Et	H	N	CH	O	H	OH	7.64	7.64	7.62	–	–
11	Et	H	N	CH	NCH_3	OH	H	7.17	6.93	6.94	4.24	4.16
12	Et	H	N	CH	NCH_3	H	OH	6.60	7.11	7.12	4.23	4.16
13	*i*-Pr	H	N	CH	O	OH	H	7.31	7.27	7.26	–	–
14	*i*-Pr	H	N	CH	O	H	OH	7.57	7.45	7.43	4.70	4.71
15	*i*-Pr	H	N	CH	NCH_3	OH	H	6.75	6.74	6.76	4.72	4.71
16	*i*-Pr	H	N	CH	NCH_3	H	OH	6.75	6.92	6.93	4.72	4.71
17	*t*-Bu	H	N	CH	O	OH	H	7.19	6.97	6.88	4.72	4.71
18	*t*-Bu	H	N	CH	O	H	OH	7.26	7.15	7.05	4.69	4.71
19	*t*-Bu	H	N	CH	NCH_3	OH	H	6.22	6.44	6.37	4.72	4.71
20	*t*-Bu	H	N	CH	NCH_3	H	OH	6.51	6.62	6.54	4.72	4.71
21	CH3	H	CH	N	O	OH	H	7.04	7.12	7.04	–	–
22	CH3	H	CH	N	O	H	OH	7.29	7.30	7.21	–	–
23	CH3	H	CH	N	NCH_3	OH	H	6.32	6.59	6.53	–	–
24	CH3	H	CH	N	NCH_3	H	OH	6.92	6.77	6.70	–	–
25	H	CH_3	CH	N	O	OH	H	7.06	6.99	6.88	–	–
26	H	CH_3	CH	N	O	H	OH	7.24	7.17	7.05	–	–
27	H	CH_3	CH	N	NCH_3	OH	H	6.39	6.46	6.37	–	–
28	H	CH_3	CH	N	NCH_3	H	OH	6.89	6.64	6.54	–	–
29	H	CH_3	CH	N	NH	OH	H	6.20	6.33	6.24	–	–
30	H	OCH_3	N	CH	O	OH	H	7.41	7.39	7.44	4.19	4.16
31	H	OCH_3	N	CH	O	H	OH	7.31	7.57	7.61	4.13	4.16
32	H	OCH_3	N	CH	NCH_3	OH	H	7.00	6.86	6.94	4.22	4.16
33	H	OCH_3	N	CH	NCH_3	H	OH	7.12	7.04	7.11	4.06	4.16
34	OCH_3	H	CH	N	O	OH	H	7.22	7.16	7.15	–	–
35	OCH_3	H	CH	N	O	H	OH	7.34	7.34	7.32	–	–
36	OCH_3	H	CH	N	NCH_3	OH	H	6.57	6.63	6.64	–	–
37	OCH_3	H	CH	N	NCH_3	H	OH	6.80	6.81	6.82	–	–
38	H	OCH_3	CH	N	O	OH	H	7.27	7.51	7.44	–	–
39	H	OCH_3	CH	N	O	H	OH	7.72	7.69	7.61	4.26	4.16
40	H	OCH_3	CH	N	NCH_3	OH	H	7.20	6.98	6.94	4.24	4.16
41	H	OCH_3	CH	N	NCH_3	H	OH	7.17	7.16	7.11	4.22	4.16
42	H	NH_2	N	CH	O	OH	H	6.07	6.38	6.37	–	–
43	H	NH_2	N	CH	O	H	OH	6.16	6.56	6.54	–	–
44	H	NH_2	N	CH	NCH_3	OH	H	6.11	5.85	5.86	–	–
45	H	NH_2	N	CH	NCH_3	H	OH	6.47	6.03	6.04	–	–
46	NH_2	H	CH	N	O	OH	H	6.46	6.36	6.36	–	–

47	NH_2	H	CH	N	O	H	OH	6.43	6.53	6.53	–	–
48	H	NH_2	CH	N	O	OH	H	5.12[d]	6.78	6.37	4.05	4.16
49	H	NH_2	CH	N	O	H	OH	5.02[d]	6.68	6.54	–	–
50	H	H	N	N	O	OH	H	6.70	6.68	6.88	–	–
51	H	H	N	N	O	H	OH	6.96	6.85	7.05	–	–
52	CH_3	H	N	N	O	OH	H	7.01	6.95	7.04	–	–
53	CH_3	H	N	N	O	H	OH	7.15	7.13	7.21	4.23	4.16
54	CH_3	H	N	N	NCH_3	OH	H	5.84[d]	6.42	6.53	–	–
55	CH_3	H	N	N	NCH_3	H	OH	6.59	6.60	6.70	–	–
56	H	CH_3	N	N	O	OH	H	6.66	6.82	6.88	–	–
57	H	CH_3	N	N	O	H	OH	7.00	7.00	7.05	4.02	4.16
58	CH_3	CH_3	N	N	O	OH	H	7.15	7.10	7.04	–	–
59	CH_3	CH_3	N	N	O	H	OH	7.41	7.28	7.21	–	–
60	CH_3	CH_3	N	N	O	OH	H	6.18	6.57	6.53	–	–
61	CH_3	CH_3	N	N	O	H	OH	6.92	6.75	6.70	–	–

[a] Using the set 1 activity contributions of substituents as given in Table V.
[b] Using the set 2 activity contributions of substituents as given in Table V.
[c] Using activity contributions of Table IX.
[d] Not used in derivation of set 1 and set 2 values of Table V.

TABLE II P3 Five-membered heterocyclic analogues (6) and their antiviral potency and cytotoxicity data studied by Kempf *et al.*[16]



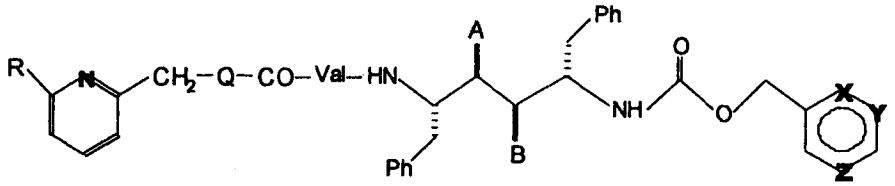
No.	R	X	Y	Z	Q	A	B	log (1/EC_{50})			log (1/$CCIC_{50}$)	
								Obsd	Calcd[a]	Calcd[b]	Obsd	Calcd[c]
1	H	N	S	CH	NCH_3	OH	H	6.13	6.14	6.48	4.08	4.13
2	H	N	S	CH	NCH_3	H	OH	6.31	6.30	6.67	–	–
3	CH_3	N	S	CH	NCH_3	OH	H	6.68	6.63	6.48	–	–
4	CH_3	N	S	CH	NCH_3	H	OH	6.74	6.79	6.67	–	–
5	Et	N	S	CH	O	OH	H	7.19	7.17	7.09	4.23	4.26
6	Et	N	S	CH	O	H	OH	7.38	7.34	7.27	4.19	4.26
7	Et	N	S	CH	NCH_3	OH	H	6.54	6.59	6.48	4.22	4.13
8	Et	N	S	CH	NCH_3	H	OH	6.59	6.76	6.67	4.22	4.13
9	*i*-Pr	N	S	CH	O	OH	H	7.96	7.76	7.72	–	–
10	*i*-Pr	N	S	CH	O	H	OH	8.00	7.92	7.90	4.72	4.60
11	*i*-Pr	N	S	CH	NCH_3	OH	H	6.54[d]	7.45	7.11	4.49	4.47
12	*i*-Pr	N	S	CH	NCH_3	H	OH	7.54	7.35	7.29	4.33	4.47
13	*t*-Bu	N	S	CH	NCH_3	OH	H	6.43	6.43	6.48	4.72	4.75
14	*t*-Bu	N	S	CH	NCH_3	H	OH	7.54[d]	6.86	6.66	4.77	4.75
15	*i*-Pr	N	O	CH	O	OH	H	7.75	7.74	7.72	–	–
16	*i*-Pr	N	O	CH	O	H	OH	7.70	7.91	7.90	4.28	4.32
17	*i*-Pr	N	O	CH	NCH_3	OH	H	7.48	7.17	7.11	4.21	4.19
18	*i*-Pr	N	O	CH	NCH_3	H	OH	7.22	7.33	7.29	4.20	4.19
19	*i*-Pr	N	S	CH	CH_2	OH	H	6.80	6.90	6.89	4.36	4.36
20	*i*-Pr	N	S	CH	CH_2	H	OH	7.17	7.07	7.08	4.00[e]	4.34
21	$MeOCH_2$	N	S	CH	O	OH	H	7.02	7.22	7.09	–	–
22	$MeOCH_2$	N	S	CH	O	H	OH	7.12	7.38	7.27	–	–

TABLE II (*continued*)

No.	R	X	Y	Z	Q	A	B	log (1/EC_{50})			log (1/$CCIC_{50}$)	
								Obsd	Calcd[a]	Calcd[b]	Obsd	Calcd[c]
23	$MeOCH_2$	N	S	CH	NCH_3	OH	H	6.48	6.64	6.48	4.22	4.13
24	$MeOCH_2$	N	S	CH	NCH_3	H	OH	7.12	6.80	6.67	4.19	4.13
25	$(Me)_2N$	N	S	CH	O	OH	H	7.05	7.05	7.09	4.28	4.26
26	4-morph	N	S	CH	O	OH	H	7.48	7.40	7.09	4.54[e]	4.26
27	4-morph	N	S	CH	O	H	OH	7.48	7.56	7.27	4.22	4.26
28	H	N	CH	S	O	OH	H	7.43	7.08	7.09	–	–
29	H	N	CH	S	O	H	OH	6.89	7.24	7.27	–	–
30	Et	S	N	CH	O	OH	H	6.54	6.63	6.70	4.18	4.26
31	Et	S	N	CH	O	H	OH	7.04	6.79	6.88	–	–
32	*i*-Pr	S	N	CH	O	OH	H	7.00	7.22	7.33	–	–
33	*i*-Pr	S	N	CH	O	H	OH	7.12	7.38	7.51	–	–
34	H	O	N	CH	O	OH	H	6.27	6.39	6.37	–	–
35	H	O	N	CH	O	H	OH	6.66	6.55	6.56	–	–
36	$MeOCH_2$	S	N	CH	O	OH	H	6.75	6.67	6.70	–	–
37	$MeOCH_2$	S	N	CH	O	H	OH	7.07	6.84	6.88	–	–
38	H	S	N	$C(CH_3)$	CH_2O	OH	H	6.96	6.94	6.70	–	–
39	H	S	N	$C(CH_3)$	CH_2O	H	OH	7.09	7.11	6.88	4.22	4.26
40	*t*-Bu	CH	O	N	O	OH	H	7.17	6.91	7.09	–	–
41	*t*-Bu	CH	O	N	O	H	OH	7.14	7.07	7.27	4.72	4.50
42	*t*-Bu	CH	O	N	NCH_3	OH	H	6.00	6.33	6.48	4.25	4.37
43	*t*-Bu	CH	O	N	NCH_3	H	OH	7.32[d]	6.76	6.67	4.26	4.37
44	CH_3O	CH	N	O	O	OH	H	6.66	6.79	7.09	4.22	4.26
45	CH_3O	CH	N	O	O	H	OH	7.08	6.95	7.27	–	–

[a] Using the set 1 activity contributions of substituents as given in Table VI.
[b] Using the set 2 activity contributions of substituents as given in Table VI.
[c] Using activity contributions of Table X.
[d] Not used in derivation of set 1 and set 2 values of Table VI.
[e] Not used in the derivation of activity contributions of Table X.

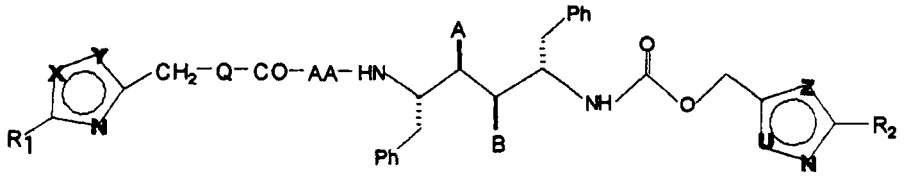
TABLE III P2′ heterocyclic analogues (7) with antiviral potency and cytotoxicity studied by Kempf *et al.*[16]



No.	R	Q	X–Y	Z	A	B	log (1/EC_{50})		$CCIC_{50}$
							Obsd	Calcd[a]	Obsd (μm)
1	H	O	CH–N	N	OH	H	6.60	6.59	>100
2	H	NCH_3	CH–N	N	OH	H	6.17	6.27	>100
3	H	NCH_3	CH–N	N	H	OH	6.35	6.27	>100
4	H	O	CH	O	OH	H	6.64[b]	6.28	>100
5	H	NCH_3	CH	O	OH	H	5.97	5.96	58
6	CH3	O	CH	O	OH	H	7.05	7.09	>100
7	CH3	O	CH	O	H	OH	7.13	7.09	>100
8	H	O	S	N	OH	H	6.28	6.28	>100
9	H	O	S	N	H	OH	6.26	6.28	>100

[a] Using the set 2 activity contributions of substituents as given in Table VII.
[b] Not included in derivation of activity contributions of Table VII.

TABLE IV Ritonavir analogues (9) and their antiviral potency and cytotoxic activity studied by Kempf *et al.*[16]



No.	R_1	X	Y	Q	AA	Z	U	R_2	A	B	log (1/EC_{50})			log (1/$CCIC_{50}$)	
											Obsd	Calcd[a]	Calcd[b]	Obsd	Calcd[c]
1	*i*-Pr	S	CH	O	Val	S	CH	H	OH	H	8.00[d]	6.04	6.13	4.00[e]	4.55
2	*i*-Pr	S	CH	O	Val	S	CH	H	H	OH	8.31[d]	6.80	6.91	4.29	4.55
3	*i*-Pr	S	CH	NCH_3	Val	S	CH	H	OH	H	6.72	6.49	6.60	4.25	4.41
4	*i*-Pr	S	CH	NCH_3	Val	S	CH	H	H	OH	7.60	7.25	7.39	4.24	4.41
5	*i*-Pr	O	CH	NCH_3	Val	S	CH	H	OH	H	6.01	6.43	6.60	4.24	4.42
6	*i*-Pr	O	CH	NCH_3	Val	S	CH	H	H	OH	7.37	7.20	7.39	4.22	4.42
7	*i*-Pr	S	CH	NCH_3	Val	O	CH	H	H	OH	7.89	7.89	7.39	4.22	4.41
8	*i*-Pr	O	CH	NCH_3	Val	O	CH	H	H	OH	–	–	–	–	–
9	CH3	S	CH	NCH_3	Val	S	CH	H	H	OH	6.10[d]	7.17	7.39	4.21	4.23
10	CH3	O	CH	NCH_3	Val	S	CH	H	H	OH	7.11	7.11	7.39	–	–
11	Et	S	CH	NCH_3	Val	S	CH	H	OH	H	6.55	6.70	6.60	4.20	4.23
12	Et	S	CH	NCH_3	Val	S	CH	H	H	OH	7.28	7.47	7.39	4.23	4.23
13	Et	O	CH	NCH_3	Val	S	CH	H	H	OH	7.75	7.41	7.39	4.00	4.03
14	3-pent	S	CH	NCH_3	Val	S	CH	H	H	OH	6.85	6.85	7.39	4.72[e]	4.23
15	*c*-Pr	S	CH	NCH_3	Val	S	CH	H	OH	H	6.82	6.69	6.60	4.35	4.23
16	*c*-Pr	S	CH	NCH_3	Val	S	CH	H	H	OH	7.33	7.46	7.39	4.32	4.23
17	*c*-Bu	S	CH	NCH_3	Val	S	CH	H	OH	H	6.64	6.64	6.60	4.70	4.65
18	*c*-Bu	S	CH	NCH_3	Val	S	CH	H	H	OH	7.41	7.41	7.39	4.59	4.65
19	4-morph	S	CH	O	Val	S	CH	H	OH	H	7.10	6.97	6.96	–	–
20	4-morph	S	CH	O	Val	S	CH	H	H	OH	7.60	7.73	7.74	4.22	4.37
21	*i*-Pr	CH	S	O	Val	S	CH	H	OH	H	5.96	6.02	6.13	4.44	4.55
22	*i*-Pr	CH	S	O	Val	S	CH	H	H	OH	6.85	6.79	6.91	4.66	4.55
23	*i*-Pr	S	CH	NEt	Val	S	CH	H	OH	H	5.85	6.42	6.13	4.85	4.55
24	*i*-Pr	S	CH	NEt	Val	S	CH	H	H	OH	6.60	7.19	6.91	4.62	4.55
25	*i*-Pr	S	CH	N*c*Pr	Val	S	CH	H	H	OH	6.55	6.26	6.26	4.72	4.55
26	*i*-Pr	S	CH	N*c*Pr	Val	S	CH	H	OH	H	5.57	5.50	5.49	4.70	4.55
27	*i*-Pr	S	CH	O	Ala	S	CH	H	OH	H	5.77	5.97	6.13	4.20	4.34
28	*i*-Pr	S	CH	O	Ala	S	CH	H	H	OH	7.00	6.73	6.91	4.25	4.34
29	*i*-Pr	S	CH	NCH_3	Ala	S	CH	H	OH	H	5.87	6.42	6.60	4.22	4.20
30	*i*-Pr	S	CH	NCH_3	Ala	S	CH	H	H	OH	7.41	7.18	7.39	4.28	4.20
31	*i*-Pr	O	CH	NCH_3	Ala	S	CH	H	H	OH	7.23	7.12	7.39	4.00	4.00
32	*I*-Pr	S	CH	NEt	Ala	S	CH	H	OH	H	7.05	6.35	6.13	–	–
33	*i*-Pr	S	CH	NEt	Ala	S	CH	H	H	OH	7.86	7.11	6.91	4.28	4.34
34	*i*-Pr	S	CH	NPr	Ala	S	CH	H	H	OH	6.85	6.99	6.91	4.72	4.60
35	*i*-Pr	S	CH	N*i*Bu	Ala	S	CH	H	H	OH	6.75	6.90	6.91	4.77	4.81
36	*i*-Pr	S	CH	N*n*Bu	Ala	S	CH	H	H	OH	6.54	6.54	6.91	4.77	4.77
37	*i*-Pr	S	CH	N*c*Pr	Ala	S	CH	H	OH	H	4.85	5.42	5.49	4.35	4.34
38	*i*-Pr	S	CH	N*c*Pr	Ala	S	CH	H	H	OH	6.26	6.18	6.26	4.39	4.34
39	*i*-Pr	S	CH	NCH_3	*β*-Ala	S	CH	H	OH	H	5.66	5.36	5.41	4.24	4.19
40	*i*-Pr	S	CH	NCH_3	*β*-Ala	S	CH	H	H	OH	5.70	6.12	6.20	4.26	4.19
41	*i*-Pr	S	CH	N*c*Pr	*β*-Ala	S	CH	H	OH	H	4.87	4.36	4.29	4.23	4.33
42	*i*-Pr	S	CH	N*c*Pr	*β*-Ala	S	CH	H	H	OH	4.75	5.13	5.08	4.32	4.33
43	*i*-Pr	S	CH	NEt	Gly	S	CH	H	H	OH	5.68	5.98	5.87	4.39	4.31
44	*i*-Pr	S	CH	NPr	Gly	S	CH	H	H	OH	6.00	5.86	5.87	4.46	4.58
45	*i*-Pr	S	CH	N*i*Bu	Gly	S	CH	H	H	OH	5.92	5.77	5.87	4.82	4.78
46	*i*-Pr	S	CH	CH_2O	Val	S	CH	H	H	OH	7.00	7.00	6.91	–	–

TABLE IV (*continued*)

No.	R_1	X	Y	Q	AA	Z	U	R_2	A	B	log (1/EC_{50})			log (1/$CCIC_{50}$)	
											Obsd	Calcd[a]	Calcd[b]	Obsd	Calcd[c]
47	*i*-Pr	S	CH	CH_2O	Val	S	CH	H	OH	H	7.77[d]	6.24	6.13	4.60	4.55
48	*i*-Pr	S	CH	Val	S	CH	H	OH	H	6.09	6.04	6.13	–	–	–
49	*i*-Pr	S	CH	Val	S	CH	H	H	OH	6.68	6.80	6.91	–	–	–
50	*i*-Pr	S	CH	CH_2NMe	Val	S	CH	H	H	OH	6.74	6.95	6.91	4.60	4.55
51	*i*-Pr	S	CH	CH_2NEt	Val	S	CH	H	H	OH	6.26	6.02	5.98	4.75	4.79
52	*i*-Pr	S	CH	CH_2NPr	Val	S	CH	H	H	OH	5.96	5.79	5.75	4.75	4.86
53	*i*-Pr	S	CH	CH_2NMe	Ala	S	CH	H	H	OH	7.09	6.88	6.91	4.23	4.34
54	*i*-Pr	S	CH	CH_2NEt	Ala	S	CH	H	H	OH	5.70	5.94	5.98	4.62	4.58
55	*i*-Pr	S	CH	CH_2NPr	Ala	S	CH	H	H	OH	5.54	5.71	5.75	4.75	4.64
56	*i*-Pr	S	CH	NCH_3	Val	S	CH	CH_3	OH	H	7.13	6.66	6.60	4.66	4.71
57	*i*-Pr	S	CH	NCH_3	Val	S	CH	CH_3	H	OH	6.96	7.43	7.39	4.75	4.71
58	*i*-Pr	S	CH	NCH_3	Val	S	CH	*i*-Pr	OH	H	5.82	5.88	5.87	4.72	4.72
59	*i*-Pr	S	CH	NCH_3	Val	S	CH	*i*-Pr	H	OH	6.70	6.64	6.65	4.72	4.72
60	*i*-Pr	S	CH	NCH_3	Val	CH	O	H	H	OH	8.05	7.85	7.39	4.55	4.41
61	*i*-Pr	O	CH	NCH_3	Val	CH	O	H	H	OH	7.92	7.79	7.39	4.24	4.22
62	*i*-Pr	O	CH	NCH_3	Ala	CH	O	H	H	OH	7.38	7.71	7.39	–	–

[a] Using the set 1 activity contributions of substituents as given in Table VIII.
[b] Using the set 2 activity contributions of substituents as given in Table VIII.
[c] Using activity contributions of substituents as given in Table XI.
[d] Not used in the derivation of set 1 and set 2 values of Table VIII.
[e] Not included in the derivation of activity contributions of Table XI.

TABLE V Activity contributions of substituents of analogues of (**5**) (Table I)
The parent structure is defined as: $R_1 = R_2 = H$, (X = N, Y = CH), Q = O, (A = OH, B = H)

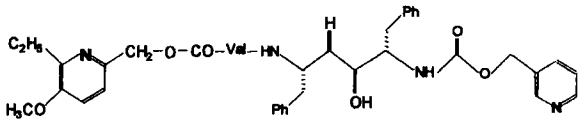
R_1	R_2	X, Y	Q	A and B	μ
$CH_3 = 0.279(\pm 0.180)$	$CH_3 = 0.146(\pm 0.196)$*	(X = CH, Y = N) $= 0.120(\pm 0.182)$*	$NCH_3 = -0.531(\pm 0.118)$	**(A = H, B = OH) = 0.179(± 0.109)**	6.724
$C_2H_5 = 0.735(\pm 0.284)$	**$OCH_3 = 0.667(\pm 0.227)$**	(X = N, Y = N) $= -0.049(\pm 0.215)$*	$NH = -0.661(\pm 0.276)$		
i-Pr $= 0.547(\pm 0.284)$	$NH_2 = -0.345(\pm 0.284)$				
t-Bu $= 0.247(\pm 0.284)$*					
$NH_2 = -0.488(\pm 0.354)$					
$OCH_3 = 0.315(\pm 0.286)$					
			$n = 58$, $r = 0.914$, $s = 0.178$, $F = 15.537$		
Activity contributions ignoring statistically insignificant substituents					
$CH_3 = 0.157(\pm 0.153)$	**$OCH_3 = 0.563(\pm 0.189)$**		$NCH_3 = -0.506(\pm 0.119)$	**(A = H, B = OH) = 0.171(± 0.113)**	6.879
$C_2H_5 = 0.571(\pm 0.242)$	$NH_2 = -0.509(\pm 0.241)$		$NH = -0.645(\pm 0.266)$		
i-Pr $= 0.383(\pm 0.242)$					
$NH_2 = -0.520(\pm 0.326)$					
$OCH_3 = 0.271(\pm 0.242)$					
			$n = 58$, $r = 0.897$, $s = 0.193$, $F = 19.425$		

* Insignificant at 95% confidence intervals.

as shown in the next set of data in Table V. These results do not appear much different from the previous ones. The statistical parameters appear to be little affected.

Both sets of results reveal that the highest activity contributions are associated with $R_1 = C_2H_5$, $R_2 = OCH_3$, and A = H, B = OH, and thus the most potent compound is predicted to be (**A**) with

$\log(1/EC_{50})$ equal to 8.30 (from set 1) or 8.18 (from set 2).



A (pEC_{50}=8.18)

In a similar manner, the activity contributions of substituents were obtained for the compounds of Tables II–IV and are reported in Tables VI–VIII, respectively. In each case, the most favourable substituents are indicated in bold face. Thus in the series of Table II the most active compound is predicted to be (**B**), in the series of

TABLE VI Activity contributions of substituents of analogues of (**6**) (Table II)
The parent structure is defined as: R = H, (X = N, Y = S, Z = CH), Q = O, (A = OH, B = H)

R	X, Y, Z	Q	A, B	μ
$C_2H_5 = 0.456(\pm 0.414)^{\dagger}$	(X = S, Y = N, Z = CH) = −0.548(± 0.263)	$NCH_3 = -0.579(\pm 0.228)$	**(A = H, B = OH) = 0.166(± 0.152)**	6.984
***i*-Pr = 1.043(± 0.434)**	(X = O, Y = N, Z = CH) = −0.334(± 0.528)	$CH_2 = -0.856(\pm 0.419)$		
t-Bu = 0.293(± 0.588)*	**(X = S, Y = N, Z = $C(CH_3)_2$) = 0.772(± 0.549)**†			
$MeOCH_2 = 0.503(\pm 0.414)^{\dagger}$	(X = CH, Y = N, Z = O) = 0.071(± 0.528)*			
4-morpho = 0.681(± 0.528)†	(X = N, Y = CH, Z = S) = 0.361(± 0.528)*			
$N(CH_3)_2 = 0.334(\pm 0.631)$*	(X = N, Y = O, Z = CH) = −0.015(± 0.335)*			
$CH_3 = 0.490(\pm 0.476)^{\dagger}$	(X = CH, Y = O, Z = N) = −0.101(± 0.572)*			
		$n = 42$, $r = 0.926$, $s = 0.176$, $F = 8.526$		
	Activity contributions ignoring statistically insignificant substituents			
***i*-Pr = 0.628(± 0.200)**	(X = S, Y = N, Z = CH) = −0.392(± 0.224)	$NCH_3 = -0.607(\pm 0.192)$	**(A = H, B = OH) = 0.182(± 0.162)**	7.090
	(X = O, Y = N, Z = CH) = −0.716(± 0.393)	$CH_2 = -0.824(\pm 0.419)$		
		$n = 42$, $r = 0.926$, $s = 0.176$, $F = 8.526$		

*Insignificant at 95% confidence intervals.
†Not significant in each iteration.

TABLE VII Activity contributions of substituents of analogues of (**7**) (Table III)
The parent structure is defined as: R = H, (X–Y = S, Z = N), Q = O, (A = OH, B = H)

R	X–Y, Z	Q	A, B	μ
CH_3 = 0.700(± 0.583)	**(X–Y = CH–N, Z = N) = 0.370(± 0.393)**	$NCH_3 = -0.380(\pm 0.393)$	(A = H, B = OH)* = 0.080(± 0.248)	6.230
	(X–Y = CH, Z = O) = 0.120(± 0.583)*			
		$n = 8$, $r = 0.996$, $s = 0.038$, $F = 48.528$		
	Activity contributions ignoring statistically insignificant substituents			
CH_3 = 0.813(± 0.189)	**(X–Y = CH–N, Z = N) = 0.310(± 0.171)**	$NCH_3 = -0.320(\pm 0.171)$		6.277
		$n = 8$, $r = 0.992$, $s = 0.054$, $F = 79.673$		

*Insignificant at 95% confidence intervals.

TABLE VIII Activity contributions of substituents of analogues of (9) (Table IV)
The parent structure is defined as: R_1 = H, (X = S, Y = CH), Q = O, (A = OH, B = H), (Z = S, U = CH), AA = Val

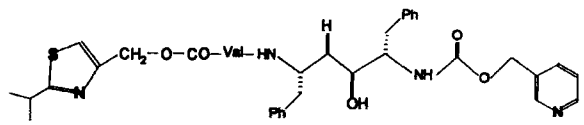
R_1	X, Y	AA	Q	Z, U	R_2	A, B	μ
i-Pr = 0.403(± 0.925)*	(X = CH, Y = S) = −0.018(± 0.739)*	Ala = −0.075(± 0.315)*	**NCH_3 = 0.449(± 0.543)**	(Z = CH, U = O) = 0.594(± 0.591)	**CH_3 = 0.173(± 0.698)***	**(A = H, B = OH) = 0.763(± 0.273)**	5.638
C_2H_5 = 0.617(± 0.981)*	(X = O, Y = CH) = −0.058(± 0.470)*	β-Ala = −1.129(± 0.522)	NEt = 0.381(± 0.573)*	(Z = O, U = CH) = 0.637(± 0.925)	*i*-Pr = −0.612(± 0.698)		
c-Pr = 0.607(± 1.029)*		Gly = −1.210(± 0.671)	NcPr = −0.546(± 0.563)				
c-Bu = 0.557(± 1.029)*			NPr = 0.264(± 0.856)*				
CH_3 = 0.318(± 1.269)*			N*i*Bu = 0.174(± 0.801)*				
4-morpho = 1.331(± 1.164)			N*n*Bu = −0.189(± 0.953)*				
			CH_2O = 0.196(± 0.956)*				
			CH_2NMe = 0.148(± 0.734)*				
			CH_2NEt = −0.787(± 0.734)				
			CH_2NPr = −1.017(± 0.734)				
					n = 57, r = 0.929, s = 0.299, F = 7.215		
			Activity contributions ignoring statistically insignificant substituents				
4-morpho = 0.830(± 0.581)		β-Ala = −1.188(± 0.435)	**NCH_3 = 0.475(± 0.252)**		*i*-Pr = −0.735(0.571)	**(A = H, B = OH) = 0.786(± 0.231)**	6.127
		Gly = −1.046(± 0.492)	NcPr = −0.649(± 0.399)				
			CH_2NEt = −0.933(± 0.584)				
			CH_2NPr = −1.163(± 0.584)				
					n = 57, r = 0.899, s = 0.352, F = 22.029		

*Insignificant at 95% confidence intervals.

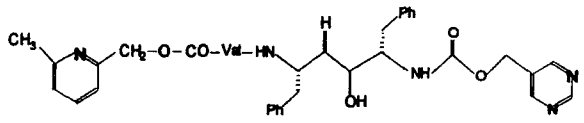
TABLE IX Cytotoxic activity contributions of statistically significant substituents of analogues of (5) (Table I)
The parent structure is defined as:
$R_1 = R_2 = H$, (X = N, Y = CH), Q = O, (A = OH, B = H)

R_1	μ
i-Pr = 0.551(± 0.093)	4.163
t-Bu = 0.548(± 0.083)	
$n = 23$, $r = 0.968$, $s = 0.068$, $F = 146.47$	

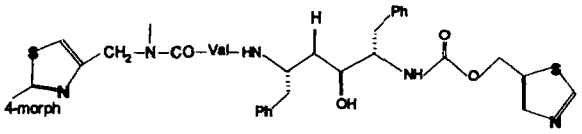
Table III (C), and in the series of Table IV (D). The prediction in each case has been based on the activity contributions listed in set 2 in each of Tables VI–VIII.



B (pEC50 = 7.90)



C (pEC_{50}=7.40)



D (pEC50 = 8.22)

The calculation shows that except (C), the other three compounds predicted (A, B, D) may have a higher activity than ritonavir (8, $pEC_{50} = 7.60$) (compound 4 in Table IV, the compound which has been licensed for use). All four compounds are predicted to have much higher activity than 3 and 4, which were found by Kempf *et al.*[14,15] to possess adequate anti-HIV activity ($pEC_{50} = 6.60$ for each).

We also analyzed using the Fujita–Ban approach the cytotoxic activity (CC_{50}) of the compounds in all the Tables, except those of Table III where the CC_{50} values reported for all compounds except one (5, 58 μM) were uncertain (>100 μM for all). The activity contributions of substituents for the different series of compounds are listed in Tables IX–XI. Only statistically significant contributions are reported. From this data, the pCC_{50} values for compounds

TABLE X Cytotoxic activity contributions of statistically significant substituents of analogues of (6) (Table II)
The parent structure is defined as: R = H, (X = N, Y = S, Z = CH), Q = O, (A = OH, B = H)

R	X, Y, Z	Q	μ
i-Pr = 0.341(± 0.134)	(X = N, Y = O, Z = CH) = −0.283(± 0.167)	NCH_3 = −0.135(± 0.092)	4.262
t-Bu = 0.618(± 0.165)	(X = CH, Y = O, Z = N) = −0.380(± 0.190)	CH_2 = −0.244(± 0.244)	
	$n = 24$, $r = 0.910$, $s = 0.083$, $F = 13.611$		

TABLE XI Cytotoxic activity contributions of statistically significant substituents of analogues of (9) (Table IV)
The parent structure is defined as: R_1 = H, (X = S, Y = CH), Q = O, (A = OH, B = H), (Z = S, U = CH), AA = Val

R_1	X, Y	AA	Q	R_2	μ
c-Bu = 0.419(± 0.199)	(X = O, Y = CH) = −0.193(± 0.133)	Ala = −0.215(± 0.094)	NCH_3 = −0.139(± 0.098)	CH_3 = 0.291(± 0.196)	4.365
i-Pr = 0.187(± 0.119)		β-Ala = −0.221(± 0.139)	**N*i*Bu = 0.471(± 0.203)**	**i-Pr = 0.305(± 0.184)**	
		Gly = −0.241(± 0.188)	N*n*Bu = 0.432(± 0.259)		
			NPr = 0.266(± 0.203)		
			CH_2NEt = 0.240(± 0.184)		
			CH_2NPr = 0.305(± 0.184)		
			$n = 52$, $r = 0.901$, $s = 0.103$, $F = 11.360$		

(**A**), (**B**), and (**D**), belonging to the series of Tables I, II and IV, respectively, are found to be 4.16, 4.60 and 4.36, respectively. If we compare the toxicity of (**D**), which is a member of the ritonavir series (Table IV), with that of ritonavir $pCC_{50} = 4.24$ (4, Table IV), both are found to have essentially the same toxicity, but the anti-HIV activity of (**D**) (8.22) is predicted to be much higher than that of ritonavir (7.60). Therefore, compound (**D**) appears to be a good drug candidate and should be further examined. Compound (**B**) could also be a better choice, as its anti-HIV activity is slightly higher than that of ritonavir and its toxicity slightly lower than that of ritonavir.

From this analysis, it is found that in each structure (**5**), (**6**), (**7**) or (**9**), the substitution A = H and B = OH has an edge over A = OH and B = H as shown in the parent structure. It may be, therefore, suggested that the OH group may form a hydrogen bond with the receptor and thus strengthen the drug-receptor binding when it is nearer to, and on the same side of, the NH moiety in the side chain. Both OH and NH may be involved in hydrogen binding either with two different hydrogen bond acceptor sites on the receptor, or with only one forming a three-centre hydrogen bond. In either case, the drug-receptor binding would be strengthened and a closer examination of the activity contribution of A = H and B = OH in all the cases shows that it is much larger in (**9**) than in any other series and this may probably be the reason that compound (**D**), which can be derived from (**9**), shows a supremacy over all the compounds studied.

It is, however, also observed that only in the series of (**9**) is $Q = NCH_3$ found to have a positive contribution, while in all other series it has a negative contribution. Ritonavir (**8**) also has $Q = NCH_3$. The negative contribution of this moiety in the other cases, i.e., in (**5**), (**6**), and (**7**) may be due to its steric hindrance in the interaction of the nearby heterocyclic ring with the receptor. The orientation and positioning of the heterocyclic ring may depend on the conformational flexibility of the molecule.

The other substituents giving a positive contribution in any series appear to contribute either through hydrophobic interaction or hydrogen bonding.

Most compounds in each series are well predicted by Fujita–Ban analysis, but a few are not. Such compounds have been deleted from the regression analysis (see footnotes to Tables I–IV) in order to have statistically most significant results. The appreciable differences in the observed and predicted activities of these compounds cannot be explained as no physico-chemical or structural properties, which really govern the activity, have been used in the regression analysis. Since there were not much variations in the substituents at any position in any series of compounds, the situation was not ideal for QSAR analysis using physico-chemical or structural properties, hence the Fujita–Ban analysis was performed. The drawback with the Fujita–Ban approach is that predictions cannot be made outside the substituents used for the regression analysis.

References

[1] N.E. Kohl, N.A. Emini, W.A. Schleif, L.J. Davis, J.C. Heimbach, R.A.F. Dixon, E.M. Scolnick and I.S. Sigal (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4686.
[2] S. Seelmeier, H. Schmidt, V. Turk and K. von der Helm (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6612.
[3] A. Wlodawer and J.W. Erickson (1993) *Annu. Rev. Biochem.*, **62**, 543.
[4] K. Applet (1993) *Perspec. Drug Disc. Des.*, **1**, 23.
[5] R.E. Boehme, A.D. Borthwick and P.G. Wyatt (1995) *Annu. Rep. Med. Chem.*, **30**, 139.
[6] K.R. Romines and S. Thaisrivongs (1993) *Drugs Future*, **20**, 377.
[7] G.L. Olson, D.R. Bolin, M.P. Bonner, M. Bos, C.M. Cook, D.C. Fry, B.J. Graves, M. Hatada, D.E. Hill, M. Kahn, V.S. Madison, V.K. Rusiecki, R. Sarabu, J. Sepinwall, G.P. Vincent and M.E. Voss (1993) *J. Med. Chem.*, **36**, 3039.
[8] See for example: P.E. Maligres, V. Upadhyay, K. Rossen, S.J. Cianciosi, R.M. Purick, K.K. Eng, R.A. Reamer, D. Askin, R.P. Volante and P.J. Reider (1995) *Tetrahedron Lett.*, **36**, 2195.
[9] D.J. Kempf, K.C. Marsh, J.F. Denissen, E. McDonald, S. Vasavanonda, C.A. Flentge, B.E. Green, L. Fino, C.H. Park, X.-P. Kong, N.E. Wideburg, A. Saldivar, L. Ruiz,

W.M. Kati, H.L. Sham, T. Robinn, K.D. Stewart, A. Hau, J.J. Plattner, J.M. Leonard and D.W. Norbeck (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 2484.

[10] J.P. Vacca, B.D. Dorsey, W.A. Schleif, R.B. Levin, S.L. McDaniel, P.L. Darke, J. Zugae, J.C. Quintero, O.M. Blahy, E. Roth, V.V. Sardana, A.J. Schlabach, P.I. Graham, J.H. Condra, L. Gotlib, M.K. Holloway, J. Lin, L.-W. Chen, K. Vastag, D. Ostovic, P.S. Anderson, E.A. Eimini and J.R. Huff (1999) *Proc. Natl. Acad. Sci. USA*, **91**, 4096.

[11] V. Kalish, S. Kaldor, B. Shetty, J. Tatlock, J. Davies, M. Hammend, B. Dressman, J. Fritz, K. Applet, S. Reich, L. Musick, B.-W. Wu and K. Su (1995) *Eur. J. Med. Chem.*, **30**, 201.

[12] E.E. Kim, C.T. Baker, M.D. Dwyer, M.A. Murcko, B.G. Rao, R.D. Tung and M.A. Navia (1995) *J. Am. Chem. Soc.*, **117**, 1181.

[13] D.J. Kempf, D.W. Norbeck, L. Codacovi, X.C. Wang, W.E. Kohlbrenner, N.E. Wideburg, D.A. Paul, M.F. Knigge, S. Vasavanonda, A. Craig-Kennard, A. Saldivar, W.J. Rosenbrook, J.J. Clement, J.J. Plattner and J. Erickson (1990) *J. Med. Chem.*, **33**, 2687.

[14] D.J. Kempf, L. Codacovi, X.C. Wang, W.E. Kohlbrenner, N.E. Wideburg, A. Saldivar, S. Vasavanonda, K.C. Marsh, P. Bryant, H.L. Sham, B.E. Green, D.A. Betebenner, J. Erickson and D.W. Norbeck (1993) *J. Med. Chem.*, **36**, 320.

[15] D.J. Kempf, K.C. Marsh, D.A. Paul, M.F. Knigge, D.W. Norbeck, W.E. Kohlbrenner, L. Codacovi, S. Vasavanonda, P. Bryant, X.C. Wang, N.E. Wideburg, J.J. Clement, J.J. Plattner and J.J. Erickson (1991) *Antimicrob. Agents Chemother.*, **35**, 2209.

[16] D.J. Kempf, H.L. Sham, K.C. Marsh, C.A. Flentge, D. Betebenner, B.E. Green, E. McDonald, S. Vasavanonda, A. Saldivar, N.E. Wideburg, W.M. Kati, L. Ruiz, C. Zhao, L. Fino, J. Patterson, A. Molla, J.J. Plattner and D.W. Norbeck (1998) *J. Med. Chem.*, **41**, 602.

[17] T. Fujita and T. Ban (1971) *J. Med. Chem.*, **14**, 148.