J. Enzyme Inhibition, 1998, Vol. 13, pp. 125–134
Reprints available directly from the publisher
Photocopying permitted by license only

INHIBITORS OF THE EPIDERMAL GROWTH FACTOR RECEPTOR PROTEIN TYROSINE KINASE: A QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIP ANALYSIS

P. SINGH * and R. KUMAR

Department of Chemistry, S.K. Government College, Sikar-332 001, India

(Received 21 July 1997; In final form 15 September 1997)

Hansch and Free-Wilson analyses are described on a data set, 4-anilinoquinazolines [the analogues of 4-(3-bromo-anilino)-6,7-dimethoxy quinazoline: PD 153035], as inhibitors of the epidermal growth factor receptor protein tyrosine kinase. These analyses have helped to ascertain the role of different substituents in explaining the observed inhibitory activities. From both approaches, it is concluded that the combined electron-donating nature of R_1- and R_2-substitutions of the quinazoline ring and the electron-withdrawing nature of the X-substitution of the anilino-ring are beneficial for increasing the inhibition activity of a compound. Further, the symmetrical alkoxy substituents present at the R_1- and R_2-positions are also engaged in a steric interaction which was determined quantitatively through the parabolic relationship between the activity and combined molar refraction parameter, ΣMR of the substituents.

Keywords: Quantitative structure–activity relationship QSAR; EGFR inhibitors of the epidermal growth factor; Tyrosine kinase enzymes; Hansch and Free-Wilson analyses

INTRODUCTION

The importance of protein tyrosine kinase (PTK) enzymes as mediators of proliferative as well as metabolic signals have been realized in recent years. A number of proliferative diseases may precipitate due to the generation of mitogenic signals from abnormally expressed or deregulated PTK's. One of

* Corresponding author.

the important targets for the treatment of proliferative diseases is the epidermal growth factor receptor (EGFR) protein tyrosine kinase family of transmembrane growth factor receptors. The EGFR is a transmembrane glycoprotein which mediates the mitogenic response of cells to the epidermal growth factor family of mitogenic polypeptides, transforming growth factor [1,2] and amphiregulin.[3] The EGFR and its ligands are involved in malignant tumor growth[4–6] and skin diseases such as psoriasis.[7] Enzymatic activity of the intracellular EGFR is for signal transduction via the EGFR.[8,9] The involvement of EGFR protein tyrosine kinase enzymes in epithelial proliferation suggests that enzymes inhibitors could have therapeutic potential in the treatment of malignant and non-malignant epithelial diseases. Since tyrosine kinase enzymes are involved in many signal transduction pathways, it would be highly beneficial to develop compounds with high selectivity at the enzyme level. Several classes of such inhibitors have been reported[10–16] but most have proved to be of limited use in cellular assays and more advanced models. More recently the analogues of 4-(3-bromoanilino)-6,7-dimethoxyquinazoline (PD 153035) have been reported[7] as possessing inhibitory properties against the epidermal growth factor receptor.

The initial structure–activity (SAR) studies on these compounds were directed to only alteration of the substituents at various positions of the structure. Although their study resulted in the discovery of 4-(3-bromoanilino)-6,7-diethoxyquinazoline as a lead compound, no rationale has yet been provided to reduce the trial and error factors. Hence, a quantitative SAR (QSAR) on these analogues was conducted since QSAR not only provides the rationale for drug design but also illuminates the mechanism of action of drugs.

MATERIALS AND METHODS

QSAR analysis was made on the analogues of PD 153035 (Figure 1 for structures) studied by Bridges and co-workers.[17] Their tyrosine kinase inhibitory activities of the EGFR, binding competitively at the ATP site are listed in Table I along with some physicochemical parameters. The most appropriate physicochemical parameters were found to be the electronic constant, σ and the molar refraction parameter, MR, taken from the compilation of Hansch *et al.*[18] Additionally, some dummy parameters were also employed to reflect the effect of some specific alteration. Both Hansch and Free-Wilson types of calculations were carried out for these compounds.

All tyrosine kinase inhibition data refers to the ability of the compounds to inhibit the phosphorylation of a 14-residue fragment of phopholipase $C_\gamma 1$ by

TABLE I 4-Anilinoquinazolines and their tyrosine kinase inhibitory activity data and physicochemical parameters (see Figure 1 for structures)

Compound	R_1	R_2	X	$\Sigma\sigma$	σ_X	$-\log IC_{50}(M)$[a]			
						Obsd.	*Calcd.*		
							Eq. (2)	Eqs. (3) & (5)[d]	F.W.
1	H	H	H	0.00	0.00	6.46	6.02	6.16	6.01
2	H	H	F	0.00	0.34	7.25	7.35	7.39	6.80
3	H	H	Cl	0.00	0.37	7.64	7.47	7.50	7.62
4	H	H	Br	0.00	0.39	7.57	7.54	7.57	7.59
5	H	H	I	0.00	0.35	7.10	7.39	7.42	7.38
6	H	H	CF_3	0.00	0.43	6.24	6.42	6.10	6.65
7	OMe	H	H	−0.27	0.00	7.26	6.41	6.57	6.98
8	OMe	H	Br	−0.27	0.39	7.52	7.94	7.98	8.57
9	NH_2	H	H	−0.66	0.00	6.11	6.98	7.16	6.23
10	NH_2	H	CF_3	−0.66	0.43	6.24	7.38	7.10	6.88
11	NH_2	H	Br	−0.66	0.39	9.11	8.50	8.57	7.82
12	NO_2	H	H	0.78	0.00	5.30	4.88	4.98	4.87
13	NO_2	H	Br	0.78	0.39	6.05	6.41	6.39	6.46
14	H	OMe	H	0.12	0.00	6.92	5.84	5.98	7.02
15	H	OMe	Br	0.12	0.39	8.00	7.37	7.39	8.61
16	H	NH_2	H	−0.16	0.00	7.00	7.69	7.83	7.77
17	H	NH_2	F	−0.16	0.34	8.70	9.02	9.06	8.57
18	H	NH_2	Cl	−0.16	0.37	9.60	9.13	9.17	9.38
19	H	NH_2	Br	−0.16	0.39	10.00	9.21	9.24	9.36
20	H	NH_2	I	−0.16	0.35	9.46	9.06	9.09	9.14
21	H	NH_2	CF_3	−0.16	0.43	8.48	8.09	7.77	8.42
22	H	NO_2	H	0.71	0.00	4.92	4.99	5.09	4.63
23	H	NO_2	F	0.71	0.34	5.21	6.32	6.31	5.42
24	H	NO_2	Cl	0.71	0.37	6.09	6.43	6.42	6.24
25	H	NO_2	Br	0.71	0.39	6.00	6.51	6.49	6.21
26	H	NO_2	I	0.71	0.35	6.27	6.35	6.35	6.00
27	OMe	OMe	H	−0.15	0.00	7.54	8.30	8.08	8.00
28	OMe	OMe	F	−0.15	0.34	8.42	9.63	8.04	8.79

TABLE I (*Continued*)

Compound	R_1	R_2	X	$\Sigma\sigma$	σ_X	$-\log IC_{50}(M)$[a]			
						Obsd.	*Calcd.*		
							Eq. (2)	Eqs. (3) & (5)[d]	F.W.
29	OMe	OMe	Cl	−0.15	0.37	9.51	8.71	9.70	9.61
30	OMe	OMe	Br	−0.15	0.39	10.60	9.83	10.36	9.58
31	OMe	OMe	I	−0.15	0.35	9.05	9.67	—[c]	9.37
32	OMe	OMe	CF_3	−0.15	0.43	9.62	8.70	9.42	8.64
33	H	NHAc	Br	0.07	0.39	7.40	7.44	7.46	8.04
34	H	NHMe	Br	−0.30	0.39	8.15	7.98	8.02	8.46
35	H	NMe_2	Br	−0.15	0.39	7.96	7.76	7.80	7.12
36	NH_2	NH_2	Br	−0.82	0.39	9.92	10.17	10.24	9.58
37	NH_2	NHMe	Br	−0.96	0.39	9.16	8.94	9.02	8.69
38	NH_2	NMe_2	Br	−0.81	0.39	6.80	—[c]	—[c]	7.35
39	NH_2	OMe	Br	−0.54	0.39	8.42	8.33	8.39	8.83
40	NH_2	Cl	Br	−0.29	0.39	8.19	7.97	8.01	8.57
41	NO_2	NH_2	Br	0.62	0.39	7.28	8.08	8.06	8.22
42	NO_2	NHMe	Br	0.48	0.39	7.17	6.85	6.84	7.33
43	NO_2	NMe_2	Br	0.63	0.39	5.70	6.63	6.62	5.99
44	NO_2	NHAc	Br	0.85	0.39	7.55	6.31	6.28	6.91
45	NO_2	OMe	Br	0.90	0.39	7.82	—[c]	—[c]	7.47
46	NO_2	Cl	Br	1.15	0.39	7.60	—[c]	—[c]	7.22
47	-OCH_2O-		Br	−0.32	0.39	7.82	8.01	8.05	
48	OH	OH	Br	−0.25	0.39	9.77	9.97	9.77	
49	OEt	OEt	Br	−0.14	0.39	11.22	9.81	11.18	
50	OPr	OPr	Br	−0.15	0.39	9.77	9.83	10.02	
51	OBu	OBu	Br	−0.22	0.39	6.98	7.86	6.86	
52	NHMe	H	Br	−0.84	0.39	8.40	8.77	8.84	
53	NMe_2	H	Br	−0.83	0.39	7.08	—[c]	—[c]	
54	NHCOOMe	H	Br	−0.15	0.39	7.92	7.76	7.80	
55	H	OH	Br	0.12	0.39	8.33	7.37	7.39	
56	H	NHEt	Br	−0.24	0.39	7.92	7.89	7.93	

[a]*in vitro* Tyrosine kinase inhibition activity on a molar basis. [b]Taken from Reference 17. [c]'Outlier' compounds in the present study. [d]The values in this column are calculated using Eq. (3) for compounds **27–32** and **48–51** and Eq. (5) for the remainder.

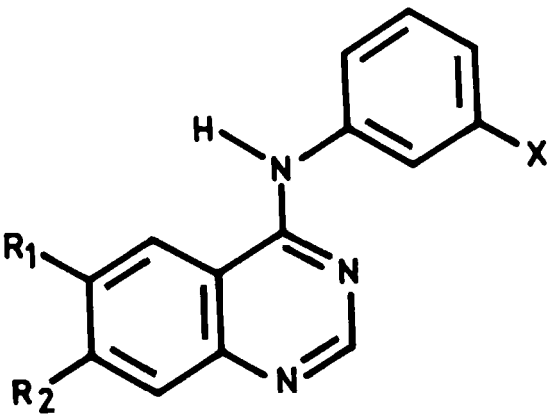

FIGURE 1 Structures of 4-anilinoquinazolines (the analogues of PD 153035: $R_1 = R_2 =$ OMe and X = Br).

EGFR (prepared from human A 431 carcinoma cell vesicles by immunoaffinity chromatography). It has been measured in terms of the IC_{50} value, the minimum concentration of the compound required to achieve 50% inhibition of the EGFR. For the present study, the same are expressed as $-\log IC_{50}$ on a molar basis.

RESULTS AND DISCUSSION

Table I lists the compounds where the alterations in substituents were made at the R_1- and R_2-positions of the quinazoline moiety and the X-position of the 4-anilino ring. For these variations, the electronic parameters, $\Sigma\sigma$ and σ_X have emerged as the best correlative parameters. The R_1 and R_2 were taken respectively at the para- and meta-positions with respect to the quinazoline ring N-1 and the added value of the electronic effects of these R_1- and R_2-positions is indicated by $\Sigma\sigma$. For substituents in the X-position of the 4-anilino ring, the same is given by σ_X. When a multiple regression analysis was performed, using these electronic variables along with three dummy variables, a significant correlation, Eq. (1), was obtained,

$$-\log IC_{50} = -0.973(\pm 0.209)\Sigma\sigma + 1.523(\pm 0.299)I_1 + 2.161(\pm 0.285)I_2 + 3.991(\pm 0.767)\sigma_X - 1.191(\pm 0.408)I_X + 5.963$$
$$n = 56 \quad r = 0.857 \quad s = 0.762 \quad F_{5,50} = 27.773 \qquad (1)$$

in which, n is the number of data points, r is the correlation coefficient, s is the standard deviation, F is the F-ratio between the variances of calculated and observed activities, and the $\pm$ data within the parentheses are standard errors of regression coefficients. The dummy variable I_X was assigned a value 1 for $X = CF_3$ and 0 otherwise. Similarly, the presence or absence of an $-NH_2$ substituent at R_2, in that order, is assigned a value 1 or 0 for the dummy variable I_1. Further, the 6,7-dimethoxy substitutions of the quinazoline moiety, similar to PD 153035, was highlighted through the third dummy variable I_2. Thus, $I_2 = 1$ for symmetric dimethoxy derivatives ($R_1 = R_2 =$ OMe) and $= 0$ for the remaining ones.

The F-value, obtained for Eq. (1), is significant at 99% level, and the r^2-value accounts for 73% of the variance in observed activity values. It is obvious that for a fairly large data size as this, this equation is significantly sound in statistical parlance. Compounds **38**, **45**, **46** and **53** have a little higher calculated activity value than the observed ones but no appropriate reason immediately becoming apparent for this 'outlier' behaviour. These compounds rather requires further intense experimental evaluation of their activity data. On eliminating them, regression analysis provided a much improved correlation, Eq. (2).

$$\begin{aligned} -\log IC_{50} = &-1.454(\pm 0.212)\Sigma\sigma + 1.436(\pm 0.262)I_1 + 2.067(\pm 0.251)I_2 \\ &+ 3.913(\pm 0.673)\sigma_X - 1.283(\pm 0.356)I_X + 6.017 \\ &n = 52, \quad r = 0.902, \quad s = 0.661, \quad F_{5,46} = 40.261. \end{aligned} \tag{2}$$

The r^2-value has now increased to 81%, the s-value has been lowered and the F-value, still significant at the 99% level, has significantly improved. The calculated activity values obtained using Eq. (2) and listed in Table I are in close agreement with observed ones. From the r^2-correlation matrix, given in Table II, it is clear that the variables on the right hand side of Eq. (2) are mutually orthogonal. Further, the above equation has indicated that the electronic effect of various substituents at the 6- and 7-positions in the quinazoline moiety operates in an opposite direction to that of the meta-substituents in anilino ring. For a more active compound, electron-donors both

TABLE II The r^2-intercorrelation matrix amongst the independent variables of Eq. (2)

	$\Sigma\sigma$	I_1	I_2	σ_X	I_X
$\Sigma\sigma$	1.000	0.012	0.017	0.011	0.019
I_1		1.000	0.038	0.001	0.006
I_2			1.000	0.002	0.003
σ_X				1.000	0.047
I_X					1.000

at R_1 and R_2 and electron-acceptors at X are beneficial. Similarly, the numeral weights associated with the dummy variables I_1 and I_2 favour $R_1 = OMe$ or NH_2 and $R_2 = OMe$ substitutions, while that of I_X disfavours $X = CF_3$ substitution.

Though the combined electron-donor effect has emerged as the important criterion for the R_1- and R_2-positions, the involvement of these positions in steric interaction is also possible. This is realized for those compounds where alkoxy or hydroxy substituents are present symmetrically at these positions (**27**–**32** and **48**–**51**: Table I). For such compounds, the molar refraction parameter, MR (H = 0.103, F = 0.092, Cl = 0.603, Br = 0.888, I = 1.394, CF_3 = 0.502, OH = 0.285, OMe = 0.787, OEt = 1.247, OPr = 1.706, OBu = 2.166) was found to be the most appropriate variable. The regression analysis, with one 'outlier' compound **31**, has resulted in a highly significant correlation, Eq. (3),

$$-\log IC_{50} = 7.765(\pm 0.765)\Sigma MR - 1.176(\pm 0.113)(\Sigma MR)^2 + 2.582(\pm 0.461)HD - 1.634$$
$$n = 9, \quad r = 0.978, \quad s = 0.360, \quad F_{3,5} = 37.373 \qquad (3)$$

where the parameter ΣMR represents the added values of molar refractivities of R_1-, R_2- and X-substitutions and HD stands for the hydrogen-donor property of $R_1 = R_2$-substitution. This parabolic equation has resulted in an optimum value of MR = 3.302, which is almost attained by the substitutions in compound **49**. This compound has evolved as the most potent compound of the present series and may serve as a lead in the synthesis of new analogues. Inclusion of the variable, HD for the sake of one compound (HD = 1 for compound **48** and 0 for the remaining ones) is unjustified. Its realibility may, however, be ascertained further when a few more compounds having $R_1 = R_2 = OH$ and X = H, F, Cl, I, CF_3 etc. become available. Removal of the variable HD (hence compound **48**) still maintained the parabolic dependence of ΣMR on the $-\log IC_{50}$, as shown in Eq. (4).

$$-\log IC_{50} = 7.766(\pm 0.765)\Sigma MR - 1.176(\pm 0.113)(\Sigma MR)^2 - 1.634$$
$$n = 8, \quad r = 0.978, \quad s = 0.360, \quad F_{2,5} = 55.005. \qquad (4)$$

The remaining compounds having unsymmetric substitutions at the R_1- and R_2-positions have resulted in the correlation in Eq. (5)

$$-\log IC_{50} = -1.514(\pm 0.190)\Sigma\sigma + 1.426(\pm 0.235)I_1 + 3.606(\pm 0.646)\sigma_X + 1.615(\pm 0.366)I_X + 6.162$$
$$n = 42, \quad r = 0.899, \quad s = 0.590, \quad F_{4,37} = 39.044. \qquad (5)$$

This equation has similar statistical results and significance to that of Eq. (2). The calculated $-\log IC_{50}$ values, using Eqs. (3) and (5) are also included in Table I.

Similar conclusion regarding substitutional requirements were drawn from the Free-Wilson analysis. In this approach, compounds **47–56** (Table I) were not included in the formulation of Free-Wilson matrix as the frequency of occurrence of certain groups at R_1 and R_2 was one. This matrix is avoided here for the sake of brevity. A total number of 46 linear equations in 16 unknowns including the parent contribution, μ, were generated and the equations were solved by the method of least squares. The results obtained are summarized in Table III. The 3 additional variables were computed using the restricted equations (i–iii), listed in Table IV, one corresponding to each of the varying positions of the parent structure. The statistical parameters of this study are

$$n = 46, \quad r = 0.926, \quad s = 0.654, \quad F_{16,29} = 10.878.$$

These parameters are, therefore, in tune with highly significant results. The F-value obtained is significant at the 99% level and the r^2-value accounts for 86% of variance in observed $-\log IC_{50}$'s. The calculated values of $-\log IC_{50}$, obtained by adding the requisite substituents contribution to μ, are in close agreement with the observed ones (Table I). From Table III, the substituents

TABLE III Substituents contributions to the EGFR tyrosine kinase inhibitory activities of 4-anilinoquinazolines

Positions	*Substitutions*	*Substituent contributions*
R_1	H	−0.012
	OMe	0.963
	NH_2	0.214
	NO_2	−1.145
R_2	H	−0.456
	OMe	0.558
	NH_2	1.310
	NO_2	−1.836
	NHAc	−0.005
	NHMe	0.416
	NMe_2	−0.924
	Cl	−0.302
X	H	−1.141
	F	−0.347
	Cl	0.468
	Br	0.442
	I	0.228
	CF_3	−0.499
	Parent contribution, $\mu = 7.616$	

TABLE IV The restricted equations corresponding to substitutional variations in Figure 1

Positions	*Equations*	
R_1	$H = -0.364(OMe + NH_2 + NO_2)$	(i)
R_2	$H = -0.769OMe - 0.615NH_2 - 0.385NO_2 - 0.231(NHMe + NMe_2)$ $- 0.154(NHAc + Cl)$	(ii)
X	$H = -0.5(F + Cl + I + CF_3) - 2.75Br$	(iii)

making the highest positive contribution to tyrosine kinase activity in the parent moiety have the following substitutional pattern:

R_1	R_2	X
OMe	NH_2	Cl
NH_2	OMe	Br
	NHMe	

From these approaches, similar conclusions regarding substitutional requirements are evident. For a compound to be potent, the Hansch analysis suggests that both the R_1- and R_2-substitutions of the quinazoline ring should possess high electron donor properties, the Free-Wilson approach, in conformity with this, assigns a more positive activity-contribution to such substituents. Likewise, the anilino substitution $X = CF_3$ adding negatively to activity is undesirable while Cl and Br having positive activity-contributions are the preferred substitutions at this position.

Thus the two analyses in the present study provide the grounds for rationalizing the substituent selection in designing more potent compounds of the series.

Acknowledgement

The financial assistance provided by the UGC, Bhopal, through the sanction of a minor research grant to PS, is thankfully acknowledged.

References

[1] Carpenter, G. (1987). *Annu. Rev. Biochem.*, **56**, 881–914.
[2] Yarden, Y. and Ullrich, A. (1988). *Annu. Rev. Biochem.*, **57**, 443–478.
[3] Shoyab, M., Plowman, G., McDonald, V., Garrett Bradley, J. and Todaro, G. (1989). *Science*, **234**, 1074–1076.
[4] Ullrich, A. and Schlessinger, J. (1990). *Cell*, **61**, 203–212.

[5] Gullick, W.J. (1991). *Brit. Med. Bull.*, **47**, 87–98.
[6] Aaronson, S.A. (1991). *Science*, **254**, 1146–1152.
[7] Elder, J.T., Fisher, G.J., Lindquist, P.B., Bennett, G.L., Pittelkow, M.R., Coffey, R.J., Ellingsworth, L., Derynck, R. and Voorhees, J.J. (1989). *Science*, **243**, 811–814.
[8] Chen, W.S., Lazar, C.S., Poenie, M., Tsien, R.J., Gill, G.N. and Rosenfeld, M.G. (1987). *Nature (London)*, **328**, 820–823.
[9] Honegger, A.M., Dull, T.J., Felder, S., Van Obberghen, E., Bellot, F., Szapary, D., Schmidt, A., Ullrich, A. and Schlessinger, J. (1987). *Cell*, **51**, 199–209.
[10] Burke, T.R. (1992). *Drugs Future*, **17**, 119–131.
[11] Fry, D.W. (1994). *Invest. Drugs*, **3**(6), 577–595.
[12] Barker, A.J., Davies, D.H., Eur. Patent Appl. No. 0 520 722 A1, 1992; Barker, A.J., Eur. Patent Appl. No. 0 566 226 A1, 1993; Ward, W.H.J., Cook, P.N., Slater, A.M., Davies, D.H., Holdgate, G.A. and Green, L.R. (1994). *Biochem. Pharmacol.*, **48**, 659–666.
[13] Dobrusin, E.M. and Fry, D.W. (1992). *Annu. Rep. Med. Chem.*, **27**, 169–178.
[14] Chang, C.-J. and Geahlen, R.L. (1992). *J. Nat. Prod.*, **55**, 1529–1560.
[15] Levitski, A. and Gazit, A. (1995). *Science*, **267**, 1782–1788.
[16] Palmer, B.D., Rewcastle, G.W., Thompson, A.M., Boyd, M., Showalter, H.D.H., Sercel, A.D., Fry, D.W., Kraker, A.J. and Denny, W.A. (1995). *J. Med. Chem.*, **38**, 58–67.
[17] Bridges, A.J., Zhou, H., Cody, D.R., Rewcastle, G.W., McMichael, A., Showalter, H.D.H., Fry, D.W., Kraker, A.J. and Denny, W.A. (1996). *J. Med. Chem.*, **39**, 267–276.
[18] Hansch, C. and Leo, A. (1979). *Substituent Constants for Correlation Analysis in Chemistry and Biology*. New York: John Wiley.