

用定量构效关系分析进行 酪氨酸激酶抑制剂的结构类型衍化

郭宗儒 后藤了* 杨光中 寺田弘* R977.3

(中国医学科学院药物研究所, 北京 100050; *德岛大学药学部, 德岛, 日本)

提要 用 QSAR 方法分析了取代的亚苄丙二腈类衍生物的定量构效关系, 结果表明苯环上的两个或两个以上的羟基与侧链上可极化的共轭链的存在是酪氨酸激酶抑制剂的必要结构, 与其它类型的抑制剂作构效分析, 衍化出对这类酶有抑制作用的基本结构。

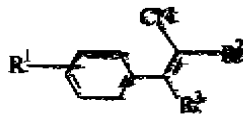
关键词 酪氨酸激酶抑制剂; 亚苄丙二腈化合物; QSAR; 摩尔折射; 基本结构衍化

酪氨酸蛋白激酶是一组酶系, 催化 ATP 的磷酸基转移到底物蛋白质的酪氨酸残基上, 使酚基磷酸化, 对传导生长因子的信号, 调节细胞的生长和繁殖起重要作用。具有酪氨酸激酶活性的生物体很多, 例如上皮细胞生长因子(EGF)受体, 血小板来源的生长因子(PDGF)受体和胰岛素生长因子受体等。某些逆转录病毒的癌基因产物以及人体的上皮癌, 皮肤癌和白血病细胞中也发现有酪氨酸激酶的活性。因此, 寻找这些酶的特异性抑制剂, 不仅有助于了解它们的生物学意义(作为工具药), 而且可能在癌的化学预防和治疗上有应用的前景。

对酪氨酸激酶有抑制作用的已知化合物类型不多, 包括合成化合物和天然产物⁽¹⁻⁴⁾。本文通过对亚苄丙二腈系列化合物的定量构效关系的研究和与其它类型化合物作结构的定性比较, 进行基本结构类型的衍化(lead generation)。

数据和方法 化合物是取代的亚苄丙二腈类衍生物^(5,6), 活性 IC_{50} 是人上皮癌细胞 A431/15 的细胞膜上皮细胞因子受体被 ATP 自磷酸化(autophosphorylation)的能力受到 50% 抑制时, 化合物的摩尔浓度, 取其倒数的对数($-\log IC_{50}$)作为活性指标。QSAR 采用 Hansch - 藤田线性自由能分析法, 参数 $\sum R$ 是指苯环上基团 R 共振效应总和数; TMR 代表与苯环呈反式键合的基团 R 的摩尔折射(molar refractivity), 数值取自 Hansch 和 Leo 的手册⁽⁷⁾和根据摩尔折射的加和性计算而得。因活性是在生理 pH 条件下测定的, 羧基的 MR 值取 COO^- 离子的 6.05, 为计算方便 MR 值缩小十倍; 指示变量(indicator variable)的定义是, 苯环上有两个或两个以上的酚羟基时, 赋值为 1, 只有一个或无羟基时, 赋值为零。化合物的结构, 活性和物化参数列于表 1。

Tab 1. The structure, biological activities, and physicochemical parameters of benzylidene malonitrile derivatives (the calculated values are from equation 3).



本文于1991年2月19日收到。

No.	R ¹	R ²	R ³	Σ R	TMR	I	-logIC ₅₀		Dev
							Obsd	Calc	
1	3,5-(OCH ₃) ₂ , 2,4(OH)	CN	H	-0.66	0.633	0	3.07	3.42	-0.35
2	4-NHCO(CH ₂) ₂ COOH	CN	H	-0.26	0.633	0	3.19	3.42	-0.23
3	3,4-OCH ₂ O-, 6-NO ₂	CN	H	-0.86	0.633	0	3.25	3.42	-0.17
4	4-OH	CN	H	-0.64	0.633	0	3.34	3.42	-0.17
5	3,5-(t-Bu) ₂ , 4-OH	CN	H	-0.90	0.633	0	3.34	3.42	-0.08
6	3-OH, 4-NO ₂	CN	H	-0.48	0.633	0	3.35	3.42	-0.07
7	4-CH=CHCOOH	CN	H	-0.21	0.633	0	3.35	3.42	-0.07
8	4-COOH	CN	H	0.13	0.633	0	3.37	3.42	-0.05
9	3-OH	CN	H	-0.64	0.633	0	3.43	3.42	0.01
10	3,4-OCH ₂ O-	CN	H	-1.02	0.633	0	3.46	3.42	0.04
11	3-OCH ₃ , 4-OH	CN	H	-1.15	0.633	0	3.70	3.42	0.28
12	3-OCH ₃ , 4-OH, 5-NO ₂	CN	H	-0.99	0.633	0	3.80	3.42	0.38
13	3-OCH ₃ , 4-OH, 5-Br	CN	H	-1.32	0.633	0	3.82	3.42	0.40
14	3-NO ₂ , 4-OH	CN	H	-0.48	0.633	0	3.90	3.42	0.48
15	3-OH, 6-NO ₂	CN	H	-0.48	0.633	0	4.02	3.42	0.60
16	4-CHO*	CN	H	0.13	0.633	0	4.22	3.42	0.80
17	3,5-(OH) ₂	CN	H	-1.28	0.633	1	4.43	4.88	-0.45
18	3,4-(OH) ₂	CN	H	-1.28	0.633	1	4.46	4.88	-0.42
19	3,4-(OH) ₂ , 5-OCH ₃	CN	H	-1.79	0.633	1	5.22	4.88	0.34
20	3,4,5-(OH) ₃	CN	H	-1.92	0.633	1	5.52	4.88	0.64
21	H	COOH	H	0.00	0.605	0	3.07	3.41	-0.34
22	4-OCH ₃	COOH	H	-0.51	0.605	0	3.08	3.41	-0.33
23	4-F	COOH	H	-0.34	0.605	0	3.08	3.41	-0.33
24	3-OCH ₃ , 4-OH	COOH	H	-1.15	0.605	0	3.49	3.41	0.08
25	3,5-(OCH ₃) ₂ , 4-OH	COOH	H	-1.66	0.605	0	3.58	3.41	0.17
26	4-OH	COOH	H	-0.64	0.605	0	3.78	3.41	0.37
27	4-CHO*	COOH	H	0.13	0.605	0	4.33	3.41	0.92
28	3,4-(OH) ₂	COOH	H	-1.28	0.605	1	4.60	4.87	-0.27
29	3-OCH ₃ , 4-OH	COOH	CH ₃	-1.15	0.605	0	2.96	3.41	-0.45
30	4-OH	CONH ₂	H	-0.64	0.981	0	3.10	3.61	-0.51
31	4-OH*	CONHCN	H	-0.64	1.990	0	3.20	4.13	-0.93
32	4-NHCO(CH ₂) ₂ COOH	CN	CH ₃	-0.26	0.633	0	3.30	3.42	-0.12
33	4-OCH ₃	CN	OH	-0.51	0.633	0	3.35	3.42	-0.07
34	4-OH	CN	OH	-0.64	0.633	0	3.64	3.42	0.22
35	3,4,5-(OH) ₃	CN	OH	-1.92	0.633	1	4.87	4.80	-0.07
36	3,5-(t-Bu) ₂ , 4-OH	CONH ₂	H	-0.90	0.981	0	4.26	3.61	0.65
37	3,4-(OH) ₂	CONH ₂	H	-1.82	0.981	1	5.00	5.03	-0.03
38	3,4-(OH) ₂	CSNH ₂	H	-1.82	1.828	1	5.62	5.53	0.09
39	4-OH*	MALOCN**	H	-0.64	2.992	0	6.90	4.48	2.42
40	3-NO ₂ , 4-OH	MALOCN**	H	-0.48	2.992	0	4.22	4.48	-0.26
41	3,4-(OH) ₂	MALOCN**	H	-1.28	2.992	1	5.60	5.94	-0.34
42	3,4,5-(OH) ₃	MALOCN**	H	-1.92	2.992	1	6.10	5.94	0.16
43	3,4-(OH) ₂ , 5-OCH ₃	MALOCN**	H	-1.45	2.992	1	5.92	5.94	-0.02
44	3,4-(OH) ₂ , 5-Br	MALOCN**	H	-2.81	2.992	1	5.30	5.94	0.36

*Not included in the derivation of equation 9. **NHC=C(CN)₂

结果与讨论 经逐步展开的多重回归计算, 44 个化合物的 QSAR 方程(或 1 ~ 7)列于表 2 中。

Tab 2. The stepwise developing equations 1 ~ 7

Eq No.	TMR	$\sum R$	I	Const	n	r	s	F
1	1.009 (0.322)			3.096 (0.386)	44	0.695	0.744	39.25
2		-0.892 (0.414)		3.262 (0.460)	44	0.553	0.862	18.55
3			1.683 (0.471)	3.621 (0.246)	44	0.741	0.695	55.11
4	0.837 (0.301)	-0.594 (0.334)		2.719 (0.402)	44	0.777	0.659	31.40
5		-0.131 (0.456)	1.556 (0.655)	3.536 (0.392)	44	0.743	0.701	25.30
6	0.699 (0.247)		1.263 (0.387)	3.054 (0.274)	44	0.864	0.528	60.37
7	0.696 (0.251)	-0.049 (0.355)	1.217 (0.514)	3.025 (0.350)	44	0.864	0.534	39.37

In the equations, n represents the number of data points used to derive the equation, r is the correlation coefficient, s is the standard deviation, F is the value of the F ratio, representing the significance of the equations. The figures in parentheses are the 95% confidence intervals of the regression coefficient.

44 个数据未加舍弃, 方程 6 用两个参数可以对全部化合物活性变化的 74.6% 加以解释 ($r^2 = 0.746$), 表明化合物活性变化的规律用 TMR 和 I 这两个结构参数可以得到大部分说明。加入 $\sum R$ 以考察电性效应对活性的影响, 回归分析表明该参数是不显著的。

检查方程 6 的计算残差, 可以发现化合物 No. 31 和 39 的残差值很大(未列出), 31 的计算值大于观测值, 39 的计算值则小于观测值, 误差均大于标准偏差的两倍。从结构上分析, No. 31 的侧链上 R_2 是 CH_2CN 基团, 故 TMR 值较大。虽然氰基是可极化性基团, 却因经 SP^3 杂化的亚甲基隔开, 使该基团未与双键乃至整个分子成共轭体系。方程 6 的 TMR 系数为正值, 故使计算值增大, 比实际活性高。关于共轭效应与活性贡献的关系, 将于后面讨论。No. 39 的活性特别高, 苯环上却只有一个羟基, 结构上无特殊变异, 故将 39 和 31 共同作为奇异点(outliers)舍弃。42 个化合物的 QSAR 最佳方程为 8:

$$-\log \text{IC}_{50} = 0.525(0.204)\text{TMR} + 1.412(0.293)\text{I} + 3.151(0.200) \quad [8]$$

$n = 42, r = 0.923, s = 0.372, F_{2,39} = 111.64$

与方程 6 比较, 方程 8 有了明显的改善, 表明剔除了结构上或活性上奇异的化合物后, 8 式可以解释化合物活性的 85% 变异。进一步考察方程 8 的计算残差, No. 16 和 27 的计算值显著低于实测值。这两个化合物的苯环上都没有酚羟基, 均为芳醛类化合物。醛基是化学上比较活泼的功能基, 可以推想 16 和 27 与酶相互作用时, 结合部位和作用机理可能有别于其它化合物。因而可将这两个化合物除去, 再经回归处理, 得方程 9:

$$-\log \text{IC} = 0.537(0.181)\text{TMR} + 1.460(0.261)\text{I} + 3.085(0.181) \quad [9]$$

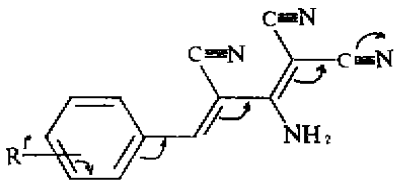
$n = 40, r = 0.943, s = 0.329, F_{2,37} = 148.20$

式 9 为最佳方程。改换或增加其它参数, 均会使显著性降低(引入 $\sum R$ 项, 方程稍有改善, 但 I 值与 $\sum R$ 有较高的共线性, $r = 0.67$, 因而不能使用)。TMR 项的系数在方程 9 中

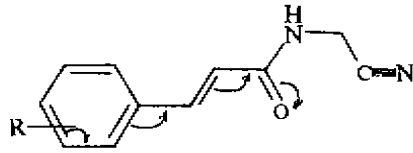
为正值,意味着与双键相连的处于苯环的反式的基团 R, 摩尔折射值越大, 对酶的抑制活性越高。摩尔折射 MR 的定义是:

$$MR = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{MW}{d}$$

式中 n 代表化合物的折射率, MW 为分子量, d 为密度。该式表明, MR 与分子或基团的折光性质呈正相关关系。皆知, 折光性质与分子或基团的可极化性 (polarizability) 密切相关。在 QSAR 分析中, 方程的 MR 项系数为负值时, 通常表示分子或基团的体积或空间效应与活性呈负性相关, 即体积越大, 活性越低; 当 MR 系数为正值时, 则表示分子或基团的可极化度与活性呈正性相关, 即极化性越大, 活性越高。分子中共轭链越长, 共轭性能越好, 可极化度越高。因此, 化合物 40 ~ 44 的侧链 R 为 $-\text{CH}=(\text{CN})_2$, 有较大的 TMR 值, 生物活性较高; 前述的 No.31 虽有较高的 TMR 值, 但 CH_2CN 基未参与共轭体系, 实际上没有极化性贡献, 因此活性的计算值高偏于实测值。下图说明了 No.40 ~ 44 的完全共轭系统和 31 的非完全共轭系统。



No. 40 ~ 44

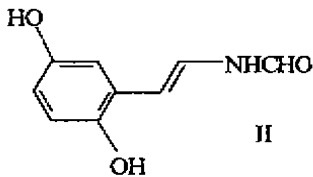


No. 31

指示变量 I 在方程 9 中的系数为 1.46, 根据定义, 化合物的苯环上含有两个或更多的酚羟基时, 有利于对酪氨酸激酶活性的抑制作用。其活性较只含一个或不含酚羟基的相应化合物活性高 27 倍 (anti log 1.46), 说明芳环上酚羟基的数目对活性的影响很大。

从以上 QSAR 分析可以看出, 分子中具有可极化基团的共轭系统和苯环上至少有两个酚羟基是呈现活性的重要结构因素。这也可从 erbstatin, 异黄酮和桂皮酰胺类化合物对酪氨酸激酶抑制作用的构效关系得到进一步证明。

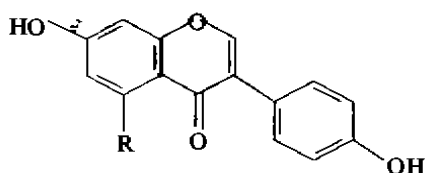
梅泽等合成的 erbstatin (II) 类似物的构效关系辅证了以上结果:⁽⁴⁾



II

Erbstatin (II) 是从霉菌 *Streptomyces meiyagawaensis* var. *crobolere* 的培养液中分离的抗菌素, 对上皮细胞生长因子受体的酪氨酸激酶的活性有很强的抑制作用 ($\text{IC}_{50} = 0.6 \mu\text{g}/\text{ml}$)⁽⁸⁾。II 的构效关系表明, 苯环上存在两个羟基也是呈现抑制活性的必要条件, 只含一个羟基化合物都没有活性。

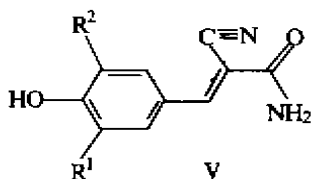
异黄酮类化合物对 EGF 受体激酶的抑制作用与结构的关系也表明, 色酮环上至少需有两个羟基方有活性 ($\text{IC}_{50} = 0.7 \text{ pg}/\text{ml}$), 例如染料木黄酮 (genistein, III) 有强抑制活性 ($\text{IC}_{50} = 0.7 \text{ pg}/\text{ml}$), 而大豆甙原 (daidzein, IV) 的色酮环上只有一个羟基, 没有活性 ($\text{IC}_{50} > 100 \text{ pg}/\text{ml}$)⁽¹⁾。



III R=OH

IV R=H

Shiraichi 等合成的单羟基取代的桂皮酰胺衍生物 (V) 却有抑制活性^(3,9), 似乎同上述的结果相悖。

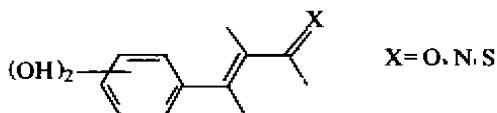


V 类化合物的构效关系表明, 当 $R^1 = R^2 = H$ 时, 不呈现活性, R^1 或 R^2 至少有一个基团是较强的疏水基团时, 例如 $R^1 = CH_2SC_6H_5$, $R^2 = OC_2H_5$ 的化合物活性很高。表明这类化合物与酪氨酸激酶的相互作用除酚羟基作用外, 还有疏水作用。而前述的 QSAR 结论是两个或多个酚羟基是必需的。因此, 可以认为酪氨酸激酶对抑制剂的结构变化的互补性(或适应性)较大。两个酚羟基可能作为氢键结合的给体与酶结合, 也可能通过一个氢键结合和一个疏水键结合。这个推论可以由化合物 36 的实测值高于计算值得到佐证。36 只含一个酚羟基和两个疏水性较强的叔丁基, 叔丁基可能参与了结合作用, 在推导方程 9 时使用疏水参数未获得显著性贡献, 是因大多数化合物的苯环上取代基的疏水性变异很小。

所以 No. 36 缺少的酚基 ($I = 0$) 可能得到了疏水结合的补偿, 因而实测值高于计算值。

两个醛基化合物 16 和 27 作为奇异点被舍弃, 未参与 QSAR 计算。舍弃无随意性。用上述的方法, 强调和突出了含醛基的亚苄丙二腈与酪氨酸激酶结合时, 可能有不同的作用位点或结合方式。

结论 由以上 QSAR 分析和对不同类型化合物的 SAR 比较, 可以认为参与同酪氨酸激酶作用的抑制剂的基本结构是: 含有两个或两个以上的酚羟基的芳环, 连结一个共轭的可极化的侧链, 侧链上存在杂原子有利于抑制作用, 如下图所示:



X=O, N, S

参 考 文 献

1. Ogawara H. et al. Inhibition of tyrosine protein kinase activity by synthetic isoflavones and flavones. *J Antibiot* 1989; 42:340.
2. Cahlen RL. et al. Inhibition of protein-tyrosine kinase activity by flavones and related compounds. *J Nat Prod* 1989; 52:982.
3. Shiraichi T. et al. Specific inhibitors of tyrosine-specific protein kinase. I. Synthesis and inhibitory activity of α -cyanocinnamamides. *Chem Pharm Bull* 1988; 36:974.
4. Isshiki K. et al. Inhibition of tyrosine protein kinase by synthetic erbstatin analogues. *J Antibiot* 1987; 40:1209.

5. Yaish P, et al. **Block**ing of EGF-dependent cell differentiation by EGF receptor kinase inhibitors. *Science* 1988; 242:933.
6. Gazit a, et al. Typhostins I. Synthesis and biological activity of protein tyrosine kinase inhibitors. *J Med Chem* 1989; 32:2344.
7. Hansch C and Leo A. *Substituent constants for correlation analysis in chemistry and biology*. New York: John Wiley & Sons. 1979.
8. Umezawa H, et al. Studies on a new epidermal growth factor-receptor kinase inhibitors, erbstatin, produced by MH435-1F3. *J Antibiot* 1986; 39:170.
9. Shiraichi T, et al. Specific inhibitors of tyrosine-specific protein kinases: properties of 4-hydroxycinnamamide derivatives *in vitro*. *Cancer Res* 1989; 49:2374.

THE LEAD GENERATION OF PROTEIN TYROSINE KINASE INHIBITORS BY QUANTITATIVE STRUCTURE - ACTIVITY RELATIONSHIP ANALYSIS

ZR Guo, S Goto*, GZ Yang and H Terada*

(*Institute of Materia Medica, CAMS, Beijing 100050; *Faculty of Pharmaceutical Sciences, University of Tokushima, Tokushima, Japan*)

ABSTRACT Analysis of the quantitative relationship between a series of substituted benzylidenemalonitriles and the inhibitory action on protein tyrosine kinases were carried out using the Hansch - Fujita method. The results indicate that the presence of two or three hydroxy groups attached to a phenyl ring, which connects an extended (trans form) conjugated chain, is essential to the inhibition. Compared with the structure of other kinds of inhibitors, a presumable basic structure for inhibiting the enzyme was generated.

Key words Protein tyrosine kinase inhibitors; Benzylidenemalonitriles; QSAR; Molar refractivity; Lead generation

更 正

本刊 1991 年第 26 卷第 12 期第 943 ~ 944 页

图 2($\times 6$)应更正为($\times 12$)

图 3 上($\times 15$)应更正为($\times 30$); 下($\times 90$)应更正为($\times 120$)