

# 维甲类化合物构效关系研究—— 维甲 X 受体特异性配体的二维和三维定量构效关系研究

黄牛 褚凤鸣 郭宗儒\*

(中国医学科学院、中国协和医科大学药物研究所, 北京 100050)

**摘要** 定向地设计作用于特定受体的维甲配体分子对于阐明维甲作用的分子机制、合成新作用类型的维甲类化合物有重要意义。本文在研究系列维甲 X 受体特异性配体的二维与三维构效关系的基础上, 考察了维甲 X 核受体对其配体的结构及电性需求, 以进一步指导设计和合成特异性的维甲 X 受体激动剂。

**关键词** 维甲类化合物; 维甲酸核受体; 维甲 X 核受体; Hansch 分析; 比较分子场分析

维甲类化合物(retinoids)有调节细胞增殖、分化以及细胞生理凋亡(apoptosis)的功能, 这归因于可激活相应的维甲核受体蛋白, 从而调控特定核基因的转录活性, 产生生物学效应。维甲核受体分为维甲酸核受体(retinoic acid receptor, RAR)和维甲 X 核受体(retinoid X receptor, RXR)两类, 为胞内独立而又相互关联的两大受体信号系统<sup>[1]</sup>。已知的 RAR 天然配体全反式维甲酸(ATRA)及合成的芳维甲 TTNPB(图 1)等临床应用时会出现毒副作用, 阻碍了它们临床上化学预防和治疗皮肤病及癌症的应用。有证据说明这种毒副作用与维甲分

子激活 RAR 的能力有关<sup>[2]</sup>。近年来 RXR 激动剂的研究引起了人们的普遍关注, 它们与 RAR 激动剂相比较少出现副作用, 使其有可能摆脱“维甲过高症”的阴影, 安全有效性得到提高。RXR 天然配体 9-顺式维甲酸(9-*cis*-RA, 为 ATRA 的体内代谢活化物)和合成的强效激动剂 LGD 1069(图 1), 已有临床研究的报道, 在治疗白血病、淋巴瘤、肺癌和其他实体瘤效果较好, 且很少有惯常维甲化合物的毒副作用, 尤其是 LGD 1069 不影响血液中脂类的正常代谢, 疗效令人满意<sup>[3]</sup>。

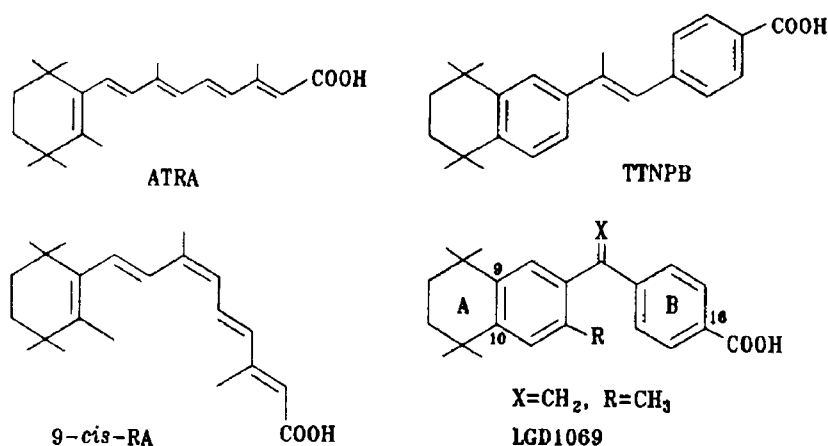


Fig 1 The ATRA and TTNPB are antagonists of RAR, and the 9-*cis*-RA and LGD 1069 is agonists of RXR.

本文于 1997 年 6 月 24 日收到。

本研究为国家自然科学基金重点项目, No 39430130

\* 联系人

本文研究了系列 RXR 激动剂的结合活性和结构之间的关系, 着重对二维定量构效关系研究: Hansch 分析和三维定量构效关系研究: 比较分子场分析(CoMFA)结果, 加以分析和比

较。并基于考察维甲 X 受体特异性配体的结构特征,进行新分子设计方面的尝试。

## 方法及结果分析

本工作中计算化学和分子图形学模拟工作是在 Silicon Graphics IRIS Indigo XZ4000 图形工作站上,用 SYBYL 6.04 软件<sup>[4]</sup>完成,计算过程中除特殊指明外皆选取其缺省值。Hansch 分析是在 HP Vectra-VL2 4/50 微机上,用回归分析软件 Bilin 完成的。

### 1 二维定量构效关系研究

定量构效关系是研究一组化合物的活性与其本身的疏水性质、立体性质和电子效应等的关系,利用取代基理化参数作为描述分子性质的指标,用数理统计的方法揭示化合物的活性与

上述结构特征的变化规律,并以数学模型概括和表达,为设计新药提供结构和物理化学依据。

本文研究的维甲 X 受体的特异性配体来源于文献<sup>[5]</sup>,结合活性和有显著性的物化参数列于表 1。采用线性自由能相关分析的 Hansch 方法。Kd 是各化合物与 RXR 形成复合物的离解常数,其负对数值越大,与受体的结合能力越强。各个 R 基团的疏水参数  $\pi$ 、立体参数 Es 等取自 Hansch 和 Leo 合编的手册<sup>[6]</sup>。指示变量 I 是为了区分两类化合物的结构特征:取代的二苯酮类(X=O)I=0(No 01-08);1,1-二苯乙烯类(X=CH<sub>2</sub>)I=1(No 09-18)。LUMO (ev)为半经验量子化程序 MOPAC 6.0(AM1)<sup>[7]</sup>计算的分子最低空轨道值。通过对多个参数的分步展开回归分析,有意义的参数和方程列于表 2。

**Tab 1 The physical and chemical parameters of substituents in Hansch analysis, and the observed and calculated activities by Hansch analysis and CoMFA**

No	R	X	$\pi$	Es	I	LUMO	Activity -log Ki(nmol·L <sup>-1</sup> )				
							(obs.)	(cal.) <sup>*</sup>	(resd.) <sup>*</sup>	(cal.) <sup>**</sup>	(resd.) <sup>**</sup>
01	Me	O	0.56	-1.24	0	-1.036	6.860	7.054	-0.194	6.912	-0.052
02	Et	O	1.02	-1.31	0	-1.031	6.710	6.895	-0.185	6.829	-0.119
03	i-Pr	O	1.53	-1.71	0	-1.024	6.950	6.654	0.296	6.878	0.072
04	n-Pr	O	1.55	-1.60	0	-1.023	6.570	6.548	0.022	6.470	0.100
05	F	O	0.14	-0.46	0	-1.034	6.150	6.465	-0.315	6.226	-0.076
06	Cl	O	0.71	-0.97	0	-1.025	6.890	6.821	0.069	6.825	0.065
07	Br	O	0.86	-1.16	0	-1.006	6.980	6.889	0.091	7.000	-0.020
08	OH	O	-0.67	-0.55	0	-1.184	6.090	5.873	0.217	6.066	0.024
09	H	CH <sub>2</sub>	0.00	0.00	1	-0.639	6.820	6.696	0.124	6.778	0.042
10	Me	CH <sub>2</sub>	0.56	-1.24	1	-0.615	7.850	7.671	0.179	7.678	0.172
11	Et	CH <sub>2</sub>	1.02	-1.31	1	-0.605	7.510	7.512	-0.002	7.396	0.117
12	i-Pr	CH <sub>2</sub>	1.53	-1.70	1	-0.607	7.260	7.271	-0.011	7.358	-0.098
13	n-Pr	CH <sub>2</sub>	1.55	-1.60	1	-0.592	6.860	7.165	-0.305	6.969	-0.109
14	F	CH <sub>2</sub>	0.14	-0.46	1	-0.618	7.160	7.082	0.078	7.112	0.048
15	Cl	CH <sub>2</sub>	0.71	-0.97	1	-0.606	7.570	7.438	0.132	7.594	-0.024
16	Br	CH <sub>2</sub>	0.86	-1.16	1	-0.603	7.550	7.506	0.044	7.717	-0.167
17	OH	CH <sub>2</sub>	-0.67	-0.55	1	-0.497	6.310	6.490	-0.180	6.345	-0.035
18	OCH <sub>3</sub>	CH <sub>2</sub>	-0.02	-0.55	1	-0.558	7.020	7.079	-0.059	6.956	0.064

\* Predicted by the Hansch equation (5); \*\* Predicted by the resulting CoMFA model.

Tab 2 The regression analysis equation

$\pi$	$\pi^2$	$E_s$	I	C	n	r	s	F	Eq. No
0.33			0.59	6.42	18	0.737	0.351	8.90	(1)
		-0.44	0.62	6.16	18	0.721	0.360	8.10	(2)
0.78	-0.48			6.88	18	0.654	0.392	5.62	(3)
0.79	-0.46		0.57	6.53	18	0.881	0.254	16.19	(4)

$$\log 1/K_d = -0.599(\pm 0.23)\pi^2 + 0.493(\pm 0.33)\pi - 0.716(\pm 0.50)E_s + 0.617(\pm 0.21)I + 6.079(\pm 0.37) \\ n=18, r=0.933, s=0.200, F=21.93, \pi_{(opt)}=0.41 \quad (5)$$

式(5)是最具有统计学显著意义的方程, C 为回归方程中常数项; n 为回归分析的样本数; r 为相关系数; s 为标准偏差; F 值是统计学检验值; 括号中的数字是回归系数的 95% 可信限;  $\pi_{(opt)}$  是呈现最大结合常数的最适基团疏水常数。方程表明此类化合物的受体结合能力与 R 基团的疏水性呈抛物线关系, 基团 R 的疏水性低于 0.41 时, 随疏水性的增加而结合力增高, 到达 0.41 (接近于甲基的  $\pi$  值) 达最大值, 再增加疏水性则结合力减低。Es 系数为负值, 由于 Es 常数除 H 外均为负值, 因而取代基的立体效应对受体结合作用呈正贡献, 表明具有一定体积的 R 基团有利于受体的结合。化合物中的最大基团是丙基, 体积的变动范围不大, 所以结合力与基团体积呈直线关系, 可以预料如果引入更大的基团未必会一直呈线性关系, 增加结合性。指示变量 I 的系数为 0.617 说明 X=CH<sub>2</sub> 化合物是相应的 X=O 化合物的 4 倍活性。由以上结果可以推测这类化合物在受体结合部位存在一个容纳适当大小疏水性基团的腔穴, 以甲基为宜。增大基团的疏水性, 可能会超出此腔穴的容积, 从而不利于结合。酮基被亚甲基取代, 结合能力均普遍增高的事实说明氧原子的存在不利于与受体的结合。虽然指示变量 I 与量化计算得到的分子电性参数 LUMO 值之间显著相关 ( $r=0.980$ ), 但这仍然难以推测 I 值的本质意义。各参数之间的相关分析, 显示疏水参数  $\pi$  和立体参数 Es 之间有一定的相关性 ( $r=0.881$ )。

量化计算结果表明这系列分子最高能量占据轨道 HOMO 主要发生在 A 环部位, R 基团的变化引起 HOMO 值的变化; 最低能量空轨道

LUMO 主要发生在 B 环部位, X 基团的变化引起 LUMO 值变化。图 2 为这种轨道分布情况的示意图。受体与配体小分子相结合的主要作用力是结构互补的非共价键合作用, 分子间前沿轨道的匹配、电子从给予体分子的 HOMO 向接受体的 LUMO 的单纯转移形成电荷迁移复合物也是常见的作用和传能方式, 配体小分子的 HOMO 和 LUMO 轨道如能与受体结合部位的 LUMO 及 HOMO 轨道适配, 将增强配体结合受体的能力。虽然在 (5) 式中完全可用 LUMO 值代替指示变量 I, 但在本系列化合物中 X 基团只有两种取代, 故不能得出 B 环处的 LUMO 值与结合受体的能力相关的结论, 但可考虑修饰 X 基团和 B 环, 使这个部位的 LUMO 值发生改变, 来考察配体与受体之间是否存在这种电荷迁移作用。

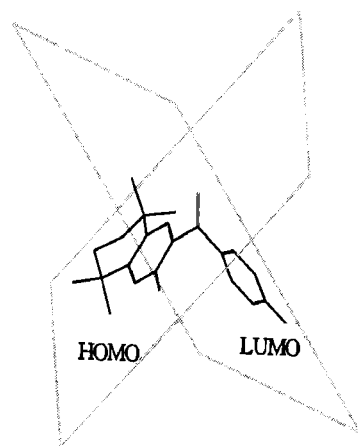


Fig 2 The orbit diagrammatic sketch calculated by MOPAC program (LGD 1069 shown as reference molecule).

## 2 三维定量构效关系研究

化合物的起始构象在 SYBYL 的“Build/Edit”下构建的。使用“系统搜寻法”(system

search)进行分子构象搜索,原子点电荷选取 Gasteiger-Huckle 方法计算,旋转单键的步长限定为  $15^\circ$ ;对获得的低能量构象进一步在 Tripos 力场中用 MAXIMIN2 进行能量优化,得到此类化合物的 2 个能量相近 ( $\Delta E$  一般在  $2 \text{ kcal}\cdot\text{mol}^{-1}$ 左右)的较稳定的低能量构象(local minimals),这两种低能量构象的能量差别列于表 3 中,由于 A 环与 X 基团之间空间位阻的影响,X 基团为亚甲基的化合物一般采取羧基尾端“向下折叠”的构象能量最低,X 基团为酮基

的化合物一般采取羧基尾端“向上伸展”的构象能量最低,通过与维甲 X 受体的天然配体 9-*cis*-RA 的最低能量构象比较,确定此类化合物与受体结合时的结合构象,空间形状为“向下折叠式”,疏水头部与极性尾部的距离在  $9.0\text{\AA}$  左右( $C_{10}-C_{16}$ ),与 Dawson 等<sup>[8,9]</sup>对其它类型有维甲 X 活性化合物优势构象的研究结果相近。图 3 为 LGD 1069 的两种低能量构象的叠合图,图 4 为 9-*cis*-RA, LGD 1069 最低能量构象的叠合图。

**Tab 3 The energy difference of two local minimals.  $E_{\text{local}}$  is the energy values presenting “upward folding” conformation,  $E_{\text{bind}}$  presents “downward folding” shape,  $\Delta E = E_{\text{local}} - E_{\text{bind}}$**

No.	01	02	03	04	05	06	07	08	09
$E_{\text{local}}$	4.159	4.641	4.701	4.002	1.225	3.204	2.601	0.525	6.446
$E_{\text{bind}}$	5.757	6.760	6.603	5.254	0.960	3.417	2.943	4.667	6.968
$\Delta E$	-1.598	-2.299	-1.902	-1.252	0.265	-0.213	-0.342	-4.142	-0.522
No.	10	11	12	13	14	15	16	17	18
$E_{\text{local}}$	7.353	8.473	8.387	8.085	4.337	5.717	6.026	6.808	8.492
$E_{\text{bind}}$	6.949	7.477	7.128	6.624	5.009	5.603	5.751	7.252	7.431
$\Delta E$	0.404	0.966	1.259	1.461	-0.672	0.114	0.275	-0.444	1.061

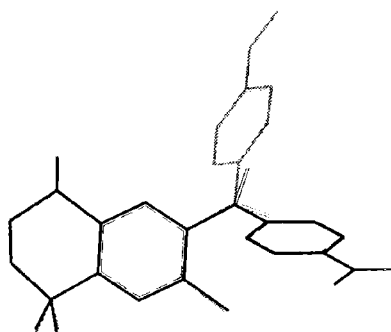


Fig 3 The superimposition of the two local minimals of LGD 1069.

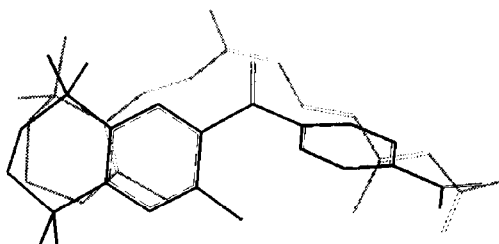


Fig 4 The overlap of the lowest energy conformation of 9-*cis*-RA and LGD 1069.

比较分子场分析法(comparative molecular field analysis, CoMFA)<sup>[10]</sup>认为药物分子与受

体之间的非共价相互作用主要是通过范德华作用和静电相互作用实现的。与同一受体相互作用的系列化合物周围的力场分布应当有一定的相似性,力场的大小和方向也应与生物活性值有相关性。这样,研究药物分子周围作用力场的分布与生物活性之间的定量关系,即可推测受体的某些性质,又可依此模型来设计新的化合物,定量地预测新分子的药效强度。

CoMFA 程序在 SYBYL 中的 QSAR 模块中运行。所有分子均采取羧基端“向下折叠”的结合构象,叠合分子头端的  $C_9$ ,  $C_{10}$  和羧基尾端的  $C_{16}$  (原子编号如图 1 中所示),空间场和静电场效应能的域值分别设为  $4.0 \text{ Kcal}\cdot\text{mol}^{-1}$  和  $15.0 \text{ Kcal}\cdot\text{mol}^{-1}$ 。围绕所有分子的三维网格是用 CoMFA 中的“自动建造”(creat automatically)选项完成,在所有叠合的化合物的三维方向延伸  $4\text{\AA}$  的距离,建造的三维网格中相邻格点之间的距离为  $2.0\text{\AA}$ ;选取依赖于距离的介电常数,探针原子选用  $sp^3$  杂化  $C^+$ ,探

针在三维网格中行进的步长设为  $2.0\text{\AA}$ 。

CoMFA 方法采用偏最小二乘法 (partial least square, PLS), 对分子周围分布的势场值与其生物活性进行回归分析。首先用交叉验证 (cross-validation) 抽一法 (leave-one-out), 检验所得模型的预测能力; 然后用非交叉验证方法 (non-validation) 计算出回归模型的表达式。由于自变量大大多于因变量, 难以用数学函数式表达, 采用直观的三维图形表示。首先进行 X

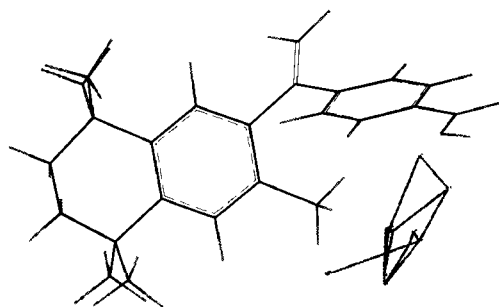


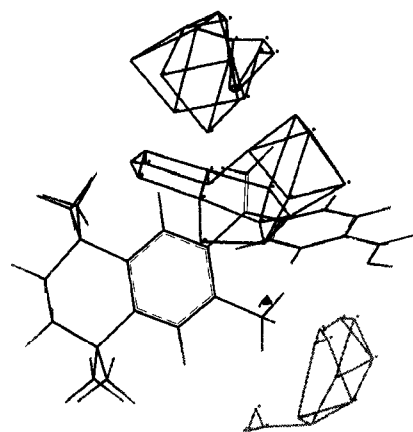
Fig 5 (a) View of the CoMFA steric field graph. More bulk in black regions, less bulk in dim regions is favorable.

CoMFA 方法中仅包括可精确计算场值的静电场和空间场, 在 Hansch 方程中采取的经验参数是疏水参数  $\pi$ 、立体参数  $E_s$  和指示变量  $I$ ; 比较 10 个化合物的 CoMFA 模型和 18 个化合物的 CoMFA 计算结果, X 基团的变化导致这个部位静电场能的变化与生物活性显著相关, 负电性增加对活性不利, 在 Hansch 方程中 X 基团的变化用  $I$  值表征,  $I$  值的真实含义可能是表征 X 基团电性的贡献; Kim 等<sup>[11]</sup>在许多研究中用 CoMFA 方法的计算结果来表征系列化合物的疏水性质, 同样 R 基团的疏水作用对化合物活性变化的贡献在 CoMFA 计算中很大程度被包含在空间场和静电场的贡献中。在表 1 中可明显看到 CoMFA 计算结果比 Hansch 方法计算的精确。

### 结果讨论

比较 ATRA 与 9-*cis*-RA, TTNPB 与 LGD 1069 的构象差别, 可设想 RAR 与 RXR 对其特

=CH<sub>2</sub> 的 10 个化合物的 CoMFA 计算, 结果为:  $R_{cv}^2 = 0.793$ , steric/electrostatic = 0.680/0.320,  $R^2 = 0.998$ ,  $F(5, 4) = 324.158$ ,  $S = 0.034$ 。然后将 X=O 的 8 个化合物添入, 建立最终的 CoMFA 模型, 结果为:  $R_{cv}^2 = 0.616$ , steric/electrostatic = 0.515/0.485,  $R^2 = 0.964$ ,  $F(5, 12) = 64.459$ ,  $S = 0.110$ , 图 5 为其空间场和静电场的三维图像显示。



b) View of the CoMFA electrostatic field graph. More positive potential in black regions, less positive in dim regions is favorable.

异性配体的主要结构需求差异是分子疏水头端与羧基尾端的长度及空间上采取的形状, 疏水头端的形状及大小需求可能是相似的, 四甲基取代的四氢萘的头端可能最适宜于与受体之间的疏水和空间作用<sup>[12]</sup>。参考 RAR 特异性配体的分子设计和构效分析, 在不带来额外的空间障碍时, 尽可能的限制单键旋转, 以降低在配体与受体相互结合时因配体的构象受到限制而发生的熵值减少, 以有利于整个系统自由能的降低。以 LGD 1069 为先导化合物, 设计了限制单键旋转的系列化合物 (图 6), 保留 R 基团为甲基、X 基团为亚甲基的结构特征, 并引入正电性的 N 杂原子, 以考察此处的电性需求。进行 MOPAC 计算, 发现这些新分子的分子轨道的显示图像与图 2 相符, 数值与 LGD 1069 的相近。这些分子的最低能量构象与 LGD 1069 进行叠合, 如图 7 所示, 发现疏水头部与尾部羧基叠合的较好, 可能会有较好的 RXR 结合活性。

并且对叠合情况最好的分子 1 和分子 2 用建立的 CoMFA 模型进行了活性的预测, 分别为  $7.368 \pm (-0.107)$ ,  $7.748 \pm (0.533)$ , 有可能比

先导物 LGD 1069 有更强的 RXR 结合活性, 同时分子脂溶性的增加可能在细胞水平和整体水平上也改善药物分子的药代动力学过程。

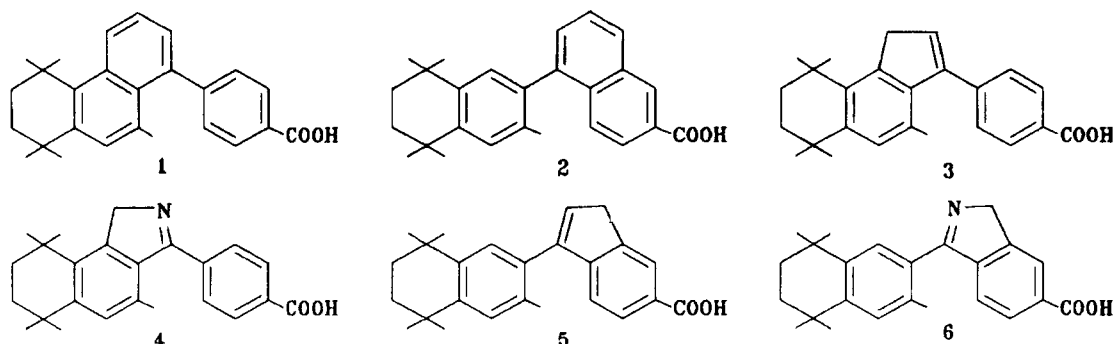


Fig 6 Structure of the designed compounds predicted to bind RXR well.

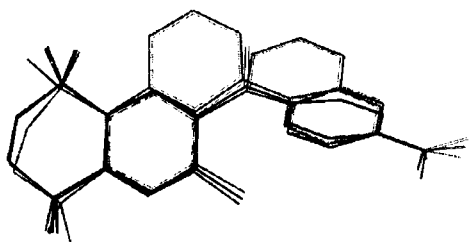


Fig 7 Alignment of designed molecules and LGD 1069.

最近 Moras 等成功的解析出 RXR 的晶体结构数据<sup>[13]</sup>, 单独的受体蛋白结晶生成的晶体结构与受体和配体形成复合物的结构可能会有些差异, 但通过分子生物学对受体定点突变的实验数据, 将较易确定 RXR 的配体结合腔穴的情况。今后将有效结合 RXR 特异性配体构效关系研究的结果与受体提供的结构信息, 期待演化出新结构类型的有特异性激活 RXR 能力的维甲类分子。

### 参 考 文 献

- Mangelsdorf DJ, Umesono K, Evans RM. The retinoid receptors. In: Sporn MB ed. *Biology, Chemistry and Medicine*. New York: Raven Press, Ltd, 1994:319~349
- Dolle P, Fraulob V, Kastner P, *et al.* Developmental expression of murine retinoid X receptor (RXR) genes. *Mech Dev*, 1994, **45**:91
- Scrip, 1994, 1932:32
- SYBYL [Computer Programme]. Version 6.04, St. Louis (MO): Tripos Associates
- Boehm MF, Zhang L, Badea BA, *et al.* Synthesis and structure-activity relationships of novel retinoid X receptor-selective retinoids. *J Med Chem*, 1994, **37**:2930
- Hansch C, Leo A. Substituent constants for correlation analysis in chemistry and biology. New York: John Wiley & Sons Press Inc, 1979:65~169
- Stewart JJP. MOPAC: A semiempirical molecular orbital program. *J Comput-Aided Mol Des*, 1990, **4**:1
- Jong L, Lehmann JM, Hobbs PD, *et al.* Conformational effects on retinoid receptor selectivity. 1. Effect of 9-double bond geometry on retinoid X receptor activity. *J Med Chem*, 1993, **36**:2605
- Lehmann JM, Jong L, Fanjul A, *et al.* Retinoids selective for retinoid X receptor response pathways. *Science*, 1992, **58**:1944
- Cramer R III, Patterson D, Bunce J. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc*, 1988, **110**:5959
- Kim KH. Description of the reversed-phase high-performance liquid chromatography (RP-HPLC) capacity factors and octanol-water partition coefficients of 2-pyrazine and 2-pyridine analogues directly from the three-dimensional structures using comparative molecular field analysis (CoMFA)

- approach. *Quant Struct-Act Relat*, 1995, **14**:8
- 12 Huang N, Wang MM, Guo ZR, *et al.* 3D-QSAR studies on retinoids of anticarcinogenic activity. *J Chin Pharm Sci*, 1996, **5**:121
- 13 Sourguet W, Ruff M, Chambon P, *et al.* Crystal structure of the ligand-binding domain of the human nuclear receptor RXR- $\alpha$ . *Nature*, 1995, **375**:377

## STUDIES ON THE STRUCTURE-ACTIVITY RELATIONSHIP OF RETINOIDS — HANSCH ANALYSIS AND 3D-QSAR STUDIES ON SPECIFIC LIGANDS OF RETINOID X RECEPTOR

Huang Niu(Huang N), Chu Fengming(Chu FM) and Guo Zongru(Guo ZR)

(*Institute of Materia Medica, Chinese Academy of Medical Sciences  
and Peking Union Medical College, Beijing 100050*)

**ABSTRACT** Retinoids (Vitamin A, its metabolites and synthetic analogues) play important roles in a variety of biological processes, including cellular differentiation, proliferation and apoptosis. The many diverse actions of retinoids attribute to the ability of regulating transcription of different target genes through activation of multiple retinoid nuclear receptors (RAR or RXR). So, retinoids with selective binding ability to specific receptor may not only have improved therapeutic indices, but may also be invaluable for elucidating the molecular mechanism of retinoidal transcriptional activation. Based on the two dimensional and three dimensional quantitative structure-activity relationships of specific ligands of RXR, we carried out mimesis of environment of ligands interacting with their receptor and, to some extent, mapping the topological and physico-chemical characteristics of receptor. The knowledge of the QSAR study will offer detailed molecular information for design, synthesis and biological evaluation in drug research and development.

**KEY WORDS** Retinoids; RAR; RXR; Hansch analysis; CoMFA