

基于 DFT 的二噁英脂-水分配系数 QSPR 研究

谷成刚, 蒋新*, 卞永荣, 余贵芬

(中国科学院南京土壤研究所土壤与农业可持续发展国家重点实验室, 南京 210008)

摘要: 基于密度泛函理论, 利用多氯二苯并对二噁英(PCDDs)和多氯二苯并呋喃(PCDFs)几何构型全优化而得的量子化学和延伸计算参数: 前线轨道能、熵、无迹四极矩、分子绝对硬度、电负性和亲电指数等, 对两类化合物在(虹鳟)脂-水间的平衡分配性质进行了定量化模型研究, 并由多元线性回归分析(MLR)建立了二次函数型的定量结构-性质相关关系(QSPR): $\lg K_{lw} = 5.343 - 0.001(S - 125.480)^2 - 0.355(\omega - 3.239)^2 + 0.006(Q_z - 2.950)^2 - 22.728(\eta - 2.365)^2$. 结果表明, 模型具有较高的拟合优度与稳健性, 其决定系数和交叉验证相关系数分别为 0.943 和 0.908, 且模型的内外预测性能理想, 总体质量明显优于 SOFA (solubility parameter for fate analysis) 计算结果, 可用于未知脂-水分配性质的预测; PCDD/Fs 脂-水平衡分配(系数)主要与分子体积以及由电荷分布决定的芳烃分子间相互作用有关, 也受潜在的生物转化和分子反应性的影响.

关键词: 二噁英; 虹鳟; 脂-水分配系数; 密度泛函理论(DFT); 定量结构-性质相关关系(QSPR)

中图分类号: X131 文献标识码: A 文章编号: 0250-3301(2008)05-1330-06

QSPR Study on the Lipid-Water Partition Coefficients of Dioxins Based on DFT

GU Cheng-gang, JIANG Xin, BIAN Yong-rong, YU Gui-fen

(State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China)

Abstract: With computational method of density functional theory (DFT), quantified model study of equilibrium partitioning properties of polychlorinated dibenzo-*p*-dioxins (PCDDs) and dibenzofurans (PCDFs) between lipid phase of organism (*Poecilia reticulata*) and water phase was carried out based on quantum chemical and further calculated parameters, namely frontier orbital energies, entropies, traceless quadrupole moments as well as molecular absolute hardness, electronegativities and electrophilicity indices, which all were derived from full geometry optimization of PCDD/Fs. Through multiple linear regression (MLR) analyses, quantitative structure-property relationship (QSPR) was successfully proposed in the form of multi-parameter quadratic function: $\lg K_{lw} = 5.343 - 0.001(S - 125.480)^2 - 0.355(\omega - 3.239)^2 + 0.006(Q_z - 2.950)^2 - 22.728(\eta - 2.365)^2$. It was shown that the obtained QSPR had higher goodness of fitting and robustness, determination coefficient and cross-validated correlation coefficient being 0.943 and 0.908 respectively, and it was also provided with ideal interior and exterior predictive abilities so that it could be used for prediction of unknown lipid-water partitioning properties. By comparison, QSPR in this research was superior to that from previous SOFA (solubility parameter for fate analysis) method on the whole. Lipid-water partitioning properties (coefficients) of PCDD/Fs should be mainly related to molecular volume and aryl hydrocarbon molecular interactions determined by charge distribution. To a certain degree, they also might be influenced by potential biotransformation and molecular reactivity.

Key words: dioxins; *Poecilia reticulata*; lipid-water partition coefficients; density functional theory (DFT); quantitative structure-property relationship (QSPR)

二噁英(PCDD/Fs)已成为环境中广泛分布、持久存在,能够对生态系统造成严重危害的典型有机污染物,多数表现出较强的毒性、环境内分泌干扰、亲脂性和生物累积,以及自然条件下难以降解消除等特性,属诸多环境问题关注的焦点.资料显示^[1-3],当二噁英水溶解度 < 350 ng/L或辛醇-水分配系数($\lg K_{ow}$) > 6 时,被称作“超亲脂”类化合物,在水环境中易浓缩于水生生物有机体内或吸附在有机碳载体颗粒上.因此,研究二噁英在水生生物体内的生物浓缩性质将为其生态风险评价和毒性作用机制解释等提供重要的科学依据.

生物浓缩可用生物浓缩系数(BCF)或脂-水分配系数(K_{lw})予以表征,其中 BCF 的计算方法很多,如 $BCF = c_{f, \exp} / c_w$, $c_{f, \exp}$ 是暴露期间实验测定的生物体

内污染物浓度($\mu\text{g} \cdot \text{kg}^{-1}$), c_w 是污染物在水体中的平均浓度($\mu\text{g} \cdot \text{L}^{-1}$); 已有研究涉及的生物测试系统多以鱼类为主,如关于虹鳟(*Poecilia reticulata*)、黑头呆鱼(*Pimephales promelas*)或虹鳉(*Salo gairdneri*)的测试研究^[3,4]. 由于生物浓缩系数测定较为复杂、PCDD/Fs 异构体繁多,前人主要针对多氯代(不小于四氯)化合物而开展了有限的工作,探究了生物浓缩系数与辛醇-水分配系数间的线性或非线性关系^[1,3,4]; 除辛醇-水分配系数外,有机碳吸附系数、分

收稿日期: 2007-05-30; 修订日期: 2007-09-07

基金项目: 国家重点基础研究发展规划(973)项目(2002CB410805); 国家杰出青年科学基金项目(40325001)

作者简介: 谷成刚(1979 ~), 男, 博士研究生, 主要研究方向为持久性有机污染物的量子化学计算及其结构-性质的相关性, E-mail: cggui@issas.ac.cn

* 通讯联系人, E-mail: jiangxin@issas.ac.cn

子体积和分子连接性指数也曾用于与 PCDDs 生物浓缩的相关性研究^[3]. Govers 等^[5]曾以脂-水分配系数为基础,利用各种实验参数建立了类似定量结构-性质相关性(QSPR)的 SOFA(solubility parameter for fate analysis)模型,并对 PCDD/Fs 脂-水分配性质进行了预测,但更多的 QSPR 相关研究还未见到.由于低氯代二噁英化合物缺乏实验测定值^[4~6],有研究^[7]以辛醇-水分配系数代替脂-水分配系数,即 $K_{lw} = K_{ow}$ 用于发展模型,但所建模型变量多依赖于不同条件或方法确定的实验参数以及分子拓扑参数等.

为建立关于 PCDD/Fs 脂-水平衡分配性质的 QSPR 模型,提高模型对未知分配性质预测的精度,本研究选择密度泛函理论(DFT)较精确计算方法,对污染物异构体几何构型全优化并获得结构参数,结合 25℃ 下虹鳉(*Poecilia reticulata*)生物浓缩的实测数据^[4]建立了完全基于电子结构参数的 QSPR,并以模型预测值与 Govers 等^[5]的结果作了比较分析.

1 材料与方法

1.1 结构参数

运用 DFT 中杂化密度泛函 B3LYP 方法^[8],在 6-311G(d,p)基组水平上^[9,10],全优化无对称性限定的 PCDD/Fs,其中包括母体化合物二苯并对二噁英(DD)和二苯并呋喃(DF);后经频率振动分析,优化结构均无虚频,表明相应构型对应于势能面上的能量极小点.所有计算均以 Gaussian 03 程序^[11]完成.

在结构参数的计算中,由于 PCDD/Fs 分子的极性相对较小,因此不再考虑偶极矩的作用,而选择了高阶电子张量(无迹)四极矩在对角线上的分向量(Q_{xx} 、 Q_{yy} 、 Q_{zz}),用以表征分子电荷在三维空间的分布或相对于球形对称的偏离程度^[12].同时,分子的最高占有轨道能级(E_{HOMO}),最低空轨道能级(E_{LUMO}),以及与分子体积有关^[13]的热力学参数——熵(S)也用作结构描述符.根据 Koopmans 理论^[14],分子的绝对硬度(η)和电负性(χ)可用以表示分子的反应性和选择性. η 越大,越不利于前线轨道能级上的电子跃迁,反应难以进行; χ 表示分子从外界吸收电子的能力.它们分别由式(1)和(2)给出:

$$\eta = \frac{E_{LUMO} - E_{HOMO}}{2} \tag{1}$$

$$\chi = -\frac{E_{HOMO} + E_{LUMO}}{2} \tag{2}$$

相应地,通过计算得到了反映最大电子转移情况下分子能量下降程度的亲电指数 ω ^[15]:

$$\omega = \frac{\chi^2}{2\eta} \tag{3}$$

用于 QSPR 模型建立的实验值取自 Loonen 等^[4]以虹鳉(*Poecilia reticulata*)为测试对象,对部分 PCDD/Fs 生物浓缩的测定,测定值以脂-水分配系数的对数形式($\lg K_{lw}$)表示;为方便在相同训练集的基础上与 SOFA 模型^[5]作比较研究,模型发展包含了部分低氯代二噁英,其实验值源自 Govers 等^[6]的近似处理,即借助了 $K_{lw} = K_{ow}$ 的关系^[7].相关数据共同列于表 1.

1.2 数据转换与统计分析

PCDD/Fs 在虹鳉体内生物浓缩、富集的影响因素众多,难以用简单的线性模型来表达脂-水分配系数与结构参数间的关系. $\lg K_{lw}$ 与 $\lg K_{ow}$ 的函数曲线表明,当 PCDD、PCDF 的 $\lg K_{ow}$ 分别等于 6.61、6.13 时,其曲线上存在 1 个拐点^[6];比较不同实验背景条件下的 $\lg K_{ow}$ 测定值,均对应于四氯取代化合物.为此,以 2,3,7,8-四氯代二苯并对二噁英(TCDD)为参考分子,对表 1 所示化合物的脂-水分配系数实验值和相应量化参数进行了转换.

$$\lg K_{lw}' = \lg K_{lw} - \lg K_{lw, TCDD} \tag{4}$$

式中, $\lg K_{lw, TCDD}$ 和 $\lg K_{lw}'$ 分别为 TCDD 原始的和其它化合物转换后的脂-水分配系数.

$$X' = (X - X_{TCDD})^2 \tag{5}$$

式中, X_{TCDD} 、 X 、 X' 分别为 TCDD 原始的和其它化合物转换前后的 DFT 计算参数.

$\lg K_{lw}'$ 和 X' 间的定量结构-性质相关关系经逐步多元线性回归分析建立,其拟合优度由决定系数(R^2)、调整决定系数(R^2_{adj})、标准误(SE)、Fischer 检验值(F)和显著性水平(p)来衡量;模型的稳健性和预测性能主要通过内部验证,即逐一去除(LOO)交叉验证相关系数(q^2_{cv})和标准预测误差(SEP)以及外部验证(模型的测试集)予以表征;交叉验证的过程如下:从样本集剔除一个样本 y_i ,由其余($n - 1$)个样本所构建的模型对剔除样本进行预测,得预测值 $\hat{y}_i$ 并与实验值 y_i 加以比较;将剔除样本放回样本集,重新选择 1 个样本进行剔除,再由剩余($n - 1$)个样本所构建的模型对相应剔除样本预测.重复以上过程,直到每个样本均被剔除且仅剔除 1 次; q^2_{cv} 由式(6)计算求得^[16].统计分析和交叉验证分别由 SPSS 13.0、MINITAB 14 软件完成.

$$q^2_{cv} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{6}$$

表 1 PCDD/Fs 的脂-水分配系数和 DFT 计算结构参数

Table 1 Lipid-water partition coefficients and DFT-calculated structural parameters of PCDD/Fs

PCDD/Fs	化合物 ¹⁾	lg <i>K</i> _{lw}	lg <i>K</i> _{lw}	残差	<i>E</i> _{HOMO} /eV	<i>E</i> _{LUMO} /eV	<i>S</i>	<i>Q</i> _{ex} /Debye•Å	<i>Q</i> _{yy} /Debye•Å	<i>Q</i> _{zz} /Debye•Å	<i>η</i> /eV	<i>ω</i> /eV	<i>χ</i> /eV
		实验值	预测值				×4.1868 /J•(mol•K) ⁻¹						
多氯二苯 并对二噁 英 (PCDDs)	DD	4.49	4.48	0.01	-5.608	-0.688	97.347	8.349	1.043	-9.392	2.460	2.014	3.148
	1	5.17	5.12	0.05	-5.833	-0.932	104.533	8.276	0.540	-8.816	2.450	2.334	3.382
	2*	5.09	4.95	0.14	-5.813	-0.955	104.807	1.469	5.055	-6.524	2.429	2.357	3.384
	2,3,7,8	5.24	5.34	-0.10	-6.279	-1.549	125.480	-14.640	11.690	2.950	2.365	3.239	3.914
	1,2,3,7,8	5.27	5.26	0.01	-6.401	-1.699	132.292	-11.655	9.134	2.521	2.351	3.489	4.050
	1,2,3,4,7,8*	5.01	5.00	0.01	-6.515	-1.851	138.948	-8.061	5.623	2.438	2.332	3.752	4.183
	1,2,3,6,7,8	4.94	5.01	-0.07	-6.523	-1.841	139.237	-8.993	6.485	2.507	2.341	3.735	4.182
	1,2,3,7,8,9	4.93	5.03	-0.10	-6.514	-1.832	138.905	-9.083	7.326	1.757	2.341	3.720	4.173
	1,2,3,4,6,7,8*	4.68	4.61	0.07	-6.628	-1.974	145.824	-5.702	3.678	2.024	2.327	3.976	4.302
多氯二苯 并呋喃 (PCDFs)	1,2,3,4,6,7,8,9	4.13	4.06	0.07	-6.734	-2.099	152.477	-2.500	0.651	1.849	2.317	4.209	4.417
	DF	3.68	3.63	0.05	-6.018	-0.938	90.762	6.569	1.996	-8.565	2.540	2.382	3.478
	1*	4.33	4.46	-0.13	-6.185	-1.158	97.749	8.354	-0.286	-8.068	2.514	2.681	3.671
	2	4.37	4.45	-0.08	-6.211	-1.234	98.017	0.976	5.126	-6.102	2.489	2.784	3.722
	3	4.35	4.55	-0.20	-6.128	-1.204	98.000	0.106	5.930	-6.036	2.462	2.729	3.666
	4*	4.44	4.39	0.05	-6.230	-1.181	97.994	4.730	3.188	-7.917	2.525	2.719	3.705
	1,4	4.99	5.08	-0.09	-6.319	-1.388	105.021	8.208	-1.616	-6.592	2.466	3.012	3.854
	1,8	4.99	5.14	-0.15	-6.308	-1.434	104.956	5.185	0.851	-6.036	2.437	3.074	3.871
	2,7*	5.02	4.86	0.16	-6.334	-1.480	105.211	-8.992	10.683	-1.692	2.427	3.145	3.907
	2,8	5.04	4.86	0.18	-6.406	-1.509	105.215	-2.636	5.399	-2.764	2.449	3.198	3.958
	4,6	5.11	4.97	0.14	-6.416	-1.401	105.171	6.432	1.094	-7.526	2.507	3.047	3.909
	2,3,4,7,8*	5.14	5.09	0.05	-6.697	-2.017	125.690	-10.785	8.183	2.602	2.340	4.057	4.357
	1,2,3,6,7,8	4.95	4.89	0.06	-6.794	-2.131	132.179	-7.828	5.676	2.151	2.331	4.272	4.463
	1,2,3,4,6,7,8	4.46	4.47	-0.01	-6.871	-2.256	138.985	-5.048	2.636	2.413	2.308	4.513	4.564
	1,2,3,4,6,7,8,9*	3.90	4.01	-0.11	-6.914	-2.388	140.068	-2.563	0.593	1.970	2.263	4.780	4.651

1)* 代表模型外部预测性能的测试集

式中, $\bar{y}_i$ 为脂-水分配系数实验值的平均值.

2 结果与讨论

2.1 QSPR 模型构建

以转换后的 $\lg K_{lw}'$ 为模型因变量,以转换后的 DFT 计算结构参数为模型自变量构建 QSPR. 为防止过度拟合,样本数(脂-水分配系数的实验值)应大于自变量数目的 4~5 倍;在既定显著性水平上($\alpha = 0.05$),经逐步多元线性回归分析,最优模型如式(7):

$$\lg K_{lw}' = 0.103(\pm 0.007) - 0.001(\pm 0)S'$$
$$- 0.355(\pm 0.084)\omega' + 0.006(\pm 0.002)Q_{zz}'$$
$$- 22.728(\pm 9.879)\eta' \tag{7}$$
$$R^2 = 0.943, R_{adj}^2 = 0.932, q_{cv}^2 = 0.908,$$
$$SE = 0.116, SEP = 0.132, F = 79.271,$$
$$p = 0.000, n = 24$$

括弧中数值为常数项和各回归系数 95% 的置信区间. 为检验各变量间的共线性问题,表 2 给出了偏相关系数(R')和方差膨胀因子(VIF).

表 2 偏相关系数、方差膨胀因子和回归系数 *t* 检验值

Table 2 Partial correlation coefficients, variance inflation factors and *t* values of variables and constant

预测变量	<i>R'</i>				VIF	<i>t</i>
	<i>S'</i>	ω'	<i>Q</i> _{zz} '	η'		
<i>S'</i>	1	0.060	0.387	0.484	2.865	-9.370
ω'		1	-0.084	0	1.009	-8.882
<i>Q</i> _{zz} '			1	0.414	2.657	7.964
η'				1	2.939	-4.815
常数项						2.231

尽管 S' 和 η' 的简单相关系数 R 达 0.768,但二者的偏相关系数 R' 仅为 0.484,说明相关性不大;其

它变量间的 R' 也均小于 0.5,而且,相应变量的方差膨胀因子(VIF)远小于 10,充分表明式(7)中多个变

量间不存在多重共线性.各变量回归系数和常数项的 t 检验值也列于表 2,其绝对值均大于 $t_{0.95}(19) = 1.729$,说明变量具有显著的统计学意义.

PCDD/Fs 脂-水分配系数的变异由转换变量,即与熵、亲电指数、四极矩和绝对硬度有关的参数 S' 、 ω' 、 Q_{zz}' 和 η' 予以解释,变异解释能力达 94.3%,且标准误差 SE 较小,仅为 0.116.结合其它统计量可判定模型的拟合优度较好,如图 1(a),复相关系数 R 为 0.971,标准差 SD 为 0.105.同时,残差分析[图 1(b)]显示,实验值与预测值之残差均匀分布在 ± 0.20 范围,表明残差的基本正态分布性.交叉验证相关系数 q_{cv}^2 高达 90.8%,标准预测误差 SEP 比标准误差 SE 仅高 0.016 个单位,说明模型的稳健性和内部预测性能较好.

为验证模型的外部预测性能,从表 1 中任意选择近 1/3 化合物构成测试集,以剩余化合物构成的训练集重新建立模型,如式(8)所示.与式(7)相比,常数项、相应变量的回归系数基本一致,尤其对于 S' 和 Q_{zz}' .以式(8)对训练集和测试集中化合物的 $\lg K_{bw}$ 预测,如图 1(a),对于训练集 R 为 0.978;对于测试集 R 为 0.958,均表明预测值与实验值良好吻合.由此可以判断,以 4 个转换变量构建的模型具备理想的外部预测性能.

$$\lg K_{bw}' = 0.064 - 0.001S' - 0.283\omega'$$
$$+ 0.005Q_{zz}' - 13.715\eta' \quad (8)$$
$$R^2 = 0.958, R_{adj}^2 = 0.942, SE = 0.109,$$
$$F = 62.134, p = 0.000, n = 16$$

2.2 模型解释

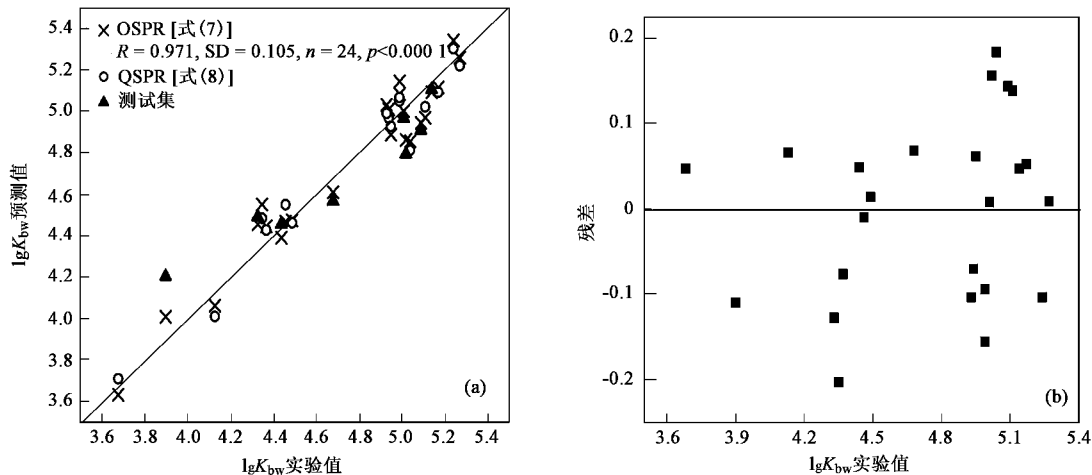


图 1 PCDD/Fs 脂-水分配系数实验值与预测值的比较($\lg K_{bw} = \lg K_{bw}' + 5.24$)和残差分析

Fig. 1 Comparison between observed and predicted $\lg K_{bw}$ of PCDD/Fs ($\lg K_{bw} = \lg K_{bw}' + 5.24$) and residual analysis

PCDD/Fs 在虹鲚体内的生物浓缩主要由脂-水间的平衡分配来决定,而影响脂-水平衡分配系数的因素可能有:脂溶解度,有效膜渗透所需的适宜分子体积和形状,生物有效性,生物转化以及与芳烃受体(AhR)的亲合等^[4].式(7)中, S' 、 ω' 、 Q_{zz}' 和 η' 的标准化回归系数分别为 -0.865 、 -0.487 、 0.708 和 -0.450 ,可说明 S' 和 Q_{zz}' 对脂-水分配系数的影响要大于其余 2 个变量,即 PCDD/Fs 在脂-水间的分配主要受到分子体积^[13]和电荷分布^[12]的影响:相对于 TCDD 而言,随着转换变量 S' 的增大,其它异构体的原始 DFT 计算变量 S 趋于 2 个极端,更小或更大,脂-水分配系数随之下降,不利于 PCDD/Fs 在脂相中的溶解;由于 S 是与分子体积有关的参数^[13],分子体积较小时,氯代程度和亲脂性低,实测分配系数应

较小;分子体积较大时,分子不易穿过生物膜而进入脂相,且溶质分子在脂相中形成“空穴”的能量较大^[4],同样导致分配系数较小.描述电荷分布在矢量 Z 方向上(垂直于芳环平面)的变化 Q_{zz} ,随间位(β -)氯取代数的变化而“分层”^[12],暗示了可能存在的芳香烃(脂)分子之间相互作用(静电、诱导或色散作用)^[3,17]对分配性质有显著的影响.当 Q_{zz} 越正(Q_{xx} 越负),PCDD/Fs 易与 AhR 亲和而发生静电吸引作用^[12],使脂-水分配系数增大.另外,脂-水分配系数随 ω' 和 η' 的增大而减小,即相对于 TCDD,PCDD/Fs 的亲电指数 ω 越大、绝对硬度 η 越小,均利于电子转移并有可能促进生物转化^[18]而使得脂-水分配系数下降;对于非 2,3,7,8-位取代的化合物,就因极易生物转化而使脂-水分配系数较小^[19].当 PCDD/Fs

的亲电指数 ω 越小、绝对硬度 η 越大,均不利于电子转移的情况下,PCDD/Fs 与脂分子形成复合物的能力下降而亦使得脂-水分配系数下降.

由于 X' 与原 DFT 计算参数 X 间存在二次函数转换关系,脂-水分配系数 $\lg K_{\text{bw}}$ 与 X 亦满足二次函数关系,如与 S 的相关系数达 66.8% (见图 2). 结合式(7),QSPR 模型还原为式(9).

$$\lg K_{\text{bw}} = 5.343 - 0.001(S - 125.480)^2 - 0.355(\omega - 3.239)^2 + 0.006(Q_{\text{z}} - 2.950)^2 - 22.728(\eta - 2.365)^2 \quad (9)$$

与先前 $\lg K_{\text{bw}}$ 和 $\lg K_{\text{ow}}$ 间的非线性模型类似^[4,5],式(9)为定量结构-性质相关的二次函数型 QSPR;其预测变量来源于 DFT 计算的结构参数,内部误差水平一致,因此能更准确地用于 PCDD/Fs 异构体未知分配性质的预测.

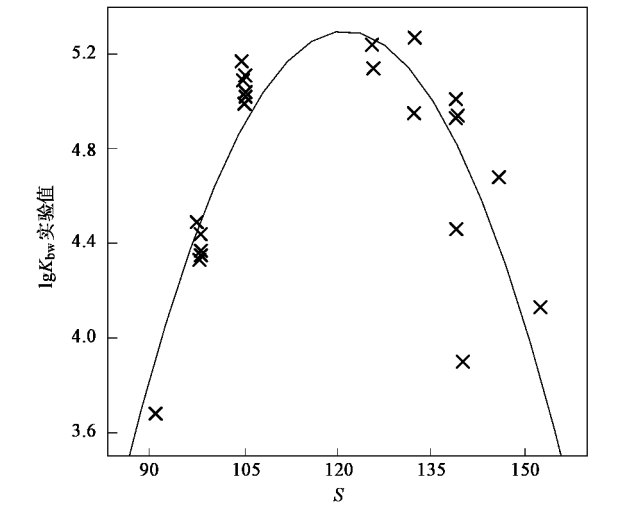


图 2 脂-水分配系数($\lg K_{\text{bw}}$)与熵(S)的抛物线关系
Fig.2 Parabolic relationship between experimental $\lg K_{\text{bw}}$ and entropies (S)

2.3 模型比较分析

由式(9)对所有 PCDD/Fs 在虹鳟体内的脂-水分配系数进行了预测.由于异构体繁多,仅列出了如表 1 所示的 24 个样本化合物的预测值.

经残差比较发现,多数低氯代化合物的残差绝对值相对较大,如 3-氯代二苯并呋喃(3-MCDF),其实验值与预测值之差为 -0.20 .这是因为低氯代化合物的亲脂性较低,在鱼体内的浓度往往低于或接近实验的检出限^[4],造成脂-水分配系数实验测定值的匮乏,并以 K_{ow} 代替 K_{bw} 做了近似处理^[6].残差绝对值经平均化后,仅为 0.087 ,远小于前人由 SOFA 方法得出的残差计算结果(0.18)^[5],说明模型的误

差水平下降.

与本研究 QSPR 具有可比性的 SOFA 模型^[5],复相关系数 R 仅为 0.87 ;相比之下,QSPR 对实验值的拟合优度显著提高,达 0.971 ,交叉验证与外部验证共同表明模型具有很高的稳健性能和预测性能,可用于其它未知脂-水分配系数的预测.然而,与 SOFA 模型的预测值相比,存在一定差别,如图 3,式(9)对多数 PCDFs 的预测值要大于 SOFA 模型的预测值;式(9)对多数 PCDDs 的预测值要小于 SOFA 模型的预测值. PCDFs 与 PCDDs 间的预测差别应归因于 Govers 等^[5]的模型研究,即对多数 PCDDs 的 $\lg K_{\text{bw}}$ 预测值大于 PCDFs.另外,用以描述分配性质的各种参数的来源不同,可能是导致 2 种模型间预测值存在差别的主要原因. SOFA 模型的重要参数,即摩尔体积 V 和聚集能密度 δ 等全部来自于不同文献的报道结果,数据的准确性有待考察;而式(9)不含经验参数,是 DFT 高精度能量计算基组 B3LYP/6-311G^{**} 优化的结果.这无疑会增加模型对性质预测的准确度而使其优于 SOFA 模型.

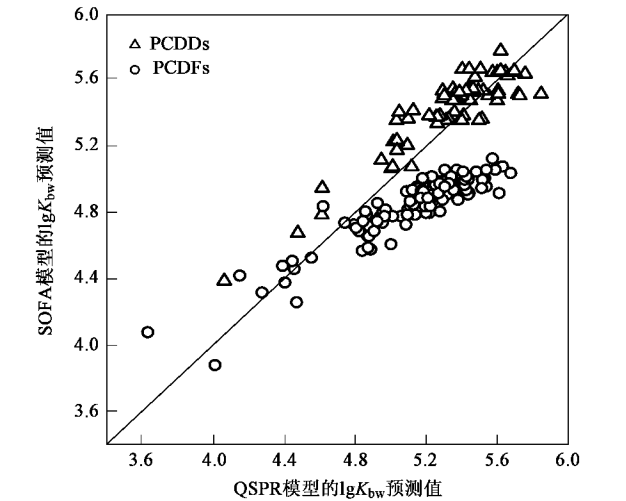


图 3 SOFA 与 QSPR 对 PCDD/Fs 脂-水分配性质预测值的比较
Fig.3 Comparison between the predicted $\lg K_{\text{bw}}$ values of PCDD/Fs by SOFA and QSPR respectively

3 结论

(1) 以 DFT-B3LYP/6-311G^{**} 基组水平上全优化几何构型而得的量化参数用于 PCDD/Fs(虹鳟)脂-水分配系数的 QSPR 研究,经多元线性回归分析成功构建的模型 $\lg K_{\text{bw}} = 5.343 - 0.001(S - 125.480)^2 - 0.355(\omega - 3.239)^2 + 0.006(Q_{\text{z}} - 2.950)^2 - 22.728(\eta - 2.365)^2$,稳健性和预测能力较高,可用于其它

化合物未知脂-水分配系数的预测, 为评估生物浓缩性质和进一步的生态风险评价提供科学依据.

(2) QSPR 模型分析表明, PCDD/Fs 脂-水分配系数主要与分子体积、电荷分布决定的芳烃(脂)分子间相互作用(静电、诱导或色散作用)有关; 在一定程度上, 也与异构体的电子转移能力有关, 即受到亲电指数和绝对硬度所描述的分子反应性和生物转化的影响.

(3) 与 SOFA 模型相比, QSPR 的总体质量大为提高. 2 种模型间预测值的差异主要与描述分配性质的参数来源, 即预测变量的选择有关. 不以文献报道值为模型参数, 而选择 DFT 计算的量子化学参数, 有助于降低模型的误差水平, 提高模型的准确度.

参考文献:

[1] Muir D C G, Marshall W K, Webster G R B. Bioconcentration of PCDDs by fish: effects of molecular structure and water chemistry [J]. Chemosphere, 1985, **14**: 829-833.

[2] Parsons J R, Storms M C M. Biodegradation of chlorinated dibenzo-*p*-dioxins in batch and continuous cultures of strain JB1 [J]. Chemosphere, 1989, **19**: 1297-1308.

[3] Voogt P de, Muir D C G, Webster G R B, *et al.* Quantitative structure-activity relationships for the bioconcentration in fish of seven polychlorinated dibenzodioxins [J]. Chemosphere, 1990, **21**: 1385-1396.

[4] Loonen H, Tonkes M, Parsons J R, *et al.* Bioconcentration of polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans in guppies after aqueous exposure to a complex PCDD/PCDF mixture: relationship with molecular structure [J]. Aquatic Toxicolgy, 1994, **30**: 153-169.

[5] Govers H A J, Krop H B. Partition constants of chlorinated dibenzofurans and dibenzo-*p*-dioxins [J]. Chemosphere, 1998, **37**: 2139-2152.

[6] Govers H A J, Loonen H, Parsons J R. Nonlinear dependence of

bioconcentration factors on *n*-octanol-water partition coefficients of chlorinated dibenzofurans and dibenzo-*p*-dioxins [J]. SAR QSAR Environ Res, 1996, **5**: 63-78.

[7] Mackay D. Correlation of bioconcentrations factors [J]. Environ Sci Technol, 1982, **16**: 274-278.

[8] Parr R G, Yang W. Density functional theory of atoms and molecules [M]. Oxford: Oxford University Press, 1989.

[9] Becke A D. Density-functional thermochemistry. III. The role of exact exchange [J]. J Chem Phys, 1993, **98**: 5648-5652.

[10] Lee C, Yang W, Parr R G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density [J]. Phys Rev B, 1988, **37**: 785-789.

[11] Frisch M J, Trucks G W, Schlegel H B, *et al.* Gaussian 03, Revision B.03 [M]. Pittsburgh PA: Gaussian Inc, 2003.

[12] Mhin B J, Lee J E, Choi W. Understanding the congener-specific toxicity in polychlorinated dibenzo-*p*-dioxins: chlorination pattern and molecular quadrupole moment [J]. J Am Chem Soc, 2002, **124**: 144-148.

[13] 李新华, 朱龙观, 俞庆森. 分子连接性指标对羧酸类疏水常数的研究 [J]. 计算机与应用化学, 2000, **17**(1): 35-36.

[14] Koopmans T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms [J]. Physica, 1934, **1**: 104-113.

[15] Parr R G, Szentpály L V, Liu S. Electrophilicity index [J]. J Am Chem Soc, 1999, **121**: 1922-1924.

[16] Wold S. Cross-validation estimation of the number of components in factor and principal components analysis [J]. Technometrics, 1978, **24**: 397-405.

[17] Kim K S, Tarakeshwar P, Lee J Y. Molecular cluster of π -systems: theoretical studies of structures, spectra, and origin of interaction energies [J]. Chem Rev, 2000, **100**: 4145-4185.

[18] Gobas F A P C, Schrap S M. Bioaccumulation of some polychlorinated dibenzo-*p*-dioxins and octachlorodibenzofuran in the guppy (*Poecilia reticulata*) [J]. Chemosphere, 1990, **20**: 495-512.

[19] Pluess N, Poiger H, Schlatter C. The metabolism of some selected pentachlorodibenzofurans in the rat [J]. Xenobiotica, 1987, **17**: 209-216.