

80种硝基多环芳烃结构及其对鼠伤寒沙门氏菌致变性能的关系

白 乃 彬

(中国科学院生态环境研究中心)

摘 要

80种硝基多环芳烃因拓扑差异而对鼠伤寒沙门氏菌致变性能不同的规律可以用以模式识别为基础的分子结构获取方法从电子效应、空间效应和分子形状大小三方面得到解释。在制约致变活性的诸因素中,电子因素影响最为重要。

在我们以前的工作中曾提出^[1]:对于27种硝基省系列化合物,由于分子拓扑差异而引起的对鼠伤寒沙门氏菌致变性能 TA_{98} 和 TA_{100} 之变化,可从电子效应、空间效应和分子大小等三方面因素得到解释。本文中,我们企图应用以模式识别为基础的计算机结构特征获取方法^[2],进一步处理扩及省、联苯、茚、萤蒽等80种硝基多环芳烃对鼠伤寒沙门氏菌致变性能 TA_{98} 数据,以求论证该观点。

1. 对鼠伤寒沙门氏菌致变性能数据

首先使用化学物质毒性数据库做硝基多环芳烃系列化合物致变性能数据联机检索^[3, 4]。以其为基础,再参阅其它文献^[5-11],理出扩及省、联苯、茚、萤蒽等80种硝基多环芳烃致变性能指标 TA_{98} 。其中,部分数据样本虽多,但分散性较大;部分数据仅有一两个实验值;有几个数据仅是区间值。作为一种数据质量评价方法^[12],可将数值转换为等级指标,以更能概括此数据集合的离散特点。同时,在定量的构效关系研究中,易于增加数据间的可比性和可比程度的真实性。根据同一文献^[1],我们将 TA_{98} 小于100者归属于非活性、弱活性类;将 TA_{98} 大于100者归属于中活性、强活性类。见表1。

从分子拓扑与 TA_{98} 数值对应关系中,我们分析得到如下经验规律:致变性能随硝基数目增多而增强;致变性能随芳烃环数增多而增强,但5环以上渐弱;当L区或湾区有硝基、或两个硝基邻位时,致变性能受到强烈抑制;比起硝基来,甲基对致变性能的激活作用似不明显。

2. 分子结构编码

在定量的构效关系研究中,设计与选择比较切题的分子结构编码是成功的重要前提。基于上述结构活性定性经验规律所提供的线索,我们设计出7种结构编码。除了以分子量、分子内所含芳环数、硝基数及甲基数为分子结构编码外,我们还建议用分子连接指数表征分子图形拓扑差异;用分子障碍面积表征分子大小;用硝基与邻近氢原子距

* 本课题得到国家自然科学基金资助。

表 1 硝基多环芳烃分子结构编码与对鼠伤寒沙门氏菌致变能力数值TA₉₈比较Table 1 Comparison of molecular structure descriptor and directacting mutagenicity in salmonella typhimurium TA₉₈ of nitrated PAHs

№	化 合 物	分 子 结 构 编 码							TA ₉₈		
		分子连接指数	分子阻碍面积	分子空间障碍	分子量	硝基数	芳环数	甲基数	均值 (REV/nmol)	实验等级	预报等级
1	硝基苯	3.21	57	1	123	1	1	0	0	—	—
2	邻位二硝基苯	3.93	58	3	168	2	1	0	0	—	—
3	间位二硝基苯	3.92	69	1	168	2	1	0	0	—	—
4	对位二硝基苯	3.92	65	1	168	2	1	0	0	—	—
5	邻位硝基甲苯	3.62	57	1	137	1	1	1	0	—	—
6	间位硝基甲苯	3.60	62	1	137	1	1	1	0	—	—
7	对位硝基甲苯	3.60	64	1	137	1	1	1	0	—	—
8	2,4-二硝基甲苯	4.35	73	1	182	2	1	1	0.06	—	—
9	3,4-二硝基甲苯	4.35	73	3	182	2	1	1	0	—	—
10	2,6-二硝基甲苯	4.36	81	1	182	2	1	1	0.02	—	—
11	4-硝基二甲苯	4.03	73	1	151	1	1	2	0	—	—
12	1-硝基萘	4.70	79	2	173	1	2	0	0.05	—	—
13	2-硝基萘	4.69	75	1	173	1	2	0	0.20	—	—
14	1,3-二硝基萘	5.50	90	2	218	2	2	0	0.90	—	—
15	1,5-二硝基萘	5.43	89	2	218	2	2	0	3.30	—	—
16	1,8-二硝基萘	5.43	77	5	218	2	2	0	0	—	—
17	2-甲基-1-硝基萘	5.11	89	2	187	1	2	1	0	—	—
18	3-甲基-2-硝基萘	5.10	75	1	187	1	2	1	1.00	—	—
19	1-甲基-2-硝基萘	5.11	88	1	187	1	2	1	0.20	—	—
20	1,3,6,8-四硝基萘	6.08	98	5	308	4	2	0	0.20	—	—
21	5-硝基二氢萘	5.18	75	2	199	1	2	2	2.50	—	—
22	2-硝基芴	5.68	91	1	211	1	2	1	14	—	—
23	2,7-二硝基芴	6.40	103	1	256	2	2	1	471	+	—
24	2-硝基联苯	5.71	92	3	199	1	2	0	0	—	—
25	3-硝基联苯	5.69	94	1	199	1	2	0	0	—	—
26	4-硝基联苯	5.69	83	1	199	1	2	0	0.51	—	—
27	2,2'-二硝基联苯	6.44	92	6	244	2	2	0	0	—	—
28	2,3'-二硝基联苯	6.42	94	3	244	2	2	0	0	—	—
29	2,4'-二硝基联苯	6.42	103	3	244	2	2	0	0	—	—
30	2,4-二硝基联苯	6.42	103	3	244	2	2	0	0	—	—
31	2,3'-二硝基联苯	6.41	97	1	244	2	2	0	0	—	—
32	3,5-二硝基联苯	6.41	106	1	244	2	2	0	0	—	—
33	4,3'-二硝基联苯	6.41	105	1	244	2	2	0	1.68	—	—
34	4,4'-二硝基联苯	6.41	92	1	244	2	2	0	14.78	—	—
35	2,4,2'-三硝基联苯	7.16	103	6	289	3	2	0	0.64	—	—
36	2,4,3'-三硝基联苯	7.15	105	3	289	3	2	0	1.07	—	—
37	2,4,6-三硝基联苯	7.16	122	5	289	3	2	0	0	—	—
38	2,4,4'-三硝基联苯	7.15	114	3	289	3	2	0	4.53	—	—
39	2,6,2'-三硝基联苯	7.17	109	6	289	3	2	0	0	—	—
40	2,6,3'-三硝基联苯	7.16	113	5	289	3	2	0	0	—	—
41	2,6,4'-三硝基联苯	7.16	122	5	289	3	2	0	0	—	—
42	3,4,3'-三硝基联苯	7.15	103	5	289	3	2	0	23.35	—	—

续表 1

№	化 合 物	分 子 结 构 编 码							TA ₉₈		
		分子连接指数	分子障碍面积	分子空间障碍	分子量	硝基数	芳环数	甲基数	均 值 (REV/ nmol)	实验 等级	预报 等级
43	3,4,4'-三硝基联苯	7.15	114	5	289	3	2	0	400.29	+	+
44	2,4,2',4'-四硝基联苯	7.88	103	6	334	4	2	0	457.19	+	-
45	2,4,2',6'-四硝基联苯	7.89	122	6	334	4	2	0	0.86	-	-
46	2,4,3',4'-四硝基联苯	7.88	103	5	334	4	2	0	33.44	-	-
47	2,6,2',6'-四硝基联苯	7.91	109	6	334	4	2	0	0	-	-
48	3,4,3',4'-四硝基联苯	7.88	103	5	334	4	2	0	715.72	+	-
49	9-硝基菲	6.19	103	2	223	1	3	0	0.5	-	-
50	2-硝基萤蒽	6.68	113	1	247	1	3	0	1030	+	+
51	3,4-二硝基萤蒽	7.43	100	5	292	2	3	0	4830	+	+
52	3,7-二硝基萤蒽	7.43	103	2	292	2	3	0	14450	+	+
53	3,9-二硝基萤蒽	7.41	122	2	292	2	3	0	121500	+	+
54	8-硝基萤蒽	6.68	109	1	247	1	3	0	11125	+	+
55	7-硝基萤蒽	6.69	104	2	247	1	3	0	544	+	+
56	3-硝基萤蒽	6.69	100	2	247	1	3	0	5439	+	+
57	1-硝基萤蒽	6.69	113	3	247	1	3	0	740	+	+
58	2-硝基菲	6.18	99	1	223	1	3	0	0.5	-	-
59	9-硝基蒽	6.19	98	3	223	1	3	0	<0.5	-	-
60	2-硝基蒽	6.17	93	1	223	1	3	0	892	+	-
61	7-硝基苯并[a]蒽	7.35	125	3	273	1	4	0	0.30	-	-
62	3-硝基芘	8.18	131	2	297	1	4	0	<0.30	-	-
63	1-硝基芘	6.68	110	2	247	1	4	0	453	+	+
64	2-硝基芘	6.67	116	1	247	1	4	0	2225	+	+
65	1,3-二硝基芘	7.41	115	2	292	2	4	0	144760	+	+
66	1,6-二硝基芘	7.41	121	2	292	2	4	0	183570	+	+
67	1,8-二硝基芘	7.41	116	2	292	2	4	0	254000	+	+
68	1,3,6-三硝基芘	8.15	144	2	337	3	4	0	40700	+	+
69	1,3,6,8-四硝基芘	8.88	144	2	382	4	4	0	15590	+	+
70	2-硝基蒾	7.67	107	1	273	1	4	0	<0.6	-	-
71	5-硝基蒾	7.68	121	3	273	1	4	0	<0.6	-	-
72	6-硝基蒾	7.68	112	2	273	1	4	0	269	+	+
73	6-硝基苯并[a]芘	8.18	134	3	297	1	5	0	0	-	-
74	1-硝基苯并[a]芘	8.17	134	2	297	1	5	0	1367	+	-
75	3-硝基苯并[a]芘	8.15	131	1	297	1	5	0	1070	+	-
76	1-硝基苯并[e]芘	9.26	145	3	297	1	5	0	39	-	-
77	3-硝基苯并[e]芘	9.35	145	2	297	1	5	0	890	+	-
78	4-硝基苯并[g,h,i]芘	8.67	131	2	321	1	6	0	<0.6	-	-
79	7-硝基苯并[g,h,i]芘	8.67	133	3	321	1	6	0	<0.3	-	-
80	1-硝基晕苯	9.15	148	2	345	1	7	0	28	-	-

离或相邻硝基最短距离表征原子间空间排斥程度。

分子连接指数是定量构效关系研究在环境化学用得最多的编码,也是较有成效的分子拓扑编码^[13]。其定义为点价乘积平方根倒数和,而点价为各自原子的价电子数减去与该原子连接的氢原子数。计算时,没有考虑芳环不饱和度的修正,但每出现一个环结

构, 便减去0.5, 以校正环状结构之影响^[13]。

分子障碍面积计算方法同文献^[1]。当分子结构为非平面结构时, 以投影方法按平面分子处理。因为依文献^[1]所述, 所谓不共面系指硝基与母体烃不共面; 母体烃本身一般共面, 且面积远大于硝基平面。投影结果的近似引起的误差不超过10%, 不会影响所讨论问题的性质。

依硝基与邻位氢原子或邻位硝基空间排斥程度, 分析80种硝基多环芳烃分子图形, 有7种情况, 可标为5个等级。如图1。前两等级中, 硝基与邻位氢原子或邻位硝基的距离大于它们的范氏半径之和, 空间排斥作用不大, 硝基与母体烃共面; 后三个等级中, 硝基与邻位氢原子或邻位硝基的距离小于它们的范氏半径之和, 存在空间排斥作用; 硝基与邻位氢原子或邻位硝基距离愈小于它们的范氏半径之和, 空间排斥作用应较大。特别是后两等级, 为了突出它们排斥作用, 予以升高一级, 记为5和6, 而不是4和5, 据此排出等级。后三等级的化合物硝基(至少有一个)与母体烃应不共面。计算所得7种分子结构编码均列入表1。

3. 特征选择

将规模为80×7的分子编码矩阵, 加权归一后, 进行Karhunen-Loeve展开, 所得特征值及其对应特征函数列于表2。

表2 分子结构编码矩阵特征值与特征函数

Table 2 Eigenvalues and Eigenfunctions for molecular structure descriptor matrix

特征值 P_i		1	2	3	4	5	6	7
		2.282	0.811	0.215	0.177	0.052	0.028	0.005
特征函数	1	-0.738*	0.187	0.073	0.059	-0.045	0.591*	0.244
	2	0.504*	0.116	0.165	-0.198	0.491	0.427	0.492
	3	-0.072	-0.353	0.833*	-0.331	-0.043	0.051	-0.250
	4	0.029	-0.374	0.030	-0.034	-0.531*	-0.189	0.732*
	5	-0.115	0.732*	0.318	-0.082	0.005	-0.507*	0.228
	6	0.410	0.383	0.072	0.007	-0.680*	0.415	-0.212
	7	0.118	-0.066	0.365	0.917*	0.077	0.016	0.050

* 系数权重较大者

将特征值依大小顺序排队, 取第一、二、三主成份, 它们对问题的贡献为:

$$P = \frac{\sum_{i=1}^3 P_i}{\sum_{i=1}^7 P_i} = 0.927$$

足以表征问题的性质。表2特征函数诸项中权重最大项可为该主成份所代表的因素。因此, 第一、二、三主成份近似代表1, 2, 5, 3四个因素, 即分子连接指数、分子障碍面积、硝基数和空间障碍因素。于是, 可淘汰三个对问题贡献小的因素, 即分子量、芳环数和甲基数。最后, 在前4个对问题中肯程度较大的因素参数基础上建立特征空间。

4. 非线性映照

在特征空间内, 进行80个样本点群的非线性映照, 所得活性二维分布图示于图2。图2中坐标 Y_1 , Y_2 系由上述第一、二主成份转化而来, 无量纲。

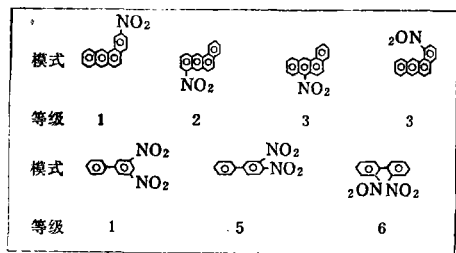


图1 硝基多环芳烃空间障碍等级图

Fig. 1 Stereo-barrier scales in nitrated polycyclic aromatic hydrocarbon molecule

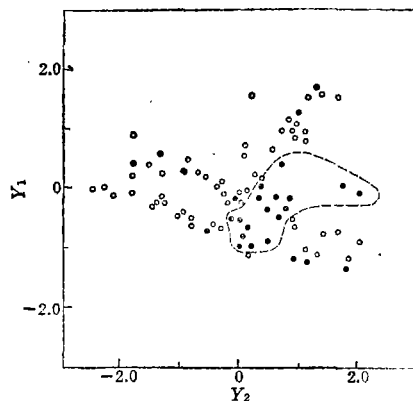


图2 80种硝基多环芳烃 TA_{98} 非线性映照图

○: $TA_{98} \leq 100$ ●: $TA_{98} > 100$

Fig. 2 Nonlinear map of TA_{98} of 80 nitrated PAHs

从图2可见: 非活性、弱活性化合物与中活性、强活性化合物明显各自归类, 类间可用曲线分开, 正确分类率达到0.913, 良好地满足统计分类和预报要求。据图2模式预报的80种硝基多环芳烃致变性能 TA_{98} 的非活性、弱活性与中活性、强活性两类分级亦列于表1中最末一项。最强致变活性区大致落于分子连接指数大于0.65, 且分子障碍面积在120平方埃附近, 即图2的偏中心位置。

主成份分析和非线性映照的原理、算法、语言及程序均在文献[14]中作了说明。计算系在PC/AT微机上实现。

5. 讨论

5.1 特征选择淘汰了三个因素: 分子量、芳环数和甲基数。前二者或因不能反映同分异构差异, 或因不能反映同环异构差异, 均对拓扑差异所引起的致变活性 TA_{98} 之明显变化反应迟钝; 而甲基对致变活性的影响, 至少在目前所能获取的数据范围内不如硝基那样明显地起激活作用。这可能是由于甲基不能象硝基那样提供较多的 π 电子, 以降低致变前兆体的能量, 使其易于产生; 再者, 起致变作用的关键部位可能就是硝基^[1]。

5.2 特征选择选中了分子障碍面积为代表分子几何因素的独立作用因素。依照Arcos提出的假设^[15], 可以解释为: 平面芳烃分子在细胞受体上的活动是落于某一定大小和一定形状区域内的; 分子面积过大, 将不利于反应部位与其相接触; 分子面积过小, 易于从反应区脱附, 使反应机会减少。该参数实际上反映的是致变化学品与生物目标受体间反应的熵因素。表1和图2示出的致变活性 TA_{98} 最强处的分子障碍面积约在

120平方埃左右, 有趣的是, 该值恰恰与我们以前的工作相符合[1, 15]。

5.3 由于分子拓扑差异, 硝基与邻近的硝基或氢原子会产生空间排斥作用, 造成分子中的硝基(至少有一个)与母体烃不共面, 破坏了整体共轭效应, 使致变活性锐减。表1列出的30个存在空间障碍的化合物中, 有29个落于非活性或弱活性类内, 占0.967。即使是余下的3,4-二硝基萤蒽(TA_{88} 为4830 REV/nmol), 也比它的同分异构体3,7-二硝基萤蒽(TA_{88} 为14450 REV/nmol)和3,9-二硝基萤蒽(TA_{88} 为121500 REV/nmol)的致变活性要小得多。这说明空间效应的确是一个制约该系列化合物致变活性的独立因素。

5.4 分子连接指数是根据分子的构型计算出来的参数, 反映的是分子中原子连接情况及分支的多寡。现已证明, 它与若干环境化学参数有关, 诸如化学品的生物体与水体间分配系数、溶剂化参数、急性毒性数据、水生生物毒性、遗传毒性等数据与其分子连接指数有线性关系[13, 16]。这似乎表明: 分子连接指数可以间接地反映电子结构因素。有趣的是, 单、双、叁硝基蒽的分子连接指数和最低空轨道能级有较好的线性关系, 见图3。图3中最低空轨道能级计算值取自文献[1]。但是, 从整体看来, 分子连接指数不似最低空轨道能级反映拓扑差异那样灵敏。

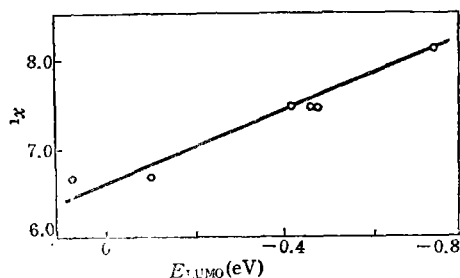


图3 硝基蒽系列化合物分子连接指数与 E_{LUMO} 关系

Fig. 3 Relationship between molecular connectivity index and their E_{LUMO} of nitrated pyrene compounds

5.5 硝基数也可间接地反映电子结构因素, 因为每增加一个硝基, 就增加几个 π 电子。另一方面, 硝基数的增加也可能导致反应机会的增加。

应该指出: 特征选择数据表明, 以分子连接指数为代表的主成份是贡献最大的第一主成份, 对整个问题的贡献达到0.639; 再加上以硝基数为代表的第二主成份, 那么电子因素对整个问题的贡献达到0.866。显然, 在制约致变性能的诸因素中电子因素最为重要。

本文工作表明: 80种硝基多环芳烃因拓扑差异而对鼠伤寒沙门氏菌致变性能不同的规律可用以模式识别为基础的分子结构获取方法从电子效应、空间效应和分子形状大小三方面得到解释。在制约致变活性的诸因素中, 以电子因素的影响最为重要。

参 考 文 献

- [1] 白乃彬, 王忠成, 1989. 硝基多环芳烃结构对鼠伤寒沙门氏菌致变性能关系研究. 环境化学, 8(6): 15
- [2] Klopman G et al., 1984. *Mutation Research*, 126(3):227
- [3] 白乃彬, 1986. 化学物质毒性库. 计算机世界, (41):20
- [4] 王忠成, 1988. 化学物质毒性效应登录数据库在VAX/VMS上联机检索的实现. 计算机与应用化学, 5(5): 218
- [5] Tokiwa H et al., 1981. *Mutation Research*, 91(5):321
- [6] Rosenkranz H S, 1983. *Mutation Research*, 114(3):217
- [7] Villiam A V et al., 1985. *Mutation Research*, 143(3):213
- [8] Motoski Nohara et al., 1985. *Mutation Research*, 149(1):9
- [9] Teruhisa Hirayma et al., 1986. *Mutation Research*, 163(2):101
- [10] Barbar Zielinska et al., 1987. *Mutation Research*, 190(4):259
- [11] Reiko Nukagawa et al., 1987. *Mutation Research*, 191(2):85
- [12] 森口郁生, 1979. 化学领域(日), 增刊(122):281
- [13] 王连生, 1989. 关于分子连接性指数与理化参数相关性探讨. 环境科学丛刊, 10(3):1
- [14] 白乃彬, 1987. 计算机在环境化学中的应用. 环境化学, 6(1):78
- [15] 白乃彬, 1980. 多核芳烃致癌K, L和Bay区理论研究. 生物化学与生物物理进展, (5):1
- [16] 白乃彬, 1990. 定量构效关系(QSAR)在环境化学中的进展与方法. 环境科学, 11(2):62

1990年1月4日收到.

THE STUDY ON THE RELATIONSHIP BETWEEN THE STRUCTURES OF 80 NITRO-PAHs AND THEIR FUNCTIONS TO SALMONELLA TYPHIMURIUM MUTATION

Bai Naibin

(Research Center for Eco-Environmental Sciences, Academia Sinica)

ABSTRACT

The regularity that the topological difference of 80 nitro-PAHs leads to the different mutation of salmonella typhimurium can be explained by the method of molecular structure attaining based on pattern recognition from the three aspects of electron effect, spatial effect and molecular shape. Among all the restricting factors the electron effect is the most important.