



预测亲核反应位点方法的比较

曹静思^①, 任庆^①, 陈飞武^{①*}, 卢天^②

① 北京科技大学化学与生物工程学院化学与化学工程系, 北京 100083

② 北京科音自然科学研究中心, 北京 100022

*通讯作者, E-mail: chenfeiwu@ustb.edu.cn

收稿日期: 2015-06-29; 接受日期: 2015-07-22; 网络版发表日期: 2015-10-23

doi: 10.1007/s11426-015-5494-7

摘要 预测分子不同位点发生亲核取代反应的活性具有重要的理论和实际意义. 目前已提出了许多基于反应物自身电子结构特点的预测方法. 本文将碳基化合物、芳香族化合物和吡啶衍生物这 3 类分子作为测试体系, 对 14 种预测方法的可靠性进行了详细的比较分析. 结果表明, 体现局部电子软度的方法能很好地预测反应位点, 如简缩双描述符方法, 但表现静电效应的预测方法, 如原子电荷分析和静电势分析, 整体表现很差. 对于本文中所用的分子体系来讲, 简缩双描述符和 Hirshfeld 电荷分析方法对分子反应位点预测最准确.

关键词

亲核反应
分子表面
福井函数
双描述符
静电势
原子电荷
轨道成分
相对亲电性

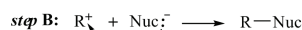
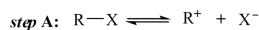
1 引言

亲核反应是有机化学反应中最基本的反应类型之一. 亲核反应有两种: 亲核取代反应和亲核加成反应. 亲核取代反应指有机化合物分子中的原子或原子团被亲核试剂所取代的反应, 包括两种主要的反应机理: 单分子亲核取代(S_N1)和双分子亲核取代(S_N2)机理. 图 1 所示为两种反应机理的示意图^[1]. 从图中可知, S_N1 反应分为两步进行, 首先是卤代烷(图 1 中 X 表示为卤素原子)缓慢解离出碳正离子, 然后碳正离子迅速与亲核试剂结合生成亲核取代产物. S_N2 反应是一个同步过程, 即亲核试剂进攻反应物形成过渡态, 紧接着过渡态分解成产物和卤素阴离子. 图 2 总结了亲核加成反应的两种反应机理: 碱催化和酸催化机理. 在碱催化亲核加成中, 首先是一个强亲核试剂进攻羰基上的碳原子形成带一个负电荷的中间体, 然后带负电荷的中间体进攻亲核试剂中的氢原子形成一个羟基产物. 而在酸催化反应中, 首先是

氢离子进攻羰基上的氧原子, 然后弱亲核试剂进攻羰基上的碳原子. 亲核反应的取代效应已在文献[2]中有详细的讨论, 本文不再赘述.

准确预测可能的亲核反应位点及反应位点的相对活性大小是理论化学上的一个重要课题. 预测反应位点的最准确可靠的方法是做从头算动力学模拟^[3], 它可以全面地考虑各种影响反应过程的效应. 但是从头算动力学计算非常耗时, 即使对于中等大小的分子体系, 其计算量也巨大, 因此从头算动力学很难用于实际问题的计算. 第二种预测反应位点的

S_N1 mechanism

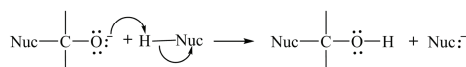
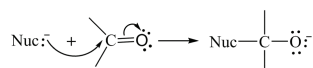


S_N2 mechanism



图 1 S_N1 和 S_N2 反应机理示意图

Basic Conditions



Acidic Conditions

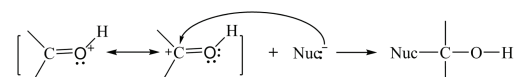
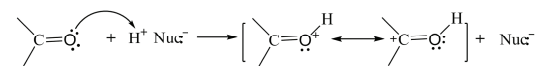


图2 亲核加成反应机理示意图

方法是过渡态理论^[4], 然而它本身就是一个近似性的理论, 其结果并不准确, 而且在一个复杂的反应中会有很多的中间体和过渡态, 利用过渡态理论考察所有的过渡态和可能的反应路径非常不易. 第三类方法对反应位点的预测完全基于反应物自身的性质, 如反应物的软度和硬度. 第三类方法有很多种, 如轨道成分分析^[5]、福井函数^[6]、双描述符^[7]、静电势^[8]、原子电荷^[9]和相对亲电性^[10]等方法均属于此类. 相对于从头算动力学模拟和过渡态理论方法, 第三类方法在预测分子的反应位点时计算量很小, 这是因为这些方法都仅考虑了化合物自身的电子结构特点, 而对于亲核试剂的电子结构、位阻效应等其他可能影响反应位点的因素都完全忽略了. 虽然这类方法的使用已经相当普遍, 但是它们的可靠性一直以来缺乏系统的检验和对比. 与我们前期对于亲电反应^[11]做的工作类似, 本文的目的在于通过对羰基化合物、芳香族化合物以及吡啶衍生物这 3 种化合物进行测试比较和分析, 最终找到最适合预测亲核反应位点的方法.

2 预测亲核反应位点的方法

2.1 福井函数

福井函数(Fukui function, FF)是由 Parr 和杨伟涛^[6]在 1984 年所提出的, 它描述了当体系电子数变化时各个位置上电子密度的变化程度. 福井函数被严格定义为:

$$f(\mathbf{r}) = \left[\frac{\delta \mu}{\delta \nu(\mathbf{r})} \right]_N = \left[\frac{\partial \rho(\mathbf{r})}{\partial N} \right]_{\nu(\mathbf{r})} \quad (1)$$

式中 μ 和 N 分别为当前体系的化学势和电子数, $\nu(\mathbf{r})$ 代表原子核对电子产生的吸引势. 由于电子密度相对于 N 的偏导数在 N 为整数时是不连续的, 所以福井函数经常利用有限差分来近似计算.

$$\begin{aligned} f^-(\mathbf{r}) &= \rho_N(\mathbf{r}) - \rho_{N-1}(\mathbf{r}) \approx \rho^{\text{HOMO}}(\mathbf{r}) \\ f^+(\mathbf{r}) &= \rho_{N+1}(\mathbf{r}) - \rho_N(\mathbf{r}) \approx \rho^{\text{LUMO}}(\mathbf{r}) \\ f^0(\mathbf{r}) &= [f^-(\mathbf{r}) + f^+(\mathbf{r})]/2 \\ &\approx [\rho^{\text{HOMO}}(\mathbf{r}) + \rho^{\text{LUMO}}(\mathbf{r})]/2 \end{aligned} \quad (2)$$

在上述等式的推导过程中, 采用了轨道冻结近似. $\rho_N(\mathbf{r})$ 、 $\rho_{N-1}(\mathbf{r})$ 和 $\rho_{N+1}(\mathbf{r})$ 分别代表体系在原始状态(N 电子)、电离掉一个电子状态($N-1$ 电子)和结合一个电子状态($N+1$ 电子)下的电子密度. $\rho^{\text{HOMO}}(\mathbf{r})$ 和 $\rho^{\text{LUMO}}(\mathbf{r})$ 分别为分子最高占据轨道(highest occupied molecular orbital, HOMO)和最低非占据轨道(lowest unoccupied molecular orbital, LUMO)的电子密度. 式(2)中的三个福井函数在本文中分别用来描述亲电、亲核和自由基反应. 通常情况下, 一个反应位点上福井函数的值越大, 其越可能是相应类型反应的活性位点.

将福井函数收缩到原子上, 我们就得到了所谓的简缩福井函数(condensed Fukui function, CFF)^[13], CFF 是预测分子反应位点的一个更简单的方法. 原子 A 上的 CFF 可被写为:

$$\begin{aligned} f_A^- &= q_{N-1}^A - q_N^A \\ f_A^+ &= q_N^A - q_{N+1}^A \\ f_A^0 &= (q_{N-1}^A - q_{N+1}^A)/2 \end{aligned} \quad (3)$$

式中 q^A 为分子中原子 A 的电荷. 值得注意的是, 原子电荷不是一个可观测的量, 而且它有多种定义方法, 因此, 不同种类原子电荷所得到的 CFF 往往存在较大差异^[13~15]. 计算 CFF 时, 本文采用 Hirshfeld 方法计算原子电荷^[15].

2.2 双描述符

双描述符(Dual descriptor, DD)^[7]是在 2005 年由 Morell、Grand 和 Toro-Labbé 提出的, 定义为:

$$f^{(2)}(\mathbf{r}) = \left[\frac{\delta \eta}{\delta \nu(\mathbf{r})} \right]_N = \left[\frac{\partial f(\mathbf{r})}{\partial N} \right]_{\nu(\mathbf{r})} \quad (4)$$

式中 η 是体系的化学硬度^[12]. 与福井函数一样, 通过有限差分的方法, 可以得到双描述符的计算公式:

$$f^{(2)}(\mathbf{r}) = f^+(\mathbf{r}) - f^-(\mathbf{r}) \approx \rho^{\text{LUMO}}(\mathbf{r}) - \rho^{\text{HOMO}}(\mathbf{r}) \quad (5)$$

Morell 等^[7]指出, 双描述符 $f^{(2)}(\mathbf{r})$ 正值(负值)的部分易受亲核试剂(亲电试剂)的进攻. 与福井函数相比, 双描述符可以同时展现亲核与亲电反应位点, 因此在预测反应位点时双描述符方法较福井函数方法更为便利. 类似于 CFF, 对于每个原子 A 上的简缩双描述符(condensed dual descriptor, CDD)也可以定义为:

$$f_A^{(2)}(\mathbf{r}) = f_A^+(\mathbf{r}) - f_A^-(\mathbf{r}) = 2q_N^A - q_{N+1}^A - q_{N-1}^A \quad (6)$$

与计算 CFF 一样, 计算 CDD 时, 本文采用 Hirshfeld 方法计算原子电荷.

2.3 前线轨道理论

1952 年 Fukui^[16]提出了前线轨道理论(frontier molecular orbitals, FMO). 前线轨道理论认为, 体系 LUMO (HOMO)成分越高的原子发生亲核(亲电)反应的可能性越高. 有多种方法可计算分子轨道中原子的成分^[16], 如 Mulliken 方法和自然原子轨道方法(natural atomic orbital method, NAOMO). 本文分析 LUMO 成分使用的是基于自然原子轨道方法, 因为这种方法对基组的依赖性小, 而且计算效率高^[17].

2.4 相对亲电性

Roy 等^[18]在 1998 年提出了相对亲电性(relative electrophilicity)和相对亲核性(relative nucleophilicity)理论, 其综合了局部软度和福井函数, 定义如下:

$$s(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{v(\mathbf{r})} \left(\frac{\partial N}{\partial \mu} \right)_{v(\mathbf{r})} = f(\mathbf{r})S = \left(\frac{\partial \mu}{\partial v(\mathbf{r})} \right)_N S \quad (7)$$

式中 $f(\mathbf{r})$ 为福井函数, $s(\mathbf{r})$ 为局部软度, S 为全局软度. 将福井函数式(2)代入到式(7)中, 可以得到:

$$\begin{aligned} s_k^+ &= [\rho_k(N+1) - \rho_k(N)]S \\ s_k^- &= [\rho_k(N) - \rho_k(N-1)]S \\ s_k^0 &= \frac{1}{2}[\rho_k(N+1) - \rho_k(N-1)]S \end{aligned} \quad (8)$$

式中 $\rho_k(N)$ 表示在 N 个电子的体系中原子 k 的电子布居数. 相对亲电性(亲核性)定义为 s_k^+/s_k^- (s_k^-/s_k^+), 分子中某个位点的 s_k^+/s_k^- (s_k^-/s_k^+) 越大, 其越容易被亲核试剂(亲电试剂)进攻, 从而发生亲核(亲电)反应.

2.5 原子电荷

原子电荷(atomic charge)是对化学体系中电荷分布最简单、最直观的描述方式之一. 它可以非常直观地被理解为原子电荷越正(越负)的原子越易发生亲

核(亲电)反应. 计算原子电荷的方法非常多, 如 Mulliken^[19]、CHELPG^[20]、NPA^[21]、Hirshfeld^[22]、ADCH^[23]. 这些方法在文献[9]中已有详细的讨论, 本文中我们将计算以上原子电荷, 并检验其预测亲核反应位点的能力.

2.6 静电势

静电势(electrostatic potential, ESP)是由库仑作用力产生的, 它描述的是电子之间以及电子与核电荷之间的相互作用. 在化学体系中, 静电势常被描述为^[24,25]:

$$V_{\text{tot}}(\mathbf{r}) = V_{\text{nuc}}(\mathbf{r}) + V_{\text{ele}}(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (9)$$

式中 Z_A 是原子 A 的核电荷, $\mathbf{R}_A$ 为原子核坐标, $\rho(\mathbf{r})$ 为电子密度. 由式(9)可见, 静电势由两部分组成, 第一部分是原子核与电子的相互作用, 第二部分为电子之间的相互作用. 在给定的区域上, $V_{\text{tot}}(\mathbf{r})$ 为正值还是负值取决于此处静电势是由核的电荷主导还是电子的电荷主导. 分子范德华(van der Waals, vdW)表面上的静电势分布长期以来被用于预测亲核反应位点^[8,25-27], 在此表面上, 静电势越正的区域越易吸引亲核试剂进攻. 我们所选用的范德华表面是由 Bader 定义的^[28], 即电子密度等于 0.001 a.u. 的等值面. 原子距离分子范德华表面上静电势最大点(最小点)位置越近, 越可能是亲核(亲电)反应的反应位点^[29].

Pearson 的硬软酸碱理论(hard-soft acid-base theory, HSAB)认为, 较硬的路易斯酸容易与较硬的路易斯碱相互作用, 较软的路易斯酸容易与较软的路易斯碱相互作用^[30]. 在亲核反应过程中, 路易斯酸和路易斯碱分别对应于缺乏电子的底物和电子丰富的亲核试剂. 基于 HSAB 的观点, 预测反应位点的方法可以分为两大类: (1) 基于静电效应并且着重表现的是不同位点局部硬度的方法, 如静电势、原子电荷; (2) 对局部软度的描述, 其体现了不同区域电子密度的变化, 如福井函数和双描述符.

亲核反应的过程可以假定分为两个阶段. 第一阶段, 通过静电吸引将亲核试剂拉到反应位点; 第二阶段, 电子发生转移并伴随着新键生成的过程. 因此, 有人建议将体现静电效应和体现局部软度的两类预测方法结合使用, 以达到更准确预测反应位点的目的^[31]. 但是如果同时考虑这两种基于不同原理的方法, 当两者的结果出现冲突时, 则很难判断哪一种方

法的预测结果是正确的。因此, 探寻预测反应位点能力最佳的单一方法仍很有必要。除了以上介绍的方法外, 还有一些其他方法可以预测亲核反应不同位点的反应活性大小, 如 Parr 函数^[32]、电子定域化函数^[33]、电子密度拉普拉斯函数^[34]、局部极化能^[35]、局部极化率^[36]。由于这些方法有的不适合本文所研究的体系, 有的很少使用, 因此本文中不再做进一步的讨论。

3 测试集和计算方法

本文的测试体系由 3 类化合物组成^[37-46], 分别为羰基类化合物、芳香烃类化合物以及吡啶衍生物。分别用“ $\gg$ ”、“ $>$ ”和“ $\approx$ ”表示各个反应位点之间活性的相对大小。“ $A \gg B$ ”表示位点 A 的反应活性要远远大于位点 B 点的反应活性; “ $A > B$ ”表示位点 A 的反应活性大于位点 B 的反应活性; 而“ $A \approx B$ ”表示位点 A 和 B 的反应活性几乎相等。为了表示出预测结果与实验结果是否一致, 我们用“ $\checkmark$ ”和“ $\times$ ”符号分别表示预测结果的正确与错误。如果一个分子中有多个反应位点, 则用符号“ $\backslash$ ”表示这种方法只能预测出最主要的反应位点。

从第二部分的讨论中可知, 测试所考虑的预测亲核位点的方法可分为体现局部软度的方法和基于静电效应的方法两类。体现局部软度的方法我们将检验以下 6 种: 福井函数(FF)、简缩福井函数(CFF)、双描述符(DD)、简缩双描述符(CDD)、LUMO 轨道成分分析以及相对亲电性(s_k^+/s_k^-)。基于静电效应的方法, 我们考虑以下 8 种, 分别为 5 种原子电荷分析方法(Hirshfeld、ADCH、CHELPG、Mulliken 和 NPA 电荷), 以及 3 种基于静电势的方法(ESP^{max}、ESP(1.6)和 ESP(reg)), 其中 ESP^{max} 表示分子范德华表面静电势极大点的数值, ESP(1.6)为原子在垂直于环平面上方 1.6 Å 处的静电势数值, ESP(reg)为分子范德华表面上每个原子区域的平均静电势。

本文中, 分子的结构优化和电子波函数的获得是通过 Gaussian 03 程序^[47]用 B3LYP/6-31+G**^[48]方法完成的。在分析部分, 除了 Mulliken 和 NPA 电荷是由 Gaussian03 程序计算得到外, 其他所有的计算包括图形绘制都是由 Multiwfn^[49,50]程序得到的。

需要指出是, 许多方法在计算化学等价的碳原子时得到的结果并不相等。为了解决这一问题, 我们

将这些等价原子的结果取平均值, 以作为每个等价原子的计算值。

4 结果与讨论

4.1 羰基化合物

羰基化合物的亲核反应位点在羰基碳上。而对于 α, β 不饱和羰基化合物, 亲核反应位点也可以在 β 位上的碳原子上发生, 但羰基上的碳原子比 β 位上的碳原子更易被亲核试剂进攻, 发生亲核反应。通常情况下, 醛类分子比酮类分子更易发生亲核加成反应。在图 3 中, 我们选取了 10 种羰基化合物进行测试。羰基化合物的亲核反应通常在 NaCN 溶液或醇溶液中进行。

表 1 给出了第一类方法对羰基类化合物预测结果。从表中可以看出, 第一类方法可以很好地预测出 C1-C5 分子的亲核反应位点。但是若分子中含有苯环或者 C=C 双键, 如 C6-C9 分子, 除了福井函数和双描述符方法外, 其他方法均不能正确地指出亲核反应的反应位点。第一类方法中, 福井函数表现最出

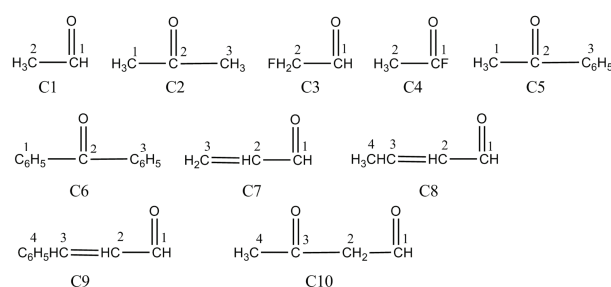


图 3 用于测试第一和第二类方法亲核反应位点预测能力的 10 种羰基化合物

表 1 第一类方法对于羰基化合物反应位点的预测结果

Mol.	Exp ^{a)}	FF	CFF	DD	CDD	LUMO	s_k^+/s_k^-
C1	1 $\gg$ 2	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
C2	2 $\gg$ 1 $\approx$ 3	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
C3	1 $\gg$ 2	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
C4	1 $\gg$ 2	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
C5	2 $\gg$ 1 $\approx$ 3	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
C6	2 $\gg$ 1 $\approx$ 3	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
C7	1 > 3 > 2	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\times$	$\times$
C8	1 > 3 > 4, 2	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\times$	$\times$
C9	1 > 3 > 2, 4	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$
C10	1 > 3 > 2, 4	$\backslash$	$\backslash$	$\checkmark$	$\times$	$\checkmark$	$\times$

a) 实验结果

色, 对图 3 中所有分子反应位点的预测结果均与实验结果一致. 虽然前线轨道理论可以视为对福井函数的近似, 但不能预测含有 C=C 双键分子的反应位点. 同样地, 相对亲电性方法在预测此类分子的反应位点时表现也很不尽人意. 双描述符和简缩双描述符方法的预测结果很不错, 但其相较福井函数的预测结果还是稍逊一筹.

表 2 列出了第二类方法对于碳基化合物亲核反应位点的预测结果. 由于 ESP(1.6)仅适用于平面分子, 因此表 2 中没有 ESP(1.6)方法的预测结果. 对这 7 种方法的预测结果进行比较发现, 5 种原子电荷方法比基于静电势的 3 种方法表现出色. 分子范德华表面静电势极大点分析方法预测结果最差, 这有两方面的原因: (1) 极大点出现在氢原子附近, 而不是在正确的反应位点, 如图 4(a)所示的丙酮体系(C2), 静电势极大点(红色的点)错误的分布在了氢原子附近; (2) 极大点往往出现在不合适的位置而导致无法判断哪个位点更可能参与反应, 如图 4(b)中苯乙酮(C5)的情况, 静电势极大点恰恰出现在苯环中心处. 若改用分

表 2 第二类方法对于碳基化合物反应位点的预测结果^{a)}

Mol.	a	b	c	d	e	f	g
C1	√	√	√	√	√	√	√
C2	√	√	√	√	√	×	√
C3	√	√	√	√	√	√	×
C4	√	√	√	√	√	×	√
C5	√	√	√	√	√	×	√
C6	√	√	√	√	√	×	√
C7	\	\	\	\	\	×	√
C8	√	√	√	√	\	×	√
C9	√	\	√	√	\	×	√
C10	×	×	×	×	×	×	×

a) a: NPA; b: Mulliken; c: Hirshfeld; d: ADCH; e: CHELPG; f: ESPmax; g: ESP(reg)

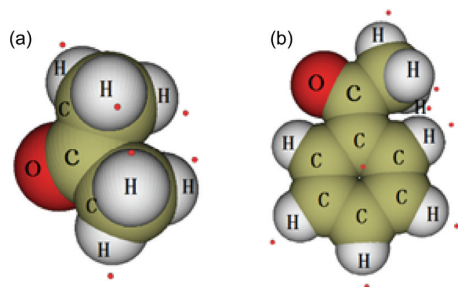


图 4 丙酮(a)和苯乙酮(b)范德华表面上静电势极大点(红色的点)的分布图(网络版彩图)

子范德华表面的平局局部静电势(ESP(reg))来讨论反应位点, 上述两方面的问题可以避免, 但 ESP(reg)的预测结果仍然很不理想. 对于 C10 分子, 第二类方法的预测结果很差, 其预测的反应位点均与实验结果不一致.

4.2 芳香烃类化合物

图 5 中所示为本文中测试的 6 种芳香类化合物. 亲核试剂可以取代卤代芳香类化合物中的卤素原子, 尤其是当卤素原子的邻位或对位上有强吸电子基时, 卤素原子更易被亲核试剂取代. 对于亲核取代芳香化合物中氢的反应, 由于 4-X-硝基苯中卤素原子(X 表示卤素原子)的存在, 亲核试剂不易进攻卤素原子所连接的碳原子, 更易进攻这类分子中未取代的 2-位上的氢原子, 从而发生亲核取代氢的反应. 芳香烃类化合物的亲核反应通常在乙醇溶液或液氨溶液中进行.

第一类方法对于芳香烃分子各个位点亲核活性的预测结果列在表 3 中. 在表 3 的 6 种方法中, LUMO 和 CDD 方法表现出色, 其他方法都出现了至少一次的判断错误, 如福井函数和双描述符均无法准确判断出氯苯(A1)分子的反应位点. 图 6 中为氯苯(A1)分子福井函数和双描述符的等值面图, 图中绿色和蓝色部分分别对应了其正值和负值区域. 从图中可以

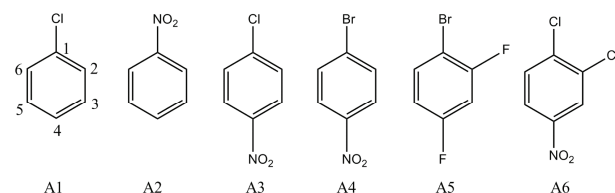


图 5 用于测试第一和第二类方法亲核反应位点预测能力的 6 种芳香烃化合物

表 3 第一类方法对于芳香烃化合物反应位点的预测结果

Mol.	A1	A2	A3	A4	A5	A6
Exp ^{a)}	4>2≈6>3	4>2≈6>3	3≈5>2	3≈5>2	2>4	1>2
FF	×	\	√	√	×	√
CFF	×	√	√	√	×	√
DD	×	\	√	√	×	√
CDD	√	√	√	√	√	√
LUMO	√	√	√	√	√	√
s_k^+/s_k^-	×	√	√	√	√	√

a) 实验结果

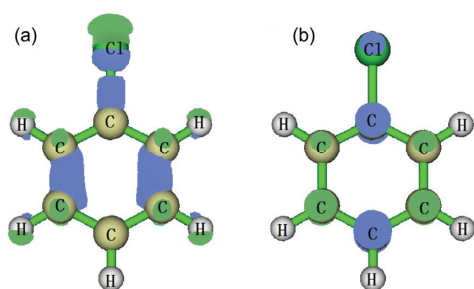


图6 氯苯的福井函数(a)和双描述符(b)的等值面图. 绿色和蓝色的部分分别对应于福井函数和双描述符的正值和负值区域(网络版彩图)

看出这两种方法给出的氯苯分子的主要反应位点为邻位碳原子, 但是实验结果表明, 氯苯的主要反应位点是对位上的碳原子.

表4列出了第二类方法对于芳香烃化合物亲核反应位点的预测结果. 可以看出, ESP(1.6)和Hirshfeld电荷表现良好. 这两种方法能较好地预测出所有分子的各个位点亲核反应活性的相对大小. 与碳基化合物的结果类似, 分子范德华表面静电势极大点分析方法结果最差.

4.3 吡啶衍生物

6种吡啶衍生物列在图7中. 吡啶和其衍生物有很明显的缺电子特性, 因此, 吡啶与其衍生物可以与

表4 第二类方法对于芳香烃化合物反应位点的预测结果

Mol.	A1	A2	A3	A4	A5	A6
NPA	√	√	√	√	×	√
Mulliken	×	√	√	√	×	√
Hirshfeld	√	√	√	√	√	√
ADCH	√	\	√	√	×	×
CHELPG	√	×	×	×	√	×
ESP ^{max}	×	×	×	×	×	√
ESP(1.6)	√	√	√	√	√	√
ESP(reg)	×	\	√	√	×	√

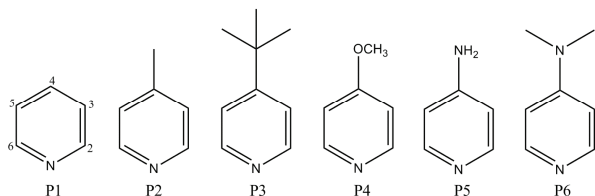


图7 用于测试第一和第二类方法亲核反应位点预测能力的6种吡啶衍生物

亲核试剂反应. 亲核试剂可以进攻并取代图7中所示吡啶衍生物的2-位或4-位. 吡啶衍生物的亲核反应通常在铵盐溶液或醇盐溶液中进行.

对于吡啶及其衍生物, 第一类测试方法的预测结果列在表5中. 可以看出, 除福井函数外, 其他所有方法都可以准确指出所有分子的反应位点并给出各位点的反应活性顺序. 福井函数并不能有效地指出反应位点有两方面的原因: (1) 福井函数最正的区域往往不能出现在正确的反应位点上, 如4-甲基吡啶(P2), 其福井函数的等值面图列在图8(a)中. 从图中可看到, 福井函数最正的区域仅出现在了氮原子和吡啶环的4-位碳原子上, 而这两个位点都不是实验给出的反应位点, 实验给出的4-甲基吡啶(P2)的亲核反应位点在吡啶环的2-位碳原子上. (2) 福井函数最正的区域平均分布在了所有的碳原子上, 致使我们无法判断哪个位点更可能参与反应, 如图8(b)所示的4-氨基吡啶的福井函数等值面图. 然而, 我们从表5的结果中发现, 如果改用简缩福井函数来预测反应位点, 上述两个问题是可以避免的. 这说明福井函数在任一点处的函数值还不足以用来预测该点的反应活性.

表6中给出了第二类预测方法对于吡啶类化合

表5 第一类方法对于吡啶衍生物反应位点的预测结果

Mol.	P1	P2	P3	P4	P5	P6
Exp ^{a)}	2 ≈ 6 > 3 ≈ 5	2 ≈ 6 > 3 ≈ 5	2 ≈ 6 > 3 ≈ 5	2 ≈ 6 > 3 ≈ 5	2 ≈ 6 > 3 ≈ 5	2 ≈ 6 > 3 ≈ 5
FF	√	×	√	√	×	×
CFF	√	√	√	√	√	√
DD	√	√	√	√	√	√
CDD	√	√	√	√	√	√
LUMO	√	√	√	√	√	√
s_k^+/s_k^-	√	√	√	√	√	√

a) 实验结果

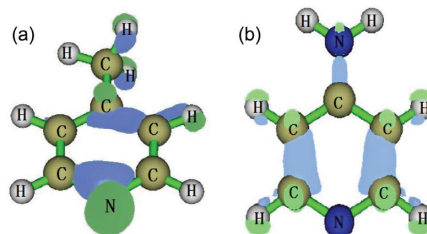


图8 4-甲基吡啶(a)和4-氨基吡啶(b)的福井函数等值面图(网络版彩图)

表 6 第二类方法对于吡啶衍生物反应位点的预测结果

Mol.	P1	P2	P3	P4	P5	P6
NPA	√	√	√	√	√	√
Mulliken	√	×	×	√	√	√
Hirshfeld	√	√	√	√	√	√
ADCH	√	√	√	√	√	√
CHELPG	√	√	√	√	√	√
ESP ^{max}	×	×	×	×	×	×
ESP(1.6)	√	×	×	\	√	\
ESP(reg)	×	×	×	\	\	\

物反应位点预测的结果. 与羰基体系中得到的结果类似, 除 Mulliken 电荷外, 其余 5 种原子电荷方法都较好地预测出分子中各个位点亲核活性的相对大小, 但 ESP^{max}、ESP(1.6)和 ESP(reg)这 3 种基于静电势的方法几乎完全预测失败, 其原因与羰基体系中讨论的一样.

5 总结

预测亲核反应活性位点具有重要的理论和实际意义. 本文介绍了各种常用亲核反应位点预测方法, 并通过对 3 类化合物, 共 22 个分子, 进行测试、比较和分析, 我们得出以下结论: (1) 对于给出的 3 种体系, 即羰基类、芳香烃类和吡啶类, 虽然不同体系有不同的适用方法, 但描述电子局部软度的方法整体上比描述静电效应的方法明显更可靠. (2) 在描述电子局部软度的方法中, 简缩双描述符方法表现最好, 它不仅预测出亲核反应位点, 通常还可以准确预测出不同位点之间相对反应活性的大小; 简缩福井函数也有很好的预测能力, 不过较简缩双描述符方法, 简缩福井函数的预测能力还是稍逊一筹. (3) 描述静电效应的预测方法通常无法正确指出反应位点. 比较本文中所用的 8 种描述静电效应方法所得到的预测结果发现, 最常用的原子电荷模型 Mulliken

电荷的预测能力非常差, 因此在预测分子的反应位点时, 要尽量避免使用 Mulliken 电荷. 出乎意料的是, 尽管 CHELPG 电荷是通过拟合分子范德华表面静电势得到的, 原理上讲其预测能力应该比其他电荷预测方法更好, 但从结果上看, CHELPG 电荷表现并不好. 在所有的原子电荷方法中, Hirshfeld 电荷表现最好. (4) 对于芳香烃和吡啶类化合物来讲, 参与亲核反应的原子可能处在一个平面上, 分析分子范德华表面静电势的极大值点虽然看似是一个很好的预测反应位点的方法, 然而我们发现, 这种方法经常得不到有意义的结果, 因为静电势的极大值常常不出现在任何反应位点的附近. 从实际效果来看, 不如直接分析环平面上方 1.6 Å 处的静电势的数值, 虽然它的预测结果在吡啶类化合物中并不理想. (5) 综合来看, CDD 是最可靠的预测各个位点的反应活性大小的方法. 因此, 在处理实际问题时, 建议使用 CDD 方法来预测化合物的反应位点.

虽然本文讨论的方法只考虑了反应物自身的电子结构特征, 而其他因素, 如亲核试剂的电子结构、位阻效应、溶剂效应、反应温度等都被忽略了, 但是对于当前的测试体系, 依然有许多方法准确地预测了分子的反应位点, 甚至各个位点的反应活性顺序. 因此, 对本文所讨论的测试体系, 被忽略的因素可认为是相对次要的. 然而在某些情况下, 它们的影响, 尤其是位阻的影响不可忽略. 这就要求要对现有的测试方法加以改进. 这样的改进也可能进一步提高这些方法的预测精度, 以至于改进后的方法不仅可以定性预测出不同位点反应活性的次序, 还有可能定量地预测出各种产物的产率.

在本文中, 我们只考察了亲核反应体系中 3 类简单的分子体系, 但亲核反应还包含了很多复杂的体系, 这些复杂体系的亲核反应位点的预测可能会给现有的预测方法带来更大的挑战, 我们将在后续工作中作进一步讨论.

致谢 本工作得到国家自然科学基金(21173020, 21473008)资助, 特此致谢. 此外, 作者衷心感谢黎乐民先生长期以来对我们的鼓励、支持和悉心指导.

参考文献

- 1 Wade Jr LG. *Organic Chemistry*. 6th Ed. New Jersey: Pearson Education International, 2006
- 2 Cao CZ, Wu YX. Recent progress in quantifying substituent effects. *Sci China Chem*, 2013, 56: 883–910

- 3 Marx D, Hutter J. *Ab Initio Molecular Dynamics-Basic Theory and Advanced Methods*. Cambridge: Cambridge University Press, 2009
- 4 Esteves PM, de M. Carneiro JW, Cardoso SP, Barbosa AGH, Laali KK, Rasul G, Surya Rrakash GK, Olah GA. Unified mechanistic concept of electrophilic aromatic nitration: convergence of computational results and experimental data. *J Am Chem Soc*, 2003, 125: 4836–4849
- 5 Jensen F. *Introduction to Computational Chemistry*. 2nd Ed. West Sussex: John Wiley & Sons, 2007. 487–492
- 6 Parr RG, Yang W. Density functional approach to the frontier-electron theory of chemical reactivity. *J Am Chem Soc*, 1984, 106: 4049–4050
- 7 Morell C, Grand A, Toro-Labbé A. New dual descriptor for chemical reactivity. *J Phys Chem A*, 2004, 109: 205–212
- 8 Sjöberg P, Politzer P. Use of the electrostatic potential at the molecular surface to interpret and predict nucleophilic processes. *J Phys Chem*, 1990, 94: 3959–3961
- 9 Lu T, Chen FW. Comparison of computational methods for atomic charges. *Acta Phys Chim Sin*, 2012, 28: 1–18
- 10 Roy RK. Stockholders charge partitioning technique. A reliable electron population analysis scheme to predict intramolecular reactivity sequence. *J Phys Chem A*, 2003, 107: 10428–10434
- 11 Fu R, Lu T, Chen FW. Comparing methods for predicting the reactive site of electrophilic substitution. *Acta Phys Chim Sin*, 2014, 30: 628–639
- 12 Liu SB. Conceptual density functional theory and some recent developments. *Acta Phys Chim Sin*, 2009, 25: 590–600
- 13 Yang W, Mortier WJ. The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines. *J Am Chem Soc*, 1986, 108: 5708–5711
- 14 Jin JL, Li HB, Lu T, Duan YA, Geng Y, Wu Y, Su ZM. Density functional studies on photophysical properties and chemical reactivities of the triarylboranes: effect of the constraint of planarity. *J Mol Model*, 2013, 19: 3437–3446
- 15 Oláh J, van Alsenoy C, Sannigrahi AB. Condensed Fukui functions derived from stockholder charges: assessment of their performance as local reactivity descriptors. *J Phys Chem A*, 2002, 106: 3885–3890
- 16 Fukui K. *Orientation and Stereoselection*. Berlin: Springer, 1970. 1–85
- 17 Lu T, Chen FW. Calculation of molecular orbital composition. *Acta Chim Sin*, 2011, 69: 2393–2406
- 18 Roy RK, Krishnamurti S, Geerlings P, Pal S. Local softness and hardness based reactivity descriptors for predicting intra- and intermolecular reactivity sequences: carbonyl compounds. *J Phys Chem A*, 1998, 102: 3746–3755
- 19 Mulliken RS. Electronic population analysis on LCAO-MO molecular wave functions. I. *J Chem Phys*, 1955, 23: 1833–1840
- 20 Breneman CM, Wiberg KB. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J Comput Chem*, 1990, 11: 361–373
- 21 Weinhold F. Natural bond orbital methods. In: Schleyer PVR. *Encyclopedia of Computational Chemistry*. West Sussex: John Wiley & Sons, 1998. 1792–1811
- 22 Hirshfeld FL. Bonded-atom fragments for describing molecular charge densities. *Theor Chem Acc*, 1977, 44: 129–138
- 23 Lu T, Chen F. Atomic dipole moment corrected hirshfeld population method. *J Theor Comp Chem*, 2012, 11: 163–183
- 24 Murray JS, Politzer P. The electrostatic potential: an overview. *WIREs Comp Mol Sci*, 2011, 1: 153–163
- 25 Lipkowitz KB, Cundari TR, Boyd DB. *Reviews in Computational Chemistry*. New York: John Wiley & Sons, 1991. 273–312
- 26 Geerlings P, Langenaeker W, De Proft F. Molecular electrostatic potentials vs. DFT descriptors of reactivity. In: Murray JS, Sen K, Eds. *Molecular Electrostatic Potentials: Concepts and Applications*. Amsterdam: Elsevier Science BV, 1996, 3: 587–617
- 27 Politzer P, Murray JS. The fundamental nature and role of the electrostatic potential in atoms and molecules. *Theor Chem Acc*, 2002, 108: 134–142
- 28 Bader RFW, Carroll MT, Cheeseman JR, Chang C. Properties of atoms in molecules: atomic volumes. *J Am Chem Soc*, 1987, 109: 7968–7979
- 29 Lu T, Chen FW. Quantitative analysis of molecular surface based on improved marching tetrahedra algorithm. *J Mol Graph Model*, 2012, 38: 314–323
- 30 Pearson RG. Hard and soft acids and bases. *J Am Chem Soc*, 1963, 85: 3533–3539
- 31 Politzer P, Murray JS. The average local ionization energy: concepts and applications. In: Toro-Labbé A, Ed. *Theoretical Aspects of Chemical Reactivity*. Amsterdam: Elsevier, 2007. 119–137
- 32 Domingo LR, Pérez P, Sáez JA. Understanding the local reactivity in polar organic reactions through electrophilic and nucleophilic Parr functions. *RSC Adv*, 2013, 3: 1486–1494

- 33 Lu T, Chen FW. Meaning and functional form of the electron localization function. *Acta Phys Chim Sin*, 2011, 27: 2786–2792
- 34 Bader RFW, Chang C. Properties of atoms in molecules: electrophilic aromatic substitution. *J Phys Chem*, 1989, 93: 2946–2956
- 35 Murray JS, Peralta-Inga Z, Politzer P, Ekanayaka K, LeBreton P. Computational characterization of nucleotide bases: molecular surface electrostatic potentials and local ionization energies, and local polarization energies. *Int J Quantum Chem*, 2001, 83: 245–254
- 36 Ehresmann B, Martin B, Horn AHC, Clark T. Local molecular properties and their use in predicting reactivity. *J Mol Model*, 2003, 9: 342–347
- 37 Parthasarathi R, Padmanabhan J, Elango M, Subramanian V, Chattaraj PK. Intermolecular reactivity through the generalized philicity concept. *Chem Phys Lett*, 2004, 394: 225–230
- 38 Morell C, Grand A, Toro-Labbé A. New dual descriptor for chemical reactivity. *J Phys Chem A*, 2005, 109: 205–212
- 39 Oláh J, van Alsenoy C, Sannigrahi AB. Condensed Fukui functions derived from stockholder charges: assessment of their performance as local reactivity descriptors. *J Phys Chem A*, 2002, 106: 3885–3890
- 40 Politzer P, Murray JS. Molecular electrostatic potentials and chemical reactivity. In: Lipkowitz KB, Boyd DB. *Reviews in Computational Chemistry. Volume 2*. New York: John Wiley & Sons, 1991. 273–312
- 41 Wang JT, Hu Q, Zhang B, Wang Y. *Organic Chemistry*. 2nd Ed. Tianjin: NanKai University Press, 1993
- 42 Morrison RT, Boyd RN. *Organic Chemistry*. 6th Ed. New Jersey: Prentice Hall, Inc., 1992
- 43 Smith MB, March J. *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*. New York: John Wiley & Sons, 2007
- 44 Chupakhin ON, Charushin VN, van der Plas HC. *Nucleophilic Aromatic Substitution of Hydrogen*. London: Academic Press, 2012
- 45 Terrier, François. *Modern Nucleophilic Aromatic Substitution*. New York: John Wiley & Sons, 2013
- 46 Deuri S, Phukan P. A DFT study on nucleophilicity and site selectivity of nitrogen nucleophiles. *Comput Theor Chem*, 2012, 980: 49–55
- 47 Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Menucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Namayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03. Version B.02. Wallingford, CT: Gaussian, Inc., 2004
- 48 Becke AD. A new mixing of Hartree-Fock and local density-functional theories. *J Chem Phys*, 1993, 98: 1372–1377
- 49 Lu T, Chen FW. Multiwfn: a multifunctional wavefunction analyzer. *J Comput Chem*, 2012, 33: 580–592
- 50 Multiwfn website: <http://Multiwfn.codeplex.com> (accessed on 2014-10-10)

Comparative study on the methods for predicting the reactive site of nucleophilic reaction

Jingsi Cao¹, Qing Ren¹, Feiwu Chen^{1*}, Tian Lu²

¹ Department of Chemistry and Chemical Engineering, School of Chemical and Biological Engineering, University of Science and Technology Beijing, Beijing 100083, China

² Beijing Kein Research Center for Natural Sciences, Beijing 100022, China

*Corresponding author (email: chenfeiwu@ustb.edu.cn)

Abstract: Predicting the reactivity of nucleophilic reaction at different sites has important theoretical and practical significance. There are many prediction methods so far that solely based on the electronic structure of reactants have been proposed. In this paper, detailed comparative analyses on the reliability of 14 methods are carried out and three series of molecules, carbonyl compounds, aromatic hydrocarbons and pyridine derivatives are exploited as test systems. It is found out that the methods reflecting local electronic softness, such as condensed dual descriptor, have satisfactory prediction ability; while the ones reflecting electrostatic effect, such as atomic charge analysis and electrostatic potential analysis, have evidently worse overall performance. For all systems of interest, condensed dual descriptor and Hirshfeld charge display the most robust predictive capacity.

Keywords: nucleophilic reaction, molecular surface, Fukui function, dual descriptor, electrostatic potential, atomic charge, orbital composition, relative electrophilicity