

[Review]

doi: 10.3866/PKU.WHXB2012281

www.whxb.pku.edu.cn

原子电荷计算方法的对比

卢 天 陈飞武*

(北京科技大学化学与生物工程学院, 北京 100083)

摘要: 原子电荷是对化学体系中电荷分布最简单、最直观的描述形式之一, 在理论和实际应用中都有重要意义. 本文介绍了 12 种重要的原子电荷计算方法的原理和特点, 通过大量实例从不同角度比较了它们的优缺点. 这些方法包括 Mulliken、分子环境中的原子轨道(AOIM)、Hirshfeld、原子偶极矩校正的 Hirshfeld 布居(ADCH)、自然布居分析(NPA)、Merz-Kollmann (MK)、分子中的原子(AIM)、Merck 分子力场 94 (MMFF94)、AM1-BCC、Gasteiger、电荷模型 2(CM2)以及电荷均衡(QEq)方法. 最后本文对如何在实际应用中选择合适的计算方法给出了建议.

关键词: 原子电荷; 计算化学; 布居分析; 电负性; 静电势

中图分类号: O641

Comparison of Computational Methods for Atomic Charges

LU Tian CHEN Fei-Wu*

(School of Chemical and Biological Engineering, University of Science and Technology Beijing, Beijing 100083, P. R. China)

Abstract: Atomic charge is one of the simplest and the most intuitive description of charge distribution in chemical systems. It has great significance in theory and in practical applications. In this article we introduce the basic principles and special characteristics of twelve important computational methods for the determination of atomic charges and compare their pros and cons from various aspects by considering a large number of instances. These methods include Mulliken, atomic orbitals in molecules (AOIM), Hirshfeld, atomic dipole moment corrected Hirshfeld population (ADCH), natural population analysis (NPA), Merz-Kollmann (MK), atom in molecules (AIM), Merck molecular force field 94 (MMFF94), AM1-BCC, Gasteiger, charge model 2 (CM2), and charge equilibration (QEq). Finally some general suggestions on how to choose a proper method for practical applications are given.

Key Words: Atomic charge; Computational chemistry; Population analysis; Electronegativity; Electrostatic potential

1 引言

原子电荷, 即位于原子中心的点电荷, 是对化学体系中电荷分布最简单、最直观的描述方式之一. 它有很多重要意义, 比如帮助化学工作者研究原子在各种化学环境中的状态、¹ 考察分子性质、^{2,3} 预测反应位点.⁴ 另外原子电荷模型在计算化

学中也有很多实用价值, 如作为分子描述符用于药物虚拟筛选、⁵ 在分子对接和分子动力学/蒙特卡罗模拟中描述静电作用、⁶ 在量子力学(QM)与分子力学(MM)结合计算中表现 QM 区域原子对 MM 区域原子的静电作用势.⁷ 在其它一些理论方法中也需要借助原子电荷, 如 Pipek-Mezey 轨道定域化方法、⁸

Received: September 13, 2011; Revised: October 25, 2011; Published on Web: October 31, 2011.

*Corresponding author. Email: chenfeiwu@ustb.edu.cn; Tel: +86-10-62332689.

The project was supported by the National Natural Science Foundation of China (20773011).

国家自然科学基金(20773011)资助项目

电荷自洽的紧束缚密度泛函方法(SCC-DFTB)、⁹ Truhlar 课题组开发的一系列溶剂模型(SM)¹⁰⁻¹³等。

原子电荷并不是可观测量,也没有客观、唯一的定义,因此获得原子电荷的方法多种多样.通过一些实验数据可以间接地考察原子电荷,¹⁴如通过分子的多极矩、红外光谱强度和频率、配体场分裂能、核磁共振(NMR)位移等,但是数据的获取相对不便,与原子电荷的对应关系存在较大经验性,也不能分析不稳定的体系和状态.计算化学的发展使原子电荷方便、快速、可靠的获得成为了可能.¹⁵⁻¹⁸自从1955年 Mulliken 电荷¹⁹⁻²¹被提出以来,迄今已有不下60种原子电荷计算方法被相继提出,近年仍有许多研究者在改进计算方法.

虽然在许多计算方法的原文和部分文献²²⁻²⁵中都对不同类型原子电荷进行了比较和讨论,但是涉及的方法并不全面,测试分子和测试内容缺乏统一性.因此,很有必要对较为重要的、文献中常涉及的原子电荷计算方法进行综合的比较,使它们的特点、优缺点能够清晰地展现,这将有助于在实际问题研究中选择适合的计算方法.我们选取 Mulliken、分子环境中的原子轨道(AOIM)、²⁶ Hirshfeld、²⁷ 原子偶极矩校正的 Hirshfeld 布居(ADCH)、²⁸ 自然布居分析(NPA)、²⁹ Merz-Kollmann (MK)、³⁰ 分子中的原子(AIM)、³¹ Merck 分子力场 94 (MMFF94)、^{32,33} AM1-BCC、^{34,35} Gasteiger、³⁶ 电荷模型 2 (CM2)³⁷ 和电荷均衡(QEq)³⁸共12种方法.在文中将首先介绍这些方法的基本原理和特点,之后通过实际体系测试和比较它们的各方面性能,最后讨论计算量的大小,并给出方法选择上的建议.

2 各种原子电荷计算方法的原理和特点

2.1 Mulliken 方法

Mulliken 方法¹⁹⁻²¹是最古老的原子电荷计算方法.它的算法简单,计算量可忽略不计,几乎出现在所有量子化学软件中.首先考虑分子轨道波函数归一化条件(假设为实函数,后同)

$$\int \phi_i(\mathbf{r})^2 d\mathbf{r} = 1 \quad (1)$$

\mathbf{r} 代表空间坐标.将分子轨道 ϕ_i 按以原子为中心的基函数 χ_m 展开

$$\phi_i = \sum_m C_{m,i} \chi_m \quad (2)$$

这里 \mathbf{C} 代表系数矩阵.将式(2)代入式(1),积分后得

$$\sum_m C_{m,i}^2 + 2 \sum_m \sum_{m>n} C_{m,i} C_{n,i} S_{m,n} = 1 \quad (3)$$

其中 $S_{m,n} = \int \chi_m(\mathbf{r}) \chi_n(\mathbf{r}) d\mathbf{r}$. 上式中第一项可视为各个基函数对轨道独立贡献之和,称定域项;第二项是交叉项,体现了由于每一对基函数之间的耦合对轨道产生的联合贡献. Mulliken 将分子轨道 i 中基函数 m 所占成分定义为

$$\Theta_{m,i} = C_{m,i}^2 + \sum_{n \neq m} C_{m,i} C_{n,i} S_{m,n} \quad (4)$$

也就是说,定域项被完全划归到相应基函数,而交叉项被平分给对应的两个基函数.将所有轨道中属于相同原子的基函数的占据数加和就得到了原子占据数,并直接得到原子电荷

$$q_A = Z_A - \sum_i \eta_i \sum_{m \in A} \Theta_{m,i} \quad (5)$$

其中 Z 是原子核电荷数, η 是轨道占据数.

Mulliken 电荷存在一些严重问题: (1) 交叉项平分的做法物理意义不严格,没有考虑到原子间的差异性; (2) 基组依赖性非常大.尤其是使用大基组、含弥散函数基组的情况下问题更为明显.这主要是因为弥散函数延展范围大,容易破坏基组的平衡性,后有一些研究者提出了不同方法试图改进交叉项划分的不合理性,³⁹⁻⁴³ 其中也包括使用相对较多的 Löwdin 方法,³⁹ 但是严重的基组依赖性却并没消除,也并未普及开来.

2.2 NPA 方法

自然布居分析(NPA)²⁹的关键是将带有随意性的原始基组(一般为扩展基)描述的波函数转化到物理意义清晰的正交极小基下描述,使基函数与原子轨道有明确对应关系,很大程度避免了基组不平衡性问题对结果的影响,也避免了划分交叉项的困难.

NPA 的原理如下.首先根据原始基函数与原子的对应关系将单粒子密度矩阵中对应于各个原子的对角块依次提取并对角化,所得本征向量就是初自然原子轨道(PNAO),其本征值就是 PNAO 的占据数.根据占据数可将 PNAO 分为两类: (1) 极小集轨道.它们拥有较高占据数,是对原子上的电子密度最主要、最紧凑的描述,与基态原子的内层和价层原子轨道一一对应. (2) 里德堡集轨道.这是指 PNAO 中除极小集轨道之外的、电子占据数很少的轨道,它们对原子的电子密度的描述不起主要作用.随后对这些 PNAO 进行占据数权重的对称正交

化(OWSO), 使它们不仅在原子内正交也在原子间正交. 相对于其它正交化方法, OWSO方法可以使有意义的极小集轨道变形尽量小, 而允许意义不大的里德堡集轨道自由变化, 物理意义更明确. PNAO在正交化后被称为自然原子轨道(NAO). 以NAO为基重新构建体系单粒子密度矩阵, 将对应于相应原子的矩阵对角元相加就得到了原子占据数.

值得一提的是, 由于NPA是自然键轨道(NBO)方法⁴⁴框架内的一个组成部分, 而且NPA电荷通常由Weinhold等开发的NBO程序⁴⁵计算, 所以NPA电荷经常在文献中被称为NBO电荷.

2.3 AOIM方法

和NPA的出发点相近, 黎乐民等^{26,46}提出的分子环境中的原子轨道(AOIM)方法也是将原始扩展基下的波函数转化为极小基描述后再做布居分析. AOIM有两种不同的具体实现方法:

第一种方法⁴⁶认为分子环境会使得原子轨道相对于原子在自由状态时收缩或扩张, 分子环境的影响通过平均化的球对称有效外势来表达

$$V_A^*(r) = (1/4\pi) \int V(r, \theta, \varphi) \sin\theta d\theta d\varphi \quad (6)$$

其中 r, θ 和 φ 是球坐标, 球坐标系以原子A的原子核为中心, V 是分子势场. 在这样的有效势下解原子的径向薛定谔方程, 所得波函数结合角度波函数就是分子环境下的原子轨道. 实际求解时使用变分法, 将计算分子时用的扩展基的径向函数作为变分函数, 就得到了原先扩展基与新的极小基之间的变换关系.

第二种方法²⁶利用Sanchez-Portal投影方法^{47,48}将扩展基波函数投影到Slater极小基轨道(STO). 设Löwdin正交化后的扩展基为 φ^{ortho} , 系数矩阵为 C' ; 所期望的极小基在Löwdin正交化后为 χ^{ortho} , 系数矩阵为 $C'^{(\text{mini})}$, 则变换关系为 $C'^{(\text{mini})} = \sum_n \langle \chi_n^{\text{ortho}} | \varphi_n^{\text{ortho}} \rangle C'_{n,i}$. 基函数的约化会导致基组完备性下降, 从而使波函数信息丢失, 衡量丢失量的参数为 $\Delta = \sum_i^{\text{occ}} [1 - \sum_m (C'^{(\text{mini})}_{m,i})^2]$. 它是STO的指数与方向的函数. 为了最大程度避免因约化带来的基组不完备性增加, 通过有限内存大尺度限制性优化(L-BFGS-B)方法最小化 Δ 就得到在分子环境下膨胀、收缩和旋转后的原子轨道. 通过上述过程得到极小基下的系数矩阵后, 就可以照常按照Mulliken方法计算原子电荷. 下文中涉及的AOIM方法都特指第二种AOIM方法.

2.4 AIM方法

Bader的分子中的原子(AIM)方法^{31,49}是典型的基于实空间划分的计算原子电荷的方法. AIM方法将电子密度零通量面定义为原子间的分界面, 划分出的每个原子独立的空间被称为原子盆. 在分界面上没有电子密度梯度线穿过, 即满足 $\nabla\rho(\mathbf{r}') \cdot \mathbf{n}(\mathbf{r}') = 0$, 这里 \mathbf{r}' 为原子界面上的任意点, \mathbf{n} 为界面上的单位法矢量, ρ 是电子密度函数. 这样的空间划分从量子力学角度来看理论意义明确, 在每个原子盆内维里定理得到满足. 对原子盆内电子密度积分并与核电荷求差值即得到原子电荷

$$q_A = Z_A - \int_{\Omega_A} \rho(\mathbf{r}) d\mathbf{r} \quad (7)$$

这里 Ω_A 代表A原子盆. 注意以AIM的划分, 在特殊条件下可能会出现不含原子核的原子盆, 被称为膺原子. 例如锂金属的两个锂原子间就存在膺原子, 这是由金属键所导致的. 另外还可能出现一个盆内包含不止一个原子核的情况, 如 KrH^+ 体系中Kr由于电子分布范围过广而淹没了氢. 这些情况都无法使用AIM方法来计算原子电荷.

2.5 Hirshfeld方法

Hirshfeld电荷写为

$$q_A = - \int w_A(\mathbf{r}) \Delta\rho(\mathbf{r}) d\mathbf{r} \quad (8)$$

其中

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \sum_A \rho_A^0(\mathbf{r})$$

$$w_A(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_A \rho_A^0(\mathbf{r})} \quad (9)$$

上式中 $\sum_A \rho_A^0(\mathbf{r})$ 代表所有原子在自由状态下的电子密度的加和, 也称promolecule密度. $\Delta\rho$ 被称为变形密度, 表现了原子形成分子后电子密度在弛豫过程中的变化. $w_A(\mathbf{r})$ 是A原子权重函数, 它定义了整个实空间中属于A原子的区域. 与AIM的离散式原子空间划分相比, Hirshfeld对原子空间的划分是模糊的, 原子空间彼此间存在交叠. Hirshfeld的划分方式有比较清楚的物理意义, Nalewajski和Parr⁵⁰从信息论角度证明了这种划分是允许电子密度在原子间转移的情况下, 使原子在分子中和在自由状态下信息的距离(Kullback-Leibler相对熵)最小化的划分方式.

2.6 ADCH方法

Hirshfeld电荷数值普遍偏小,⁵¹ 而且偶极矩、静

电势重现性比较差,²⁵ 我们认为这主要是忽略了原子偶极矩所引起的. 原子偶极矩可定义为

$$\mu_A = -\int w_A(\mathbf{r})\Delta\rho(\mathbf{r})(\mathbf{r}-\mathbf{r}_A)d\mathbf{r} \quad (10)$$

其中 \mathbf{r}_A 代表 A 原子核坐标. 原子偶极矩是对原子空间内电子密度各向异性分布最重要的描述, 它对分子可观测性质有直接贡献. 然而式(8)的积分却相当于把原子空间内电子密度的分布球对称化了, 完全掩盖了原子偶极矩的效应. 因此提出了原子偶极矩校正的 Hirshfeld 布居(ADCH).²⁸ 在这个方法中首先计算各个原子的 Hirshfeld 电荷及原子偶极矩, 然后将每个原子偶极矩按下式展开成为周围原子的校正电荷

$$\mu_A = \sum_B \Delta q_{AB} \mathbf{r}_B \quad (11)$$

其中 Δq_{AB} 代表展开 A 原子偶极矩后对 B 原子产生的校正电荷. 把所有原子偶极矩展开成校正电荷后, 将校正电荷累加到原始 Hirshfeld 电荷上就得到了 ADCH 电荷. 为了保证分子净电荷在校正过程中不变, 展开任一原子 A 的偶极矩时需满足这个条件

$$\sum_B \Delta q_{AB} = 0 \quad (12)$$

同时, 为了让原子偶极矩只展开到离当前原子较近的原子, 在实际计算时令式(11)和式(12)以拉格朗日乘子方式作为限制条件, 使下式最小化来获得 A 原子对其它原子产生的校正电荷

$$F_A(\Delta q_{A1}, \Delta q_{A2}, \dots, \Delta q_{AN}) = \sum_B \frac{\Delta q_{AB}^2}{v_{AB}(r_{AB})} \quad (13)$$

其中 r_{AB} 是 A、B 原子间距离, N 是总原子数, v 是依赖于原子间距离的函数. 它呈倒 S 形状, 当 r_{AB} 由 0 变化到 A、B 原子范德华半径之和的过程中 v_{AB} 由 1 变化为 0. 因此, 在原子偶极矩展开时, 离当前原子越近的原子越容易获得校正电荷, 与当前原子无相互作用的原子(距离超过范德华半径和)的电荷不会受到影响.

2.7 MK 方法

在化学体系中, 静电势(ESP)由原子核电荷和电子密度两部分贡献构成

$$V(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{r}-\mathbf{r}_A|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (14)$$

对于离原子核较近的区域, 总是由第一项主导, 静电势数值为正, 这是化学上不感兴趣的. 而体系范德华表面以外的区域静电势可正可负, 具有显著的化学意义, 对于研究受体-配体结合方式、晶体堆叠方式, 预测亲核/亲电反应位点, 通过比较场分析

(CoMFA)进行药物筛选等问题至关重要. 若在分子范德华表面外侧一定范围内选取一批格点, 通过最小二乘法令原子电荷在这些点产生的静电势

$$V_q(\mathbf{r}) = \sum_A \frac{q_A}{|\mathbf{r}-\mathbf{r}_A|} \quad (15)$$

对式(14)计算的静电势拟合, 就得到了拟合静电势电荷. 拟合静电势电荷计算方法很多, 主要差异在于格点位置的选取和拟合过程的细节. 最常见的方法有 Merz-Kollman (MK)方法,³⁰ 静电势获得的电荷 (CHELP)方法⁵² 和基于格点的 CHELP (CHELPG)方法.⁵³ CHELPG 电荷在三者中旋转不变性最好, 而通过多极矩重现性分析发现 MK 电荷质量比 CHELP 和 CHELPG 略好,⁵⁴ 因此本文选取 MK 电荷作为拟合静电势电荷的代表在后文中与其它类型电荷进行比较.

2.8 AM1-BCC 方法

传统拟合静电势方法存在一些弊端, 例如: (1) 电荷对构象依赖性较大; (2) 单一构象拟合的原子电荷不能体现原子的等价性, 例如一氯丙烷中甲基的三个氢是化学等价的, 但任何构象下拟合出来的这三个氢的电荷都不是全同的; (3) 内部被包埋的原子电荷拟合不准确, 这是由于这些原子距离拟合点过远所导致的. 这三个问题使得拟合静电势电荷用于分子动力学模拟等问题中可能导致错误的结果. 限制性静电势拟合(RESP)方法⁵⁵ 可以很大程度解决这些弊端, 它对内部原子在拟合时施加限制势以降低数值不稳定性, 对等价原子在拟合时强制相等, 并且将拟合分为两步或者多步以提高拟合质量. 然而计算 RESP 电荷过程相对复杂, 而且需要较为准确的量子化学方法计算静电势, 用于大批量或者较大分子体系较为耗时.

AM1-BCC^{34,35} 是一种简单的近似计算 RESP 电荷的方法, 其电荷是在 AM1 电荷(即 AM1 半经验方法⁵⁶ 获得的波函数产生的 Mulliken 电荷)基础上进行键电荷校正(BCC)得到的. 校正过程只依赖于原子类型和原子间连接关系. 校正参数拟合自 2755 个种类多样的有机分子的 HF/6-31G* 下的 RESP 电荷. 例如甲醇中的碳原子形成三个“C_{sp³}-single-H”键和一个“C_{sp³}-single-O_{sp³}”键, 根据查寻事先拟合的 BCC 参数就可得知校正电荷应为 $3 \times 0.0274 + 1 \times 0.0835 = 0.1657$ (本文电荷单位均为原子单位), 加到碳的 AM1 电荷上就是 AM1-BCC 电荷.

2.9 MMFF94 方法

Merck 分子力场 94 (MMFF94)^{32,33} 获得电荷的方式在形式上与 AM1-BCC 十分相似, 同样使用键电荷校正过程, 即考虑到每个原子形成的化学键的极性来对初始电荷校正. 与 AM1-BCC 不同之处是 MMFF94 的初始电荷由原子类型直接确定, 且参数来自于拟合 HF/6-31G* 下计算的分子偶极矩及相互作用能.

2.10 CM2 方法

电荷模型 (CM) 系列方法目前包括 CM1、⁵⁷ CM2、^{37,58} CM3⁵⁹⁻⁶¹ 和 CM4,^{62,63} 它们都是对低级别方法下的 Mulliken 或 Löwdin 电荷³⁹ 进行简单校正的方法, 目的是使原子电荷计算出的偶极矩尽可能精确地重现气相实验值或者高精度量子化学计算值. 这一类电荷也被用于不同版本的 SM 系列溶剂模型中.

CM2 校正的公式与 AM1-BCC 和 MMFF94 的相比略为复杂, CM2 电荷可写为

$$q_A^{\text{CM2}} = q_A^{\text{Löwdin}} + \sum_{B \neq A} M_{A,B} (D_{A,B} + C_{A,B} M_{A,B}) \quad (16)$$

其中 M 是原子间 Mayer 键级.⁶⁴ C 与 D 是对实验气相偶极矩拟合的参数, 它们取决于化学键的种类, 而键的种类仅由它相连的两个原子的元素决定. CM2 方法对使用 AM1、PM3、HF/MIDI!、HF/6-31G* 等级别计算 Löwdin 电荷的情况分别拟合了校正参数.

2.11 Gasteiger 方法

Gasteiger 电荷³⁶ 也被称为轨道电负性部分均衡 (PEOE) 电荷, 这种方法利用了电负性均衡概念,⁶⁵ 即电负性不同的原子成键时, 电负性较小的原子附近电子密度会流向电负性较大的原子. 在这个过程中原先电负性小的原子电负性会增大. 当所有原子间电负性相等时, 电子密度的分布就是平衡状态分布.

PEOE 将原子电负性与原子电荷关系表达为

$$\chi_A = a + bq_A + cq_A^2 \quad (17)$$

其中 a 、 b 和 c 为预设参数, 对于相同元素但价层轨道处于不同杂化状态的情况参数不同. 参数通过实验测定的相应元素的基态和电离态在相应杂化状态下的电离势与电子亲和势来计算. PEOE 的迭代过程如同电荷不断在原子间转移的过程, 当迭代收敛时电荷分布也就平衡了. 在每一步中按照下式计算每一对相连原子间电荷转移量

$$\Delta q^{(n)} = f^n (\chi_B^{(n-1)} - \chi_A^{(n-1)}) / \chi_A^{q=1} \quad (18)$$

这里假设 $\chi_B^{(n-1)} > \chi_A^{(n-1)}$, n 是当前步数, f 是阻尼项, 一

般设为 0.5. f 意味着随着电子不断转移, 会逐渐形成电场使电荷继续转移变得愈发困难. 将转移电荷累加到第 $n-1$ 步的原子电荷上就得到了第 n 步的原子电荷, 根据式 (18) 可计算第 n 步的原子电负性, 由此再计算 $n+1$ 步的电荷转移量, 迭代反复进行. 由于阻尼项的存在, 一般四五步即可收敛. 对于中性分子在计算开始时通常将每个原子的电荷都设为 0. 注意由于 PEOE 引入了阻尼项, 所以实际上 PEOE 电荷并没有达到完全的电负性均衡, 而只达到了部分均衡.

2.12 QEq 方法

电荷均衡方法 (QEq)³⁸ 得到的是在电负性完全均衡下的原子电荷. 电负性表达式为

$$\chi_A = \chi_A^0 + J_{AA}^0 q_A + \sum_{B \neq A} J_{AB} q_B \quad (19)$$

电负性均衡条件要求 $\chi_1 = \chi_2 = \dots = \chi_N$, 它提供了 $N-1$ 个条件, 再将所有原子电荷总和等于分子净电荷作为剩余条件, 通过解线性方程即可到得全部原子电荷. 式 (19) 中 χ_A^0 是广义化的 Mulliken 电负性, J_{AA}^0 为自库仑积分, 是原子电离势与电子亲和势的差值. J_{AB} 描述了原子间静电作用, 若两原子距离较远, 则 $J_{AB} = 1/r_{AB}$; 若原子间距离较近, 电子云静电屏蔽作用不能忽视, 此时 J_{AB} 被定义为两原子间库仑积分. 每个原子用单个 Slater 函数描述电子密度分布, 其中指数通过拟合碱金属卤化物实验偶极矩数据确定, 对所有原子都相同. 普适型力场 UFF⁶⁶ 使用的正是 QEq 电荷模型.

值得一提的是 QEq 方法可以视为另一种使用广泛的方法电负性均衡方法 (EEM)⁶⁷ 的变体. 然而 EEM 的经验性较大, 不同研究者以不同的目标数据拟合了参数, 结果有很大差异, 因此不纳入本文的比较.

除了上面涉及的方法外, 广义原子极化张量 (GAPT) 电荷^{68,69} 也常出现在文献当中. 然而 GAPT 电荷计算过于耗时 (和振动分析耗时相近), 实际应用意义很有限, 因此并不在本文进行介绍和比较. 读者可参看文献²⁵ 对它的测试和讨论.

3 计算细节和测试体系

下文将对前面介绍的 12 种方法得到的原子电荷与化学经验相符程度、对偶极矩和静电势的重现性以及基组依赖性进行检验, 最后讨论计算耗时以及计算方法的选择. 测试体系以研究得最为普遍的有机小分子为主, 晶体以及包含重元素的体系不在

本文讨论范围内. 测试分子皆处于稳定构型, 过渡态、激发态、受外场作用等状态也不属于本文涉及范围. 实际上, Gasteiger、CM2 等包含经验参数的方法也不适用于这些特殊问题. 本文所测试的方法分为两组, 第一组不依赖于任何参数, 包括 Mulliken、AOIM、Hirshfeld、ADCH、NPA、MK 和 AIM, 第二组依赖于拟合的参数, 包括 MMFF94、AM1-BCC、CM2、Gasteiger 和 QEq.

除了理论方法和基组依赖性部分的讨论外, 下文的中性分子、阳离子的结构优化和波函数计算均在 HF/6-31G*⁷⁰⁻⁷² 级别下进行, 阴离子的结构优化和波函数计算在 HF/6-31+G*⁷³ 级别下进行. 波函数的计算和结构优化, 以及 Mulliken、NPA、MK、CM2、QEq 电荷的计算都通过 Gaussian 03 程序⁷⁴ 完成. 计算 CM2 电荷时的初始电荷由 HF/6-31G* 级别获得.

为了使拟合质量较好, 计算 MK 电荷时通过 IOp(6/42=6) 选项将每单位面积的拟合点由默认的 1 个增加到 6 个. Hirshfeld 和 ADCH 电荷通过我们开发的 Multiwfn 2.1.2 程序⁷⁵ 计算. AIM 电荷通过 AIMALL 10.05.04 程序⁷⁶ 计算. AOIM 电荷由 AOIM 1.1 程序⁷⁷ 计算(目前没有其它公开的程序可以计算 AOIM 电荷). MMFF94 电荷通过 Avogadro 1.0.3 程序⁷⁸ 获得. AM1-BCC 和 Gasteiger 电荷由 AmberTools 1.5 程序⁷⁹ 计算. 由于计算 AM1-BCC 电荷前 AmberTools 会自动在 AM1 下进行结构优化, 我们通过修改代码将此步骤取消.

4 普通小分子的原子电荷

我们首先检验不同方法计算的种类各异的小分子电荷, 表 1 列出了第一组方法计算的结果. 容易

表 1 第一组方法计算的一些小分子的原子电荷

Table 1 Atomic charges of some small molecules calculated by the first group of methods

Molecule	Atom ^a	Mulliken	AOIM	Hirshfeld	ADCH	NPA	MK	AIM
CH ₃ COCH ₃	C(Me)	-0.579	-0.291	-0.055	-0.210	-0.734	-0.514	0.026
	H	0.182	0.066	0.031	0.086	0.231	0.129	0.006
	H	0.208	0.097	0.035	0.091	0.244	0.139	0.030
	C	0.524	0.697	0.216	0.287	0.685	0.815	1.215
	O	-0.513	-0.571	-0.301	-0.392	-0.629	-0.582	-1.348
CH ₃ NO ₂	C(Me)	-0.347	-0.187	0.029	-0.139	-0.464	-0.371	0.417
	H	0.230	0.113	0.055	0.122	0.243	0.156	0.074
	H	0.234	0.120	0.061	0.127	0.247	0.165	0.070
	N	0.560	0.951	0.306	0.459	0.650	0.864	0.463
	O	-0.454	-0.555	-0.253	-0.345	-0.460	-0.485	-0.548
NH ₂ CHO	N	-0.884	-0.599	-0.163	-0.324	-0.927	-0.985	-1.457
	H(<i>trans</i>)	0.389	0.233	0.129	0.223	0.412	0.410	0.448
	H(<i>cis</i>)	0.398	0.249	0.133	0.227	0.419	0.442	0.462
	C	0.508	0.863	0.215	0.216	0.676	0.728	1.920
	H(-C)	0.143	-0.068	0.040	0.109	0.126	0.000	0.017
	O	-0.555	-0.678	-0.355	-0.451	-0.707	-0.595	-1.391
	N	-0.419	-0.307	-0.106	-0.237	-0.379	-0.216	-0.935
(CH ₃) ₂ NNH ₂	C(Me)	-0.281	0.106	-0.005	-0.090	-0.413	-0.273	0.527
	H	0.167	0.026	0.025	0.080	0.210	0.131	-0.011
	H	0.160	0.015	0.021	0.071	0.204	0.129	-0.021
	H	0.152	0.009	0.010	0.065	0.188	0.137	-0.019
	N(center)	-0.419	-0.307	-0.106	-0.237	-0.379	-0.216	-0.935
	N(amino)	-0.659	-0.449	-0.186	-0.618	-0.730	-0.775	-0.747
	H(amino)	0.341	0.222	0.095	0.302	0.367	0.371	0.364
C ₆ H ₅ SO ₃ H	S	1.615	2.579	0.598	0.775	2.702	1.072	3.553
	O=	-0.650	-1.017	-0.368	-0.439	-1.052	-0.556	-1.516
	-O-	-0.766	-0.792	-0.220	-0.474	-0.973	-0.650	-1.341
	H(O)	0.497	0.403	0.200	0.424	0.528	0.513	0.643
	C(-SO ₃ H)	-0.325	-0.382	-0.018	-0.106	-0.396	0.035	-0.168
	C(<i>ortho</i>)	-0.159	0.067	-0.001	0.018	-0.169	-0.111	0.028
	C(<i>meta</i>)	-0.206	-0.077	-0.025	-0.161	-0.238	-0.169	0.013
C ₂ H ₂	C(<i>para</i>)	-0.175	0.009	-0.003	-0.091	-0.179	-0.077	-0.008
	C	-0.276	-0.086	-0.094	-0.251	-0.239	-0.293	-0.176 ^b
	H	0.276	0.086	0.094	0.251	0.239	0.293	0.177
H ₂ O	O	-0.869	-0.684	-0.319	-0.798	-0.952	-0.816	-1.181
	H	0.434	0.342	0.160	0.399	0.476	0.408	0.591

to be continued

continued Table 1

Molecule	Atom ^a	Mulliken	AOIM	Hirshfeld	ADCH	NPA	MK	AIM
HCN	C	0.066	0.239	0.080	0.033	0.120	0.133	1.238
	H	0.313	0.047	0.132	0.288	0.232	0.234	0.242
	N	-0.379	-0.286	-0.212	-0.321	-0.352	-0.366	-1.481
CLi ₄	C	-0.508	-3.556	-0.789	-1.833	-2.021	-2.766	-3.445
	Li	0.127	0.889	0.197	0.458	0.505	0.692	0.866
BeO	Be	0.403	1.877	0.624	1.090	1.243	1.089	1.732
	O	-0.403	-1.877	-0.624	-1.090	-1.243	-1.089	-1.722
CH ₂ PH	C	-0.544	-0.354	-0.081	-0.205	-0.820	-0.280	-1.055
	H	0.213	0.093	0.043	0.135	0.231	0.146	0.063
	H	0.209	0.093	0.042	0.133	0.233	0.177	0.058
	P	0.188	0.252	0.028	-0.157	0.441	-0.151	1.515
ClF ₃	H	-0.066	-0.084	-0.032	0.095	-0.085	0.108	-0.582
	Cl	1.338	1.413	0.563	0.663	1.472	0.708	1.669
	F(axial)	-0.515	-0.566	-0.254	-0.295	-0.559	-0.321	-0.605
	F(equat.)	-0.308	-0.281	-0.055	-0.072	-0.354	-0.066	-0.459
CH ₃ NH ₃	C	-0.381	-0.067	0.076	-0.159	-0.414	-0.041	0.359
	H(Me)	0.271	0.143	0.076	0.146	0.259	0.120	0.102
	N	-0.847	-0.301	0.089	-0.140	-0.793	-0.277	-1.203
	H(amino)	0.472	0.313	0.202	0.287	0.477	0.320	0.513
CH ₃ COO ⁻	C(Me)	-0.685	-0.341	-0.089	-0.187	-0.739	-0.377	0.013
	H	0.163	0.032	0.000	0.034	0.203	0.052	-0.046
	H	0.170	0.019	-0.008	0.032	0.200	0.045	-0.044
	C	0.717	1.208	0.167	0.579	0.942	1.109	2.121
	O	-0.768	-0.968	-0.530	-0.745	-0.903	-0.937	-1.499
Sum of abs. ^c		25.194	29.290	9.920	18.636	33.327	26.009	44.190

^aOnly symmetry-unique atoms are shown; ^bThe charge of pseudo-atom is equally divided between the two carbon atoms;

^cThe sum of absolute value of all atomic charges; the unit of charge in this paper is a.u.

看出, 不同方法得到的原子电荷的数值大小在整体上有所不同. Hirshfeld 电荷整体偏小, 许多原子电荷的大小都不足 0.1. ADCH 是对 Hirshfeld 方法的校正, 校正后原子电荷普遍增大, 电荷绝对值加和比校正前增加了一倍. 对于碳、氢、氮、氧构成的中性有机分子, Mulliken、AOIM、MK 和 NPA 电荷在数量级上比较接近, 对于个别分子它们也存在着明显差异, 例如 ClF₃ 中氯的 MK 电荷仅为 Mulliken、AOIM 和 NPA 方法计算的约一半. 很多原子的 AIM 电荷明显大于其它方法所得电荷, 例如丙酮的氧的电荷达到 -1.35, 氰化氢的氮的电荷高达 -1.48. 从电荷绝对值加和来看, AIM 计算的电荷值(44.19)达到了 Hirshfeld 计算的电荷值(9.92)的将近 5 倍.

Mulliken 电荷对于 CLi₄ 和 BeO 这样高离子性的体系明显偏小, 甚至小于 Hirshfeld 方法. 这和交叉项的平分处理有直接关系, 它导致了本应属于阴离子的电子被过多地划归到了阳离子上. 而对这样的体系 AOIM 电荷则比较大, 这是由于在分子环境中优化 STO 的指数后, 阴离子的 STO 较为弥散, 一定程度侵入到阳离子空间内, 而阳离子的 STO 较为紧缩, 因此在新的基函数下做 Mulliken 分析时阴离子能够获得更多的电子.

电荷的大小与其合理性并没有绝对关系, 考察原子电荷的合理性关键之一是看它能否与一般化学经验、化学观念相吻合. 磺酸基是间位定位基, 因此苯磺酸的间位碳的原子电荷应当比邻对位更负, 第一组方法中除了 AIM 以外都正确地给出了预期的结果, 而 AIM 方法却认为间位碳的电荷大于对位的. 可见 AIM 电荷对于研究反应位点并不适合. 对于偏二甲胍, 由于碳的电负性大于氢, 因此氨基氮的电荷应当比中间的氮原子的电荷更负, 而只有 AIM 方法给出的结论是相反的. 磷的 Pauling 电负性比碳略小, 因此在 CH₂PH 中磷的 Mulliken、AOIM、NPA、Hirshfeld 电荷为不很大的正值并不意外. 而 ADCH 和 MK 电荷更侧重于等效描述电子密度的实际分布, 由于磷的孤对电子产生的极化效应, 使得磷的 ADCH 和 MK 电荷都为不大的负值, 这也是合理的. 而 AIM 给出的磷的电荷高达 1.51, 明显缺乏合理性.

在 HF/6-31G* 波函数级别下计算乙炔的 AIM 电荷时出现了前文谈到的“膺原子”, 它位于两个碳原子之间. 由于乙炔的对称性, 我们将膺原子盆内电荷积分值平分给两个碳原子来获得碳的 AIM 电荷, 但是当体系缺乏这样的对称性时就不能这样计算

AIM 电荷了. 很多含有炔基的分子也都存在这样的膺原子, 这使得 AIM 方法的适用性受到一定限制. 值得一提的是, 将基组替换为与 6-31G* 级别较为相似的 SVP 基组⁸⁰后膺原子随即消失, 并变为碳-碳间的键临界点, 这显示出在特殊情况下 AIM 的适用性明显受基组影响.

表 2 列出了第二组方法计算的小分子的原子电荷, 由于缺少参数, 其中删去了表 1 中的 BeO 和 CLi. 很明显 Gasteiger 电荷相对其它电荷明显整体偏小, 而 MMFF94 和 AM1-BCC 电荷通常比 CM2 和 QEq 电荷稍大. 从所列分子上看, AM1-BCC 和 CM2 电荷都与经验观念基本一致. 由于 MMFF94 电荷计

表 2 第二组方法计算的一些小分子的原子电荷

Table 2 Atomic charges of some small molecules calculated by the second group of methods

Molecule	Atom ^a	MMFF94	AM1-BCC	Gasteiger	CM2	QEq
CH ₃ COCH ₃	C(Me)	0.061	-0.205	-0.006	-0.286	-0.332
	H	0.000	0.057	0.031	0.105	0.116
	H	0.000	0.064	0.031	0.114	0.177
	C	0.448	0.572	0.126	0.284	0.300
	O	-0.570	-0.513	-0.300	-0.360	-0.453
CH ₃ NO ₂	C(Me)	0.240	-0.103	0.081	-0.125	-0.105
	H	0.000	0.091	0.051	0.130	0.164
	H	0.000	0.089	0.051	0.133	0.158
	N	0.800	0.251	0.055	0.044	0.110
	O	-0.520	-0.209	-0.144	-0.156	-0.246
NH ₂ CHO	N	-0.800	-0.701	-0.332	-0.613	-0.640
	H(<i>trans</i>)	0.370	0.311	0.145	0.312	0.245
	H(<i>cis</i>)	0.370	0.305	0.145	0.314	0.305
	C	0.570	0.679	0.196	0.339	0.342
	H(-C)	0.060	-0.003	0.127	0.063	0.166
	O	-0.570	-0.590	-0.282	-0.415	-0.418
(CH ₃) ₂ NNH ₂	C(Me)	0.270	0.133	0.001	-0.102	-0.253
	H	0.000	0.044	0.041	0.087	0.130
	H	0.000	0.042	0.041	0.081	0.144
	H	0.000	0.027	0.041	0.071	0.158
	N(center)	-0.540	-0.497	-0.253	-0.291	-0.316
	N(amino)	-0.720	-0.664	-0.269	-0.529	-0.570
	H(amino)	0.360	0.335	0.138	0.273	0.264
C ₆ H ₅ SO ₃ H	S	1.493	1.497	-0.037	0.620	0.019
	O=	-0.650	-0.641	-0.463	-0.353	-0.254
	-O-	-0.684	-0.706	-0.051	-0.427	-0.524
	H(O)	0.500	0.475	0.356	0.376	0.415
	C(-SO ₃ H)	-0.009	-0.297	0.353	-0.090	0.311
	C(<i>ortho</i>)	-0.150	-0.022	0.064	-0.061	-0.060
	C(<i>meta</i>)	-0.150	-0.147	-0.050	-0.104	-0.149
	C(<i>para</i>)	-0.150	-0.068	-0.061	-0.061	-0.119
C ₂ H ₂	C	-0.177	-0.159	-0.124	-0.161	-0.143
	H	0.177	0.159	0.124	0.161	0.143
H ₂ O	O	-0.860	-0.789	-0.412	-0.666	-0.743
	H	0.430	0.394	0.206	0.333	0.371
HCN	C	0.380	0.199	0.046	0.149	0.117
	H	0.177	0.174	0.156	0.177	0.115
	N	-0.557	-0.373	-0.202	-0.326	-0.232
CH ₂ PH	C	-0.180	0.250	-0.096	-0.365	-0.296
	H	0.150	-0.022	0.054	0.124	0.144
	H	0.150	-0.009	0.054	0.124	0.165
	P	-0.216	-0.387	-0.080	0.052	-0.136
ClF ₃	H	0.096	0.168	0.067	0.064	0.124
	Cl	- ^b	- ^b	- ^b	1.246	0.387
	F(axial)	- ^b	- ^b	- ^b	-0.481	-0.136
	F(equat.)	- ^b	- ^b	- ^b	-0.283	-0.115

to be continued

continued Table 2

Molecule	Atom ^a	MMFF94	AM1-BCC	Gasteiger	CM2	QEq
CH ₃ NH ₃ ⁺	C	0.503	0.126	- ^c	-0.101	-0.204
	H(Me)	0.000	0.111	- ^c	0.149	0.289
	N	-0.853	-0.838	- ^c	-0.426	-0.738
	H(amino)	0.450	0.460	- ^c	0.360	0.359
CH ₃ COO ⁻	C(Me)	-0.106	-0.232	-0.025	-0.303	-0.308
	H	0.000	0.000	0.029	0.066	0.001
	H	0.000	0.004	0.029	0.063	0.017
	C	0.906	0.932	0.038	0.285	0.532
	O	-0.900	-0.852	-0.550	-0.587	-0.629
Sum of abs. ^d		16.516	15.437	6.614	11.324	12.180

^aOnly symmetry-unique atoms are shown; ^bparameters are unavailable; ^cthe charges calculated by various programs differ obviously, they are not shown; ^dthe sum of absolute value of all atomic charges except for ClF₃ and CH₃NH₃⁺

算方法过于简单, 没有考虑电子效应的影响, 所得苯磺酸的邻、间、对碳原子电荷相同, 均为-0.15, 是不合理的. Gasteiger 和 QEq 方法计算的苯磺酸的硫原子电荷都不合理, 数值接近 0. 由于硫原子与三个电负性明显比它更大的氧原子相连, 故硫原子电荷理应为较大的正值, 其它电荷计算方法的结果也验证了这一点. 另外 Gasteiger 电荷并没有正确预测出苯磺酸的间位碳具有比对位的碳更负的电荷, 对这种体系有必要改用明确考虑共轭效应的 Gasteiger 电荷改进方法, 即 PEPE 方法.^{81,82}

我们发现不同程序计算的 Gasteiger 电荷对于 CH₃NH₃⁺ 差异较大, 例如用 AmberTools 1.5、OpenBabel 2.3.0⁸³ 和 Python Molecule Viewer 1.5.6⁸⁴ 计算的氮的原子电荷分别为 0.229、0.376 和 -0.083, 而不同程序计算的中性分子电荷则相符较好. 这可能是由于对于带电体系上述三个程序所设定的初始原子电荷不同, 从而影响了迭代收敛后的电荷. 因此, 使用上述三个程序计算离子的电荷时, 对计算结果应

该慎重考虑.

5 原子电荷与电负性的相关性

对于由 σ 键构成的典型分子, 合理的原子电荷计算方法应当能够与电负性规则基本吻合. 对甲烷以不同电负性原子进行取代, 各种方法计算的甲基的电荷(甲基各原子电荷的加和)如图 1 和图 2 所示. Mulliken、AOIM、NPA、Hirshfeld、ADCH、Gasteiger、CM2 和 QEq 虽然在曲线整体斜率以及曲线形状上有别, 但是都正确显示出甲基电荷随着取代基原子电负性增大而逐渐增加. MMFF94 和 AIM 方法计算的甲烷中甲基电荷基本为 0, 然而从碳、氢的电负性上来看此时甲基电荷理应为负值. MK 方法对于氮、氧、氟取代甲烷体系给出的结果与电负性关系不相符, 随着取代原子电负性的增加甲基的 MK 电荷不增反降. 出现这种情况主要原因在于在这三种甲烷取代物中氮、氧、氟的电子极化程度依次减小, 如果用原子偶极矩, 即式(10)来表示极化程度, 则数值分

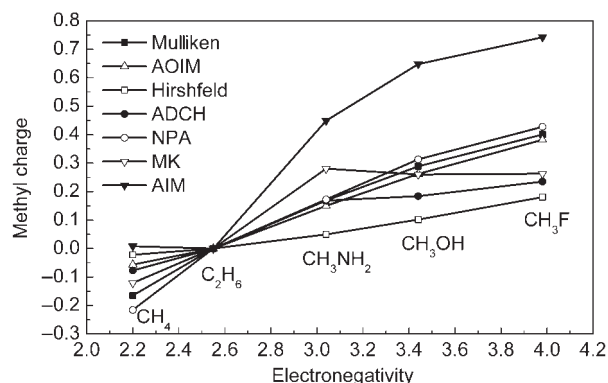


图 1 第一组方法计算的甲烷的甲基电荷与取代原子电负性的关系

Fig.1 Relationship between methyl charges calculated by the first group of methods and electronegativity of substituent atoms of methane

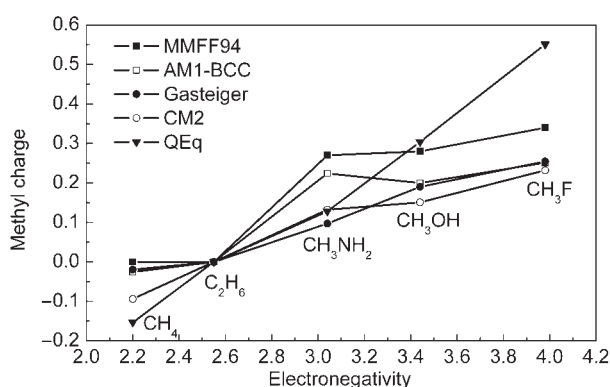


图 2 第二组方法计算的甲烷的甲基电荷与取代原子电负性的关系

Fig.2 Relationship between methyl charges calculated by the second group of methods and electronegativity of substituent atoms of methane

别为 0.296、0.237 和 0.143 a.u.。原子核附近电子极化对分子外侧静电势产生的影响会等价地体现在拟合静电势方法获得的电荷中, 并且由于在这三种体系中电子极化效应对 MK 电荷的影响超过了电负性差异的影响, 因此 MK 电荷表现出与电负性规则相反的变化趋势。由于 AM1-BCC 是对拟合静电势电荷的近似, 故它对于氮、氧取代甲烷体系给出的结果与电负性关系相违背也是预料之中的。由于 ADCH 对 Hirshfeld 电荷的校正过程中使电子密度极化效应得以等效体现, 与拟合静电势方法有一定共通之处, 这导致对于氮、氧、氟取代时 ADCH 的甲基电荷的变化曲线的斜率略小于 Hirshfeld 的斜率。然而 ADCH 并没有像 MK 电荷那样违背电负性规律, 因此在这一点上比 MK 具有更强的合理性。

对甲烷以不同数目的氟进行取代时每个氟原子的电荷应有所不同。1 至 4 取代时每个氟原子所连基团分别相当于 $-\text{CH}_3$ 、 $-\text{CH}_2\text{F}$ 、 $-\text{CHF}_2$ 和 $-\text{CF}_3$, 基团电负性依次增加, 因此合理的方法计算出的氟原子电荷应当依次减小。由图 3、图 4 可见, Mulliken、Hirshfeld、ADCH、CM2、Gasteiger、QEq 方法计算的结果都完全符合这个趋势, 而且氟原子电荷的减小是接近线性的。而其它几种方法则表现出不同程度的不合理性。AOIM 电荷与期望的趋势不符, 1 取代时氟原子的电荷值比 2、3、4 取代时的都要小, 这说明 AOIM 在理论上, 或者计算程序的数值算法上有待修正。从 1 取代变为 2 取代时, MK 电荷仅减小了

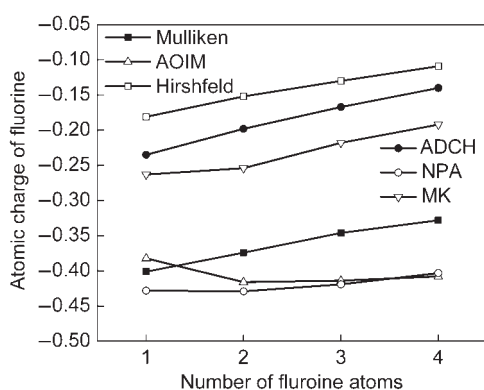


图3 第一组方法计算的不同取代数目的氟代甲烷的氟原子电荷

Fig.3 Atomic charge of fluorine atom calculated by the first group of methods of fluomethane with different numbers of substitutions

The AIM charge of fluorine atom is too large to be shown on the graph, the values are -0.742 , -0.744 , -0.744 , and -0.737 respectively for 1 to 4 substitutions.

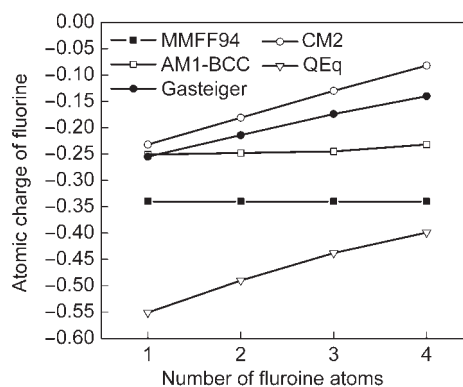


图4 第二组方法计算的不同取代数目的氟代甲烷的氟原子电荷

Fig.4 Atomic charge of fluorine atom calculated by the second group of methods for fluomethane with different numbers of substitutions

0.09, 而 NPA 电荷则基本没有变化。AIM 电荷在 1 至 4 取代过程中几乎都没有发生变化, 甚至 1 取代变为 2 取代过程中电荷甚至还增大了 0.002 (可能由于数值积分误差导致)。MMFF94 的氟原子电荷始终为 -0.340 , 这是因为 MMFF94 原子电荷只能考虑相邻原子对它的影响, 很明显, 这种过于简单的计算电荷的方式并不能准确描述原子在分子中的状态。AM1-BCC 也未能很好表现氟原子电荷应有的变化趋势, 虽然随取代数目增加氟原子电荷确实依次减小, 但是曲线却过于平坦。

6 偶极矩的重现性

偶极矩是小分子电荷分布的最简单也是最重要的描述, 分子偶极矩重现性的好坏是判断原子电荷合理性重要的客观标准。我们选取了 20 种具有结构特征不同、有代表性, 并且已知实验偶极矩的小分子, 对各种方法所得电荷的偶极矩重现性进行了测试, 结果列于表 3 和表 4。注意对于第一类电荷计算方法, 只有将原子电荷计算的偶极矩与计算电荷时所用的波函数的偶极矩期望值相互比较才是有意义的, 直接与实验值相比较是没有意义的, 表 4 显示当前所用的 HF/6-31G* 波函数偶极矩期望值普遍大于实验值 (平均大 0.26 Debye ($1 \text{ Debye} = 3.338 \times 10^{-30} \text{ C} \cdot \text{m}$)), 这主要是由于忽略了相关作用导致的。ADCH 是偶极矩保守的方法, 因此 ADCH 电荷计算的偶极矩与偶极矩期望值完全相符。由于静电势、偶极矩都是分子实际电荷分布的客观反映, 因此以重现静电势为目的的 MK 电荷也显示出很好的偶极矩重现性。有人提出在拟合静电势的过程中将偶极

表3 第一组方法计算的20种小分子的分子偶极矩(单位为 Debye)以及与 HF/6-31G*级别下偶极矩算符期望值($\langle\mu\rangle$)的比较
 Table 3 Molecular dipole moment (unit in Debye) of twenty small molecules calculated by the first group of methods, compared to the expected value of dipole moment ($\langle\mu\rangle$) operator at the HF/6-31G* level

Molecule	$\langle\mu\rangle$	Mulliken	AOIM	Hirshfeld	ADCH	NPA	MK	AIM
H ₂ O	2.20	2.39	1.88	0.88	2.20	2.62	2.25	3.25
CH ₃ OH	1.87	2.68	1.81	0.82	1.87	2.99	1.86	4.12
CH ₃ COCH ₃	3.12	3.80	2.98	2.21	3.12	4.66	3.15	8.10
CH ₃ COOH	1.79	2.00	1.98	1.77	1.79	2.40	1.82	5.40
NH ₃	1.92	1.77	1.15	0.51	1.92	1.97	1.97	1.97
HCN	3.21	3.65	1.80	1.82	3.21	3.09	3.18	9.29
CH ₃ NH ₂	1.53	1.60	0.87	0.36	1.53	1.81	1.50	2.32
NH ₂ CHO	4.10	4.27	3.92	2.99	4.10	5.18	4.09	7.24
CHF ₃	1.99	3.41	2.59	1.35	1.99	3.67	2.00	4.90
C ₆ H ₅ OH	1.45	2.32	1.14	0.55	1.45	2.33	1.45	4.48
CH ₃ PH ₂	1.27	0.38	0.99	0.33	1.27	1.13	1.21	6.48
PF ₃	1.38	4.95	6.49	1.93	1.38	7.17	1.41	9.75
tetrahydrofuran	1.94	2.92	2.19	1.14	1.94	3.35	1.91	5.10
3-oxetanone	0.88	0.16	0.18	0.94	0.88	0.31	0.92	2.27
γ -butyrolactone	5.01	6.48	5.01	3.66	5.01	7.17	5.05	12.07
2-methylpyrimidine	1.86	2.13	0.95	0.71	1.86	2.18	1.88	4.91
3-chloropropene	2.42	2.14	2.46	1.75	2.42	2.25	2.51	3.35
dimethylthioether	1.80	0.51	0.51	0.66	1.80	0.12	1.82	0.00
thiophene	0.90	0.05	0.40	0.08	0.90	1.48	0.85	0.26
vinylsilane	0.80	0.26	0.19	0.50	0.80	0.76	0.79	1.40
MSE ^a		0.32	-0.10	-0.82	0.00	0.76	0.01	2.76
MUE ^b		0.79	0.72	0.89	0.00	1.03	0.03	3.00

^a mean signed error; ^b mean unsigned error; 1 Debye=3.338×10⁻³⁰ C·m

矩能够精确重现作为限制条件,⁸⁵但实际上这么做已没有必要. Mulliken 方法的偶极矩重现性从统计结果上看并不理想, 对个别分子其偶极矩误差很大, 例如噻吩偶极矩期望值为 0.90 Debye, 而 Mulliken 方法却认为这是一个几乎无极性的分子. 需要指出的是 Mulliken 电荷的基组依赖性很大, 因此 Mulliken 电荷的偶极矩重现性也可能会随基组的不同而有一定变化. AOIM 的偶极矩重现性比 Mulliken 稍有提高. Hirshfeld 方法计算的偶极矩从平均无符号误差(MUE)上看比 Mulliken 方法更差, 而平均含符号误差(MSE)显示 Hirshfeld 的偶极矩是显著偏低的, 这与 Hirshfeld 电荷偏小有直接关系. NPA 电荷的偶极矩重现性很不理想, 有高估偶极矩的倾向, 尤其是对 PF₃ 的偶极矩高估了 4.2 倍, 而对二甲硫醚的偶极矩却低估了 10 倍以上. AIM 电荷对分子偶极矩基本没有重现能力, MUE 高达 3.00 Debye, 这个值已经大于大多数小分子偶极矩. AIM 的偶极矩误差主要来自于高估, 这很大程度上是由于 AIM 电荷数值过大所导致的.

在拟合参数过程中, MMFF94 的参数主要来自于使 MMFF94 电荷产生的偶极矩逼近 HF/6-31G*密

度产生的分子偶极矩, AM1-BCC 的参数来自于使 AM1 电荷在校正后逼近 HF/6-31G*级别下计算的 RESP 电荷. 它们都使用 HF/6-31G*下的数据作为目标数据是因为这个波函数级别对偶极矩的高估被认为可以等效反映出分子在凝聚相中存在的极化效应, 因此 MMFF94 和 AM1-BCC 的偶极矩应当与 HF/6-31G*下的偶极矩期望值相比较. 从表 4 可见这两种电荷的偶极矩重现性对大部分分子基本可以令人满意, MMFF94 的 MUE 为 0.36 Debye, 虽然 AM1-BCC 和 MMFF94 在键电荷校正过程中很相似, 但 AM1-BCC 的 MUE 更小, 为 0.25 Debye, 其改善在一定程度上可能是由于 AM1-BCC 使用了更有意义的初始电荷. Gasteiger 的偶极矩重现性无论以气相实验值为参考还是以 HF/6-31G*偶极矩为参考都显得一般, 并且对多数分子有低估偶极矩的倾向. CM2 电荷能够很好地重现气相实验偶极矩, MUE 仅为 0.15 Debye, 对所有分子都没有出现过大大误差, 这与 CM2 提出的目标完全相一致. QEq 对实验偶极矩和 HF/6-31G*下的偶极矩重现性都很不理想, 对许多分子偶极矩高估了近一倍甚至一倍以上, 严重低估偶极矩的情况同样存在, 如 2-甲基噻

表4 第二组方法计算的20种小分子的分子偶极矩(单位为Debye)以及与HF/6-31G*级别下偶极矩算符期望值和气相实验值的比较

Table 4 Molecular dipole moment (unit in Debye) of twenty small molecules calculated by the second group of methods, compared to the expected value of dipole moment operator at the HF/6-31G* level and gas-phase experimental value

Molecule	Exp. ^a	$\langle\mu\rangle$	MMFF94	AM1-BCC	Gasteiger	CM2	QEq
H ₂ O	1.85	2.20	2.37	2.17	1.13	1.83	2.05
CH ₃ OH	1.70	1.87	2.14	1.71	1.53	1.58	2.74
CH ₃ COCH ₃	2.88	3.12	3.73	3.01	2.57	2.91	3.46
CH ₃ COOH	1.70	1.79	1.65	1.64	1.31	1.85	2.52
NH ₃	1.47	1.92	1.92	1.75	0.61	1.49	1.34
HCN	2.99	3.21	3.93	2.91	1.89	2.68	1.85
CH ₃ NH ₂	1.31	1.53	1.79	1.45	0.73	1.18	1.39
NH ₂ CHO	3.73	4.10	4.30	3.98	2.22	3.59	3.06
CH ₃ F	1.86	1.99	2.23	1.80	1.97	1.94	4.47
C ₆ H ₅ OH	1.45	1.45	1.98	1.33	1.47	1.11	2.70
CH ₃ PH ₂	1.10	1.27	1.55	2.32	0.58	0.84	1.16
PF ₃	1.03	1.38	1.97	1.14	1.71	1.01	4.23
tetrahydrofuran	1.63	1.94	2.14	1.80	1.89	1.69	2.23
3-oxetanone	0.89	0.88	1.41	0.39	0.63	0.68	0.96
γ -butyrolactone	4.27	5.01	5.09	4.87	3.83	4.48	5.22
2-methylpyrimidine	1.68	1.86	2.55	1.87	1.14	1.30	0.64
3-chloropropene	1.94	2.42	2.45	2.37	1.28	1.95	5.93
dimethylthioether	1.50	1.80	2.57	2.12	1.16	1.37	3.06
thiophene	0.55	0.90	1.34	1.63	1.06	0.37	2.97
vinylsilane	0.66	0.80	0.25	1.26	- ^b	0.78	0.84
MSE (vs $\langle\mu\rangle$)			0.29	0.00	-0.63	-0.34	0.57
MUE (vs $\langle\mu\rangle$)			0.36	0.25	0.68	0.35	1.03
MSE (vs Exp.)		0.26	0.56	0.27	-0.36	-0.08	0.83
MUE (vs Exp.)		0.26	0.60	0.35	0.52	0.15	1.13

^a experimental dipole moments in gas phase, the values are obtained from supplemental material in literature⁵⁷;

^b The parameter for silicon is unavailable.

啞.

综上所述, Mulliken、AOIM、Hirshfeld、NPA、Gasteiger、QEq 电荷, 尤其是 AIM 电荷, 对偶极矩重现性都不好. ADCH 和 MK 都能很准确重现波函数的偶极矩期望值, 若波函数完全精确, 则这两种电荷计算的偶极矩将与气相的实际值相一致, 而利用 CM2 方法则可以在较低波函数级别下很准确地预测气相偶极矩. MMFF94 和 AM1-BCC 能够较好地重现 HF/6-31G* 的偶极矩.

7 静电势的重现性

当研究的问题涉及分子间较近距离的静电相互作用时偶极模型显得过于简单, 只有分子范德华表面外侧精确静电势也能被较好重现时, 才说明这种原子电荷适合研究分子间近距离静电作用. 又由于静电势是可观测量, 因此静电势的重现性与偶极矩一样也是检验原子电荷合理性的重要、客观的指标. 其重现性常以原子电荷在分子范德华表面外侧

所取的格点位置上产生的静电势 V^{Model} 与波函数计算的精确静电势 V^{Real} 之间的相对均方根偏差 (RRMSE) 来反映, 计算公式为

$$\text{RRMSE} = \sqrt{\frac{\sum_i [V^{\text{Real}}(\mathbf{r}_i) - V^{\text{Model}}(\mathbf{r}_i)]^2}{\sum_i V^{\text{Real}}(\mathbf{r}_i)^2}} \quad (20)$$

我们选取了 21 种中性分子和 4 种离子用于测试不同原子电荷计算方法的静电势重现性, 见表 5 和表 6. 其中既包含一部分测试偶极矩重现性所用的小分子, 也包含结构更复杂的生物、药物分子. 计算 RRMSE 所用格点设定与计算 MK 电荷时的一致, 参考静电势在 HF/6-31G* 级别下获得.

由于 MK 电荷正是通过最小化静电势重现性误差得到的, 因此它的 RRMSE 拥有理论最低值. 从 RRMSE 的平均值上看, 其余 11 种方法的静电势重现能力可以大致分为四个档次: (1) 重现性很好, 平均 RRMSE 在 0.25 上下. 这一类包括 ADCH、MMFF94、AM1-BCC 和 CM2. 如果改用能够比较准

表5 第一组方法产生的静电势相对于HF/6-31G*级别下精确静电势的相对均方根偏差
 Table 5 Relative root mean square error (RRMSE) of the ESP produced by the first group of methods relative to the exact ESP at the HF/6-31G* level

Molecule	Mulliken	AOIM	Hirshfeld	ADCH	NPA	MK	AIM
CH ₃ CH ₂ CH ₃	2.148	1.075	0.983	1.199	2.810	0.725	1.064
CH ₃ NO ₂	0.480	0.157	0.297	0.142	0.366	0.065	0.908
NH ₂ CHO	0.199	0.215	0.317	0.145	0.338	0.071	1.095
(CH ₃) ₂ NNH ₂	0.443	0.466	0.746	0.309	0.473	0.235	0.967
CH ₃ PH ₂	1.252	1.580	1.063	0.551	1.713	0.513	5.276
C ₆ H ₅ SO ₃ H	0.390	0.526	0.450	0.133	0.380	0.066	1.297
OMS(CH ₃ SO ₃)	0.186	0.362	0.340	0.139	0.375	0.063	1.213
tetrahydrofuran	0.696	0.238	0.479	0.254	0.879	0.162	1.583
uracil	0.378	0.276	0.297	0.131	0.335	0.059	1.473
adenine	0.249	0.284	0.396	0.194	0.283	0.059	0.724
alanine	0.240	0.365	0.568	0.230	0.336	0.091	1.299
3-chloropropene	0.318	0.276	0.459	0.215	0.347	0.177	0.538
λ -butyrolactone (C ₄ H ₆ O ₂)	0.390	0.094	0.326	0.189	0.471	0.076	1.443
3-oxetanone (C ₃ H ₄ O ₂)	0.509	0.305	0.465	0.266	0.560	0.127	1.490
TTF (C ₆ H ₄ S ₄)	0.804	0.706	0.659	0.322	0.918	0.273	0.886
TCNQ (C ₁₂ H ₄ N ₄)	0.306	0.217	0.336	0.173	0.485	0.060	2.560
aspirin (C ₉ H ₈ O ₄)	0.367	0.242	0.351	0.106	0.436	0.043	1.710
dopamine (C ₈ H ₁₁ NO ₂)	0.527	0.477	0.685	0.206	0.466	0.091	1.740
glucose (C ₆ H ₁₂ O ₆)	0.522	0.246	0.572	0.176	0.589	0.062	1.063
ephedrine (C ₁₀ H ₁₅ NO)	0.587	0.486	0.603	0.209	0.780	0.120	1.225
naphthoquinone (C ₁₀ H ₆ O ₂)	0.547	0.475	0.408	0.263	0.574	0.080	2.196
CH ₃ COO ⁻	0.032	0.015	0.068	0.029	0.052	0.010	0.128
PO ₄ (CH ₃) ₂ ⁻	0.073	0.040	0.055	0.030	0.076	0.018	0.127
CH ₃ NH ₃ ⁺	0.033	0.020	0.018	0.011	0.021	0.011	0.055
morpholine ⁺ (C ₄ H ₁₀ NO ⁺)	0.060	0.020	0.045	0.018	0.077	0.012	0.119
mean ^a	0.428	0.337	0.461	0.200	0.494	0.104	1.337

^a except for ions, CH₃CH₂CH₃, and CH₃PH₂

确地重现气相偶极矩的波函数, 即使用 B3LYP/cc-pVTZ 波函数代替 HF/6-31G* 去计算参考静电势, 则 CM2 的平均 RRMSE 值会从 0.268 降低到 0.228. (2) 重现性不错. AOIM 属于此类, 平均 RRMSE 为 0.337. (3) 重现性一般, RRMSE 平均值在 0.4–0.6. 包括 Mulliken、Hirshfeld、NPA、Gasteiger 和 QEq. (4) 重现性很差, 这一类只包括 AIM, 平均 RRMSE 高达 1.53.

通过检验 MK 电荷的数据, 可以看出原子电荷模型本身对于一些体系存在明显局限性. 对于丙烷和 CH₃PH₂ 分子, 即便是 MK 电荷, 其 RRMSE 也分别高达 0.72 和 0.51, 因此我们未将这两个分子纳入 RRMSE 平均值的统计中. 若想更准确地描述静电相互作用, 即降低 RRMSE, 就必须增加一些非原子中心的点电荷, 或者在原子中心引入点多极矩.²⁵ 若在计算 MK 电荷时在每个原子中心都引入一个点偶极矩(点偶极矩向量与原子电荷一起被拟合), 则丙烷和 CH₃PH₂ 分子的 RRMSE 会分别降至 0.22 和

0.13, 静电势重现性大大提高. 而对于一些体系用原子电荷模型表现静电势是足够准确的, 如 MK 电荷对 Aspirin 的 RRMSE 仅有 0.043.

从表中可见所有方法对离子的静电势重现性都显著好于中性分子, RRMSE 降低了约一个数量级. 出现这种情况是由于带电分子外围静电势主要由分子单极矩所主导, 而所测试的所有原子电荷计算方法都是单极矩保守的, 即原子电荷加和等于分子净电荷. 不同方法对带电分子静电势重现性的优劣关系和中性分子情况是相似的.

8 波函数质量的依赖性

图 5 比较了第一类方法计算的乙酸的甲基碳原子电荷的基组依赖性, 从左到右基组完备性逐渐增加. 从 STO-3G⁸⁶ 提升至 3-21G⁸⁷ 的过程中各种方法计算的结果都有不同程度的变化, 通常被认为具有很好基组稳定性的 NPA 电荷的变化值甚至大于公认基组稳定性差的 Mulliken 电荷. 注意由于极小基的

表6 第二组方法产生的静电势相对于HF/6-31G*级别下精确静电势的相对均方根偏差

Table 6 RRMSE of the ESP produced by the second group of methods relative to the exact ESP at the HF/6-31G* level

Molecule	MMFF94	AM1-BCC	Gasteiger	CM2	QEq
CH ₃ CH ₂ CH ₃	1.000	1.054	1.049	1.381	1.994
CH ₃ NO ₂	0.150	0.349	0.412	0.322	0.318
NH ₂ CHO	0.112	0.142	0.516	0.171	0.331
(CH ₃) ₂ NNH ₂	0.329	0.337	0.658	0.340	0.590
CH ₃ PH ₂	0.582	0.880	0.728	1.152	0.594
C ₆ H ₅ SO ₃ H	0.342	0.214	1.129	0.238	0.671
OMS(CH ₃ SO ₃)	0.155	0.119	1.257	0.264	0.505
tetrahydrofuran	0.208	0.217	0.211	0.286	0.451
uracil	0.229	0.119	0.561	0.228	0.373
adenine	0.262	0.147	0.736	0.217	0.694
alanine	0.217	0.172	0.553	0.301	0.733
3-chloropropene	0.263	0.250	0.534	0.307	1.246
λ-butyrolactone	0.093	0.110	0.341	0.234	0.266
3-oxetanone	0.219	0.240	0.323	0.311	0.326
TTF(C ₈ H ₄ S ₄)	0.424	0.482	0.609	0.394	0.604
TCNQ (C ₁₂ H ₄ N ₄)	0.332	0.169	0.520	0.187	0.439
aspirin (C ₉ H ₈ O ₄)	0.247	0.154	0.365	0.188	0.434
dopamine (C ₈ H ₁₁ NO ₂)	0.267	0.171	0.569	0.297	0.717
glucose (C ₆ H ₁₂ O ₆)	0.204	0.154	0.433	0.240	0.473
ephedrine (C ₁₀ H ₁₅ NO)	0.355	0.248	0.492	0.244	0.536
naphthoquinone (C ₁₀ H ₆ O ₂)	0.375	0.155	0.433	0.329	0.400
CH ₃ COO ⁻	0.023	0.024	0.056	0.053	0.077
PO ₄ (CH ₃) ₂	0.065	0.027	0.047	0.058	0.118
CH ₃ NH ₃ ⁺	0.016	0.017	0.024	0.012	0.114
morpholine ⁺ (C ₄ H ₁₀ NO ⁺)	0.047	0.030	0.052	0.018	0.126
mean ^a	0.252	0.208	0.561	0.268 ^b	0.532

^aexcept for ions, CH₃CH₂CH₃, and CH₃PH₂; ^bIf ESP is produced at the B3LYP/cc-pVTZ level, the value is 0.228.

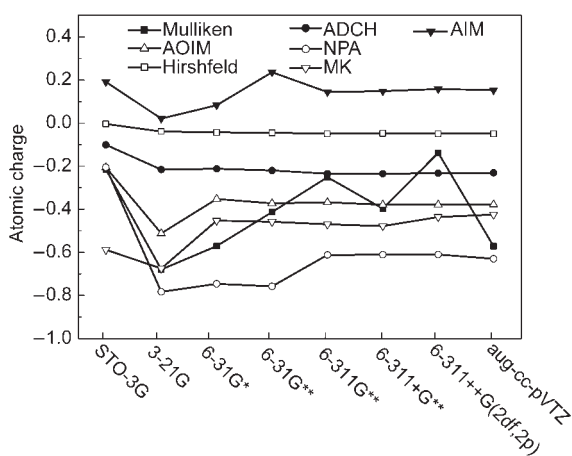


图5 不同方法结合不同基组在Hartree-Fock级别下获得的乙酸的甲基碳原子电荷

Fig.5 Atomic charge of methyl carbon of acetic acid calculated by different methods under various basis-sets at the Hartree-Fock level

完备性很低,所以提升至分裂价基后结果的明显变化并不能说明方法的合理性差.从6-311G**⁸⁸开始除了Mulliken方法外所有方法的结果都已经收敛,

提升至aug-cc-pVTZ^{89,90}的过程中结果变化都不显著.其中Hirshfeld和ADCH电荷收敛得最早,从3-21G开始就基本不发生变化.同时也说明ADCH对Hirshfeld电荷的校正并没有破坏Hirshfeld电荷良好的基组稳定性. MK和AOIM从6-31G*开始收敛,而NPA和AIM收敛得相对稍迟.唯一不能随基组增加而收敛的方法是Mulliken,曲线在基组增大过程中始终强烈波动,对6-311G**基组下重原子增加弥散函数成为6-311+G**⁷³后,其它方法得到的甲基碳电荷变化均不超过0.01,而Mulliken电荷数值则降低了0.15,这是因为碳的弥散函数严重侵入到了氢的原子空间内,原本在6-311G**基组下属于氢的电子密度现在有一部分被碳的基函数所描述,因此氢的电子占据数一部分被归属到了碳.而进一步提升到6-311++G(2df,2p)⁷³后,由于氢的弥散函数也明显侵入到了碳的原子空间内,碳的电子占据数有不少被划归给了氢,因此碳的Mulliken电荷又增加了许多.6-311++G(2df,2p)已经是完备性很高的基

组, 继续提升质量后原子电荷应当不出现明显改变, 然而提升至 aug-cc-pVTZ 后 Mulliken 电荷变化却仍高达 0.43. 因此从基组依赖性的角度上讲, Mulliken 方法是存在严重缺陷的, 没有办法“唯一”地获得 Mulliken 电荷, 基组的轻微改变就可能使结果有定性的变化. 同时也说明提升基组质量并不会提升 Mulliken 电荷的合理性, 用大基组计算 Mulliken 电荷是没有意义的. 实际上大基组, 尤其是包含弥散函数的情况, 基函数的化学意义往往比较差, 与原子轨道缺乏对应性, 而极小基组则能够与原子轨道相互对应, 会使 Mulliken 方法的计算过程更有物理意义. 以上的讨论对于乙酸的其它原子的电荷也同样适用.

第一类方法对计算波函数所用理论方法的依赖性如图 6 所示, 其中包括不含相关效应的 Hartree-Fock 方法、广义梯度近似(GGA)泛函 BP86、^{91,92} 杂化泛函 B3LYP,⁹³ 以及 MP2 和 CCSD⁷⁰ 方法, 基组均使用 cc-pVDZ.⁸⁹ 由于相关效应对多重键影响较大, 所以这里选取乙酸的羰基氧原子用于测试. 由于 AOIM 1.1 程序只能读取 SCF 波函数, 因此 MP2 和 CCSD 的结果未在图中显示, 但实际上利用自然轨道形式不难将 AOIM 的应用范围扩展到后 HF 波函数上. 从图 6 可见对于任何一种计算原子电荷的方法, 无论通过何种方式引入相关效应, 所得结果都很相近, 数值差异最大不超过 0.05. 而 Hartree-Fock 波函数和各种相关波函数下的计算结果却差异明显, 引入相关效应后所有方法给出的羰基氧原子电荷均明显减小, 这反映出相关效应会减小化学键的极性这一普遍现象. 通过检验乙酸各个原子的电荷会发现

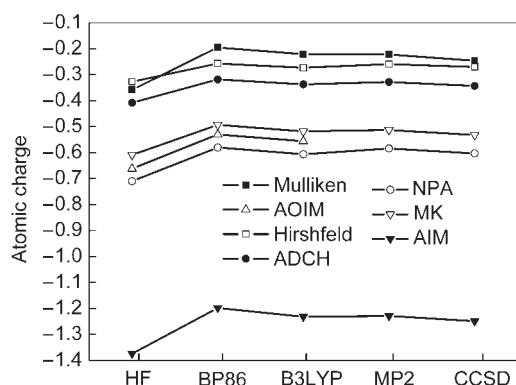


图 6 不同方法结合 cc-pVDZ 基组 and 不同哈密顿得到的乙酸的羰基氧原子电荷

Fig.6 Atomic charge of carbonyl oxygen of acetic acid calculated by different methods in combination with cc-pVDZ basis-set and various Hamiltonians

Hirshfeld 和 ADCH 电荷受电子相关效应影响相对于其它方法更弱, 比如 BP86 下 Hirshfeld 和 ADCH 的羰基氧原子电荷相对于 Hartree-Fock 下的结果差异分别为 0.07、0.09, 而 Mulliken、AOIM、NPA、MK 和 AIM 电荷变化分别为 0.16、0.13、0.13、0.12 和 0.18.

由上可见, 至少对于普通有机分子, 获得原子电荷并不需要高级别后 HF 方法计算波函数, 选择适当的密度泛函方法计算的波函数就可以满足要求. 对于 NPA、MK、AIM, 尤其是 Hirshfeld 和 ADCH 方法, 只需要中等质量的基组就已经足够. 在此之上继续提高波函数质量意义不大.

9 计算的耗时

计算方法的耗时大小直接关系到它的适用领域. 计算耗时不仅取决于方法本身的复杂度, 也明显依赖于计算程序的代码效率、数值方法和编译、运行时的软硬件环境. 另外, 很多方法的可变参数也会影响耗时, 比如 AIM、Hirshfeld 方法中对积分的期望精度、AOIM 的 L-BFGS-B 步骤中的收敛阈值、静电势拟合时拟合点的密度等. 本节对各种方法计算耗时的讨论只是定性的, 计算耗时以当前普通实验室的计算能力作为参照.

MMFF94、Gasteiger 和 QEq 方法由于不依赖于波函数, 方法本身也比较简单, 所以计算这三种电荷总耗时极小, 即使对于几千个原子的体系也可以在数秒内给出结果, 它们也能够用于需要高效计算大批量小分子电荷的情况, 比如药物高通量筛选. 由于 QEq 计算量小, 而且是依赖于几何结构的, 能反映出原子电荷对周围环境变化的响应, 因此 QEq 可以作为浮动电荷力场⁹⁴的电荷模型, 即在动力学/蒙特卡罗模拟过程中每隔一定步数或构象改变一定程度后自动更新原子电荷.

AM1-BCC 电荷的计算总耗时完全来自于 AM1 单点计算, BCC 校正过程几乎不需要耗时. 对于几十个原子体系的 AM1 计算, 通常在数秒内能够完成. 对于数百个原子的体系, AM1 计算就较为耗时, 而上千个原子的体系则十分困难. 但是如果使用线性标度方法, 如 MOZYME 方法⁹⁵来加速 AM1 步骤, 则 AM1-BCC 电荷的适用范围可以被大大扩展.

得到波函数后, 计算 Mulliken 电荷不需要额外耗时, 而计算 NPA、MK、Hirshfeld 和 ADCH 电荷的计算耗时也只占波函数计算耗时的很小比例. 因此

可以近似认为计算这几种电荷的总耗时就是计算波函数的耗时。

使用 AOIM 1.1 程序计算基组较大或原子数稍多体系的 AOIM 电荷比较慢。例如在 HF/6-31G* 下计算 Ephedrine ($C_{10}H_{15}NO$) 分子的 AOIM 电荷耗费约 8 min, 而使用默认参数、单线程下使用 Gaussian 03 计算这个分子波函数耗时 40 s。我们认为计算速度缓慢不是 AOIM 理论本身的原因, 而是因为 AOIM 1.1 代码中的 L-BFGS-B 最小化模块效率有待改善。对代码和算法进行优化和并行化后有望使 AOIM 电荷计算速度显著加快。

AIM 电荷计算耗时极长。我们使用计算效率较高的 AIM 程序 AIMALL, 并采用默认的积分精度和积分算法时, 计算 Ephedrine 分子全部 AIM 电荷总耗时为 10.5 min, 是计算波函数时间的近 16 倍。AIM 电荷计算费时是因为 AIM 的原子边界复杂, 导致原子盆不容易被准确积分, 尽管已有不少研究者探索出更有效的原子盆积分方法,⁹⁶⁻⁹⁸ 但 AIM 电荷计算缓慢的问题始终没有被彻底解决。

CM2 的校正步骤几乎不需要额外时间, 总耗时取决于使用何种理论级别计算初始电荷。如果要求计算快速, 可以用半经验方法, 如果要求对实际气相偶极矩有更好的重现能力, 则可以用从头算方法。实际上从原文献的统计数据来看, 半经验与从头算方法所得 CM2 电荷的偶极矩重现能力对于多数体系差距并不很大。

10 计算方法的选择

结合上文对各种原子电荷计算方法的对比结果、方法特点以及目前应用的现状, 我们对原子电荷计算方法的选择给出以下建议。

对于分子体系的量子化学研究, 建议使用 NPA 和 ADCH 电荷。NPA 电荷目前应用十分广泛, 合理性总体表现较好。但注意不宜使用 NPA 电荷分析静电相互作用, 同时注意 NPA 用于过渡金属、镧系铜系金属时存在一定问题。^{23,99} 问题主要是由于这些元素的极小集/里德堡集划分标准比较含糊, 而不同的划分又会影响 OWSO 过程, 进而影响电荷数值。最近提出的 ADCH 方法无论在电荷合理性、基组依赖性、偶极矩重现性和静电势的重现性上都表现不错, 电荷计算时间也远小于波函数计算时间, 是十分有前景的原子电荷计算方法, 其普适性、稳定性有待在更多的实际应用中进行检验。

基于原子电荷的分子模拟任务需要原子电荷能够较好地描述静电相互作用。通常来说以 MK 为代表的拟合静电势电荷是最适合用于分子模拟的。但如果模拟的是有机分子, 更适宜使用 RESP 电荷, 因为它解决了普通拟合静电势方法存在的构象依赖性、内部原子电荷数值不稳定性和原子等价性问题, 若这些问题存留则可能会影响模拟过程的合理性。注意由于在拟合力场参数时范德华作用、1-4 位相互作用和静电作用是相互耦合的, 故范德华参数和二面角参数对电荷模型有依赖性, 因此若模拟任务依赖于已有力场参数, 则原子电荷的计算方法应当尽量与力场原文中的方法一致。

对于较大分子体系或者大量有机小分子体系的原子电荷的计算建议使用 AM1-BCC 方法, 此方法表现静电作用较好且计算速度较快。如果需要进一步降低计算耗时, 可以改用 MMFF94 方法。另外, 如果所研究的大分子体系结构具有规律性, 如蛋白质、多糖、核酸、高分子, 也可以使用能够更好重现静电势的 MK 方法获得每个片段的电荷再进行拼接。

估算气相分子偶极矩最适合使用 CM2 方法, 可以以较低的计算花费获得很精确的气相偶极矩。

最后, 本文对目前使用广泛的 Mulliken、AIM 和 Gasteiger 电荷进行简要评述。

Mulliken 方法尽管存在原理不严格、低估离子性化合物的键的极性、基组依赖性大、偶极矩和静电势重现性差等诸多缺陷, 但是从本文的测试看到, Mulliken 电荷对于多数小分子并没有出现过于明显的不合理性, 又由于 Mulliken 电荷几乎被所有量子化学软件支持, 不耗费额外计算时间, 因此 Mulliken 电荷虽然不推荐使用, 但仍可以作为定性的参考。在极小基下产生的 Mulliken 电荷比在更高级别基组下产生结果在原理上更合理。如果所用基组包含弥散函数, 则应当舍弃 Mulliken 电荷。另外, 由于 Mulliken 电荷受基组影响明显, 比较 Mulliken 电荷时应注意必须处于在相同基组下, 否则结果没有意义。

AIM 方法一般不建议使用。尽管它拥有严格的物理解释, 但是在测试中它的结果往往明显违背化学观念, 对偶极矩和静电势重现性都很差, 当体系存在赝原子时无法使用, 而且计算耗时极长。

Gasteiger 电荷由于具有计算方便、快速的特点, 被很多药物设计、分子对接软件所采用, 常被用于

描述受体-配体静电相互作用、生成 CoMFA 任务所需要的静电势场. 然而从前文的对比结果可见它的静电势重现性并不理想, 因此并不推荐 Gasteiger 电荷用于这些场合. 相比之下使用 AM1-BCC 和 MMFF94 电荷更为合适.

11 总 结

原子电荷是十分古老、简单的模型, 可以追溯到化合价概念的提出. 随着理论化学的发展以及人们对电子结构认识的加深, 原子电荷模型不仅没有被遗弃, 新的计算方法还在不断涌现, 这是由于它有着重要的理论和实用意义. 本文介绍了目前使用最为广泛的 12 种计算原子电荷的方法, 通过大量分析对它们的特点进行了全面的比较和评述. 对比分析中显示, 许多计算方法的结果间存在很大差异, 错误地选用计算方法可能得到不可靠甚至没有意义的原子电荷. 只有充分掌握了众多计算方法各自的原理和特点, 才能根据实际问题选择最为适用的一种.

References

- (1) Qian, B. H.; Ma, W. X.; Lu, L. D.; Yang, X. J.; Wang, X. *Acta Phys. -Chim. Sin.* **2010**, *26*, 610. [钱保华, 马卫兴, 陆路德, 杨绪杰, 汪 信. 物理化学学报, **2010**, *26*, 610.]
- (2) Zheng, W. R.; Xu, J. L.; Xiong, R. *Acta Phys. -Chim. Sin.* **2010**, *26*, 2535. [郑文锐, 徐善利, 熊 瑞. 物理化学学报, **2010**, *26*, 2535.]
- (3) Shen, T.; Du, F. P.; Liu, T.; Yao, G. W.; Wu, Z.; Fang, M. M.; Xu, X. J.; Lu, H. Z. *Acta Phys. -Chim. Sin.* **2011**, *27*, 1831. [申 涛, 杜凤沛, 刘 婷, 姚广伟, 吴 峥, 方萌萌, 徐筱杰, 路慧哲. 物理化学学报, **2011**, *27*, 1831.]
- (4) Zhou, J. J.; Chen, H. M.; Xie, G. R.; Ren, T. R.; Xu, Z. H. *Prog. Chem.* **1998**, *10*, 55. [周家驹, 陈红明, 谢桂荣, 任天瑞, 许志宏. 化学进展, **1998**, *10*, 55.]
- (5) Ji, G. D.; Zhao, Y. H.; Yuan, X. *J. Northeast Normal Univ. (Natural Science Edition)* **1998**, *47*. [籍国东, 赵元慧, 袁 星. 东北师范大学学报(自然科学版), **1998**, *47*.]
- (6) Ding, Y. F.; Zhang, Y.; Zhang, D. H.; Li, Z. P. *Acta Phys. -Chim. Sin.* **2010**, *26*, 1651. [丁元法, 张 跃, 张大海, 李仲平. 物理化学学报, **2010**, *26*, 1651.]
- (7) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300.
- (8) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916.
- (9) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
- (10) Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137.
- (11) Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1997**, *98*, 85.
- (12) Li, J. B.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293.
- (13) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6532.
- (14) Meister, J.; Schwarz, W. H. E. *J. Phys. Chem.* **1994**, *98*, 8245.
- (15) Cramer, C. J. *Essentials of Computational Chemistry*, 2nd ed.; John Wiley & Sons: West Sussex, 2004; pp 309–324.
- (16) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons: West Sussex, 2007; pp 293–304.
- (17) Young, D. C. *Computational Chemistry*; John Wiley & Sons: New York, 2001; pp 99–105.
- (18) Cioslowski, J. Electronic Wavefunction Analysis. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R. Ed.; John Wiley & Sons: West Sussex, 1998; Vol. 2, pp 892–905.
- (19) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1841.
- (20) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
- (21) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 2338.
- (22) Bachrach, S. M. Population Analysis and Electron Densities from Quantum Mechanics. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B. Eds.; VCH Publishers: New York, 1994; Vol. 5, pp 171–227.
- (23) Clark, A. E.; Sonnenberg, J. L.; Hay, P. J.; Martin, R. L. *J. Chem. Phys.* **2004**, *121*, 2563.
- (24) Martin, F.; Zipse, H. *J. Comput. Chem.* **2005**, *26*, 97.
- (25) Wiberg, K. B.; Rablen, P. R. *J. Comput. Chem.* **1993**, *14*, 1504.
- (26) Lu, H. G.; Dai, D. D.; Yang, P.; Li, L. M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 340.
- (27) Hirshfeld, F. L. *Theor. Chem. Acc.* **1977**, *44*, 129.
- (28) Lu, T.; Chen, F. W. *J. Theor. Comput. Chem.* accepted.
- (29) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
- (30) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
- (31) Bader, R. F. W.; Beddall, P. M. *J. Chem. Phys.* **1972**, *56*, 3320.
- (32) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.
- (33) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616.
- (34) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
- (35) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623.
- (36) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
- (37) Li, J. B.; Zhu, T. H.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820.
- (38) Rappe, A. K.; Goddard, W. A. *J. Phys. Chem.* **1991**, *95*, 3358.
- (39) Cusachs, L. C.; Politzer, P. *Chem. Phys. Lett.* **1968**, *1*, 529.
- (40) Stout, E. W.; Politzer, P. *Theor. Chem. Acc.* **1968**, *12*, 379.
- (41) Doggett, G. *J. Chem. Soc. A* **1969**, 229.

- (42) Christoffersena, R. E.; Baker, K. A. *Chem. Phys. Lett.* **1971**, *8*, 4.
- (43) Bickelhaupt, F. M.; van Eikema Hommes, N. J. R.; Fonseca Guerra, C.; Baerends, E. J. *Organometallics* **1996**, *15*, 2923.
- (44) Weinhold, F. Natural Bond Orbital Methods. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R. Ed.; John Wiley & Sons: West Sussex, 1998; Vol. 2, pp 1792–1811.
- (45) Glendening, E. D.; Badenhop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. *NBO*, Version 5.0, 2001. <http://www.chem.wisc.edu/~nbo5/>.
- (46) Liu, W.; Li, L. *Theor. Chem. Acc.* **1997**, *95*, 81.
- (47) Sanchez-Portal, D.; Artacho, E.; Soler, J. M. *Solid State Commun.* **1995**, *95*, 685.
- (48) Sanchez-Portal, D.; Artacho, E.; Soler, J. M. *J. Phys.: Condens. Matter* **1996**, *8*, 3859.
- (49) Bader, F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: New York, 1994.
- (50) Nalewajski, R. F.; Parr, R. G. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8879.
- (51) Davidson, E. R.; Chakravorty, S. *Theor. Chem. Acc.* **1992**, *83*, 319.
- (52) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.
- (53) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
- (54) Sigfridsson, E.; Ryde, U. *J. Comput. Chem.* **1998**, *19*, 377.
- (55) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
- (56) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (57) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. -Aided Mol. Des.* **1995**, *9*, 87.
- (58) Li, J. B.; Williams, B.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 724.
- (59) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291.
- (60) Winget, P.; Thompson, J. D.; Xidos, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2002**, *106*, 10707.
- (61) Kalinowski, J. A.; Lesyng, B.; Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2545.
- (62) Olson, R. M.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 2046.
- (63) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133.
- (64) Mayer, I. *Chem. Phys. Lett.* **1983**, *97*, 270.
- (65) Sanderson, R. T. *Science* **1951**, *114*, 670.
- (66) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
- (67) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, *108*, 4315.
- (68) Cioslowski, J. *Phys. Rev. Lett.* **1989**, *62*, 1469.
- (69) Cioslowski, J. *J. Am. Chem. Soc.* **1989**, *111*, 8333.
- (70) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*, 1st rev ed.; Dover Publications: New York, 1989.
- (71) Hariharan, P. C.; Pople, J. A. *Theor. Chem. Acc.* **1973**, *28*, 213.
- (72) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
- (73) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
- (74) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; et al. *Gaussian 03*, Revision E.01; Gaussian Inc.: Wallingford, CT, 2004.
- (75) Lu, T. *Multifwfn*, Version 2.1.2; 2011. <http://Multiwfn.codeplex.com>.
- (76) Keith, T. A. *AIMALL*, Version 10.05.04, 2010.
- (77) Lu, H. G. *AOIM*, Version 1.1, 2006; <http://faculty.sxu.cn/luhg/aoim.html>.
- (78) *Avogadro: an Open-Source Molecular Builder and Visualization Tool*, Version 1.0.3, 2011.
- (79) Case, D. A.; Darden, T. A.; Cheatham, T. E. C., III; et al. *AmberTools*, Version 1.5; 2011.
- (80) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.
- (81) *PETRA Manual*. <http://www2.ccc.uni-erlangen.de/software/petra/manual> (accessed Sep 12, 2011).
- (82) Marsili, M.; Gasteiger, J. *Croat. Chem. Acta* **1980**, *53*, 601.
- (83) *The Open Babel Package*, Version 2.3.0; 2010; <http://openbabel.sourceforge.net>.
- (84) Sanner, M. F. *J. Mol. Graph. Model.* **1999**, *17*, 57.
- (85) Woods, R. J.; Khalil, M.; Pell, W.; Moffat, S. H.; Smith, V. H. *J. Comput. Chem.* **1990**, *11*, 297.
- (86) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.
- (87) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
- (88) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- (89) Dunning, J. T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (90) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (91) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (92) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
- (93) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (94) Patel, S.; Brooks, C. L. *Mol. Simul.* **2006**, *32*, 231.
- (95) Stewart, J. J. P. *Int. J. Quantum Chem.* **1996**, *58*, 133.
- (96) Biegler-König, F. W. *J. Comput. Chem.* **2000**, *21*, 1040.
- (97) Biegler-König, F. W.; Bader, R. F. W.; Tang, T. H. *J. Comput. Chem.* **1982**, *3*, 317.
- (98) Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. *J. Comput. Chem.* **2007**, *28*, 899.
- (99) Maseras, F.; Morokuma, K. *Chem. Phys. Lett.* **1992**, *195*, 500.