

·“双清论坛”专题:理论化学家视角中的仪器创制·

## 国内计算化学方法与软件开发

蒋帆 吴云东\*

(北京大学深圳研究生院,深圳 518055)

**[摘要]** 当前理论与计算化学正处于蓬勃发展时期,其中计算方法和软件的开发具有基础性地推动作用,类似实验科学中的先进科学仪器的创制。近年来,我国已有几十个研究小组在电子结构计算、动力学模拟、分子力场、药物设计、材料结构与性质预测等领域开发了各种新的计算方法,其中有些已经达到国际先进水平,且很多都有较高的实际应用价值。另一方面,很多研究工作仍然依赖国外的计算软件作为基础或平台,开发具有自主知识产权的计算化学软件平台将提升这些国内新方法的影响力并促进它们的广泛应用。

**[关键词]** 量子化学;分子模拟;计算方法;软件开发

随着信息技术迅速且持续地发展,计算在科学研究中发挥着越来越重要的作用,成为连接科学理论和实验现象的关键桥梁。当前,国际上理论与计算化学领域正在蓬勃发展,有大量与材料科学或生命科学相关的计算研究。而这些成就的基础是优秀的计算化学软件,正如实验科学中使用的各种“科学仪器”。其中最著名的包括 GAUSSIAN(1998 年诺贝尔奖得主 J. A. Pople 等开发)、GAMESS 等量子化学软件,CHARMM (2013 年诺贝尔奖得主 M. Karplus 等开发)、AMBER、GROMACS、NAMD 等分子模拟软件,以及常用于材料计算的 VASP 等软件。另外,还有一些商业化的药物设计或材料模拟软件,包括 Schrodinger, MOE, Materials Studio 等,比较昂贵,但整合了各种不同的功能模块和图形界面。

得益于国家对科研投入的逐年增长,我国计算化学在近年来有很大的发展,发表论文的数量和质量都有大幅度提升,有不少工作发表在相关方向的顶级期刊上。然而,大部分此类研究使用的都是上述的国外计算软件。缺少独立完整、功能强大、且被广泛使用的计算软件很大程度上限制了我国理论与计算化学在国际学术界的地位。尽管如此,我国仍然有一些开展各种计算方

法研究的科研小组,在国家科研基金资助下,开发出了独立的计算软件模块,与一些国外的计算程序对接,或者作为独立运行的程序,实现新的功能、更准确的结果、更高的计算效率等。下面介绍一些有代表性的国内工作。

### 1 量子化学计算方法与软件

量子力学是人类认识微观世界的理论根基,也是理论与计算化学的基石。在唐敖庆、卢嘉锡、徐光宪等老一辈杰出化学家的带领下,我国在量子化学领域长期以来就有较好的基础,近年来在开发新的计算方法和软件上也有不少进展。

#### 1.1 低标度及线性标度算法

基于分子片断的方法只需要计算各个分子片断和片断之间的相互作用能,可以对较大的体系进行高效的量子化学计算。南京大学黎书华、李伟、马晶等发展了此类基于能量的分块(EBF)方法<sup>[1]</sup>。推广的 GEBF 方法<sup>[2]</sup>通过引入背景点电荷,来近似描述长程静电和极化作用,可以较准确地计算高度极性的大体系如水分子簇。GEBF 可以与各种 QM 方法结合使用(GEBF-X),计算复杂体系的能量、几何结构、谱学性质,进行从头分子动力学模拟等<sup>[3]</sup>。

收稿日期:2017-10-18;修回日期:2017-12-13

\* 通信作者,Email: wuyd@pkusz.edu.cn

厦门大学徐昕(现复旦大学)、郭文平等,通过排斥原理将分子片方法与 ONIOM 方法结合起来,提出扩展的 ONIOM 方法 XO<sup>[4]</sup>。此方法在体系划分和计算方法组合上有更大的自由度,可针对不同体系定制计算方案。另外,徐昕、吴安安等,结合混合基组方法与“分而治之”的策略,发展了 DCMB 方法<sup>[5]</sup>,分别对体系的不同部分使用大基组而其余部分用小基组,最后再整合。尽管不是严格意义上的低标度方法,但效率较高且适合并行计算。这些方法已与 GAUSSIAN 等软件嫁接。此外,华南师范大学顾凤龙与杜克大学杨伟涛等,合作发展了利用非正交局域分子轨道(NOLMO)进行自洽场计算的方法<sup>[6]</sup>,可用于较大体系的线性标度计算,已连接到 GAMESS 量化软件上。

材料体系的电子结构计算存在周期性边界条件等,中国科学技术大学杨金龙等发展了基于密度矩阵微扰理论计算固体中电场响应(如光频介电常数)的线性标度算法<sup>[7]</sup>,以及利用数值原子基组的线性标度算法<sup>[8]</sup>。在此基础上,近来他们开发了一个线性标度 DFT 电子结构计算的开源程序包 HONPAS<sup>[9]</sup>,它与 SIESTA 计算软件对接。

## 1.2 后自洽场电子相关方法

黎书华等开发了能实现线性标度的电子相关方法“cluster-in-molecule”(CIM)<sup>[10-12]</sup>。此方法在局域分子轨道(LMO)上求解多体微扰(如 MP2)或耦合簇(如 CCSD、CCST(T))方程,将总电子相关能分解为对若干子系统的计算。他们的 CIM-X 和 GEBF-X 方法都在其自主开发的低标度量子化学(Low Scaling Quantum Chemistry, LSQC)软件<sup>[13]</sup>中实现,需要的电子结构计算利用 GAUSSIAN, GAMESS 等软件。

南京大学马晶、张洪江等引入耦合电子对近似(CEPA),改进了局域收缩组态相互作用(LC-CI),特别是包含了 TQ 效应的 LC-CEPA-3 方法可以得到很精确的键解离势能面<sup>[14]</sup>。他们还在从头算水平实现了重整化激子方法(REM)<sup>[15]</sup>,为准确和高效地计算大体系的激发态提供了新的可能性。

西北大学文振翼等开发了一组计算电子相关能的程序包 Xi'an-CI,包含多参考组态相互作用(MR-CISD)<sup>[16]</sup>及其各种近似,以及多参考微扰理论(MRPT2)<sup>[17]</sup>。程序基于图形酉群方法(GUGA),并采用他们提出的空穴一粒子对称<sup>[18]</sup>来提高计算效率。程序部分功能包含在 MOLCAS 软件中。

## 1.3 密度泛函理论

徐昕与 Gorrdard 合作,通过结合 B88 和 PW91

交换泛函,开发了 X3LYP 杂化密度泛函方法<sup>[19]</sup>,能更好地描述弱相互作用,是当时最准确的 DFT 方法之一。徐昕、张颖等还继续在 2009 年提出了双杂化泛函 XYG3<sup>[20]</sup>,是首个依照绝热途径理论建立的双杂化泛函,预测生成焓和反应势垒的准确度明显好于 MP2 方法和一些常见的双杂化泛函,预测弱相互作用的精度也整体上较好。为了解决 XYG3 缺少长程色散力的 R<sup>-6</sup> 渐进行为,他们近期发展了 Irc-XYG3<sup>[21]</sup> 双杂化泛函。这些泛函已不同程度地在 NWChem, Q-Chem, Gaussian 等计算软件中实现。另外,徐昕、吴剑鸣等近十年来引入了人工神经网络算法对 B3LYP 的计算结果进行矫正(X1 系列<sup>[22, 23]</sup>),大大提升了其预测热化学数据(生成焓、键焓、异构化焓)的准确性。

## 1.4 价键理论计算

价键理论与化学家所熟悉的电子配对成键的概念相吻合,能给出直观的化学图像,一出现就得到迅速发展,然而从头算价键理论由于使用了非正交轨道导致计算困难,后来发展缓慢。厦门大学的张乾二、吴玮、宋凌春等长期致力于发展基于价键理论的从头算方法,早期编写了 VB-XIAMEN99 程序。吴玮等于 2003 年发布了 XMVB (XiaMenValence-Bond)程序的 1.0 版<sup>[24]</sup>,近年来又推出了 XMVB 2.0 版本<sup>[25]</sup>,今年 7 月份又推出了 3.0 版。XMVB 与 GAMESS 量化软件嫁接,能够执行价键理论计算方法如 VBSCF, BOVB, VB CI, VBPT2, DFVB, VB-PCM, VBEPF 等。XMVB 软件被用来研究一些重要的理论化学问题<sup>[26]</sup>。

## 1.5 相对论性量子力学计算

北京大学的刘文剑等早年就开发了能进行相对论性密度泛函理论计算的程序 BDF(Beijing Density Functional)<sup>[27, 28]</sup>,能实现四分量全相对论、二分量或一分量准相对论性、以及非相对论性 DFT 计算。他们又发展了精确二分量(X2C)<sup>[29]</sup>、准四分量(Q4C)<sup>[30]</sup>、自旋分离 X2C<sup>[31]</sup>等相对论性电子哈密顿量。他与合作者还提出了计算核磁共振性质相对论理论<sup>[32]</sup>,以及发展了高效的舍时密度泛函理论(TD-DFT)方法<sup>[33]</sup>等。这些不同方法均在具有自主知识产权的 BDF 软件包中实现,可以提供基准性计算结果以验证其他近似方法,其中部分模块具有很高的计算效率。

除 DFT 方法外,耦合簇(CC)等后自洽场方法也是电子相关计算的重要方法。四川大学王繁等多年来发展了带自旋-轨道耦合(SOC)的耦合簇方法

SOC-CC<sup>[34, 35]</sup>。此方法在 SCF 计算使用标量(实)分子自旋轨道,而在 CC 计算时才考虑 SOC 相对论性效应,可以大大提供计算效率。在此基础上采用运动方程(EOM)方法,可以较准确计算前五周期元素的激发能。此方法已与 BDF 程序嫁接。

### 1.6 量子动力学和非绝热动力学

在量子力学水平上描述原子核运动,可以作为更近似的基元化学反应理论方法(如经典过渡态理论)的参考基准,而且在很多情况下核量子效应是不可忽略的。另外,当不同电子态的势能面相互作用较强或发生交叉时,可能会发生非绝热动力学过程,即电子与核运动的耦合。大连化学物理研究所韩克利等发展了准经典轨线(QCT)方法,研究简单化学反应的立体动力学(stereo-dynamics, SD-QCT)。韩克利等还发展了利用含时波包方法<sup>[36]</sup>来计算非绝热量子散射动力学(QSD)的程序 NAC-TDWP。由于 QSD 的计算量很大,韩克利与张佩宇还开发了利用 GPU 计算的 GQSD 程序<sup>[37]</sup>,与 CPU 相比可以提速两个数量级。另外,刘剑(现北京大学)在美国时发展了 William Miller 课题组提出的线性化半经典初值表示理论(LSC-IVR)<sup>[38]</sup>,可以研究较复杂分子体系的量子动力学,此方法与路径积分分子动力学(PIMD)方法被植入 AMBER 软件中<sup>[39]</sup>。

北京师范大学的崔刚龙与 Walter Thiel 合作,发展了一套能同时模拟内转换(IC)和系间窜跃(ISC)过程的广义非绝热动力学的方法 GTSH<sup>[40]</sup>。这是唯一能实现 QM/MM 广义非绝热动力学(IC和ISC)的程序包。最近,崔刚龙、方维海等人,在前人的惩罚函数方法的框架下,发展了可以在 QM/MM 水平下优化三态势能面圆锥交叉点的方法<sup>[41]</sup>,在其 BNU-CI 程序包中实现,与 Molpro, MNDO99, 以及 ChemShell 程序包对接。另外,青岛生物能源与过程研究所兰峥岗等最近开发了基于在线(on-the-fly)电子结构计算的广义非绝热动力学模拟程序包 JADE<sup>[42]</sup>,与 Turbomole, GAUSSIAN, GAMESS (US)中的激发态量化方法对接。

### 1.7 量子化学分析软件

北京科技大学的卢天、陈飞武开发了波函数分析的 Multiwfn 程序<sup>[43]</sup>,集成了各种功能,包括各种布局分析、轨道分解、键级计算、静电势、电子密度、自旋密度分析等,较为用户友好,自带可视化模块的独立软件,且能与 VMD 分子可视化软件对接。同时,Multiwfn 可以使用多线程并行计算,效率较高。用户群涵盖国内外各地,被 SCI 引用 1900 余次。

## 2 分子力场与动力学模拟

分子力场是描述原子(或原子团)之间相互作用的解析势能函数,基于力场的模拟已经成为生物和材料研究中的重要工具。

### 2.1 全原子分子力场

北京大学的蒋帆、吴云东等近年来开发了一系列残基特异性力场(Residue-Specific Force Fields, RSFF),其中 RSFF1 力场<sup>[44]</sup>是 OPLS-AA/L 力场的改进,RSFF2 力场<sup>[45]</sup>是 Amber-99SB 力场的改进。他们利用蛋白晶体结构的统计结果优化了对骨架和侧链二面角的描述。新力场在不改变计算效率的情况下,能更准确地模拟多肽和蛋白质体系<sup>[46]</sup>。最近,在 RSFF2 力场的基础上做了进一步的改进(RSFF2+)<sup>[47]</sup>,可以同时很好地描述无序态(非折叠态)和折叠态。这一系列的 RSFF 力场借助 GROMACS 模拟软件实现。最近,使用 CMAP 势函数将此力场在 AMBER 软件中实现。

孙淮(现上海交通大学)早年在美国 *Molecular Simulations Inc.* 工作期间,开发了 COMPASS 力场<sup>[48]</sup>,被广泛地使用。回国后孙淮与合作者继续致力于通用力场的开发,特别是与材料相关的体系。近年来发展了一个 TEAM(Transferable, Extensible, Accurate and Module,可移植、可扩展、准确、模块化)力场框架<sup>[49]</sup>。在此框架下的某些力场在 GROMACS 软件中实现。孙淮等还与国外的公司合作,开发了 COMPASS II 通用力场<sup>[50]</sup>,可以用来模拟类药小分子和高分子等材料体系。此力场在 Forcite, Discover, Materials Studio 这些软件中实现。

### 2.2 粗粒化分子力场与动力学模拟

许多生物分子体系在行使功能时,涉及到跨越多个数量级的时间和空间尺度,常常超出当前的全原子分子动力学(MD)模拟的能力范围。各种不同的粗粒化(coarse-graining, CG)方法能大大提升计算效率,成为重要研究手段。韩伟、吴云东等,开发了与 Marrink 的粗粒化水模型结合使用的联合原子蛋白质力场 PACE<sup>[51]</sup>。由于生物分子体系的全原子 MD 模拟中主要的计算量通常都消耗在处理水分子上,使用粗粒化水模型可以大大提高模拟速度,同时又能保持对溶质的细节描述。

大连化学物理研究所的李国辉等,开发了新的粗粒化模型 GBEMP<sup>[52]</sup>。对多肽和蛋白体系,使用了各向异性的 Gay-Berne 粒子来代表若干(通常 4—

11)个原子的集合。与之前绝大部分CG模型不同,GBEMP对静电相互作用的描述更为精细,使用了包括四极矩在内的电多极矩(EMP)。去年,李国辉等又报道了将他们的GBEMP力场扩展到核酸(DNA, RNA)体系<sup>[53]</sup>。

对于凝聚态软物质体系的粗粒化模拟,耗散粒子动力学(DPD)相对于通常的MD有其独特的优势。北京师范大学高靛辉、方维海等最近开发了针对多种生物分子的CG-PDP力场<sup>[54, 55]</sup>,能成功模拟脂双层体系的结构特征,以及多肽与生物膜的相互作用。

近年来,GPU强大的计算能力被用来加速MD模拟,但绝大部分是基于全原子模型。吉林大学的吕中元、朱有亮(现长春应化所)等开发了利用GPU进行大尺度分子模拟的GALAMOST软件<sup>[56]</sup>。支持LJ势和查表函数的非键相互作用、软各向异性Janus粒子模型、杂化粒子一场模型、以及链增长聚合反应模型等,可以进行粗粒化MD模拟、布朗动力学、以及耗散粒子动力学模拟等。全部计算都在GPU上完成,效率高,且可以模拟数百万粒子的体系。

### 2.3 可极化力场与电荷模型

可极化力场考虑了分子内外的环境对分子电荷分布(如原子电荷)的影响,可看作是精细程度介于电子结构方法和经典分子力场方法之间的模型,有特殊的意义。华东师范大学的张增辉、季长鸽、梅晔等利用他们开发的线性标度量子化学计算方法MFCC(分子碎片共轭帽子)可以得到蛋白特异性极化电荷(PPC),利用AMBER软件可以进行MD模拟和自由能计算<sup>[57]</sup>,并且可以在模拟中根据情况改变原子电荷<sup>[58]</sup>。

辽宁师范大学的杨忠志、赵东霞、宫利东等,在其前期原子—键电负性均衡模型ABEEM的基础上,发展了浮动电荷的可极化力场ABEEM/MM,可以处理水分子<sup>[59]</sup>、有机分子、蛋白质、含金属离子<sup>[60]</sup>的体系等,还发展了更精细的ABEEM $\sigma\pi$ 可极化力场模型<sup>[61]</sup>。目前此力场的MD研究与Tinker程序嫁接。

### 2.4 增强构象抽样方法

增强抽样方法可以极大地扩展MD模拟所能研究的时间尺度,具有很重要的意义。北京大学的高毅勤、杨立江等开发了温度积分抽样(integrated tempering sampling, ITS)方法<sup>[62]</sup>,及其扩展选择性温度积分抽样(SITS)方法<sup>[63]</sup>。ITS/SITS方法不需

预先定义反应坐标,非常有效地提高了抽样方法在复杂体系中的适用性,可以用来快速扫描复杂体系的构象空间和计算多维自由能面。被用来研究从化学反应到蛋白质折叠的各种问题<sup>[64]</sup>。

## 3 药物与材料等相关应用软件

### 3.1 药物设计与分子对接软件

北京大学来鲁华等多年来开发了基于靶点三维结构的药物设计软件:LigBuilder<sup>[65, 66]</sup>,这是一个整合了多种方法(包括他们自己发展的一系列方法)的独立程序,可以实现基于片段生长和连接的配体分子设计,以及已知分子的优化,并可进行分子的可合成性分析及类药性分析。其中包含了来鲁华、裴剑锋、苑亚夏等开发的对生物分子表面可能的结合位点进行预测和可药性分析的独立程序Cavity<sup>[67]</sup>。此软件得到国内外的较广泛应用。来鲁华等还开发了进行蛋白质—配体以及蛋白—蛋白分子对接的程序PSI-DOCK<sup>[68]</sup>和SDOCK<sup>[69]</sup>,后者与国际上流行的ZDOCK相比,准确率相当或略优。此外,北京工业大学龚新奇、王存新等开发了预测蛋白—蛋白结合位点的新方法,并在此基础上发展了蛋白—蛋白对接的HoDock程序<sup>[70]</sup>,此方法在蛋白相互作用预测国际竞赛CAPRI中有不错的表现。

上海药物所郑明月、蒋华良等利用对打分函数的“知识基础”进行扩充,发展了新的蛋白—配体亲和力和打分函数IPMF<sup>[71]</sup>,与常见的分子对接打分函数相比,在亲和力预测方面有较明显改进。他们还发展了一种新型的二维氢键统计势,能显著提高蛋白—配体结合构象的识别能力<sup>[72]</sup>。针对水的重要作用,上海药物所熊兵等发展了可以对蛋白质结合水分子的亲和力进行预测的统计势方法wPMF<sup>[73]</sup>。针对卤键在药物设计研究中的重要性,上海药物所朱维良等还发展了卤键打分函数XBPMF<sup>[74]</sup>。

### 3.2 材料性质计算软件

有机光电材料可应用于有机显示与照明、有机场效应器件、有机太阳能电池、有机光检测、生物传感、有机光通讯等领域。帅志刚早期在比利时工作时就利用密度矩阵重正化群(DMRG)方法来对共轭聚合物进行激发态电子结构计算<sup>[75]</sup>,来得到线性和非线性光学(如双光子吸收)响应性质。回国后,帅志刚与彭谦、邵久书等在谐振子模型下发展了一套考虑Duschinsky转动效应的定量预测分子荧光量子效率的方法<sup>[76, 77]</sup>,能计算较大的分子。另外,帅志刚、易院平等发展了Correction Vector方法,与

MRDCI(多参考组态相互作用)<sup>[78]</sup>或 EOM-CCSD(运动方程耦合簇)<sup>[79]</sup>等方法结合,来计算较大分子的多光子吸收特性。基于他们发展的一系列方法,清华大学帅志刚与中科院化学所彭谦、耿华等开发了一套计算分子材料性质的软件包 MOMAP(Molecular Materials Property Prediction Package)。它包含光物理模块<sup>[80]</sup>和电子输运模块<sup>[81]</sup>,可以定量预测发光效率和发光光谱、以及载流子迁移率等。MOMAP 软件支持并行计算、效率高,计算所需的分子基态、激发态信息可利用 GAUSSIAN、TURBOMOLE 等成熟的量化软件得到。此软件已被成功地商业化,推出了友好的操作界面 Device Studio,但同时也提供免费试用。

对于 d/f 电子体系(如含过渡金属和镧系锕系元素的无机和金属有机材料)的第一性原理计算,北京大学的蒋鸿与德国 Fritz-Haber 研究所合作,开发了基于全势一缀加平面波基组的 GW 程序 GAP(GW with Augmented Planewaves)<sup>[82, 83]</sup>,使用 LDA/GGA+U 方法。此程序可以作为 WIEN2K 的插件使用。

### 3.3 材料晶体结构预测软件

获得晶体结构对于理解无机固体材料的性质和设计新材料有很重要的意义。吉林大学马琰铭等发展了晶体结构预测新技术,将全局最优化的粒子群算法应用到晶体结构预测领域,发展了拥有自主知识产权的晶体结构预测程序 CALYPSO<sup>[84, 85]</sup>,为 40 个国家近 1000 研究者使用。此软件当前与 VASP、CASTEP、Quantum Espresso、GULP、SIESTA、CP2K 程序都有接口。

复旦大学刘智攀等近年来发展了随机势能面行走(stochastic surface walking, SSW)方法<sup>[86]</sup>,可以跨越势垒进行全局搜索来预测晶体结构以及不同晶相的转变路径。最近,他们将 SSW 与神经网络算法,发展了一种能进行快速材料发现与筛选的新方法,并用它发现了两个新的 TiO<sub>2</sub> 多孔晶相<sup>[87]</sup>。此软件已经与 VASP, SIESTA 等常用软件连接,使用其中的电子结构计算方法。

## 4 总结与展望

开发新的计算方法是理论与计算化学学科发展的重要推动力。综上对国内一些相关工作的介绍,可见在各个重要的子领域都有我国各地的研究小组开发出各种计算方法和软件,其中不少都有很高的学术价值和实际应用价值,有些可以在一些关键指

标上超越国外同类方法或软件。然而,由于缺乏我国自己的计算化学软件平台,许多计算方法没有得到很好的实际应用。不少国内的新计算方法(包括低标度电子结构算法、密度泛函、分子力场、增强构象抽样方法等)主要都在国外的软件平台上实现。

近年来,我国在高性能计算的硬件上有很大的投入和很好的成果,包括目前世界上性能最强的两台超级计算机,特别是排名第一的“神威·太湖之光”采用了中国自主研发的处理器。然而,目前较缺乏与高性能计算机相配套的各种应用软件,对相关应用型高性能计算人才的培养也很不足。开发我国自主的计算化学软件平台,能推动高性能计算在我国新材料和新药开发中的基础和应用研究及人才培养。

另外,相对于实现计算化学方法的软件,在化学信息与图形处理等实用方面,我国的优秀软件也很少。例如,化学界广泛使用的 ChemDraw 和 ChemOffice 软件都是由外国公司开发。全世界的科学家最常使用的前三名生物分子结构可视化软件 PyMol, VMD, Chimera, 都是美国公司或大学的研究组开发的。

综上所述,我们需要开发自主知识产权的计算化学软件平台,其中能够整合各种国内外开发的先进理论与计算化学方法,并且能方便大家在此基础上开发新的计算方法。最终,此计算软件平台将包括各种与分子科学相关的计算方法,如电子结构计算、量子动力学、分子力场与分子动力学、材料微观结构与物理性质计算、药物开发相关的分子对接和分子设计方法、基于大数据技术的化学信息学方法、以及图形图像处理软件等。

计算化学软件平台的发展是一个长期的、需要多方合作的工作。高水平的软件开发需要稳定的有高度专业知识和技能的研发团队,包括理论化学家和计算机程序员。利用此平台,有些现有软件可以进一步优化以提高计算效率等,特别是充分利用现在最新的硬件和软件技术,包括 GPU 和我国自主研发的处理器等;有些软件可以改进其用户界面,提升易用性。这些重要的工作很难直接在高影响因子期刊上发表研究论文,相关开发人员在学术上发展空间可能受限,且由于计算机行业人才的现实情况,一般要求较高的待遇。这可能也是我国至今没有被广泛使用的自主计算化学软件平台的一个重要原因,因此很需要国家对基础计算软件平台开发提供专门支持。

**致谢** 本文在中国理论与计算化学发展战略研究组关于建立“计算化学软件平台专项”的建议,以及国家自然科学基金委员会第181期《理论化学家视角中的仪器创制》双清论坛的基础上撰写。

### 参 考 文 献

- [1] Li S, Li W, Fang T. An efficient fragment-based approach for predicting the ground-state energies and structures of large molecules. *J Am Chem Soc*, 2005, 127(19): 7215—7226.
- [2] Li W, Li S, Jiang Y. Generalized energy-based fragmentation approach for computing the ground-state energies and properties of large molecules. *J Phys Chem A*, 2007, 111(11): 2193—2199.
- [3] Li S, Li W, Ma J. Generalized energy-based fragmentation approach and its applications to macromolecules and molecular aggregates. *Acc Chem Res*, 2014, 47(9): 2712—2720.
- [4] Guo W, Wu A, Xu X. XO: An extended ONIOM method for accurate and efficient geometry optimization of large molecules. *Chem Phys Lett*, 2010, 498: 203—208.
- [5] Wu A, Xu X. DCMB that combines divide-and-conquer and mixed-basis set methods for accurate geometry optimizations, total energies, and vibrational frequencies of large molecules. *J Comput Chem*, 2012, 33(16): 1421—1432.
- [6] Peng L, Gu FL, Yang W. Effective preconditioning for ab initio ground state energy minimization with non-orthogonal localized molecular orbitals. *Phys Chem Chem Phys*, 2013, 15(37): 15518—15527.
- [7] Xiang H, Yang J, Hou JG, et al. Linear-scaling density matrix perturbation treatment of electric fields in solids. *Phys Rev Lett*, 2006, 97(26): 266402.
- [8] Shang H, Xiang H, Li Z, et al. Linear scaling electronic structure calculations with numerical atomic basis set. *Int Rev Phys Chem*, 2010, 29(4): 665—691.
- [9] Qin X, Shang H, Xiang H, et al. HONPAS: A linear scaling open-source solution for large system simulations. *Int J Quantum Chem*, 2015, 115(10): 647—655.
- [10] Li S, Ma J, Jiang Y. Linear scaling local correlation approach for solving the coupled cluster equations of large systems. *J Comput Chem*, 2002, 23(2): 237—244.
- [11] Li S, Shen J, Li W, et al. An efficient implementation of the “cluster-in-molecule” approach for local electron correlation calculations. *J Chem Phys*, 2006, 125(7): 74109.
- [12] Li W, Piecuch P, Gour JR, et al. Local correlation calculations using standard and renormalized coupled-cluster approaches. *J Chem Phys*, 2009, 131(11): 114109.
- [13] Li W, Chen C, Zhao D, et al. LSQC: Low scaling quantum chemistry program. *Int J Quantum Chem*, 2015, 115(10): 641—646.
- [14] Zhang H, Malrieu JP, Reinhardt P, et al. Improved version of a local contracted configuration interaction of singles and doubles with partial inclusion of triples and quadruples. *J Chem Phys*, 2010, 132(3): 034108.
- [15] Zhang H, Malrieu J, Ma H, et al. Implementation of renormalized excitonic method at ab initio level. *J Comput Chem*, 2012, 33(1): 34—43.
- [16] Suo B, Zhai G, Wang Y, et al. Parallelization of MRCI based on hole-particle symmetry. *J Comput Chem*, 2005, 26(1): 88—96.
- [17] Lei Y, Wang Y, Han H, et al. New implementation of the configuration-based multi-reference second order perturbation theory. *J Chem Phys*, 2012, 137(14): 144102.
- [18] Wang Y, Zhai G, Suo B, et al. Hole-particle correspondence in CI calculations. *Chem Phys Lett*, 2003, 375: 134—140.
- [19] Xu X, Goddard WA. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proc Natl Acad Sci USA*, 2004, 101(9): 2673—2677.
- [20] Zhang Y, Xu X, Goddard WA. Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics. *Proc Natl Acad Sci USA*, 2009, 106(13): 4963—4968.
- [21] Zhang IY, Xu X. Reaching a uniform accuracy for complex molecular systems: long-range-corrected XYG3 doubly hybrid density functional. *J Phys Chem Lett*, 2013, 4(10): 1669—1675.
- [22] Wu J, Xu X. The X1 method for accurate and efficient prediction of heats of formation. *J Chem Phys*, 2007, 127(21): 214105.
- [23] Wu J, Zhou Y, Xu X. The X1 family of methods that combines B3LYP with neural network corrections for an accurate yet efficient prediction of thermochemistry. *Int J Quantum Chem*, 2015, 115(16): 1021—1031.
- [24] Song L, Mo Y, Zhang Q, et al. XMVB: a program for ab initio nonorthogonal valence bond computations. *J Comput Chem*, 2005, 26(5): 514—521.
- [25] Chen Z, Ying F, Chen X, et al. XMVB 2.0: A new version of Xiamen Valence Bond program. *Int J Quantum Chem*, 2015, 115(11): 731—737.
- [26] Mo Y. Computational evidence that hyperconjugative interactions are not responsible for the anomeric effect. *Nat Chem*, 2010, 2(8): 666—671.
- [27] Liu W, Hong G, Dai D, et al. The Beijing four-component density functional program package (BDF) and its application to EuO, EuS, YbO and YbS. *Theor Chem Acc*, 1997, 96(2): 75—83.
- [28] Liu W, Wang F, Li L. The Beijing Density Functional (BDF) program package: methodologies and applications. *J Theor Comput Chem*, 2003, 02(02): 257—272.
- [29] Liu W, Peng D. Exact two-component Hamiltonians revisited. *J Chem Phys*, 2009, 131(3): 31104.
- [30] Liu W, Peng D. Infinite-order quasirelativistic density functional method based on the exact matrix quasirelativistic theory. *J Chem Phys*, 2006, 125(4): 44102.
- [31] Li Z, Xiao Y, Liu W. On the spin separation of algebraic two-component relativistic Hamiltonians. *J Chem Phys*, 2012, 137(15): 154114.
- [32] Sun Q, Liu W, Xiao Y, et al. Exact two-component relativistic theory for nuclear magnetic resonance parameters. *J Chem Phys*, 2009, 131(8): 81101.
- [33] Wu F, Liu W, Zhang Y, et al. Linear-scaling time-dependent density functional theory based on the idea of “from fragments to molecule”. *J Chem Theory Comput*, 2011, 7(11): 3643—3660.
- [34] Wang F, Gauss J, Van Wullen C. Closed-shell coupled-cluster theory with spin-orbit coupling. *J Chem Phys*, 2008, 129(6): 64113.
- [35] Cao Z, Wang F, Yang M. Coupled-cluster method for open-shell heavy-element systems with spin-orbit coupling. *J Chem Phys*, 2017, 146(13): 134108.

- [36] Chu T, Zhang Y, Han K. The time-dependent quantum wave packet approach to the electronically nonadiabatic processes in chemical reactions. *Int Rev Phys Chem*, 2006, 25: 201—235.
- [37] Zhang P, Han K. GQSD: The program for the graphic processing units accelerated quantum scattering dynamics. *Int J Quantum Chem*, 2015, 115(11): 738—743.
- [38] Liu J, Miller WH. Using the thermal Gaussian approximation for the Boltzmann operator in semiclassical initial value time correlation functions. *J Chem Phys*, 2006, 125(22): 224104.
- [39] Liu J, Miller WH, Paesani F, et al. Quantum dynamical effects in liquid water: A semiclassical study on the diffusion and the infrared absorption spectrum. *J Chem Phys*, 2009, 131(16): 164509.
- [40] Cui G, Thiel W. Generalized trajectory surface-hopping method for internal conversion and intersystem crossing. *J Chem Phys*, 2014, 141(12): 124101.
- [41] Liu X, Cui G, Fang W. Three-state conical intersection optimization methods: development and implementation at QM/MM level. *Theor Chem Acc*, 2017, 136(1): 8.
- [42] Du L, Lan Z. An On-the-Fly Surface-hopping program jade for nonadiabatic molecular dynamics of polyatomic systems: implementation and applications. *J Chem Theory Comput*, 2015, 11(4): 4522—4523.
- [43] Lu T, Chen F. Multiwfn: A multifunctional wavefunction analyzer. *J Comput Chem*, 2012, 33(5): 580—592.
- [44] Jiang F, Zhou C, Wu Y. Residue-specific force field based on the protein coil library. RSFF1: modification of OPLS-AA/L. *J Phys Chem B*, 2014, 118(25): 6983—6998.
- [45] Zhou CY, Jiang F, Wu YD. Residue-specific force field based on protein coil library. RSFF2: modification of AMBER ff99SB. *J Phys Chem B*, 2015, 119(3): 1035—1047.
- [46] Li S, Elcock AH. Residue-specific force field (RSFF2) improves the modeling of conformational behavior of peptides and proteins. *J Phys Chem Lett*, 2015, 6(11): 2127—2133.
- [47] Wu H, Jiang F, Wu Y. Significantly improved protein folding thermodynamics using a dispersion-corrected water model and a new residue-specific force field. *J Phys Chem Lett*, 2017, 8(14): 3199—3205.
- [48] Sun H. COMPASS: An ab initio force-field optimized for condensed-phase applications: overview with details on alkane and benzene compounds. *J Phys Chem B*, 1998, 102(38): 7338—7364.
- [49] Li X, Li F, Shi Y, et al. Predicting water uptake in poly(perfluorosulfonic acids) using force field simulation methods. *Phys Chem Chem Phys*, 2010, 12(43): 14543—14552.
- [50] Sun H, Jin Z, Yang C, et al. COMPASS II: extended coverage for polymer and drug-like molecule databases. *J Mol Model*, 2016, 22(2): 47.
- [51] Han W, Wan C, Jiang F, et al. PACE force field for protein simulations. 1. full parameterization of version 1 and verification. *J Chem Theory Comput*, 2010, 6(11): 3373—3389.
- [52] Shen H, Li Y, Ren P, et al. Anisotropic coarse-grained model for proteins based on gay-berne and electric multipole potentials. *J Chem Theory Comput*, 2014, 10(2): 731—750.
- [53] Li G, Shen H, Zhang D, et al. Coarse-grained modeling of nucleic acids using anisotropic gay-berne and electric multipole potentials. *J Chem Theory Comput*, 2016, 12(2): 676—693.
- [54] Chen L, Li X, Gao L, et al. Theoretical insight into the relationship between the structures of antimicrobial peptides and their actions on bacterial membranes. *J Phys Chem B*, 2015, 119(3): 850—860.
- [55] Li X, Gao L, Fang W. Dissipative particle dynamics simulations for phospholipid membranes based on a four-to-one coarse-grained mapping scheme. *PLoS One*, 2016, 11(5).
- [56] Zhu Y, Liu H, Li Z, et al. GALAMOST: GPU-accelerated large-scale molecular simulation toolkit. *J Comput Chem*, 2013, 34(25): 2197—2211.
- [57] Ji C, Mei Y, Zhang JZH. Developing polarized protein-specific charges for protein dynamics: MD free energy calculation of pKa shifts for Asp26/Asp20 in Thioredoxin. *Biophys J*, 2008, 95(3): 1080—1088.
- [58] Duan LL, Mei Y, Zhang D, et al. Folding of a helix at room temperature is critically aided by electrostatic polarization of intraprotein hydrogen bonds. *J Am Chem Soc*, 2010, 132(32): 11159—11164.
- [59] Yang Z, Wu Y, Zhao D. Atom-bond electronegativity equalization method fused into molecular mechanics. I. A seven-site fluctuating charge and flexible body water potential function for water clusters. *J Chem Phys*, 2004, 120(6): 2541—2557.
- [60] Gong L. Development and applications of the ABEEM fluctuating charge molecular force field in the ion-containing systems. *Sci China Chem*, 2012, 55(12): 2471—2484.
- [61] Zhao D, Liu C, Wang F, et al. Development of a polarizable force field using multiple fluctuating charges per atom. *J Chem Theory Comput*, 2010, 6(3): 795—804.
- [62] Gao YQ. An integrate-over-temperature approach for enhanced sampling. *J Chem Phys*, 2008, 128(6): 64105.
- [63] Yang L, Gao YQ. A selective integrated tempering method. *J Chem Phys*, 2009, 131(21): 214109.
- [64] Yang L, Liu C, Shao Q, et al. From thermodynamics to kinetics: enhanced sampling of rare events. *Acc Chem Res*, 2015, 48(4): 947—955.
- [65] Wang R, Gao Y, Lai L. LigBuilder: A multi-purpose program for structure-based drug design. *J Mol Model*, 2000, 6: 498.
- [66] Yuan Y, Pei J, Lai L. LigBuilder 2: A practical de novo drug design approach. *J Chem Inf Model*, 2011, 51(5): 1083—1091.
- [67] Yuan Y, Pei J, Lai L. Binding site detection and druggability prediction of protein targets for structure-based drug design. *Curr Pharm Design*, 2013, 19(12): 2326—2333.
- [68] Pei J, Wang Q, Liu Z, et al. PSI-DOCK: towards highly efficient and accurate flexible ligand docking. *Proteins*, 2006, 62(4): 934—946.
- [69] Zhang C, Lai L. SDOCK: A global protein-protein docking program using stepwise force-field potentials. *J Comput Chem*, 2011, 32(12): 2598—2612.

- [70] Gong X, Wang P, Yang F, et al. Protein-protein docking with binding site patch prediction and network-based terms enhanced combinatorial scoring. *Proteins*, 2010, 78(15): 3150—3155.
- [71] Shen Q, Xiong B, Zheng M, et al. Knowledge-based scoring functions in drug design: 2. Can the knowledge base be enriched? *J Chem Inf Model*, 2011, 51(2): 386—397.
- [72] Zheng M, Xiong B, Luo C, et al. Knowledge-based scoring functions in drug design: 3. A two-dimensional knowledge-based hydrogen-bonding potential for the prediction of protein-ligand interactions. *J Chem Inf Model*, 2011, 51(11): 2994—3004.
- [73] Zheng M, Li Y, Xiong B, et al. Water PMF for predicting the properties of water molecules in protein binding site. *J Comput Chem*, 2013, 34(7): 583—592.
- [74] Liu Y, Xu Z, Yang Z, et al. A knowledge-based halogen bonding scoring function for predicting protein-ligand interactions. *J Mol Model*, 2013, 19(11): 5015—5030.
- [75] Shuai Z, Bredas J, Saxena A, et al. Linear and nonlinear optical response of polyenes: A density matrix renormalization group study. *J Chem Phys*, 1998, 109(6): 2549—2555.
- [76] Peng Q, Yi Y, Shuai Z, et al. Towards quantitative prediction of molecular fluorescence quantum efficiency: role of Duschinsky rotation. *J Am Chem Soc*, 2007, 129(30): 9333—9339.
- [77] Peng Q, Yi Y, Shuai Z, et al. Excited state radiationless decay process with Duschinsky rotation effect: formalism and implementation. *J Chem Phys*, 2007, 126(11): 114302.
- [78] Yi Y, Zhu L, Shuai Z. The correction vector method for three-photon absorption: The effects of  $\pi$  conjugation in extended rylenebis (dicarboximide) s. *J Chem Phys*, 2006, 125(16): 164505.
- [79] Yi Y, Li Q, Zhu L, et al. Effects of donor/acceptor strengths on the multiphoton absorption: an EOM-CCSD correction vector study. *J Phys Chem A*, 2007, 111(38): 9291—9298.
- [80] Shuai Z, Wang D, Peng Q, et al. Computational evaluation of optoelectronic properties for organic/carbon materials. *Acc Chem Res*, 2014, 47(11): 3301—3309.
- [81] Wang L, Nan G, Yang X, et al. Computational methods for design of organic materials with high charge mobility. *Chem-Soc Rev*, 2010, 39(2): 423—434.
- [82] Jiang H, Gomezabal RI, Rinke P, et al. First-principles modeling of localized d states with the GW@LDA+U approach. *Phys Rev B*, 2010, 82(4): 45101—45108.
- [83] Jiang H, Gomezabal RI, Li X, et al. FHI-gap: A GW code based on the all-electron augmented plane wave method. *Comput Phys Commun*, 2013, 184(2): 348—366.
- [84] Wang Y, Lv J, Zhu L, et al. Crystal structure prediction via particle-swarm optimization. *Phys Rev B*, 2010, 82(9): 94116.
- [85] Wang Y, Lv J, Zhu L, et al. CALYPSO: a method for crystal structure prediction. *Comput Phys Commun*, 2012, 183(10): 2063—2070.
- [86] Shang C, Zhang X, Liu Z. Stochastic surface walking method for crystal structure and phase transition pathway prediction. *Phys Chem Chem Phys*, 2014, 16(33): 17845—17856.
- [87] Huang S, Shang C, Zhang X, et al. Material discovery by combining stochastic surface walking global optimization with a neural network. *Chem Sci*, 2017, 8(9): 6327—6337.

## Developments of computational chemistry methods and software in China

Jiang Fan      Wu Yundong

(Peking University Shenzhen Graduate School, Shenzhen 518055)

**Abstract** Currently, theoretical and computational chemistry is progressing rapidly, in which the development of new methods and software has a fundamental role, similar to the development of advanced scientific instruments in experimental sciences. In recent years, dozens of research groups in China have developed various new computational methods in electronic structure calculations, dynamic simulations, molecular force fields, drug design, and prediction of material structures and properties. Some of them are comparable or better than state-of-the-art methods, in certain aspects. Many of them also have great practical values. On the other hand, most of them rely on foreign software as a foundation or platform. Developing computational chemistry software platforms with independent intellectual properties will enhance the influence of these new methods and promote their widespread applications.

**Key words** quantum chemistry; molecular simulation; computational methods; software development