

·“双清论坛”专题：理论化学家视角中的仪器创制·

## 理论化学家视角中的仪器创制

——第181期双清论坛学术综述

徐 昕<sup>1</sup>      吴云东<sup>2</sup>      方维海<sup>3</sup>      帅志刚<sup>4</sup>  
高飞雪<sup>5\*</sup>      张守著<sup>6</sup>      孟庆峰<sup>7</sup>      雷惊雷<sup>5,8</sup>

(1. 复旦大学化学系, 上海 200433; 2. 北京大学深圳研究院, 深圳 518055;

3. 北京师范大学化学学院, 北京 100875; 4. 清华大学化学系, 北京 100084;

5. 国家自然科学基金委员会化学科学部, 北京 100085;

6. 国家自然科学基金委员会数学物理科学部, 北京 100085;

7. 国家自然科学基金委员会政策局, 北京 100085; 8. 重庆大学化学化工学院, 重庆 400044)

**[摘要]** 回顾历史, 科学史上很多重要仪器的发明, 都源于理论上新概念和新原理的提出。新概念和新原理可以催生具有原创思想的科学仪器的创制, 最终促进我国基础科学研究创新能力的整体提升。当前理论与计算化学正处于蓬勃发展阶段, 我国近年来在该学科方向发展迅速, 已形成了强大的研究团队, 积累了良好的研究基础。计算方法的发展是理论化学发展的重要部分, 相关计算程序是理论方法得以付诸应用的桥梁, 而众多理论研究又是建立在大量具体计算的基础上, 所以计算化学软件的发展是理论与计算化学的重要前沿。然而, 我国目前主要使用国外商业化软件, 自主研发的很少。这不利于提升我国理论与计算化学的国际影响力, 迫切需要弥补。为此, 2017年6月11—13日, 化学科学部、数学物理科学部、信息科学部与政策局联合举办主题为“理论化学家视角中的仪器创制”的第181期双清论坛。专家们通过充分而深入的研讨, 形成了一个重要的共识: 即缺乏具有自主知识产权的通用计算化学软件平台是制约中国理论化学发展的瓶颈。因而, 基于新理论、新方法和新算法, 自主系统开发、推进和创新计算化学软件是中国理论与计算化学学科今后能做出世界顶级贡献的必要条件。这和实验科学中基于新原理、新方法和新技术的自主研发大型仪器设备一样, 需要长期和稳定的支持。

**[关键词]** 仪器创制; 软件; 计算化学; 理论化学

计算化学软件不仅是理论化学研究者的主要工具, 也越来越成为化学实验研究组常备、甚至必备的有力工具。高效成熟的计算化学程序, 是研究人员运用成熟理论方法理解实验现象和探索微观机理的重要手段; 而先进的计算化学程序是新理论、新方法得以付诸应用的桥梁, 是进一步理论创新的技术支撑。

目前中国理论化学研究在诸多领域已经处于世界前列, 但主流计算化学程序却无一由中国发展。

究其原因, 功能完善、架构先进和高效、用户友好的大型计算程序的发展、维护和升级需要长期的积累, 而现有的基金资助模式, 在相当程度上都不利于计算化学软件的发展。由国家资助机构主导和支持, 国内众多单位通过协作进行自主知识产权的计算化学软件研发是当下较好的选择。

与会专家围绕论坛主题作报告或发言, 深入探讨了中国理论与计算化学所面临的机遇与挑战, 集中讨论了理论化学自身发展需重点关注和优先解决

的核心科学与技术问题;并进一步思考理论化学发展如何与实验发展互利共生、相互促进,如何对大型创新仪器设备的研制提供原创性的思想和技术支持;系统探讨了计算化学软件的建设发展模式、时间规划和管理制度,研讨了今后5—10年中国计算化学软件建设的目标。

现代科技的重大突破越来越依靠科学仪器的进步。真正有望做出世界领先水平的实验科学研究工作,必须依靠有自己特点的先进科学仪器,特别需要那些基于新原理、新方法和新技术的自主研发仪器。本论坛的主旨之一是发展新的理论化学方法,为原创科学仪器的创制提供理论基础和思想;同时,与会专家一致认同,理论化学家手中的计算机软件即等同于实验化学家手中的科学仪器,因而理论化学家视角中的“仪器创制”即等同于计算机软件的开发。杨学明院士在大会报告中提出了“软仪器”的概念。他认为“创新科学仪器”是科学发展的原动力,与实验科学使用的硬件仪器相对应,计算机软件是理论化学研究工作者的“软仪器”,其开发同样应该积极倡导和大力扶持。有专家建议,我们应当将“基于新理论、新方法和新算法的自主研发软件”提高到与“基于新原理、新方法和新技术的自主研发仪器”同等重要的地位予以重视。

会议重点就通用计算化学软件平台的搭建,融合我国具有特色功能的计算方法,面向国家重点需求和国际前沿领域,发展创新型理论化学方法,计算化学软件开发人才的培养和稳定支持等议题进行了深入讨论。

## 1 理论与计算化学在化学科学中的重要性不断提升

作为对人类生活和发展具有重要意义的中心学科,化学是由纯实验科学发展而来。随着20世纪初,量子力学的诞生与发展,使得我们得以真正从微观视角探索化合物的结构、性质与反应机理。结合量子力学、平衡与非平衡态热力学、统计力学和分子动力学的理论成果,化学理论从实验归纳总结为主,上升到有能力进行未知实验的理论模拟与预测。比如1929年Dirac说到“大部分物理学和全部化学的基本规律已经完全知道了,困难只是在于运用这些规律得到的数学方程太复杂,无法求解”<sup>[1]</sup>。借助数十年来计算机技术的高速发展,化学家已逐步能够运用上述理论定性和定量处理相当复杂的化学问题,对化学实验现象给予科学解读,甚至模拟和预测

目前实验技术还无法涉及或探测的化学反应,使得理论与计算化学成为化学学科不可缺少的重要组成部分。

化学学科已经发展为依靠实验—计算—理论三根支柱推动的科学,理论计算在理论模型与实验结果之间架起桥梁,是两者之间联系的纽带。吴云东院士在其大会引导性报告中指出,国际上,理论与计算化学学科也越来越得到重视。C&E News称美国从事理论与计算化学的研究人员占总体人数的10%—15%。其中哥伦比亚大学化学系23位教授中有5位从事理论与计算化学研究(22%)。美国科学院的约180位从事化学相关的院士中,理论与计算化学家有近30位(16%)。美国化学会会志(JACS)2010—2016涉及理论与计算的论文在25%以上。计算模拟在实际化学研究中已经成为强有力的工具,与实验研究配合互动,触及化学、生命科学、材料等各个分子科学方面。计算对象逐渐逼近复杂的真实化学体系和其中发生的过程。计算精度不断提高,对有些体系已有非常准确的定量预测能力。计算复杂度从简单基元过程扩展到非平衡过程和多种(可能是非线性的)过程耦合。理论计算化学已经成为了数学、物理、化学、药物、生命科学和信息科学的交叉点。

## 2 我国理论与计算化学领域的现状与国际地位

我国理论与计算化学起步较晚,20世纪50年代初唐敖庆、徐光宪先生等回国后才开始逐步开展量子化学等研究,培养了一批理论化学骨干人才。改革开放后计算机的逐步引进,才开始较多地开展对化学问题的计算模拟研究。近二十年来,随着我国经济实力的不断增强,计算机的普及与巨大发展,计算化学得到了迅猛的发展。目前已有几百个科研小组从事理论与计算方法的发展与应用,对化学的贡献与国际现状一致,约为15%左右。中国化学会下设25个专业委员会中就有理论化学、计算机化学两个理论计算化学领域的专业委员会,其它委员会如化学动力学、物理有机化学、晶体化学、催化等领域都不可或缺化学理论与计算。2008年第十届全国量子化学大会会有600多人参加,2011年有1200多人,2014年有1400多人,而刚刚结束的2017年全国量子化学大会参会人员达到了1700多人。

此外,目前我国有几十个研究组在从事各种计

算化学方法的发展研究<sup>[2]</sup>。已有很多小组将自己发展的方法编写成计算模块,甚至有独立完整的计算程序,如电子结构计算方面有 BDF<sup>[3]</sup>、XMVB<sup>[4]</sup>、LSQC<sup>[5]</sup>、HONPAS<sup>[6]</sup>、xDH(XYG3 系列泛函<sup>[7]</sup>)、XO<sup>[8]</sup>和 X1<sup>[9]</sup>系列、CIM<sup>[10]</sup>等等,量子动力学软件有 GQSD<sup>[11]</sup>、LSC-IVR<sup>[12]</sup>等等,力场与经典动力学有 ABEEM 与<sup>[13]</sup>、GALAMOST<sup>[14]</sup>、GBEMP<sup>[15]</sup>、RS-FF<sup>[16]</sup>等等,分子结构优化软件 SSW<sup>[17]</sup>,分子材料性能预测软件 MOMAP<sup>[18]</sup>,以及药物设计软件 Lig-Builder<sup>[19]</sup>、DOX<sup>[20]</sup>等。同时有一批非常优秀的中青年学者,在国际上已崭露头角。已有国际量子分子科学院院士四名,每三年一届的国际量子化学大会的大会报告人最早从第十二届(2006 年)的 2 人次,增加到目前的总共 13 人次。2015 年第十五届国际量子化学大会在北京成功召开,标志着我国在这一学科的发展水平已经得到国际同行的认可。

### 3 计算软件是理论化学研究工作者的“软仪器”

“工欲善其事,必先利其器”,现代科技的重大突破越来越依靠科研仪器的进步。大型科学仪器设备的创新对科学研究的创新性发展的作用不言而喻。高端科学仪器的创制能力是反映一个国家科学技术水平的重要标志之一。真正有望做到世界领先水平的实验科学研究工作,必须依靠有自己特点的先进科研仪器,特别是需要那些运用到许多新原理、新方法和新技术的自主研发仪器。欧美日等发达国家长期以来始终把“发展一流的科学仪器、支撑一流的科研工作”作为国家战略重点发展。近年来仪器及关键技术创新也逐渐成为我国关注的科研热点,而且《国家中长期科学和技术发展规划纲要(2006—2020 年)》已明确将科学仪器创制列为优先发展的战略领域。国家自然科学基金委员会从 2011 年起,设立“国家重大科研仪器研制项目”,着重面向科学前沿和国家需求,以科学目标为导向,鼓励和培育具有原创性思想的探索性科研仪器研制,着力支持通过关键核心技术突破或集成创新,用于发现新现象、揭示新规律、验证新原理、获取新数据的原创性重大科研仪器设备研制,以为科学研究提供更新颖的手段和工具,全面提升我国的原始创新能力。

目前理论计算化学已在药物、生命、催化以及材料等重要科学领域发挥着重要的作用,而理论计算化学工具的不断发展,是计算化学得以在各个科学领域广泛使用的关键所在。正如科学仪器是利用物

理原理来表征和探索具体物理和化学体系的工具一样,计算化学软件作为理论和计算方法付诸实践的重要手段,是连接理论方法与化学实验的纽带。例如:Walter Kohn 在 20 世纪 60 年代初提出了密度泛函理论(DFT)<sup>[21,22]</sup>,直到 20—30 年后以 B3LYP<sup>[23-26]</sup>和 PBE<sup>[27]</sup>等为代表的 DFT 方法被整合到计算化学软件后才获得广泛应用,随之该领域也获得飞速发展,并成为了当代应用最广,也是唯一对复杂体系实际可用的第一性原理方法。Walter Kohn 和发展 GAUSSIAN<sup>[28]</sup>(高斯)计算软件的 John Pople 一起获得了 1998 年的诺贝尔化学奖。1998 年的诺贝尔化学奖颁奖公报说<sup>[29]</sup>“90 年代快结束的时候,我们看到,化学理论和计算的研究有了很大的进展,其结果使整个化学正在经历着一场革命性变化”。此外,2013 年的诺贝尔化学奖颁给了 Martin Karplus, Michael Levitt 和 Arieh Warshel,以表彰他们为复杂化学系统的计算机模拟,创立了多尺度模型。通告中说<sup>[30]</sup>“电脑和试管对化学家一样重要”,“通过模拟,化学家能获得比传统实验更快速、更精准的预测结果”。

不仅如此,理论发展与实验发展是互利共生、相互促进的,特别是新理论(尤其是量子化学)方法以及新实验方法的深入结合和广泛运用,极大地推动了化学反应动力学等化学前沿学科研究的发展;理论发展也将对大型创新仪器设备的研制提供原创性的思想和技术支撑。事实上,实验与理论计算紧密配合互动已成为学科发展新阶段的工作新常态,而突破量子化学计算的基本矛盾,即计算量和精度难以兼顾的问题,开发兼顾高效率和高精度的新一代量子化学软件是促进理论与实验完美结合的必须,也应是未来理论与计算化学自身发展的核心攻坚任务。

### 4 目前国际上计算化学软件发展现状

目前被维基百科载入的有 80 余种量子化学与固体物理计算软件<sup>[31]</sup>。其中源于美国的 GAUSSIAN<sup>[28]</sup>和源于欧洲的 VASP<sup>[32]</sup>,它们几乎囊括了大多数的用户群体;其次有来自美国的 NWCHEM<sup>[33]</sup>、GAMESS<sup>[34]</sup>、Q-CHEM<sup>[35]</sup>等,以及来自欧洲的 MOLPRO<sup>[36]</sup>、TURBOMOL<sup>[37]</sup>、ADF<sup>[38]</sup>等为代表的具有相当用户群的软件;还有一些国内外各研究小组自主开发的理论化学计算程序。当然还有很多由于知名度不高,用户群体较少而未被维基百科列入其中的<sup>[31]</sup>。目前主流的具有

相当用户群体的通用或多功能计算化学软件包均由欧美国家资助或商业公司所有。

美国杜克大学杨伟涛教授在其大会报告中详细分析了目前国际上计算化学软件发展现状。GAUSSIAN<sup>[28]</sup>是美国 Gaussian Inc 的商业化计算化学软件包,侧重于簇模型体系的基于分子轨道理论的计算,在化学界有广泛的影响力,是目前国内外最为通用的计算化学软件包。从 70 版开始至今,GAUSSIAN 得到了长期的稳定发展,最早主要由美国 NSF 资助,于 1987 年商业化。其从 G70、G76、G80、G82、G86、G88、G90、G92、G92/DFT、G94、G98、G03、G09 到目前最新的 GAUSSIAN 16 共发布了 14 个版本。得益于众多国际主流前沿理论科学家的倾力支持与贡献,及其先发优势与易用性,GAUSSIAN 一直是国际上,同时也是国内最为通用的计算化学软件包。它包括了自洽场、后自洽场和 DFT 等在内的主流第一性原理电子结构理论方法、多种半经验方法、分子力学(MM)等经验方法以及他们的组合方法;可以处理气相、溶液和周期性边界(PBC)等体系;可进行体系构型与性质预测、轨道分析、各种光谱分析、分子动力学模拟等等。它不仅是理论化学研究工作者的的重要研究平台,也是实验科学家最常用的理论预测工具。GAUSSIAN 早期即使在软件商业化后,其软件的源代码仍是公开发售的。但近 20 年来由于贸易与技术壁垒,以及其公司自身政策等原因,GAUSSIAN 的源代码一直禁止向包括我国研究人员在内的科研人员开放。此外,GAUSSIAN 商业化过程中还将包括其创始人 John Pople 等在内的一些著名理论化学家列入到禁用该软件的名单中。John Pople 等由此于 20 世纪 90 年代另外开发了 Q-CHEM<sup>[35]</sup> 软件,其主要功能与 GAUSSIAN 接近,但在并行算法和微扰理论加速等多方面更胜一筹,目前也已成功商业化运作。

GAMESS<sup>[34]</sup>始于 1977 年,是较早与 GAUSSIAN 一同发展起来的通用计算化学软件,其全称为 General Atomic and Molecular Electronic Structure System。1981 年分出 US 和 UK 两个版本,目前 US 版由美国爱荷华大学的 Mark Gordon 量子理论研究组维护,是 1977 年起由美国能源部(DOE)和美国国家自然科学基金(NSF)共同资助建立的国家计算化学资源的项目之一。

为了适应快速发展的超级计算机,以充分利用其并行运算能力,美国以国家资助的形式,由西北太平洋国家实验室(Pacific Northwestern National

Laboratory)开发 NWCHEM<sup>[33]</sup> 软件。其跨生物分子、纳米结构和固体体系,从量子到经典或者组合法,可以处理基态和激发态,可用高斯基组和平面波基组,可计算各种性质包括相对论效应,更重要的是即可以采用单线程计算,也可以同时高效利用数千个 CPU 核心进行并行运算以处理复杂体系。为此,NWCHEM 往往只针对美国国内的超算中心进行优化,在不同的超算平台的计算效率存在较大差异。由于是国家持续资助,NWCHEM 没有商业化,在 ECL(Educational Community License)2.0 的协议下实现源代码开放。

来自于欧洲的 Molpro<sup>[36]</sup>是一个优秀的专业级电子结构量化软件,由德国斯图加特大学 H.-J. Werner 教授和英国加的夫大学 P. J. Knowles 教授等研究组设计和共同维护。其功能强大,使用灵活,尤其擅长各种后自洽场的理论计算,是专业级理论计算研究人员的有力工具。但用户入门门槛较高,较不适合初级用户或者广大实验组用户。最近, Molpro 开始控制其程序源码的对外销售。

Turbomole<sup>[37]</sup>是最初由德国卡尔斯鲁厄大学的 Reinhart Ahlrichs 教授开发,并于 2007 年开始商业化运作。正如其创始人所说的它更注重程序的高效,以便能够处理数百个原子及以上的第一性原理计算,因此它在 DFT 和微扰算法的加速以及并行化效率具有优势,其图形界面功能强大,更贴近于应用型的计算化学软件使用者。Turbomole 不开放源码。

ADF<sup>[38]</sup>(Amsterdam Density Functional)于 20 世纪 70 年代由荷兰阿姆斯特丹自由大学的 Evert Jan Baerends 教授和加拿大卡尔加里大学 Tom Ziegler 教授等开发,是少有的基于 STO 型基组的量化程序,特别擅长处理光谱、过渡金属及重元素的相对论问题。ADF 于 1995 年商业化,由 Baerends 研究组孵化出来的 Scientific Computing & Modeling (SCM)公司运营。2004 年后,算法开始超出 DFT 算法,向更大、更复杂体系发展,提供图形界面与半经验的 MOPAC<sup>[39]</sup>、DFTB<sup>[40]</sup> 以及反应力场(RelaxFF<sup>[41]</sup>)相结合,可以处理周期性体系和溶剂化效应。

VASP<sup>[32]</sup>(Vienna Ab initio Simulation Package)主要方案由奥地利维也纳大学的 Jürgen Furthmüller 撰写,目前的开发者是 George Kress。VASP 采用赝势及平面波基组进行周期性边界条件下的材料模拟,主要进行 DFT 级别的第一性原理计

算材料的结构与性质,在物理与材料领域的应用面最广。VASP的发展始于1991年,它基于英国剑桥1989年版的CASTEP<sup>[42]</sup>,1995年成为一个稳定的版本。1996年开始MPI(message passing)并行,2004年开始支持Hartree-Fock、GW和线性响应的计算。目前的一个重要发展方向是试图引入量子化学中处理分子的常用方法,如MP2和CCSD(T)等方法,以处理周期性固体<sup>[43]</sup>。

GAUSSIAN<sup>[28]</sup>和VASP<sup>[32]</sup>作为用户最广的程序,两者有着非常显著的区别,分别在学术观点、应用领域、甚至是计算精度、并行效率上均有较大差异。各有各的优势,同时也各有各的短板。这也给予了许多其它软件存在的理由和空间。例如,GAUSSIAN对入选其软件包的各种基础算法和理论方法均有严苛的流程和标准,这使它具有优异的稳定性和可靠性,得到广大用户(尤其是非专业用户)的信赖;但由于其基础算法未能充分适应新硬件的发展而导致并行运算效率没能与时俱进。VASP虽起步较晚,其算法有较好的并行能力,可以在大型超算中心进行较大规模的并行计算,同时其主要是针对周期性模型的计算工具,有能力处理大块的固体体系,满足了近年来对材料计算高速增长的需求;但是它由单一的学术开发组维护与更新,理论方法的更新速度有限,方法的选用上受到一定的限制。

上述众多主流著名计算化学软件,全部都是由美国和欧洲开发或控制的软件。实际上,从70年代初开始一直到现在,国家的支持一直是这些计算化学软件发展的最重要基础。即使是已经商业化的软件如GAUSSIAN,上百名来自学术界的开发者都长期得到国家的支持软件开发的项目经费。其它知名度低一些的通用型计算化学程序也鲜见由非美欧国家开发的软件。纵观40—50年来的计算化学软件发展,这些成功的软件有如下的特点:

(1) 一个强大软件包的健康发展,往往得益于数十位甚至上百名理论计算科学家的长期的投入和贡献。如:(a) 使用GAUSSIAN 16版的要求文献引用如下<sup>[28]</sup>: Gaussian 16, Revision A. 03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Menucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young,

F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016。列表体现了对GAUSSIAN代码有主要贡献的作者,其中大部分与以往版本(如G09和G03)的作者相同;(b) ADF<sup>[38]</sup>在其官网上声明有超过100位的作者;(c) Molpro<sup>[36]</sup>官方列出的主要作者如下:H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Sch. S, P. Celani, W. GyW. Gy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. K G., Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklae, D. P. O'Neill, P. Palmieri, D. Peng, K. Pfl. Pf, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang。各个软件的作者国际化程度一般较高,从中也可以看出有不少中国人的参与。

(2) 这些软件在非商业化阶段,均得到了美欧等国家长期的持续资助。

(3) 著名的计算化学软件包大多起步较早,往往从20世纪70年代起开始由某个著名的理论化学研究组设计开发,经过数十年来不间断的积累和改进。即使是NWChem这样近几年才起步的软件,也是建立在以往丰厚的技术积累的基础上,在美国政府的大力资助下,聚多方顶尖人才进行集中开发而发展起来的。

(4) 此外,软件的使用群体也需要长期的培育。一般来讲,使用这些计算软件的用户首先要求对理论和计算化学领域有较好的基础,其次对所要研究的体系有基本的认识,对所希望采用的计算方法的适用性有所了解。系统学习一个计算软件包的使用是一个较长期的过程。往往一位较资深的研究生用户,也仅仅熟悉一个软件包中少数的功能和方法。如果不是因为该软件无法实现某种必需的功能,一般用户往往不愿意切换并选择一个新的软件。因为那样意味着需要牺牲更多宝贵的科研时间,而且还不能保证新软件一定能够得到其所期望的结果。

## 5 创制我国自有著作权的通用计算化学软件势在必行

国家自然科学基金委员会和中国科学院自2012年起共同组织了学科发展研究。由黎乐民院士负责,撰写了中国学科发展战略“理论与计算化学”专集<sup>[2]</sup>。由于早期我国缺少计算条件,导致理论计算化学研究的重点不是计算方法和计算程序的发展。但改革开放以来,得益于我国国力的持续增强,在科技领域的持续投入,计算能力逐渐处于世界领先水平,同时我国的理论与计算化学研究快速发展,达到国际先进水平。量子化学计算研究队伍快速成长壮大,部分研究成果达到国际先进水平,少数达到国际领先水平。尽管如此,我国仍缺乏自主的通用计算化学软件平台,不仅与当前的学科发展不匹配,同时也不利于相关学科和产业的健康发展。创制我国自有著作权的通用计算化学软件已势在必行。

(1) 发展国有通用计算化学软件包,提升我国理论化学的国际影响力

我国理论与计算化学工作者使用的计算软件工具,绝大多数仍是以 GAUSSIAN、VASP 为代表的由美欧主导开发的通用软件包。虽然这些软件包已具有很好的稳定性和广泛的用户群体,但是其应用与开发、特别是对于我国国内用户而言仍存在诸多的限制,如高斯很早就实行歧视性销售,即中国的用户支付比欧美更贵的价格还不带源程序。尤其是我国大多具有特色功能的软件和模块难以集成进这些主流计算软件包,而往往需要以这些软件包为平台,形成衍生的程序或模块。这是目前国内研究者所普遍采用的方式。如南京大学开发的低标度量化学程序(LSQC<sup>[5]</sup>)可运用 GEBF<sup>[44]</sup>等分块方法实现

大体系电子结构的计算,需与 GAUSSIAN、Molpro 和 GEBF 等程序嫁接;复旦大学开发的 XYG3 系列双杂化密度泛函可以比昂贵的高等级从头算小 2—3 数量级的计算量,精确计算生成热、电离能、电子亲和能、键焓、能垒、非键相互作用和过渡态构型等,需要与 GAUSSIAN、NWCHEM、Q-CHEM 等程序包对接。

虽然这些外挂的程序模块可以实现所对接的通用软件包无法实现的功能,但由于在易用性、无法与通用软件包实现算法同步更新、宣传不足,甚至是自身的代码稳定性不足等各种原因,其影响力往往不如标准内置于著名通用软件包的功能。如 X3LYP<sup>[45]</sup>泛函被 GAUSSIAN 选用并内置为其中标准关键字功能,其 SCI 引用目前已达到 636 次,其中近 5 年 SCI 引用为 280 次。

因此,如果我国拥有自主通用计算化学软件包,将可为国内学者所发展的理论方法及特色功能软件和模块提供可持续的、便利的开发平台;可及时将最新的、优秀的功能与算法集成入软件包,以快速扩大新算法和新功能的国际影响力。

(2) 鼓励发展特色的计算软件,促进学科平衡快速发展

目前由我国自主开发的完全独立的程序包较少,虽在国内外有一定程度的影响,但还相对不足。如在电子结构理论方面有北京大学开发的 BDF<sup>[3]</sup>和厦门大学开发的 XMVB<sup>[4]</sup>。其中 BDF 主要特色是相对论量子化学,其原创的方法包括精确二分量(X2C<sup>[46]</sup>)及相关的扩展,部分模块效率很高,但整体功能不如 ADF<sup>[38]</sup>、Dirac<sup>[47]</sup>等相对论程序。XMVB<sup>[4]</sup>是国际上少有的,也是最好的价键理论程序之一。尽管价键理论更契合传统的化学概念,但是目前主流的电子结构理论体系是分子轨道理论,主流发展方向是大体系的高效率和高精度,而价键理论不适合处理大分子体系,使得该软件的用户数难以扩大。如果将这些优秀、成熟的程序集成入通用软件包,可便于非该领域的研究人员接触和使用这些程序模块,扩大不同领域间的相互交流与融合,亦可形成我国理论计算软件的一个重要特色。

在集中国家力量建立若干个通用软件包的同时,仍需继续鼓励发展多层次、不同领域的、有特色的计算化学程序。不仅要赶超、甚至引领国际同行,也要促进国内学术届的百家争鸣,有利于在理论、方法和算法等各方面发现新的创新点。

### (3) 节省大量外购软件经费

通常来说,国外的软件授权主要按学术机构和商业机构划分。学术机构一般是以课题组为单位采购的。商业机构采购一个软件往往需要数倍于学术机构的价格。单个软件的学术研究组的授权费用在数千美元或欧元。购买后,如果为了获得更新版本,每年需要支付大约 20% 的更新费用,或者需要额外支出 20%—60% 不等的版本升级费用。我国每年为购买这些软件授权,耗费了大量的科研经费(据保守统计,每年平均耗费约数千万元)<sup>[2]</sup>。如果是非学术团体,甚至是超算中心为这些软件授权所支付的费用更加巨大。为此,我国应该建立具有自主知识产权的通用计算化学软件平台。

### (4) 完善学术与产业生态

计算软件作为理论方法联系乃至指导实验研究的桥梁与纽带,其自身的发展也需要在不断的理论联系实际的过程中发现问题和修缮,而理论方法和实验研究本身的发展也将极大得益于计算软件的不断完善。没有先进的理论与计算化学软件就难以引领诸如第一性原理材料设计、蛋白质折叠等重要科学研究。例如:对于新一代的 OLED 显示技术,我国已经或年内将投入的资金超过 4000 亿元,但是几乎所有信息显示核心材料都被国外垄断。由清华大学为主研发的分子发光效率预测程序 MOMAP<sup>[18]</sup>(嫁接于 Turbomole<sup>[37]</sup>或 GAUSSIAN<sup>[28]</sup>),是面向激发态结构与过程的计算化学软件,可以为我国自主研发的发光主体材料设计提供重要的帮助。

近年来中国的计算机硬件水平得到了飞速的提升,目前中国超级计算机的硬件水平在国际上已处于领先地位,并成立了若干国家超级计算中心。其中无锡超算的“太湖之光”和广州超算的“天河二号”分别以 93.01PFlops 和 33.86PFlops 高居 2016 年第 48 届全球 TOP500 超级计算机的前 2 名,远远领先处于第 3 位的美国橡树岭国家实验室的泰坦 17.59PFlops。可是与之相比,我国科学计算软件的发展大大滞后,无法与硬件发展水平相匹配,而且鲜与超算中心进行有效对接。而美国与日本均有针对各自的超算设计或优化的计算化学程序,如 NWCHEM 针对当前的 Intel 众核架构进行了深度优化,使之最高可以同时高效利用 62 560 个核心的 Intel 至强处理器(Xeon)和 Intel 融合协处理器(Phi)。最近美国西北太平洋国家实验室与 IBM Almaden 研究中心合作,正将 NWCHEM 针对杂化

CPU-GPU 架构进行优化,以适应下一代高达 300PFlops 的超级计算机 Summit。

尽管美国已经在这一领域已经处于领先地位,但是仍旧非常重视理论方法和计算软件的集成与开发,以进一步维护其在新能源、新材料、生命科学、新药研制乃至空间探索等各领域的领先地位。美国能源部近期建立了分子科学软件研究所(The Molecular Sciences Software Institute, MolSSI<sup>[48]</sup>),专注于长期发展计算化学软件,将与全美和国际的计算分子科学(CMS)研究组一起,为整个领域设计、开发、测试、部署、和维护关键代码的基础和框架。并将与工业界、NSF 超算中心、国家实验室、和国际机构中的合作伙伴互动,来确认新兴的硬件趋势并采取行动、使用最前沿的计算架构。

因此,如果想成为分子科学与技术强国,发展中国自己的计算化学软件势在必行。这不仅有利于提升我国理论化学的国际地位,有利于节省宝贵的基础科研和企业研发经费,同时也将利于建立更完整的学术和高科技产业生态。与会者一致倡议我国应创制自有著作权的计算化学软件,尤其是包含各种传统和最新的量子化学计算方法的通用计算化学软件平台。

## 6 建设计算化学软件平台、发展新时代通用计算化学软件

华盛顿大学李晓松教授在大会报告中,深入地探讨了理想软件开发应该具有的组织架构,分享了一些国际主流软件开发的成功经验。与会专家一致认为,中国应建立基于新理论、新方法和新算法的具有自主知识产权的通用计算化学软件平台。简单重复开发国际主流计算化学软件程序意义不大,中国的目标应当致力于发展各方面性能优秀并至少有部分特性优于 GAUSSIAN 或 VASP 等国际主流计算化学软件的通用计算化学程序包。作为国家发展目标的计算化学软件平台和现代化先进通用计算化学软件包应具备以下特征:

(1) 软件的功能范围可覆盖大部分主流理论与计算化学的需求。需要发展以第一性原理为核心基础,涵盖电子结构理论、动力学和统计力学等多个学术方向,适用于气、固、液等不同物质形态乃至多相界面的复杂体系的高精度计算和多尺度模拟方法,以适应当前不断增长的在新能源、新材料、新药研制等领域的计算需求。

(2) 软件应具有先进性和独特的功能,在基本性能方面(如自洽场)要超过通用程序;在特色方面,应具有国际领先水平的功能或性能,或者具有原创性的高效算法。在理论方法上的创新并程序化,可用于研究特定的、以前无法解决的重要科学问题,“有特色才有生命力”。在算法效率上的创新,使得在软件层面上尽量实现功耗的降低,同时节省计算资源与时间,可用于研究更复杂的体系。

(3) 结合成熟的数据库技术,甚至是云技术,在一定范围内实现计算数据共享,避免不必要的重复计算。融入最新的机器学习与人工智能技术,一方面充分挖掘已有数据的价值,产生新的认识;另一方面实现目标导向的智能和最优计算。

(4) 拥有标准或统一的对外接口,促进多个专业领域顶尖团队间的协作和成果集成,允许各课题组接入自己发展的新方法,及时收纳国际顶级和通用算法,重点突出我国科学家的优势理论和算法成果。我国理论化学研究呈现出百家争鸣的可喜局面,相当多的课题组为开展特色研究需要发展自己的代码,但即使是通用计算化学软件也不可能涵盖各方向、各类型的理论计算功能。较合理的设计是在通用计算程序中留出接口,方便接入自定义代码,使得软件包的功能随着 5—10 年以上的发展,最终可涵盖分子科学多个方面,包括生命、材料、化学/化工等。

(5) 适应计算机硬件的新变革,尤其是提高对众核构架或异构计算机系统的支持与效率。现代大型计算系统核数往往成千上万,众核处理器也已成为高性能计算的新宠。但是,一些国际主流程序包在超过 10 个核时并行效率已经大大下降,它们对 GPU 加速功能的支持也尚处于初步发展中。中国发展的计算化学程序若能实现对众核并行以及 GPU 加速的支持,甚至是实现在我国配置自主研发的神威 CPU 的“太湖之光”上高效并行,可望在高性能计算平台上获得显著超越国际主流程序包的计算速度,实现中国计算软件的“弯道超车”。

(6) 通用计算化学程序的发展有利于普及计算模拟在一般化学研究中的应用,这尤其需要兼顾非计算化学专业科研人员的需求。通过发展用户友好的操作界面,特别是可视化操作界面,可大大增加计算化学程序的受众范围,提高影响力。注重用户体验,提供有一定的内置纠错能力和可检查机制,提供用户可定制的计算脚本等。

(7) 拥有多层次多学科交叉背景的,专业稳定的软件设计开发与培训队伍。首席科学家(或团队)从国家需求、科学与技术的前瞻性角度指导与把握软件的发展方向;开发人员对接应用、算法、系统和硬件等各部分功能性的需求,实现代码编写与优化;建立程序交换中心和计算结果数据库;与使用用户保持长期联系,及时响应并解决用户的问题,适时组织用户培训活动,不断培育新生用户以扩大软件影响力,维持软件的长期发展生态。

### 参 考 文 献

- [1] Dirac PAM. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London A: Containing Papers of a Mathematical and Physical Character*, 1929, 123 (792): 714–733.
- [2] 国家自然科学基金委员会, 中国科学院. 中国学科发展战略·理论与计算化学. 北京: 科学出版社, 2016.
- [3] Liu WJ, Wang F, Li L. The Beijing density functional (BDF) program package; Methodologies and applications. *Journal of Theoretical and Computational Chemistry*, 2003, 2(02): 257–272.
- [4] Song LC, Mo YR, Zhang Q, et al. XMVB: a program for ab initio nonorthogonal valence bond computations. *Journal of Computational Chemistry*, 2005, 26(5): 514–521.
- [5] Li W, Chen CH, Zhao DB, et al. LSQC: Low scaling quantum chemistry program. *International Journal of Quantum Chemistry*, 2015, 115(10): 641–646.
- [6] Qin XM, Shang HH, Xiang HJ, et al. HONPAS: A linear scaling open-source solution for large system simulations. *International Journal of Quantum Chemistry*, 2015, 115 (10): 647–655.
- [7] Zhang IY, Xu X. Doubly hybrid density functional for accurate description of nonbonded interactions, thermochemistry, and thermochemical kinetics. *International Reviews in Physical Chemistry*. 2011, 30(1): 115–160.
- [8] Guo WP, Wu AA, Xu X. XO: An extended ONIOM method for accurate and efficient geometry optimization of large molecules. *Chemical Physics Letters*, 2010, 498(1): 203–208.
- [9] Wu JM, Xu X. The X1 Method for Accurate and Efficient Prediction of Heats of Formation. *J. Chem. Phys.*, 2007, 127, 214105–214113.
- [10] Li W, Li SH. Cluster-in-molecule local correlation method for large systems. *Science China Chemistry*, 2013, 57(1): 78–86.
- [11] Zhang PY, Han KL. GQSD: The program for the graphic processing units accelerated quantum scattering dynamics. *International Journal of Quantum Chemistry*, 2015, 115 (11): 738–743.



- [12] Liu J. Recent advances in the linearized semiclassical initial value representation/classical Wigner model for the thermal correlation function. *International Journal of Quantum Chemistry*, 2015, 115(11): 657 – 670.
- [13] Zhao DX, Liu C, Wang FF, et al. Development of a polarizable force field using multiple fluctuating charges per atom. *Journal of Chemical Theory and Computation*, 2010, 6(3): 795 – 804.
- [14] Zhu YL, Liu H, Li ZW, et al. GALAMOST: GPU-accelerated large-scale molecular simulation toolkit. *Journal of Computational Chemistry*, 2013, 34(25): 2197 – 2211.
- [15] Shen HJ, Li Y, Xu PJ, et al. An anisotropic coarse-grained model based on Gay-Berne and electric multipole potentials and its application to simulate a DMPC bilayer in an implicit solvent model. *Journal of Computational Chemistry*, 2015, 36(15): 1103 – 1113.
- [16] Jiang F, Wu Y-D. Folding of fourteen small proteins with a residue-specific force field and replica-exchange molecular dynamics. *Journal of the American Chemical Society*, 2014, 136(27): 9536 – 9539.
- [17] Zhang XJ, Shang CQ, Liu ZP. From atoms to fullerene: stochastic surface walking solution for automated structure prediction of complex material. *Journal of Chemical Theory and Computation*, 2013, 9(7): 3252 – 3260.
- [18] MOMAP (Molecular Materials Property Prediction Package) <http://www.shuaigroup.net/index.php/home-build-programs>.
- [19] Yuan YX, Pei JF, Lai LH. LigBuilder 2: a practical de novo drug design approach. *Journal of Chemical Information and Modeling*, 2011, 51(5): 1083 – 1091.
- [20] Rao L, Chi B, Ren YL, et al. DOX: A new computational protocol for accurate prediction of the protein-ligand binding structures. *Journal of Computational Chemistry*, 2016, 37(3): 336 – 344.
- [21] Hohenberg P, Kohn W. Inhomogeneous electron gas. *Physical Review*, 1964, 136(3B): B864 – 871.
- [22] Kohn W, Sham LJ. Self-consistent equations including exchange and correlation effects. *Physical Review*, 1965, 140(4A): A1133 – 1138.
- [23] Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 1988, 38(6): 3098 – 3100.
- [24] Lee CT, Yang WT, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Physical Review B* 1988, 37(2): 785 – 789.
- [25] Becke AD. Density-functional thermochemistry 3: the role of exact exchange. *The Journal of Chemical Physics*, 1993, 98: 5648 – 5652.
- [26] Stephens PJ, Devlin FJ, Chabalowski CF, et al. Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields. *The Journal of Physical Chemistry*, 1994, 98(45): 11623 – 11627.
- [27] Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Physical Review Letters*, 1996, 77(18): 3865 – 3868.
- [28] GAUSSIAN. <http://gaussian.com/>.
- [29] [http://www.nobelprize.org/nobel\\_prizes/chemistry/laureates/1998/](http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/).
- [30] [http://www.nobelprize.org/nobel\\_prizes/chemistry/laureates/2013/](http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/).
- [31] [https://en.wikipedia.org/wiki/List\\_of\\_quantum\\_chemistry\\_and\\_solid-state\\_physics\\_software](https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software).
- [32] VASP. <http://www.vasp.at/>.
- [33] NWCHEM. [http://www.nwchem-sw.org/index.php/Main\\_Page](http://www.nwchem-sw.org/index.php/Main_Page).
- [34] GAMESS. <http://www.msg.chem.iastate.edu/gamess/>.
- [35] Q-CHEM. <http://www.q-chem.com/>.
- [36] MOLPRO. <http://www.molpro.net/>.
- [37] TURBOMOL. <http://www.cosmologic.de/turbomole/support-download.html>.
- [38] ADF. <https://www.scm.com/about-us/our-people/>.
- [39] MOPAC. <http://www.openmopac.net/MOPAC2016.html>.
- [40] DFTB+. <http://www.dftbplus.org/>.
- [41] ReaxFF. <https://www.scm.com/product/reaxff/>.
- [42] CASTEP. <http://www.castep.org/>.
- [43] Booth GH, Grüneis A, Kresse G, et al. Towards an exact description of electronic wave functions in real solids. *Nature*, 2013, 493(7432): 365 – 370.
- [44] Li W, Li SH, Jiang YS. Generalized energy-based fragmentation approach for computing the ground state energies and properties of large molecules. *The Journal of Physical Chemistry A*, 2007, 111(11): 2193 – 2199.
- [45] Xu X, Goddard WA. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(9): 2673 – 2677.
- [46] Liu WJ, Peng DL. Exact two-component hamiltonians revisited. *The Journal of chemical physics*, 2009, 131: 031104.
- [47] Dirac. <http://www.diracprogram.org/doku.php>.
- [48] MolSSI-olSSI. <http://molssi.org>.

## Creative instrumentation in the perspective of theoretical chemists

Xu Xin<sup>1</sup>      Wu Yundong<sup>2</sup>      Fang Weihai<sup>3</sup>      Shuai Zhigang<sup>4</sup>  
Gao Feixue<sup>5</sup>      Zhang Shouzhu<sup>6</sup>      Meng Qingfeng<sup>7</sup>      Lei Jinglei<sup>5,8</sup>

(1. *Department of Chemistry, Fudan University, Shanghai 200433* 2. *Shenzhen Graduate School, Peking University, Shenzhen 518055*  
3. *College of Chemistry, Beijing Normal University, Beijing 100875* 4. *Department of Chemistry, Tsinghua University, Beijing 100084*  
5. *Department of Chemical Sciences, National Natural Science Foundation of China, Beijing 100085*  
6. *Department of Mathematical and Physical Sciences, National Natural Science Foundation of China, Beijing 100085*  
7. *Policy Bureau, National Natural Science Foundation of China, Beijing 100085*  
8. *School of Chemistry and Chemical Engineering, Chongqing University, Chongqing 400044*)

**Abstract** Looking back at the history of science, the invention of any important instrument always has its origin in new concepts and principles put forward by theory. New concepts and principles can bring about original ideas for creative instrumentation, and ultimately help to promote the innovation of China's basic scientific research. Currently, theoretical and computational chemistry is in the boom years. This discipline has also undergone rapid development in China, where strong research teams have formed and many significant research progresses have been made. Development of novel computational methodology is an important component of theoretical development. Computer software is the key in translating theory into applications. There are many software platforms in the world, the majority of which are, nonetheless, made in the US or European countries. The consensus achieved from the 181th Shuang-Qing Forum is that, to meet the need of experimental science for the discovery and design of new chemical and materials systems, next-generation software developments are essential. China has great potential to lead the world efforts in theory and software development with the stable long-term support in parallel with the innovative development of large instrument and equipment.

**Key words** instrumentation; software; computational chemistry; theoretical chemistry