

## 量子化学文献データベースの学術的評価

(京大・福井センター<sup>1</sup>, お茶大・院人間文化<sup>2</sup>, 大阪市大・院理学<sup>3</sup>, 産総研・ナノシステム<sup>4</sup>, 名古屋大・院情報科学<sup>5</sup>, 首都大東京・院理工<sup>6</sup>, 千葉工大・工<sup>7</sup>, 岐阜大・地域科学<sup>8</sup>, 横浜市大・院国際<sup>9</sup>, 計算化学振興財団<sup>10</sup>, 大阪大・院基礎工<sup>11</sup>)  
 山邊信一<sup>1</sup>, 細矢治夫<sup>2</sup>, 松下叔夫<sup>3</sup>, 長嶋雲兵<sup>4</sup>, 鷹野景子<sup>2</sup>, 古賀伸明<sup>5</sup>, 波田雅彦<sup>6</sup>, 松澤秀則<sup>7</sup>, 橋本健朗<sup>6</sup>, 和佐田裕昭<sup>8</sup>, 立川仁典<sup>9</sup>, 西川武志<sup>10</sup>, 重田育照<sup>11</sup>, 森寛敏<sup>2</sup>

### Academic Evaluation of Quantum Chemistry Literature Data Base

(<sup>1</sup>Kyoto Univ., <sup>2</sup>Ochanomizu Univ., <sup>3</sup>Osaka City Univ., <sup>4</sup>AIST, <sup>5</sup>Nagoya Univ., <sup>6</sup>Tokyo Metropolitan Univ., <sup>7</sup>Chiba Institute of Technology, <sup>8</sup>Gifu Univ., <sup>9</sup>Yokohama City Univ., <sup>10</sup>FOCUS, <sup>11</sup>Osaka Univ.)  
<sup>1</sup>S. Yamabe, <sup>2</sup>H. Hosoya, <sup>3</sup>T. Matsushita, <sup>4</sup>U. Nagashima, <sup>2</sup>K. Takano, <sup>5</sup>N. Koga, <sup>6</sup>M. Hada, <sup>7</sup>H. Matsuzawa, <sup>6</sup>K. Hashimoto, <sup>8</sup>H. Wasada, <sup>9</sup>M. Tachikawa, <sup>10</sup>T. Nishikawa, <sup>11</sup>Y. Shigeta, <sup>2</sup>H. Mori

#### §1. はじめに

標題のデータベースは、quantum chemistry literature data base を略して QCLDB と称されて来た。これは、1975 年の北海道大学の 大野公男教授の呼びかけが契機となり、現在まで国内の量子化学研究会のメンバーが構築・運用を継続して来た。QCLDB のシステムの構成や、データ収集・検索方法については、過去に一度、分子構造総合討論会で発表した。この時期以降、量子化学には、理論化学・計算化学としてめざましい発展があった。同時にインターネット時代となり各学術雑誌の電子化及び商用データベースの普及があり、情報検索には不自由の無い環境が実現したと言える。この環境下で、改めて QCLDB のたどってきた経緯と実績を総括し、その役割・意義を吟味・自己評価し、次世代への継承・発展を期したい。

#### §2. QCLDB の経過と QCLDB II の発足

1978 年から、所謂 *ab initio* 計算を含む文献を、国際的な主要雑誌から決められた手順で QCLDB のデータとして抄録する作業が開始された。収集項目が指定され、電子データとして収集され、査読により、データの質を保障する体制が取られた。この体制は現在も変わらず、収集・査読が QCLDB での中心作業となっている。

2004 年 4 月より、web 版の QCLDB II を公開した。QCLDB II には、単にインターネット時代への対応及び激増する該当文献への措置の意味だけでなく、化学の学術データベースの性格の明確化の意図があった。その性格を「理念」として、以下の (1), (2) に掲げた。

- (1) データ収集・査読が研究活動の基礎としての文献調査の意義を持つ。収集査読が研究活動であることと、QCLDB II の検索利用両立を図る。このため、前者の支障・負担となる恐れのある後者側の利便性の追及には必ずしも応じない。
- (2) QCLDB II は化学及びその関連分野の雑誌を対象としたデータで構成されること。原子・分子・分子集団・錯体等を通じ、物質の個性が見える研究成果を収集対象とする。関連分野の雑誌を対象とするか否かは研究会が判断する。対象雑誌を *core journals* として、収集・査読体制をとり、該当データを選出する。

また、収集項目は、QCLDB I からの改訂として、(3) に挙げた 9 個とした。改訂には、文献情報の補充と同時に、理念 (1) の文献調査での負担軽減の意図があった。

### (3) QCLDB II での収集項目

I Author(s), II Title, III Journal (6 桁 coden), vol, number, page range, year, IV Molecular formula(e), V Compound name(s), VI Method(s), VII Basis set(s), VIII Properties (onel や reac 等の 4 桁指定の 14 個の keywords), IX Comments

QCLDB I 運用末期には、生物化学の巨大分子での反応や触媒化学の表面吸着等をモデル化した計算が増加し、一体どの範囲が *ab initio* 計算対象なのか判断困難な事態が頻発した。これは項目 IV 分子式欄入力の困難につながり、収集・査読の大きな負担となった。このため、個性が見えない反応系では、この欄での一部入力省略可とし、新たに、中かっこ { } を用いての IUPAC 命名法及び慣用名以外の物質の構造を表す自由な用語入力欄を項目 V に設けた。(これらの用語は、検索側にヒットし易いとの条件が付く)。IX コメント欄が、最重要項目である点は、QCLDB I 及び QCLDB II を通じて変わらない。この欄では、著者視点の密度高い情報入力を求めている。

### (4) 収集・査読から公開まで

(2) での core journals に対し、収集・査読者が割り当てられて、半年単位で、QCLDB II データが集められる。それらに対して、研究会メンバーがデータ全体を通して、誤入力の有無、コメント欄の過不足等を吟味・修正を行う。これらの作業を経たデータが以下の §3 での web サイトで順次公開されて行く。

## §3. 検索システム

収集された QCLDB II のデータへの検索機能強化のため、新しい検索システムを作成中である。これについては当日説明する。

## §4. 現状での成果と課題

- (1) 収集に関わる若手研究者の文献調査の促進と査読者側からの指導・アドバイスの体制は、曲がりなりにも理念 (1) を具体化している。
- (2) 理念 (2) に従い、収集されるデータの網羅性は無くなって来ている。だが、集められるデータの質は維持されている。母集団のデータ量がコンパクトであるため、他のデータベースでの欲しい情報の絞り込み作業は必要なく検索可能である。また、ヒットしたデータには著者目線の情報が記述されている。併用すべきデータベースの一角として、QCLDB II の意義は変わらないと考えられる。

## 謝辞

本研究は、科学研究費補助金（研究成果公開促進費）データベース（研究成果データベース）（課題番号 247003）の支援を受けると共に、大学共同利用機関法人自然科学研究機構分子科学研究所との共同事業として実施している。