

演 題	CML/XML を用いた多配座データベースシステムの開発	
発 表 者 (所 属)	豊橋技術科学大学 吉野 智、吉住 亮、後藤仁志	
連 絡 先	441-8580 愛知県豊橋市天伯町雲雀が丘 1 - 1 豊橋技術科学大学 知識情報工学系 後藤研究室 TEL: 0532-47-0111 内線5730 FAX:0532-48-5588	
キ ー ワ ー ド	XML, CML, SQL Database Server, CONFLEX, Conformational space search	
開 発 意 図 適 用 分 野 期 待 効 果 特 徴 な ど	CONFLEX の出力データから配座情報を自動的に登録し、ユーザが必要とする情報を Web 上から閲覧できるデータベースシステムを開発した。	
環 境	適 応 機 種 名	パーソナルコンピュータ及びワークステーション
	O S 名	Red Hat Linux7.2 (他のOSでも可)
	ソ ー ス 言 語	C, Java, PHP
	周 辺 機 器	
流 通 形 態 (右 の い ず れ か に を つ け て く だ さ い)	<ul style="list-style-type: none"> ・日本コンピュータ化学会の無償利用ソフトとする ・独自に頒布する ・ソフトハウス、出版社等から市販 ・ソフトの頒布は行わない ・その他 未定 	具 体 的 方 法

1. はじめに

近年、パーソナルコンピュータの処理能力の向上や高速計算処理技術の一般化により、これまで以上に大きな分子の高精度理論計算が行えるようになった。また、当研究室で開発している配座空間探索プログラム CONFLEX においても、並列技術や Grid 技術が導入され、生体高分子を含む超巨大有機化合物の配座探索も実現可能になりつつある。しかし、計算対象となる分子が大きくなれば、総配座数も多くなり、それらに応じて、理論計算や配座探索の出力ファイルのデータサイズもまた膨大になる。従って、必然的に出力ファイルの管理は難しくなり、そして、それらのファイルの中から必要な情報のみを適切に取り出すことも困難になっていくだろう。

そこで、本研究では CONFLEX の出力データの中からあらかじめ設定された重要な分子情報、計算法に関する情報、および配座情報を取り出し、データベースに登録するまでを自動的に取り扱い、また、ネットワークを介してデータベースの中から必要な情報のみを取り出し、閲覧できるような、統合的データベースシステムを開発する。

2. データの構成

本システムで取り扱うデータの形式は XML に準拠することとし、基本的には化学情報を扱うために標準化が進められている XML/CML を可能な限り採用することにした。ただし、計算手法の特定と配座異性体を記述する部分に関してのみ拡張した。また、CONFLEX が生成する配座情報、特に全ての座標データを XML/CML に変換してデータベースに登録すると、分子が大きくなるにつれて膨大なサイズになってしまうことは明らかである。そこで本システムでは、物質名、分子式、文献等の検索に必要な分子情報、計算法やエネルギーなど基本配座情報のみを登録し、座標データに関してはその格納位置を示すリンク情報として登録することにした。

尚、本システムでは、多くのプラットフォームにおいて広く利用されているデータベースシステム PostgreSQL を使用した。また、表形式のデータを扱う SQL と木構造である XML/CML の相違を補完するために「XMLPGSQL」という拡張関数群を導入した(*1)。

3. システムの構成

本システムの基本構成を図 1 に示した。ここで、CONFLEX から出力された計算結果は、別のシステムによって自動的に本システムのサーバに送られることとする。本システムでは、データを受け取ると、そこからあらかじめ設定された情報を抽出し、XML/CML ファイルを作成し、データベースに格納する。

一方、格納されたデータを参照する場合、検索用の Web ページからキーワードを用いて必要なデータを絞り込むことも可能であるが、あらかじめ参照分子と配座が特定されている場合は、現在開発中の Java アプレットを使って連続的に特定の分子の配座異性体座標データを取り出し、また、三次元構造を表示することも可能にする予定である。

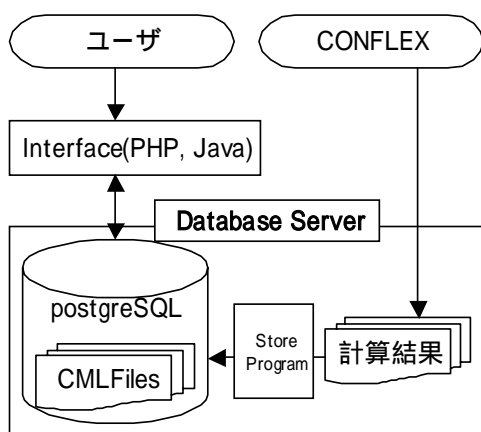
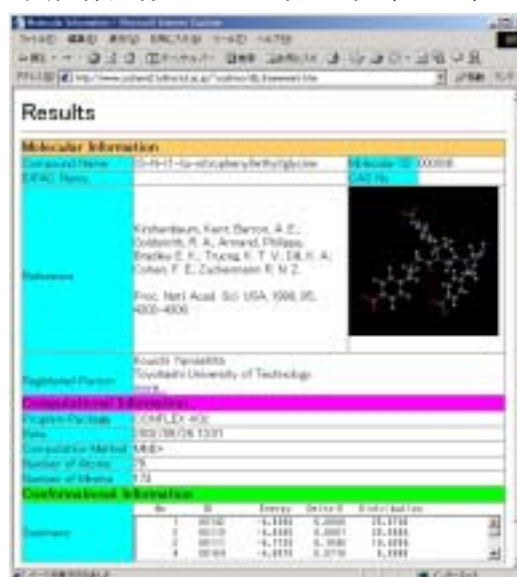


図 1：本システムの基本構成



Energy	Level	Occupancy
0.0000	4.0000	2.0000
0.0010	4.0000	2.0000
0.0020	4.0000	2.0000
0.0030	4.0000	2.0000

図 2：検索結果の出力例

謝辞：

本研究は科学技術振興事業団「計算科学技術活用型特定研究開発事業」の援助を受けて行われています。

参考文献：

(*1) 小松 誠、「XML+PostgreSQL で作る万能データベース」、Software Design、2001 年 3 月号、技術評論社、pp.15-69.