高速低コスト大規模分子計算の実現に向けての試み

工業技術院産業技術融合領域研究所計算科学グループ　○長嶋　雲兵

1．はじめに

非経験的分子軌道法は、その原理から考えてさまざまな機能性分子の設計や開発に対して最も基盤的でかつ重要な手法である。しかし、その計算量が膨大であるため、設計したい大きさの大規模分子ではなくそれを簡素化した小規模モデル分子への適用がせいぜいである。そのため、分子設計に必要とする情報をなかなか十分与える事ができないでいる。ここ数年の計算機の進歩、特にスーパーコンピュータや高性能ワークステーションのような大規模並列計算機の進歩により、ようやく（生体）高分子を意識した1,000基底を超える分子軌道計算が行われつつあるが、これらの計算機は高価であり設備の維持管理も大変であるため、研究者が研究室レベルで大規模計算を実現することは容易なことではない。そこで、我々は、「現実を反映した大規模分子系」の分子軌道計算を、「低コスト＝パーソナルユース」で実現すべく、Fig.1に示すような大規模分子をターゲットとした分散小メモリ型並列アーキテクチャである分子軌道計算専用並列計算機MOE（Molecular Orbital calculation Engine）の開発に現在取り組んでいる。

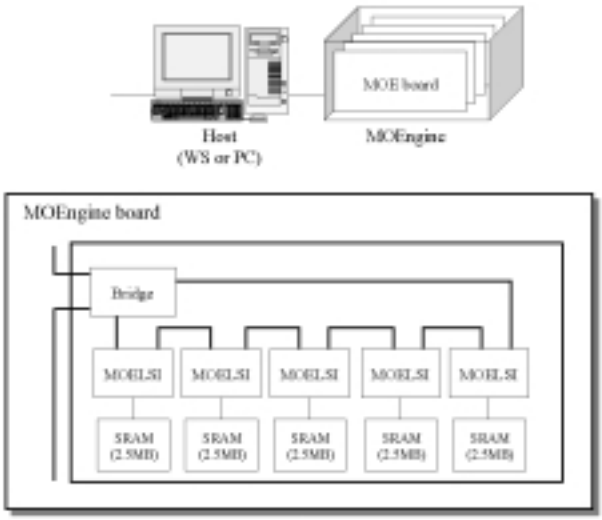

Fig.1 MOEngine, a special purpose machine for accelerating molecular orbital calculations.

非経験的分子軌道計算において最も時間を要するのが二電子積分を計算しフォック行列を生成する部分であるため、大規模計算を高速化する際にはこの部分の並列化が肝要である。これまでに、スーパーコンピュータや大規模ワークステーションクラスタを用いた多くの並列分子軌道計算が行われているが、その殆どは、（並列）各マシンに全ての密度行列及び全てのフォック行列を持たせて従来のフォック行列作成アルゴリズムをそのまま並列化した、いわゆる共有メモリ型アルゴリズムを採用している。しかし、このような方法だと、例えば100残基程度の蛋白質に相当する10,000次元の計算に必要な密度行列及びフォック行列全てを保持するために、各プロセッサ毎に最低800MBものメモリが必要になる。しかも、必要となるメモリ容量は系のサイズの二乗に比例して増大するため、莫大なメモリ量を必要とする。このように、従来のアルゴリズムでは、計算機も高価なものとなりスケーラビリティに乏しい方法である。

現在開発中の専用計算機MOEngineでは、安価なコストで高速計算を実現するため搭載メモリ量（や設備の規模）を小さくしている。そのため各並列プロセッサに全ての行列要素を搭載できないので、MOEngineで大規模フォック行列を計算すつためには、ホストと各プロセッサの間で行列要素の通信を行なう並列分散アルゴリズムが必要になる。そこで我々は、各専用計算機MOEngineのような各プロセッサのメモリが小容量（数MByte）しかない分散メモリ型並列アーキテクチャのもとで大規模フォック行列を計算するための新規アルゴリズムを考案した。

2．研究成果

本研究で開発された新規アルゴリズムを用いた場合の計算される積分の数をTable 1に示す。

Table 1 Number of basis *N* and the number of cutoff survival basis *M*.

molecule	6-31G		6-31G*		Extended	
	N	M	N	M	N	M
H2O	13	13	25	25	55	55
C6H6	66	66	120	117	270	262
Arginine	136	119	250	196	560	457
GAQ	207	180	375	263	845	671
GAQMY	427	314	769	455	1733	1176
d(CGC)	1203	479	2106	635	4797	2069
Gramicidin	2910	710	5295	869	11910	3525
BPTI	4694	739	8504	891	19138	4121
IGG	7940	803	14438	962	32476	4806
Lysozyme	10967	881	19850	1044	44705	5578

6-31G,6-31G*,Extendedというのは基底関数の種類であり、N,Mは、それぞれ全基底関数、新しいアルゴリズムにおいて計算すべき基底関数の平均である。これをみると、大規模な系に関しては、計算すべき基底関数が1/20となっており、計算コストは(1/20)**4=1/160000に減少している。Fig.2にサイズ減少の度合いのイメージを示した。本アルゴリズムは、特に巨大分子系に有効であることがわかる。

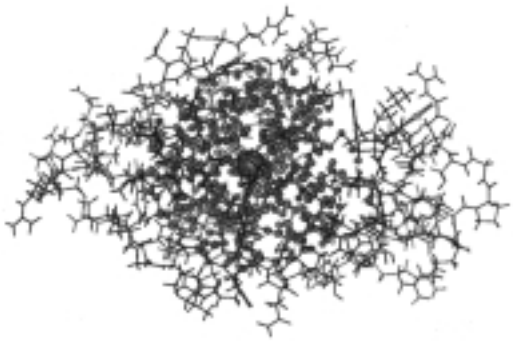

Fig.2 Cutoff scheme in Lysozyme molecule.

3．まとめ

本研究では、分子軌道計算専用機に向けたフォック行列計算アルゴリズムを開発した。それは巨大な分子系に関して有効であり、Lysozymeの場合では、16万分の１に計算コストが減少している。

研究成果

1)高島　一、山田　想、小原　繁、北村一泰、稲深二郎、宮川宣明、田辺和俊、長嶋雲兵、「分散メモリ型並列計算機に適した新しい大規模フォック行列生成アルゴリズム。（１）積分カットオフとの関連」、J Chem. Software, 印刷中.

2)高島　一、山田　想、小原　繁、北村一泰、稲深二郎、宮川宣明、田辺和俊、長嶋雲兵、「分散メモリ型並列計算機に適した新しい大規模フォック行列生成アルゴリズム(２)RT並列アルゴリズム」応用数理学会論文誌、印刷中

これは平成１２年３月９日に開催した
計算科学技術活用型特定研究開発推進事業
研究報告会（主催 科学技術振興事業団）
の予稿集から抜粋したものです。

識別記号 OY_H10S_???_H11