

5-2 計算科学研究センターの現状と将来

分子科学研究所・電子計算機センターは1978年に設立され、2000年4月より、岡崎国立共同研究機構（岡崎機構）計算科学研究センターに改組された。電子計算機センターは日本全国の分子科学研究者に大規模計算を実行する環境を提供する計算機センターとして設立され、22年を経た今日においても所内外の分子科学研究の基盤施設としての重要性は変わらない。実際、「分子研レポート 94」に報告されている通り、外部評価委員、運営委員、所内外の利用者の多くは、本センターが分子科学理論計算分野へ貢献してきた歴史的経緯を高く評価しており、当初の目標を高い水準で達成できていることを認めている。

5-2-1 現在の計算機システム

2000年3月から導入されたスーパーコンピュータシステムを図1の左側に、昨年度に更新を終えた汎用高速演算システムを右側に示す。新スーパーコンピュータシステムは、富士通製 VPP5000 と SGI 製 Origin2800 から構成される。

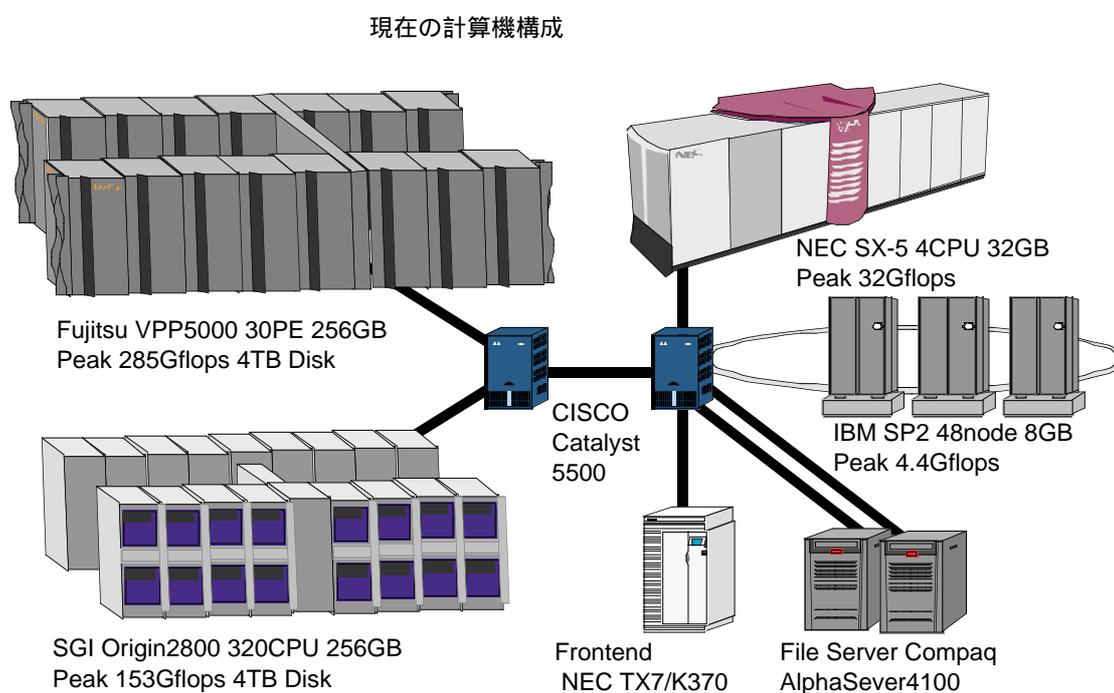


図1 . 2000年11月以降の計算機構成

VPP5000 は1 CPU 当たりの最高演算性能が9.5 Gflops のベクトル演算装置30台から構成され、各 CPU に8 ~ 16 GB の主記憶装置を持つベクトル並列計算機である。一方、SGI Origin2800 は1 CPU 当たりの最高演算性能が0.59 Gflops のスカラ演算装置256 CPU から構成され、CPU 当たり1 GB の主記憶をそれぞれの CPU から共有メモリとしてアクセスが可能な分散共有方式の超並列計算機である。2000年11月にはSGI側システムは320 CPU に増強され現在に至っている。VPP5000 では高速なベクトル演算能力を活かした大型ジョブの逐次演算処理はもちろん、例えば8台以上のベクトル演算装置を使った大規模なベクトル並列演算が可能となる。Origin2800 は Non Uniform Memory Access (NUMA)方式と呼ばれる論理的な共有メモリ機構を有する。NUMA は主記憶装置が各 CPU に分散して配置されているため CPU から主記憶へのアクセス速度が非等価ではあるが、利用者プログラムから大容量のメモリを容易に利用することが出来るため、大規模な並列ジョブの実行が可能となる。1 昨年度、導入された SX-5 は1 CPU 当たり8Gflops の最高演算能

力を持つ共有メモリ型ベクトル計算機であり、SP2は分散メモリ型スカラ並列計算機である。更に、2001年3月中旬より、分子科学およびバイオサイエンス関連の計算処理を目的とした高速シミュレーションシステムが稼働を始める。高速シミュレーションシステムは、日立製作所SR8000とSGI 3200から構成される中規模なシステムであり、当面の間は主に機構内における利用を目的として運用を行う。今後は、個々の計算機の特徴を活かしつつ、大規模な分子科学計算が実行できる基盤施設としての機構外施設利用として、および機構内の多様な計算処理に適応できるセンターとして、分散処理環境をさらに充実させ、利用者ジョブの効率的な実行環境を構築することがこれからの課題である。

2002年（平成14年度）には、機構から約1km離れた場所（E地区と呼ぶ）に、新たに統合バイオサイエンス棟と2つの付属施設棟が完工する。統合バイオサイエンス棟内には、E地区の計算科学の研究基盤施設およびネットワークのノード室として計算科学研究センターの分室が設置される予定であり、上述の高速シミュレーションシステムと関連周辺装置を移設する。更に平成14年度末には、計算科学研究センターの汎用高速演算システム（図1、右側）の拘束レンタル期間が終了するためシステムの更新が可能となるが、その更新時期については、施設利用を目的とした汎用システムとしての役割、近い将来予想される機構内の計算処理需要の増大等を熟慮して、現在（本稿原稿執筆時点）基礎生物学研究所、生理学研究所の関係者および外部運営委員と協議中である。

5-2-2 計算機施設利用枠の新設「特別申請」

本計算科学研究センターの前身である「分子研電算センター」はこれまで全国700人におよぶ分子科学者に対して文字どおり「共同利用施設」としてサービスを提供してきた実績をもっている。これは、他の研究機関の「スパコンセンター」がその利用者の大部分を事実上その機関内に閉じていることを思うとき、「分子研電算センター」が誇るべき偉大な実績であり、今後も「計算科学研究センター」が継承すべき特色である。しかし、一方、ワークステーションや高性能のパーソナルコンピュータの普及に伴って、これまで「計算機センター」が果たしてきた役割の一部が変更を迫られていることも確かである。これまで計算機センターを利用して行われていた計算のかなりの部分がワークステーションやパソコンで簡便に行えるようになり、「煩わしい手続きをして大型センターを利用するまでもない」と考えるユーザーも増えている。他方、国際的には米国を中心に超並列マシンの性能を極限まで使って初めて可能になるような計算が報告されつつあり、このままでは我が国の理論化学が国際的に遅れをとってしまうという危機感も生まれている。すなわち、一方では「できるだけ多くの研究者へのサービスの提供を維持」しながら、他方では「世界のピークを目指すような大規模計算を可能にする」という「二兎を追う」ことを要求されている。このような要請に応えるため、2000年春の運営委員会において計算資源の利用枠を「一般利用」と「特別利用」に2本化することが提案された。承された。「一般利用」はこれまでとほぼ同様であり、同様の手続きで申請を行う。他方、「特別利用」は毎年小数の大規模計算プロジェクトに供するものであり、特別の申請手続きと審査を経て許可されるものである。「特別利用」は他の多くの利用者の「犠牲」の上に行うわけであるから、その申請者の責任の自覚が不可欠であり、審査もそのような高い基準で行われる必要がある。

平成13年度春から試行予定の「特別利用」の概要を以下に示す。

毎年、少数（10件程度）の特別利用申請を公募する。このカテゴリーは他のユーザーの利用をある程度制限することを前提に行うもので、それを正当化する唯一の根拠は、その利用の結果、一般の利用ではできない、際立った研究成果を挙げる事以外にない。そのような成果を期待するには申請に関する審査と研究結果報告に対する評価をできるだけ厳しく行う必要がある。まず、申請書に基づいて審査を行い、点数に基づいて順位をきめる。その上位順位から、「特別利用」の枠内で採用を決定する。審査は通常の論文審査の場合と同様に複数の審査委員に依頼する。審査委員は

あらかじめ定められた審査基準に基づいて審査を行い採点すると同時に審査意見を述べる。最終的な可否は審査結果に基づき運営委員会において行う。また、研究結果の評価に関しては論文の発表を原則とし、その論文は計算科学研究センター特別利用申請と明記する。

2001年度中に特別利用に供する計算機資源は以下の2機種である。

超並列計算機 SGI3800 128 CPU (128 GB).....特別枠専用利用

(ただし2001年6月までは主記憶は64GBに制限される)

ベクトル計算機 VPP3000 最大16 CPU (128 GB).....一般施設利用ジョブと共存利用

VPP3000については特別枠の利用期間を限定する。

5-2-3 機構化の現状と将来展望

岡崎機構の計算科学研究センター(旧分子科学研究所電子計算機センター)はこの1年の間にいくつかの点で大きな変貌を遂げた。その第一はスーパーコンピュータの更新であり、この更新によって、演算処理能力は従来の約10倍に向上し、これまで滞りがちだった計算処理速度を大幅に改善することができた。

第二は従来の分子科学研究所のセンターから機構のセンターへの組織的変更である。この変更は、一方では、将来、岡崎機構内で分子科学と生物科学を融合した計算科学の新しい展開に大きな展望を与えるものであると同時に、他方では、もしそのような展開が十分な裏付けもなく性急に行われた場合、分子科学分野において国際的に高い評価を確立している計算センターの存立基盤そのものを危うくする契機を孕んでいる。ここでは「計算科学研究センター」の運営に主として責任を負っている分子科学研究所電算機室の立場から、特に、「機構化」に伴って生じる諸問題について言及し、「センター」の将来を展望する。

「計算科学研究センター」の将来を展望する上で二つの「座標軸」が必要である。ひとつは旧「分子研電算センター」が共同利用施設として分子科学の発展に果たしてきた役割であり、他のひとつは計算科学の発展において生物関係の問題が占める比重の増大である。「分子研電算センター」は約22年前に設立以来、国内における分子科学者の共同利用施設として、一方では国際的にもピークをなす計算を行い、他方では国内の計算化学全体の底上げを行うという二重の役割を果たしてきた。まだ、ワークステーションが普及していなかった時期に地方の大学などで計算機にアクセスできなかった研究者に計算資源を提供し、その中から国際的にも高い評価を受けている分子科学のリーダーを数多く輩出してきたことは周知の事実である。このことは、現在、世界で発表されている量子化学関連論文全体の2、3パーセントが旧「分子研電算センター」を使って行われたものであることから伺われる。さらに、最近、センター利用者の分布に見られる特徴は、有機化学や固体物性分野を中心として実験研究グループの利用の比重が増大していることである。その理由のひとつは量子化学、分子シミュレーションを始めとする理論化学における(計算プログラムを含む)各種方法論が成熟期を迎え、実験家自身ないしは実験家と理論家との協力によって、実験条件に近い系の理論計算が可能になったことが上げられる。このような傾向はナノサイエンスを始めとする物質科学の最近の展開によって今後さらに拍車がかげられることが予想される。最近のセンター利用状況に見られるこれらの特徴は実験研究において理論計算が有効であり得ることを実証する有力な根拠であり、まさに、旧「分子研電算センター」がその創設において企図した目的のひとつであった。したがって、分子科学分野における「計算科学研究センター」の役割が今後益々増大することは必至であり、この分野でのサービスを引き続き強化していくことが必要である。

もう一方の「座標軸」である生物関連分野の位置付けであるが、これまでもこの分野ではprotein data bankなどに象徴されるような国際的に公開された生物情報データを利用して、一方ではいわゆるバイオインフォマティクスに代

表されるようなデータベース処理において、他方では分子動力学 (MD) やモンテカルロ法 (MC) といった分子シミュレーションにおいて、計算科学が重要な役割を演じてきた。もちろん、分子科学自身、その発展に少なからぬ貢献をしてきたことは、MD や MC といった計算科学の手法が分子科学の分野で開発され、発展してきたことから見ても明らかである。このような生物分野における計算科学の位置付けは「人ゲノム計画」の進展とともに以前にも増した高まりを見せている。ゲノムの全解析によってもたらされる膨大な情報を巡って、医学から基礎科学に至る様々な学問分野だけでなく薬品工学や生命工学を含む広大な工業分野が創出されるであろうことは火を見るより明らかである。今後、ゲノムの情報解析において計算科学が有用な学問的知見を引き出し、また、新たな社会的資産を生み出すために重要な役割を果たすことは疑いないであろう。生物科学系の 2 研究所を有する機構内の「計算センター」がこのような学問の進展に手をこまねていることは許されないし、まさに、このことが「計算センター」機構化の重要な契機ともなった。したがって、「計算科学研究センター」が将来展望として生物関連分野の利用を視野に入れた展開を図るべきであることはいうまでもない、その方向については今後十分な検討が必要である。その検討の材料として以下の要素が考慮されるべきであろう。そのひとつは生物関連の計算科学分野の国内における現在の状況である。この分野は、現在、大きくふたつの流れで進みつつある。その流れのひとつはゲノム情報など生物「データベース」から情報処理の手法により生物進化や医学における有用な情報を得ることを目的にしたいわゆるバイオイオンフォ - マテイクスの分野である。この分野はこれまで京大 (化研)、東大、名大 (生物)、阪大 (蛋白研) を中心にいくつかの研究拠点が形成されており、これらに岡崎の基生研の「情報生物学研究センター」が新たに加わろうとしている。また、この分野の共同利用計算センターとしては遺伝研が機能している。生物関連計算科学のもうひとつの重要な流れは分子レベルから蛋白質の構造の成り立ちや揺らぎ、あるいはその機能を解明しようとする分野であり、いわば「生物分子科学」とも呼ぶべき研究領域である。

以上述べたきた諸点から「計算科学研究センター」における生物関連分野への展開は、分子科学分野における共同利用研究者へのサービスを低下させないことを前提に、次のように、短期、長期に分けて展望するのが最善であると思われる。

まず、短期的な展望としては、特に、生物科学と分子科学の境界に位置する分野、すなわち、「生物分子科学」分野における計算科学を発展させる。その理由のひとつは岡崎機構が分子科学研究所と生物関連の 2 研究所を有する国際的にもまれな地理的条件にある点である。このような条件を最大限に活かして、生物学と分子科学との「境界領域」に新しい計算科学の分野を構築していくことは「センター」に課せられた責務であるとさえ言える。もうひとつの理由は歴史的に見て「計算科学研究センター」の前身である「分子研電算センター」がもともと分子科学研究者の共同利用施設であり、計算科学におけるノウハウの蓄積の面でも、人的資源においても分子科学研究所にその大部分の基盤をおいている点である。そもそもスーパーコンピュータのような大型施設を使って行うような計算の多くは非常に大きく、かつ、高度に組織化されたソフトウェアを要求する。量子化学計算で使われる Gaussian や分子シミュレーションで使われる CHARMM などはその代表例である。そのようなソフトウェアはその製作だけでなくその維持、管理においても高度のノウハウと莫大な人的資源を要求し、一朝一夕にして出来上がるものでない。将来、生物独自の情報処理への展開を展望する場合には、まず、生物関連 2 研究所にその基盤となる研究グループを創出し、同時に、計算科学研究センターにおける財政面、人材、ハードウェア、ソフトウェアなどにおいてその分野の拡充を図ることが前提となる。すべての面においてその基盤を分子研電算センター時代の「資源」に依存している現状では、それらの「資源」を最大限に活用できる生物の分野に限定した展開を行うことは極めて自然であり、分子科学と生物の「境界領域」にその可能性を求める所以である。以下に、例として、生物・分子科学境界領域で未解決になっているいくつかの重

要問題について挙げておく。

(1)ゲノム情報から得られるアミノ酸配列に基づく未知蛋白質の立体構造予測。

ゲノム情報は生体内で活躍するすべての蛋白質の配列情報を含むはずであるが、それらの中には、未だ、構造どころか機能も明らかになっていないものが数多く存在する。このような蛋白質の構造を一次配列から予測することはその生体内での役割や生物進化における位置付けなどを明らかにする上で基本的な課題である。このような問題は蛋白質の熱力学的安定性と密接に関係しているが、それを溶媒をも含めて分子レベルで解明することは生物学と分子科学の境界に位置する重要課題である。

(2)生体系における光化学反応過程。

光化学反応はいうまでもなく生物が光エネルギーを物質に固定するもっとも重要な機能であり、すぐれて生物学上の問題である。一方、光の収集や電子移動（酸化・還元）といった機能を担うのは蛋白質という「分子」であり、その全プロセスに「水」というこれまた「分子」が極めて重要な役割を演じている。方法論的には量子化学、統計力学、分子シミュレーションなど分子科学において発展してきた方法が主役を演じる格好の生物学上の舞台だといえる。

(3)神経伝達系におけるイオンチャンネル。

細胞膜内外のイオンの移動は脳を含む神経伝達系におけるもっとも重要なプロセスのひとつであり、その解明は生理学における重要課題である。このプロセスには細胞膜や膜蛋白質の構造と揺らぎ、イオンの個性とその溶媒和などが複雑に関係しており、まさに分子レベルでの解析が本質的な意味をもっている。

この例の他にも、「DNAとペプチドの分子認識」、「受容体とリガンドの結合」、「各種酵素反応」、「人工酵素」などなど、分子科学と生物の「境界領域」に位置する計算科学の重要問題は数え上げればきりがなく、この分野において「計算科学研究センター」が果たすべき役割は極めて重要である。

長期的な展望としては上記の「生物分子科学」および「生物情報」を含む総合的な計算・情報処理拠点の創出を行うため基生研および「情報生物学研究センター」の協力のもとに特に物情報関連の計算・情報処理パワーの大幅な拡充をハード、ソフト、スタッフの面で行う。