

3.1.3 生命科学分野、物質科学分野、ものづくり分野の分野横断連携

生命現象は、ナノメートルの非常に微細なスケールから我々日常生活のスケールに至るまで、非常に幅広いスケールにまたがって存在し、本質的にマルチスケールな対象である。タンパク質やDNAなどに代表される生体分子は、生命科学的側面から見ると「生命」を構成する基本単位に位置づけられるが、物質科学的側面から見ると対称性の少ない非常に複雑な「物質」であり、「生命」と「物質」の二面性を持つ存在である。したがって、生体分子の研究、特にその立体構造に基づく解析は、生命科学と物質科学という大きな2分野の境界に位置する課題であり、それぞれの分野で培ってきた方法を横断的に集約し駆使することで、大きなブレークスルーが期待される。細胞、臓器といった高次の生命現象へのつながりを意識した生命科学的アプローチと、電子や原子・分子、更には分子集団における物質的側面を重視した物質科学的アプローチは相補的であり、この二つを融合することによって、創薬や生体分子を活用したものづくりなど社会的に重要な課題において飛躍的な発展が可能になるであろう。ここでは、生命科学分野、物質科学分野、ものづくり分野の分野横断連携の例として、生体分子・複合体の立体構造に基づくシミュレーション解析を取り上げ、具体的な対象、すなわち、創薬、バイオナノ境界ものづくり、ウイルスや細胞動態などの巨大系のシミュレーションについて詳述する。

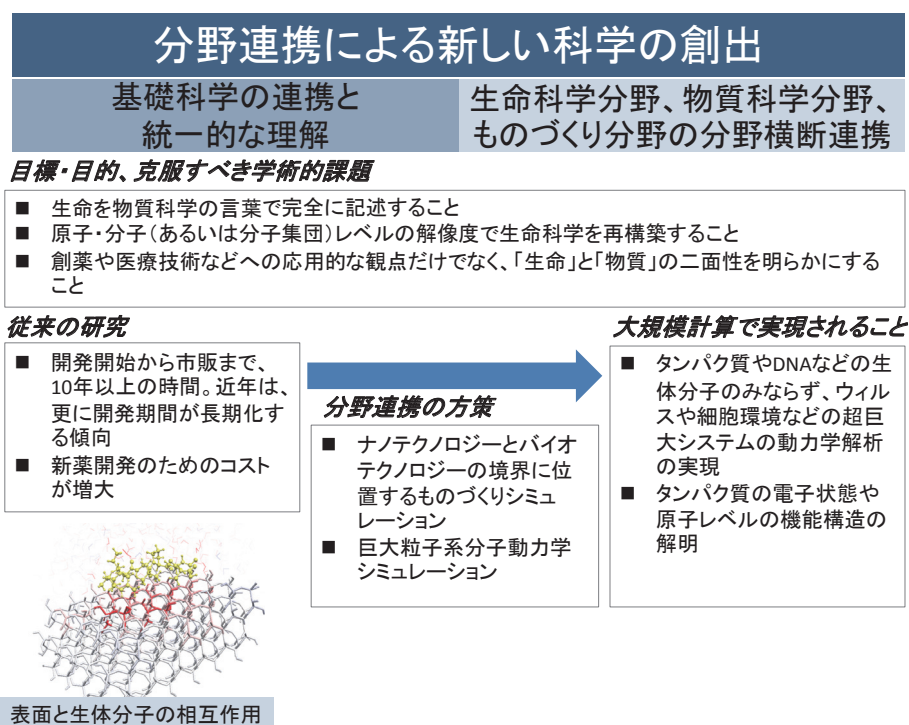


図 3.1.3-1 生命科学分野、物質科学分野、ものづくり分野の分野横断連携

3.1.3.1 計算創薬における連携例

(1) 課題概要

ほとんどの人は「健康で長生きしたい」と願っている。この願いを少しでも実現するために薬は非常に重要であるが、すべての病気に対応できているわけではない。更に、新型インフルエンザのように、人間にとって脅威となる病気は日々増えていく。こうしたことを考えると、今後も多くの薬を開発する必要がある。また、薬の開発手法そのものも効率化・加速化が期待される。

新薬の開発には、経済的な効果も重要である。通常、新薬は開発開始から市販まで10年以上の時間がかかるが、近年はさらに開発期間が長期化する傾向にあり、新薬開発のためのコストが増大している。実際、一つの新薬開発に数百億円規模の投資が必要と言われている。こうした観点からも、薬開発に技術的革新を引き起こすブレークスルーが望まれている。良い薬の市場は1兆円規模である。困難を乗り越え新しい薬を一つでも開発できれば、海外から日本国内にこうした規模の利益が流れ込み、大きな経済効果が見込まれる。こうした戦略のために、日本の新薬開発の国際競争力を高めていく技術革新を強く推進していく必要がある。

上述の技術革新の実現に向けたブレークスルー技術として期待されている新規手法の一つがコンピュータ支援によるドラッグデザイン(CADD: computer aided drug design)である。CADDの代表的手法は定量的構造活性相関であり、1960年代から始まった長い歴史を持つ。1990年以降、X線構造解析やNMR分光によるタンパク質の立体構造が多数解明されたことから、CADDは、立体構造に基づくドラッグデザイン(SBDD: structure based drug design)という新段階を迎えた。SBDDでは、タンパク質の結合部位に医薬品候補化合物を計算機上で結合させ、立体構造とエネルギーに基づいて化合物の結合の強さを評価して分子設計を進めてゆく。この手法の成功例として1997年に開発されたインフルエンザ治療薬であるタミフルがよく知られている。しかし、現在の標準的な手法では結合構造の予測計算やエネルギー計算の信頼性の不十分さから、CADDは限定的な役割に留まっている。しかし、「京」コンピュータが登場し、これまでにない規模の計算資源を利用できるようになったことで、計算精度が大きく改善されている。また、情報科学分野、物質科学分野、生命科学分野で発展した手法が創薬研究のさまざまな課題に活用できる状況が生まれている(ここではすべてを含めて計算創薬と呼ぶ)。このような状況から、今後、計算創薬の効率・信頼度が高まり、イノベーションが起きると考えられる。

計算創薬の研究には、非常に多岐にわたる技術・知識・知見を集約し、組み上げる必要がある。設計する薬が標的とするのは一つのタンパク質の機能であるから、原子レベルでのタンパク質と医薬品の相互作用を解析する必要がある。このように物質科学と見なせる側面がある一方で、タンパク質の機能は細胞内の化学反応ネットワークの一部であり、薬の作用が細胞、臓器、個体へと現れることを考えると、生命科学としての一面もある。現に、物質科学系・生命科学系それぞれで標準的なCADDの手法を超えた計算創薬手法を開発しようとさまざまな基礎・応用研究が進んでいる。しかし、実践の計算創薬で世界をリードしていくためには、両分野の技術を結集していくことが必須である。

現状の計算創薬においては、典型的な新薬の開発プロセス（ここでは SBDD を想定）は、以下のようになっている。

- ① 標的探索：疾患に関連したタンパク質の中から標的とするタンパク質（ターゲットタンパク質）を決める。
- ② タンパク質の構造決定：以下の分子設計や阻害活性予測のために必須であるタンパク質の 3 次元構造を解く。
- ③ 分子設計・化合物ライブラリの作成：タンパク質の構造を参照した化合物のアイデアを創出する。もしくは、化合物探索のための化合物ライブラリを準備する。
- ④ 医薬品候補の選出：阻害活性を定量的に予測することで、医薬品候補を選出する。
- ⑤ 動物実験・臨床試験：細胞や動物による薬効と毒性試験を行い、次いで臨床試験を行う。

これらの研究ステージのうち、①ではバイオインフォマティクス技術による遺伝子レベルでの標的タンパク質の探索が行われている。②では、X 線結晶構造解析に多く依存している。しかし、100 残基程度の天然構造を分子動力学（MD：Molecular Dynamics）計算で再現することが可能になっていることから、実験の構造で精度の悪い部分を計算で補正して、分子設計や候補化合物の選択に利用できるようになっている。

③に関しては、1960 年頃に始まるドッキングと呼ばれる手法が中心的役割を担っている。実際、さまざまなタイプのドッキングソフトが開発されている。ドッキングとは、計算機で化合物ライブラリの化合物をしらみつぶしに標的タンパク質に結合させ、ドッキングスコアと呼ばれる経験的な親和性の指標で医薬品候補を選び出す技術である。また、創薬支援ソフトウェアでは、このドッキングの機能だけでなく、化合物の体内動態や毒性（いわゆる ADMET；Absorption, Distribution, Metabolism, Excretion, Toxicity）を経験的に見積もる機能（Lipinski の法則など）も搭載され、分子設計を支援するための重要な情報を提供できるようになっている。このように、機能的には創薬支援ソフトウェアは広範囲をカバーしているにもかかわらず、現状では創薬研究において補助的、限定的な位置にとどまっている。この理由として、必要な物理量の予測精度が低いことが挙げられる。特に、結合自由エネルギーとドッキングスコアの相関が不十分であることは致命的な欠点であった。④は、このようなドッキングスコアに代わり、定量的に結合自由エネルギーを見積もることで阻害活性を予測しようとする試みである。現在、「京」の戦略分野の重点課題の一つとして創薬応用シミュレーションという課題が立ち上がっているが、これは MP-CAFEE 法という MD 計算に基づく定量的結合自由エネルギー評価法を用いることで、計算機による新しい医薬品候補の選出法を確立することを目的としている。計算創薬が創薬プロセスの中心となるためのコアとなる部分で、世界屈指のスーパーコンピュータの莫大な計算資源を惜しみなく投入することで、イノベーションにつながりつつある。

⑤では実験的なアプローチがいまだ主流であり、新しい計算科学的なアプローチの開発が望まれているところである。これは、ドラッグデリバリの問題として議論されており、全原子レベルの MD シミュレーションだけでは不十分で、粗視化モデルや細胞・臓器・人体に関するマクロスケールのシミュレーションと連成するようなマルチスケールのアプローチが重要であると考えられている。

現在の計算創薬は、未熟で創薬プロセスで中心的役割を果たせていない。これは基盤となる学問分野の理論・計算法の未熟に起因する場合と、計算理論・手法はすでに存在するが、コンピュータの性能が不十分なために現実にあわせて計算が簡便化されていることに起因する場合とが考えられる。計算資源を投入すれば、実践の創薬において中心的な役割を担える計算技術に関しては、「京」コンピュータをはじめとする国内スーパーコンピュータの計算資源をしっかりと投入して、ノウハウの蓄積と新薬開発の中心技術として確立していくことが喫緊の課題であり、この課題を乗り越えることでスーパーコンピュータによるイノベーションへと展開が期待される。また、創薬を加速・効率化できる理論・計算手法に関しても、新たに開発させたり、現存法を成熟させたりしていくことも重要である。常に新しい技術が生まれる環境があって、初めて、日本が常に世界をリードできる医薬品開発力を持つと言える。

(2) サイエンスの質的な変化と長期的目標

現在、「京」の戦略分野 1 の 1 課題として創薬応用シミュレーションが推進されている。これは、「京」コンピュータのようなペタフロップス級の計算機を用いて、計算科学が主体となる創薬プロセスを確立しようとするものである。

こうした計算科学を積極的に利用した創薬は、世界各国が自国の創薬開発力の優位性を獲得すべく注力している。特に、米国の D. E. Shaw らは、MD 計算専用のチップを搭載した計算機 Anton を開発し、通常の汎用計算機の 100 倍以上の速度でタンパク質の MD 計算をすることを可能にしている。彼らは、米国メガファーマと強力なコンソーシアムをつくり、計算機を用いた創薬応用へと乗り出している。マイクロソフト社の Bill Gates もこのコンソーシアムを支持し、巨額の資金援助を行っている。

上述のように、現在、MD 計算を基盤とする結合自由エネルギーの定量的予測方法を中心として、スーパーコンピュータの創薬応用が進んでいる。しかし、それ以外にも将来的に重要になると考えられる方法が生命科学分野・物質科学分野で発展してきている。その中でも特に重要と考えられるのが、タンパク質の電子状態を露わに考慮する QM/MM 法やフラグメント分子軌道 (FMO) 法や Divided-and-Concur (D&C) 法である¹。現在の MD 計算で標準的に用いられている力場は、金属錯体やラジカルのような特殊な原子団のよいモデルを与えておらず、また、化学反応のような動的な化学結合の切り替えを記述できないという欠点が残っている。将来、複雑なタンパク質を標的にする際や新しい機能阻害機構を持つ医薬品開発を目指す際には、こうした特殊な原子団の扱いを避けられない。こうした問題を解決する最も重要な方法が電子状態計算の手法である。

今後 10 年で実験にかかる負担を計算機科学の力で軽くしていくことが可能になっていくだろう。しかし、実験手順を省略できる段階ではない。こうした観点から、広い分野の実験グループとの連携を構築し、実践的計算創薬のノウハウを蓄積し進化させていくことが重要である。

¹QM/MM 法は、重要な部位を量子化学計算、環境としての役割しかない部分は古典的なモデルを使い計算を効率化させる方法である。D&C 法や FMO 法は、領域分割をつなげてタンパク質全体を計算する方法である。領域分割の考え方は並列化に向けたアルゴリズムになる。

次に、今後5年間の進展の見通しについて述べる。①標的探索、②タンパク質の構造決定、⑤動物実験・臨床試験に関しては、4.1節で改めて議論することとし、ここでは、③分子設計・化合物ライブラリの作成および④医薬品候補の選出に焦点を当てる。

現在、「京」コンピュータを利用した戦略分野が開始され、生命科学分野の中で、計算科学主体の医薬品開発を実践段階に移す創薬応用シミュレーションが行われている。現在は、従来法で粗く絞り込んだ数百個の化合物からMP-CAFEE法という定量的結合自由エネルギー評価法を用いて阻害活性のある化合物を探し出すことを行っている。今後5年間で、ペタフロップス級の計算資源を使った技術が確立されていくと考えられる。生命科学分野では、分子設計への重要な知見を与える方法として、統計力学の積分方程式法の一つである3D-RISM法を用いたフラグメント探索についても、研究が進んでおり将来的な応用が期待される。

また、物質科学分野では、以下の課題が進行している。

課題1) シード化合物（薬設計の元となる化合物）の探索範囲の拡大

現状ではバーチャルスクリーニング（計算機を用いて有力な化合物候補を化合物ライブラリーから取り出す技術）は、化合物データベース中の数十万程度の化合物を探索するレベルで行われてきた。ペタフロップス級の計算機を用いれば、1000万化合物を対象に活性化化合物探索を実施することが可能になり、従来知られている化合物群とは全く異なった新規骨格を持つ医薬品候補化合物群が発見される可能性が高い。多くの化合物を同時に計算していくアレイジョブ化が可能であり、高効率化が見込める。

課題2) タンパク質と化合物の複合体構造の予測

タンパク質と化合物間の相互作用を解析するために、タンパク質の3次元構造が正確に解けていることは非常に重要である。現在は、実験構造にほとんど依存しているが、実験構造の解像度不足を量子化学計算（QM/MM法、FMO法、D&C法）レベルでの構造最適化を用いて、高い精度で補うことができるようになると考えられる。

課題3) タンパク質と化合物間の分子間相互作用の解明

タンパク質の特異的な分子認識は多部位かつ同時の分子間相互作用に起因している。したがって、特異性を高める分子設計を行うためには相互作用部位間の協同効果の知見が有用である。ペタフロップス級の計算機を利用することで、タンパク質の複合体全系の量子化学計算（FMO法、D&C法）がルーチンの可能になり、多部位同時の相互作用やタンパク質内部の環境が化合物との結合に及ぼす効果が解明され、医薬品分子設計の指針となるタンパク質の特異的な分子認識機構を理解できるようになる。こうした計算を種々の化合物に適用し、実際の化合物探索を行えるようにしていく。この課題においても多数の化合物について調べるという観点から、単純並列計算で多くの試行錯誤ができる環境が好ましい。

特に、課題2)と3)は電子状態計算に基づくもので、一般に、MD計算で用いられる経験的力場の精度より適用範囲が広い。しかし、計算コストの問題から十分な統計集団に基づく議論はいまだに難しい。こうした点から、将来の計算機の発展にともない重要になる手法であり、ま

た、今後 10 年の間においては、経験的力場に基づく MD 計算と相補的な役割を持つものと考えられる。

今後 10 年後の展望としては、比較的単純な水溶性タンパク質に関しては、MP-CAFEE 法を基盤とした創薬プロセスがさらに実践されるフェーズに入ることが見込まれる。現在では、従来法で絞り込んだ数百の化合物を対象に MP-CAFEE 法を適用しているが、10 年後は数千から数万の化合物を直接的に MP-CAFEE 法で調べられるようになる。このように定量的探索の範囲を広げることで、これまで見落としていた化合物も確実に発見することができるようになる。

更に、計算機資源が増大することで G タンパク質共役受容体 (GPCR) のような膜タンパク質などさらに複雑なタンパク質を標的にした計算創薬が始まる。この重要性は、現在の薬の 50% が GPCR ファミリーを標的としているためであることを考えると自明であろう。2012 年には GPCR の構造解明に対してノーベル化学賞が送られたことが示すように、GPCR に対する医薬品設計の可能性が注目され競争が始まっている。しかしながら、膜タンパク質の結晶化は依然として非常に難しい問題であり、せっかく解かれた結晶構造の中にも精度の悪い部分が存在する可能性がある点が挙げられる。こうした問題は、物質分野やインフォマティクス分野で発展している構造予測法、長時間 MD 計算や拡張アンサンブル法などにより、うまく構造補正を行えると予測される。こうした構造に基づき、定量的自由エネルギー計算を行うことで、複雑なタンパク質の阻害薬を開発することが可能になっていくと思われる。

こうした計算創薬の発展にともない、新しい分子設計の在り方も十分に議論していく必要がある。上述の発展にともない、これまで十分に解き明かせていない阻害活性を上げる物理的条件が見えてくるであろう。このような知見に基づき、化合物の修飾を行うことは論理的ドラッグデザインの目指すべきゴールであろう。現在、さまざまな重要化合物の物性 (pKa や logP など) は化合物中に含まれる原子団の種類に基づき定性的・経験的に見積もられている。もし計算資源が潤沢に与えられるのであれば、これらの物性も高精度量子化学計算に基づき実験を行う前に定量的に予測してしまうことが望ましい。さまざまな計算技術を投入し、創薬プロセスはますます効率化されていくと考えられるが、こうした計算科学が創薬をリードできる時代には、化合物ライブラリの在り方も変わってくると思われる。これまでは、とにかく多くの化合物の種類をカバーすることが重要であったが、代表的な化合物から残り多数の化合物の情報を計算機で予測・派生させる方法を確立できれば、化合物ライブラリに含めるべき化合物の数は圧倒的に少なくなることが見込まれる。

さまざまな種類のタンパク質がターゲットになると、金属錯体やラジカルのような原子団を含むタンパク質が創薬対象として含まれてくると想定される。これら特殊な原子団の力場パラメータが、十分に確立されていないという問題は重要である。例えば、核内受容体タンパク質には亜鉛錯体が多く見られるし、呼吸鎖を形成するタンパク質には鉄イオンを含むヘムや銅-硫黄錯体を持つ。当然、これらのタンパク質に異常が起こると疾病を引き起こす。このような特殊な構造を含むタンパク質を正確に扱うことは、創薬応用シミュレーションにとって重要である。こうした特殊な原子団の取り扱い、物質科学で育まれてきた分子軌道法計算によって解決されていくだろう。

こうした力場パラメータは、時に、環境に大きく依存する。したがって標的タンパク質の環境下でパラメータを決めるのが望ましい。現在、標準的な分子軌道法計算では、巨大なタンパク質をそのまま扱うことはできないが、QM/MM 法や FMO 法が発展してきており、こうした手法を用いることで MD 計算のための高精度パラメータを決定することが考えられる。

従来の医薬品は、分子間力で標的タンパク質に結合することでタンパク質機能を阻害する。それゆえ、古典的経験力場を用いた MD シミュレーションが非常に強力な開発解析ツールであった。しかし、標的タンパク質と化合物の親和性を高めるために、化学結合で複合体形成をさせるという新しいアイデアが考えられている。こうした新しいタイプの医薬品開発を、QM/MM 法や FMO 法でタンパク質内化学反応を直接的に計算・解析することで強力に支援することが可能になる。

また、従来のドラッグデザインは安定構造のタンパク質を標的としているが、反応経路にそって変形した構造をターゲットとした医薬品分子設計の可能性を提供する。これは、キナーゼなど大きなタンパク質ファミリーで、疾患に関係する特定のキナーゼに対する特異性を高める分子設計に有用な可能性がある。

3.1.3.2 ナノバイオ境界ものづくりシミュレーションにおける連携

(1) 課題概要

ナノテクノロジーとバイオテクノロジーの境界は、このところ高い注目を集めて活発に研究開発が展開されている領域であり、バイオセンサーや安全性の高いインプラントの調製、バイオミネラライゼーション²に基づくナノ構造・粒子の形成などが含まれている。端緒となったのは、米国国立がん研究所（National Cancer Institute）の Brown によって 1992 年に報告された、酸化鉄表面への特異的結合能を持つペプチドの遺伝子工学による創成である。また、半導体表面ではマサチューセッツ工科大学（MIT）の Belcher がガリウムヒ素で先鞭を付け、酸化チタン表面やナノチューブ系では日本でもがん研究所の芝らが先導的な研究を行ってきている。

ナノバイオの研究開発は、遺伝子工学では“総当たり探索”が比較的容易なこともあり、これまでは試行錯誤的な要素が多かったように見える。しかし、今後はシミュレーションに基づく生体分子と固体表面の界面での相互作用の原子レベルあるいは電子状態レベルでの把握が重要になってくると思われる。ただし、計算への要求は必然的に高くなる。古典力学に基づく分子動力学（MD）計算では、金属などの異種原子を含む力場セットの新規準備は不可欠であるが、複雑な界面での電子移動や広域での分極の記述あるいは化学反応の扱いを考えると、タンパク質の長時間の動的挙動の把握などを別にすれば、適用可能な問題はむしろ限られるかもしれない。一方、量子論系の密度汎関数（DFT）法では周期境界条件を課した計算を行うが、周期的に隣接する生体分子間の相互作用が十分に小さくなるように単位格子の大きさを設定する必要があることもあり、タンパク質の規模によっては非常に負荷の重い計算になってしまう。規模の大きい対象に対しては、低スケーリング系（いわゆる $O(N)$ の手法）の DFT プログラムの開発・利用がこれまで以上に必要とされる。分子軌道計算においても、FMO 法などの

²生物が鉱物を作る作用のこと

分割&統合系の方法に期待がかかるが、タンパク質では実応用も含めて成功を収めてきたこれらの手法も固体系への適用は緒についたばかりである。いずれにせよ、単一の計算手法ですべての問題をカバーすることは難しい。どの方法を使うにせよ、ペタ~エクサ級の計算資源があってこそ実用的になると思われる。

固体界面の問題ではないが、生体分子の精密分光も今後の発展が望まれる分野である。例えば、GFPなどの蛍光タンパク質を使ってがんなどの病変部を特定することはすでに広く用いられているが、生体組織透過性の高い赤外域の蛍光を使う試みがなされている。また、2次高調波発生(SHG)などの非線形応答に基づく組織の可視化技術も開発されている。さらに最近では、振動分光の一種であるテラヘルツ波を用いた非侵襲型の測定をタンパク質の畳込みの解析、基質・タンパク質の相互作用解析に応用しようとする試みが報告されている。これらの分光スペクトルの解析にはタンパク質の量子計算が必要不可欠である。

ここでは、DFT系の手法、MO系の手法によるナノバイオ境界問題の扱い、ならびに生体分子のテラヘルツ分光について概説する。

(i) 密度汎関数(DFT)法によるタンパク質 —表面反応の計算—

吸着効果(吸着の強さや吸着にともなうタンパクやペプチドの構造変化、吸着ペプチド間の相互作用の強さなど)は溶媒の種類やpHにも敏感である。吸着させる材料とペプチドの種類、溶媒環境などの変数の組み合わせは膨大であり、試行錯誤によって探索できるのは極一部の組み合わせでしかない。そのため、シミュレーションによる探索の加速が期待されている。シミュレーション手法は大きく分けて、①量子論的に行うもの、②経験的力場を使って行うもの(MM、MC、MDなど)、③粗視化手法により行うもの、の三つがある。②あるいは③のシミュレーションは、創薬の分野で盛んに使われている。①、②、③の順に計算負荷が重い、③、②、①の順により長時間の動力学計算が可能である。①の方法には、量子化学的手法と、主として固体物理に適用されてきたDFTに基づいた量子論的手法がある。DFTに基づくシミュレーションは材料研究において果たす役割がますます大きくなっており、現在もさまざまな改良を施されながら発達している。最近では電池材料中の電極・電解質界面における反応計算などに適用対象を広げている。計算負荷が重いため、長時間の動力学計算を行って結合自由エネルギーを求めたり、大きな分子量のタンパク質が材料表面上にどのように吸着するかを求めることは難しい。一方、②あるいは③の方法は経験論的パラメータに依存しており、表面、タンパク、および溶媒の組み合わせごとに最適なパラメータセットを用意しないと信頼性のある結果が得られない。また、そういったパラメータの組をつくるのは表面の反応系では難しい。

(ii) 分子軌道法による表面 —ペプチド相互作用解析—

分子軌道(MO)法のスキームでは、基底状態を扱う場合、基本のハートリーフォック(HF)計算によって軌道を決めた後、2次摂動論から結合クラスター展開に至る階級的な近似法によって電子相関を導入して計算の信頼性を向上させる。その前提は、十分な広さのバンドギャップがあることで、DFT法に比せば、クラスターモデルを使うにせよ固体側の取り扱い種には自ずと限界がある。他方、ペプチド・タンパク質やDNAなどの生体分子では静電相互作用や水素結合とともに分散力による弱い相互作用が重要である。分散力の寄与は2電子励起で記述さ

れるが、摂動論系の相関計算では直接かつ適切に取り込める。更に、励起状態についても相関を含めた配置間の相互作用的な取り扱いが可能で、界面での電荷移動励起についても処方箋はある。

これまでのところ、MO 計算によってナノバイオの境界領域の問題を扱った例はきわめて限定的で、摂動論の代わりに相関を実効的に取り込める DFT を用い、シリカやアパタイトを対象に小規模モデルでの計算が報告されている程度である。上記のバンドギャップの問題もあるが、きちんと実問題を扱おうとすると、生体分子そのものの大きさと人工的な周縁効果を見捨てるほどのサイズのクラスターモデルの大きさがともに課題となる。この二つの大きさ要求を水和条件を課しつつ満たすことは、通常のアプローチでは計算コスト的にほぼ不可能である。そこで、並列処理と高相性の FMO に代表されるフラグメント分割&統合系の手法の適用が強く求められる。FMO 計算では、フラグメント間の相互作用エネルギーが得られるため、単なる高速化の手段ではなく対象系の内部解析ツールとしての利用もでき、特に創薬分野で好んで使われてきていることは本書に既述のとおりである。FMO 法を固体系に適用する試みは始まっており、GAMESS では軌道緩和に制限を課した 2 体のやり方で、一方の ABINIT-MP(X) では制限無しでフラグメント展開を 4 体まで進めたやり方 (FMO4) による。FMO4 では、コストは相対的に AFO よりも大きくはなるが、エネルギーの精度が適切に担保され得るメリットがあり、超並列処理の条件下で本質的なアドバンテージが出てくると思われる。

ナノバイオ分野の人造ペプチドでは、特定のアミノ酸残基が固体表面の認識に重要であることが議論されるのが通例である。例えば、芝らの酸化チタン結合タンパク質 mini TBP-1 では正荷電を帯びたアルギニンと負荷電を持ったアスパラギン酸が表面の荷電ポイントと相補的に引き合うことが表面への固着能を与えると実験を基に論じられている。しかしこれだけでは、アミノ酸残基と表面の活性ポイントとの相互作用の描像は不明であり、荷電性でない他残基の役割や水和の効果も含めた詳細な理解があれば、ペプチドの最適化デザインには有益である。こうした目的には、相互作用解析に長けた FMO 計算はまさに好適である。固体側としては、有バンドギャップが条件であることからシリカ、炭酸カルシウム、アパタイトなどが考えられるが、タンパク質を用いた次世代デバイスやバイオセンサー、ドラッグデリバリーや極微量抗体検知用のナノ粒子の模倣バイオミネラリゼーション、生体親和性の高いインプラント（人造の骨や歯）、逆に細菌汚染の足場となるバイオフィルム形成を阻害する清浄表面などの研究開発をカバーできると考えられる。

(iii) 生体分子分光

量子力学と分光学は現代の物質科学を支える二つの主要な柱であり、固体物理学においては、その相互連携の重要性は十分認識されている。本節では、ナノバイオ分野における量子力学と分光学の連携を確立することを視野に記述する。

生体分光は、生体分子の挙動を分子レベルで観測することを可能とすることから、今後の発展が望まれる分野である。これまでに、可視分光、非線形光学分光、核磁気共鳴スペクトル法等の分光技術が生体分子の計測に用いられ、タンパク質をはじめとする生体分子の分子構造ならびに電子状態の解析に適用されている。これらの分光スペクトルは、基本的に分子の電子状態に起因するので、そのスペクトルの解析は、分子の電子状態を解析することにより可能であ

る。しかし、生体分子、特に水中の生体分子の電子状態には、溶媒である水の電子状態が関与するため、そのスペクトルの解析には水中での生体分子の電子状態を知ることが必要である。水中における分子の電子状態解析には、これまで水を誘電体として近似する平均場近似が用いられてきた。この近似では、水を連続的な誘電体と近似し、水と生体分子の相互作用を分子と誘電体として近似された水の間の静電相互作用として記述する。しかし、実際の水と生体分子の相互作用は、必ずしも静電相互作用のみで記述できるものではなく、水素結合ならびに電荷移動相互作用等の量子力学的な相互作用が重要な働きをする。したがって、水溶液中での生体分子の分光スペクトル解析には、水と生体分子からなる系の量子状態を明らかにし、その量子状態より生体分子のスペクトルを解析する技術を開発することが必要である。

生体分光のもう一つの大きな課題に、テラヘルツ分光がある。テラヘルツ分光は遠赤外光を用いる赤外分光の一種である。生体分子は巨大分子の一種であり、その骨格振動に特異的な振動を有する。テラヘルツ分光は、この骨格振動にともなう電磁波吸収を計測することにより、生体分子の種類、構造、機能を明らかにする。例えば、タンパク質と基質の相互作用系の骨格振動をテラヘルツ分光法により解析することにより、相互作用の強さを分光学的に評価する試みがなされている。しかし、現状のテラヘルツ分光法には、実測されたスペクトルを解析することができないという問題がある。低分子に関しては、分子軌道計算により分子の振動準位を解析した例があるが、タンパク質をはじめとする生体分子の骨格振動を量子力学的に解析した例はない。またテラヘルツスペクトルは、一般に水の影響を強く受けるので、水中での生体分子のテラヘルツスペクトルの解析には、水分子の存在を露わに考慮した分子の骨格振動計算が必要である。密度汎関数法に基づく水中生体分子の電子状態計算ならびにそれに基づく骨格振動計算は、これまでの不可能であった生体分子のテラヘルツスペクトルの分子論的な解析を可能とするものであり、テラヘルツ分光学および生体分子の研究を行ううえで、重要な寄与をすることが期待できる。

(2) サイエンスの質的变化と長期的目標

(i) 密度汎関数 (DFT) 法によるタンパク質 —表面反応の計算—

今後、DFT 計算は次のように利用されていくと考えられる。

- (1) 生体分光や STM、ESR などの実験観測と照応させながら、材料表面におけるペプチドなどの吸着構造を同定する。
- (2) 材料の表面構造を求めるのに用いる (溶媒中)。
- (3) 材料表面に吸着するペプチドの安定構造を探索するのに用いる。
- (4) 吸着構造探索や自由エネルギー計算は前述の②、③の方法で行うが、そのための経験的力場パラメータの高精度化は DFT 計算を行って得た結果を用いて行う。
- (5) 自由エネルギー計算自体も比較的小さいペプチドであれば、DFT 計算を用いて行う。DFT 計算は電子間クーロン相互作用の高精度化 (ハイブリッド汎関数など) やファンデルワールス相互作用のような分散力を取り入れた計算が必要になる。

今後は、高精度・高効率化のための手法の開発とともに高速・高並列計算機の利用が必須である。また、DFT 計算に適用できる加速分子動力学法（Accelerated MD 法。メタダイナミクス法もそのうちの一つ）の開発も構造最適化を加速するであろう。

(ii) 分子軌道法による表面 ―ペプチド相互作用解析―

現時点での計算資源では、大型のクラスターモデルを組み、それにタンパク質を合わせる FMO 計算は単一から少数の構造サンプルでの実行に限られそうである（mini TBP-1 と千原子程度のシリカクラスターとの相互作用解析の先導的な研究事例は有り）。しかし、エクサ級の超並列機であれば統計的な算定が可能な数の計算が容易に処理できるので、シミュレーションとしての信頼性が向上する。また、多種のタンパク質を同時並行的に流して得られる多様なデータを解析して新しい最適化指針を得ることも意図され得るが、実用性という点ではこうした“アレイジョブ”という形態も有効であろう。ルシフェラーゼ（550 残基）とシリカの相互作用をモデル化する場合、基板側が 1 万個以上の原子を含むクラスターを使う必要があり、1 サンプル構造としても上記 mini TBP-1 モデルの FMO4 計算の百倍程のコストを要し、複数構造を高速処理するにはエクサ級の計算機の利用が望まれる。計算法の改良やプログラムの整備は不可欠であるが、ナノバイオの境界問題は FMO 法にとっても開拓すべきフロンティアを提供してくるものと期待している。

(iii) 生体分子分光

密度汎関数 (DFT) 法による数百から数千の水分子と生体分子からなる系の量子状態計算は、水中の生体分子の挙動を明らかにするとともに生体分光技術を発展させるうえで重要である。このような計算は必然的に数万原子規模の量子計算となり、現行のペタフロップス級を超えるエクサフロップス級のスーパーコンピュータを必要とする。エクサフロップス級スーパーコンピュータによる数万原子規模の量子計算は、それ自身が計算科学上重要なテーマであり、生体分子の分光スペクトル解析は計算科学に及ぼすインパクトが大きい。

3.1.3.3 ウイルスや細胞動態などの巨大系シミュレーションにおける連携例

(1) 課題概要

MD 計算は、液体や気体、固体、ガラスや超臨界流体といった物質の種々の状態のミクロな性質からマクロな物性に至るまで、分子間相互作用から力学・統計力学法則をもとに明らかにする「物質科学」に起源をもつ研究手法である。これまで MD 計算はタンパク質や DNA、脂質分子、あるいは膜タンパク質、生体膜やウイルスといった生体分子系に適用され、その物質的側面についての情報を与えてきた。しかしながら、これらの研究は生命科学との連携が薄く、より高次の細胞や臓器といったレベルのシミュレーションとの連携が十分になされてきたとは言いがたい。今後「物質科学」分野の方法論や物理・化学的アプローチと「生命科学」分野の生命現象に直結したシミュレーションとの連成による分野横断的な研究テーマが待ち望まれる。

そもそも生体分子は、溶媒である水やイオンと共存する環境の下で、弱い秩序構造を保ちつつ柔軟に変形可能な巨大分子集合体であり、熱揺らぎ程度のエネルギーを利用しつつ機能を発現させている。したがって生体分子のみならず溶媒やイオンの自由度についても露わに取り扱う、正確な MD 計算が不可欠である。そのためには高い精度の力場が必要であり、すでに開発されている AMBER (米)・CHARMM (米)・GROMOS (蘭) などとともに、タンパク質分子については更なる改良が加えられ精度が向上している。また巨大分子集団系の構造やその揺らぎをシミュレートするためには、溶媒も含めて莫大な数の原子を含む系を高速で計算する必要があるが、大規模系に適した MD 計算のアプリケーションについても、AMBER (米)・CHARMM (米)・GROMACS (蘭)・NAMD (米)・LAMMPS (米) が広く知られている。また国内でも精力的に開発が進められており、Marble、Modylas、Genesis といったアプリケーションが「京」において高い性能を発揮している。さらに系の熱力学的安定性の評価のための自由エネルギー計算については、熱力学的積分法、自由エネルギー摂動法、さまざまな拡張アンサンブル法に加え、溶媒和エネルギーを高速に計算可能なエネルギー表示法も提案され、すでに生体分子系に対してもその有効性が示されている。

(2) サイエンスの質的变化と長期的目標

(i) 生体膜のシミュレーション

生体膜の MD 計算は、1980 年代以降に全原子モデルによる単一脂質組成の系を中心に二分子膜構造の一部を切り出した、平面膜のパッチ構造について広く行われてきた。現在では、生体系と同一の複雑な脂質組成のリアルなモデル生体膜の全原子 MD 計算とともに、リポソーム（球殻状に閉じた膜構造を有する小胞）のような巨大系を丸ごと対象とするために脂質分子に粗視化モデルを用いたシミュレーションへと展開が図られている。前者は生体内の器官ごとの生体膜物性の違いを明らかにすることが可能であり、後者は平面膜のみならずリポソーム全体の物性評価を行える。これらにより得られる知見を高次の細胞シミュレーションへと連成させることにより、細胞環境を模倣しその力学構造をより正確に再現することが可能になるものと期待される。

また、リポソームはミセルなどとともにドラッグデリバリーシステムの搬送媒体として薬学的関心も高い。リポソームのカプセルとしての安定性や、搬送薬物や細胞膜との親和性の評価のためには、数千万原子～一億原子規模のリポソーム全体の MD 計算が不可欠である。また得られる知見は薬物の体内動態シミュレーションの高精度化に寄与できるものと期待される。

(ii) ウイルスの全原子シミュレーション

ウイルスの立体構造に関する実験的情報の蓄積、高精度化と高性能計算機の発展が相まって、タンパク質分子の複合体であるウイルスの全原子シミュレーションが計算科学の俎上に載ろうとしている。これにより、ウイルスの水溶液中における温度や圧力、pH、塩濃度などの熱力学条件や化学環境に対する安定性に加え、構造ゆらぎや弾性といった物理的性質の評価が行えるようになる。また、ウイルスと細胞表面のレセプターとの相互作用について自由エネルギー解析を行うことにより、レセプターのウイルス認識という生命科学的現象について物理化学的に取り扱えるようになる。このような物質科学的アプローチによって得られた微視的情報に基

づいて、ウイルスの細胞感染過程の高精度なモデル構築が可能となり、細胞レベルのシミュレーションとの連携が図れるものと期待される。

更に、細胞表面にあるタンパク質や糖質からなるレセプターをウイルスがどのように認識し選択的に結合するのか、あるいはウイルス表面上のタンパク質がウイルスと細胞との接合部位をどのように認識して切断するのかといった分子レベルの情報が MD 計算によりもたらされる。これらをもとにしたウイルス感染や増殖を阻害する抗ウイルス薬の開発、あるいは新規の作用機序の開拓を通して、創薬への貢献も期待できる。

加えて、表面にエンベロープと呼ばれる脂質膜構造を有する、さらに巨大な HIV やインフルエンザウイルスなどの病原性の高いウイルスへの展開が必要である。これらは原子数 10 億程度、数十マイクロ秒からサブミリ秒の統計が必要となるため、エクサスケールのコンピュータの登場が待たれる。

(iii) 細胞環境のシミュレーション

原子粒度での生体分子シミュレーションと、分子粒度での細胞スケールシミュレーションの間にはいまだに大きなギャップが存在する。細胞スケールのシミュレーションの多くは空間解像度を持たない反応拡散モデルによる記述を行っているが、近年 pSpatocyte のように空間解像度を持つ粒子モデルに基づく細胞スケールシミュレータも開発されてきた。このことは、近い将来に原子粒度の生体分子シミュレーションと pSpatocyte などの細胞シミュレータが連携し、分子・細胞スケールシミュレーションが実現する可能性が高いことを示唆している。実際、細胞質内の分子混雑環境 (Molecular Crowding や Confinement) でのタンパク質の水和、安定性、分子認識などに関するシミュレーションが行われるようになってきた。さらに計算機の能力が発展するならば、現在はゲノム解析や反応拡散モデルで記述されている細胞質内の反応ネットワークが、タンパク質の立体構造を用いたシミュレーションを通して理解されるようになると期待される。また、細胞内環境を記述するためには、タンパク質や RNA、DNA などの生体高分子のみならず、混雑物や微小管などの細胞骨格をつかさどる巨大生体超分子の導入も必要である。ターゲットとする細胞の種類によってその複雑さは異なるため、初めはすべてのゲノムが記述されているバクテリアなどから理解が進み、分子細胞生物学でモデル細胞としてこれまでよく用いられてきた大腸菌などの記述が進むと期待される。更に人の細胞に関する記述が実現すれば、疾患などの解明に本質的な役割を果たすことができる。計算科学的な技術としては、原子粒度の分子モデルだけでなく粗視化分子モデルを組み合わせたマルチスケールな分子モデルの記述が必要であるとともに、タンパク質とタンパク質の反応をどのように記述するかが大きな問題となる。また、細胞全体を記述するためには従来の構造生物学の枠組みにとらわれない新しい実験的手法が必須であるが、XFEL などを実現する生体系の単粒子解析は細胞スケールのシミュレーションと組み合わせることで、細胞の物理化学的記述を飛躍的に進歩させる可能性を秘めている。

3.1.3.4 コミュニティからの意見

2013年6月12日～14日に開催された第13回日本蛋白質科学会年会にて、「エクサフロップス時代の計算蛋白質科学」と題したワークショップを開催し、意見交換を行った。そこでは、以下に示すような意見が出た。

- 従来の浮動小数点演算性能を重視する HPC だけでなく、I/O 性能や整数演算性能も重視すべき（バイオインフォマティクス分野より）
- アプリケーションに関しては、GPU などのアクセラレータの利用促進や、ソフトウェアの共有化、公開
- 創薬など産業界との関係についても、今後、ますます発展させるべき
- タンパク質については酵素反応が重要な点の一つであるので、量子化学計算など物質科学で発展してきた方法を更に取り入れて、研究を行っていったほうがよい
- 今後想定される、更なる計算の高速化に対応して、大量に出力されるデータの解析法の充実も望まれる
- 次世代研究者育成について、HPC 用プログラミング方法などの教育プログラムの充実
- 生体分子分野で HPC の技術を持つ研究者の雇用先の拡充

2013年6月27日に、製薬企業関係者が多数集まる情報計算法学（CBI）学会の研究講演会にて、計算創薬の展望についての意見交換を行った。

- 放射光施設など大型実験施設との関係
- これから更なる並列数増加への対応
- ものづくり分野との連携

についての意見が出された。

3.1.3.5 必要な計算機資源

詳細は 2.1 節の創薬・医療、4.1 節の生命科学、4.2 節の物質科学を参照のこと。

課題	要求性能 (PFLOPS)	要求メモリ リバンド幅 (PB/s)	メモリ量 /ケース (PB)	ストレージ量 /ケース (PB)	計算時間 /ケース (hour)	ケース数	総演算量 (EFLOP)	概要と計算手法	問題規模	備考
創薬などMD・自由エネルギー計算	1000	400	0.0001	0	0.0012	1000000	4300000	全原子分子動力学シミュレーション	ケース数: 10万化合物×10標的蛋白質(10万原子程度)	B/F=0.4. 数百から数千ケース同時に実行することを想定している。実行時に必要な全メモリ量、各ケースの実際の実計算時間は、表の値の数百～数千倍となる。メモリ量/ケースは100ノード実行時を想定。
高精度創薬	0.83	0.1	1	0.001	1	100	300	薬品とタンパク質間相互作用の量子化学計算を統計的ゆらぎを含めた複数サンプルで行う	フラグメント分子軌道法で～500残基程度までの計算	計算要求は「物質科学」のフラグメント分子軌道法のところを参照
バイオデバイス設計	1.1	0.2	1	0.001	1	100	400	200～500残基程度のタンパク質の分光計算	電子軌道数10万超	計算要求は「物質科学」のフラグメント分子軌道法のところを参照
細胞環境・ウィルス	490	49	0.2	1.2	48	10	850000	全原子/粗視化分子動力学シミュレーション	～1億粒子	B/F=0.1