

計算化学アプリケーションパッケージと 蛋白質解析に向けた取組み

Software for Computational Chemistry Application Packages and Protein Analysis

あらまし

バイオ実験技術の進歩により、ゲノム情報や、様々な蛋白質のアミノ酸配列情報、構造情報が決定され、蛋白質の構造に基づく薬物の設計が行われるようになってきた。そして、蛋白質の様々な情報を計算化学の手法によって求め、それらを薬物設計に利用したいという要望も高まってきている。

本稿では、富士通が開発している計算化学アプリケーションパッケージMOPACおよびMASPHYCについて紹介し、MOPACを用いて蛋白質全体を分子軌道法で計算することによって得られた新たな知見や、MASPHYCによる蛋白質・水系の分子動力学シミュレーション、およびこれらの計算化学アプリケーション技術を創薬研究に有効活用するためのin silico screeningシステム開発への取組みについて紹介する。

Abstract

Recent advances in biotechnology have made it possible to determine the genome sequences and the amino-acid sequences and structural information of various proteins. As a result, more and more drugs are being designed based on the structures of proteins. Requests for information about proteins using the method of computational chemistry and the use of this information to design drugs have also increased. In this paper, we introduce the software of the MOPAC and MASPHYC packages for computational chemistry applications. Next, we introduce some of the new knowledge we have obtained by applying the molecular orbital method to the general protein system using MOPAC and introduce a molecular dynamics simulation of the protein-water system using MASPHYC. Finally, we introduce our activities in the development of the in silico screening system to effectively use this computational chemistry application technology in drug research.



紙谷 希 (かみや のぞむ)
ライフサイエンス推進室 所属
現在、計算化学アプリケーション
パッケージの開発に従事。



牟田 元 (むた はじめ)
ライフサイエンス推進室 所属
現在、計算化学アプリケーション
パッケージの開発に従事。



高橋篤也 (たかはし あつや)
ライフサイエンス推進室 所属
現在、計算化学アプリケーション
パッケージの開発に従事。

ま え が き

最近のバイオ実験技術の急速な進歩により、ゲノム情報や、様々な蛋白質のアミノ酸配列^(注1)情報、構造情報が決定されてきている。そのため従来は不可能であった、蛋白質の構造に基づく薬物の設計が行われるようになってきた。それとともに、蛋白質の様々な情報を計算化学の手法によって求め、それらを薬物設計に利用したいという要望も高まってきている。

富士通では、10年以上にわたり半経験的分子軌道法プログラムMOPACや、分子動力学法プログラムMASPHYCなどの計算化学アプリケーションパッケージを開発、販売してきており、現在は上記のような蛋白質構造に基づく薬物設計分野における要望に応えるべく、アルゴリズムの改良や精度向上のためのパラメタ改良などを行っている。

本稿では、富士通の計算化学アプリケーションパッケージによる蛋白質系解析に向けた取組みについて紹介する。

計算化学アプリケーションパッケージ

富士通で開発、販売している計算化学アプリケーションパッケージMOPACおよびMASPHYCについて、以下に紹介する。

MOPAC

分子軌道法は、分子に関するSchrödingerの波動方程式を解くことにより、様々な物理化学的特性（分子軌道、原子電荷、生成熱、双極子モーメント、反応熱、遷移状態、相互作用エネルギー）を計算によって求めることが可能な手法である。

MOPACは、世界で最も広く使われている半経験的分子軌道法アプリケーションパッケージである。高速分子軌道計算アルゴリズムMOZYME法⁽¹⁾の実装により、蛋白質などの巨大分子系の計算を可能としている。

MOPACを用いることにより、蛋白質の電子状態、蛋白質中の個々のアミノ酸残基^(注2)が持つ電荷、蛋白質-リガンド^(注3)相互作用エネルギー値の評価などが可能であ

(注1) アミノ酸配列：蛋白質は、アミノ酸のアミノ基とカルボキシル基間の脱水縮合から作られるペプチド結合の繰り返しによる高分子である。

(注2) アミノ酸残基：元のアミノ酸から水分子が1個とれた形のアミノ酸をアミノ酸残基といい、蛋白質はアミノ酸残基の配列で表現することができる。

(注3) リガンド (ligand)：「結合するもの」という意味であり、無機化学では金属錯化合物において中心金属原子に結合している原子団のことをいい、生化学では例えば酵素に対する基質や補酵素

り、バイオケミカル分野における様々な応用が考えられる。

MASPHYC

分子動力学法は、原子間に働く相互作用下でのNewtonの運動方程式を解くことによって、分子を構成する原子の位置の時間変化を与える。

MASPHYCは、富士通が開発し1992年に製品化された分子動力学法アプリケーションパッケージである。様々なポテンシャル関数、パラメタを装備し、金属、セラミックス、液晶、脂質二重層膜^(注4)など、様々な物質系への適用を可能としている。

MASPHYCを用いることにより、蛋白質のフォールディング^(注5)やアンフォールディング、水溶液中における蛋白質の立体構造とダイナミクス（構造の揺らぎ）、エネルギーなどの情報を得ることが可能である。

MOPACの蛋白質系への適用

従来、蛋白質系の分子軌道法計算では、系に含まれる原子数が数千～数十万原子と、通常の分子軌道法で扱えるサイズ（～数百原子）を大きく超えていたため、興味のある一部分の構造を切り出した小さなモデル系で近似せざるを得なかったが、MOPACに搭載のMOZYME法^(注6)により、数万原子程度であれば蛋白質系全体を実用的な時間で計算することが可能となった。蛋白質全体の電子状態をMOPACを用いて求めることにより、従来の部分構造モデル系では得られなかった新たな知見を得ることができるようになってきている。

以下で、これまでに発表されているMOPACを用いた蛋白質系の計算事例を紹介する。

(1) 酵素の活性部位とフロンティア軌道

分子軌道法の計算では、いくつかのエネルギー準位を持つ分子軌道が得られる。電子はその軌道をエネルギー準位の低い方から二つずつ対になって入っていく（占有軌道）、残りの軌道は電子が詰まっていない（空軌道）という絵を描くことができる（図-1）。このとき、占有

をリガンドという。

(注4) 脂質二重層膜：細胞や細胞内小器官と外界を区切る生体膜の構造のことで、生体膜は脂質二重層膜に蛋白質がモザイク状にはまりこんだ構造をしている。

(注5) 蛋白質のフォールディング：高分子鎖が、非共有結合性の相互作用により折りたたまって、3次元的な立体構造を形成することをフォールディング（折りたたみ）という。水溶液中の蛋白質に酸やアルカリを加えると、立体構造が壊れてランダムな主鎖構造になるが、この状態から生理的環境に戻すと、自発的に再びもとの天然状態の構造を形成する。これを蛋白質のフォールディングという。

(注6) MOZYME法：生体高分子などの大規模分子系のHartree-Fock方程式を解くためのアルゴリズムで、局在化分子軌道を基底関数として、占有軌道-非占有軌道間のFock行列要素を消去する解法。

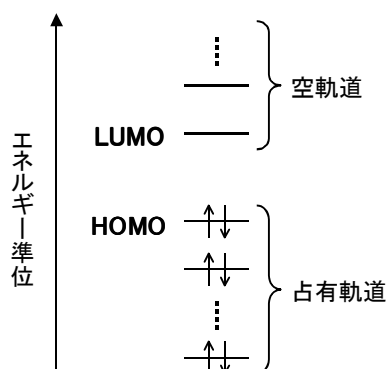


図-1 分子軌道の概念図(上向き, 下向きの矢印が, それぞれ上向きスピン, 下向きスピンの電子を表す)

Fig.1-Conceptual diagram of molecular orbital. (An upward arrow and a downward arrow express the electron of an upward spin and a downward spin, respectively.)

軌道のうちで一番エネルギーの高い軌道をHOMO (Highest Occupied Molecular Orbital), および空軌道のうちで一番エネルギーの低い軌道をLUMO (Lowest Unoccupied Molecular Orbital) と呼ぶ。これらの軌道は, ほかの軌道の最前線 (Frontier) に立って, 化学反応などの種々の化学的現象に重要な役割を果たすことから, フロンティア軌道と呼ばれている。このフロンティア軌道理論は, 1951年に福井謙一博士によって提唱され, その有用性から世界的な注目を集め, この功績により1981年にノーベル化学賞を受賞した。

フロンティア軌道理論は主に低分子間の化学反応でその重要性が認識されてきたが, 最近MOPACを用いて, 酵素蛋白質で化学反応が起こる活性部位に, 化学反応と関わりの深いフロンティア軌道が局在化しているという興味深い知見が見出されている²⁾。このことは, 蛋白質の構造と活性部位との関連からも非常に興味のあることであり, 蛋白質全体を分子軌道法で計算することの有用性を示している。さらに, これらのフロンティア軌道は, 蛋白質の周りに存在する水の効果を取り入れて初めて活性部位に局在化することが見いだされている。生物の体はその多くが水であり, 蛋白質も水の存在下で機能し, 有機溶媒などの非水系ではその機能が失われることが知られている。この計算結果は上記の事実と符合するものであり, 蛋白質の機能は溶媒としての水の役割が重要であることを示唆した非常に興味深い結果であると言える。

(2) バクテリオロドプシン中のレチナル分子の紫外・可視吸収スペクトル

ロドプシンという蛋白質は, 内部にビタミンAに類似した構造を持つレチナルという分子を持ち, 光によ

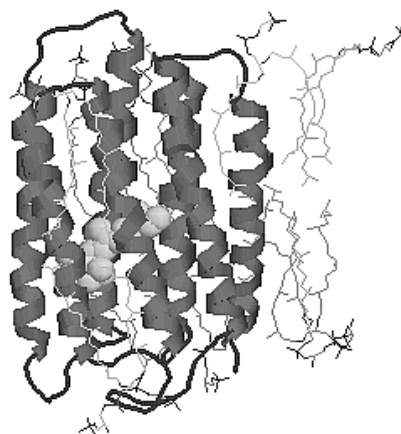


図-2 バクテリオロドプシンの構造(濃く表示されている蛋白質部分の中に, 球状モデルで示されたレチナル分子が存在している。細線は細胞膜の構成成分であるリン脂質)

Fig.2-Structure of bacteriorhodopsin. (The retinal molecule shown by the spherical model exists in the protein portion which is displayed deeply. A thin line is phosphatide which is the composition ingredient of a cell membrane.)

てレチナルの分子構造が変化するため, 光応答性を示す蛋白質として知られている。このロドプシンは人間の視神経の細胞にも存在し, 我々が物を見るために重要な機能を果たしている蛋白質である。

ロドプシンの一種であるバクテリオロドプシンは, 高度好塩菌と呼ばれるバクテリアの持つ紫色の色素蛋白質で細胞膜部分に存在し, 光によって生体膜を介したプロトン輸送を駆動する機能などを有する光受容蛋白質として知られている(図-2)。Halobacterium salinarum由来のバクテリオロドプシン中のレチナル分子の光吸収波長は568 nm (紫色)であり, レチナル分子単独の値440 nm (黄色)から, 蛋白質中に入ることによって顕著に長波長化することが見いだされている。この現象はオプシシフト (opsin shift) と呼ばれ, そのメカニズムに大きな興味を持たれてきた。

このバクテリオロドプシンに対して, レチナルの周りの蛋白質の効果をMOPACを用いて取り込むことにより, オプシシフトをよりよく説明することができることが最近になって見いだされた³⁾。この研究では, MOPACを用いたバクテリオロドプシンの分子軌道法計算によって, レチナル周辺の芳香族アミノ酸残基(トリプトファン, チロシン, フェニルアラニン)の分極効果を見積ることができるとなっており, 従来の分子力学法のモデルよりも良い結果を得ることができると示

された。

蛋白質全体を分子軌道法で計算することによって、これまででない新たな知見を得ることが可能となった。また、ONIOM法^{(4),(5)}と呼ばれる手法を用いて、MOPACと、より高精度な密度汎関数法^(注7)と組み合わせる方法が有力であるという報告もなされている。⁽⁶⁾ MOPACは、生体分子の化学現象の解明へ向けた研究の一翼を担うツールとして、今後ますます活用されると考える。

MASPHYCの蛋白質系への適用

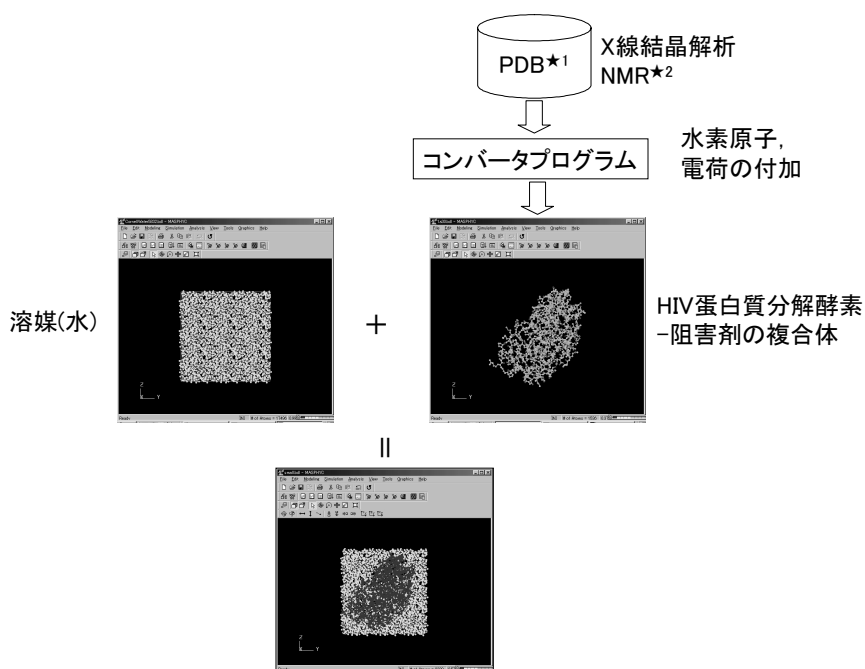
蛋白質・水系の分子動力学シミュレーション

HIV蛋白質分解酵素は、HIVが増殖する際、数珠つなぎになって作られた蛋白質を切断して部品を作る役割をしており、多くの研究がなされている。このHIV蛋白質分解酵素を題材に、MASPHYCを用いて、蛋白質・水

系の分子動力学シミュレーションを試みた(図-3)。

計算を行った系の原子数は約15,000、水分子をあらわに扱い、周期境界条件下でクーロン力の計算にはEwald法^(注8)を用いた。1ステップの計算に要した時間は、PRIMEPOWER850(675 MHz)15CPU並列計算で、0.66 sであった。時間刻みは0.2 fs (1 fs = 10⁻¹⁵ s)である。1ステップの計算に要する時間は、原子数、実空間および波数空間のカットオフなどに依存して異なるため一概には言えないが、本計算例の場合、1 ns (1 ns = 10⁻⁹ s)の現象を追跡するために必要な計算時間は約38日となり、より高速な計算手法が求められる。

そこで、計算時間短縮のために有効な、運動方程式の数値解法について調査、検証を行ったので、以下に報告する。



- 1: PDB (Protein Data Bank): 米国のブルックヘブン国立研究所により管理されている蛋白質立体構造データベース。http://www.rcsb.org/pdb/で参照できる。
- 2: NMR (Nuclear Magnetic Resonance): 原子核が磁場の中におかれたとき、それぞれの原子核に特定の周波数の電磁波を吸収する現象をいい、低分子、蛋白質、核酸の立体構造を解析することができる。

図-3 HIV蛋白質分解酵素-阻害剤の複合体の水中への配置

Fig.3-Arrangement of HIV protease and inhibitor complex into water.

(注7) 密度汎関数法: ある外場中にある多電子系の基底状態は、その電子密度の汎関数として一意的に決まるとする原理に基づく電子論的手法で、系の基底状態のエネルギーを密度に関する変分原理により求めることができる。

(注8) Ewald法: イオン結晶のクーロンポテンシャルによる凝集エネルギーを計算するためにP. P. Ewaldにより開発されたクーロン力の計算方法。クーロン力は長距離まで強い力を及ぼすため、基本セルの長さの半分でクーロン力をカットしてしまうと大きな誤差を招くことになる。Ewald法ではこのようなカットをせず、遠くのイメージセル中との相互作用も含めた形で評価するため精度が高い。

運動方程式の数値解法の調査，検証

長い時間刻みでも安定なVelocity Verlet法と，変化の速い力，遅い力など，力の種類に合わせて時間刻みを変えるMultiple Time Step法について調査，検証を行った。

(1) Velocity Verlet法⁽⁷⁾

分子動力学法では，運動方程式の時間発展は，数値積分によって求められる。微分方程式の数値解法には大きくわけて差分近似法と予測子-修正子法があり，MASPHYCでは，予測子-修正子法の一つであるGear法を採用している。この方法は，時間刻み (Δt) が小さい場合には，差分近似法よりも精度よく解が求められるという特長がある。しかし，時間刻みが大きくなると，差分近似法よりも精度が低くなるという短所がある。

蛋白質解析のためには，できるだけ時間刻みを大きくして，追跡時間を長くしたい。そこで，差分近似法の一つであるVelocity Verlet法を用いたプロトタイプを作成し，評価を行った。

直鎖アルカン C_5H_{12} の計算例を図-4に示す。 $\Delta t = 0.2$ fsの場合，Gear法とVelocity Verlet法ともに，エネルギーは保存される。精度はGear法の方が良い。 $\Delta t = 1.0$ fsの場合は，Gear法ではエネルギーが保存しなくなる。Velocity Verlet法では，平均値のまわりの揺らぎは大きいのが，保存の程度はGear法よりも良い。

Velocity Verlet法を前述の計算例に適用し， $\Delta t = 0.2$

1.0 fsとすると，1 nsの現象の追跡に約38日の計算時間を要していたものが，8日弱まで短縮されることが期待できる。

(2) Multiple Time Step法⁽⁸⁾

分子動力学法における計算時間の大部分は，原子間に働く相互作用を求めることに費やされている。通常の分子動力学計算では，相互作用関数の如何に関わらず，そ

の計算は単一の時間刻みごとに行われる。

一方，蛋白質を構成する各原子の間に働く相互作用を記述している力場には，時間に対して速く変化する力と，ゆっくり変化する力が混在している。時間刻みが単一の場合には，これらのうちで最も速く変化する力に合わせて，時間刻みを小さく調整しなければ精度の高い計算結果が得られない。

しかし多くの場合，速く変化する力よりもゆっくりと変化する力の方が多くの計算量を必要とするため，小さな時間刻みごとにすべての相互作用を計算するのは効率的でない。速く変化する力は小さな時間刻みで求め，ゆっくりと変化する力は精度を保ったままで数ステップおきに更新することができれば，計算時間の大幅な軽減を期待できる。

このようなアイデアはMultiple Time Step法として知られ，前述のVelocity Verlet法と組み合わせることにより，広く利用されつつある。

そこで，速く変化する相互作用として分子内の結合相互作用を選び，分子間の非結合相互作用をゆっくり変化する力としたプロトタイプを作成し，評価を行った。

直鎖アルカン $C_{25}H_{52}$ の計算例を図-5に示す。単一の時間刻みで計算した場合と，Multiple Time Stepで計算した場合とで，温度や圧力，ポテンシャルエネルギーのトラジェクトリがよく再現されていることを確認できる。

Multiple Time Step法を，Velocity Verlet法と組み合わせることで前述の計算例に適用し，仮に非結合相互作用の計算が2.0 fsごとで十分だとすると，1 nsの現象の追跡に約38日の計算時間を要していたものが，4日まで短縮され

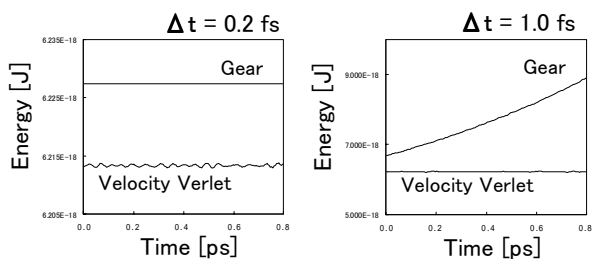


図-4 Gear法とVelocity Verlet法の比較 (直鎖アルカン C_5H_{12} の計算例)

Fig.4-Comparison of the Gear algorithm and the Velocity Verlet algorithm. (Example calculation of n-pentane.)

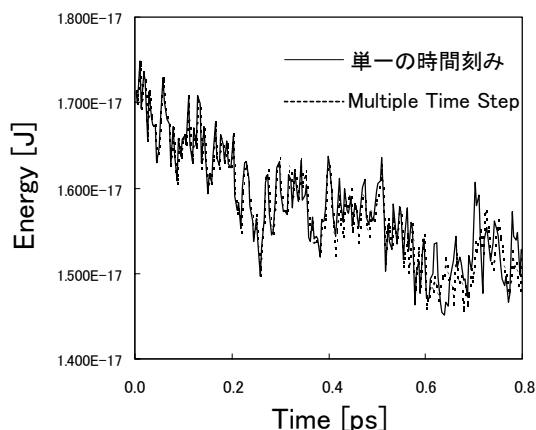


図-5 Multiple Time Stepの計算例 (直鎖アルカン $C_{25}H_{52}$)
Fig.5-Example calculation of the Multiple Time Step method. (Straight chain alkane, $C_{25}H_{52}$)

ることが期待できる。

このように、計算時間短縮のために有効な運動方程式の数値解法を採用することにより、計算可能な原子数および追跡可能な時間のオーダを更に増すことが可能となり、今後MASPHYCが蛋白質解析に有効なツールとして更に役立つものと考えられる。

in silico screeningシステムの開発

薬は一般に低分子化合物であり、生体内において蛋白質と相互作用することにより効果が発揮される。最近のゲノム情報解析のめざましい発展の結果、疾病に關与する遺伝子が次々と明らかにされてきており、それらの遺伝子から作られる蛋白質の構造と活性部位を知り、その活性部位に相互作用する低分子化合物を探し出すことにより、疾病に対する薬物候補化合物を見いだすことができる。

これら一連の操作は、

- (1) ゲノム情報からバイオインフォマティクスを通じて標的とする蛋白質を選択
- (2) ゲノム解析により読解されたアミノ酸配列の相同性に基づいて、データベースから、あるいは計算によって蛋白質の立体構造を予測
- (3) 蛋白質とリガンドの結合部位と結合様式を推定し、バーチャルスクリーニングによって、蛋白質に強く

結合すると予想される化合物を化合物ライブラリから選出

- (4) 選出された化合物について実験による評価（リアルスクリーニング）を行い、薬物候補化合物を選出となり^{(9),(10)} 最近のコンピュータの性能向上と各種ソフトウェアにおける高速アルゴリズムの開発や並列化技術により、(1)から(3)のすべてをコンピュータ上（in silico）で行えるようになってきた。これらの流れを図-6に示した。

現在富士通では、富士通の持つ計算化学ソフトウェア技術や分子モデリング技術を駆使することにより、上記一連の流れをPC上で容易に実行できるようなin silico screeningシステムの開発を検討中である。このシステムを実現しゲノム創薬研究を少しでも効率化することができれば、新しい医薬品の開発に寄与できるものと考えられる。

具体的には、上記の(2),(3)においては分子軌道法や分子動力学法などの計算化学の手法を適用することが可能であり、富士通の半経験的分子軌道計算プログラムMOPACや分子動力学計算プログラムMASPHYCを、蛋白質の構造解析や蛋白質と低分子化合物の相互作用解析（結合エネルギーの評価）に用いることができる。しかし、MOPACにおいては蛋白質系における計算精度の向上、MASPHYCにおいては蛋白質をより速く計算す

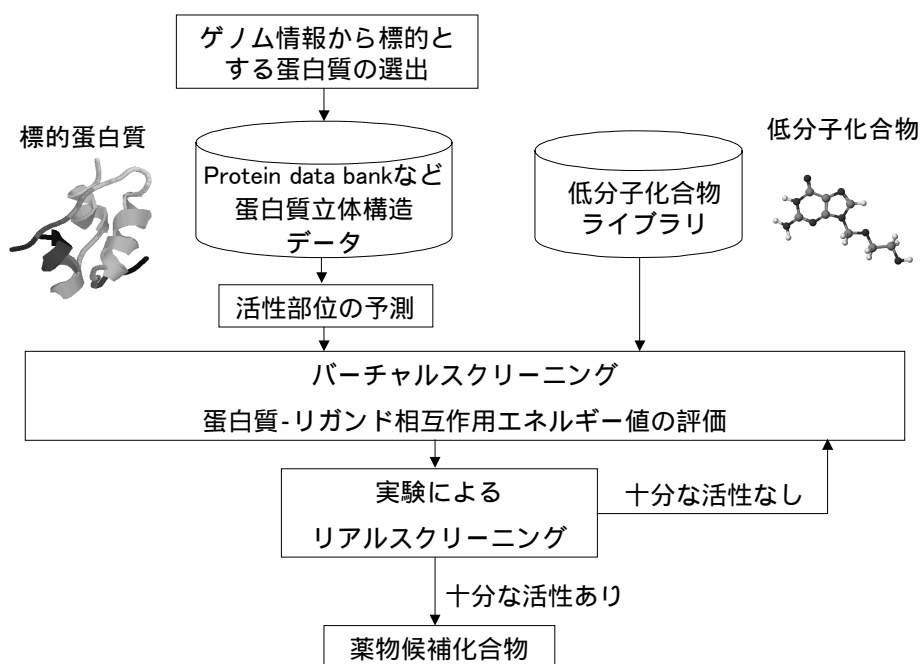


図-6 in silico screeningシステムの流れ
Fig.6-Flow chart of the in silico screening system.

るためのアルゴリズム改良や溶媒効果を取り入れる計算手法の導入などの課題があり、これらの解決に向け前章で述べたような取組みを継続していく必要がある。

む す び

MOPACを用いて蛋白質全体を分子軌道法で計算することによって得られた新たな知見や、MASPHYCによる蛋白質・水系の分子動力学シミュレーション、およびこれらの計算化学アプリケーション技術を創薬研究に有効活用するためのin silico screeningシステム開発への取組みについて紹介した。

今後もコンピュータによる蛋白質解析の実用化に向けて、高速な計算アルゴリズムの研究開発と、蛋白質系への適用を積極的に推進する。

参 考 文 献

- (1) J. J. P. Stewart : Int. J. Quant. Chem. , Vol.58 , p.133-146 (1996).
- (2) K. Ohno et al. : J. Am. Chem. Soc. , Vol.123 , p.8161-8162 (2001).
- (3) H. Houjou et al. : J. Phys. Chem. B , Vol.105 , p.867-879 (2001).
- (4) M. Svensson et al. : J. Phys. Chem. , Vol.100 , p.19357 (1996).
- (5) M. Svensson et al. : Phys. , Vol.105 , p.3654 (1996).
- (6) K. Ohno et al. : Chem. Phys. Letters , Vol.341 , p.387 (2001).
- (7) Swope et al. : Chem. Phys . Vol.76 , p.637-649 (1982).
- (8) Tuckerman et al. : J. Chem. Phys. , Vol.97 , p.1990-2001 (1992).
- (9) 宇野公之ほか：構造生物学と創薬科学．21世紀の創薬科学（辻本豪三，田中利男編），p.87（1998）．
- (10) 金井理ほか：バイオインフォマティクスからの創薬．ゲノム創薬（古谷利夫，増保安彦，辻本豪三編），p.116（2001）．

