

## 1. 序論

医薬分子の開発を分子レベルで行う場合、医薬分子と標的分子である生体高分子特にタンパク質との相互作用を精密に予測することが極めて重要である。標的分子の構造解析の質と速度が顕著に向上している状況に加え、最近になり経済的な医薬品開発に対する要求が急激に高まっている。こうした中で、標的分子と低分子の結合（ドッキング）を *in silico* 予測することへの期待が益々高まっている。

ドッキング・シミュレーションを行うためには、高い精度の構造データおよび分子間相互作用を正確に評価できる方法が必須である。しかし、最良の手段である X 線解析を用いても、理想的な構造情報を得ることは現状では容易ではない。また従来から分子間相互作用の評価に用いられてきた分子力場法は多様なタンパク質と低分子化合物との相互作用を精密に求める上では必ずしも充分でない。

これらの問題を解決するためには、十分信頼できる構造データと量子化学計算を併用することが望ましい。しかしこれまでは、必要な計算量が膨大であり、巨大分子に適するソフトウェアがないことなどが、量子化学計算の適用を阻んできた。我々は一昨年より、半経験的分子軌道法ソフトウェア LocalSCF[1]を用いて、この問題に挑戦する試みを開始した。LocalSCF は市販のソフトウェアであり、アルゴリズムは一応完成しているが、タンパク質に実際に適用する上で解決すべき問題が多く残されている。特にドッキング・シミュレーションに分子軌道法を適用するには、全タンパク質の構造最適化が必須であり、それを実行する上での戦略を整備する必要がある。

本研究では、小型のタンパク質を用い、実際に LocalSCF で試行錯誤的に全構造の最適化を行い、計算過程で遭遇する様々な問題点に対処する方法を検討した。本研究の結果得られた知見は、今後より大きなタンパク質を用い、実際にドッキング・シミュレーションを通じて創薬につなげる上で重要な戦略を与えるものと考えられる。

## 2. 計算

本研究で計算に用いたタンパク質は  $\alpha$ -conotoxin で、構造データは、X 線解析で決定された。原子パラメータは Protein Data Bank から得た (1hje; Si form)。このタンパク質は 13 残基 (ICCNPACGPKYSC) からなり、非イオン化モデルでは 175 原子 ( $C_{55}H_{84}O_{16}N_{16}S_4$ ) を含む。X 線回折の分解能は 0.75Å で、結晶学的 R 因子は 0.127 であり、十分精密な初期構造である。 $\alpha$ -conotoxin は非常に小型ながら、分子内の 2 本のジスルフィド結合によって立体構造が固定されている。また、小型であるため計算に要する時間が短く、結果の検討も容易であるため、試行錯誤する際に都合が良いので本研究の目的には適切な系である。

計算には、BioMedCACHe 6.1.10 に搭載された LocalSCF v6.1.10[2]を使用し、Microsoft Windows XP Professional SP1 搭載の PC ワークステーション (Dell Precision 650, Dual Intel Xeon 3.2GHz, 3.5GB RAM) 上で、シングル・プロセスで実行した。

ハミルトニアンには AM1, PM3, PM5 および MNDO を用いた。他の LocalSCF の計算条件は BioMedCACHe による設定 (default および coarse)、および default 設定に精密化のキーワード precise を加えた合計 3 種類を用いた。

構造最適化のサイクル数の上限を設定するキーワード cycles を、coarse については 50、default と precise では 200 に設定し、上限を越えるなどで停止した場合は、その時点での構造を用いて再度構造最適化を実行した。LocalSCF は計算の際に局在化軌道を用いるが、構造が大きく変化するとこの局在化軌道が壊れ、計算が停止することがある。この場合、再度計算を実行し、局在化軌道を作り直すことで計算を続行できる。そのため、キーワード cycles の設定は必要である。

計算の前段階における分子構造の調整 (水素付加、荷電基の編集など) および結果解析には、MOE (Molecular Operating Environment) version 2004.03[3]を用いた。モデリングの際に水分子は削除した。溶媒効果は計算には含めていない。

N 末端と Lys10 がイオン化したモデルおよび非イオン化モデルについて計算を実行し、それらの結果を比較した。両モデルについて、LocalSCF の設定は precise を用い、5 回の構造最適化 (200 サ

イクル×5回) を実行した。イオン化したモデルについては、1回目の計算が154サイクルで停止したが、2回目以降は上限の200サイクルまで計算が進行した。非イオン化モデルではこのようなことは起きなかった。

precise 設定で5回計算した場合、1回目の200サイクル計算には約10分要したが、計算が収束するにつれてサイクルあたりの計算時間は減少し、5回目の200サイクルは約5分であった。全構造最適化に要した全計算時間は、約30分であった。

### 3. 結果

#### 3.1 $\alpha$ -conotoxin の最適化構造

図1,2に、 $\alpha$ -conotoxin の全構造最適化前後の構造を示した。両方のモデルとも非常に類似した結果となり、構造最適化前後でタンパク質の全体の構造は良く保たれており、この手法がタンパク質の全構造最適化に適用できる大きな可能性を示唆している。また、表面に存在する自由度の大きい Lys10 および Ile1 の構造変化は比較的大きい。

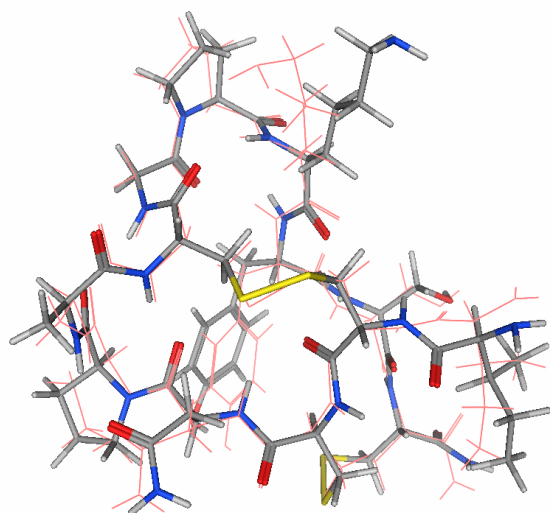


図1  $\alpha$ -conotoxin (非イオン化モデル) の最適化前後の構造比較 (赤が構造最適化前の構造)

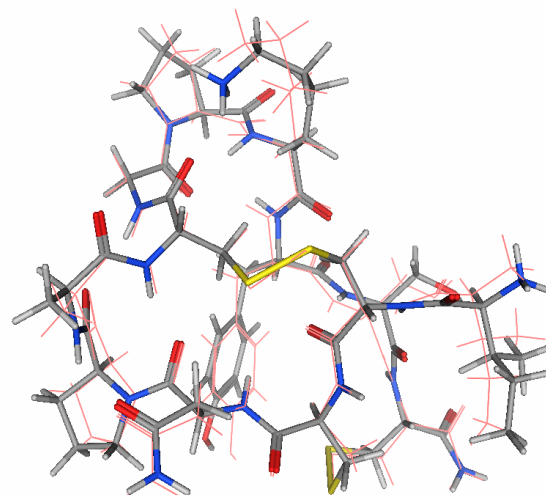


図2  $\alpha$ -conotoxin (イオン化モデル) の最適化前後の構造比較

#### 3.2 非イオン化モデルとイオン化モデルの比較

図3に、非イオン化モデルとイオン化モデルの比較を示した。イオン化モデルについては、Lys10の側鎖の $\text{NH}_3^+$ 基が、静電相互作用によってCys7の主鎖のカルボニル酸素に引き付けられ、水素結合を形成した(図4)。この構造変化は1回目の最適化前後で起きていることから、1回目の最適化が154サイクルで停止した原因はこの構造変化によるものと推定される。Lys10の側鎖がCys7側に動いたことで、隣接するPro9が接触を避けるような構造を取っている。他に、N末端のIle1とC末端のCys13構造にも違いが観察される。

今回の計算では溶媒効果を入れていないため、イオン化モデルでは、Lys10のようにイオン化され、自由度の高い官能基が表面に孤立して存在する場合、近くに相互作用可能な箇所が存在するとそちらの方に強く引き付けられる。このことによって周辺も影響を受け、構造を大きく歪めてしまう危険性がある。

結果として、今回のように溶媒効果を入れない計算の場合は、非イオン化モデルの方がより望ましいと考えられる。

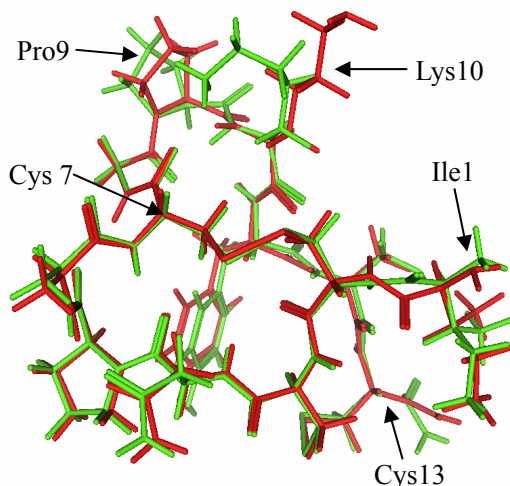


図3  $\alpha$ -conotoxin の非イオン化モデル (赤) とイオン化モデル (緑) の構造最適化結果の比較 (all-atom RMSD: 0.81)

### 3.3 構造最適化の収束過程

構造最適化の過程での全エネルギー (生成熱) 変化を図4に、gradient norm の変化を図5に示した。全エネルギーについては、イオン化したモデルでは200サイクル付近から変化量が小さくなった。一方、イオン化していないモデルでは600サイクル以降変化量が小さくなったが、900サイクルを越えたあたりから再び変化量が大きくなる傾向が見られた。Gradient norm については、双方とも600サイクル以降でおおむね3前後の値を取るようになるが、収束は非常に遅く、時折構造変化に伴うと見られる gradient norm の急な変化も観察される。その傾向はイオン化していないモデルの方がより顕著である。

タンパク質では最適化すべき変数が低分子に比較して極めて多いため、低分子の量子化学計算で通常見られるような良好な収束を得ることは非常に困難であると考えられる。Gradient norm の結果はこれを明瞭に示している。一方、エネルギーに関しては、おおむね良好な収束を得ていると見ることができる。

タンパク質の構造比較に良く用いられる root mean square deviation (RMSD) の変化を図6に示した。この図は、5回の構造最適化を行った最終結果をRMSDの基準とし、各段階の構造が最終結果とどれだけずれているかを示したものである。また、構造最適化に含まれる全原子の構造変化を追跡するため、水素原子を含めた全原子についてのRMSD (all-atom RMSD) を求めた。後半ではRMSDの変化は0.1未満であり、全体の平均的な構造の変化はこの時点で十分小さくなっていると考えられる。

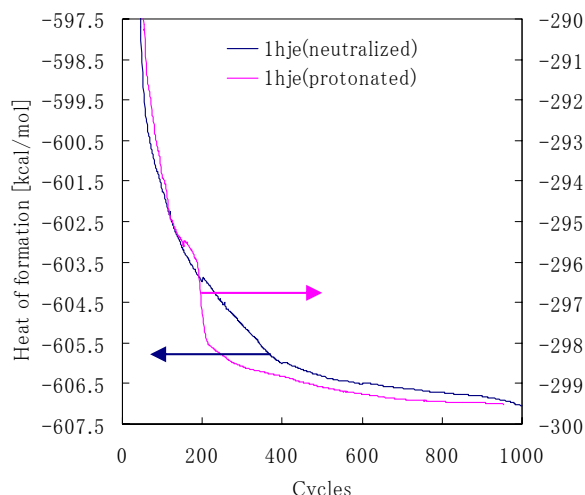


図4  $\alpha$ -conotoxin の構造最適化における全エネルギー (生成熱) 変化

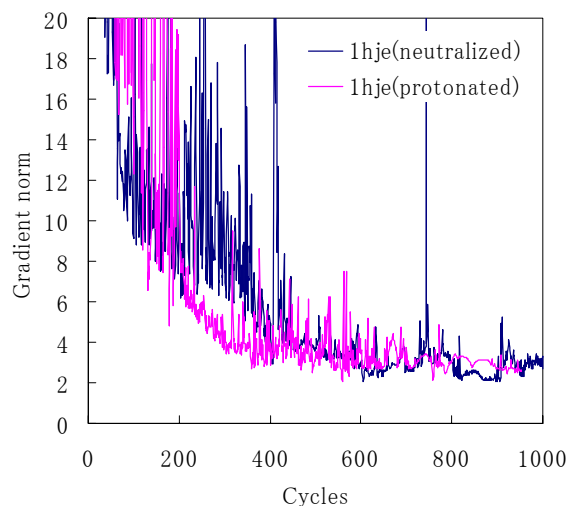


図5  $\alpha$ -conotoxin の構造最適化における gradient norm の変化

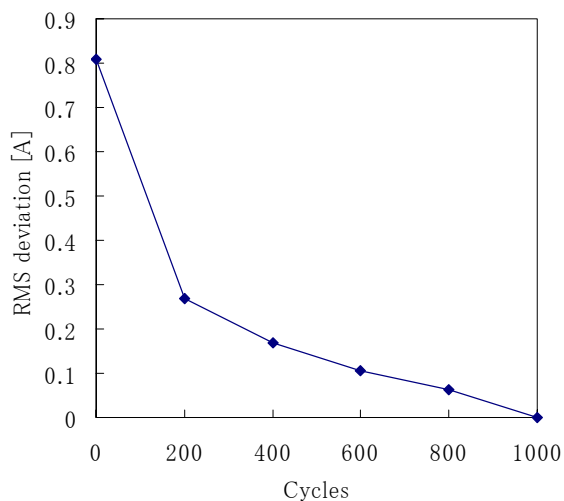


図6  $\alpha$ -conotoxin の構造最適化における all-atom RMSD の変化 (非イオン化モデル、1000サイクル目を基準)

### 3.4 半経験的手法間の比較

AM1, PM3, PM5 および MNDO の各手法間の比較を図7に示した。これらの計算では、LocalSCF の設定の coarse, default および precise を順に実行した後の結果を用いている。どれも似たような構造を取っているように見えるが、半経験的方法で懸念される水素結合距離の違いについて検討した。これを調べるため、AM1, PM3, PM5 および MNDO の各手法でのカルボニル酸素とアミド窒素間の水素結合距離のプロファイルを図8に示した。プロファイルの点の高さは、その点の前後に他の点がどれだけ集まっているかを示し、点が密集しているところほど高くなる。AM1 についてはオリジナルの構造に比べて約 0.2Å 長い方にシフトしているが、プロファイルの形は似ている。しかし、PM3 のプロファイルは大きく異なり、PM5 と MNDO については水素結合の大部分または全てが長くなってしまい、プロファイルから欠落している。このことから、タンパク質の全構造最適化には現状では AM1 法が望ましいといえる。

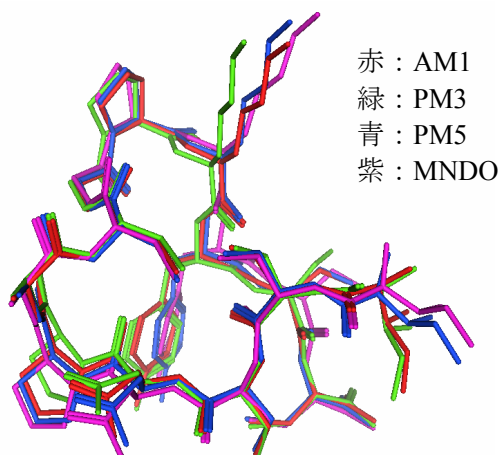


図7  $\alpha$ -conotoxin (非イオン化モデル) の AM1, PM3, PM5 および MNDO 各手法間の比較

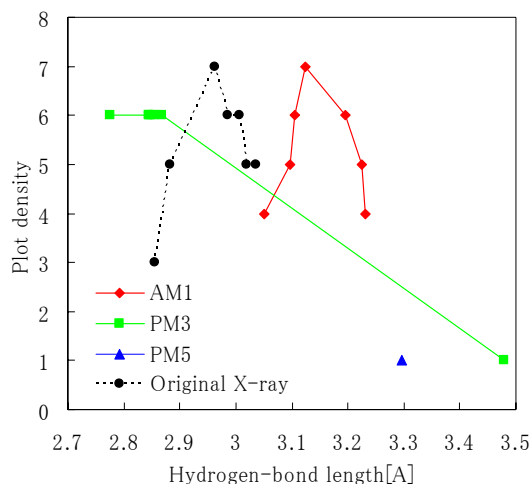


図8 カルボニル酸素とアミド窒素間の水素結合距離のプロファイル (非イオン化モデル)

## 4. 結論

$\alpha$ -conotoxin の LocalSCF を用いた半経験的分子軌道法による全構造最適化の結果から、以下のことが示唆された。

- 半経験的分子軌道法による構造最適化によってタンパク質の立体構造が大きく崩れることはないため、この手法が一般的にタンパク質の全構造最適化に適用できる可能性が高い。
- 溶媒効果を入れない場合、イオン化される可能性があり、タンパク質表面に存在する自由度の高い官能基はイオン化させない方が、元の構造が大きく歪むことを避ける上で、安全である。
- 収束の過程においては特に gradient norm の収束が遅くなった。このため、全エネルギーに加えて、all-atom RMSD などの指標を併用することが現実的である。
- 全構造最適化に用いるハミルトニアンは、タンパク質の立体構造保持に重要な役割を果たしている水素結合形成の比較から、現状では AM1 が最も妥当である。

## 謝辞

本研究を行うにあたり、東海大学総合研究機構から援助を得た。ここに記して謝意を表す。

## 参考文献

- [1] N. A. Anikin, V. M. Anisimov, V. L. Bugaenko, V. V. Bobrikov, A. M. Andreyev, J. Chem. Phys., 121, 1266-1270 (2004)
- [2] N. A. Anikin, V. M. Anisimov, V. L. Bugaenko, V. V. Bobrikov, A. M. Andreyev, BioMed CAChe v6.1.10, LocalSCF v6.1.10, Fujitsu Limited, Tokyo, Japan (2004)
- [3] MOE (Molecular Operating Environment) version 2004.03, Chemical Computing Group Inc., Montreal, Canada (2004)