

審査の結果の要旨

氏名 角越 和也

本論文は6章からなり、第1章では分子生物学におけるドッキング計算の重要性、およびその特徴と困難性を提示し、第2章ではこれまで用いられてきたドッキング計算法の概要を述べる。第3章において提出者が球面調和関数を用いて新規に設計した正規直交基底関数による級数展開を用いた、高分子-高分子、特にタンパク質-タンパク質複合体の結合配座予測(ドッキング予測)の高速計算手法を提案する。第4章では、ドッキングにおけるスコア関数としてACE(原子接触エネルギー)を導入し、その直交分解により主要項のみで近似する手法を導入する。第5章では、このアルゴリズムをいくつかのタンパク質データに適用し、計算時間、精度について比較する。第6章では提案するアルゴリズムの高速性と有効性を示し、残された課題について論じる。

広大な配座空間を探索するために従来提案されている手法には、(1)タンパク質表面の特徴的な点のマッチングによるもの、(2)分散した複数の初期配座からエネルギーの最小化を行うもの、(3)Fast Fourier Transform (FFT)による高速相関計算を用いた最小エネルギー配座探索を行うものがある。近年では、FFTによる6次元配座空間の探索に基づく手法が多く提案されているが、多くの計算時間を要するのでより高速な手法の開発が求められている。そこで論文提出者は、球面調和関数を用いて新規に設計した正規直交基底関数での級数展開による高速内積計算を使ったドッキング予測アルゴリズムを提案した。本提案手法では、スコア関数を、各分子から定義されるスカラー場の内積の線形和として表し、順位付けし、上位のものを候補配座とする。本手法では、スカラー場を上記正規直交基底関数で展開することにより、スコア関数の計算に必要な内積計算を高速に行う。また、配座空間の探索に必要な座標変換操作も高速に行えることを示した。

球面調和関数を用いたドッキング予測手法は過去にもいくつか提案されているが、展開係数によるスカラー場の表現能力が中心からの距離 r の増加に従って劇的に劣化するという問題点がある。このため、 r によって複数の領域に分け、修正 Legendre 多項式と球面調和関数とを組み合わせた基底関数を使い、また、タンパク質の結合エネルギーをよく反映すると報告されている原子レベルの統計ポテンシャルである Atomic Contact Energy (ACE)を、本手法の枠組みで効率的に計算できる形に変換しスコア関数に取り込んだ。

本手法の性能評価として、実際のタンパク質の構造データを用いて、計算時間と精度の測定を行った。本手法は、同程度の精度をもつ従来の FTDock より2桁から3桁高速である。これは探索に必要なスカラー場の座標変換操作、とくに回転操作が高速にできること、また各分子に対する一つのスカラー場の表現に必要なメモリ量が少なく、さらに一部の空間のみを重点的に探索することが可能であることなどによる。精度の評価は、各候補構造の interface RMSD、つまり、正解の複合体構造において相互作用をしている部位の構造がどれだけ予測構造のものと近いかを表す指標を計算することによって行った。

論文提出者は、ドッキング予測について新しいアルゴリズムを提案しその有効性を示したもので、生命情報科学について新しい知見を与えるものである。今後より多くの例に適用し有効性を示すことが期待される。

本研究は、寺田透、中村周吾、清水謙多郎との共同研究であるが、論文提出者が主体となって分析及び検証を行ったもので、論文提出者の寄与が十分であると判断する。

よって本論文は博士(情報理工学)の学位請求論文として合格と認められる。