

論文の内容の要旨

応用生命工学 専攻
平成 21 年度博士課程 進学
氏 名 西村 麻彦
指導教員名 清水 謙多郎

論文題目

タンパク質の動的構造解析に関する研究

タンパク質は 1 本のアミノ酸の鎖として合成されてから、そのアミノ酸配列特有の天然構造へと折り畳まれて (folding)、そのタンパク質特有の機能を発揮する。また、近年になって、タンパク質中で特定の構造を持たない disorder 領域の研究が注目されるようになり、ある標的分子が近付いてきた時のみ特定の構造へ変化して、その機能を発揮する例も発見された。これらの現象はタンパク質を構成するアミノ酸鎖の動きに起因するものであり、その動きの機構を解明しようと多くの研究が行われてきた。例えば、 Φ 値解析によって遷移状態における構造を解明する研究や、光学装置と急速混合装置の組み合わせによってミリ秒以下の単位で変化を記録し、その動きを解明する研究が行われてきた。しかし、これらの実験手法では原子レベルの挙動を知ることは出来ない。そこで、原子レベルの挙動が得られる MD (分子動力学)シミュレーションを用いて、タンパク質の動きを解析する研究が盛んになってきた。しかし、MD シミュレーションによって天然構造への folding を観測するには克服しなければならない問題がある。また、disorder に関してはその定義の仕方に問題がある。そこで本研究では、MD シミュレーションによって天然構造への folding を観測させる手法の開発と、disorder を再定義することを試みた。

1. 二次構造予測に基づいた拘束 MD シミュレーション

MD シミュレーションで天然構造への folding を観測しようとした場合、主に 2 つの問題

がある。1つ目の問題は、力場が正しくないという事である。力場とはシミュレーションで使われるパラメータで、この値の違いによって構造の安定性が変わってくる。このような事があるため、用いる力場によっては、天然構造ではシート構造であるべき部分がヘリックス構造の方が安定となってしまうという問題が出てくる。2つ目の問題は、シミュレーション中に、そのタンパク質がとり得る構造数が多いという事である。タンパク質のとり得る構造というのは、そのタンパク質を構成するアミノ酸残基数が多くなる程、増えていく傾向にある。そして、構造数が多ければ、天然構造という1つの構造状態を取る確率はその分低くなる。そのような場合、たとえ力場が天然構造を最安定構造だとしても、計算時間に限りのあるシミュレーションでは発見できないという問題が起こり得る。

このような背景を受けて、本研究では二次構造予測の情報を付加することによってこれらの問題を解決することを試みた。タンパク質の形を決定づける主な要因は主鎖二面角であり、そして、二次構造にはその二次構造特有の主鎖二面角の値がある事が分かっている。つまり、二次構造が既定されていれば、その構造特有の主鎖二面角の値になるよう拘束を加えれば適切な構造が安定となり、かつ探索すべき構造数が減ることになるという考えである。実験の手順としては、まず拘束する強さを検討するために、二次構造予測プログラム **PSIPRED** の予測精度を調べた。そして、各予測結果が出た時の実際的主鎖二面角の値の分布の確率を再現するように主鎖二面角の拘束エネルギーを決定した。その結果、ヘリックスと予測された場合はヘリックス領域が最安定となるような拘束エネルギーマップとなり、シートと予測された場合はシート領域が最安定となるような拘束エネルギーマップとなり、その信頼度が上がるにつれて、その最安定領域と他の領域のエネルギー差が大きくなるようになった。また、コイルと予測された場合は、信頼度が高くなるに従ってヘリックス領域を排斥するようになり、代わりに左巻きヘリックス領域を安定点とする拘束エネルギーマップとなった。

次に、その拘束エネルギーが適切に働いているかどうかを調べるため、**Alanine-dipeptide** を用いて調査した。その結果、ヘリックスやシート拘束はそれぞれ適切な領域に留まり、かつ信頼度が高くなる程その領域に留まる確率が高くなる傾向が見られた。このことから、二次構造拘束が望みの通りに働くことが確認された。

次に、信頼度の違いによって **folding** にどのような影響を与えるかを調べるために、**chignolin** (PDB ID: 1uao) という 10 残基のタンパク質を用いて二次構造拘束 MD シミュレーションを行った。その結果、天然構造の存在確率は、拘束が無かった場合と比べて、最も拘束エネルギーが低い信頼度 0 でも 10%程度上がり、そして最も拘束エネルギーが強い信頼度 9 では 20%以上の上昇が見られた。このような結果から、適切な二次構造拘束が加えられた場合は天然構造への **folding** の確率が上昇することが確認された。

最後に、実際の二次構造予測の結果を用いて、数種類のタンパク質に対して二次構造拘束 MD シミュレーションを行った。その結果、シミュレーション中に得られた構造をクラスタリングして最も存在確率が高いクラスに天然構造を確認することに概ね成功した。

以上の結果から、二次構造予測に基づいて拘束をかけて MD シミュレーションを行い、天然構造を探索することは可能であることが分かった。

2. Disorder の定量化

現在、disorder は X 線結晶構造解析において座標が定まらなかったアミノ酸残基 (missing residue) として定義されている。つまり、「0 か 1」という具合に定義されているという事である。しかし、disorder というものがそのアミノ酸残基の揺らぎを反映したものなら、その尺度は「0~1」のような連続値で表現された方が自然である。実際に、disorder 領域付近のアミノ酸残基の電子密度を観察すると、他のアミノ酸残基の電子密度に比べて疎であることが分かった。疎であるということはその残基は結晶中で定まった座標に留まっていないということであり、つまりはある程度 disorder していると言える。このような考えに基づくと、電子密度の高低によって disorder を連続的に表すことが可能であると考えた。

そこで、X 線結晶構造解析のデータを使って電子密度を再構成し、そのタンパク質内の各アミノ酸残基の周辺電子密度を計算することで、各アミノ酸残基の disorder の程度を求めることを試みた。実験の手順としては、最初に、結晶構造解析プログラムである CCP4 を用いて、PDB に登録されている pdb ファイルと構造因子データから電子密度を再構成した。その再構成した電子密度データと pdb ファイルの各アミノ酸残基の原子座標を用いて、van der Waals 半径内の電子密度を足し合わせて周辺電子数を求め、その周辺電子数をそのアミノ酸残基の理想電子数で割ることで Raw Order Degree という目的としていた disorder の尺度を計算した。しかし、この Raw Order Degree という値を調べてみると、タンパク質によって大小バラバラで、全体的にその値は予想していたよりも低かった。その原因を調べたところ、実験条件の違いや、CCP4 で電子密度を再構成する計算の過程で値が小さくなってしまいう事が明らかとなった。これらの影響を補正するために B-factor による補正と、総電子数による補正を加えることで最終的な disorder の尺度 Order Degree が求めた。

次に、求めた Order Degree の値を使って、様々な解析を行った。

まず、各アミノ酸残基における Order Degree の分布を調べた。その結果、Asp, Glu, Lys, Arg 等の極性残基は Order Degree が低い、つまりは disorder 寄りである事が分かり、逆に疎水性残基は高い Order Degree を示す傾向があることが分かった。また、主鎖骨格の Order Degree では Gly が低い値を取る傾向があることも分かった。

次に、各二次構造における Order Degree の分布を調べた。その結果、ヘリックスでは 1 巻に要する残基数が増える程、Order Degree が高くなる傾向があることが分かった。また、シートは二次構造の中では最も Order Degree が高い割合が多い、つまりは他に比べて構造が堅牢であることが明らかとなった。また、ターン、バンドやコイルの構造は、ヘリックスやシートに比べて低 Order Degree の傾向があり、disorder 寄りになることが分かった。

3 つ目の解析として、溶媒露出面積と Order Degree の関係を調べた。その結果、強い極性を持つアミノ酸残基は溶媒露出表面積が増えていくに従って Order Degree が低くなっている

く傾向が見えた。一方で、疎水性残基では溶媒露出表面積が増えても Order Degree は低くなりにくい傾向が見えた。

4つ目の解析は、リガンド周辺残基における Order Degree の分布に関するものである。解析の結果、リガンド周辺残基の Order Degree の値は、一般的な Order Degree の平均値に比べ高くなる傾向が見られた。例外として、硫酸イオンの場合は、その周辺残基における Arg, Lys の割合が多いため、Order Degree は他のリガンドに比べて低くなっていた。それでも、その Order Degree の値は平均的な Arg, Lys よりは大きいので、一般的に、リガンドの周辺残基はそのリガンドによる相互作用によって安定化されて Order Degree は高くなる傾向にあると言える。

最後に、低 Order Degree が続く配列の組成を調べた。低 Order Degree を Order Degree (主鎖骨格)が 0.7 未満のものと定めて、そのような値が続く配列を選び出し、その並びや各アミノ酸残基の出現頻度を調べた結果、まず圧倒的に Gly が多い事、次いで、Asp, Glu, His, Lys, Gln, Ser などの極性残基が平均よりも多く現れることが分かった。逆に、疎水性残基である Phe, Ile, Val, Trp や架橋する Cys の出現頻度は平均より少なくなることが明らかとなった。

以上の解析で見られた傾向は、従来の disorder 領域でも見られる傾向と似ており、このことから、order とされている領域においても disorder と似たような傾向を持つ配列は構造が揺らぐ傾向があることが分かった。

総括

本研究によって、MD シミュレーションによる天然構造の探索の可能性を広げること、そして、より広義な disorder に関する情報を得ることに成功した。今後は、これらの研究結果を礎として、タンパク質が天然構造へと folding する過程、または、disorder 領域が特定の構造へと変化する過程の法則を導けるような方法へと発展させていきたいと考えている。