

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年6月11日現在

機関番号：14401

研究種目：若手研究（B）

研究期間：2010～2011

課題番号：22770149

研究課題名（和文） 蛋白質の構造ドメインの自動同定と自動分類とそれに基づく相互作用データベースの構築

研究課題名（英文） Automatic determination and classification of structural domains of proteins and development of interaction database

研究代表者

金城 玲（KINJO AKIRA）

大阪大学・蛋白質研究所・准教授

研究者番号：30370117

研究成果の概要（和文）：

蛋白質の相互作用部位（低分子化合物、蛋白質、核酸の結合部位を含む）の立体構造を網羅的に比較分類し、その結果を構造モチーフとして類型化した。その結果、特に蛋白質間相互作用部位に関しては、ほぼ全てのモチーフがホモログの間でしか共有されていないことが明らかとなった。このことから、従来の構造ドメインによる分類に比べて、構造と機能のより直接的な対応を明らかにできる可能性が示唆された。さらに、各サブユニットに含まれる複数のモチーフをまとめた「複合モチーフ」を定義し、生物学的機能との相関を詳細に調べた結果、相互作用部位の構造そのものの類似性や、配列類似性に比べてより頑強な相関があることが分かった。複合モチーフを、生物学的機能を単位としてさらにまとめることで、各蛋白質の静的構造のつながりを動的に表現し、註釈する方法を提案した。

研究成果の概要（英文）：

We have conducted exhaustive comparison and classification of interaction interface structures in proteins the result of which is summarized as a list of structural motifs. In particular, we have found that the structural motifs for protein binding interfaces are shared only among homologous proteins. This suggested structural motifs may better correspond to protein functions compared to traditional approaches based on overall folds. Furthermore, we have defined “composite motifs” by grouping together the individual structural motifs present in each subunit and studied the correlation between such composite motifs and biological functions. It was found that the composite motifs had more robust correlation with function than bare structural similarities of interfaces or sequence similarity.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,600,000	480,000	2,080,000
2011年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	2,800,000	840,000	3,640,000

研究分野：生物物理学

科研費の分科・細目：生物科学・生物物理学

キーワード：相互作用モチーフ、構造比較、配列比較

1. 研究開始当初の背景

蛋白質の立体構造に基づく機能予測は、構造

ドメインの同定に基礎を置いていた。しかし、

構造の保存性は配列の保存性よりも遥かに強いために、同じ構造をもつ蛋白質でも異なる機能を持つ例が少なくなかった。また、従来のドメインデータベースは更新頻度が低いために、最新の立体構造データベースの情報を反映していないという問題があった。

2. 研究の目的

本研究では立体構造に基づく蛋白質の機能予測を進めるために必要な蛋白質構造の類型化を自動的かつ網羅的に行うための方法を開発し、その評価を行うことである。

3. 研究の方法

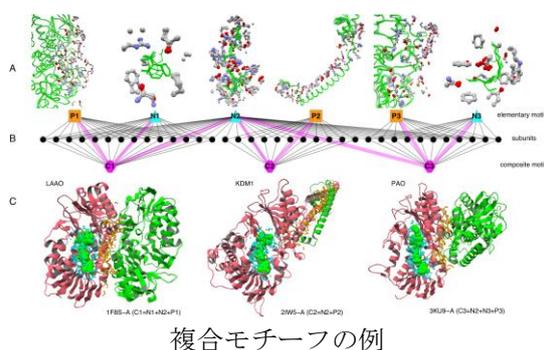
現在蛋白質立体構造データベースに存在する全ての蛋白質の(相互作用部位)構造を総当たり比較するための方法を開発した。この方法では、蛋白質立体構造の相互作用部位を予め局所座標系に変換したものを関係データベースに保存し、その幾何学的特徴に基づいて索引付けをすることにより、高速な検索を可能にする。この方法を GIRAF (Geometric Indexing and Refined Alignment Finder) と名付け、最新の立体構造データベース (Protein Data Bank, PDB) の全構造に対して適用し、相互作用部位の総当たり比較を行った。比較結果をクラスタリングすることにより、モチーフないしドメインを定義する。この方法の詳細は、BIOPHYSICS 誌 (Vol. 8, 79-94, 2012) に発表された。

4. 研究成果

低分子化合物の結合部位の分類はすでに行っていたが、これをまず蛋白質間相互作用部位の分類に拡張した。すなわち、蛋白質立体構造データバンクにその時点で存在する全ての蛋白質間相互作用部位の構造の総当たり比較を行い、クラスタ分析によって 3000 パターン以上の相互作用構造モチーフを同定した。その結果、低分子化合物相互作用部位にくらべると、蛋白質間相互作用部位のモチーフは、coiled-coil などごく少数の例外はあるものの、相同蛋白質のみで共有されている傾向がとて強いことが明らかになった。この結果は、蛋白質間相互作用の予測のためには、比較的近縁の蛋白質の複合体が解かれている必要があることを意味しており、また、数が限られていると言われている蛋白質フォールドに比べて、蛋白質間相互作用のパターンは遥かに多いことを示唆している。以上の結果は J. Mol. Biol. 誌 (Vol. 399, pp. 526-540, 2010) に発表された。

つぎに、核酸相互作用部位も含めて、各サブユニットに含まれる構造モチーフを組み合わせることで複合モチーフ (図) を定義し、

機能分類との整合性を評価した結果、フォールドレベルではなく相互作用レベルの構造モチーフが機能予測には向いているということが分かった。すなわち、一つ一つの構造モチーフは比較的遠縁の蛋白質でも共有されているが、その場合に、必ずしも同様の生物学的機能が共有されているとは言えず、複数の構造モチーフの組み合わせを考慮することによって、より類似した機能が類推できると考えられる。しかし実際には構造データベース自体まだまだ小規模であるために、機能予測よりは機能アノテーションを構造に基づいて行うことの方が生産的であることが示唆された。そのための簡便な注釈法として、メタ複合モチーフを定義し、特に転写関連蛋白質のネットワークについて調べた。ここで注意すべきことは、全く同一の蛋白質でもリガンドの結合状態や構造変化の状態によって、異なる複合モチーフをもつことがあり得るということである。つまり、複合モチーフおよびメタ複合モチーフによる蛋白質の分類は排他的ではない。これは従来の SCOP や CATH などの階層的構造分類の思想とは大きく異なるものであり、分類体系の複雑さが増してしまうという欠点がある一方で、生体内で実際にあるうる構造及び状態の動的変化を反映するという利点も有している。メタ複合モチーフのネットワーク構造を統計的に解析した結果、連結成分、蛋白質間相互作用を表現する辺、同一の蛋白質の状態遷移を表現する辺、などが配列類似性によるクラスタのネットワークに比べて大きくなる傾向が示された。これらの特性により、メタ複合モチーフにより「豊富」な機能アノテーションが可能になることが示唆される。以上の結果は PLoS One 誌 (Vol. 7, e31437, 2012) に発表された。



前述の蛋白質間相互作用の研究の一つの観察結果として、同じ蛋白質でも相互作用する相手が異なると、相互作用面に重なりがあった場合でも異なるモチーフに属するということがあげられる。たとえば、ユビキチンや Ras 蛋白質等の「ハブ蛋白質」は多くの種類の蛋白質と相互作用し、しかもその相互作用

面が重複しているが、それらの相互作用面は異なる構造モチーフとして同定された。このことから、ハブ蛋白質が異なる相手と相互作用する場合には、(1) 相互作用面が構造変化するか、または(2) 部分的に重複はしているが、重複していない部分で各リガンドに対する特異性を実現していると予想される。そこで、(2) の観点から複数の種類の相手との複合体の立体構造が解かれている蛋白質を16種類選んで、その相互作用面を重複領域と非重複領域に分割して、それらの物理化学的特性を詳細に解析した。その結果、重複領域は、より疎水性残基に富み、より埋もれた残基が多いながらも規則的な2次構造は少なく、さらに側鎖の構造が多様であることが明らかになった。特に最後の点は、ハブ蛋白質の「多重特異性」を理解するためには側鎖構造をも考慮する必要性を示している。この結果は *J. Mol. Biol.* 誌 (Vol. 411, pp. 713-727, 2011) に発表された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計12件) 全て査読あり

(1) Functional structural motifs for protein-ligand, protein-protein and protein-nucleic acid interactions and their connection to supersecondary structures.

Kinjo AR, Nakamura H. *Methods in Molecular Biology* in press:(2012)

(2) GIRAF: a method for fast search and flexible alignment of ligand binding interfaces in proteins at atomic resolution. Kinjo AR, Nakamura H. *BIOPHYSICS* 8:79-94 (2012)

[DOI:10.2142/biophysics.8.79]

(3) Composite structural motifs of binding sites for delineating biological functions of proteins.

Kinjo AR, Nakamura H. *PLoS One*

7:e31437 (2012) [PubMed]

[arXiv:1107.3278]

[DOI:10.1371/journal.pone.0031437]

(4) Distinct roles of overlapping and non-overlapping regions of hub protein interfaces in recognition of multiple partners. Dasgupta B, Nakamura H, Kinjo AR. *Journal of Molecular Biology* 411:713-727 (2011) [PubMed]

[DOI:10.1016/j.jmb.2011.06.027]

(5) Geometric similarities of protein-protein interfaces at atomic resolution are only observed within homologous families: An exhaustive structural classification study. Kinjo AR, Nakamura H. *Journal of Molecular Biology* 399:526-540 (2010)

[PubMed]

[DOI:10.1016/j.jmb.2010.04.028]

[学会発表] (計14件)

(1) 金城玲 Composite structural motifs of binding sites for delineating biological functions of proteins Recent advances in physical and chemical biology 2012年3月5-7日 Saha Institute of Nuclear Physics, India

(2) 金城玲 Composite structural motifs of binding sites for delineating biological functions of proteins 第3回生体分子研究会 2011年12月20日 名古屋工業大学

(3) 金城玲、中村春木 Composite structural motifs of interaction interfaces for

delineating biological functions of proteins 2011年9月16-18日 第49回日本生物物理学会年会 兵庫県立大学

(4) 金城玲 Exhaustive comparative study of interaction site structures in proteins: Implications for function prediction Biological Data Evaluation Workshop, Asia Pacific Bioinformatics Conference 2011年1月12日 Incheon, Korea.

(5) 金城玲 An exhaustive structural classification at atomic resolution reveals that geometric similarities of protein-protein interfaces are confined within homologous families InCoB2010 2010年9月26-28日 早稲田大学

[その他]

ホームページ等

<http://pdbj.org/giraf/>

6. 研究組織

(1) 研究代表者

金城 玲 (KINJO Akira)

大阪大学・蛋白質研究所・准教授

研究者番号：30370117