

平成22年6月1日現在

研究種目：基盤研究(C)

研究期間：2007～2009

課題番号：19500255

研究課題名（和文） 遷移金属を結合するタンパク質の探索と機能予測

研究課題名（英文） Bioinformatics analysis of transition metal-binding proteins

研究代表者

中村 建介 (NAKAMURA KENSUKE)

奈良先端科学技術大学院大学・情報科学研究科・特任准教授

研究者番号：20212095

研究成果の概要（和文）：鉄、銅、亜鉛等の遷移金属は生物にとって必須であることが知られており、その殆どがタンパク質と結合することにより機能している。近年こうした遷移金属を配位して機能するタンパク質が注目を集めるなかで、タンパク質の金属イオン結合部位のデータベースとしてMetalMineを構築しWeb上で公開することが出来た。MetalMineはタンパク質の金属結合サイトに関するデータベースとしては世界最大規模のものであり、分子進化解析等によるタンパク質の機能予測をすすめる環境が整備された。

研究成果の概要（英文）：Some of the transition metal ions such as iron, copper and zinc are known to be essential for life. Most of these transition metal ions in biological systems are functioning as complexes with specific metal binding proteins. The emergence of the terms such as metallome and metallomics, is indicating the increasing interest toward these metal binding proteins. Under such a circumstance, where a useful public bioinformatics resource for metal binding proteins is being desired, we have developed a new database system called MetalMine, which includes structural information of metal binding sites in an organized fashion. MetalMine is the largest database concerning metal binding proteins, in terms of the number of metal binding sites included, and it would serve as a quite useful resource for the functional analysis of proteins.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,300,000	390,000	1,690,000
2008年度	1,100,000	330,000	1,430,000
2009年度	1,100,000	330,000	1,430,000
総計	3,500,000	1,050,000	4,550,000

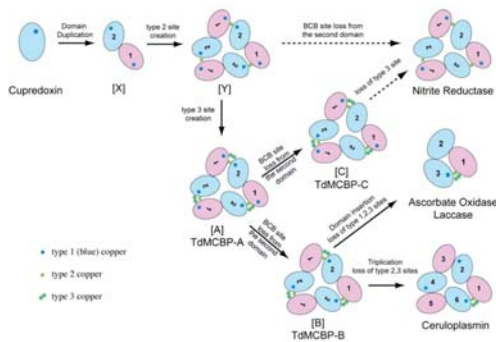
研究分野：生物情報学

科研費の分科・細目：情報学・生体生命情報学・バイオインフォマティクス

キーワード：金属タンパク質、ゲノム、生命情報学、酵素、進化、メタローム

1. 研究開始当初の背景

ゲノム配列がつぎつぎと決定される中で新しく見いだされる ORF の機能アノテーションは既知遺伝子との相同性に大きく依存しており、タンパク質機能予測手法の早急な進展が望まれている。我々は、遷移金属イオン結合性タンパク質にターゲットを絞り、未知の構造、機能を有するタンパク質遺伝子の探索を行ってきた。この試みによる成果の一つとして、マルチ銅ブルータンパク質の進化経路の解析により、バクテリアゲノム中に、二つのドメインからなる新しいタイプのマルチ銅ブルーオキシダーゼを発見し報告した。(1, 2)



この成果については最近になって、X 線結晶構造解析により、我々の予測が正しかったことを裏付ける報告が複数の研究グループによりなされている。(3, 4, 5, 6)

こうした背景の中で、同様のタンパク質機能予測を、異なる遷移金属タンパク質のグループに対して適用することと、そのために必要な方法論、計算機プログラム、データベース等の情報インフラストラクチャの整備に取り組むために、本科研費による研究課題を提案した。

先行する金属イオンを結合するタンパク質のデータベースとしては、例えば米国、Scripps 研究所の Metalloprotein Database (<http://metallo.scripps.edu/>)があるが、このデータベースは 2004 年以降更新されていない。

- (1) K. Nakamura, N. Go, *Cellular and Molecular Life Science*, 62, 2050-2066 (2005)
- (2) K. Nakamura, T. Kawabata, K. Yura and N. Go, *FEBS Letters*, 553, 239-244 (2003)
- (3) T. Skalova, J. Dohnalek, L. H. Ostergaard, P. R. Ostergaard, P. Kolenko, J. Duskova, A. Stepanokova, J. Hasek, *JMB*, 385, 1165-1178 (2009)

(4) I. S. MacPherson, Doctoral Dissertation U. British Columbia (2008)

(5) T. J. Lawton, L. A. Sayavedra-Soto, D. J. Arp, A. C. Rosenzweig, *JBC*, 284, 10174-10180 (2009)

(6) H. Komori, K. Miyazaki, Y. Higuchi *FEBS Letters*, 583, 1189-1195, (2009)

2. 研究の目的

ポストゲノム時代に入り、未知遺伝子の機能の解析がすすめられているが、実験による遺伝子 (タンパク質) の機能推定はコストのかかるプロセスであり、バイオインフォマティクス的な手法による効率化が喫緊の課題である。背景で述べた様に、申請者等は金属イオン結合タンパク質にターゲットを絞り、分子進化解析を行うことで分子機能の進化の解析が可能であることを示してきた。本研究課題は、バイオインフォマティクス解析のツール、データベースなどの技術基盤を整備することで、こうした解析を推進し発展させてゆくことを目的として開始した。

生物学は分類・整理の手法により発達してきた学問であり、進化の過程から生まれてきた多様な形態の分類により、生命機能のメカニズムに対する理解を深めてきた。タンパク質の機能の理解にもこうした博物学的な考え方が有効であると考えられ、研究の推進を支えるインフラストラクチャとして、今後も時間をかけて利用価値の高いものを作り上げてゆきたいと考えている。

3. 研究の方法

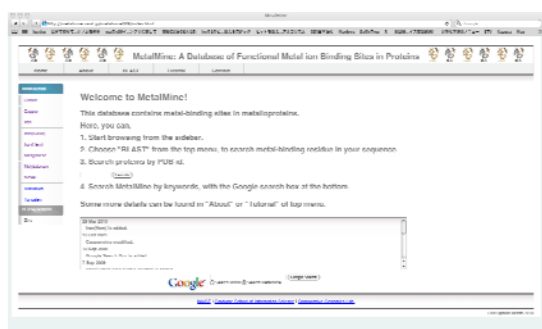
技術基盤の整備として、金属イオン結合タンパク質に関するデータベースの構築を優先しておこなった。

データ取得の方法について、銅イオンを例にとりしめすと、タンパク質立体構造デー



データベース（以下 PDB と略記）には 1503 の配位サイトがある。これを計算機プログラムによりタンパク質の立体構造による分類、配位残基の種類と数などにより、200 あまりのグループに分類し、さらに文献情報を参照しながら生物的に意味のあるサイトとして 57 のクラスを記載した。こうした手作業による分類を効率的に行うために、GUI によるデータ編集ツールを作成した。

このようにして、構築したデータベースから得られる情報をもとに、金属結合残基の予測や、金属イオン結合タンパク質の分子進化解析をすすめた。



4. 研究成果

(1) 金属イオン結合タンパク質データベース MetalMine を構築し、Web 上で一般に公開した。(発表論文①、学会発表②、③)

Web サイトで左側のサイドバーから金属イオンの種類を選択し、該当する金属イオンを含む結合サイトをブラウズすることが出来る。金属結合サイトのクラスについて、タンパク質フォールド、金属イオンの数と種類、サイトを構成する配位残基、Wikipedia へのリンクがあり、また、それぞれの金属結合サイトのインスタンスについては Jmol により、それぞれの金属配位サイトの立体構造を表示することが出来る。

pdb id やタンパク質名などのキーワードによる検索機能の他に、BLAST によるアミノ酸配列からの同源性検索機能があり、データベ



ース記載の金属結合サイトとの対比をとることができる。単純な BLAST 検索でも配列全体の同源性を確かめることは出来るが、この

機能ではさらに、相同配列の中に保存された金属イオンの配位残基の存在が保証されるので、アミノ酸配列の金属結合サイト予測として用いることもできる。

金属結合サイトのクラス分類の考え方としては、特定のタンパク質フォールドの同じ位置にある金属結合サイトは進化的な起源を共有する相同（ホモログス）なサイトであると見なしている。この考え方は、金属イオン以外にも低分子化合物の結合サイト、あるいは触媒機能等の活性サイトの分類に適用可能であると考えられる。立体構造の類似したタンパク質を重ね合わせ、金属イオンの位置を投影することで金属配位位置の分類を容易にするツールを作成した。

2010 年 3 月の時点で、MetalMine は 12,013 のサイトからなる、487 種のタンパク質金属イオン結合サイトのクラスを記載している。サイト種数の内訳は、生物に必須とされる 8 種の主要遷移金属イオンについて、コバルト：14、銅：57、マンガン：76、モリブデン：25、ニッケル：14、バナジウム：1、タングステン：4、鉄：296（単核：88、ヘム：133、鉄硫黄クラスター：75）、となっており、公開されているタンパク質金属イオン結合サイトデータベースの中では世界的に見ても最大規模のものとなっている。

このデータベースについては、今後も金属イオンとして、カルシウム、マグネシウム、亜鉛など生体に重要な役割を持つものを含めて拡充しつつ、ウェブサイトの体裁も含めてより利用価値が高いものを効率的に維持管理することができるシステムとして発展させてゆきたい。

(2) アミノ酸配列からの金属イオン結合残基の予測- 遺伝子に記述されているアミノ酸配列情報から金属結合部位を推定する方法の開発をおこなった。(学会発表④) X線蛍光スペクトルによる実験でタンパク質に結合する金属イオンの種類と数を計測する技術が報告されている。この方法により金属イオンを結合するアミノ酸配列を集めることができるが、どの残基が金属イオンを結合しているのかまでは知ることが出来ない。分子進化の情報を利用することによりこれを予測する方法を開発した。金属イオンを結合するアミノ酸配列についてゲノム配列データベースから同源性の高い配列を探索し保存度の高い金属結合性残基を抽出した。実験により報告された立体構造のない 50 本の配列のうち互いに類縁性の低い 41 本に対して解析を行った結果、21 本は金属を結合した状態での近縁タンパク質の立体構造が解析の時点ではすでに解かれており、予測の信頼性を検証することが出来た。また 17 本の配列については金属イオンを含まない相同な金属

配位サイトをテンプレートとしたモデル構造上に金属を配位すると考えられるサイトの予測を行うことが出来た。残る3本の配列についてはテンプレートとなる相同配列が見つからないため配列の保存度からの予測をおこなった。

(3) 金属イオン配位サイトの進化解析 - 金属イオン配位サイトデータベース構築の過程において、タンパク質の立体構造および配列の類似性からは明らかに相同性がありながら、異なる金属イオンを配位しているものがいくつか見られた。こうしたタンパク質は機能的にも多様性の高いタンパク質群を形成し、分子進化の観点からきわめて興味深いものであると考え、ニッケル結合タンパク質であるウレアーゼを含む、アミドヒドロラーゼスーパーファミリーの進化解析を行った。系統解析の結果、このタンパク質群の共通祖先として二核の亜鉛イオンを配位するものがあり、亜鉛イオンの配位数が一つに変異するプロセスと、亜鉛以外の金属に変異するプロセスがあったことが示唆された。また、アミドヒドロラーゼ、ウレアーゼ、ジヒドロオロターゼなど機能既知のタンパク質ファミリーとは独立した、機能未知のタンパク質ファミリーが存在することが確かめられた。

5. 主な発表論文等

[雑誌論文] (計7件)

- ① K. Nakamura, A. Hirai, Md. Altaf-Ul-Amin, H. Takahashi (4名1番目) "MetalMine: a database of functional metal-binding sites in proteins", *Plant Biotechnology*, 査読有 vol. 26, 517-521, 2009.
- ② K. Tanaka, K. Nakamura, T. Saito, H. Osada, A. Hirai, H. Takahashi, S. Kanaya, Md. Altaf Ul-Amin, (8名2番目) "Metabolic pathway prediction based on inclusive relation between cyclic substructures", *Plant Biotechnology*, 査読有 vol. 26, 459-468, 2009.
- ③ K. Nishikata, M. Wada, H. Takahashi, K. Nakamura, S. Kanaya, Md. Altaf-Ul-Amin, (6名4番目) "Predicting conformation of protein complexes by determining statistically significant domain-domain interactions", *Plant Biotechnology*, 査読有 vol. 26, 495-501, 2009.
- ④ K. Murayama, T. Tanabe, Y. Ishikawa, K. Nakamura, S. Nishiyama, (5名4番目) "A Synthetic study on gymnastatins F and Q: the tandem Michael and aldol reaction approach", *Tetrahedron Letters*, 査読有

vol. 50, 3191-3194, 2009.

- ⑤ T. Oishi, (他16名12番目) "An approach to peak detection in GC-MS chromatograms and application of KNAPSAcK database in prediction of candidate metabolites", *Plant Biotechnology*, 査読有 vol. 26, 167-174, (2009)
- ⑥ J. Shiina, M. Oikawa, K. Nakamura, R. Obata, S. Nishiyama (5名3番目) "Synthesis of pingusane-type sesquiterpenoids, acutifolone A and pinguisenol along with bisacutifolones through the Diels-Alder dimerization reaction", *European Journal of Organic Chemistry*, 査読有 vol. 2007, 5190-5197 (2007)
- ⑦ G. Xiujun, K. Nakamura, H. Yu, K. Yura, N. Go, (5名2番目) "BAAQ: An Infrastructure for Application Integration and Knowledge Discovery in Bioinformatics", *IEEE Transactions on Information Technology in Biomedicine*, 査読有 vol. 11, 428-434 (2007)

[学会発表] (計4件)

- ① K. Nakamura, A. Hirai, H. Takahashi, S. Kanaya, "Metallome in coordination with Metabolome" 8th Metabolome Informatics Workshop. 2010. 1. 29. 慶応鶴岡キャンパス
- ② K. Nakamura, A. Hirai, H. Takahashi, Md. Altaf Ul-Amin, S. Kanaya "MetalMine: a database of functional metal-ion-binding sites in proteins" GIW2009 The 20th International Conference on Genome Informatics. 2009. 12. 15. パシフィコ横浜
- ③ 中村建介、遷移金属結合サイトデータベースの構築、第八回日本タンパク質化学会年会、2008. 6. 10. タワーホール船堀 (東京)
- ④ K. Nakamura, K. Yura, N. Go "Prediction of metal-binding sites for high-throughput metalloproteomics data" VII European Symposium of the Protein Society, 2007. 5. 15. Stockholm, Sweden.

[その他]

金属結合サイトデータベース: MetalMine
<http://metalmine.naist.jp/metalmine009/index.html>

6. 研究組織

(1) 研究代表者

中村 建介 (NAKAMURA KENSUKE)
奈良先端科学技術大学院大学・情報科学研究科・特任准教授
研究者番号: 20212095