

科学研究費助成事業 研究成果報告書

平成 27 年 9 月 28 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2012～2014

課題番号：24651225

研究課題名(和文) ファットグラフを用いたクロマチンのループ構造の予測

研究課題名(英文) Chromatin loop structure prediction via fatgraph

研究代表者

児玉 大樹 (Kodama, Hiroki)

東京大学・数理(科)学研究科(研究院)・特任准教授

研究者番号：40466826

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：ペプチドユニットを頂点とするグラフの辺に、ペプチドユニットの回転を表す回転行列を書き込んだものがタンパク質の $S0(3)$ 接続モデルである。我々はこのモデルを用いてタンパク質の立体構造の違いを視覚的に表示するシステムを構築した。クロマチンの折り返し構造を表現するカントール集合を用いたモデルを考案した。pHの変化によるタンパク質変性をタンパク質のグリッドモデルにより表現した。

研究成果の概要(英文)： $S0(3)$ -connection model of protein is a graph that each vertex is a peptide unit and on each edge rotation matrix of peptide units is given. We used this model to construct a system that visualize the difference of 3D structures of proteins. We contrived a model via Cantor set that explicates the loop structure of Chromatin. We also described pH-induced protein denaturation via grid model of protein.

研究分野：トポロジー

キーワード：トポロジー $S0(3)$ -接続 ファットグラフ 生命生体情報学 国際研究者交流(オーストリア) 国際研究者交流(デンマーク)

1. 研究開始当初の背景

(1) 生体分子の三次元構造を数学的に表現し、その構造を調べることは、生命情報学の聖杯とも言われる難問であり、タンパク質の立体構造、RNA の高次構造について多くの研究がなされていたが、いまだに不十分であった。

(2) R. C. ペナー、M. クヌーセン、C. ヴィウフ、J. E. アンダーセンによる論文 “Fat-graph Models of Proteins” (タンパク質のファットグラフモデル) が雑誌 *Communications on Pure and Applied Mathematics* に 2010 年に掲載された。この論文はタンパク質の立体分子構造が、ファットグラフという組み合わせ幾何的な構造を用いてそれなりの精度で記述できることを述べており、さまざまな生体高分子の立体構造がファットグラフ、または別の組み合わせ幾何的な構造で記述できるという可能性を切り拓いたといえる。

また、この論文には後述する $SO(3)$ -接続についてのアイデアも書かれている。

2. 研究の目的

生体高分子の取りうる構造は多岐にわたり、その解析のためには極めて広大な空間を扱わなければならない。実際、タンパク質の一次構造、二次構造および三次構造の関係性をみても極めて複雑である。しかし、生体高分子の相互作用や機能を決定する因子を探り、機能を予測するためには、構造の比較および分類という数学的理論の観点からも、実験的な観点からも極めて困難な課題を解決しなければならない。

ここでは、生体高分子の構造の比較・分類を、代数的トポロジーに基礎をおく数学的な抽象化を行うアプローチを採用する。具体的には、生体高分子をバックボーンの原子の位置、水素結合らを最小単位とし、それらの並進、回転で全体構造を組み合わせ幾何的に表現する。

本研究では、タンパク質に適用したファットグラフ表現の方法論を拡張することにより、クロマチン構造にファットグラフ表現を初めて適用し、代数的トポロジーの観点から転写過程におけるクロマチンのループ構造を予測することを目標とした。

3. 研究の方法

(1) グラフの頂点と辺を、それぞれ円板と帯で置き換えて幅や表裏の概念を持たせたものをファットグラフと呼ぶ。

R. C. ペナー達は次のようにタンパク質のファットグラフモデルを定式化した [Commun. on Pure and Applied Math. LXIII, 1249 (2010)]。まず、タンパク質の主鎖に

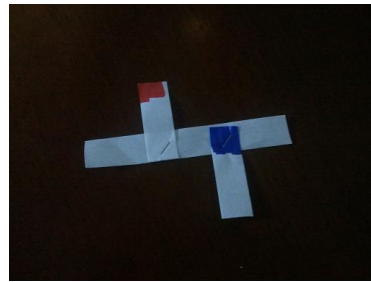


Figure 1: ペプチドユニット



Figure 2: α ヘリックス

繰り返し現れる C^α -CO-NH- C^α という構造 (ポリペプチドユニット) を Figure 1 のようなファットグラフで表す。次に、ポリペプチドユニット同士の C^α 原子での結合および水素結合を、対応するファットグラフを帯で貼り合わせることによって表す。例えば、Figure 2 はタンパク質の一次構造である α ヘリックスをファットグラフモデルの模型として表したものである。

二つの幾何学的対象が互いに相異なることを証明あるいは判定するのは人間にとってもコンピュータにとっても難しい場合が多い。例えば、与えられた結び目がほどけることを示すには、実際にほどいてみせてやればよいが、与えられた結び目がどうやってもほどけないということを数学的に直接示すのはとても難しい。

そこで不変量という概念を導入する。不変量とは図形から通常は簡単に計算できる量で、同じ図形は同じ量を定めるようなものである。例えば結び目の場合は、三色彩色数やジョーンズ多項式といった不変量を用いれば、三つ葉結び目と解けた紐とを区別することが出来る。また、ファットグラフに対しては、たとえば境界の成分数を考えればこれは不変量となる。

まず、50 個から 200 個程度のアミノからなるタンパク質の静的状態の各原子の座標の情報からファットグラフモデルを構成し、さらにそのファットグラフモデルから幾何的な不変量を取り出すような情報システムを開発する。そのような不変量はロバスト、

つまりファットグラフが少しだけ変化したときに、不変量も少しだけしか変化しないようなものが望まれる。タンパク質の一次構造、つまり α ヘリックスや β シートの存在は、そのままファットグラフモデルの構造に、さらには不変量に反映される。

(2) (1) で完成したシステムを、クロマチンにも適用できるように改良する。

(3) R. C. ペナーや J. E. アンダーセンらと緊密に連絡を取り、最新の情報のもとに研究を続ける。

4. 研究成果

(1) タンパク質のファットグラフを用いた情報システムを検討しているうちに、ファットグラフよりも $SO(3)$ -接続の情報を用いたほうがよりよくタンパク質の情報を表現できることがわかってきた。

ファットグラフはグラフの二つの頂点を結ぶ辺を帯で置き換えたもので、帯は180度ねじって結ぶかそのまま結ぶかの二つしか選択肢がないので、グラフの辺集合から $\mathbb{Z}_2 = \{0, 1\}$ への写像を定めるのに対し、 $SO(3)$ -接続の情報を用いたモデルは辺集合から三次元回転群 $SO(3)$ への写像を定めることに相当する。その分データ量は増えてしまうのだが、コンピュータで扱うぶんにはさほど問題ない量である。

実装としては、大田佳宏らと協力して、グラフィカルで直感的な表示を伴うシステムを構築した。

(2) クロマチンの繰り返し構造を自己相似構造だとみなし、コントロール集合をそのモデルとした折りたたみのモデルを考案した。

次の補題が本質的である:

補題. 任意の無理数 α に対して、円周上のコントロール集合で、 α 回転を何回施しても自分自身と交わらないようなものが構成できる。

このコントロール集合についての補題を用いた松元重則との共著論文 “Minimal C^1 -diffeomorphisms of the circle which admit measurable fundamental domains” を Proceedings of the American Mathematical Society に発表するとともに、国際数学者会議 (ICM ソウル、韓国) などの学会で発表した。

この論文の主定理は以下のようなものである:

定理. 任意の無理数 α に対して、回転数が α であり、測度論的基本領域が取れるような円周上の C^1 -級微分同相写像が構成できる。

(3) 研究教育活動の一環として、2013年7月に東京大学玉原国際セミナーハウスにおいてサマースクールを開催し、グリッドモ

デルによるタンパク質フォールディングのモデル化をあつかった。その後、東大医学部の学生である日高剛朗・島田敦の両名と pH の変化によりタンパク質フォールディングが変わる様子に着目した研究を行った。

具体的には、次のようなことを行った。まず、アミノ酸を親水性のものと疎水性のものに大別し、親水性のアミノ酸どうし、疎水性のアミノ酸どうし、親水性と疎水性のアミノ酸がそれぞれグリッド上で隣接している時のエネルギーへの寄与を pH の関数として定める。すると、pH を変化させることにより、グリッド上で最もエネルギー的に安定する配列も変化することが観察できる。

この結果を論文 “Simple exact model of pH-induced protein denaturation” にまとめた。この論文は2015年6月3日に Physical Review E に accept された。

(4) 研究教育活動の一環として、これらの研究に基づき、東京大学の大学生向けに以下の授業を行った。

- 2014年6月9日・16日 全学自由ゼミナール 教養学部1・2年生むけ
- 2014年6月24日・7月1日 現象数理 II 理学部3・4年生むけ
- 2015年5月26日・6月2日 現象数理 II 理学部3・4年生むけ

5. 主な発表論文等

[雑誌論文] (計2件)

[1] T. Hidaka, A. Shimada, Y. Nakata, H. Kodama, H. Kurihara, T. Tokihiro and S. Ihara, Simple exact model of pH-induced protein denaturation, Physical Review E, 査読あり, 2015年6月3日 accepted

[2] Hiroki KODAMA and Shigenori MATSUMOTO, Minimal C^1 -diffeomorphisms of the circle which admit measurable fundamental domains, Proceedings of the American Mathematical Society, 査読あり, Volume 141 (2013), 2061–2067, DOI: 10.1090/S0002-9939-2013-11472-3

[学会発表] (計2件)

[1] Hiroki KODAMA, Minimal C^1 -diffeomorphisms of the circle which admit measurable fundamental domains, 国際数学者会議 (ICM), 2014年8月19日, Seoul (Korea).

[2] Hiroki KODAMA, Minimal C^1 -diffeomorphisms of the circle which admit measurable fundamental domains, Geometry and topology of macromolecule folding, 2013年12月6日, CENTRE for QUANTUM GEOMETRY of MODULI

SPACES, Aarhus University, Aarhus
(Denmark).

[図書] なし

[産業財産権] なし

6. 研究組織

(1) 研究代表者

児玉 大樹 (KODAMA, Hiroki)

東京大学・大学院数理科学研究科・特任
准教授

研究者番号: 40466826

(2) 研究分担者

井原 茂男 (IHARA, Shigeo)

東京大学・先端科学技術研究センター・
教授

研究者番号: 30345136