

平成 30 年 8 月 31 日現在

機関番号：34315

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00412

研究課題名(和文)天然変性タンパク質におけるO-GlcNAc修飾の予測と検証

研究課題名(英文) Prediction and verification of O-GlcNAc modification in intrinsically disorder protein

研究代表者

伊藤 将弘 (ITO, Masahiro)

立命館大学・生命科学部・教授

研究者番号：50388112

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：O-GlcNAc修飾を担う2つの酵素であるOGTとOGAとO-GlcNAc修飾を受けるタンパク質に着目してIDRとの関係も含めて進化解析を行った。OGTとOGAの解析では、ヒトOGT/OGAの相同タンパク質をゲノムワイドに検索することで付加と除去を行う酵素の有無からO-GlcNAc修飾の獲得過程を解明した。O-GlcNAc修飾を受けるタンパク質の解析では、ヒトO-GlcNAc修飾タンパク質を進化解析の基準とし、生物間での差異を確認することでO-GlcNAc修飾によるタンパク質の機能調節の解明を行った。

研究成果の概要(英文)： We analyzed the relationship between evolutionary conservation and function in O-GlcNAcylated proteins. The result of the multiple alignments of OGT/OGA was shown as a plot pattern. The phylogenetic analysis demonstrates that, OGT was widely conserved in eukaryotes, but OGA was conserved only in animals. Additionally, in fungi and protists, OGT was conserved only in the catalytic domain. O-GlcNAcylated proteins were clustered into three classes, which called "Class 1", "Class 2", and "Class 3." Class 1 with 47 proteins, conserved mainly in vertebrates. That is, it is evolutionary "young" proteins. Class 2 with 94 proteins conserved in animals. Over half number of O-GlcNAcylated proteins belong to this class. Class 3 which 21 proteins widely conserved in eukaryotes.

研究分野：ゲノム情報学

キーワード：翻訳後修飾 O-GlcNAc修飾 リン酸化 機械学習 進化解析 天然変性タンパク質 バイオインフォマティクス ゲノムワイド解析

1. 研究開始当初の背景

翻訳後修飾の一種である *O*-GlcNAc 修飾は、セリンまたはスレオニンに単一の β -*N*アセチルグルコサミンが結合する。その一連の流れとして、*O*-GlcNAc 転移酵素(OGT) により上記 2 種類の残基に GlcNAc を グリコシド結合で付加し、*O*-GlcNAc 分解酵素 (OGA) により GlcNAc を解離する。*O*-GlcNAc 修飾を行う 2 つの酵素は核・細胞質に局在している。この事実は、小胞体やゴルジ内などのオルガネラで管理され、細胞表面へと運ばれる *N* 型やムチンタイプの *O* 型糖鎖修飾とは大きく異なる。

OGT のドナーである UDP-GlcNAc は、生体内の高エネルギー小分子として ATP に次いで二番目に多く、ヌクレオチド代謝、グルコース代謝、窒素代謝、脂肪酸代謝からの代謝物とエネルギー量に対する感度が高い栄養センサーである。また、OGT 活性が UDP-GlcNAc 量に依存することから、*O*-GlcNAc 修飾は細胞の栄養状態に密接に関係している。*O*-GlcNAc 修飾の多くは、核・細胞質のタンパク質を対象としたセリンとスレオニンを修飾することから、高頻度でタンパク質のリン酸化と競合している。そのため、*O*-GlcNAc 修飾の主要な役割としては細胞の栄養またはストレスに応じてリン酸化を競合的に阻害することでの生体機能の調節である。

近年、転写因子の多くが *O*-GlcNAc 修飾によって直接調節されていることが報告された。例えば、転写因子 Specificity protein 1 (Sp1) の *O*-GlcNAc 修飾により、複合体となるタンパク質との相互作用を阻害することで、転写調節対象遺伝子の活性化を阻害する他、Sp1 は糖尿病において *O*-GlcNAc 修飾の頻度が増加し、糖尿病に特徴的な遺伝子の転写を促進しながら他の遺伝子の転写を抑制する。転写因子 Signal transducer and activator of transcription 5 (STAT5) は転写のコアクチベータに結合するために *O*-GlcNAc 修飾を必要とする。また、インスリン転写を調節する転写因子 pancreatic and duodenal homeobox 1 (PDX-1) の *O*-GlcNAc 修飾は DNA への結合能を高め、プロインスリンの転写を活性化する。以上のように、*O*-GlcNAc 修飾は転写調節の特異性と活性の調節を行い、多く生物の生体機能に関与していることが示された。

天然変性タンパク質(IDP)は、単独では立体構造を作らない天然変性領域(IDR)を持つ。IDP は真核生物の核に豊富に存在し、転写、翻訳、および細胞周期制御を含む重要な細胞機能に関与している。実際に、IDR においてタンパク質リン酸化やムチンタイプの *O* 型糖鎖修飾が多く起こる。*O*-GlcNAc 修飾は IDP と細胞内局在および機能の面で共通点を有しており、IDR に多く見られるリン酸化との競合をもつ。以上の理由から、IDP における *O*-GlcNAc 修飾は構造的にも機能的にも、

深く関与していると考えた。

2. 研究の目的

本研究では、*O*-GlcNAc 修飾を担う 2 つの酵素である OGT と OGA と *O*-GlcNAc 修飾を受けるタンパク質に着目してこれらの 2 つの視点から進化解析を行う。*O*-GlcNAc 修飾を受けるタンパク質の解析では、ヒト *O*-GlcNAc 修飾タンパク質を進化解析の基準とし、生物間での差異を確認することで *O*-GlcNAc 修飾によるタンパク質の機能調節の解明をする。まず、ヒト *O*-GlcNAc 修飾タンパク質の相同配列検索を行った。また、マルチプルアライメントにより相同タンパク質での *O*-GlcNAc 修飾残基の変化を解析する。この解析をゲノム配列決定済みの真核生物全てを対象として行うことで真核生物における *O*-GlcNAc 修飾の対象の変化を解明する。OGT と OGA の解析では、ヒト OGT/OGA の相同タンパク質を検索することで付加と除去を行う酵素の有無から *O*-GlcNAc 修飾に対する獲得過程の解明を目指す。

3. 研究の方法

(1) *O*-GlcNAc 修飾タンパク質の進化解析 *O*-GlcNAc 修飾タンパク質データセット

O-GlcNAc 修飾タンパク質のデータセットは、OGTsite より取得した。対象は、ヒトタンパク質を基準として進化解析を行うため、ヒトの *O*-GlcNAc 修飾タンパク質のみとした。また、各タンパク質のアミノ酸配列データは UniProt データベースから、ID 領域に関するデータは DICHOT データベース[47]から、リン酸化残基に関するデータは PhosphoSitePlus から取得した。それを、本研究で用いるデータセットとした。

オーソログタンパク質の取得

O-GlcNAc 修飾タンパク質を進化的に解析するため、ヒト *O*-GlcNAc 修飾タンパク質オーソログを KEGG データベースのオーソログクラスターより提供されているオーソログ情報をもとに、327 生物種に対して取得した。この生物種の数には KEGG Organisms の Complete Genomes 一覧 (http://www.kegg.jp/kegg/catalog/org_list.html) に登録されている真核生物とした。

O-GlcNAc 修飾残基の保存

取得されたヒト *O*-GlcNAc 修飾タンパク質オーソログについて *O*-GlcNAc 修飾残基が保存されているか次の方法にて確認した。(1) ヒト *O*-GlcNAc 修飾タンパク質と他生物でのオーソログとのマルチプルアライメントを ClustalW2 により行った。(2) 配列アライメントの後に *O*-GlcNAc 修飾残基の保存度を測る指標として、ヒト *O*-GlcNAc 修飾タンパク質での修飾残基にあたるアミノ酸残基が他生物において他のアミノ酸に置換されていたもの、または欠損していたものは *O*-GlcNAc 修飾を受けないとし、それ以外は

O-GlcNAc 修飾を受けるとした。

(A)

	Human (Query)	Organism 1	Organism 2	Organism 3
Protein A				×
Protein B				×
Protein C				×
Protein D			×	

(B)

	Human (query)	Organism 1	Organism 2	Organism 3
Protein A	1	1	0.5	-1
Protein B	1	1	1	-1
Protein C	1	0	0	-1
Protein D	1	1	-1	0

図 1. 進化情報によるスコア付け方法

図 1(A) オースログタンパク質において O-GlcNAc 修飾残基の保存度が異なる 4 種類のタンパク質の例。各マス内の横長の図形はタンパク質の 1 次構造を表し、図形上のローマ字はアミノ酸残基を表す。ヒトの列における S および T は O-GlcNAc 修飾される残基、他の生物種の列ではヒトの O-GlcNAc 修飾残基にあたる残基を表す。× はオースログが存在しないことを表す。図 1(B) 四種類のタンパク質における算出スコアおよびスコアを基にしたビットパターンをクラスタリングすることで得られる系統樹。オースログタンパク質が存在し、かつ全ての O-GlcNAc 修飾残基が保存されている場合 (例: 生物 1 のタンパク質 A) のスコアは 1 で、オースログタンパク質が存在するものの半分の O-GlcNAc 修飾残基が置換されている場合 (例: 生物 2 のタンパク質 A) のスコアは 0.5、オースログタンパク質が存在するもののどの O-GlcNAc 修飾残基も保存されていない場合 (例: 生物 1 のタンパク質 C) はスコア 0、オースログタンパク質が存在しない場合はスコア -1 とした。また、スコアパターンを見やすくするためにスコア 0 から 1 の場合は赤から青のグラデーション、スコア -1 は白色としたヒートマップで表示した。

(2) OGT と OGA の進化解析

OGT と OGA のオースログタンパク質の取得

O-GlcNAc 修飾タンパク質の進化解析を行うため、ヒト OGT とヒト OGA のそれぞれのオースログを SSDB より提供されているベストヒット情報をもとに取得した。この際、閾値は SW スコア 100 を用いた。ここで、SW スコアとは Smith-Waterman アルゴリズムによるローカルアライメントのスコアのこ

とである。また、検索対象として用いた生物種は 2 - 1 - 2 と同様に KEGG Organisms より 327 種の真核生物を対象とした。

OGT と OGA のドメインの保存

取得された各生物のオースログに対して、OGT・OGA それぞれのどの領域が保存されているかを確認するため、ヒトの OGT・OGA とオースログ検索によって取得した他生物でのオースログタンパク質で ClustalW を用いたマルチプルアライメントによって OGT もしくは OGA のドメイン保存の確認を行った。マルチプルアライメントの結果、ヒト OGT/OGA と同じアミノ酸が当てはめられた場合は該当の生物で保存されているとした。

4. 研究成果

(1) O-GlcNAc 修飾タンパク質の進化解析データセット

O-GlcNAc 修飾部位のデータベース OGTsite より 408 個の O-GlcNAc 修飾残基を含むヒト O-GlcNAc 修飾タンパク質 162 個が取得され、これらをデータセットとした。データセットのタンパク質は Glycophorin-B を除くすべてがリン酸化されることが知られるタンパク質であった。タンパク質全長における IDR の割合は平均で 49.6% であった。それに対して O-GlcNAc 修飾残基は 78.2% が IDR 内に存在した。また、リン酸化と同一のアミノ酸残基で起こるものは 128 ケ所 (31.3%) であった。

進化情報によるクラスタリング

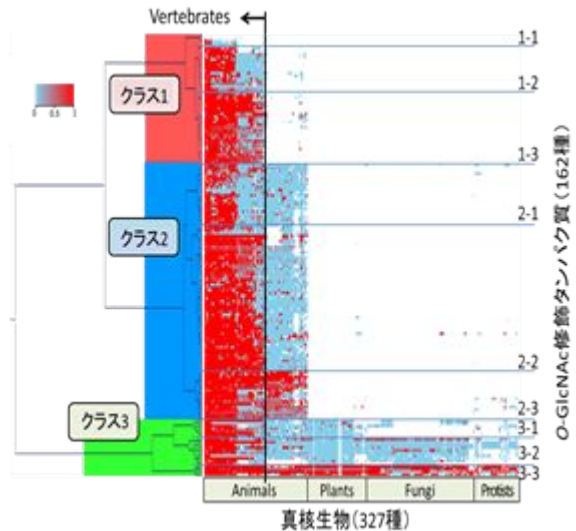


図 2. ヒト O-GlcNAc 修飾タンパク質スコア付け後のクラスタリング

図 2 の縦方向には、ヒト O-GlcNAc 修飾タンパク質がワード法によりクラスタリングされて配置されている。横軸は生物種を表す。クラスタリング後に得られる樹形図をもとに修飾残基をクラス分けした。ヒートマップで表示された保存パターンは左上のスケールに従う。

ヒト *O*-GlcNAc 修飾タンパク質は、進化情報(図1)を基にワード法により、3つのクラスに分類された(図2)。それぞれのクラスは保存されている生物種の少ない順にクラス1、クラス2、クラス3と定義された。47タンパク質からなるクラス1は、脊椎動物で、94タンパク質からなるクラス2は動物のみで、21タンパク質からなるクラス3は真核生物で保存されていた。また、各クラスはさらに3つのサブクラスを取った。4つのタンパク質からなるサブクラス1-1は哺乳類のみで保存されており、*O*-GlcNAc 修飾残基は Glycophorin-B と Perilipin-4 に起こる合計3つの修飾残基のみが霊長類に保存されていた。17個のタンパク質が含まれるサブクラス1-2は脊椎動物で保存されており、*O*-GlcNAc 修飾残基は哺乳類で保存されていた。26個のタンパク質からなるサブクラス1-3は脊椎動物で保存されており、*O*-GlcNAc 修飾残基も脊椎動物で保存されていた。23個のタンパク質からなるサブクラス2-1は動物界で保存されており、*O*-GlcNAc 修飾残基は PDZ and LIM domain protein 5、Microtubule-associated protein 4 (MAP-4)、Transcription initiation factor TFIID subunit 4B の IDR にそれぞれ存在する合計3つの *O*-GlcNAc 修飾残基以外は哺乳類で保存されていた。53個のタンパク質からなるサブクラス2-2は動物で保存されており、*O*-GlcNAc 修飾残基は脊椎動物で保存されていた。18個のタンパク質からなるサブクラス2-3はタンパク質、*O*-GlcNAc 修飾残基共に動物界で保存されていた。7つのタンパク質からなるサブクラス3-1は動物界と植物界で保存されており、*O*-GlcNAc 修飾残基は YTH domain-containing family protein 1 の IDR に存在する1つが霊長類で、その他は脊椎動物で保存されていた。8つのタンパク質からなるサブクラス3-2は真核生物で保存されていた。*O*-GlcNAc 修飾残基は脊椎動物で保存されており、特に Deoxyuridine 5'-triphosphate nucleotidohydrolase の dUTPase ドメインに存在する2つの修飾残基は真菌でも保存されていた。4つのタンパク質からなるサブクラス3-3はタンパク質、*O*-GlcNAc 修飾残基共に真核生物で保存されていた。

O-GlcNAc 修飾タンパク質が動物界で保存されていることが見出された(図2)。事実、動物界以外の真核生物では、クラス3の21タンパク質(13%)のみしか、*O*-GlcNAc 修飾タンパク質は保存されていない。それに対して、*O*-GlcNAc 修飾残基は脊椎動物で保存されていることが見出された。脊椎動物以外の生物で *O*-GlcNAc 修飾残基が保存されているのはサブクラス2-3の18タンパク質(11.1%)とサブクラス3-3の4タンパク質(2.4%)の合計22タンパク質(13.5%)であり、それ以外では *O*-GlcNAc 修飾残基は保存されていない。

天然変性領域と *O*-GlcNAc 修飾の関係先の解析結果(図2)で得られた3クラスに属するタンパク質に対し、天然変性領域(IDR)の割合を算出した(図3)。各クラスのタンパク質が有する IDR の割合はクラス1(51.3%)とクラス2(52.5%)と同等であったが、クラス3では IDR の割合が低い傾向(33.4%)であった(図3A)。さらに、*O*-GlcNAc 修飾残基を注目すると、クラス1、2そして3での修飾残基数は137残基、240残基、31残基で、それらのうち IDR 中で起こるものの割合はそれぞれ85.4%、77.0%、54.8%を示した。すなわち、進化的に新しいと定義したクラスの *O*-GlcNAc 修飾は IDR で起こりやすいと示唆された(図3B)。

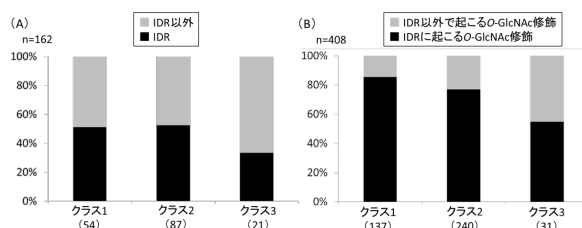


図3. 各クラスのタンパク質に含まれる IDR の割合(A)と IDR で起こる *O*-GlcNAc 修飾残基の割合(B)

(A) 各クラスのタンパク質が有する IDR と構造領域の割合を示す。(B) 各クラスのタンパク質における *O*-GlcNAc 修飾残基の領域別割合を示す。どちらも IDR を黒、構造領域をグレーで表示している。各クラスのカッコ内の数字は(A)はタンパク質、(B)は *O*-GlcNAc 修飾残基の数を示す。

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5件)

Structural analysis of neutral glycosphingolipids from the silkworm *Bombyx mori* and the difference in ceramide composition between larvae and pupae. Itonori S, Hashimoto K, Nakagawa M, Harada M, Suzuki T, Kojima H, Ito M, Sugita M, J Biochem, 査読有, Vol. 163(3), 2018, PP201-214, DOI: 10.1093/jb/mvx072

Characterization of a Novel Rhamnose-containing Acidic Glycosphingolipid from the Ascidian *Halocynthia aurantium*. Kojima H, Shinohara R, Itonori S, Ito M, J Oleo Sci, 査読有, Vol. 66(3), 2017, PP285-295, DOI: 10.5650/jos.ess16150

The glycosylation stoichiometry of EWS species in neuronal cells. Kamemura K, Abe H. Biosci Biotechnol Biochem 査読有, Vol. 81(1), 2017, PP165-167, DOI: 10.1080/09168451.2016.1230004

O-GlcNAc glycosylation stoichiometry of the FET protein family: only EWS is glycosylated with a high stoichiometry. Kamemura K, Biosci Biotechnol Biochem, 査読有, Vol 81(3), 2017, PP541-546, DOI: 10.1080/09168451.2016.1263148

Structural Characterization of Neutral Glycosphingolipids from 3T3-L1 Adipocytes. Kojima H, Suzuki Y, Ito M, Kabayama K, Lipids, 査読有, Vol. 50(9), 2015, PP913-917, doi: 10.1007/s11745-015-4035-7

Investigation of glycan evolution based on a comprehensive analysis of glycosyltransferases using phylogenetic profiling. Tomono T, Kojima H, Fukuchi S, Tohsato Y, Ito M, Biophys Physicobiol, 査読有, Vol.12, 2015, PP57-68, DOI: 10.2142/biophysico.12.0_57

〔学会発表〕(計 37 件)
海外発表 13 件、国内発表 24 件

6 . 研究組織

(1)研究代表者

伊藤 将弘 (ITO, Masahiro)
立命館大学・生命科学部・教授
研究者番号：5 0 3 8 8 1 1 2

(2)研究分担者

亀村 和生 (KAMEMURA, Kazuo)
長浜バイオ大学・バイオサイエンス学部・
准教授
研究者番号：0 0 3 9 9 4 3 7