

Elementary flux mode 型経路の 化学量論的な代謝ネットワーク構造における非分岐直鎖部分構造

太田潤^{†1,a)}

概要：代謝ネットワークにおける反応の集合としての“経路”である elementary flux mode 型経路（EFM 型経路）に可能な化学量論的な代謝ネットワーク構造のそれぞれを EFM 型経路の variant と呼び、EFM 型経路に含まれる反応に関する情報から、EFM 型経路の variant すべてを求めるアルゴリズムを開発してきた。本稿では、EFM 型経路の化学量論的な代謝ネットワーク構造における非分岐直鎖部分構造の算出と算出された非分岐直鎖部分構造の利用による化学量論的な代謝ネットワーク構造 variant 算出の高速化について述べる。

キーワード：代謝ネットワーク, elementary flux mode

Non-branched linear substructures in stoichiometric metabolic network structures of elementary flux mode type pathways

JUN OHTA^{†1,a)}

1. はじめに

代謝経路には、化学量論的に釣り合った個々の反応の集合としての経路 [1,2] がある。反応の集合としての経路の代表例は elementary flux mode [1] と extreme pathway [2] であり、本稿では反応の集合としての経路を elementary flux mode 型経路（EFM 型経路）と呼ぶ。EFM 型経路は、経路を定義する反応の集合に含まれる反応の情報から得られる経路の収支式左辺の原料分子群から経路の収支式右辺の目的分子群が生成する経路と理解できるが、それを定義する反応の集合は、原料分子群から目的分子群への道筋を明示しない。私は、その道筋を示す“化学量論的な代謝ネットワーク構造”を、EFM 型経路を定義する反応の集合に含まれる反応の情報から求める方法を開発してきた [3,4,5]。

EFM 型経路の化学量論的な代謝ネットワーク構造は、経路の化学量論も正確に表現する。仮想原料反応と仮想目的反応（後述）を追加した EFM 型経路の化学量論的な代謝ネットワーク構造においては、代謝産物ノードそれぞれが 2 つの反応ノード（代謝産物ノードに対応する代謝産物分子を直接生成する反応と直接利用する反応に対応する）のみとつながりを持つ [3,4,5]。ここで、“代謝産物ノード”は、代謝産物分子に対応するノード、“反応ノード”は、反応に対応するノードのことである。これまでの代謝産物レベルの生合成代謝経路に関する情報学的研究である、目的代謝産物の生成に必要な原料代謝産物の集合を求める手法の研究 [6]、与えられた代謝ネットワークに存在する生合成経路の列挙を試みる研究 [7] における提案手法は、化学量論を考慮した生合成経路を算出するものではなかった。

EFM 型経路に対して可能な化学量論的な代謝ネットワーク構造のそれぞれを EFM 型経路の variant と呼ぶ [8]。ペントースリン酸経路の非還元的な反応により $2n$ 分子の fructose 6-phosphate と n 分子の glyceraldehyde 3-phosphate から $3n$ 分子の ribose 5-phosphate が生成する EFM 型経路には n が 1, 2, 3, 4 のときそれぞれ 2 種類, 8 種類, 26 種類, 94 種類の variant が存在し [3,4,5], n が大きい場合にそれぞれの variant の構造を細かく把握することが困難であった。そこで、EFM 型経路の異なる複数の variant に共通して存在する可能性のある、化学量論的な代謝ネットワーク構造における非分岐直鎖部分構造を調べることを着想した。ここで、化学量論的な代謝ネットワーク構造における非分岐直鎖部分構造とは、化学量論的な代謝ネットワーク構造の、1 つの代謝産物ノードから 1 つ以上のノードを経由して 1 つの代謝産物ノードに至る非分岐の直鎖状経路とみなされる部分構造全体のことであり、以下では、化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造、あるいは略して非分岐直鎖部分構造と呼ぶ。化学量論的なネットワーク構造の非分岐直鎖部分構造は、1 個の代謝産物分子を基質として利用して 1 個の代謝産物分子を生成する反応から構成される。本稿では、与えられた EFM 型経路の化学量論的な代謝ネットワーク構造の可能な非分岐直鎖部分構造すべての算出と算出された非分岐直鎖部分構造の利用による EFM 型経路の化学量論的な代謝ネットワーク構造 variant 算出の高速化について述べる。

2. 方法

EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造算出の理論的な枠組み、EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造算

^{†1} 岡山大学 学術研究院医歯薬学域 (医) 生化学分野
a) jo25@md.okayama-u.ac.jp

出のためのアルゴリズム, EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造の異同判定, 非分岐直鎖部分構造を利用する EFM 型経路の化学量論的な代謝ネットワーク構造算出について述べる.

2.1 EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造算出の理論的な枠組み

EFM 型経路の化学量論的な代謝ネットワーク構造算出に関して [3,4] に述べて [9] に“EFM 型経路の化学量論的な代謝ネットワーク構造算出の理論的な枠組み”としてまとめた考え方を非分岐直鎖部分構造の算出に応用する.

反応の集合として与えられる EFM 型経路全体の収支から, 経路の化学量論的な原料代謝産物 (原料分子群) と目的代謝産物 (目的分子群) を求め, 原料代謝産物を何もないうところから生成する仮想原料反応と目的代謝産物を除去するが何も生成しない仮想目的反応を定義する. 以下では, この 2 つの仮想反応を, 与えられた EFM 型経路を定義する反応の集合に追加して生成する, 単一の仮想原料反応から単一の仮想目的反応への EFM 型経路を考える.

EFM 型経路の化学量論的な代謝ネットワーク構造において, 反応ノードは, EFM 型経路に現れる反応と 1 対 1 に対応し, 各代謝産物ノードは, 反応ノードからその代謝産物ノードに入るエッジの数, その代謝産物ノードから反応ノードに向けて出るエッジの数ともに 1 である. したがって, EFM 型経路の化学量論的なネットワーク構造においてエッジを介して反応ノードから出る代謝産物ノードの集合を P , エッジを介して反応ノードに入る代謝産物ノードの集合を S とすると, 集合 P と集合 S は代謝産物ノードの集合としては同一の集合である. すなわち, 集合 P の要素 (どの反応ノードから出るかにより区別される代謝産物ノード) は集合 S の要素 (どの反応ノードに入るかにより区別される代謝産物ノード) と同一ノードとして 1 対 1 に対応付けられる. P の特定の要素と S の特定の要素は, 両方が同じ種類の代謝産物のノードである場合に限り同一ノードとして対応付けできる (“対応付け条件”). 化学量論的な代謝ネットワーク構造とその非分岐直鎖部分構造はこの “対応付け条件” を満たす P の要素と S の要素の対応付けにより生成する. すなわち, “対応付け条件” を満たす P の要素と S の要素の対応付けには, それに対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造が存在する. 以下, “ P の要素と S の要素の対応付け” の意味で “ P - S 間対応付け” を用いる.

P と S を, 要素 (代謝産物ノード) の代謝産物種 (代謝産物の種類) ごとに部分集合に分割し, P と S の同じ代謝産物種の部分集合間で要素を対応付ける (同一とみなす) ことを考える. このような対応付けは “対応付け条件” を明らかに満たす. “対応付け条件” を満たす任意の 1 つの P - S 間対応付けを x とすると, “対応付け条件” を満たすすべての P - S 間対応付けは x により対応付けられた S の要素

の, 代謝産物種ごとの置換と 1 対 1 に対応する.

“ x により対応付けられた S の要素の置換により得られる P - S 間対応付け” を “ x の置換” と呼ぶと, 原理的には “対応付け条件” を満たす任意の 1 つの P - S 間対応付け x を求め, “対応付け条件” を満たす x の置換をすべて生成すると, 可能な化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造がすべて x の置換に対応して求まる.

2.2 EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造算出のためのアルゴリズム

[4] に “アルゴリズム B” として示した考え方を非分岐直鎖部分構造の算出に応用する.

P - S 間対応付け x に対して, 本項でも前項と同様に, “ x により対応付けられた S の要素の置換により得られる P - S 間対応付け” を “ x の置換” と呼び, さらに, “ x により対応付けられた S の要素の互換により得られる P - S 間対応付け” を “ x の互換” と呼ぶ. また, “ x の置換” と “ x の互換” は, それぞれ, “ x により対応付けられた S の要素の置換”, “ x により対応付けられた S の要素の互換” の意味でも用いる. 互換が置換に含まれることから, “ x の互換” は “ x の置換” でもある.

このアルゴリズムでは, 最初に “対応付け条件” を満たす P - S 間対応付けの 1 つ x を求め, x に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造を空集合 V に要素として入れる. 数学的に, 任意の置換は, 互換の組み合わせで表現できるので, “対応付け条件” を満たす P - S 間対応付けを生成する x の互換を繰り返し (x に可能な, “対応付け条件” を満たす P - S 間対応付けを生成する 1 回の互換をすべて行い, 得られた x の置換それぞれに対して可能な, “対応付け条件” を満たす P - S 間対応付けを生成する 1 回の互換をすべて行い, さらに, 得られた x の置換それぞれに対して可能な, “対応付け条件” を満たす P - S 間対応付けを生成する 1 回の互換をすべて行うことを繰り返し), “対応付け条件” を満たす x の置換 (P - S 間対応付け) を生成する. その際, 各互換で x の置換 y が得られるごとに, その x の置換 y (P - S 間対応付けである) に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造 p を求めて, V の要素の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造との異同判定を行い, 次のルールに従って, x の置換 y の互換をやめるか y に次の互換を行うかを決める.

- p が V の要素であれば, x の置換 y の互換をやめる.
- p が V の要素でなければ, p を V に要素として入れ, x の置換 y に次の互換を行う.

互換により V に含まれない新たな化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造 p が生成されなくなるまで互換を続けると, EFM 型経路のすべての可能な化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造が V の要素として得られる.

V の要素に前述の手順で得られる順に番号を付し、 i 番目に得られた V の要素を $V[i]$ と表す。 $V[i]$ は、 x の特定の置換に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造として求めたものである。この、 $V[i]$ に対応する($V[i]$ を生成した) x の特定の置換を $x(i)$ で表す。また、 x の置換 y に対応する化学量論的な代謝ネットワーク構造を $n(y)$ 、その非分岐直鎖部分構造を $p(y)$ と表す。これらの記法では、 $V[i] = p(x(i))$ 、 $V[1] = p(x(1))$ 、 $x = x(1)$ である。

以下の記述における“置換”と“互換”はそれぞれ“対応付け条件”を満たす P - S 間対応付けを生成する“置換”と“対応付け条件”を満たす P - S 間対応付けを生成する“互換”を意味するものとする。

x の置換 a 、 b に対して $p(a) = p(b)$ のとき、 $p(a)$ と $p(b)$ の同型条件を満たす全単射 f [5]により、 $p(a)$ のノードと $p(b)$ のノードを1対1に対応付けできる。 $p(a)$ に含まれない $n(a)$ のノードについては f がそれらを $p(b)$ に含まれない $n(b)$ のノードに1対1に対応付けるものとする、 f により $n(a)$ のノードと $n(b)$ のノードが1対1に対応付けられる。この f により $n(a)$ の2代謝産物ノード a_i, a_j に対応付けられた $n(b)$ の2代謝産物ノードを b_i, b_j とし(a_i, a_j, b_i, b_j は同じ種類の代謝産物分子とする)、 x の置換 a における a_i, a_j に対応する S の要素の互換を $a(a_i, a_j)$ 、 x の置換 b における b_i, b_j に対応する S の要素の互換を $b(b_i, b_j)$ と表すと、 $p(a) = p(b)$ のとき、 $p(a(a_i, a_j)) = p(b(b_i, b_j))$ なので、 $p(a) = p(b)$ であるときには $\{p(z)|zはaの互換\}$ の要素と $\{p(z)|zはbの互換\}$ の要素が、同一(の構造)であるという関係により1対1に対応付けできる。したがって、一般に、 x の置換 a, b に対して、

$$p(a) = p(b) \text{ならば } \{p(z)|zはaの互換\} = \{p(z)|zはbの互換\}$$

が成り立つ。

よって、前述の手順で最終的に得られる V と x の置換 y に対して $p(y) \in V$ ならば、 $p(y) = V[m] = p(x(m))$ となる m が存在して、

$$\{p(z)|zはyの互換\} = \{p(z)|zはx(m)の互換\}$$

が成り立つ。 $V[m]$ を生成した x の置換が $x(m)$ であるので、 $x(m)$ の互換はすべて V に含まれ、

$$\{p(z)|zはx(m)の互換\} \subset V$$

が成り立つ。したがって、

$$\{p(z)|zはyの互換\} \subset V$$

が成り立つ。すなわち、 x の置換 y に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造が V に含まれていれば、 y の互換に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造すべてが V に含まれる。

したがって、 $x = x(1)$ (x の恒等置換)に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造 $V[1]$ が

V に含まれていることと合わせて、数学的帰納法により、 x の任意の置換(互換の組み合わせで表される)に対応する化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造はすべてが V に含まれることが証明される。

2.3 EFM型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造の異同判定

前に述べたように、化学量論的な代謝ネットワーク構造における(化学量論的な代謝ネットワーク構造の)非分岐直鎖部分構造とは、化学量論的な代謝ネットワーク構造の、1つの代謝産物ノードから1つ以上のノードを経由して1つの代謝産物ノードに至る非分岐の直鎖状経路とみなされる部分構造全体のことである。EFM型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造の異同判定は、非分岐直鎖部分構造を構成する1つの代謝産物ノードから1つの代謝産物ノードに至る非分岐の直鎖状経路の比較により行う。比較されるべき非分岐の直鎖状経路は、2個以上の代謝産物分子を基質あるいは生成物とする反応および仮想原料反応と仮想目的反応に対応する反応ノードの間を結ぶがそれらの反応ノードを経由しない非分岐の直鎖状経路を計算し経路両端の反応ノードを除くことにより算出できる。

2.4 非分岐直鎖部分構造を利用するEFM型経路の化学量論的な代謝ネットワーク構造算出

EFM型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造に含まれる、1つの代謝産物ノードから1つ以上のノードを経由して1つの代謝産物ノードに至る非分岐の直鎖状経路を、仮想的に1個の代謝産物分子を基質として利用して1個の代謝産物分子を生成する反応とみなすと、EFM型経路(2.1の“単一の仮想原料反応から単一の仮想目的反応へのEFM型経路”)は、非分岐直鎖部分構造に含まれる非分岐の直鎖状経路に対応する仮想反応と2個以上の代謝産物分子を基質あるいは生成物とする反応、仮想原料反応と仮想目的反応からなる経路に変換される。この変換後の、非分岐直鎖部分構造に含まれる非分岐の直鎖状経路に対応する仮想反応を含むEFM型経路に対して化学量論的な代謝ネットワーク構造を算出する。

EFM型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造が2種類以上ある場合には、非分岐直鎖部分構造ごとに、上記の変換後の、非分岐直鎖部分構造に含まれる非分岐の直鎖状経路に対応する仮想反応を含むEFM型経路に対して化学量論的な代謝ネットワーク構造を算出する。

3. 計算実験の結果と考察

糖質代謝モデルネットワーク [10,11] のペントースリン酸経路に含まれる2種類のtransketolase反応・transaldolase反応・ribulose-5-phosphate 3-epimerase反応の計4反応(文献 [10,11] のプロセス番号 26, 28, 29, 30)の逆反応と

ribose-5-phosphate ketoisomerase 反応 [文献 [10,11] のプロセス番号 27) の組み合わせ (計 5 反応, 延べ 7 反応) により化学量論的に 2 分子の fructose 6-phosphate と 1 分子の glyceraldehyde 3-phosphate から 3 分子の ribose 5-phosphate を生成する EFM 型経路を EFM1 と呼び, EFM1 に含まれる反応を, それぞれの数を n 倍にして組み合わせた, 化学量論的に $2n$ 分子の fructose 6-phosphate と n 分子の glyceraldehyde 3-phosphate から $3n$ 分子の ribose 5-phosphate を生成する EFM 型経路を EFM n と呼ぶ。

EFM 型経路である EFM1, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造を算出した。1 で述べたように, EFM1, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造にはそれぞれ 2 種類, 8 種類, 26 種類, 94 種類の variant が存在するが, EFM1, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造は何れも 1 種類のみであった。EFM1, EFM2, EFM3, EFM4 では, 同じ非分岐の直鎖状経路が異なるつながり方をすることにより異なる variant が生じていることが明らかとなった。

EFM1, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造を利用して既報のアルゴリズム B+付加ステップ [4] (グラフ同型判定アルゴリズムは_{+ECp}GCORE [5]) による算出を行ったところそれぞれ 2 種類, 8 種類, 26 種類, 94 種類の variant がそれぞれ 0.08 秒, 1.18 秒, 15.0 秒, 140 秒の計算時間 (非分岐直鎖部分構造の計算時間を含む) で算出された。一方, EFM1, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造を利用せず既報のアルゴリズム B+付加ステップ [4] (グラフ同型判定アルゴリズムは_{+ECp}GCORE [5]) による算出を行ったところそれぞれ 2 種類, 8 種類, 26 種類, 94 種類の variant がそれぞれ 0.06 秒, 1.46 秒, 21.3 秒, 204 秒の計算時間で算出された。非分岐直鎖部分構造の利用により, EFM2, EFM3, EFM4 の化学量論的な代謝ネットワーク構造はより速く算出することが可能であった。これらの計算はすべて Matlab 上で行った (Intel(R) Core(TM) i5-3320M CPU @ 2.60GHz)。

以上より, EFM 型経路の化学量論的な代謝ネットワーク構造の非分岐直鎖部分構造の計算は EFM 型経路の variant の構造の理解と variant 算出の高速化に有用であると考えられる。

参考文献

[1] Schuster, S.; Hilgetag, C. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* **1994**, *2*, 165-182.
 [2] Schilling, C. H.; Palsson, B.O. The underlying pathway structure of biochemical reaction networks. *Proc. Nat. Acad. Sci. USA* **1998**, *95*, 4193-4198.
 [3] 太田 潤 Elementary flux mode 型代謝経路の化学量論的に釣り合った代謝産物レベルネットワーク構造の算出. 情報処理

学会研究報告, **2020**, Vol.2020-BIO-62, No. 3, 1-6.
 [4] 太田 潤 Elementary flux mode 型代謝経路の化学量論的に釣り合った代謝産物レベルネットワーク構造算出の高速化. 情報処理学会研究報告, **2020**, Vol.2020-BIO-63, No. 1, 1-7.
 [5] 太田 潤 Elementary flux mode 型代謝経路の化学量論的な代謝ネットワーク構造の算出: 算出されたネットワーク構造間の異同判定のためのアルゴリズムの最適化. 情報処理学会研究報告, **2021**, Vol.2021-BIO-65, No. 4, 1-8.
 [6] Acuña, V.; Milreu, P.V.; Cottret, L.; Marchetti-Spaccamela, A.; Stougie, L.; Sagot, M.F. Algorithms and complexity of enumerating minimal precursor sets in genome-wide metabolic networks. *Bioinformatics* **2012**, *28*, 2474-83.
 [7] Ravikrishnan, A.; Nasre, M.; Raman, K. Enumerating all possible biosynthetic pathways in metabolic networks. *Sci. Rep.* **2018**, *8*, 9932. DOI: 10.1038/s41598-018-28007-7
 [8] 太田 潤 Elementary flux mode 型経路における代謝産物分子の生成・利用状態のスナップショットとしての化学量論的な代謝ネットワーク構造. 情報処理学会研究報告, **2021**, Vol.2021-BIO-66 No.22, 1-7.
 [9] 太田 潤 Elementary flux mode 型経路の化学量論的な代謝ネットワーク構造 variant の対称性の数. 情報処理学会研究報告, **2021**, Vol.2021-BIO-68 No.4, 1-3.
 [10] Ohta, J. Connectivity matrix method for analyses of biological networks and its application to atom-level analysis of a model network of carbohydrate metabolism. *IEE Proc. Syst. Biol.* **2006**, *153*, 372-374. DOI: 10.1049/ip-syb:20060018
 [11] Ohta, J. Single-atom tracing in a model network of carbohydrate metabolism and pathway selection. *IPSJ Transactions on Bioinformatics* **2018**, *11*, 1-13.