

修士論文

KrevelenPlot の拡張による
代謝物質の悉皆的解析への応用法

松田 かおり

2009 年 2 月 5 日

奈良先端科学技術大学院大学
情報科学研究科 情報生命科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学)授与の要件として提出した修士論文である。

松田かおり

審査委員：

金谷重彦教授(主指導教員)

湊小太郎教授(副指導教員)

MD. ALTAF-UL-AMIN 准教授 (副指導教員)

KrevelenPlot の拡張による代謝物質の悉皆的解析への応用法

松田 かおり

内容梗概

KrevelenPlot は 2 次元 (2D) の散布図で、石炭の構造と化学組成を研究するために開発された。この研究の目的は KrevelenPlot の拡張によって様々な代謝物質の代謝反応を視覚的に比較可能とすることにある。メタボロミクスでは遺伝子との相関を知るために代謝物質の悉皆的な解析が大変重要であり、質量分析などの進展によって代謝物質の精密な化学的分析が可能となった。しかしながら未だに代謝反応経路における代謝物質について多くは未解明である。既知の代謝物質の構造の特徴と反応経路の分類は未知の代謝物質の同定に役立つ。先行研究によると、KrevelenPlot は代謝経路の全体像を示すことができる。ほとんどの代謝物質は主要元素の組成比率 (H/C や O/C) に基づいて KrevelenPlot 上に描画できるため、それらの分類、すなわち種類と構造的特徴を理解できる。また代謝反応の前と後の代謝物質の座標を線分で結ぶことで代謝反応を表現できる。一方、2 次元のため座標値が重複しやすく主要元素以外は把握できないという表現限界がある。KrevelenPlot を既存の 2D から 3D に拡張することで、新しいグループやサブグループを判別できる。そこで本研究では拡張した KrevelenPlot を用いて官能基の種類や反応の種類、化学種などで代謝物質のより多くの詳細な特徴を理解しようと試みた。これを用いて代謝物質の構造と種類によって特定のパターンに従う座標値と、代謝物質と酵素の反応によって特定のパターンに従うベクトルを調査した。加えて既知の代謝経路の反応パスの特徴を視覚的に比較できることを示す。最後に調査結果により特定の化学種を指す座標値とベクトルを分類できることを示す。また、精密分子量を用いた化学種の予測に本調査結果が応用可能である。

キーワード

代謝物質、代謝反応および酵素反応の整理、ケモインフォマティクス

* 奈良先端科学技術大学院大学情報科学研究科情報生命科学専攻修士論文
NAIST-IS-MT0851110, 2009 年 2 月 5 日

Applied method of widening KrevelenPlot for exhaustive analysis of metabolite

Kaori Matsuda

Abstract

The Krevelen plot is a two dimensional (2D) scatter plot which was developed for studying coal's structure and chemical composition. The objective of this work is to make possible visual comparison of reaction paths of various metabolites by extended Krevelen plots. In metabolomics the exhaustive analysis of metabolites and their relations with genes is very important. Precise chemical analysis of metabolites is possible by advanced mass spectrometry etc. However, still a lot of things are unknown about the metabolites in the context of metabolic pathways. It is useful to classify the well-known metabolites according to features of the structure and reaction pathways for the identification of unknown metabolites. According to earlier studies, the Krevelen plot can present a broad overview of metabolic pathways. Most metabolites can be represented in a Krevelen plot on the basis of the ratio of the major elements (e.g. H/C, O/C etc.) and thus their classification i.e. type and structural features can be understood. However, if a Krevelen plot is extended from 2D to 3D new groups or subgroups could be identified. Accordingly, in this study we try to apprehend more features of metabolites which might be a type of functional group, or type of reaction, chemical species, and so on, by enhanced Krevelen plots. Using it, We investigate the coordinates which follows a specific pattern by the structure and species of metabolites, and vectors which follows a specific pattern by the reaction of metabolites and enzymes. Moreover, we show that it can reveal the perspective by visual comparison of the reaction path of well-known metabolic pathways. At the end, according to the result of the survey we indicate that we can classify the coordinates and the vectors which imply a specific species. Results of this analysis can be applied to the estimation of chemical species from high-resolution molecular mass spectrum.

Keywords:

Metabolite, Classification of metabolic and enzyme response, ChemoInformatics

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0851110, February 5, 2009

目次

第1章	序論	1
第2章	背景と目的	3
2.1	背景	3
2.1.1	メタボロミクスへの期待	3
2.1.2	質量分析による代謝物質の同定	4
2.1.3	グラフによる代謝反応の解析手法	6
2.2	目的	9
第3章	材料および方法	10
3.1	材料	10
3.1.1	化学物質（化学式）データ	10
3.1.2	酵素反応分類	11
3.1.3	代謝物質の生合成経路	11
3.2	方法	12
3.2.1	KrevelenPlot の拡張による代謝物質の解析	12
(1)	3次元化と用いる軸	12
(2)	代謝物質の特徴を捉える軸	13
3.2.2	KrevelenPlot による代謝物質の特徴整理	15
(1)	母核の構造的特徴による座標の分離	15
(2)	官能基による座標の分離	22
3.2.3	KrevelenPlot による代謝反応の視覚化	23
(1)	代謝反応の整理	23
(2)	酵素反応の整理	24
3.2.4	既知の代謝反応経路による検証	25
第4章	結果と考察	26
4.1	KrevelenPlot による代謝物質の表現	26
4.1.1	データの特徴	28
4.1.2	KrevelenPlot の拡張 一次元	32
4.1.3	KrevelenPlot の拡張 一軸	35
4.2	KrevelenPlot による代謝物質の特徴整理	40
4.2.1	母核の構造的特徴による座標の分離	40
4.2.2	官能基による座標の分離	54
4.3	KrevelenPlot による代謝反応の視覚化	63
4.3.1	代謝反応の整理	63
4.3.2	酵素反応の整理	69
4.4	既知の代謝反応経路による検証	71

第 5 章 結論：今後の研究課題.....	74
謝辞	75
参考文献	76
Appendix(1) 軸の組合せによる代謝物質の表現	77
Appendix(2) 構造から見る天然物の全体像	89
Appendix(3) 代謝反応	92
Appendix(4) 酵素反応	96
Appendix(5) 生合成反応	100

第1章 序論

2000 年 12 月に植物で初めてシロイヌナズナのゲノムが解読されて以来、メタボロミクスはポストゲノム科学の 1 分野として注目されるようになった。ポストゲノム科学とは、ゲノミクス・トランスクリプトミクス・プロテオミクス・メタボロミクスなどのオミクス(-omics)研究から成り、生物界のセントラルドグマに従って、ゲノム情報から RNA 情報 (トランスクリプトーム)、タンパク質情報 (プロテオーム)、代謝物質情報 (メタボローム) への流れに沿って生物システムを網羅的な解析によって理解しようとするアプローチである。ゲノミクスは遺伝子、トランスクリプトミクスはメッセンジャーRNA、プロテオミクスは発現しているタンパク質、メタボロミクスは代謝物質を扱う。これらの研究成果は相補的であり、生体内環境の微妙な変化を体系的かつ包括的に解明出来る。

メタボロミクスでは酵素が触媒する有機化学反応の組合せによって生じた代謝物質の総体を解析する。代謝物質とは生物の生命活動により生じる化学物質であり、生命維持に必須の一次代謝物質とそれ以外の二次代謝物質がある。一次代謝物質は生物間で互換性を有するので、ゲノム情報が利用できない実用植物や実用微生物にも適用可能である。つまり代謝物質を扱うメタボロミクスはゲノム情報の発現を高解像度に解析できる手法といえ、その可能性は大きい。二次代謝物質は限られた種の生物が産出する化合物で代表的な例として植物、微生物由来のアルカロイド、テルペノイド、フラボノイド、抗生物質などがある。これら二次代謝物質の構造は多種多様なため種の個性の表現型ともいわれている。また、薬用、香料など人の生活に欠かせない物質が多く存在することもあり、新規有用化学物質の探索のためにも代謝物質の解析に寄せられる期待は大きい。植物における総代謝産物は 20 万種とも言われている中で化学構造が解明されたものはうち 5 万件、代謝経路が解明されたものは 5 千件にとどまっている。これは、観測対象の代謝物質の化学的性質が多岐に渡るため、手法の標準化が困難であることが一因である。標準技術が確立されていないことが、メタボロミクスが広く浸透するうえでの課題となっている。

メタボロミクス解析においては主に質量分析が代謝物質の解析に用いられている。近

年、質量分析機器の性能向上が目覚ましく多様な試料の解析が可能となった。またその解析結果の精密分子量から化学組成式を推定する情報処理技術も発達している。精密分子量と一致する既知の代謝物質があれば、その既知の代謝物質である可能性が高まる。また精密分子量が一致しなくても、既知の代謝物質の精密分子量に近い値であればその類縁体と予測したり、元素の組合せから考えられる理論化学式が候補として挙げられる。こうした候補を基に、質量分析機器や NMR などの別の解析結果と照らし合わせて代謝物質の化学式や構造が同定されている。

代謝物質の解析は進みつつあるが得られるデータは膨大である。データを悉皆的に解析し、代謝反応の解明などに役立てるための情報処理技術が求められている。そこで化学組成式のみを用いて解析できる KrevelenPlot は有効なツールである。KrevelenPlot は 1950 年に van Krevelen により提案された 2 次元の散布図で、石炭の組成式と構造に関する問題を研究する目的で用いられた。主要元素の組成比を X、Y 軸とし化学組成式から得た比率を座標値とする。ほとんどの代謝物質を表現でき、代謝物質の種類により座標領域が分離するため座標から化合物の種類や構造的な特徴を把握できる。また代謝反応前と後の代謝物質の座標を線分で結ぶことで代謝反応を表現できる。一方、2 次元のため座標値が重複しやすく、CHO 以外の構成元素は記述に反映されないという表現限界がある。

本修士論文では、この KrevelenPlot を既知の代謝物質に適用して代謝物質と代謝反応の整理を行い、目視による代謝物質の構造的な特徴や種類の把握と代謝反応の種類の判別を可能とし、グラフ上で反応経路の全体像を捉えられるようにすることを目標とする。まず様々な特徴を捉えられるよう KrevelenPlot を拡張した上で、代謝物質と代謝反応を整理する。その上で KrevelenPlot で代謝物質を表現することでその座標値からその構造的な特徴や官能基を目視により判別できるか否かを検証する。

第2章 背景と目的

2.1 背景

2.1.1 メタボロミクスへの期待

代謝物質は 20 万種とも言われており、未知の化学構造の宝庫と考えられている。生物の生命活動によって生じる代謝物質は 2 種類に大別できる。生命維持に必須の一次代謝物質とそれ以外の二次代謝物質である。植物の一次代謝物質の種類は細菌と同程度であり、多様性は二次代謝物質に大きく依存している[8]。図 2-1-1-1 に炭素の代謝経路と生成される代謝物質を示す。

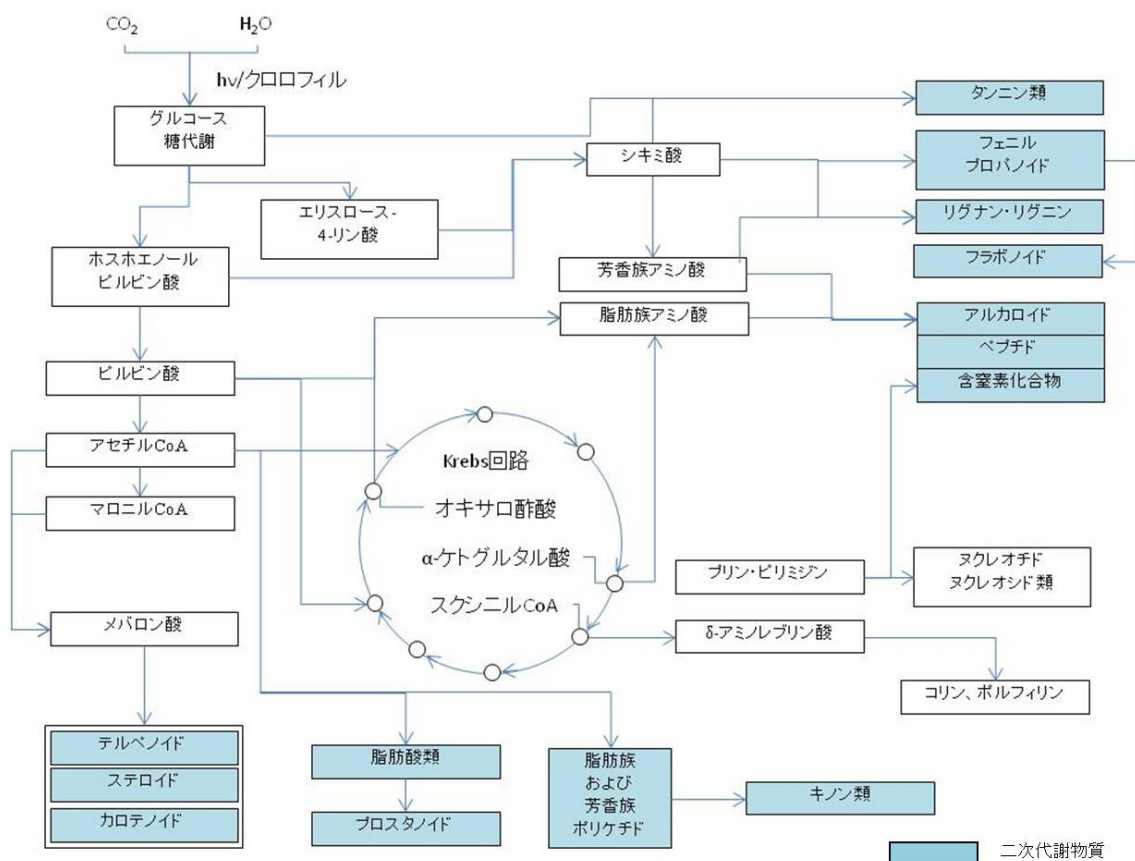


図 2-1-1-1 炭素の代謝経路と主要な二次代謝経路

この多様性に注目している例として創薬が挙げられる。古来、民間薬として用いられてきた化学物質の多くも二次代謝物質であり、既知の二次代謝物質をベースに化学的な修飾を変えたり加えたりすることで目的に応じた作用や薬効を表すように改良されてきた薬も多い。全く新しいタイプの薬を開発する場合、新たな化学物質を創薬リード物質（新薬として開発する価値のある、薬効を有する化学構造）として選定することから始まる。実験的にはコンビナトリアルケミストリーで人工的に多様な化学物質を生成して生体への適合具合をスクリーニングする手法が試みられてきた[8]。しかし、合成反応の難しさなどもあって人工的に作る化学物質の種類は限られてしまい、本来の意味で多様性を示してはいないことが指摘されている[2]（図 2-1-1-2）。このため、創薬におけるリード物質の探索対象として、種類も多く多様性を有する二次代謝物質を用いることが期待されている。

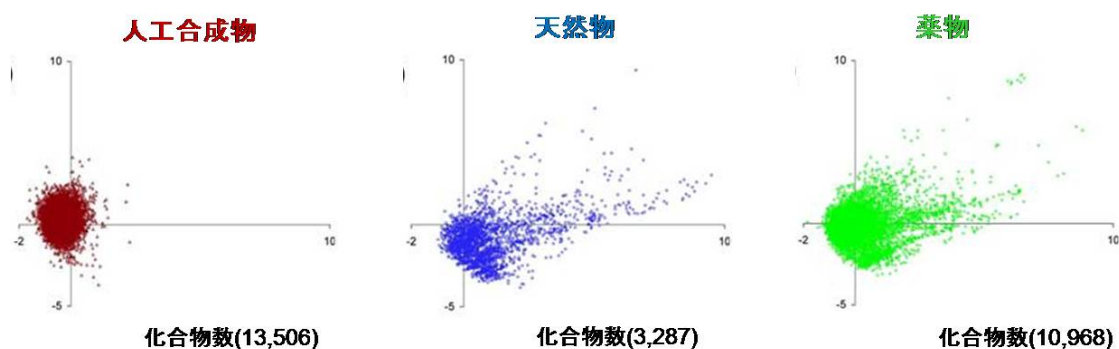


図 2-1-1-2 化合物の多様性の比較（主成分分析による）

2.1.2 質量分析による代謝物質の同定

メタボロミクスでは様々な化合物の混合したサンプルを解析する。そこで代謝物質の測定に質量分析機器が用いられる。これはサンプルに含まれる化学物質をその質量によって分離して測定する装置であるが、質量分析機器の性能向上に伴い高感度の結果が得られるようになってきた。近年開発されたFT-ICR MS（フーリエ変換イオンサイクロトロン質量分析装置）は最も高分解能なマススペクトルデータがえられる。高分解能とは

非常に近接したマスペクトルのピーク値（質量電荷比：m/z）を示す異なる元素組成で構成される別々の化学物質がそれぞれ別のピークとして捉えられることを指す。つまり FT-ICR MS では質量数の近似した化学物質をその精密分子量の差異により区別することができる。

さらに前処理の分離プロセスを省略して直接サンプルの解析も可能なため実験手続きが簡略化されるだけでなく、生体から抽出したサンプルを速やかに処理することで生体内での条件に近い状態での代謝物解析を行える。そして1回の測定でも膨大な情報を得られる。FT-IMS により得た精密分子量から該当する化学式を割り出す方法としては、既知の化学物質とのマッチングが最も確実である。この手法は KNApSAcK でも提供されている（図 2-1-2-1）。既知の化学物質のデータと精密分子量が一致しない場合も、理論化学式の導出が可能である。各元素の精密質量数の組合せにより該当する化学式を割り出せる。理論的な化学組成式の候補の中から構造を含めた詳しい情報を得るには再び質量分析にかける手法がある。質量分析ではサンプルをイオン化するため、サンプルに高エネルギーの電子ビームが照射される。サンプルの化学構造の共有結合はこのエネルギーによって切断され、あらゆる中性およびイオン化フラグメントの組合せに分解される。この時、元の物質より低質量にピーク値が現れる。このフラグメントのパターンを見ることで、元の化学組成式や構造を特定できる[6]。しかし、膨大なマスペクトルのデータを扱うことを考えると、情報処理によって質量ピークから効率よく理論的に可能な化学組成式の候補を割り出していく工夫が必要である。

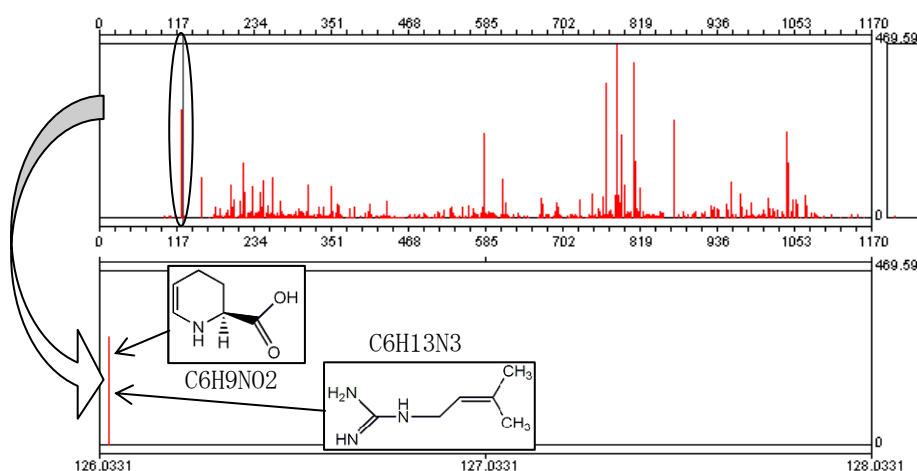


図 2-1-2-1 マスペクトルデータからの化学組成式の推定

2.1.3 グラフによる代謝反応の解析手法

代謝物質は炭素 (C) が主要元素であり、次いで水素(H)、酸素(O)がその化学組成の大部分を占めている。van Krevelen によって提唱された van Krevelen Plot は石炭³⁵に含まれる化合物群の構造と組成を研究するために開発されたグラフ表現の手法である。対象とする化学物質の化学組成式を用い、そこに含まれる C、H、O の原子数を用いて座標値を得る (図 2-1-5-1)。X 軸に (酸素原子の数/炭素原子の数 : O/C) をとり、Y 軸に (水素原子の数/炭素原子の数 : H/C) をとった 2 次元のグラフに有機物をプロットした散布図を描くことができ、わずかな化学組成式の変化を座標値の違いから捉えられる [1]。

表 2-1-3-1 化学物質の組成数と座標値

化学式	C	H	O	X 軸 (O/C)	Y 軸 (H/C)
C ₁₉ H ₂₄ O ₆	19	24	6	0.3157894	1.2631578
C ₂₀ H ₂₈ O ₅	20	28	5	0.25	1.4
C ₆ H ₆	6	6	0	0	1.0

先行研究において、van Krevelen Diagram (KrevelenPlot) は高分解能な質量分析データ (マスマスペクトルデータ) の解析に用いられている [3]。KrevelenPlot 上では既知の脂質、セルロース、リグニン、縮合芳香族多環式炭化水素のような生物由来の化合物がそれぞれの化合物の特徴的な組成比により座標内の特定の部位部位に分離される (図 2-1-3-1) [4]。

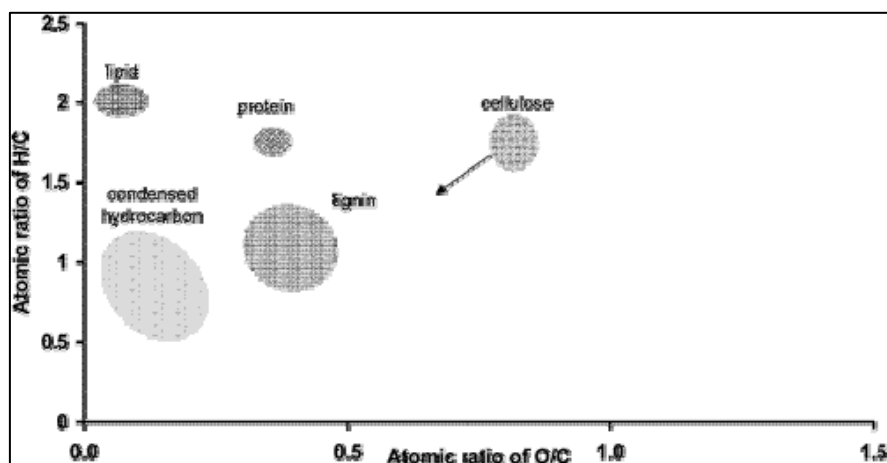


図 2-1-3-1 KrevelenPlot で表現される有機物

また、質量分析結果からはピーク強度 (peak intensities) により定量的な情報を得ることができる。図 2-1-3-2 に、マスペクトルデータから得た化学式を KrevelenPlot にプロットする際にピーク強度によって色を変えて表現する方法に加え、KrevelenPlot に Z 軸としてピーク強度を加えた 3 次元表現により KrevelenPlot による解析を拡張した例を示している。3 次元の KrevelenPlot によりサンプルに含まれる複数の類似した化合物のピーク強度の比較がしやすくなることを示している。

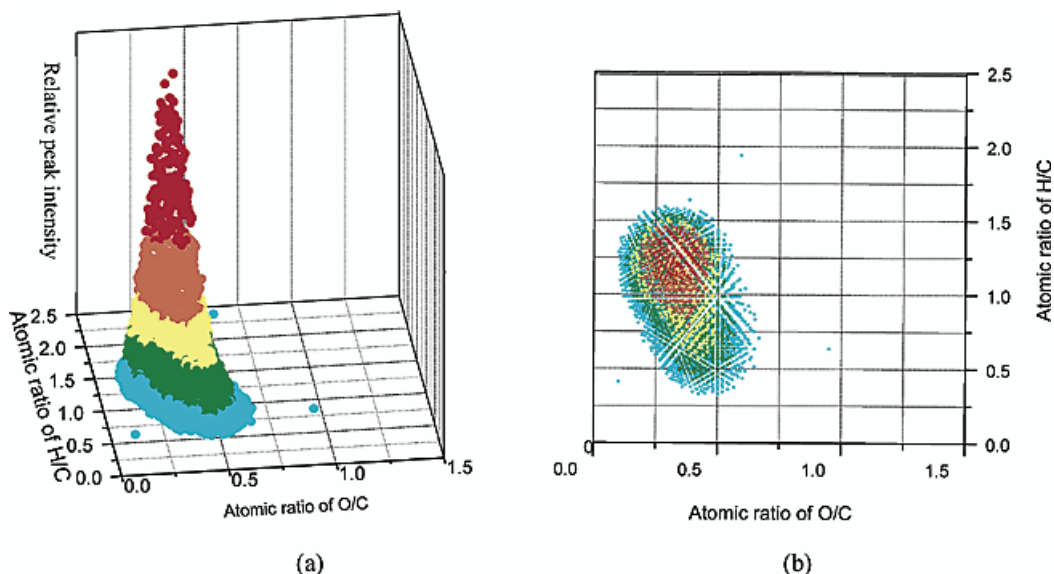


図 2-1-3-2 KrevelenPlot の 2D と 3D による表現比較

また、プロットされた化学物質同士の化学変化、つまり代謝反応を視覚化することも

可能である。その代謝反応は図 2-1-3-3 に示されたようにベクトルとして表現出来る[5]。

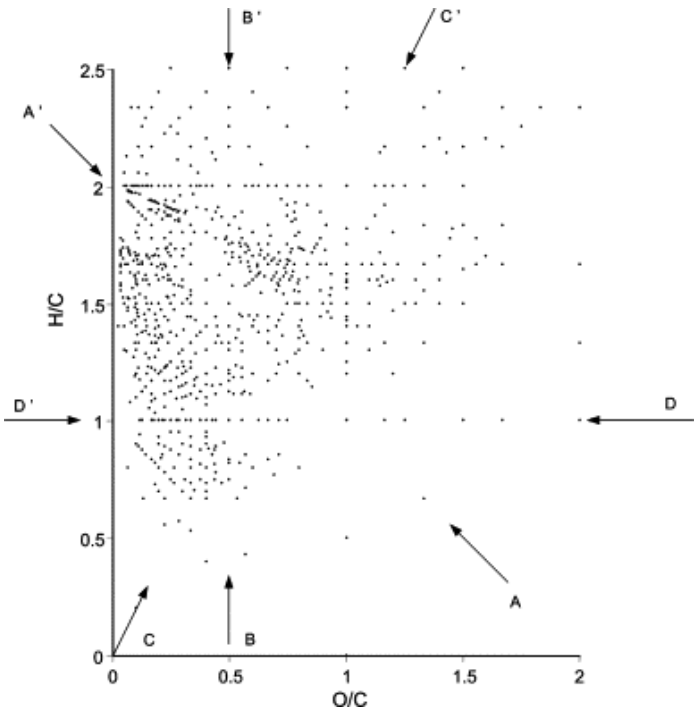


図 2-1-3-3 Krevelen Plot に現れる代謝反応経路

これらの線から目視により、様々な化学反応によって生じる生成物を推定できる。化学反応を表す線を方程式で定義した場合、各反応による特徴が傾斜と切片によって表せるとして表 2-1-3-2 の通り示されている[4]。

方程式： $H/C = -a(O/C) + b$

表 2-1-3-2 化学反応を示す方程式

図	化学反応	線の特徴
A-A'	メチル化・脱メチル化(CH3) アルキル鎖伸長	b = 2
B-B'	水素化・脱水素化(H2)	垂直
C-C'	水和・脱水(H2O)	a = 2
D-D'	酸化・還元(O)	b = 0
	脱炭酸 (CO2)	a = 0; b = 2

2.2 目的

メタボロミクスにおいて KrevelenPlot は質量分析により得られたデータの整理に用いられている。これを既知の代謝物質に適用し、代謝反応を含めて整理することは未知の代謝物質の解析にも役立つと考えられる。KrevelenPlot で整理するとその代謝物質の種類や構造的特徴、代謝反応の種類を同時に表現できるため、代謝物質の構造的変化や反応の規則性といった代謝経路の全体像を把握できる。本修士論文では KrevelenPlot を拡張して用いることで、既知の代謝物質と代謝反応、酵素反応の特徴による整理を行う。そして整理した結果に基づいて、代謝物質に応じた KrevelenPlot の拡張について考察する。

代謝物質の整理に KrevelenPlot を用いる利点は 3 つある。ほぼすべての代謝物質を 2 次元グラフ上に表現できること、その座標領域から代謝物質の種類を推定できること、化学組成式の類似した代謝物質間の反応の種類をベクトルとして視覚化できることである。一方現状の KrevelenPlot による解析の問題点としては、C、H、O の化学組成式が同じ場合は同じ座標に集約されてしまうこと、C、H、O 以外の元素については考慮されていないことの 2 つが挙げられる。代謝物質の悉皆的な解析に KrevelenPlot を用いるにはこれらの課題に対処してゆく必要がある。

まず、多様な特徴を捉えるために KrevelenPlot を拡張し、既知の代謝物質を整理する。拡張案としては 3 次元化や軸の定義、組合せを検証する。次に、様々な代謝物質の化学組成式データを用いて種類ごとの座標領域の分離を検証する。また、官能基や構造的特徴による座標領域の分離も検証する。さらに、代謝反応の規則性について整理する。代謝物質の骨格が同じで、化学組成の差異が同じであるペアを抽出し、2 点の座標の間に線分を引いて代謝反応を表現する。酵素反応についても同様に整理する。同じ反応をグラフ上に描いて規則性を整理することで、規則に合わない反応の識別が容易となる。

このように整理した結果を元に、既知の反応経路を用いて検証する。代謝物質の種類に応じた軸を用いて代謝物質と代謝反応を表現し、経路全体の全体像を示す。

最後に本研究で整理した内容を適用した未知の代謝物質の推定について述べる。

第3章 材料および方法

本章では、第 3.1 節に本研究に用いた化学物質のデータベースとそれら化学物質が生成するために必要な部品たる化学構造（部分構造）を述べる。第 3.2 節に第 3.1 節で挙げたデータを解析するための手法を述べる。

3.1 材料

3.1.1 化学物質(化学式)データ

本研究では代謝物質のデータとして、以下に述べる 4 種類のものを用いた。本研究室で開発中の生物種-代謝物質データベース KNApSACk(*1)、京都大学化学研究所バイオインフォマティクスセンター運営による生命システム情報統合データベース KEGG(*2)に含まれる drug、compound、reaction の 4 つである。それぞれの特徴と使用目的を表 3.1 に示す。なおアルキル基 (R) や複数分子（化学組成式が「・」を挟んで複数存在する）を含むデータは除外する。

(*1) KNApSACk (<http://kanaya.naist.jp/KNApSACk/>)

(*2) KEGG (http://www.genome.ad.jp/ja/gn_kegg_ja.html)

表 3-1-1-1 化学物質データベース

	データベース名	件数	特徴
1	KNApSACk	24,820	植物の二次代謝物質
2	KEGG Drug	3,725	日米の許認可医薬品
3	KEGG Reaction	4,150	生合成、代謝反応ペアとして解明された化学物質
4	KEGG Compound	11,507	生体由来の化学物質
	計	44,202	

3.1.2 酵素反応分類

酵素は国際生化学分子生物学連合により酵素命名法(*1)が決められており、EC に続く 4 桁の数字から成る EC 番号が与えられて整理されている。EC 番号は『反応形式. 供与体. 受容体. 酵素名』で構成されている。反応形式の分類を表 3-1-2-1 に示す。

(*1)酵素命名法 (<http://www.chem.qmul.ac.uk/iubmb/enzyme/>)

表 3-1-2-1 酵素番号体系

EC				Name
1	*	*	*	酸化酵素

3.1.3 代謝物質の生合成経路

既に解明されている代謝物質の生合成経路を KrefelenPlot で表現するために、代謝物質とその反応経路を KEGG (3. 1. 1 項*1) の代謝パスウェイより取得する。検証した経路を表 3-1-3-1 に示す。

表 3-1-3-1 代謝物質の反応経路

ID	MapID	経路の種類
1. 10		二次代謝物質の生合成
	00941	フラボノイドの生合成

3.2 方法

3.2.1 KrevelenPlot の拡張による代謝物質の解析

表 3-1-1-1 で示したデータはすべて化学式が既知である。まず、化学組成式に含まれる元素を抽出し、全体における存在比率を求める。その結果により KrevelenPlot の持つ C、H、O の組成比が同じデータは同じ座標に現れるという問題の改良を試みる。まず、文献上で代謝物質に多く見られる元素について、既知の代謝物質データの化学組成式に含まれる割合を調査する。その結果によりほとんどの化合物に含まれる C、H、O 以外の元素を用いて軸を定義し、Z 軸として加えた 3 次元化による座標の分離の程度を検証する。また、官能基やヘテロ環など代謝物質の特徴を捉えるための軸を定義し、軸の組合せと共に検証する。

(1) 3 次元化と用いる軸

C、H、O が同じ組成比のデータは同じ座標に表現される。他の元素を用いた Z 軸を加えて 3 次元化することでこの重なりを軽減できるか検証する (図 3-2-1-1)。まず用いるデータの化学組成式に含まれる元素を調べ、C、H、O 以外の多く含まれる元素/C 元素として座標の分離を検証する。

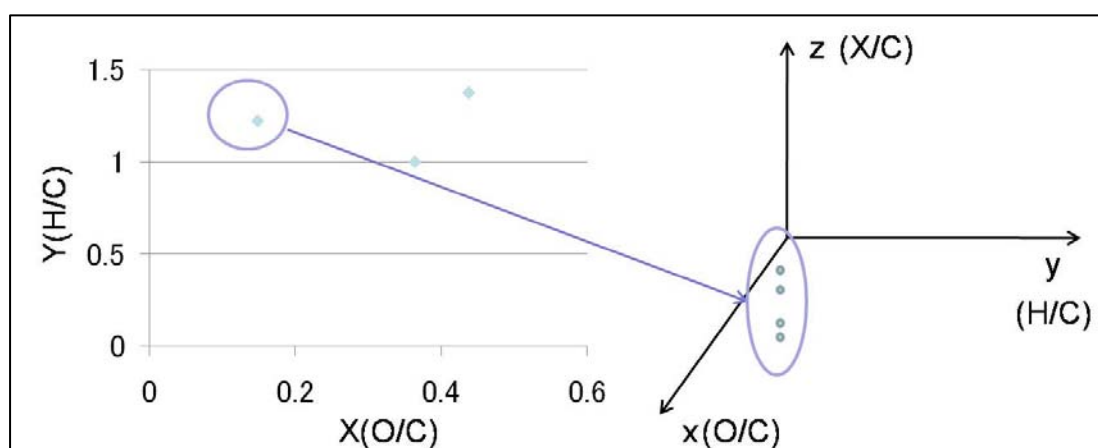


図 3-2-1-1 KrevelenPlot の 3 次元化

(2) 代謝物質の特徴を捉える軸

代謝物質の多様な特徴をとらえるために、他の元素を用いた軸を設定することで、官能基の結合の程度が把握できると考える。従来の KrevelenPlot の軸の定義とそれぞれにより捉えられる特徴を表 3-2-1-1 に示す。

表 3-2-1-1 KrevelenPlot の軸の定義

軸	定義	意図	表す特徴
X	O/C	C に結合している O の比率	O を含む構造の多少
Y	H/C	C に結合している H の比率	構造の推測（鎖状・飽和結合／環状・不飽和結合）

これに加え、多く含まれる元素や官能基の種類に応じた軸を定義し検討する。炭素 (C) 以外を含む官能基として考えられるものを図 3-2-1-2 に示す。

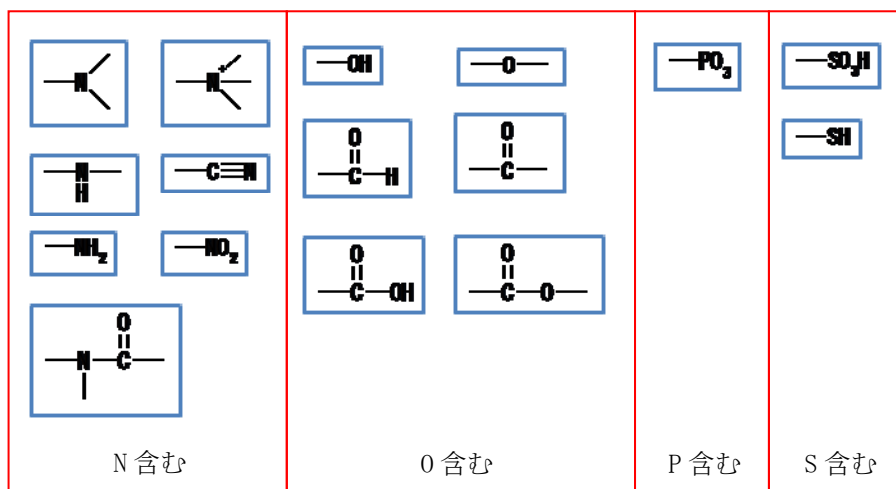


図 3-2-1-2 官能基の種類

拡張案としての軸の定義を表 3-2-1-2 に、軸の組合せを表 3-2-1-3 に示す。

表 3-2-1-2 拡張 KrevelenPlot の軸の定義

定義	意図	表す特徴
N/C	C に結合している N の比率	N を含む構造の多少
S/C	C に結合している S の比率	S を含む構造の多少
P/C	C に結合している P の比率	P を含む構造の多少
N/O	N に結合している O の比率	N と O の結合の関係（ヘテロ環／官能基）
S/O	S に結合している O の比率	S と O の結合の関係（ヘテロ環／官能基）
P/O	P に結合している O の比率	P と O の結合の関係（ヘテロ環／官能基）
O/H	O に結合している H の比率	O と H の結合の関係（種類、構造・骨格の特定）
N/H	N に結合している H の比率	N と H の結合の関係（種類、構造・骨格の特定）
S/H	S に結合している H の比率	S と H の結合の関係（種類、構造・骨格の特定）

表 3-2-1-3 拡張 KrevelenPlot の軸の組合せ

	X	Y	Z	表す特徴
1	O/C	H/C	N/C	N を含む構造の母骨格の特定
2	O/C	H/C	S/C	S を含む構造の母骨格の特定
3	O/C	H/C	P/C	P を含む構造の母骨格の特定
4	N/H	H/C	N/C	N を含む構造に結合した官能基の推定
5	S/H	H/C	S/C	S を含む構造に結合した官能基の推定
6	O/H	H/C	O/C	O を含む構造に結合した官能基の推定
7	O/C	N/O	N/C	N と O の結合の関係（ヘテロ環／官能基）と種類の特定
8	O/C	S/O	S/C	S と O の結合の関係（ヘテロ環／官能基）と種類の特定
9	O/C	P/O	P/C	P と O の結合の関係（ヘテロ環／官能基）と種類の特定

3.2.2 KrevelenPlot による代謝物質の特徴整理

KrevelenPlot 上に代謝物質をプロットすると種類によって座標領域が異なることが先行研究で示されている。そこでまず従来版の KrevelenPlot に既知の代謝物質を種類ごとにプロットすることで特徴的に現れる座標領域について整理する。また 3.2.1 項で検証した軸の組合せを用いた場合の代謝物質の種類ごとの特徴的な座標領域についても整理する。

(1) 母核の構造的特徴による座標の分離

KrevelenPlot は X 軸が O/C、Y 軸が H/C であるから、座標面上のどの領域にプロットされるかでその化学物質の特徴や母核の化学構造をある程度推測できると考えられる。元素の価電子数は元素毎に決まっており、C 元素は最大 4 つまで原子と結合できる。ほぼ全ての代謝物質は C 元素が 2 つ以上連なって構造の基本骨格を形成する。C 元素が 2 つ以上連なる場合、最大 3 つの原子と、 σ 結合と π 結合が 1 つずつの sp^2 混成軌道を生じる二重結合がある場合は 2 つの原子、 σ 結合 1 つと π 結合 2 つの sp 混成軌道を生じる三重結合では 1 つの原子と結合できる。つまり、C に対する H の結合数を表す Y 座標値からは基本骨格にどのような結合が多いかを推測できる。H の結合数が相対的に多い (y 値が大である) 代謝物質は、多くの C 元素が飽和結合 (一重結合) で連なった鎖状の構造と考えられる。また、H の結合数が相対的に少ない (y 値が小である) 代謝物質は、環構造もしくは不飽和結合 (二重結合、三重結合) が多いと考えられる (図 3-2-2-1)。

C 元素の価電子数と代謝物質が C 元素を 2 つ以上持つことを考慮すると、従来の KrevelenPlot の X 軸、Y 軸ともに 0.0~3.0 の範囲にほとんど全ての代謝物質が表現できると考えられる。この範囲内で代謝物質の種類によって座標空間が分類出来れば、座標値によっておよその構造が把握できる。まず従来版の KrevelenPlot の X 軸 (O/C)、Y 軸 (H/C) において、O 元素と H 元素の多少により 4 通りの座標領域に分け、それぞれの特徴を整理した。次にこの特徴に合致する代謝物質の種類を調査する。表 3-2-2-1 に

座標領域の特徴と該当する代謝物質の種類を整理して示す。また図 3-2-2-2 に座標領域の分離を示す。

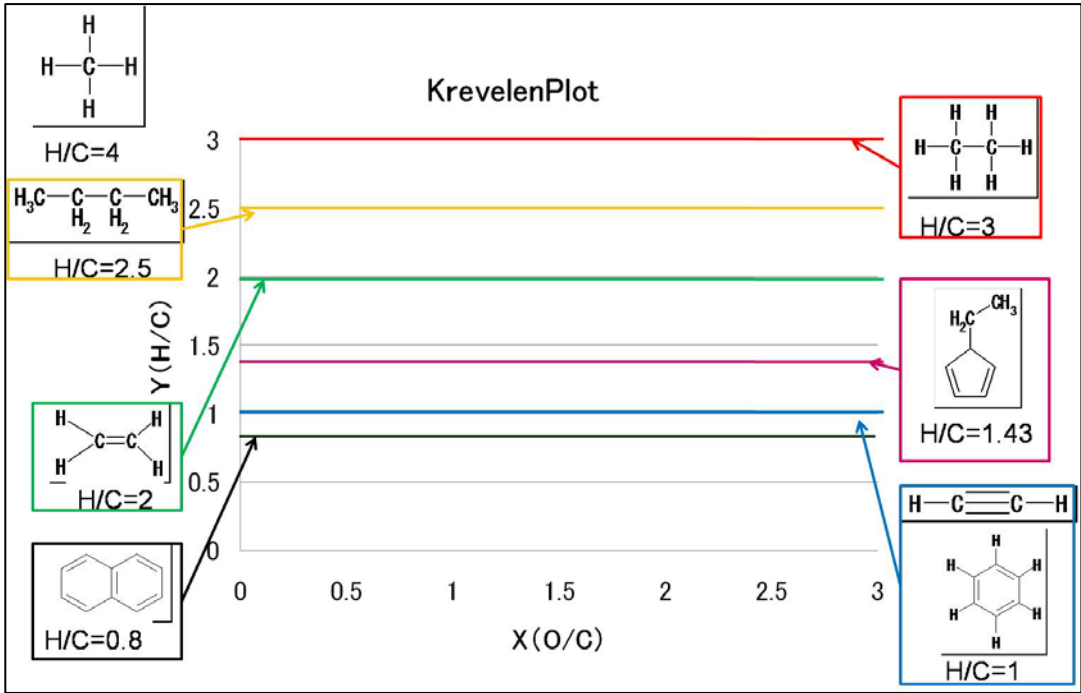


図 3-2-2-1 Y 座標値と構造の相関

表 3-2-2-1 代謝物質の特徴による座標領域の分離（従来版 KrevelenPlot）

	組成の特徴	化学構造の特徴	代謝物質の種類
(a)	O 少、H 多	飽和結合（一重結合）が多い 置換基の結合が少ない	脂質 ステロイド
(b)	O 多、H 多	飽和結合（一重結合）が多い 置換基の結合が多い	糖質
(c)	O 少、H 少	不飽和結合（二重・三重結合）が多い 置換基の結合が少ない	フラボノイド キノン
(d)	O 多、H 少	不飽和結合（二重・三重結合）が多い 置換基の結合が多い	なし

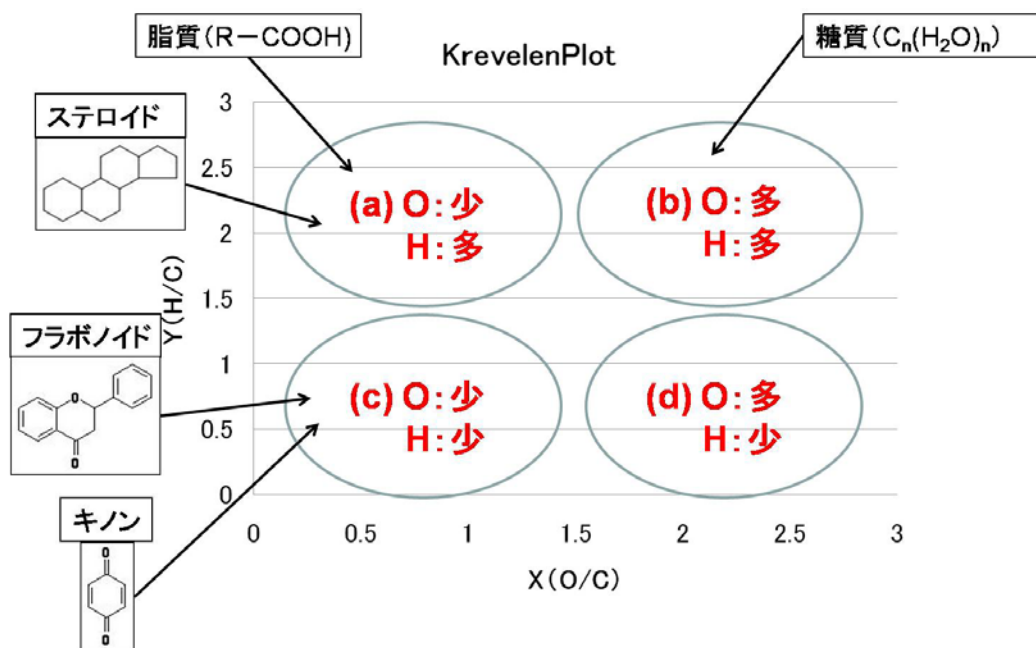


図 3-2-2-2 座標領域と代謝物質（従来版 KrevelenPlot）

代謝物質の化学組成における特徴を以下に示す[7]。

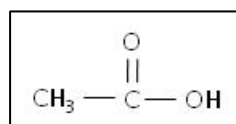
1. 脂質（脂肪酸）

炭素が直鎖状に結合した構造を持ち、カルボン酸（-COOH）が末端に結合している。飽和脂肪酸、不飽和脂肪酸、特殊脂肪酸の3つに大別される。飽和脂肪酸、不飽和脂肪酸が動植物に多く含まれるのに対し、特殊脂肪酸は主に微生物、特に細菌に含まれている。

1) 飽和脂肪酸

炭素-炭素の二重結合がないもの。偶数炭素同士の縮合反応により生成されるため、炭素数は16から18からなるものが多い。

例)



酢酸

2) 不飽和脂肪酸

炭素-炭素の二重結合が1つ以上のもの。二重結合が1つであれば、1価不飽和脂肪酸、二重結合が2つ以上であれば多価不飽和脂肪酸とよぶ。それ

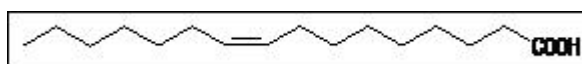
それぞれの結合には以下に示す特徴がみられる。

1 価不飽和脂肪酸 $-\text{CH}=\text{CH}-\text{CH}=\text{CH}-$ (共役 2 重結合)

多価不飽和脂肪酸 $-\text{CH}=\text{CH}-\text{CH}_2-\text{CH}=\text{CH}-$ (非共役 2 重結合)

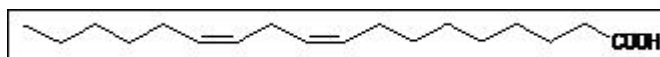
例)

1 価不飽和脂肪酸



パルミトレイン酸

多価不飽和脂肪酸

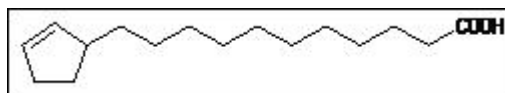


リノール酸

3) 特殊脂肪酸

枝分かれやシクロプロパン酸などを主体とするものが多い。

例)

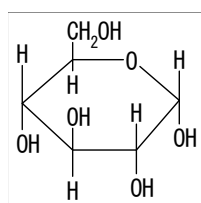


ヒドノカルピン酸

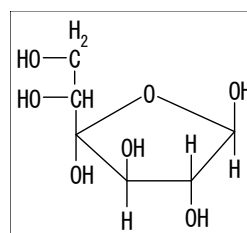
2. 糖質

基本構造として $\text{C}_n(\text{H}_2\text{O})_n$ をもつ一種のポリヒドロキシアルデヒドまたはポリヒドロキシケトンである。アルデヒド基をもつものをアルドース、ケトン基を持つものをケトースと呼び、通常環構造を成す。5 員環構造をフラノース、6 員環構造をピラノースとして分類される。

例)



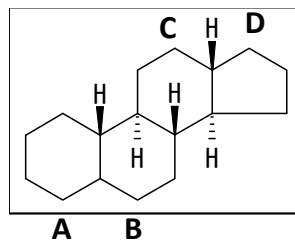
α-D-グルコピラノース



β-D-グルコピラノース

3. ステロイド類

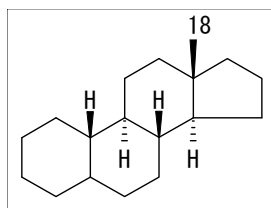
基本構造としてペルヒドロシクロペンタフェナントレン骨格を持ち、天然ステロイドの骨格炭素数は 18 から 29 個 ($\text{C}_{18} \sim \text{C}_{29}$) である。A、B、C 環はシクロヘキサン環が融合したものである。



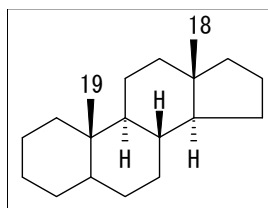
ペルヒドロシクロペンタノパーフェナントレン骨格

動植物界に広く存在し、重要な生理活性を有する。炭素骨格からの基本構造は8つに分類される。 C_{18} : エストラン estrane、 C_{19} : アンドロスタン androstane、 C_{21} : プレグナン pregnane、 C_{23} : カルダン cardane、 C_{24} : コラン cholane、 C_{27} : コレスタン cholestane、 C_{28} : エルゴスタン ergostane、 C_{29} : スティグマスタン stigmastane である。例にそれらの構造を示す。

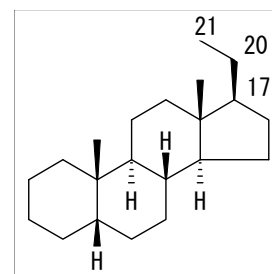
例)



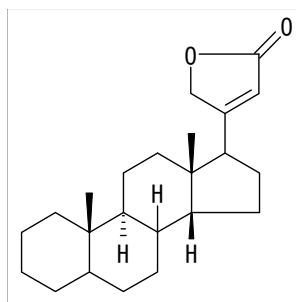
エストラン (C_{18})



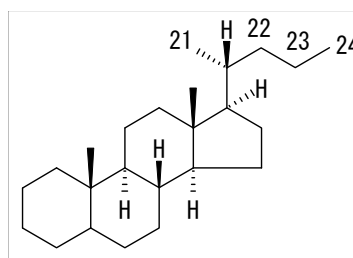
アンドロスタン (C_{19})



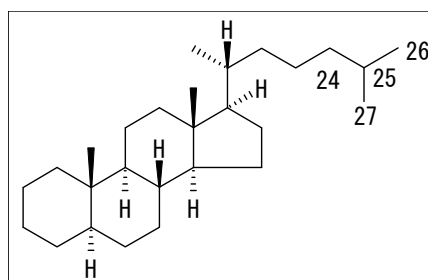
プレグナン (C_{21})



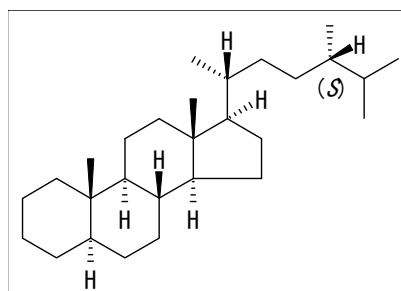
カルダン (C_{23})



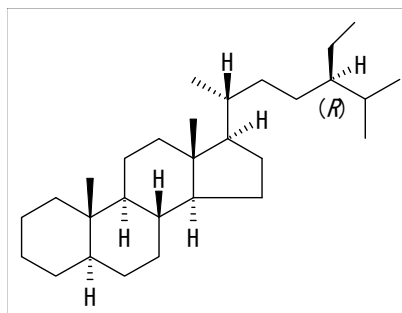
コラン (C_{24})



コレスタン (C_{27})



エルゴスタン (C_{28})



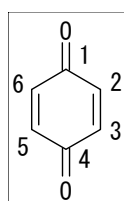
スティグマスタン(C29)

4. キノン類

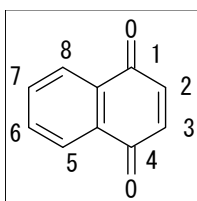
天然色素に分類される動植物・微生物由来の色素成分で非常に大きなグループとして扱われる。メチレン基とカルボニル基とが順につながった形を持つことからポリケチド polyketide と総称され、多くは環状ポリケチドとして存在する。化学構造によって5つに分類される。構造は例に示す。

- 1) ベンゾキノン benzoquinone、単環性
- 2) ナフトキノン naphthoquinone、双環性
- 3) アントラキノン anthraquinone、3 環性
- 4) 延長キノン extended quinine、多環性
- 5) フェナントラキノン phenanthraquinone、3 環性

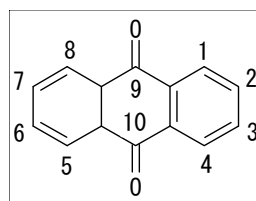
例)



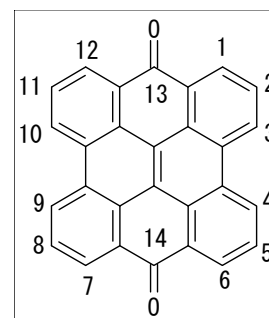
(1)



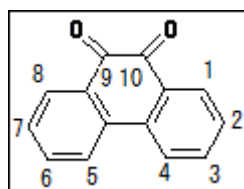
(2)



(3)



(4)

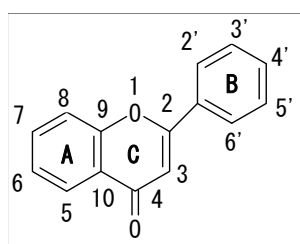


(5)

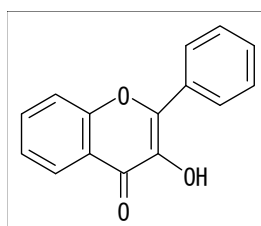
5. フラボノイド類

基本構造としてフェニルクロマン ($C_6-C_3-C_6$) 骨格を持つ芳香族化合物で、維管束植物界に広く分布し、植物色素の重要な一群を形成している。抗アレルギー作用、抗炎症作用、抗腫瘍作用など多彩な生理作用が見出されており、薬用成分として重要なものも多く存在する。また、植物病菌に対する抵抗因子として見出されたものもある。基本構造 (A、B、C 環) 中の C 環の置換様式の違いによって、10 種類に分類される。フラボン flavone、フラボノール flavonol、フラバノン flavanone、カルコン chalcone、ジヒドロフラボノール dihydroflavonol、イソフラボン isoflavone、オーロン aurone、ロイコアントシアニジン leucoanthocyanidin (flavan-3,4-diol)、フラバン-3-オール flavan-3-ol、アントシアニジン anthocyanidin である。構造を例に示す。

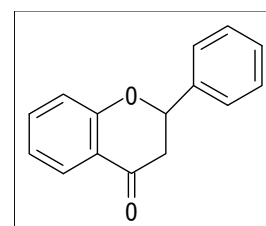
例)



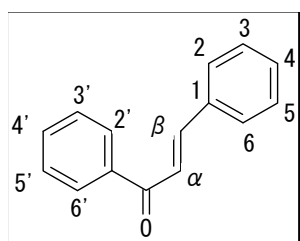
フラボン



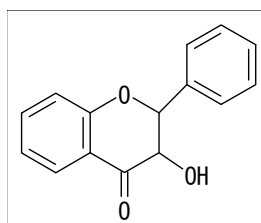
フラボノール



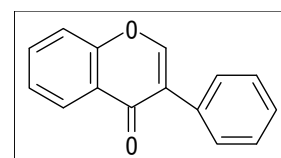
フラバノン



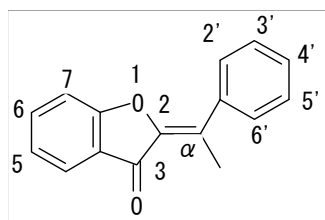
カルコン



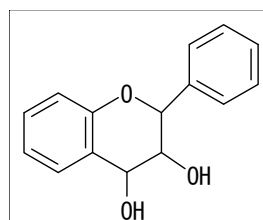
ジヒドロフラボノール



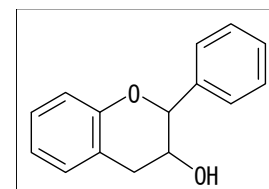
イソフラボン



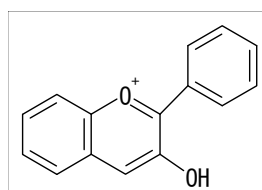
オーロン



ロイコアントシアニジン



フラバン-3-オール



アントシアニジン

(2) 官能基による座標の分離

KrevelenPlot を拡張することで C、H、O 以外の元素についての特徴も把握できる。そこで、ヘテロ元素 (N、P、S) の結合による座標値の変化を見るために、従来の KrevelenPlot 上にヘテロ元素を含む代謝物質と含まない代謝物質を凡例色で分けて表現することで、座標値が分離するなどの特徴がみられるかを検証する。元素を含むか否かにより座標領域に特徴があれば、その元素が結合することで座標軸に定義した 0 元素あるいは H 元素の増減を伴うことを示す。これにより結合している官能基の種類の特定が容易になると考える。

また、表 3-2-1-3 で定義した拡張 KrevelenPlot を用いてヘテロ元素ごとに座標領域が分離するか確認した。Y 軸方向に特徴だった分離が見られればその元素が結合しやすい構造がある程度限定できる。

さらに拡張 KrevelenPlot を用いてデータごとに凡例色を分けて表現することで、薬物や反応経路の解明済みのデータと未解明のデータの全体像を比較する。これにより薬物として利用の進んでいる構造や代謝経路の解明が進んでいる代謝物質の判別と、各方面でこれから研究する余地のある構造や種類について示す。

3.2.3 KrevelenPlot による代謝反応の視覚化

先行研究では代謝物質を KrevelenPlot で表現すると代謝物質により線が視認できるとしており、同じ反応によって変化した代謝物質のペアは特定の線上に現れると述べられている。拡張した KrevelenPlot を用いるなどして様々な代謝反応を整理する。

(1) 代謝反応の整理

KrevelenPlot 上に代謝反応を描くことで、反応経路がどのように表現されるかを整理した。まず、既知の化学物質のうち母骨格が同じで化学組成式が変化しているものを代謝反応によって変化したペアとみなし、代謝反応が同じペアをグループ化した。グループ化した代謝反応ごとに見られる特徴を確認する（図 3-2-3-1）。また、2.2.2 項で整理した代謝物質の種類や構造的特徴と照らして代謝反応について整理する。

次いで酵素反応を描いてどのように表現されるかを整理する。酵素は反応の種類によって体系づけられて整理された EC 番号を持っている。EC 番号が同じグループであれば同じ反応を触媒するため、KrevelenPlot で酵素反応を表すと視覚的に捉えられる。

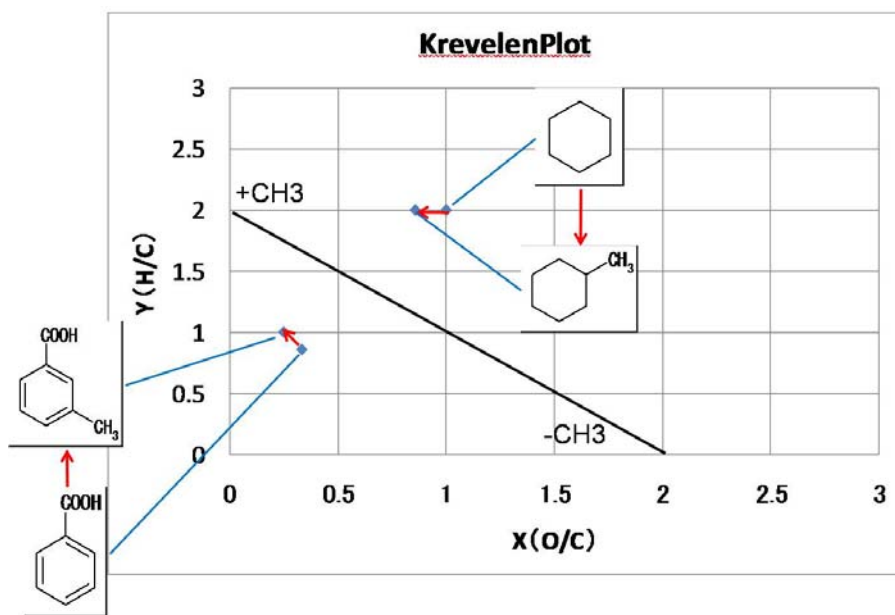


図 3-2-3-1 代謝反応の整理

以下に挙げた代謝反応の特徴について整理する。なお、その反応経路を整理するために用いる軸も反応の種類に応じて設定する。軸の定義と反応の種類を表 3-2-3-1 に示す。

表 3-2-3-1 代謝反応の種類

	拡張	X	Y	Z	代謝反応
1	Z(O/H)	O/C	H/C	O/H	CH ₃
2	Z(O/H)	O/C	H/C	O/H	OH
3	Z(O/H)	O/C	H/C	O/H	C ₂ H ₂ O
4	Z(O/H)	O/C	H/C	O/H	C ₃ H ₂ O ₃
5	Z(O/H)	O/C	H/C	O/H	C ₆ H ₁₀ O ₃
6	Z(O/H)	O/C	H/C	O/H	C ₆ H ₁₀ O ₄
7	Z(O/H)	O/C	H/C	O/H	C ₇ H ₄ O
8	Z(O/H)	O/C	H/C	O/H	H ₂
9	Z(N/H)	O/C	H/C	N/H	C ₉ H ₉ O _N
10	Z(N/H)	O/C	H/C	N/H	CH ₃ O _N
11	Z(N/H)	N/C	H/C	N/H	NH
12	Z(N/O)	O/C	H/C	N/O	C ₉ H ₉ O _N
13	Z(N/O)	O/C	H/C	N/O	CH ₃ O _N
14	Z(N/O)	N/C	H/C	N/O	NH
15	Z(P/O)	O/C	H/C	P/O	H ₂ O ₆ P ₂
16	Z(P/O)	O/C	H/C	P/O	H ₃ O ₄ P
17	Z(P/O)	O/C	H/C	P/O	H ₃ O ₃ P
18	Z(P/O)	O/C	H/C	P/O	H ₄ O ₄ P
19	Z(S/H)	O/C	H/C	S/H	O ₄ S
20	Z(S/O)	O/C	H/C	S/O	O ₄ S

(2) 酵素反応の整理

EC 番号の 1 桁目は反応の種類を指す。このため EC 番号の 1 桁目が同じ酵素反応を KnevelenPlot 上に描けば同じ特徴を示すことが予想される。そこで酵素反応の特徴も整理する。

3.2.4 既知の代謝反応経路による検証

整理した内容を既知の代謝物質およびその生合成経路に適用することで検証した。KrevelenPlot で経路の全体像を俯瞰しながら代謝物質の特徴を捉えられるかを示す。

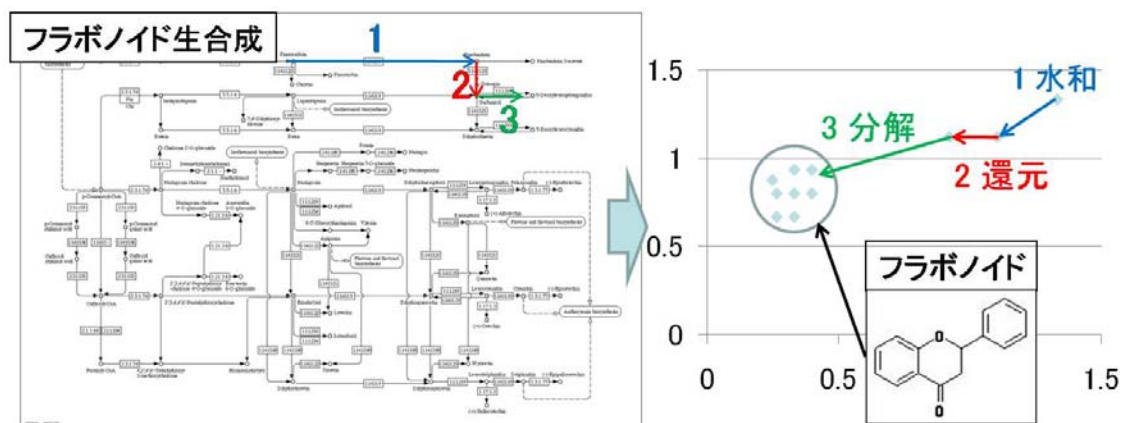


図 3-2-4-1 生合成経路の KrevelenPlot での表現

第4章 結果と考察

4.1 KrevelenPlot による代謝物質の表現

代謝物質の化学式から座標を求めて KrevelenPlot を描画し、代謝反応をベクトルで表現するプログラムを Java によって開発した。なお、理論化学式の上に、代謝物質をプロットすることで代謝物質の組成数の特徴が把握できるようにした。使用したプログラムについて以下に記す。

2D KrevelenPlot : Java (1.6.0) + JFreeChart (1.0.11) ^(*1)

3D Krevelenplot : Graph-R (2.60) ^(*2)

(*1) グラフ描画 Java フレームワーク (*2) 伊藤徹氏著作物、フリーソフト

図 4-1-1 に 4 種類のデータをプロットした例を示す。この時の座標範囲を表 4-1-1 に示す。

$$(x, y) = \{ (O \text{ の組成数} / C \text{ の組成数}), (H \text{ の組成数} / C \text{ の組成数}) \}$$

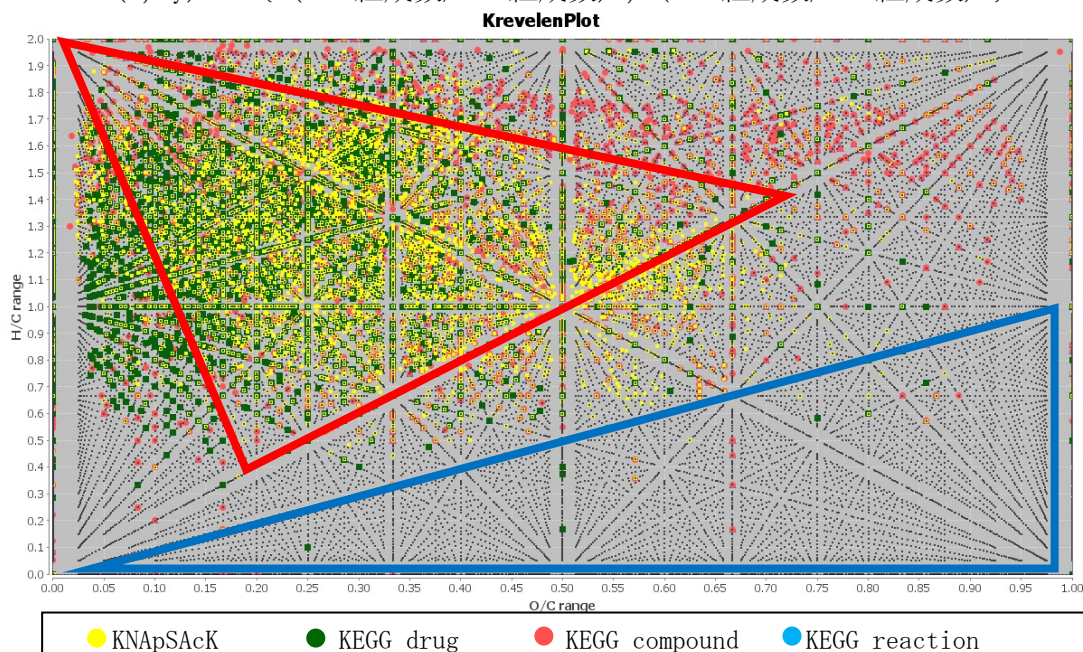


図 4-1-1 KrevelenPlot による化合物の表現

図中に青枠で示した個所は代謝物質がほとんど見られない。青枠の範囲は 0 元素が H

元素の 1 倍以上結合する組成比となる。このような組成比を持つ化学物質はあまり見られず、既知の代謝物質の特徴には一致するものがない。逆に赤枠で示した個所には KNapSAcK など多くの代謝物質が見られる。図 2-1-1-1 で示したように、代謝物質はいくつかの出発物質によって作られている。ステロイドであればメバロン酸由来、キノンはアセチル CoA 由来、フラボノイドはシキミ酸由来、アルカロイドはアミノ酸由来である。出発物質の種類が限られているため座標の重複もよく起きると考えられる。表 4-1-1 に座標範囲を示す。

表 4-1-1 各データの座標範囲

		X (O/C)		Y (H/C)	
		min	max	min	max
1	KNapSAcK	0.00	3.33	0.00	3.50
2	KEGG drug	0.00	31.99	0.00	4.00
3	KEGG reaction	0.00	4.50	0.00	4.00
4	KEGG compound	0.00	4.50	0.00	4.00

4.1.1 データの特徴

4 種類の化学物質のデータベースより取得した化学式から、その化学式に含まれる元素とその元素を持つ化学物質の数を集計した。結果を表 4-1-1-1 に示す。

表 4-1-1-1 代謝物質の元素別の集計

		KNA p SA c K	KEGG drug	KEGG compound	KEGG reaction
a	すべて [件]	24,820	3,725	11,507	4,150
b	C を含む [件]	24,807	3,643	11,355	4,083
c	b/a [%]	99.95	97.80	98.68	98.37

	元素	KrevelenPlot で解析可能なデータ [%] (*)			
		KNA p SA c K	KEGG drug	KEGG compound	KEGG reaction
1	C	100	100	100	100
2	H	100	100	100	100
3	O	97.48	92.29	93.91	96.55
4	N	11.55	57.10	29.80	29.22
5	S	2.94	24.27	10.52	10.80
6	Cl	0.05	0.01	1.88	0.12
7	P	1.10	2.08	9.49	1.88
8	Br	0.02	0.00	0.03	0.00
9	I	0.00	1.73	0.17	0.22

(*) 当該元素を含むデータ件数／b×100

c 欄の通り、C 元素は各データにおいて 98% 近く存在するため C 元素を軸の分母とする KrevelenPlot でほぼすべてのデータを表現できる。また C 元素に対する各元素の存在比をみると、H 元素、O 元素は 92% 以上の化合物に存在する。よって KrevelenPlot で表現出来るデータの 92% 以上が原点以外の座標に現れることが確認できた。その他の元素では KEGG drug の N 元素、I 元素の存在比が際立って高いが、これは生理活性や体内動態を制御するための人為的な化学修飾によるものと考えられる。KEGG compound で

P 元素の存在比が高いのは核酸、ATP などの生命活動に必須の 1 次代謝物質も含まれているためと考えられる。

X 座標値、Y 座標値をヒストグラムで表した。横軸はそれぞれの座標での値、縦軸はその座標に存在するデータ件数とした。図 4-1-1-1、図 4-1-1-2 に示す。

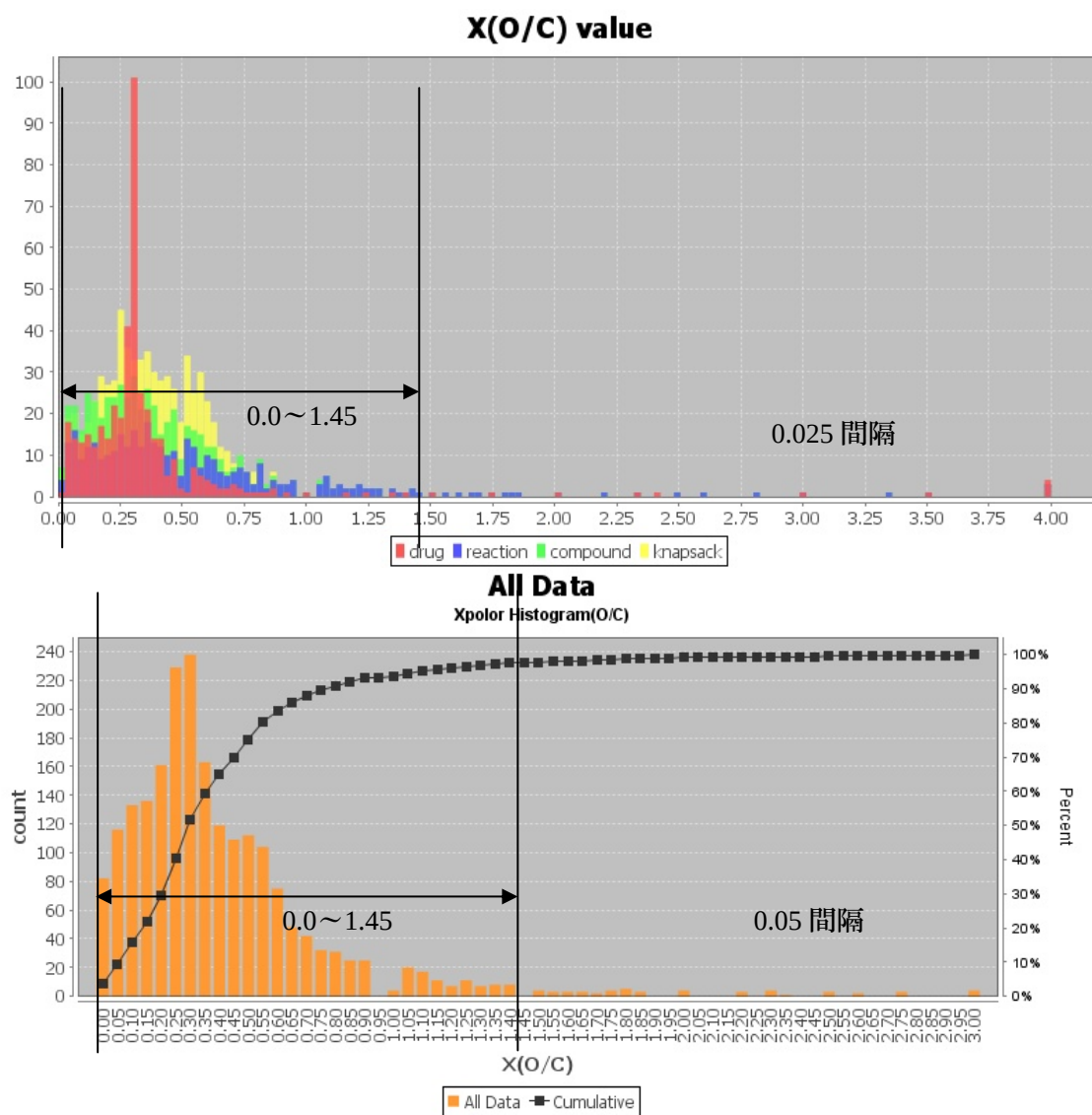


図 4-1-1-1 X(O/C)軸の分布

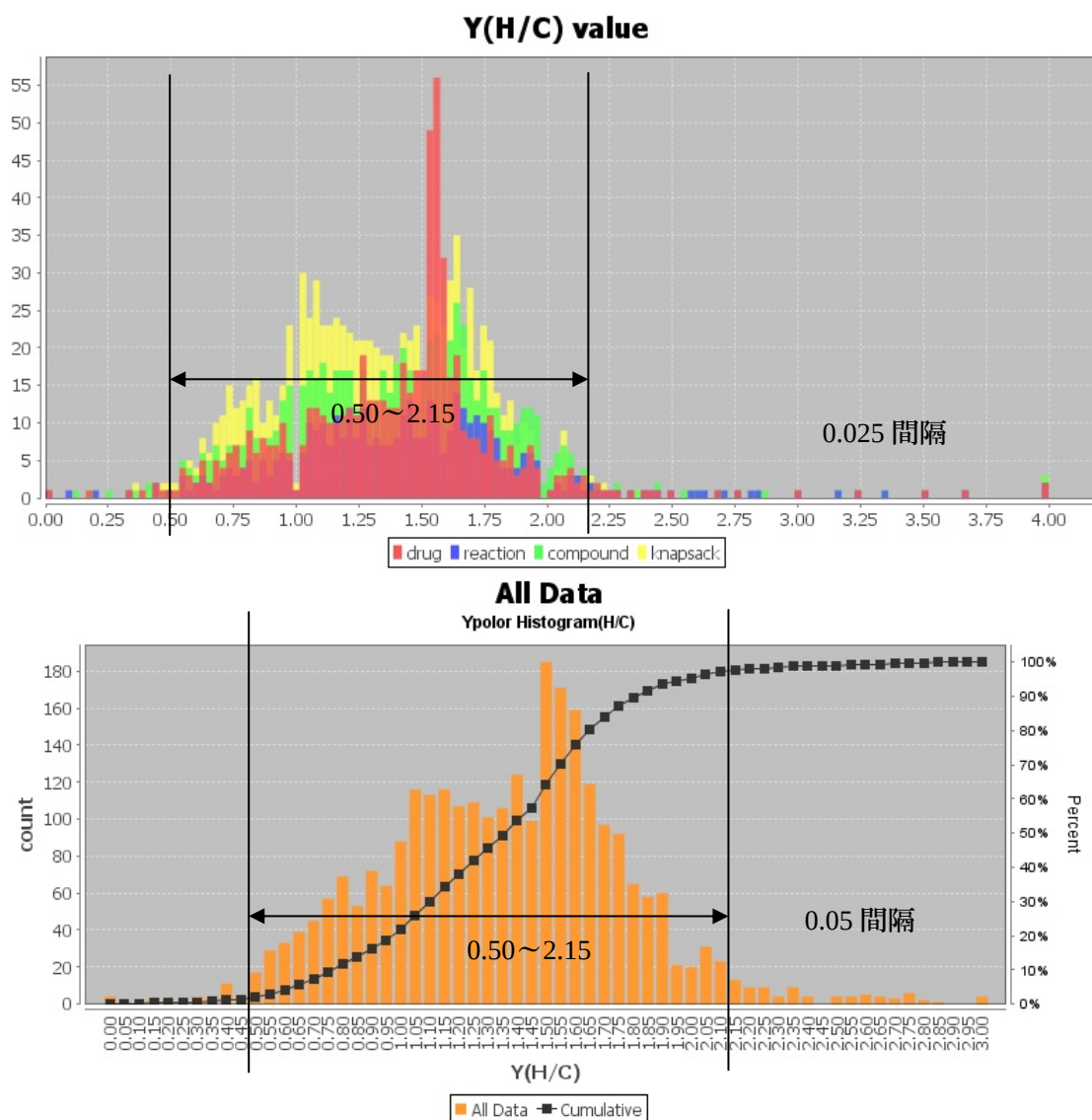


図 4-1-1-2 Y(H/C)軸の分布

データ件数が多い座標は重複が多いと考えられ、いずれのデータもデータ件数から見た重複の多い座標は同じ傾向を示して偏っている。X 座標値は 0.0~1.45 の範囲に、Y 座標値は 0.50~2.15 の範囲に集中していることから代謝物質が取りうる構造的制約による化学組成のパターンが見えてくる。たとえば、Y=1.0、2.0 となる構造のデータ件数が目立って少ない。Y=1.0 は H:C=1:1、Y=2.0 は H:C=2:1 を表す。図 3-2-2-1 でも示した通り、Y 値によって骨格が判別できる。Y=1.0 や 2.0 はいわば基準値であり簡単な構造の場合に現れる。C 元素が多く結合している代謝物質では飽和結合や不飽和

結合の割合により Y 値が基準値の間を埋める形で存在すると考えられる。O 値に関しては C 元素に O 元素が結合するには官能基としてか、ヘテロ環として環構造に O 元素が含まれる場合になる。X 座標値から O 元素は C 元素の 1.45 倍程度までとなる化学組成が多くを占めることがわかる。C 元素の価電子の数からいえば Y 軸値同様、 $O : C = 3 : 1$ ($X : O/C = 3.0$) という値も取りうるはずであるが、実際には $O : C = 1 : 2$ ($X : O/C = 0.5$) 以下までのデータの偏りが大きい。このことから O を含む官能基やヘテロ環の結合のパターンは少ないと言える。

4.1.2 KrevelenPlot の拡張 一次元

従来の KrevelenPlot で表現した場合の座標の偏りと重複の程度を調査した。構造的に取りうる座標範囲を 0.0~3.0 とし 0.1 区間ごとに集計し、xy 平面のヒートマップを作成した (図 4-1-2-1)。指標統一のためデータ存在比で座標の重複を表す。

密度：(その座標区間に存在するデータ件数／全データ件数) × 100

(凡例：密度、0.00 ~1.9 を 20 分割。1.9 以上も 1.9 とした)

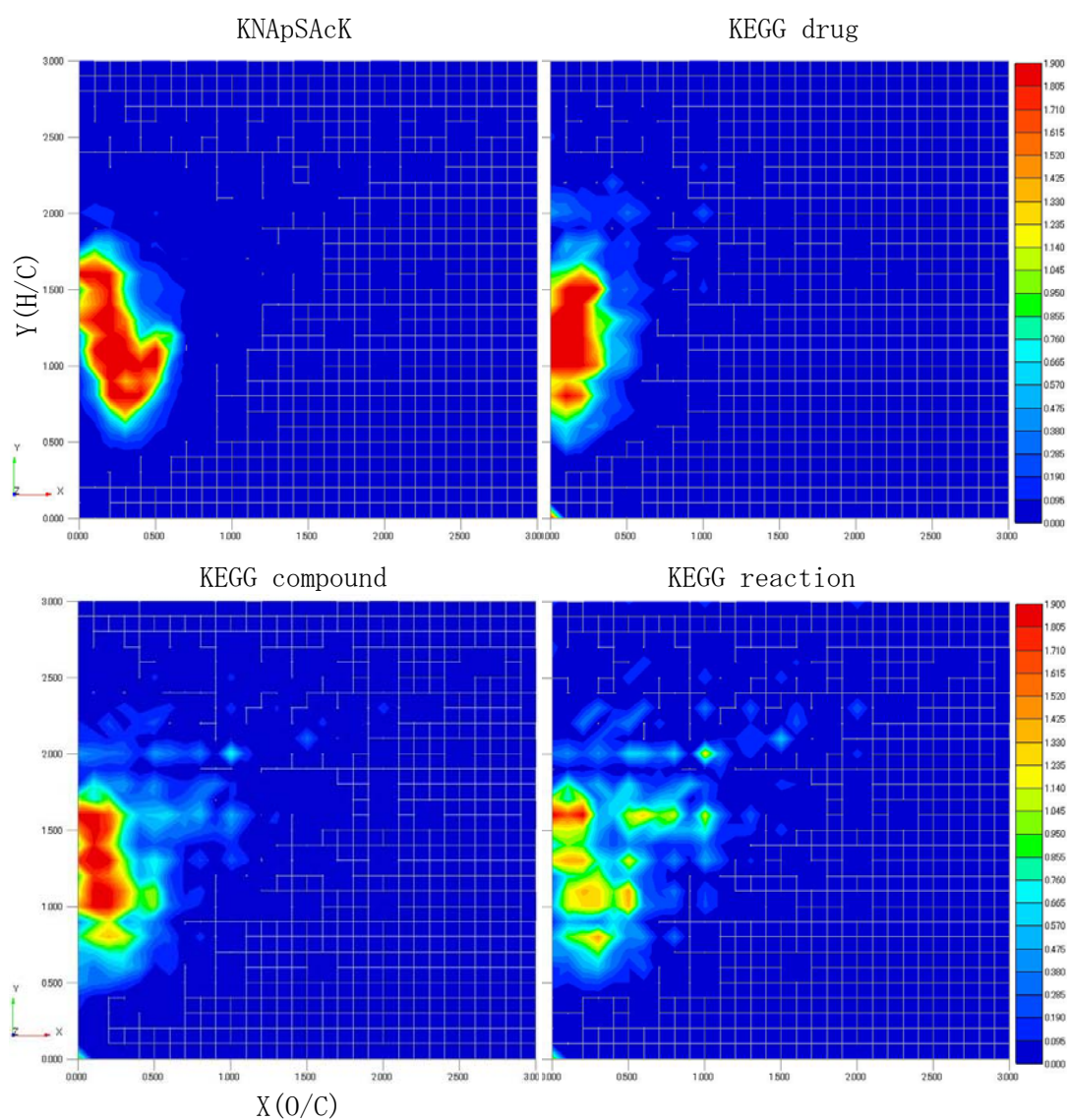


図 4-1-2-1 代謝物質の分布と出現頻度 (X : O/C、Y : H/C)

Y 軸 (H/C) 方向ではどのデータも同程度に広く分布しているが、X 軸 (O/C) 方向では KEGG drug のみ分布が小さい。また、すべてのデータで共通して X=0.0~9.5、Y=0.8~1.7 の範囲で座標の重複が多い。

3次元化することで座標の重なりをどの程度軽減できるか調査した。調査対象として存在比が高く重要な意味を持つ N、S、P を用いた。表 4-1-2-1 にユニークな座標を集計した件数と Z 軸を加えた時のユニークな座標の件数を示した。また、件数の増加率を示す。また元素の存在比を示した表 4-1-1-1 から一部抜粋し併せて示す。

増加率：(3D のユニークな件数－2D のユニークな件数)/3D のユニークな件数)*100

表 4-1-2-1 座標値のユニークパターン数

	Z	KNApSAcK		KEGG drug		KEGG compound		KEGG reaction	
		件数	増加%	件数	増加%	件数	増加%	件数	増加%
1	-	5,230	-	1,971	-	3,773	-	1,667	-
2	N/C	6,203	18.60	2,766	40.33	5,072	34.43	1,986	19.14
3	S/C	5,488	4.93	2,349	19.18	4,341	15.05	1,785	7.08
4	P/C	5,281	0.98	2,003	1.62	3,928	4.11	1,738	4.26

表 4-1-1-1 から一部抜粋

	元素	KrevelenPlot で解析可能なデータ [%] (*)			
		KNA p SA c K	KEGG drug	KEGG compound	KEGG reaction
1	C	100	100	100	100
2	N	11.55	57.10	29.80	29.22
3	S	2.94	24.27	10.52	10.80
4	P	1.10	2.08	9.49	1.88

表 4-1-2-1 から最もよく座標を分離するのは N/C 軸を用いた時と判明した。KNApSAcK では N と S の存在比、増加率とも約 4 倍である（青字箇所を参照）。元素の存在比の影響と考えられるが、データ間で比較すると存在比に応じて増加率が高まるともいえない。KEGG compound と KEGG reaction の N の存在比は約 29% とほぼ同じであるが増加率は 34% と 19% と 2 倍近い開きがある（赤字箇所を参照）。このことから N が様々な構造に

含まれており、その結果よく座標を分離すると考えられる。

S については（斜体箇所を参照）KEGG compound、KEGG reaction の存在比は約 10% とほぼ同じであるが、増加率は KEGG compound が約 15% に対し、KEGG reaction は約 7% と 1/2 となっている。KEGG compound は広く生体由来の化学物質が登録されているため、アミノ酸由来の 1 次代謝物質などが S を含む構造の多様性を大きくしていると考えられる。一方 KEGG reaction は生合成、代謝反応ペアとして特定された化学物質を収録しているため、解特定の化学構造に偏っていることを表していると考えられる。

P については（緑字箇所を参照）KEGG compound、KEGG reaction の存在比は約 9% と約 2% と開きがあるのに対し、増加率は約 4% と変わらない。P が含まれる構造はヘテロ環として環構造に P が見られる場合か、母骨格に官能基として P03 基が結合する場合の 2 通りでほぼ説明できる。多様性に乏しいための結果と考えられる。

以上のことから、N/C 軸を用いた 3 次元化がよく座標を分離することがわかった。また、O/C、N/C、S/C、P/C の軸は官能基の結合の程度を示すことができると考えられるため、解析対象とする代謝物質によって軸の組合せを 4.1.3 項で検討した。

4.1.3 KrevelenPlot の拡張 一軸一

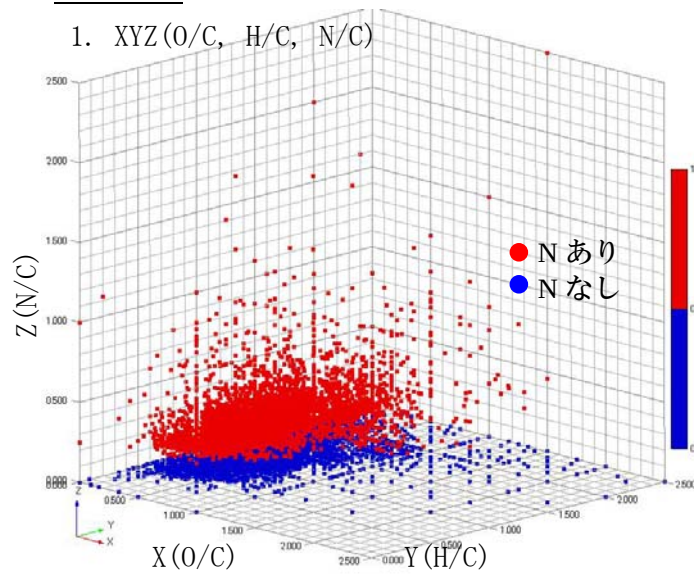
表 3-2-1-3 に示した軸の組合せを検証した。いずれも Z 軸方向に大きく分離するよう組合せと軸を定義した。3 次元の KrevelenPlot を図 4-1-3-1 から図 4-1-3-12 に示す。

表 3-2-1-3 拡張 KrevelenPlot の軸の組合せ

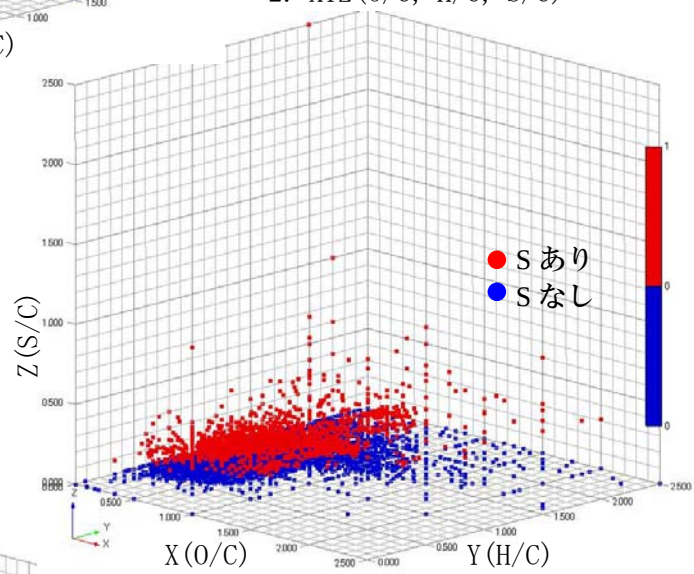
	X	Y	Z	表す特徴
1	O/C	H/C	N/C	N を含む構造の母骨格の特定
2	O/C	H/C	S/C	S を含む構造の母骨格の特定
3	O/C	H/C	P/C	P を含む構造の母骨格の特定
4	N/H	H/C	N/C	N を含む構造に結合した官能基の推定
5	S/H	H/C	S/C	S を含む構造に結合した官能基の推定
6	O/H	H/C	O/C	O を含む構造に結合した官能基の推定
7	O/C	N/O	N/C	N と O の結合の関係（ヘテロ環／官能基）と種類の特定
8	O/C	S/O	S/C	S と O の結合の関係（ヘテロ環／官能基）と種類の特定
9	O/C	P/O	P/C	P と O の結合の関係（ヘテロ環／官能基）と種類の特定

All Data

1. XYZ (O/C, H/C, N/C)



2. XYZ (O/C, H/C, S/C)



3. XYZ (O/C, H/C, P/C)

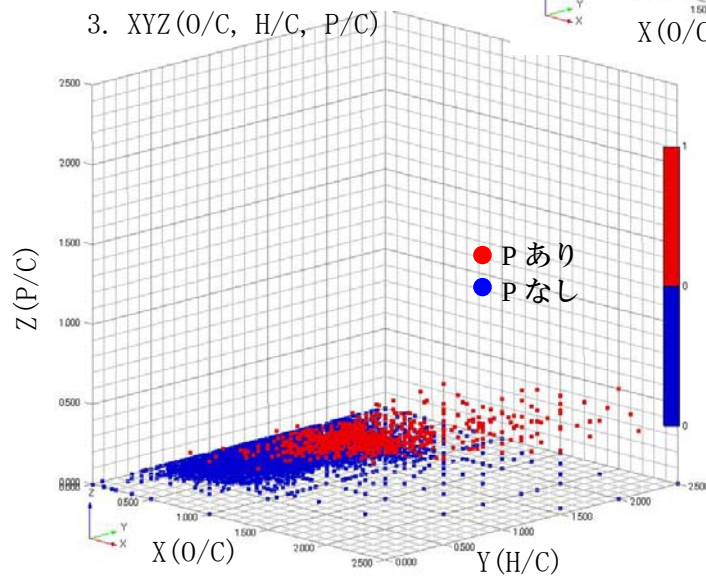
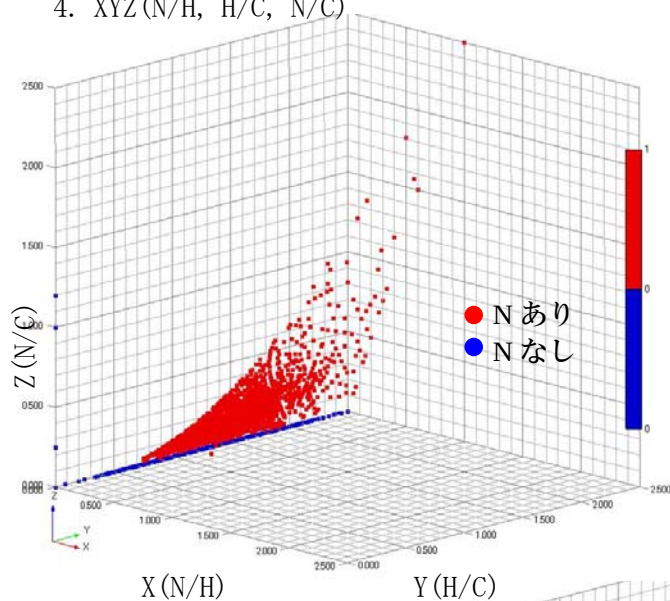


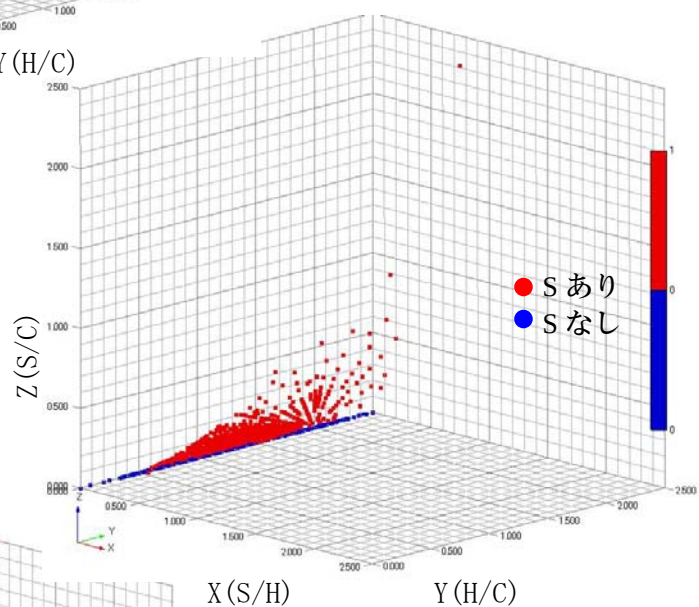
図 4-1-3-1 軸の組合せによる拡張 KrevelenPlot 表現 1

All Data

4. XYZ(N/H, H/C, N/C)



5. XYZ(S/H, H/C, S/C)



6. XYZ(O/H, H/C, O/C)

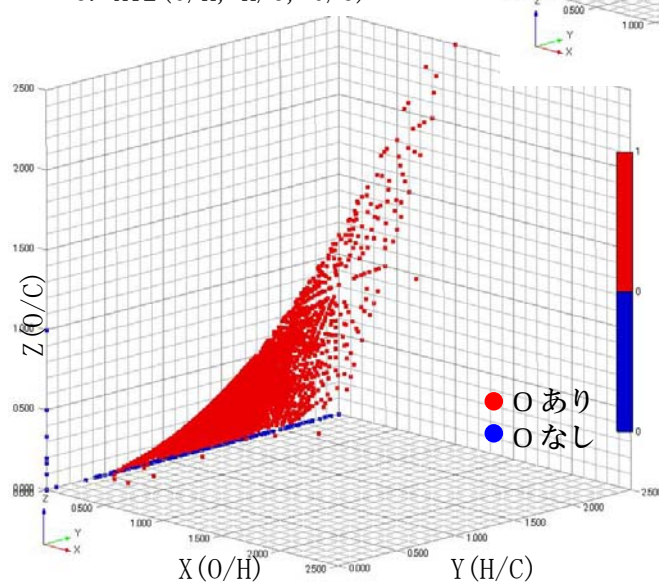
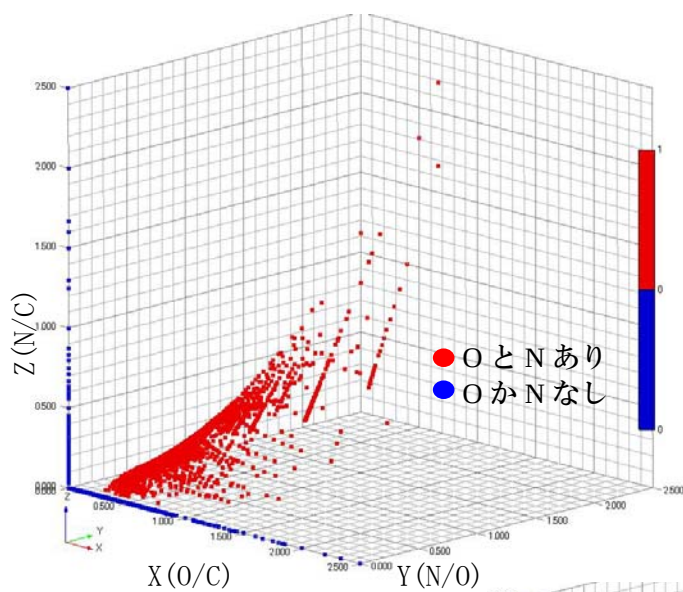


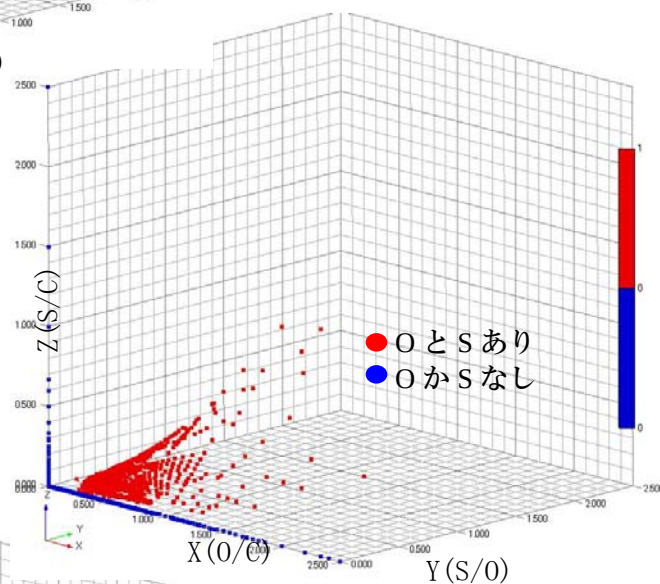
図 4-1-3-2 軸の組合せによる拡張 KrevelenPlot 表現 2

All Data

7. XYZ (O/C, N/O, N/C)



8. XYZ (O/C, S/O, S/C)



9. XYZ (O/C, P/O, P/C)

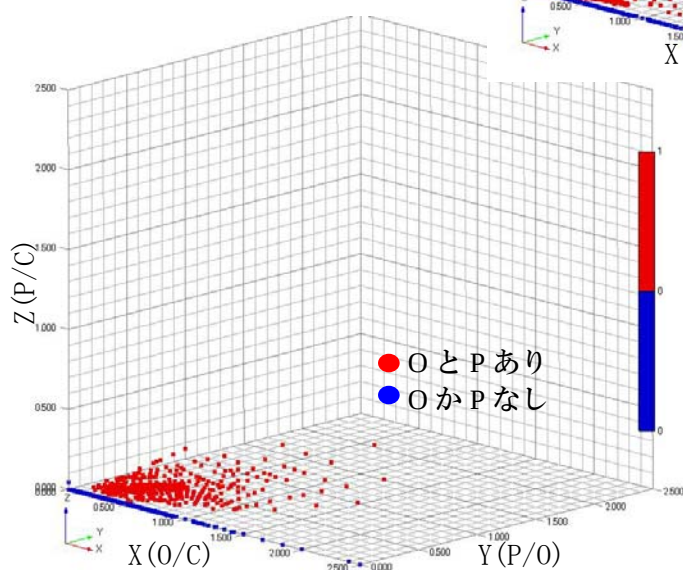


図 4-1-3-3 軸の組合せによる拡張 KrevelenPlot 表現 3

拡張パターンの No. 1 から No. 3 の $(X, Y, Z) = (O/C, H/C, N/C \text{ or } S/C \text{ or } P/C)$ では各元素が母骨格に結合した比率による座標領域の変化を見ることができる。N/C と S/C では N、S 元素を含まない組成式のグループと含む組成式のグループを比べたところ、X 軸 (O/C)、Y 軸 (H/C) の組成比に目立った特徴は見られない。N、S 元素と共に O、H 元素が多く結合することはないか、母骨格に結合する N、S を含む官能基の数が相対的に少ないため座標値に影響を与えない可能性が考えられる。一方 P/C では P 元素を含まない組成式のグループと含む組成式のグループを比べると P 元素を含むグループの方が X の値が大きくなっている。Y 値の変化は少ない。これは P 元素と共に O 元素が多く結合しているためと考えられ、これらの軸を用いることで、C、H、O 以外の元素が含まれる代謝物質の座標領域の特徴を示すと確認できる。4.2.2 項で詳しく検証する。

No. 4 から No. 6 の $(X, Y, Z) = (N/H \text{ or } S/H \text{ or } O/H, H/C, N/C \text{ or } S/C \text{ or } O/C)$ では、各元素が結合した比率と母骨格の構造との相関を見ることができる。いずれの元素においても同じ曲面上に現れるように見られる。

No. 7 から No. 9 の $(X, Y, Z) = (O/C, N/C \text{ or } S/C \text{ or } P/C, N/O \text{ or } S/O \text{ or } P/O)$ では、各元素が結合した比率と官能基の結合との相関を見ることができる。座標値が O、N、S いずれの元素でも同じ曲面上に現れるように見られる。曲面は構造との相関および官能基の種類による分布を表すと考えられる。この方程式を求めて、X、Y 座標値の取りうるパターンと併せて用いることで、質量分析データの精密分子量からの化学組成の推定に役立つと考えられる。

4.2 KrevelenPlot による代謝物質の特徴整理

4.2.1 母核の構造的特徴による座標の分離

KrevelenPlot 上に種類によって色を変えた代謝物質を描画してその座標領域の分離の程度を確認した。図 4-2-1-1 に示すように種類による分離が生じた。なお、それぞれの化合物群の基本骨格を図中に構造式で示した。

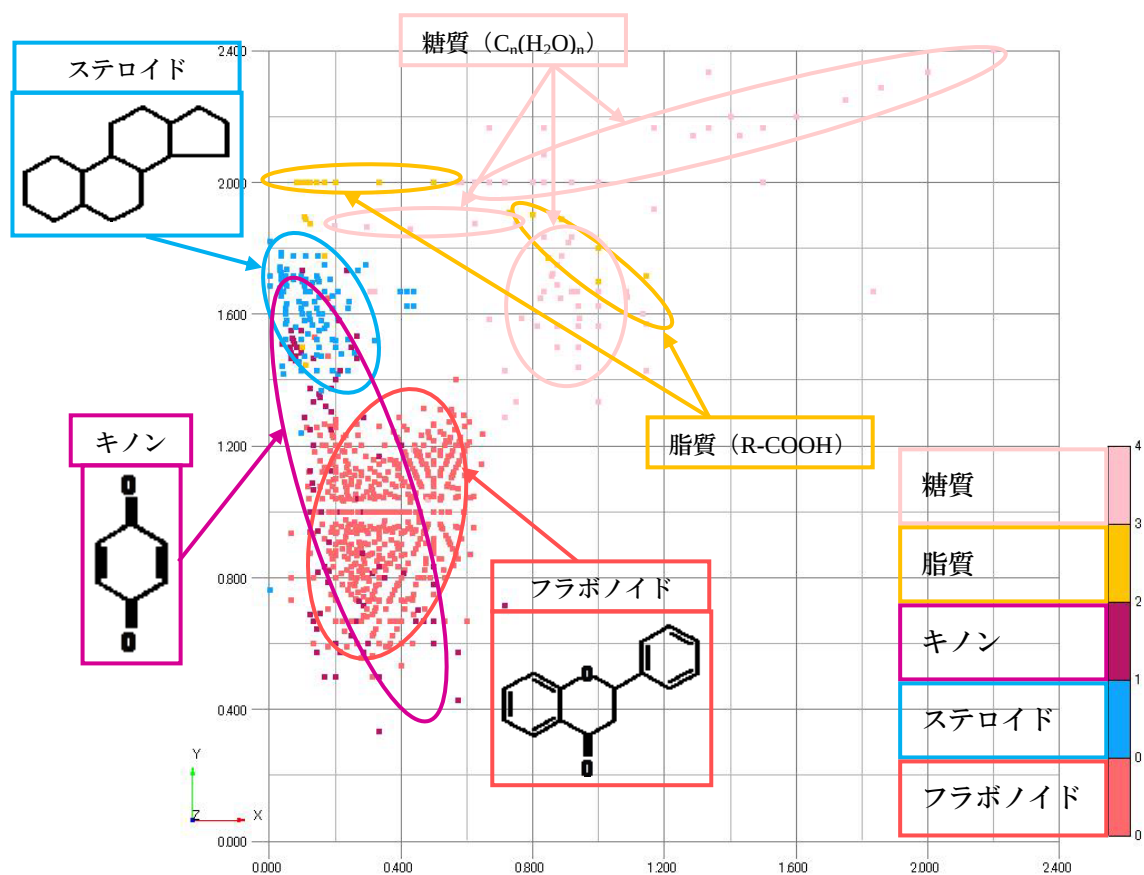


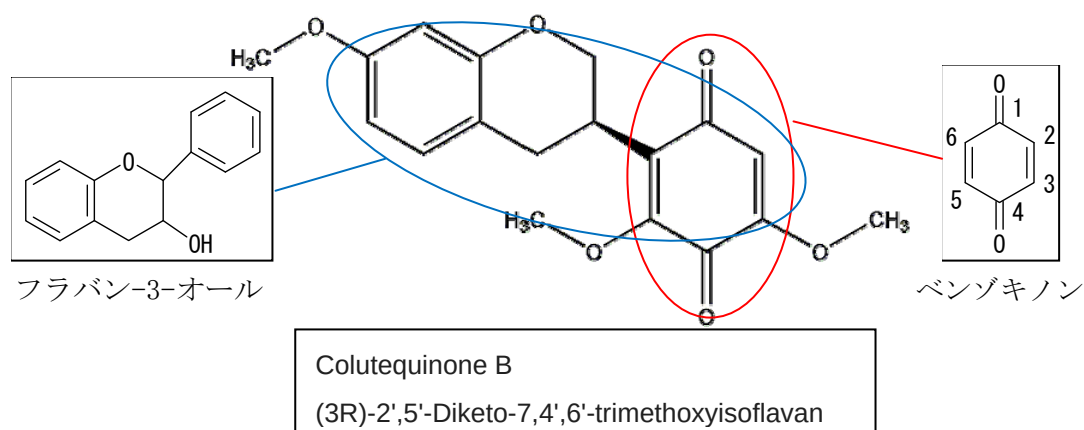
図 4-2-1-1 種類による座標の分離

	代謝物質の種類	データ件数
1	フラボノイド	2,685
2	ステロイド	194
3	キノン類	185

4	脂質	34
5	糖質	253

糖質と脂質は飽和結合が多いため H 元素の結合量も多く、Y 値が高い座標領域にのみ見られた。ステロイド、キノン類、フラボノイドは環構造が基本骨格であるため、糖質、脂質よりも H 元素の結合量が少なく、Y 値の低い座標領域に見られた。ステロイドは飽和結合の環構造を、キノン類は不飽和結合のみ、フラボノイドは不飽和結合と飽和結合を含みキノン類よりも環構造が大きい。これらを比較すると不飽和結合が多いほど、また不飽和結合を含む環構造が多く環同士が辺を共有しているほど Y 値が小さい値をとると見てとれる。ステロイドとフラボノイドはそれぞれ座標が分離しているが、両方の座標領域を含む広い範囲にキノン類が見られる。基本骨格に様々な修飾が加えられているために広範囲に分布していると考えられる。また、キノンとフラボノイド双方の母核を含むため、座標領域が重複する箇所が生じる。構造が一致している例を示す。

例)



座標値と構造的特徴の相関を探るためとりうる座標値の範囲と程度をヒストグラムで表した。基本骨格の座標値が範囲内のどの辺りに現れるかも併せて確認し、基本骨格からの構造の変化について調査した。図 4-2-1-2 から図 4-2-1-5 に示す。

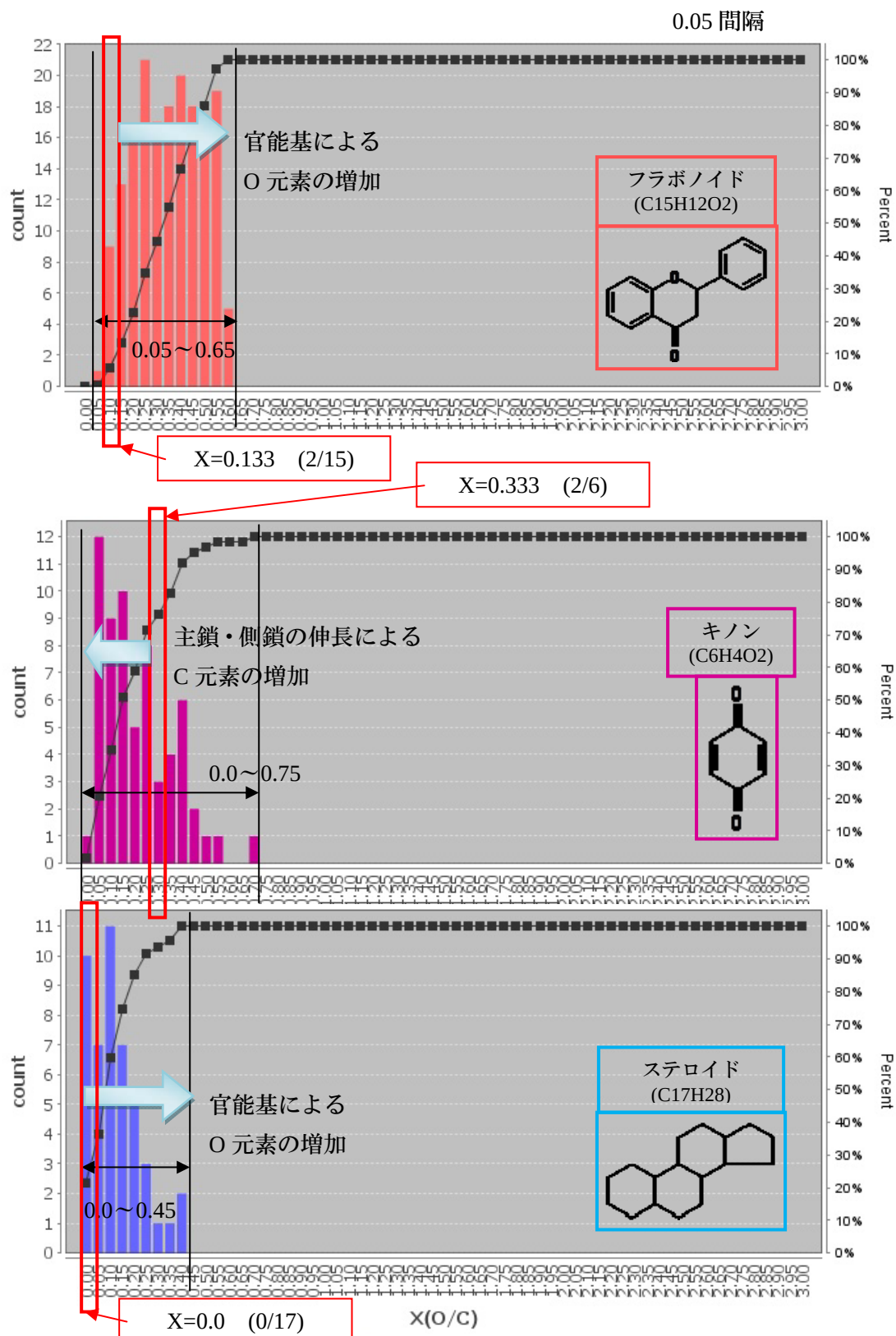


図 4-2-1-2 種類別の X (O/C) 座標値 (環構造あり)

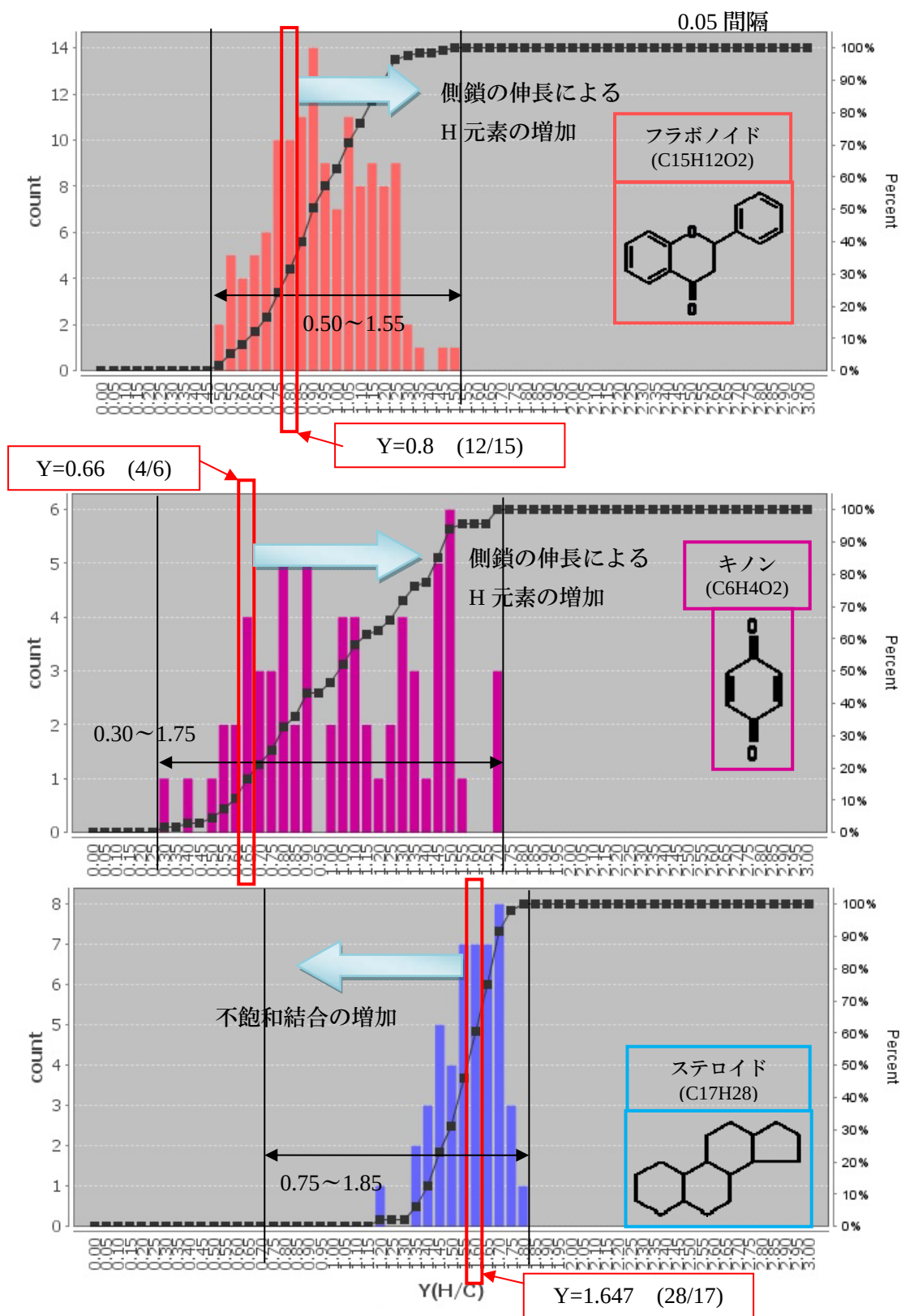


図 4-2-1-3 種類別の Y (H/C) 座標値 (環構造あり)

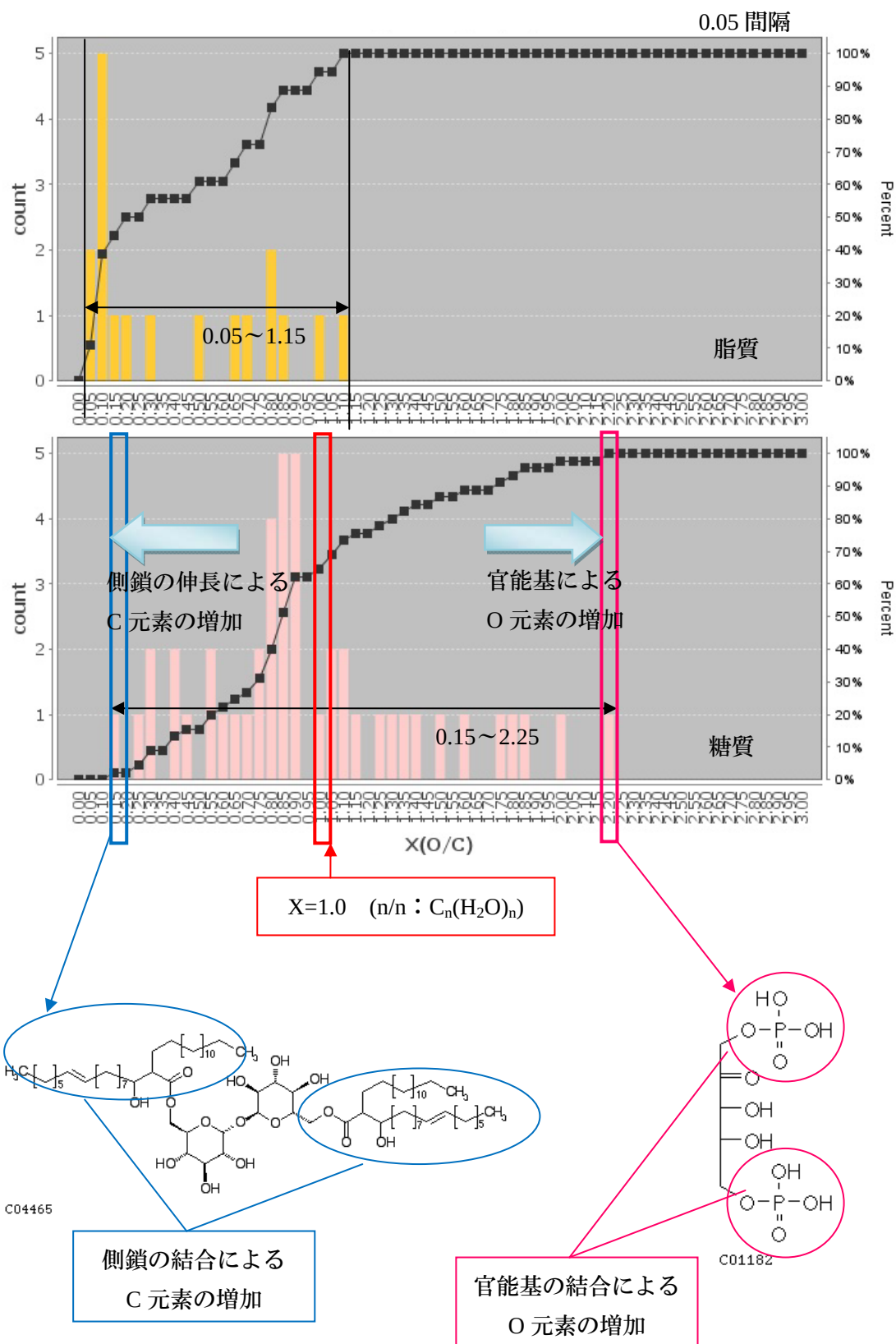


図 4-2-1-4 種類別の X (O/C) 座標値 (飽和結合)

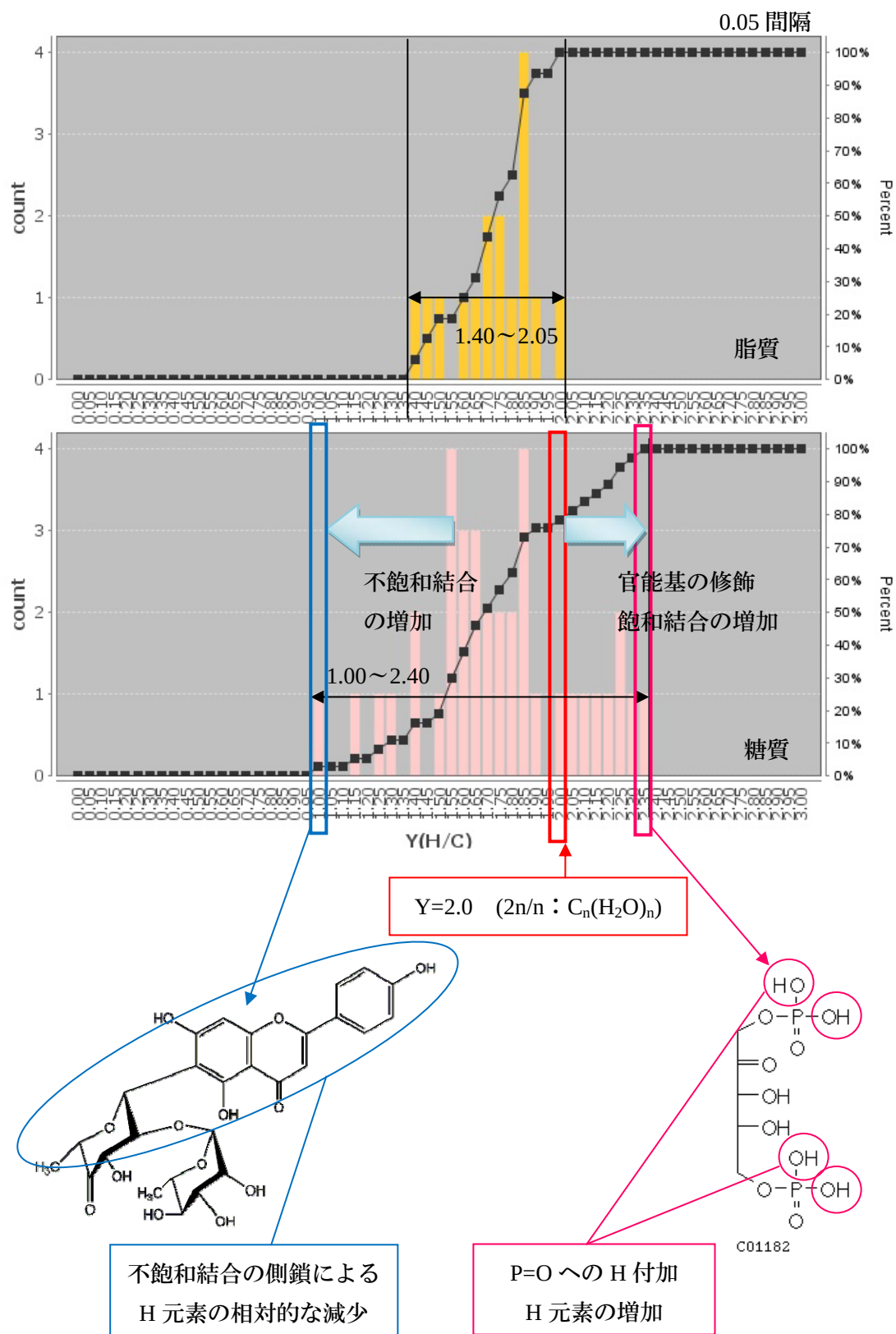


図 4-2-1-5 種類別の Y (H/C) 座標値 (飽和結合)

多環構造をとるフラボノイドやステロイドは基本骨格からの座標値の変化の幅が小さく、座標領域もある特定の範囲に現れやすい。多環構造をとることで官能基が結合できる部位や大きさの制限を受けるためと考えられる。キノン類や糖質では座標値の変化の幅が大きいため、座標領域が広がり他の種類の座標領域と重複が生じる。脂質はデータ数の少なさのため座標領域を特定できるほどの傾向が表れなかった。

図 4-2-1-6 に示すように、KrevelenPlot の XY 平面でデータが集中している範囲は O と H の結合の比率による分離ができると考える。

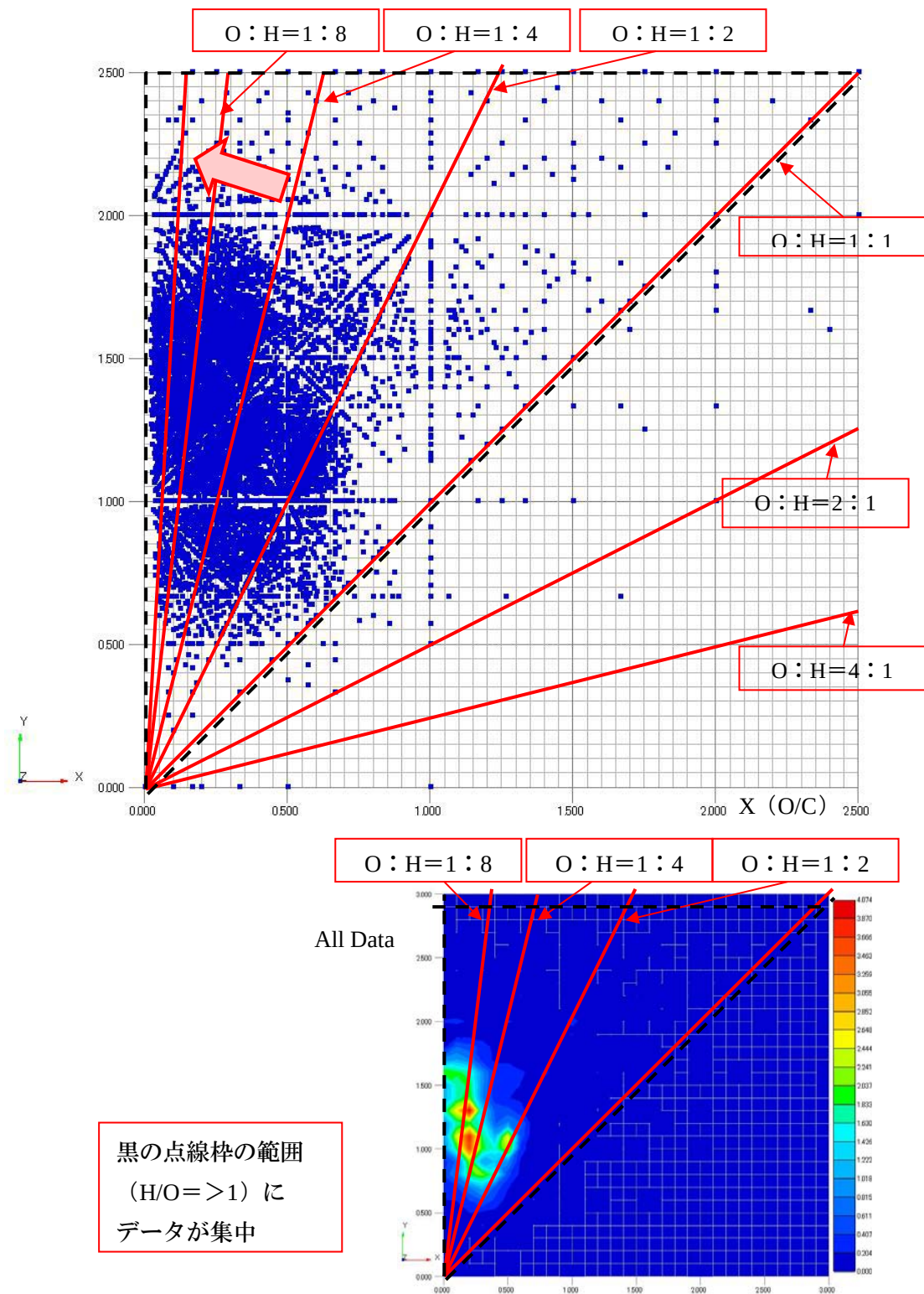


図 4-2-1-6 KrevelenPlot の XY 平面からわかる特徴

図 4-2-1-6 で示した通り、 $H/O > 1$ の範囲にデータが集中している。このため、Z 軸に H/O を用いて、データの集中する範囲の座標を Z 軸方向に分離を図った。O 元素の結合数に比べ H 元素の結合数の多いデータほど Z 値が高くなるため、酸素性官能基 ($-OH$ 、 $-COOH$ など) の結合の少ない高分子ほど Z 軸方向に高い値をとる。逆に官能基が多く結合する代謝物質は Z 軸方向に低い値をとる。構造については Y 軸から、官能基の結合の比率については X 軸と照らし合わせることで判別可能である。結果を図 4-2-1-7 に示す。

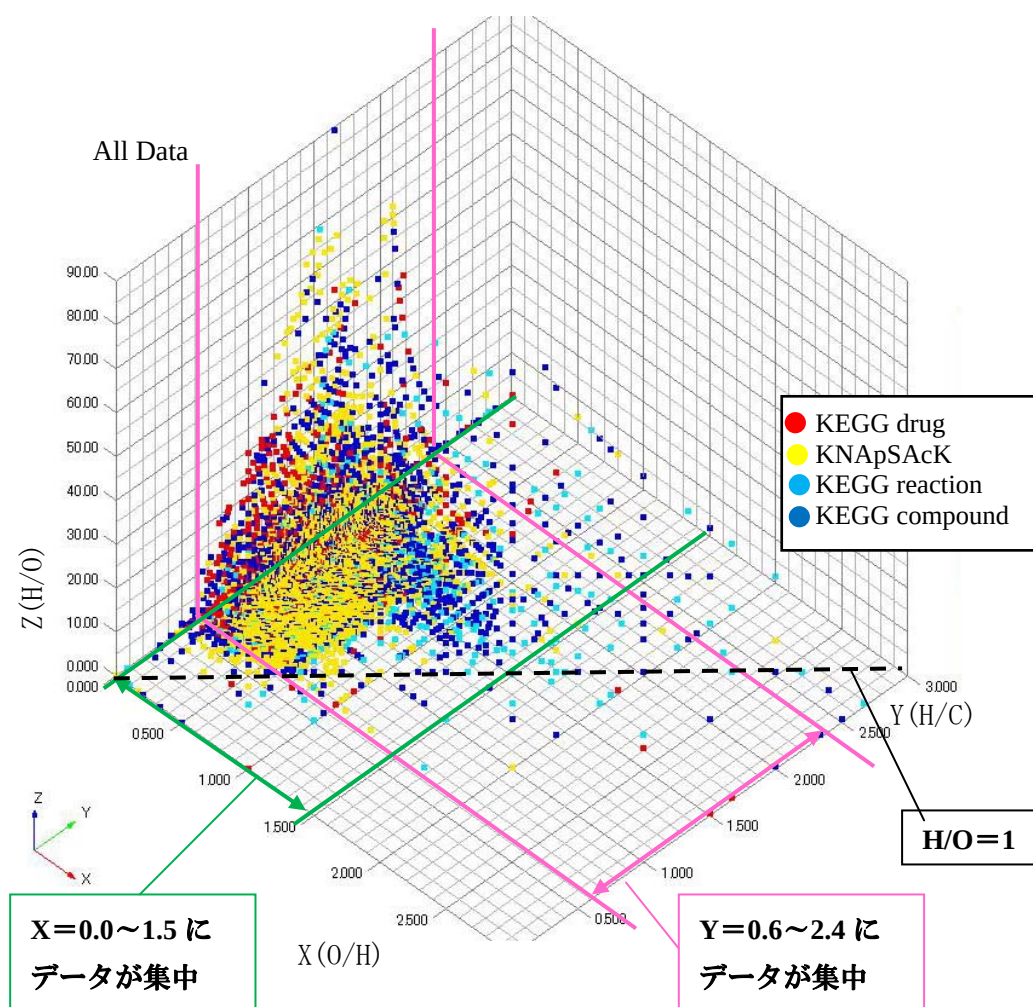


図 4-2-1-7 3DKrevelenPlot (O/C, H/C, H/O)

データごとに凡例色を変えて表示したが、目立った偏りは見られなかった。この Z 軸を用いた場合、全体的に X 軸に近づくほど Z 軸に延びる形で座標値が現われる。O 元素を含まない化学組成は $Z=0$ となるため X、Y 軸でしか分離できないが、この Z 軸により

全体の 95% (0 元素なし/C 元素あり : 1, 913/43, 888) が $Z > 0$ の値をとるため特徴の把握に役立つ。 $H/O = > 1$ のデータのほぼ全てが $0.0 < X < 1.5$ 、 $0.6 < Y < 2.4$ の範囲に収まることが分かる。Y 値と構造の相関については図 4-2-1-8 に示したように、C と H の比率は構造に依存している。代謝物質は環構造や不飽和結合を多く含む高分子であるため C 元素が複数連なって構成され、ほぼ全てのデータは $Y < 3$ の範囲に収まると考えられる。Y 値が小さくなるほど、分子量が増え構造が複雑になり、環化や不飽和結合の増加を示すと考えられる。環化や不飽和結合による H 元素の脱離が起こるにつれ Y 値も小さくなる。この組成比のパターンを質量からの理論化学式の導出に利用することで、計算時間と精度の向上が見込める。

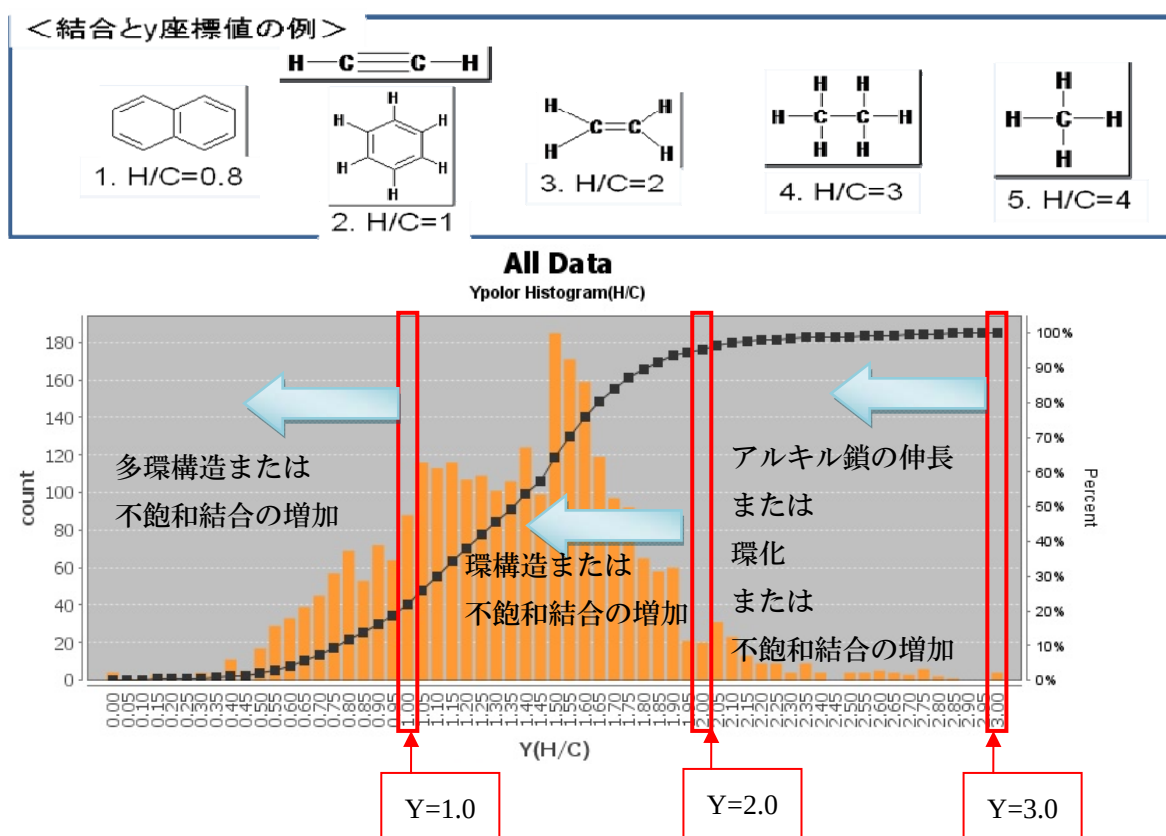


図 4-2-1-8 Y 値と構造の相関

H/C の値と、 H/O の値から構造と官能基 (OH 基) の多少、また H 元素の結合数の多少について判別できる。また、H の結合数から分子量の大小についても確認できると考えた。まず表 4-2-1-1 に示すように、 H/C の値を構造との相関をもつ 5 つの区間に分類し

た。構造と H 元素の多少について図 4-2-1-9 に示す。

表 4-2-1-1 Y 値と構造的特徴

No	Y 値	特徴	件数
1	$0.0 \leq Y < 1.0$	不飽和結合の多い多環構造	9,073
2	$1.0 \leq Y < 2.0$	不飽和結合の含まれる環構造 飽和結合の多い多環構造	32,420
3	$2.0 \leq Y < 3.0$	直鎖構造	2,233
4	$3.0 \leq Y < 4.0$	H 元素の多い官能基による修飾	147
5	$4.0 \leq Y$	H 元素の多い官能基による修飾	15
	計		43,888

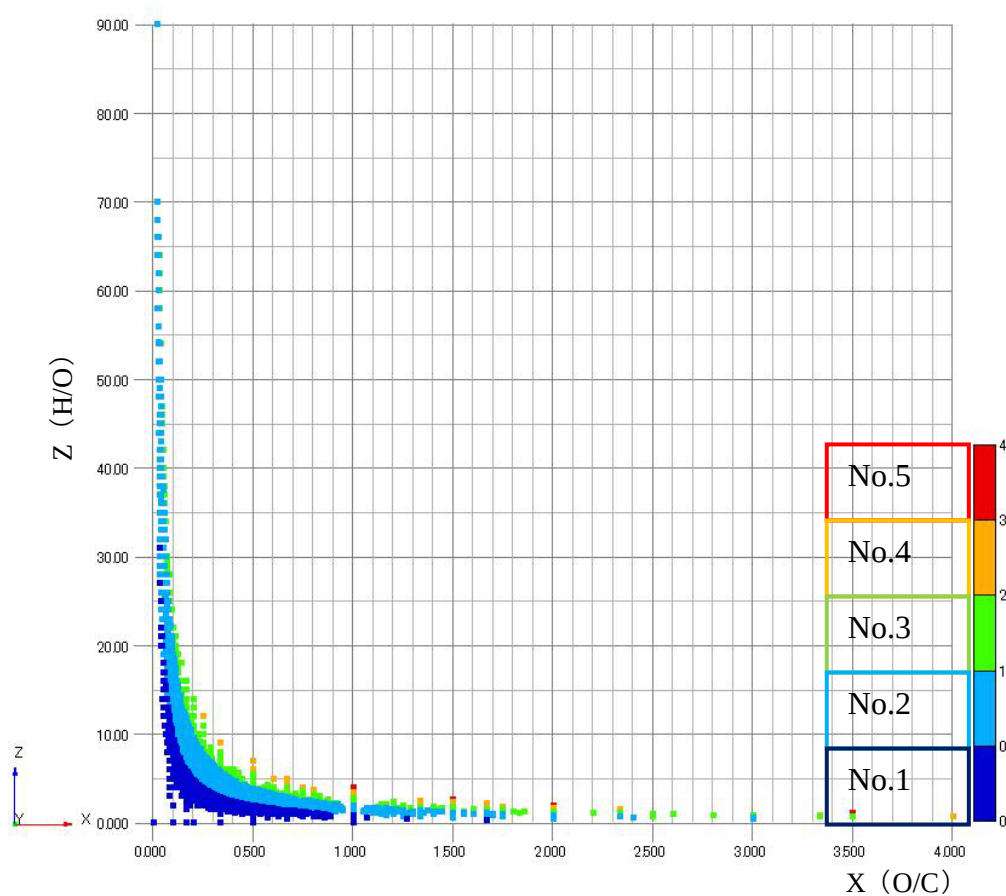


図 4-2-1-9 H 元素の結合数と母骨格の相関

図 4-2-1-9 から No.2 で示した $1.0 \leq Y < 2.0$ の範囲の代謝物質は、他の範囲の代謝物質に比べ H 元素の結合数が多い。そこで YZ 平面で捉えた場合の構造と分子量の相関について検証した。図 4-2-1-10 と表 4-2-1-2 に結果を示す。

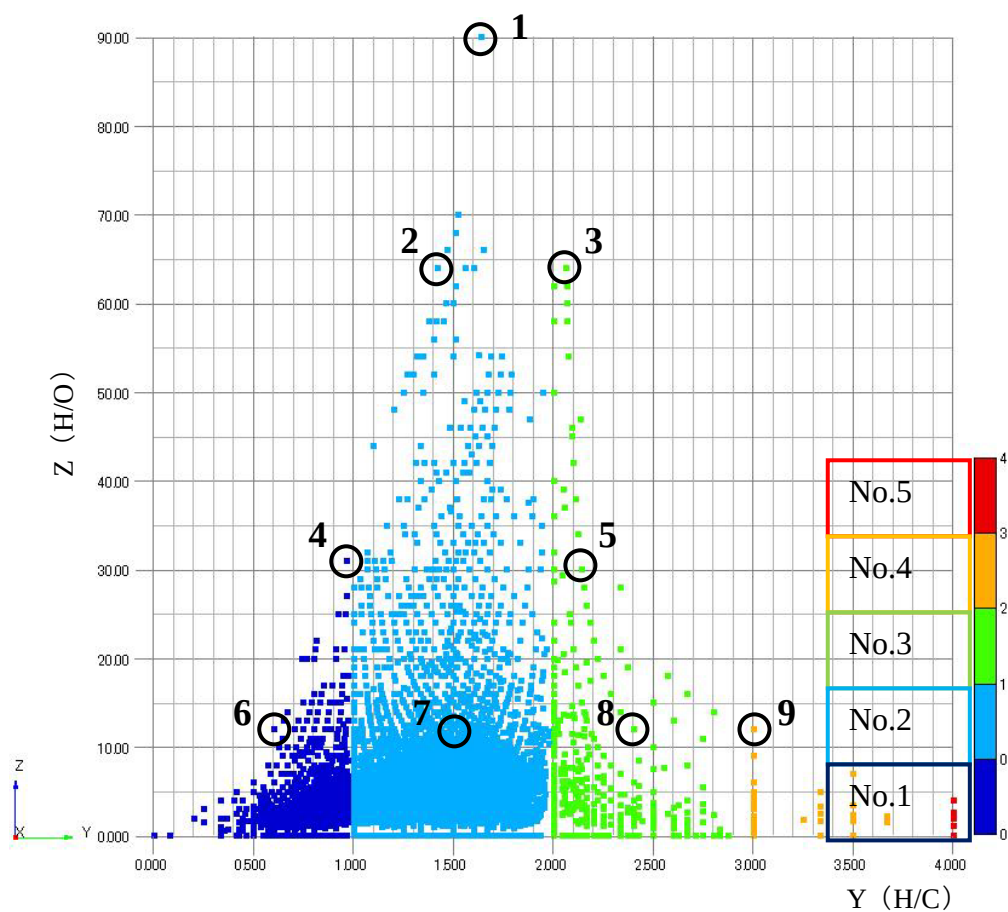
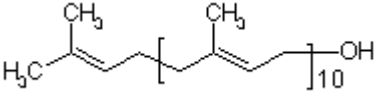
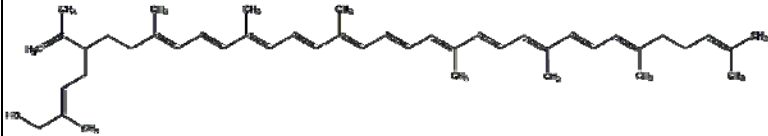
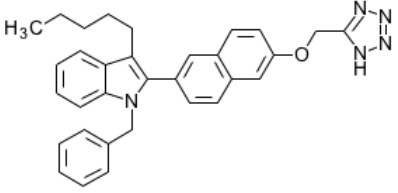
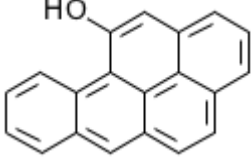
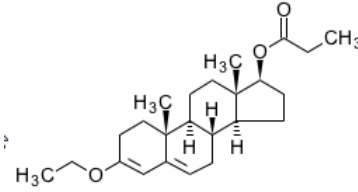
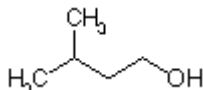
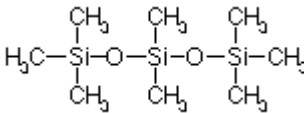


図 4-2-1-10 座標値と構造の相関

表 4-2-1-2 座標値と構造の相関

No	化学式	構造	座標値
1	C55H900	飽和結合と不飽和結合を含む 直鎖状で分子量大 	y=1.636 z=90
2	C45H640	飽和結合と不飽和結合を含む 直鎖状で分子量大 	y=1.422 z=64
3	C31H640	飽和結合のみ含む 直鎖状 	y=2.065 z=64
4	C32H31N50	環構造、不飽和結合が多い ヘテロ環を含む 	y=0.969 z=31
5	C14H300	飽和結合のみ含む 直鎖状 	y=2.143 z=30
6	C20H120	環構造、不飽和結合が多い 	y=0.6 z=12
7	C24H3603	環構造が多い 不飽和結合が少ない 	y=1.5 z=12

8	C ₅ H ₁₂ O	飽和結合のみ含む 直鎖状で低分子		y=2.4 z=12
9	C ₈ H ₂₄ O ₂ Si ₃	飽和結合のみ含む 直鎖状で低分子 母骨格に C 元素以外を多く含む		y=3.0 z=12

Z 軸を上から下に見ていくと (No. 1 から No. 9 まで) Z 値が小さいものほど低分子となる。また Y 値を右から左に見ていくと (No. 9 から No. 6 まで) Y 値が大きいものは、飽和結合の直鎖状であるが、Y 値が小さくなるに従って、まず Y=2.0 付近まではアルキル鎖が伸長し飽和結合が増えていく (No. 3、5、8)。Y=1.5 付近では直鎖状のものと環構造をとるものが混じり、直鎖状のものは不飽和結合が増えていく (No. 2、7)。Y<=1.0 では環構造内での不飽和結合が増えていく (No. 4、6)。組成比から構造と分子量との相関を捉えられるため、この結果を用いて質量分析データの質量値の中で取りうる組成比に当てはめると構造の推測ができる。

4.2.2 官能基による座標の分離

N/C、S/C、P/C 軸により置換基の種類を判別できると考えられる。特定のヘテロ元素 (N、S、P) の有無による分布の差異を見るため、凡例色で有無を示した KrevelenPlot の XY 平面を図 4-2-2-1 に示す。

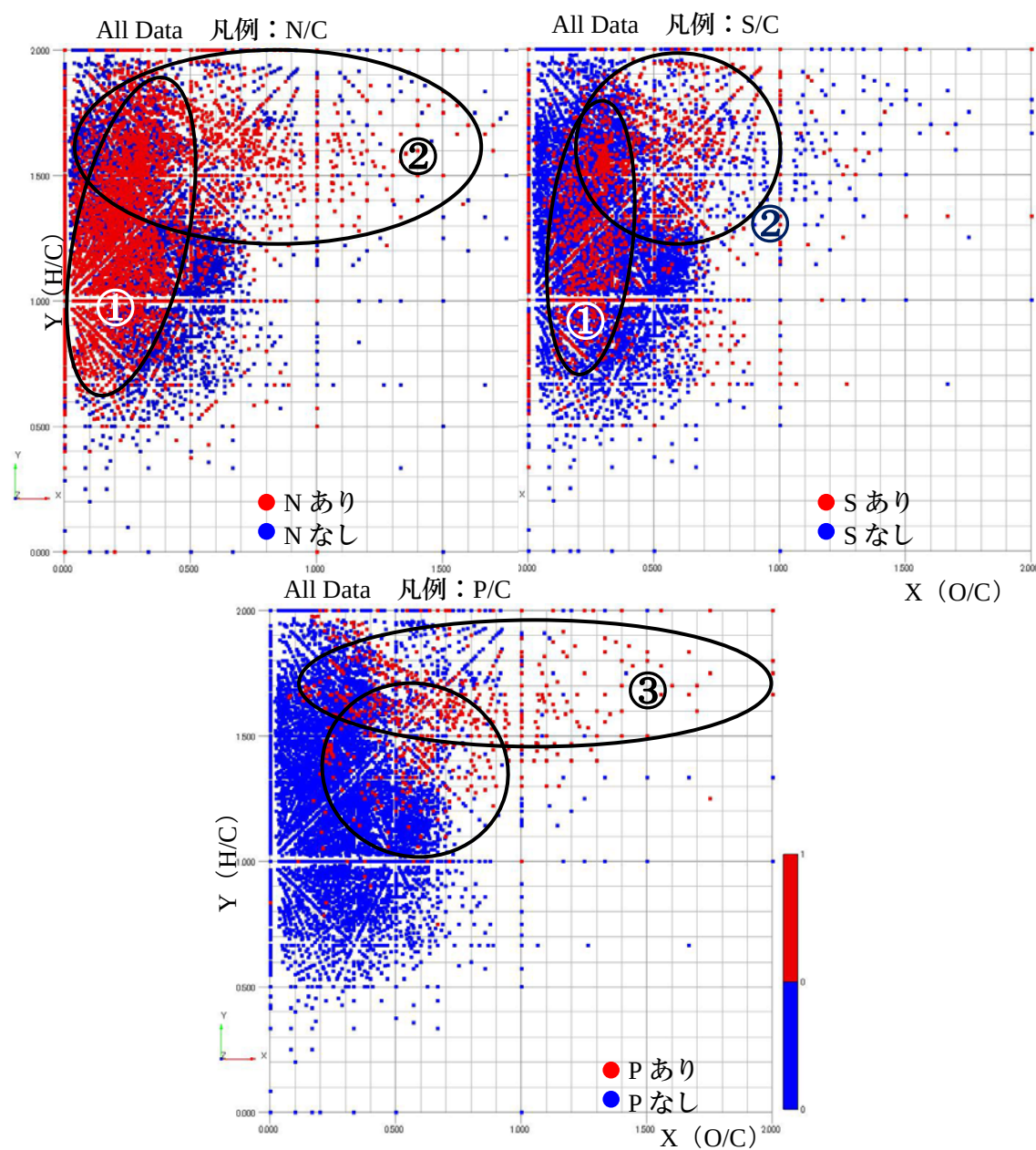


図 4-2-2-1 特定元素の有無による座標領域 (N、S、P)

N 元素、S 元素は X 値が小さい（0 の結合数が少ない）範囲では Y 値が広く分布（多様な構造に結合）している（図 4-2-2-1、①）。また Y 値が大きい（H の結合数が多い、すなわち飽和結合が多い）範囲では X 値が広く分布している（図 4-2-2-1、②）。ただし N 元素の方が S 元素より分布範囲が広い。一方 P 元素は Y 値が大きい範囲では X 値が広く分布しているが（図 4-2-2-2、③）、Y 値が小さい範囲にはほとんど見られない。

N、S、P を含む組合せで座標領域の分離を確認した結果を図 4-2-2-2 に示す。

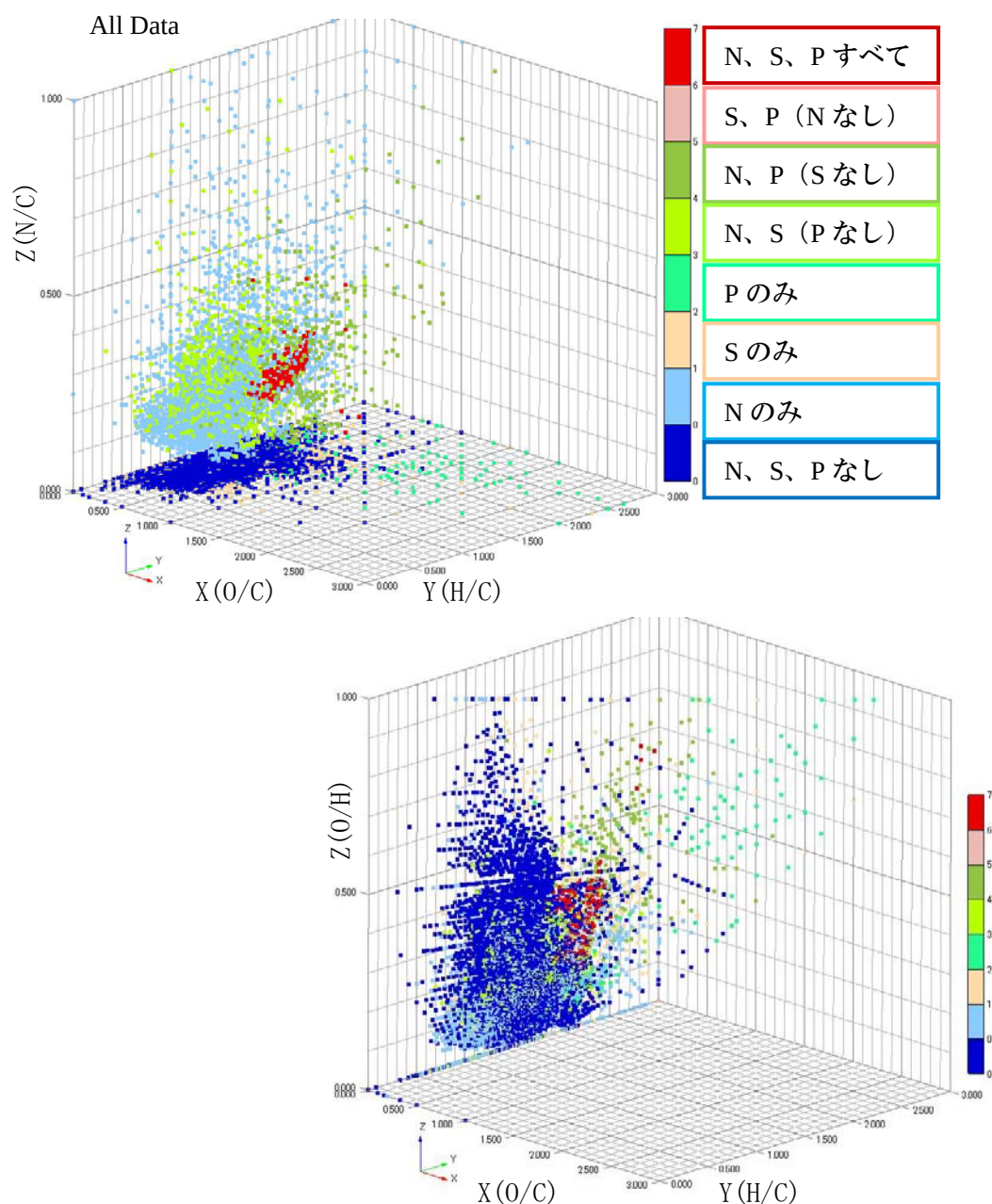


図 4-2-2-2 元素の組合せでみる座標領域（N、S、P）

N/C 軸を用いた 3 次元表現でみると、N のみを含む場合、Y 値が低い範囲に偏って分布している。N と S を含む場合は N のみの場合に比べ Y 値が高い領域に分布している。N と P を含む場合は N と S の場合よりも全体的に Y 値が高くなっており、X 値も高い。S のみを含む場合、Y 値が 1.0 付近と 2.0 付近に別れて分布している。S と P を含むものはほとんど見られない。P のみを含む場合、X 値 Y 値の高い範囲に分布している。N と S、P のすべてを含む場合、Y 値が高い辺りに集中して分布している。含まれる元素によって X、Y 座標値が大きく影響を受けることがわかる。その元素に特徴的な分布範囲があるため、拡張した KrevelenPlot 上に表現することで構造と官能基、ヘテロ環などの化学的特性を把握できる。

図 4-2-2-2 の Z 軸に O/H を定義したグラフでは、N、S、P 元素は Y 軸(+)方向へ曲面に沿って現れている。各元素と H 元素の相関が深いと考えられる。そこで、N/C、S/C、P/C と H/C 軸の組合せで表現した結果を図 4-2-2-2 に示す。N、S、P 元素が結合している母骨格の推定と多く結合している官能基の種類を把握できると考える。

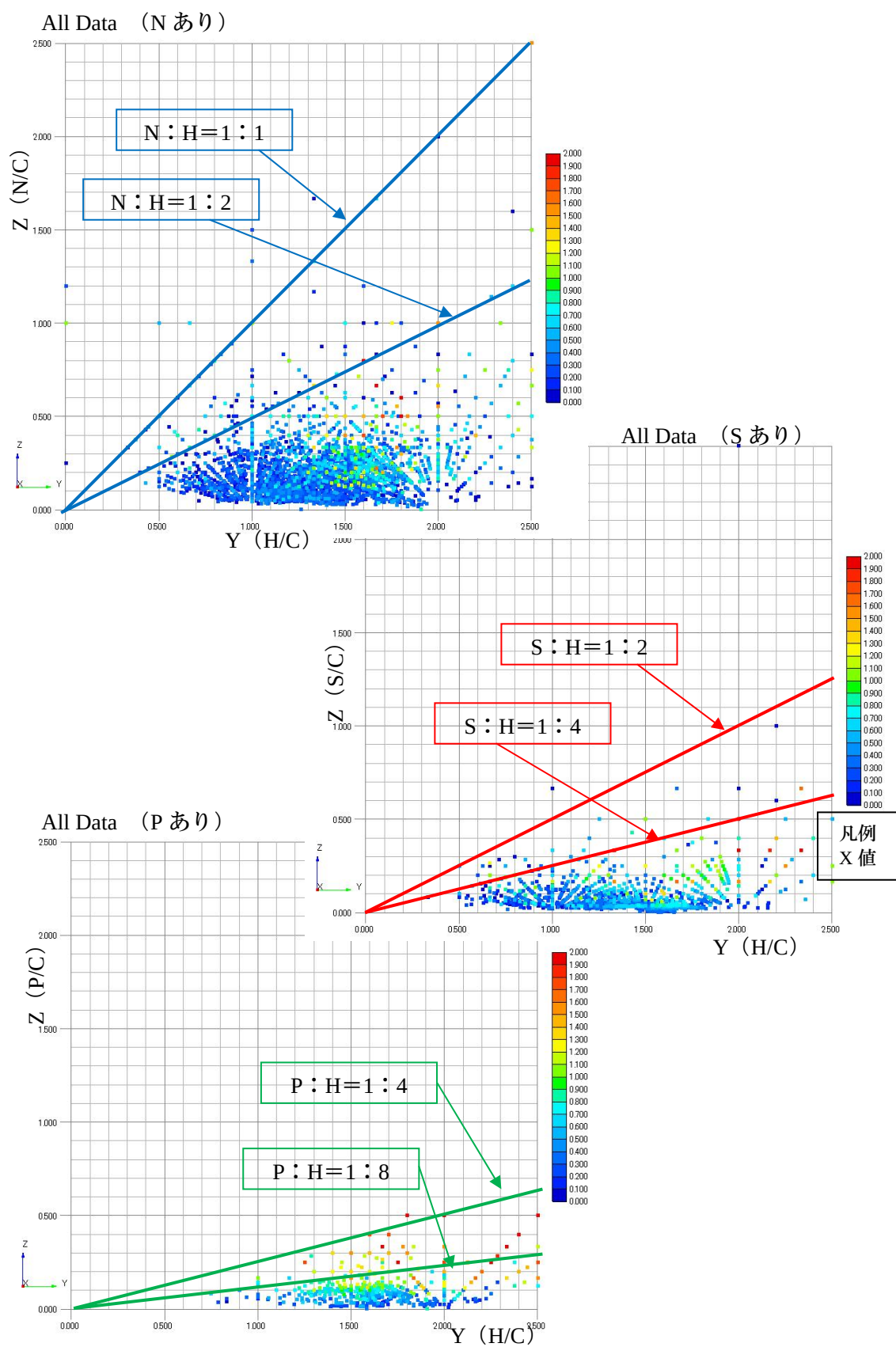


図 4-2-2-3 YZ 平面で見る座標領域 (N、S、P)

凡例色より Y 値が高く、Z 値も高い (N、S、P の結合数が多い) 化学組成式ほど O 元素の結合も多くなる傾向がある。また、N 元素では $N/H=1/2$ 以下の組成比が多く、S 元素は $S/H=1/4$ 以下の組成比、P 元素は $P/H=1/8$ 以下の組成比が多いとわかる。N 元素は H 元素の $1/2$ 倍以下の組成数が結合している代謝物質が多く、Y 値から多環構造かつ不飽和結合の多い母骨格から飽和結合と不飽和結合の含まれる母骨格への結合と多様な構造に結合している。また、N 元素は C 元素の半分以上の組成数を持つものが多い。Y 値の小さい代謝物質では N 元素の結合は H 元素の増加がないヘテロ環、もしくはアミド、アミンが考えられる。Y 値の大きい代謝物質ではアミド、アミンに加え NO_2 基、 NH_2 基などの結合も考えられる。S 元素は H 元素の $1/4$ 倍以下の組成数が結合している代謝物質が多く、Y 値から環構造で不飽和結合がやや多い母骨格に結合している。C 元素の $1/4$ 倍以下の組成数を持つものが多い。Y 値の小さい代謝物質にはヘテロ環、SH 基、 $-\text{SO}_3\text{H}$ 基などの結合が、Y 値の大きい代謝物質には SH 基 SO_3H 基の結合が考えられる。P 元素は H 元素の $1/8$ 倍以下の組成数が結合している代謝物質が多く、Y 値から不飽和結合がやや多い母骨格に結合している。C 元素の $1/8$ 倍以下の組成数を持つものも多く、Y 値の小さい代謝物質にはヘテロ環、 PO_3 基の結合が、Y 値の大きい代謝物質には PO_3 基の結合が考えられる。

官能基について調べるには軸に O 元素も加えた方がよりよく判別できる。そこで、H 元素を軸に入れて構造を見た場合と、O 元素を軸に入れて官能基を見た場合を試した。これらは 4.1.3 項で曲面が現れた軸である。結果を図 4-2-2-4 に示す。

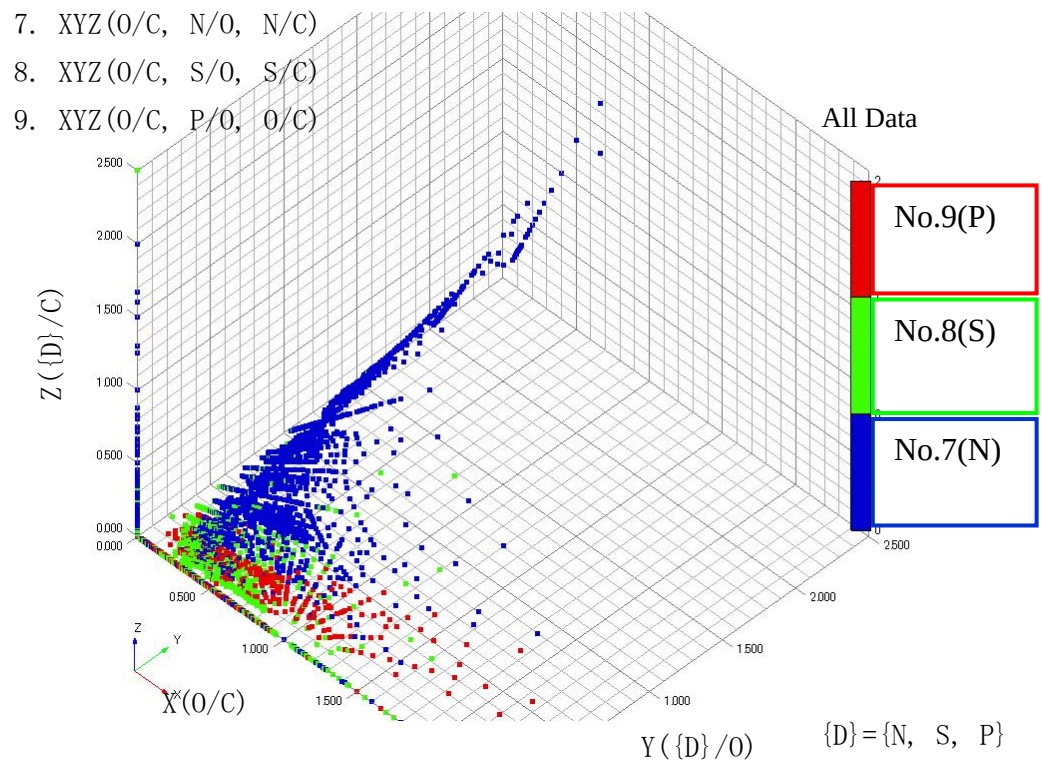
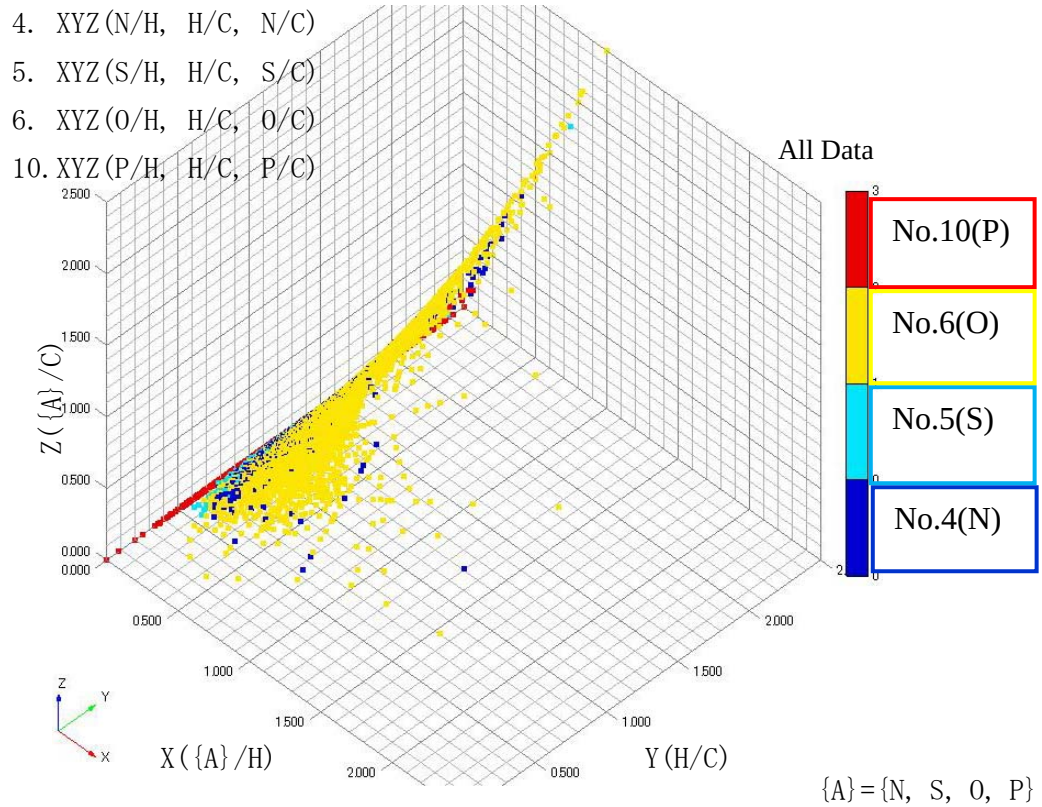


図 4-2-2-4 構造と座標値の法則性 (N、S、P、O)

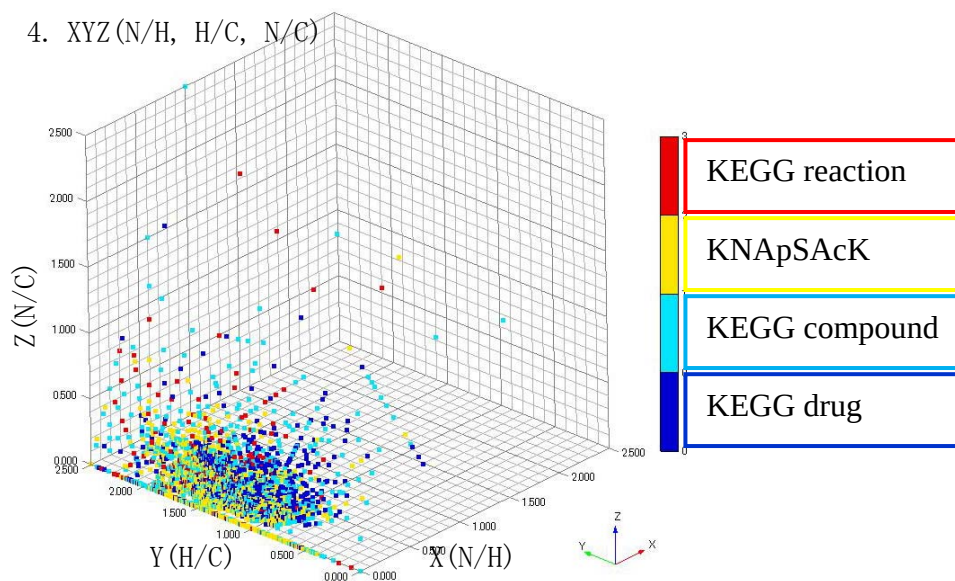
いずれの元素の場合も同一曲面に現れる。H 元素を加えた軸では構造との相関が、O 元素を加えた軸では O を含む官能基の結合のパターンとの相関があると考えられる。また、元素によって曲面のどの辺りに現れるのか特徴がある。結合の法則と天然物の構造の分布、またヘテロ元素による構造の違いをこの曲面で表せるため、この曲面の方程式を求めることで、質量分析結果の理論化学式の計算時間と精度の大きな向上が見込める。また、新規に発見した代謝物質の座標値が曲面のどの辺りに現れるかで構造や特徴の把握が可能である。あるいは合成物質の座標値と曲面との距離により天然物と合成物の構造の違いを見ることが出来る。

天然物の構造と類似度の高い合成物である KEGG drug と天然物の KNApSAcK（二次代謝物質）、KEGG compound（一次代謝物質含む）の分布、また反応ペアとして既に解明済みの KEGG reaction を比較してみた。N と S を用い No. 4 と No. 7、8 によりグラフを描いた結果を図 4-2-2-5 に示す。各データ数は表 4-2-2-1 の通りである。

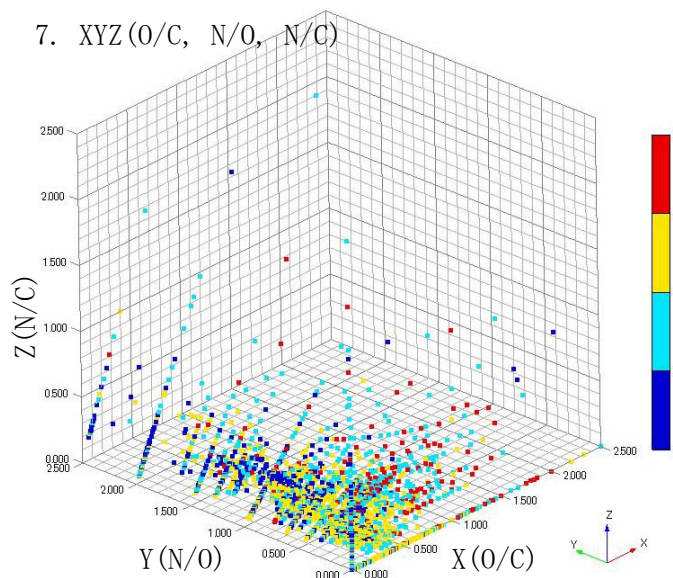
表 4-2-2-1 データ件数 (N、S)

元素	データ件数			
	KEGG drug	KNApSAcK	KEGG compound	KEGG reaction
N	2, 080	2, 865	3, 317	1, 193
S	993	729	1, 194	440

4. XYZ(N/H, H/C, N/C)



7. XYZ(O/C, N/O, N/C)



8. XYZ(O/C, S/O, S/C)

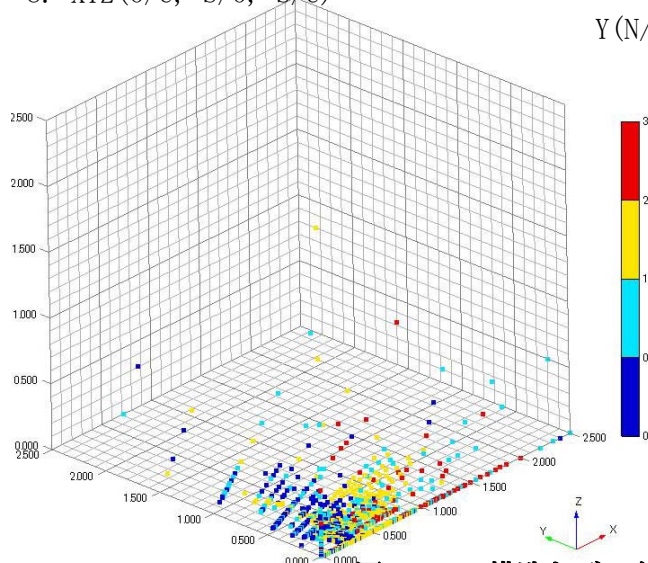


図 4-2-2-5 構造とデータ別の座標分布

KEGG reaction とその他を比べてみると反応経路が解明されていない代謝物質の特徴が把握できる。また KEGG drug と KNApSAcK、KEGG compound を比べると drug の分布が小さく、利用されていない天然物があることが判別できる。新薬開発や代謝反応の解明において未開拓の化合物空間を調べるために拡張 KRevelenPlot は有用と考えられる。

4.3 KrevelenPlot による代謝反応の視覚化

4.3.1 代謝反応の整理

KNapSAcK（二次代謝物質）データから、代謝反応によって化学組成式が変化したと考えられる代謝反応前物質（基質）と代謝反応後物質（生成物）の代謝反応ペアとして抽出し、反応が同じものをグループにした。表 3-2-3-1 に示したグループを拡張 KrevelenPlot 上で代謝反応を描き、その特徴を検証した。件数の少ないものは特徴を把握できなかったため、No. 1 から 8 までの結果を図 4-3-1-1 から図 4-3-1-3 に示す。

表 4-2-3-1 代謝反応の種類（転記）

	拡張	X	Y	Z	代謝反応
1	Z(O/H)	O/C	H/C	O/H	CH ₃
2	Z(O/H)	O/C	H/C	O/H	OH
3	Z(O/H)	O/C	H/C	O/H	C ₂ H ₂ O
4	Z(O/H)	O/C	H/C	O/H	C ₃ H ₂ O ₃
5	Z(O/H)	O/C	H/C	O/H	C ₆ H ₁₀ O ₃
6	Z(O/H)	O/C	H/C	O/H	C ₆ H ₁₀ O ₄
7	Z(O/H)	O/C	H/C	O/H	C ₇ H ₄ O
8	Z(O/H)	O/C	H/C	O/H	H ₂
9	Z(N/H)	O/C	H/C	N/H	C ₉ H ₉ O _N
10	Z(N/H)	O/C	H/C	N/H	CH ₃ ON
11	Z(N/H)	N/C	H/C	N/H	NH
12	Z(N/O)	O/C	H/C	N/O	C ₉ H ₉ O _N
13	Z(N/O)	O/C	H/C	N/O	CH ₃ ON
14	Z(N/O)	N/C	H/C	N/O	NH
15	Z(P/O)	O/C	H/C	P/O	H ₂ O ₆ P ₂
16	Z(P/O)	O/C	H/C	P/O	H ₃ O ₄ P
17	Z(P/O)	O/C	H/C	P/O	H ₃ O ₃ P
18	Z(P/O)	O/C	H/C	P/O	H ₄ O ₄ P
19	Z(S/H)	O/C	H/C	S/H	O ₄ S
20	Z(S/O)	O/C	H/C	S/O	O ₄ S

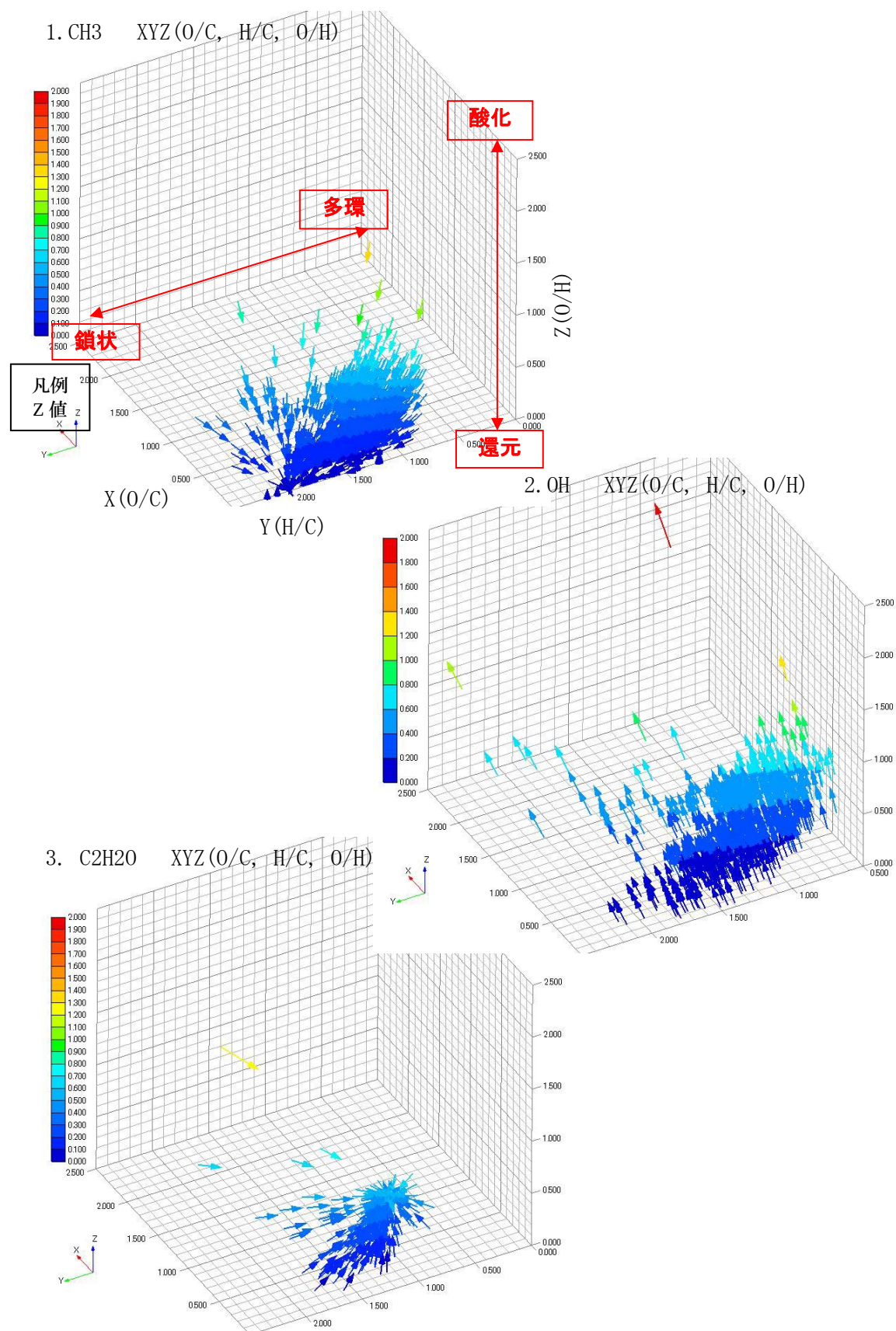
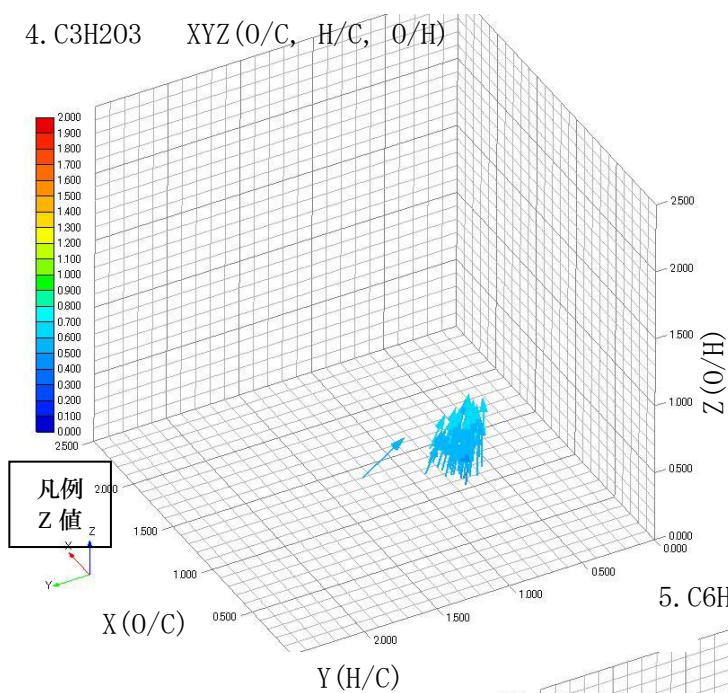
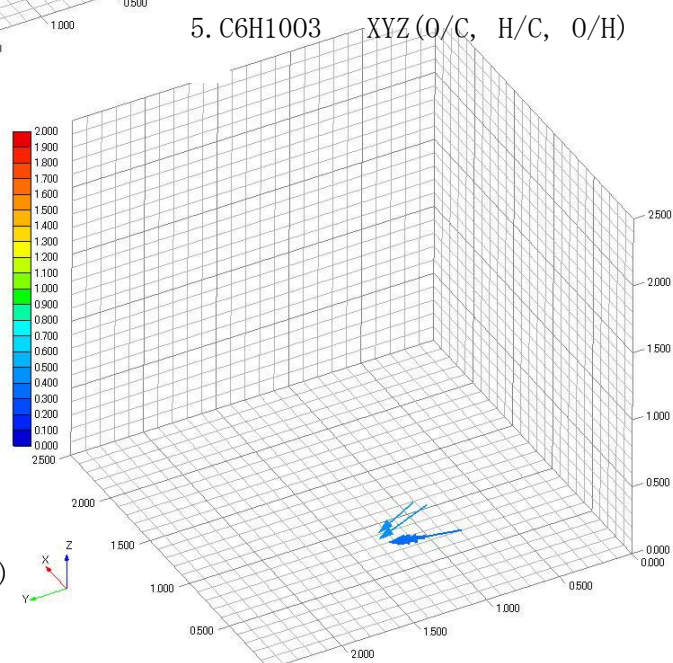


図 4-3-1-1 代謝反応とその分布 1

4. C₃H₂O₃ XYZ (O/C, H/C, O/H)



5. C₆H₁O₃ XYZ (O/C, H/C, O/H)



6. C₆H₁O₄ XYZ (O/C, H/C, O/H)

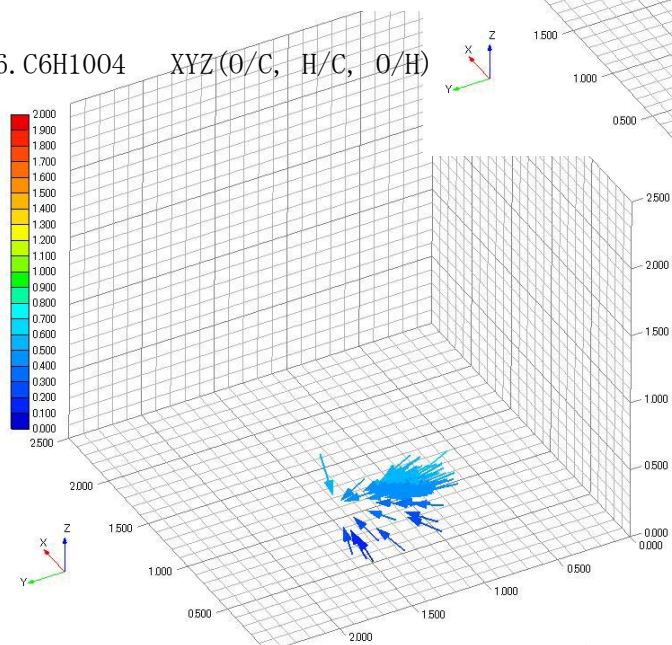
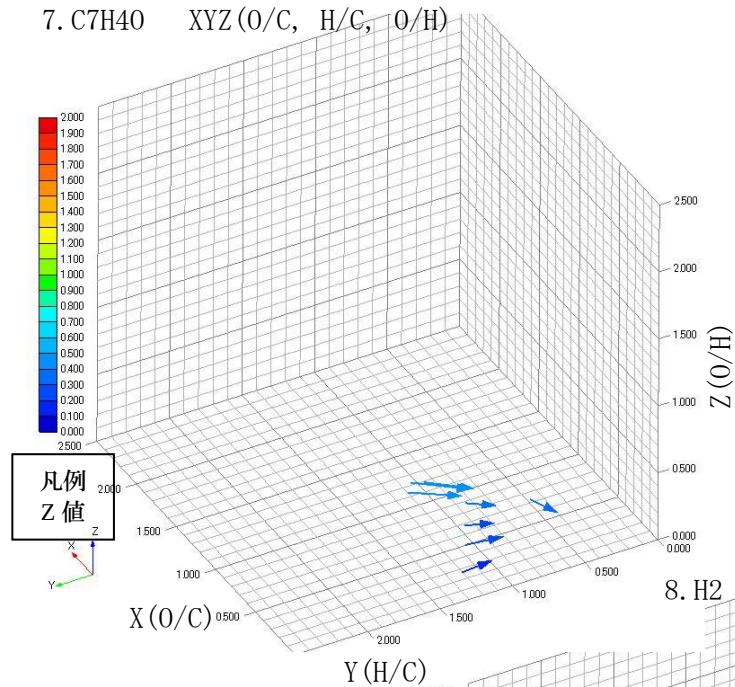
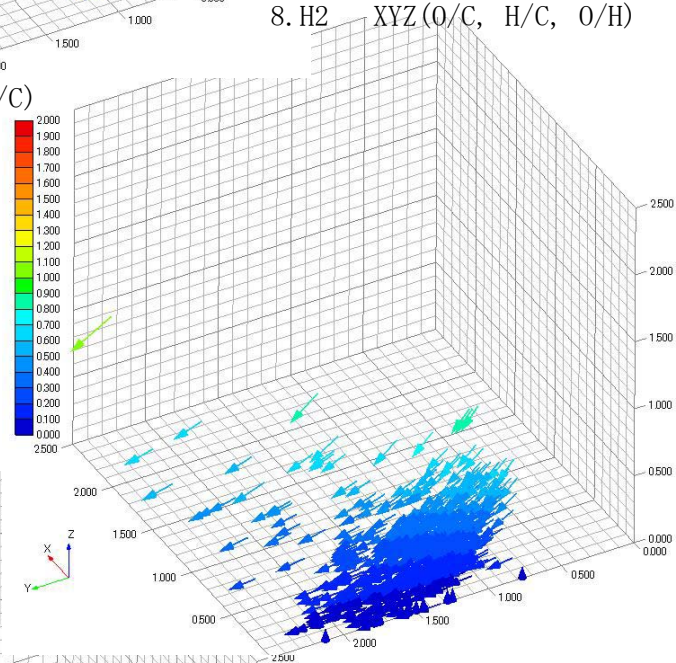


図 4-3-1-2 代謝反応とその分布 2

7. C7H4O XYZ(O/C, H/C, O/H)



8. H2 XYZ(O/C, H/C, O/H)



1-8. XYZ(O/C, H/C, O/H)

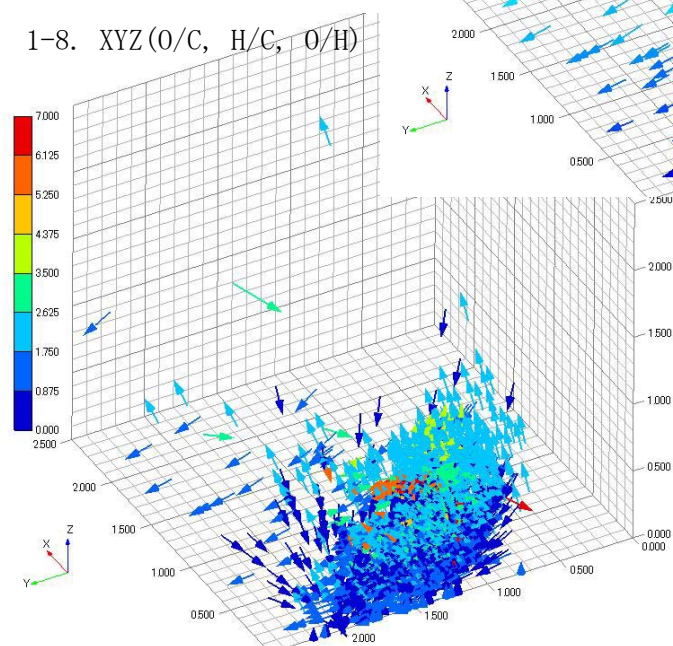
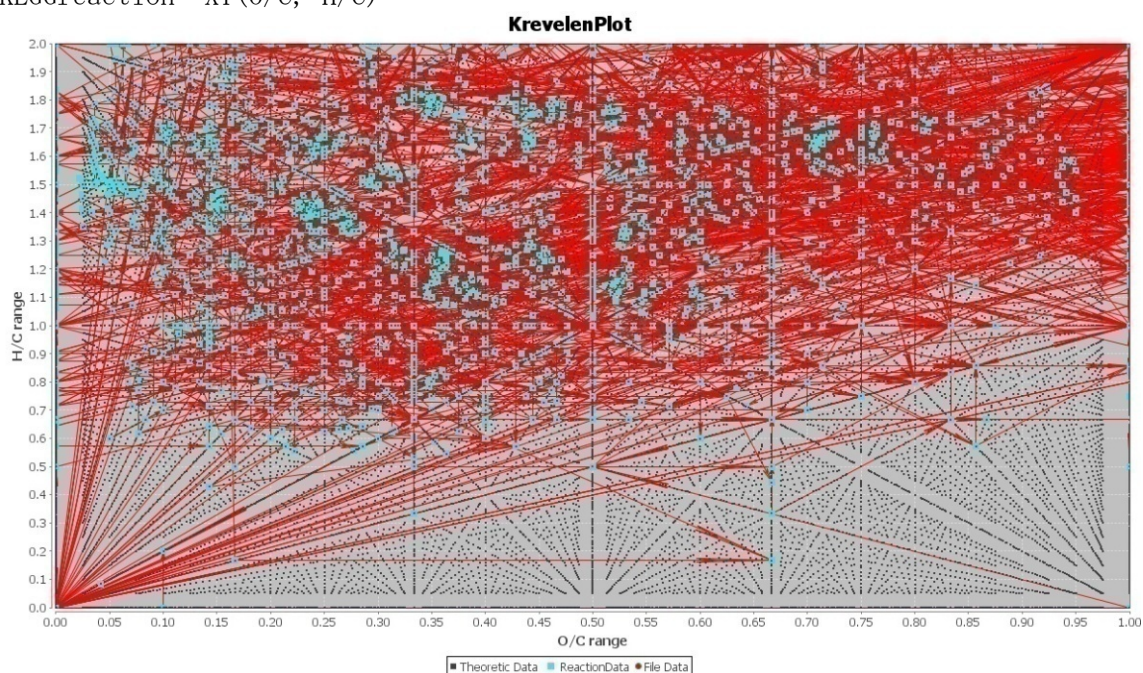


図 4-3-1-3 代謝反応とその分布 3

質量の多いものを反応前物質、少ないものを反応後物質とした。いずれの化学反応も、その変化する化学種の組成比の座標に向かって反応経路が描ける。また、の化学組成によって描かれる座標領域やベクトルの向きが異なる。Z 軸を O/H で電荷の増減を表せるようにしたため、反応によって Z 値がプラスになれば酸化、マイナスになれば還元反応であると示すことができる。Y 座標値から構造とベクトルの向きにより電荷を把握できるため、代謝反応を整理する上でも 3 次元化によってより詳細な特徴が記述できると考える。多環構造の代謝物質の場合、立体障害により起きる反応が限定されると考えられる。また、電荷が増えると求核攻撃を受けやすいなど、その代謝物質に対して起きる可能性のある反応の特徴を捉えやすくなると考えられる。

代謝反応ペア KEGG reaction の反応グループを KrevelenPlot で記述すると情報が重複して確認が難しくなる。3 次元に拡張した KrevelenPlot で記述することで、より明解に反応の特徴を表すことができる。

KEGGreaction XY(O/C, H/C)



KEGGreaction XYZ(O/C, H/C, O/H)

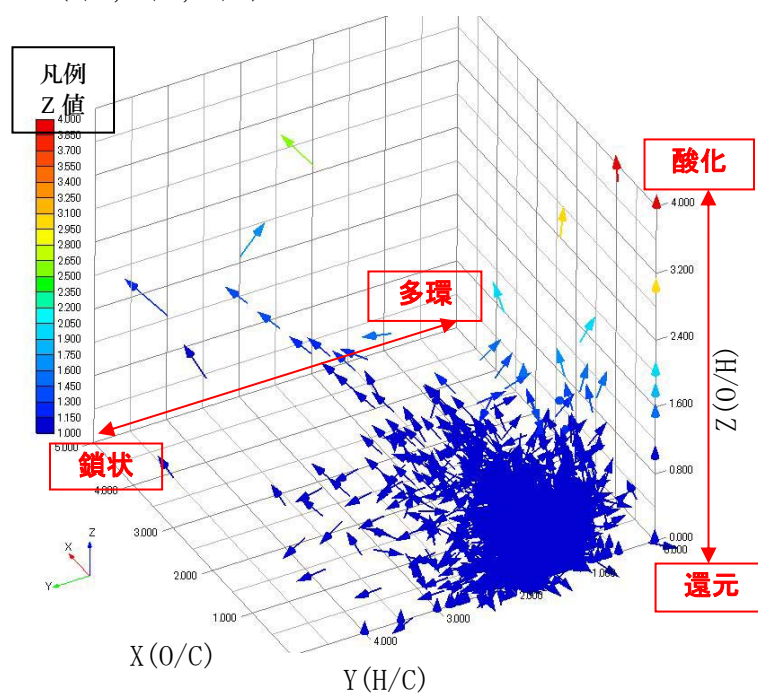


図 4-3-1-4 代謝反応の表現比較

4.3.2 酵素反応の整理

EC1. 1. 1 の酸化還元酵素の反応前後の化学物質のうち主要なものをペアとして抽出し、KrevelenPlot 上に描いた結果を図 4-3-2-1 に示す。

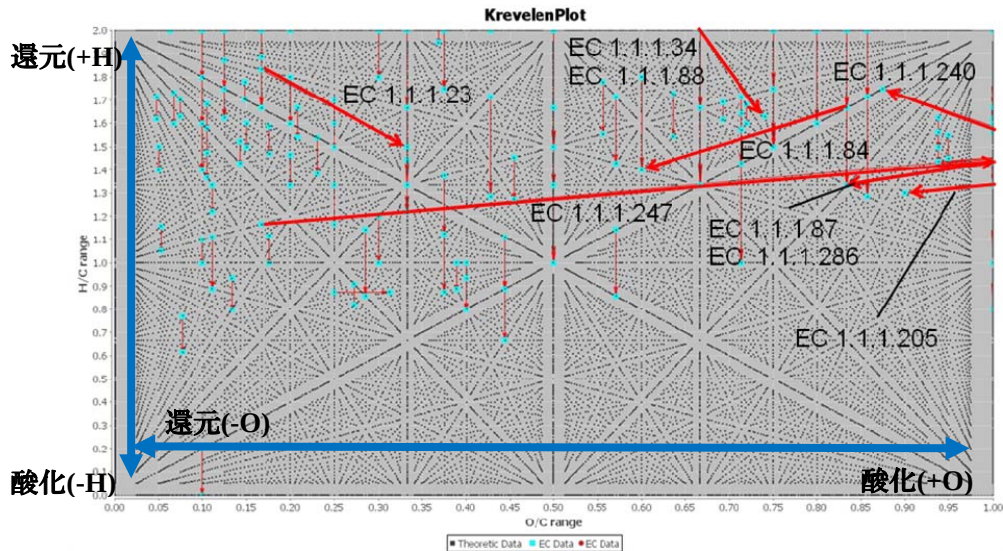


図 4-3-2-1 EC.1.1.1 グループの酵素反応

EC1. 1. 1 は酸化還元酵素であるため、反応経路は X 軸に対して垂直方向か水平方向に現れると予測されたが、実際には傾いたベクトルを示す酵素反応があった。整理した結果の一部を表 4-3-2-1 に示す。

	EC	反応（下線付きが反応ペア）	反応前	反応後	反応
1	1. 1. 1. 23	<u>L-histidinol</u> + 2 NAD ⁺ = <u>L-histidine</u> + 2 NADH + 2 H ⁺	C6H11N3O	C6H9N3O2	付加 (H ₂ O)
2	1. 1. 1. 34	<u>(R)-mevalonate</u> + CoA + 2 NADP ⁺ = <u>(S)-3-hydroxy-3-methylglutaryl-CoA</u> + 2 NADPH + 2 H ⁺	C6H12O4	C27H44N7O20 P3S	付加 (CoA)
2	1. 1. 1. 88	<u>(R)-mevalonate</u> + CoA + 2 NAD ⁺ = <u>3-hydroxy-3-methylglutaryl-CoA</u> + 2 NADH + 2 H ⁺	C6H12O4	C27H44N7O20 P3S	付加 (CoA)
3	1. 1. 1. 87	<u>(1R, 2S)-1-hydroxybutane-1, 2, 4-tricarbo</u>	C7H10O7	C6H8O5	脱離

		<u>xylate</u> + NAD ⁺ = <u>2-oxoadipate</u> + CO ₂ + NADH + H ⁺			(CH ₂ O ₂)
3	1. 1. 1. 28 6	<u>isocitrate</u> + NAD ⁺ = <u>2-oxoglutarate</u> + CO ₂ + NADH	C6H807	C5H405	脱離 (CH ₄ O ₂)
4	1. 1. 1. 84	<u>(R)-3,3-dimethylmalate</u> + NAD ⁺ = <u>3-methyl-2-oxobutanoate</u> + CO ₂ + NADH	C6H1005	C5H703	脱離 (CH ₃ O ₂)
5	1. 1. 1. 20 5	<u>inosine 5'-phosphate</u> + NAD ⁺ + H ₂ O = <u>xanthosine 5'-phosphate</u> + NADH + H ⁺	C10H15N4O14 P3	C10H13N4O9P	脱離 (H ₂ O ₅ P ₂)
6	1. 1. 1. 24 0	<u>N-acetyl-D-glucosamine</u> + NAD ⁺ = <u>N-acetyl-D-glucosamine</u> + NADH + H ⁺	C17H27N3O17 P2	C8H14N07	脱離 (C ₉ H ₁₃ N 2O ₁₀ P ₂)
7	1. 1. 1. 24 7	<u>codeine</u> + NADP ⁺ = <u>codeinone</u> + NADPH + H ⁺	C18H21N03	C17H25N3O18 P2	付加 (H ₄ N ₂ O ₁ 5P ₂) 脱離 (C)

1. 1. 1 から 1. 1. 1. 297 までの 290 の酸化還元酵素のうち 26 の酵素が付加、脱離反応を伴っていることがわかった。酵素の分類において、違う働きがあるものが混じっていることが視覚的に確認できるため、酵素の分類整理にも役立つと言える。

4.4 既知の代謝反応経路による検証

KEGG pathway には既知の代謝反応経路がまとめられている。そこで、この反応経路を KrefelenPlot 上で表現してみた。KEGG pathway から取得したフラボノイドの生合成経路 (map00941) を図 4-3-3-1 に、この生合成経路から反応ペアを抽出したものを表 4-3-3-1 に示す。また、その反応ペアを KrefelenPlot 上で表したものを図 4-3-3-2 に示す。

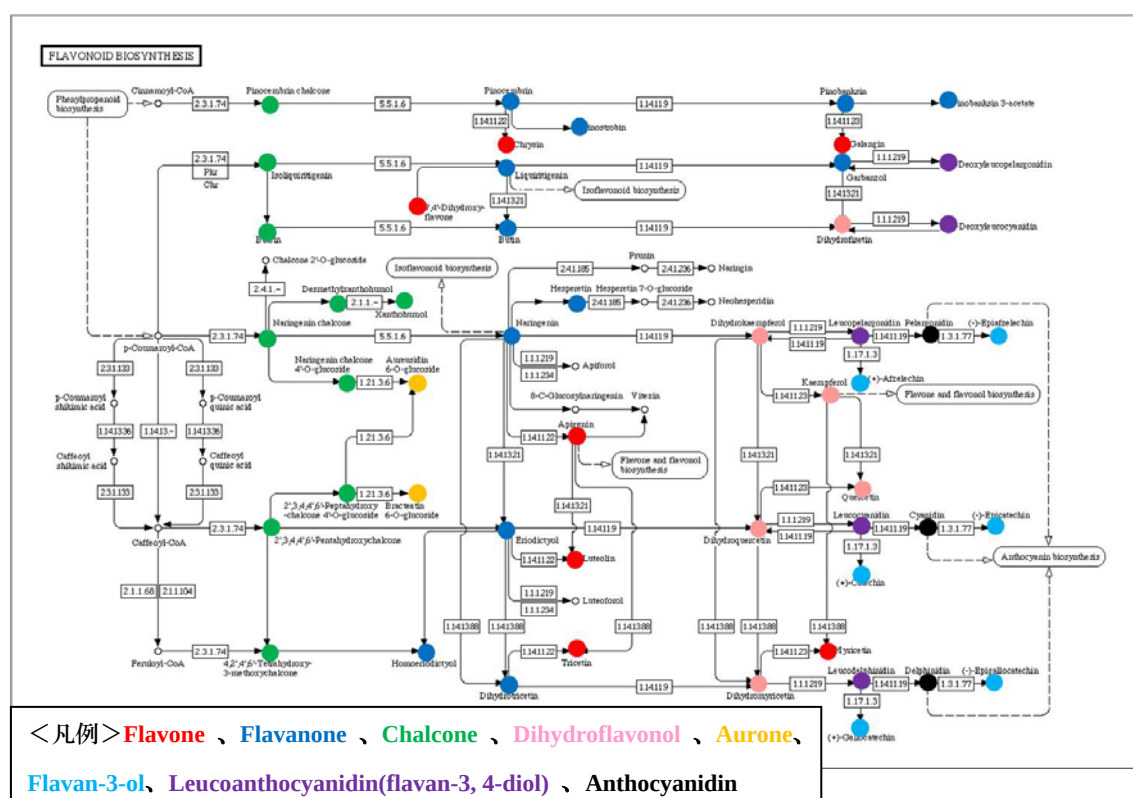


図 4-3-3-1 フラボノイド分類

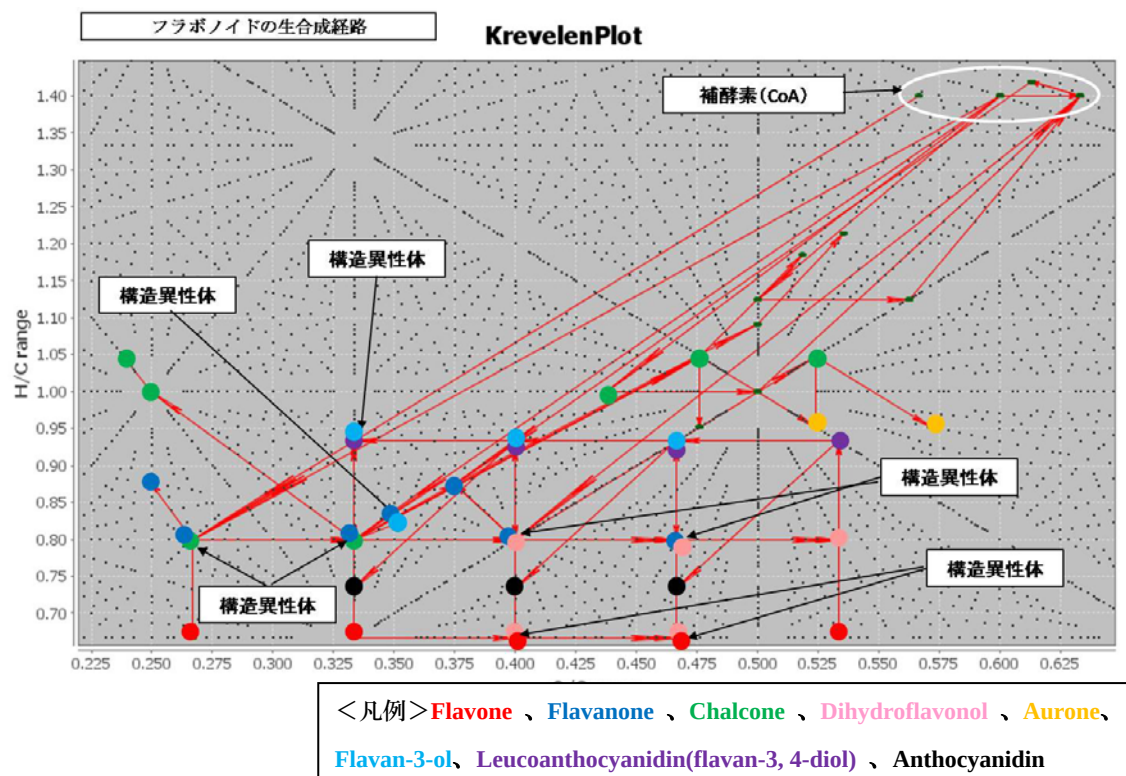


図 4-3-3-2 フラボノイドの生合成経路 (KrevelenPlot)

KrevelenPlot で生合成経路を描いた結果、フラボノイドの各分類は酸化還元反応によって生成されることが視覚的に捉えられる。また拡張 KrevelenPlot により官能基の種類と合わせた代謝反応の捉え方もできる。図 4-3-3-3 に示す。

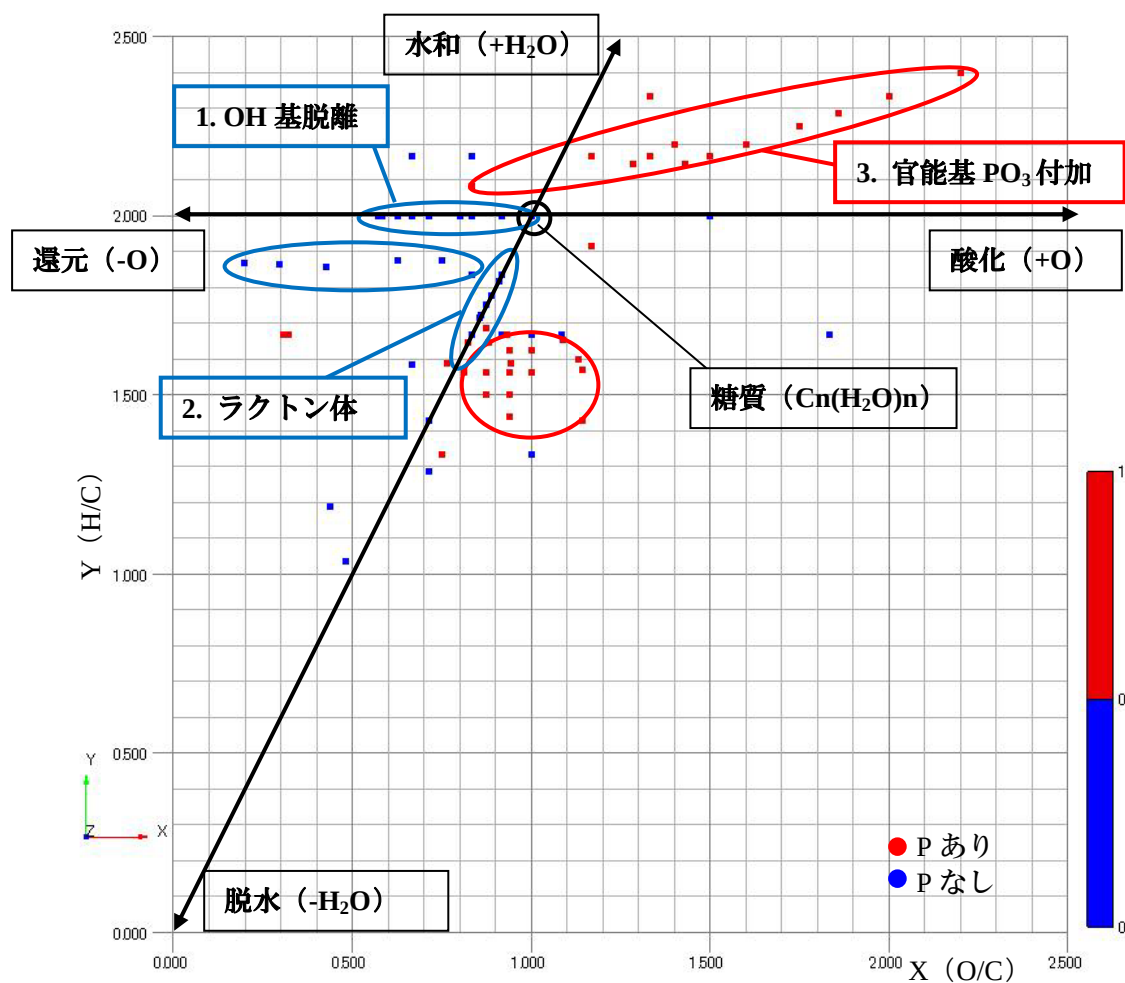


図 4-3-3-3 代謝反応と官能基による識別

図 4-3-3-3 に示したのは糖質である。代謝反応のベクトルと併せてみることで代謝反応の種類が、ヘテロ元素の軸（例では凡例色とした）を用いることでその官能基の結合した構造が判別できる。図では赤色の代謝物質はリン脂質を表している。このように KrevelenPlot の軸の設定を拡張することで、様々な視点から悉皆的に代謝物質を解析できる。

第5章 結論：今後の研究課題

本研究では、化学組成式が判明した代謝物質のデータを用い天然物全体が有する特徴を拡張した KrevelenPlot により捉えられることを示した。従来の KrevelenPlot でも代謝物質の種類やおよその構造、代謝反応経路を表現できるが、二次元であるために表現限界がある。軸の定義や組合せによっては、軸の定義に別の元素を用いても同様の法則性に則って座標値が現われた。また、天然物自体が組成比率で見ると非常に狭い座標範囲に集約されることがわかった。

拡張した KrevelenPlot を用いると天然物の構造的な特徴を俯瞰することができる。またこれを用いて天然物と合成物の化合物空間における分布を比較すれば、利用されていない天然物を把握出来る。あるいは、近年関心の高まりつつある ECO 分野に関しても素材を拡張 KrevelenPlot 上で記述し、天然物グループとの分布の比較により、自然界で代謝可能な生分解性素材の開発にも応用できる可能性がある。また、天然物の代謝経路の整理に拡張 KrevelenPlot を用いることで、まだ解明されていない代謝経路の推定と代謝物質の種類を関連付けて捉えられる。

天然物が拡張 KrevelenPlot 上で法則性をもった曲面で現れること、また座標値のパターンが限定的であること、さらに質量と構造の関係が座標値に現れることを利用することで、高分解能質量分析によって得られる精密分子量からの理論化学式の計算に応用できる。現状では元素の結合を行列式によって記述し、全ての可能な組合せを計算して理論化学式を出す方法が採られている。しかし、高分子になるほど計算時間が膨大になり計算結果も増えるため、1つの正解値に対する候補が十分絞り込めないという問題がある。本研究結果を基にし可能な理論化学式の大幅な絞り込みを行うことで、化学構造の高精度な推定を分子量の比較的大きな分子にまで適用可能とすることが見込まれる。

謝辞

本研究を遂行するにあたり終始にわたりご指導を賜りました、情報科学研究科 金谷重彦教授、並びに Md. Altaf-UL-Amin 准教授、中村建介特任准教授に心から感謝いたします。

情報科学研究科 湊小太郎教授、川端猛准教授に頂きました貴重なご意見を通して研究内容を掘り下げることができました。ここに深く感謝いたします。

本研究室の高橋弘喜氏には質量分析における注意点を含め、本研究に必要な諸知識を授けていただきました。ここに深く感謝いたします。

最後になりましたが、本研究室の皆様方には、数多くのご指導、ご助言をいただきましたことを、心より深く感謝いたします。

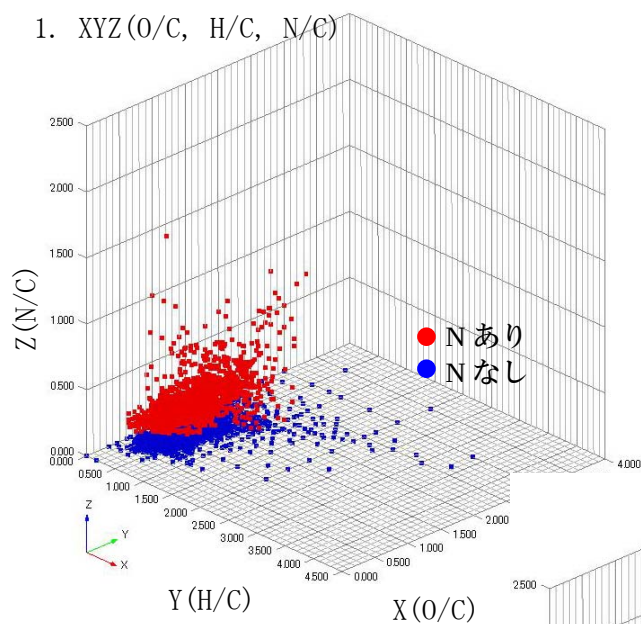
参考文献

1. Alain Reinhardt, Christian Emmenegger, Bertran Gerrits, Christian Panse, Josef Dommen, Urs Baltensperger, Renato Zenobi, Markus Kalberer, *Anal.Chem* 2007, 79(11):4074-4082
2. Miklos Feher* and Jonathan M. Schmidt, American Chemical Society, 2003, 43(1):218-227
3. N. Hertkorn, C. Ruecker, M. Meringer, R. Gugisch, M. Frommberger, E. M. Perdue, M. Witt, and P. Schmitt-Kopplin, *Anal Bioanal Chem* 2007, 389 (5):1311-1327
4. S. Kim, R.W.Kramer, P.G.Hatcher, *Anal.Chem* 2003, 75(20):5336-5344
5. Erwan Wernera, Jean-Francois Heiliera, b, Celine Ducruixa, 1, Eric Ezana, Christophe Junota, *Proc Journal of Chromatography B* 2008, 871 (2):143-163.
6. K.P.C.Volhardt, N.E.Schore, ボルハルト・ショアー現代有機化学 第4版 上下巻, 化学同人, 2004
7. 田中治, 野副重男, 相見則郎, 永井正博, 天然物化学 改訂第6版, 南江堂, 2002
8. C.G.Wermuth, 最新創薬化学 改訂第2版 上巻, テクノミック, 2004
9. 福崎英一郎, メタボロミクスの先端技術と応用, シーエムシー出版, 2008
10. 斉藤和季, メタボローム研究の最前線, シュプリンガー・フェアラーク東京 2003, p.131
11. 樹林千尋, 秋葉光雄, 新版ライフサイエンスの有機化学, 三共出版, 2004
12. John McMurry, Tadhg Begley, マクマリー生化学反応機構, 東京化学同人, 2007
13. 創薬化学, 丸善, 2000
14. 左右田健次, 生化学ー基礎と光学ー, 化学同人, 2001
15. J.Gasteiger, T.Engel, ケモインフォマティックス, 丸善, 2005
16. JFreeChart (グラフ描画Javaフレームワーク) <http://www.jfree.org/index.html>
17. Graph-R (3次元グラフ表現ツール) <http://www.iris.dti.ne.jp/~tohru/>

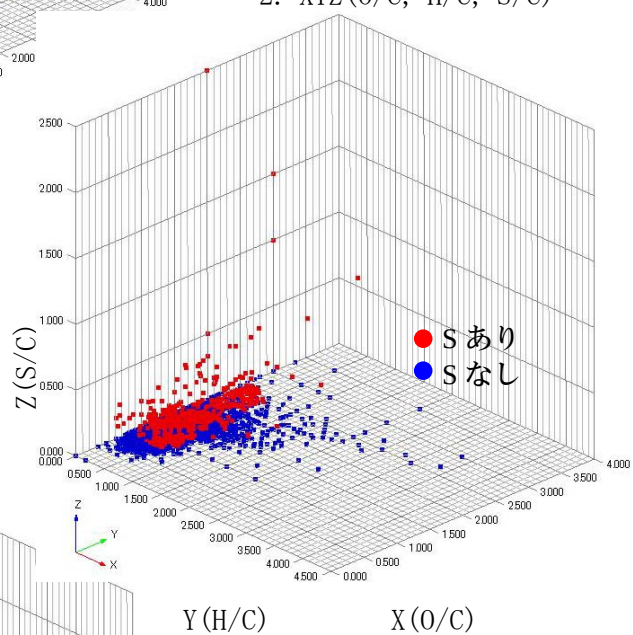
Appendix(1) 軸の組合せによる代謝物質の表現

KNapSAcK

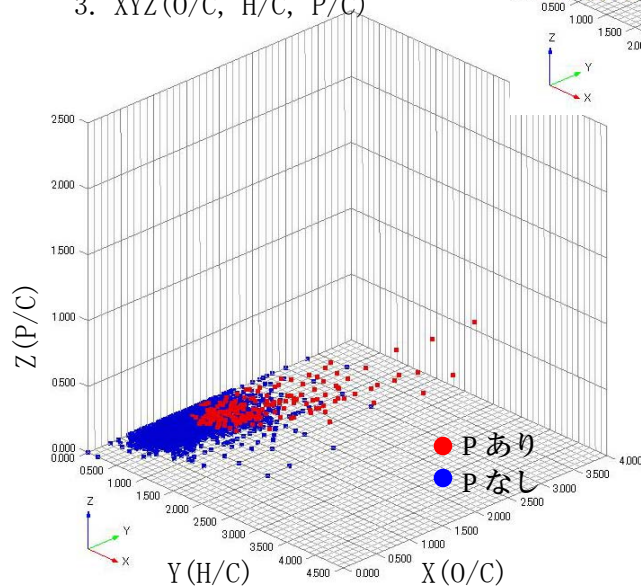
1. XYZ(O/C, H/C, N/C)



2. XYZ(O/C, H/C, S/C)

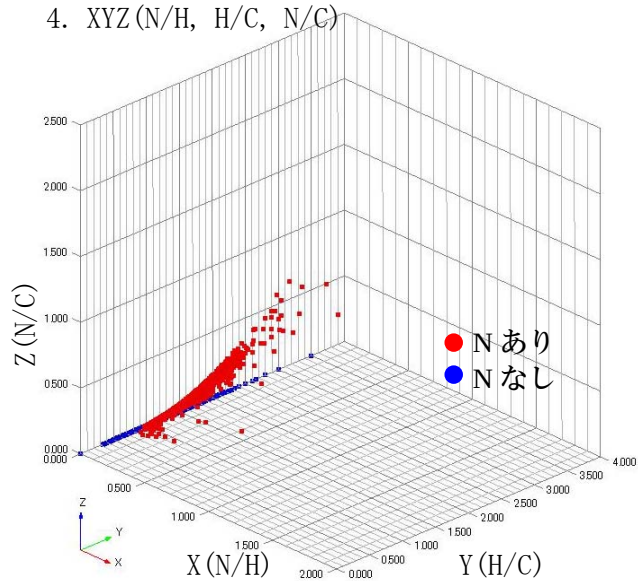


3. XYZ(O/C, H/C, P/C)

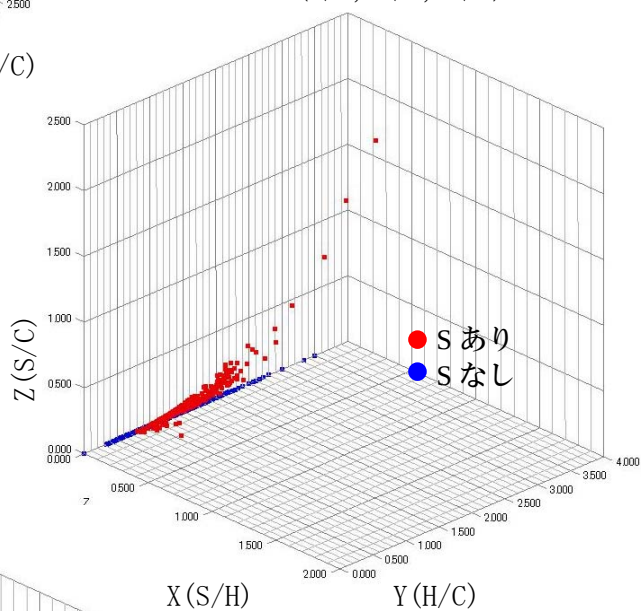


KNApSAcK

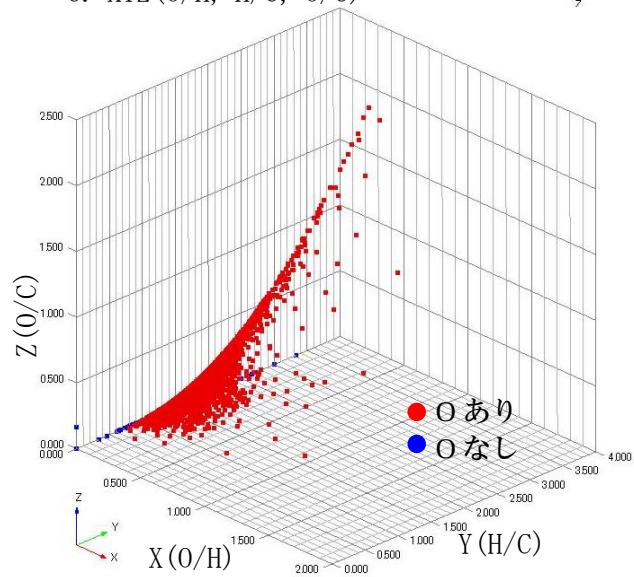
4. XYZ (N/H, H/C, N/C)



5. XYZ (S/H, H/C, S/C)

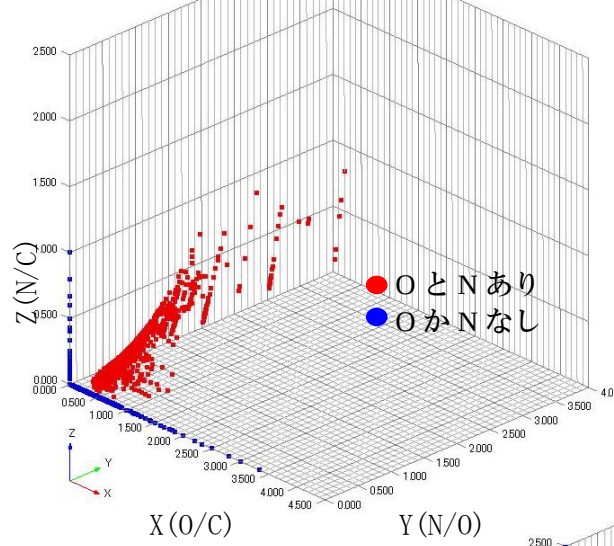


6. XYZ (O/H, H/C, O/C)

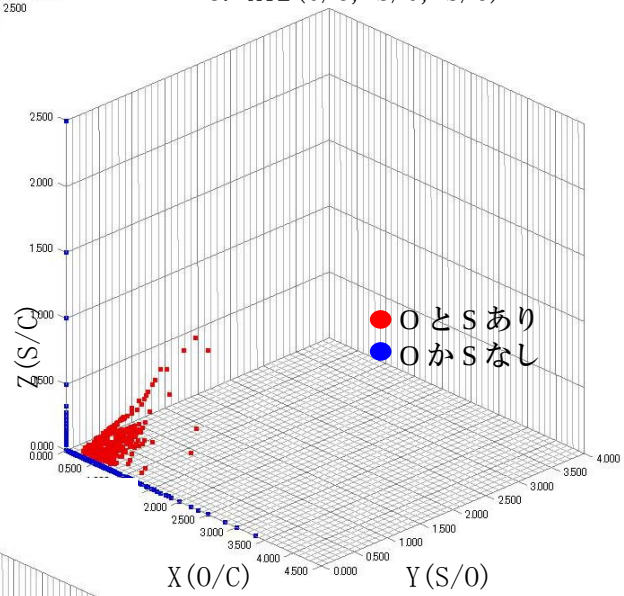


KNapSack

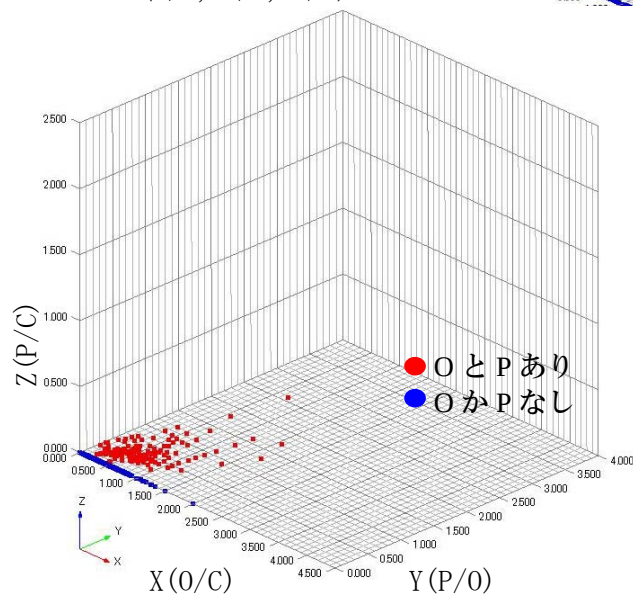
7. XYZ(0/C, N/0, N/C)



8. XYZ(0/C, S/0, S/C)

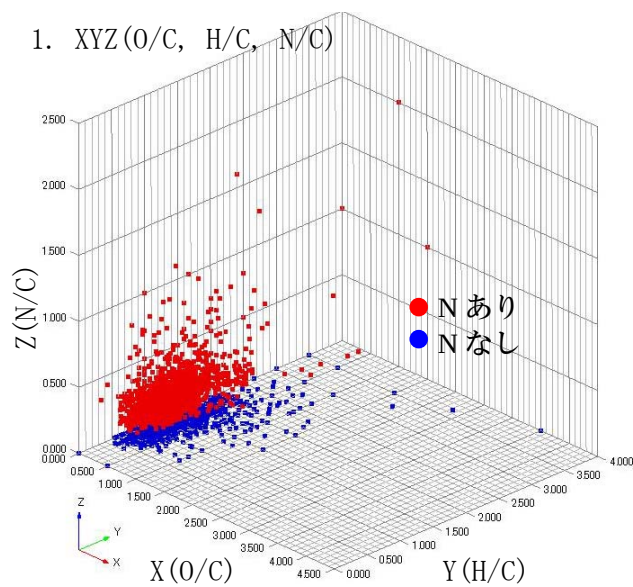


9. XYZ(0/C, P/0, P/C)

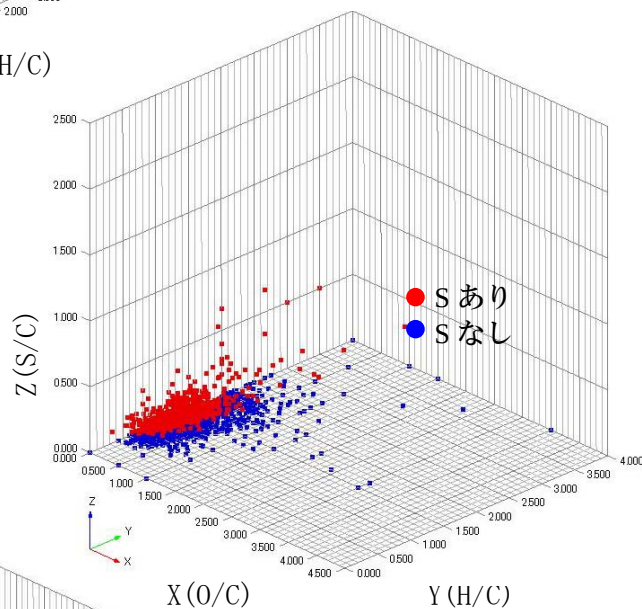


KEGG drug

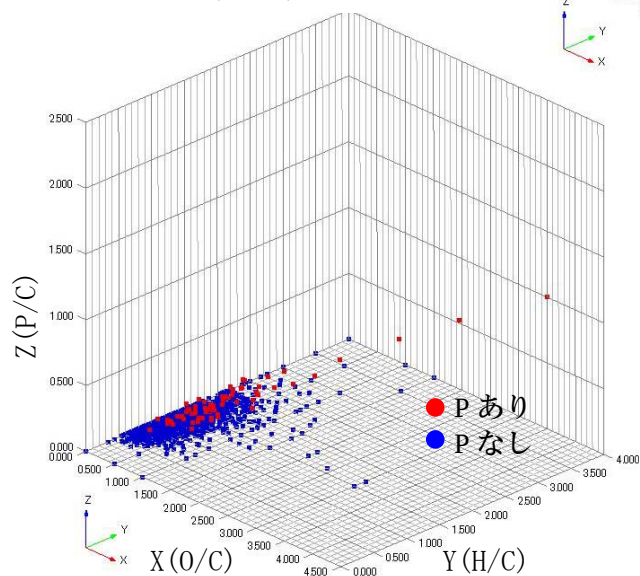
1. XYZ(O/C, H/C, N/C)



2. XYZ(O/C, H/C, S/C)

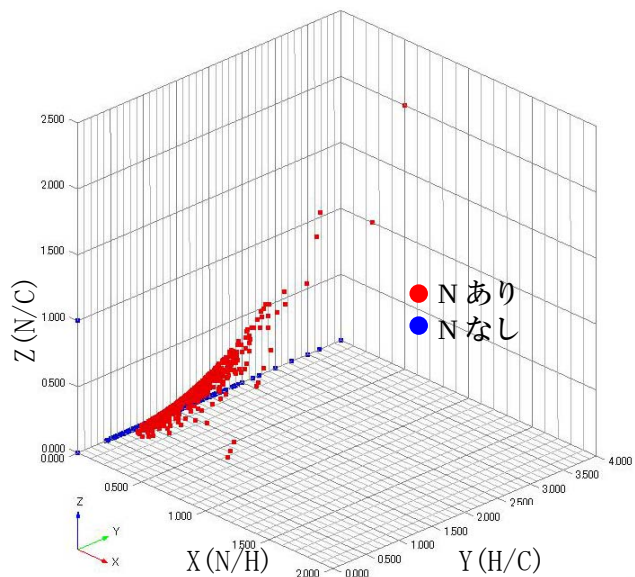


3. XYZ(O/C, H/C, P/C)

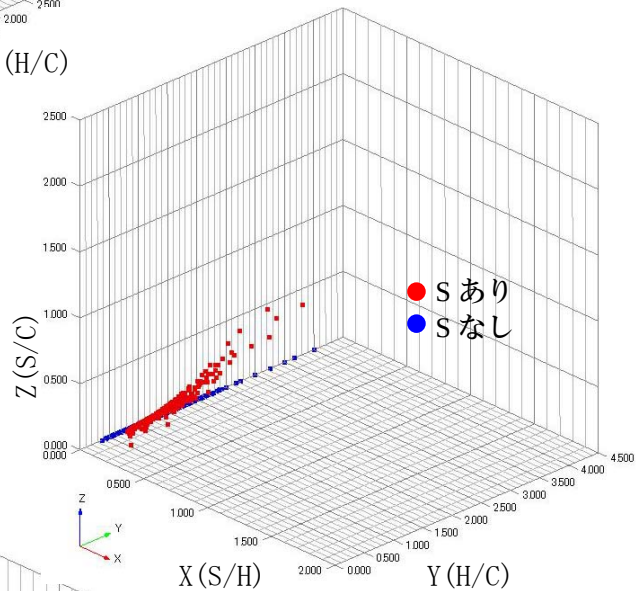


KEGG drug

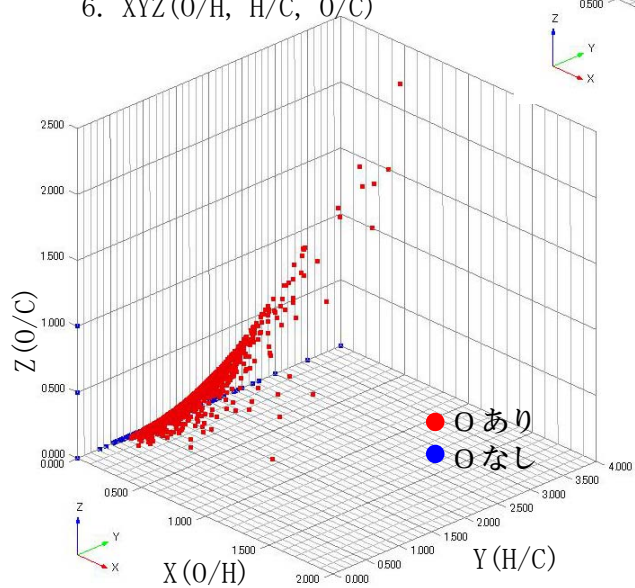
4. XYZ(N/H, H/C, N/C)



5. XYZ(S/H, H/C, S/C)

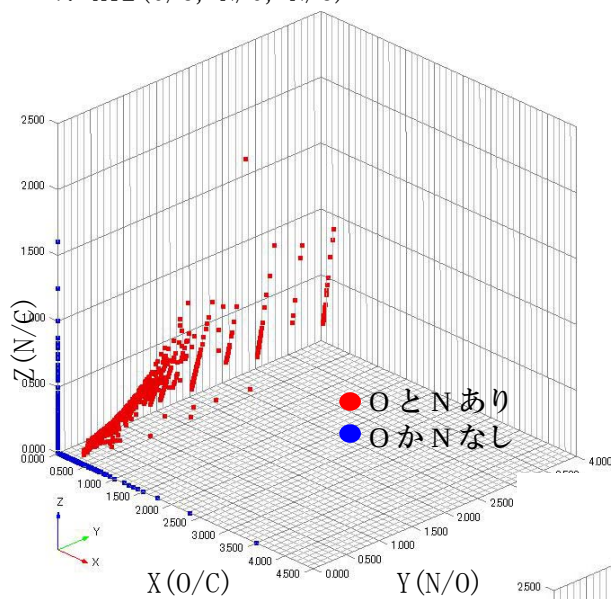


6. XYZ(O/H, H/C, O/C)

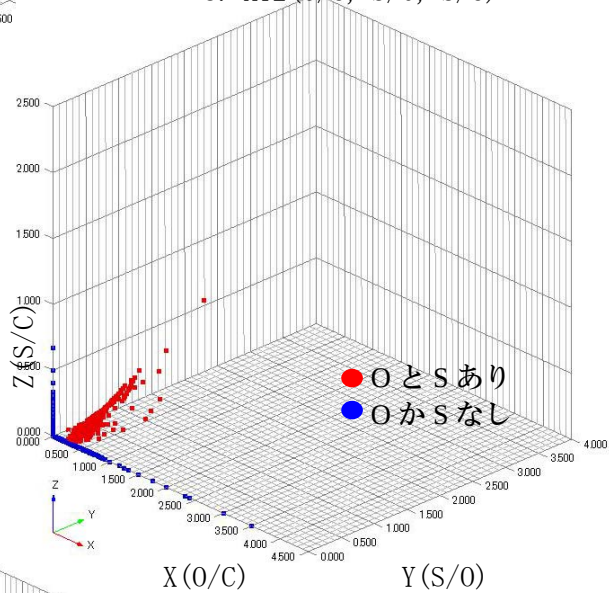


KEGG drug

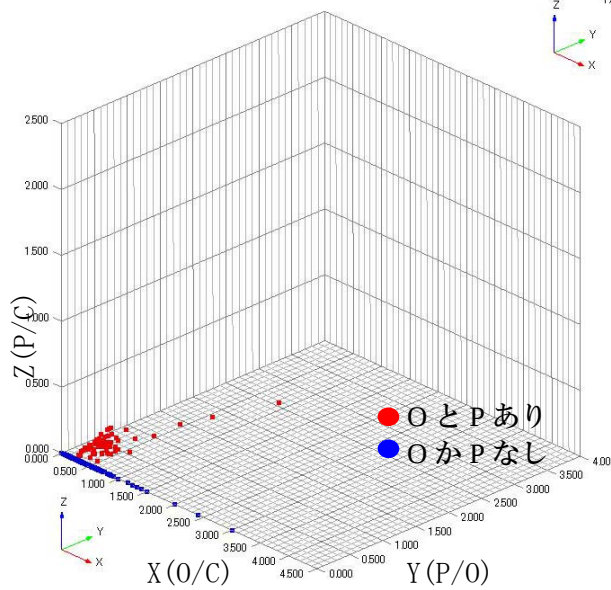
7. XYZ(O/C, N/O, N/C)



8. XYZ(O/C, S/O, S/C)

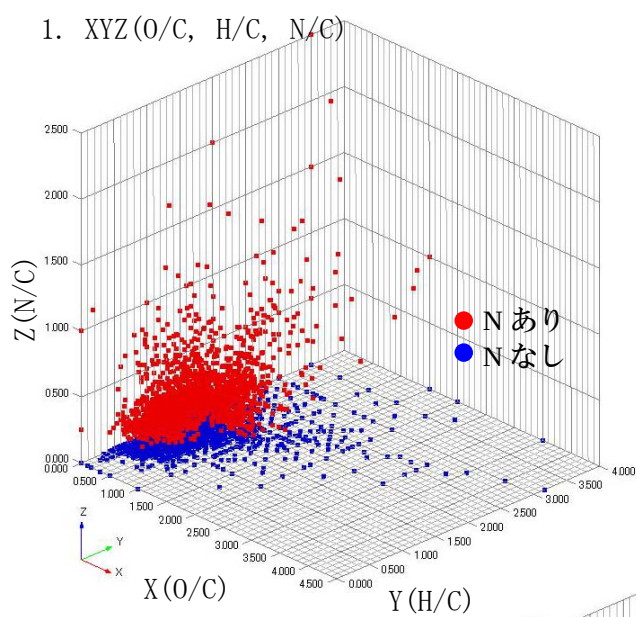


9. XYZ(O/C, P/O, P/C)

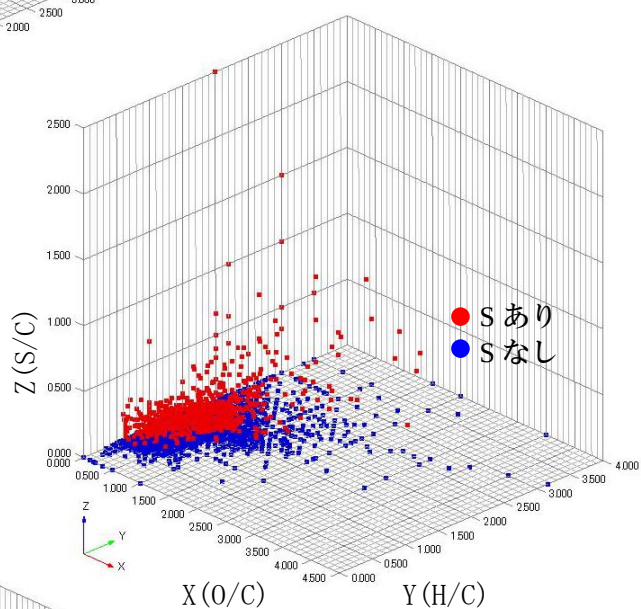


KEGG compound

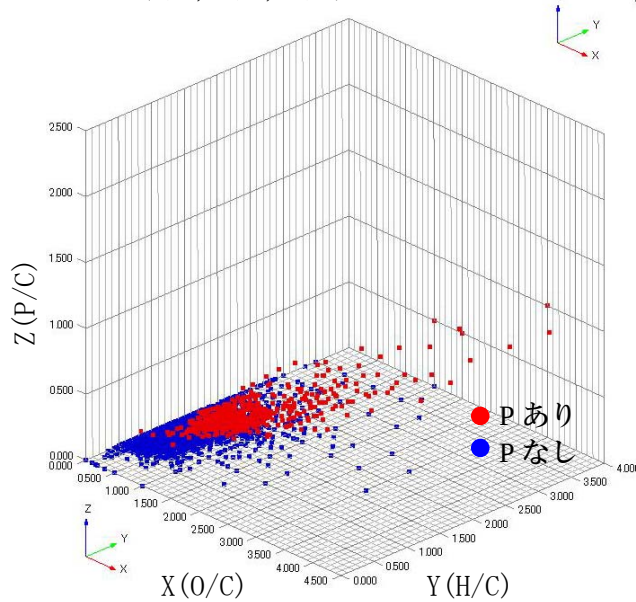
1. XYZ(O/C, H/C, N/C)



2. XYZ(O/C, H/C, S/C)

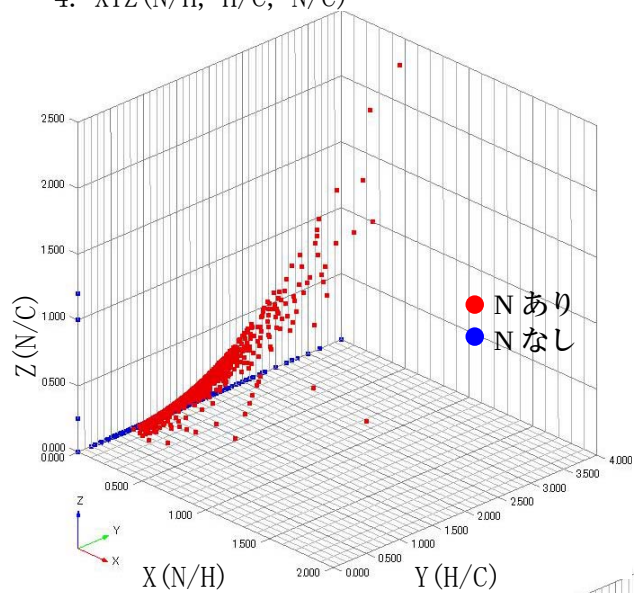


3. XYZ(O/C, H/C, P/C)

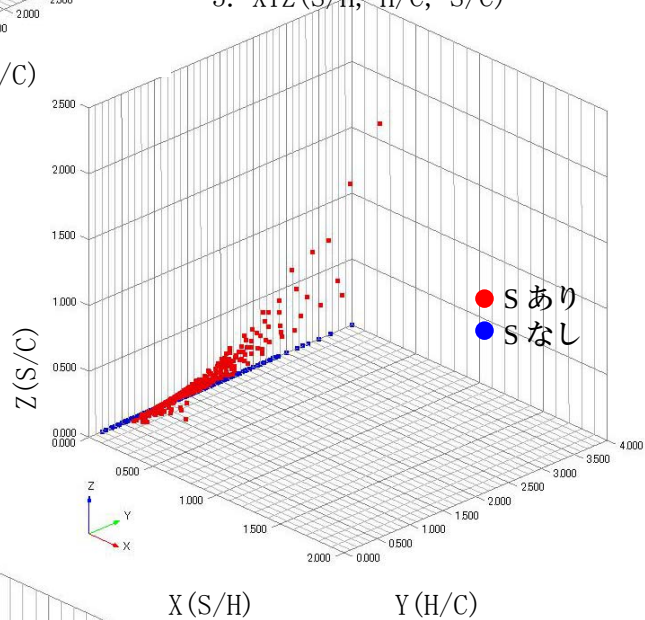


KEGG compound

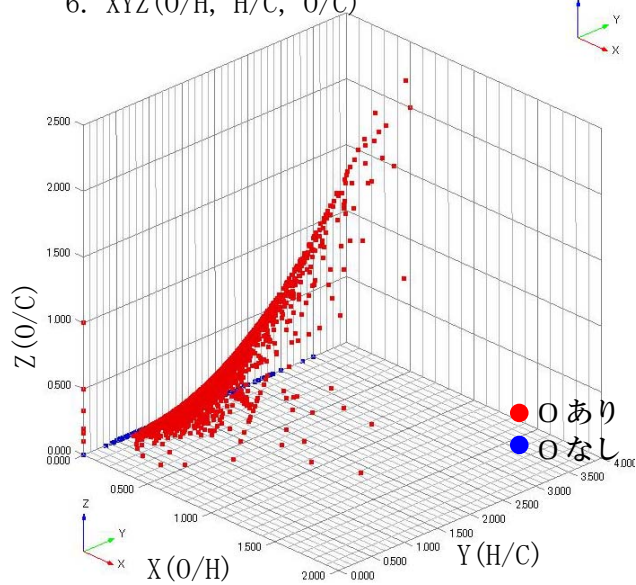
4. XYZ(N/H, H/C, N/C)



5. XYZ(S/H, H/C, S/C)

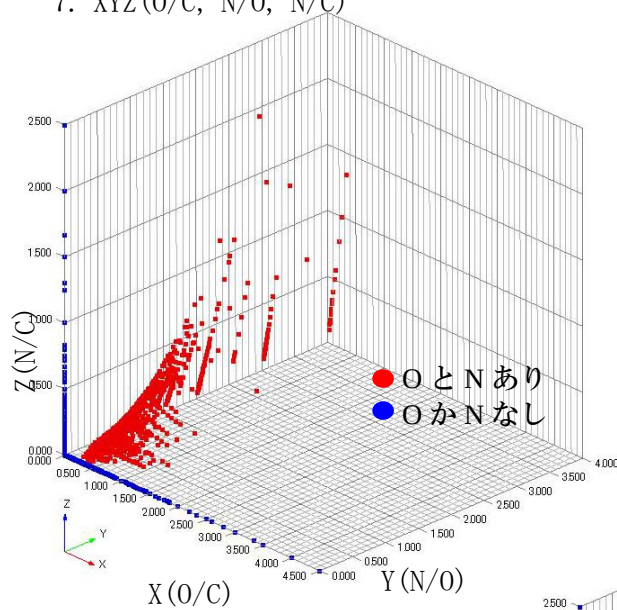


6. XYZ(O/H, H/C, O/C)

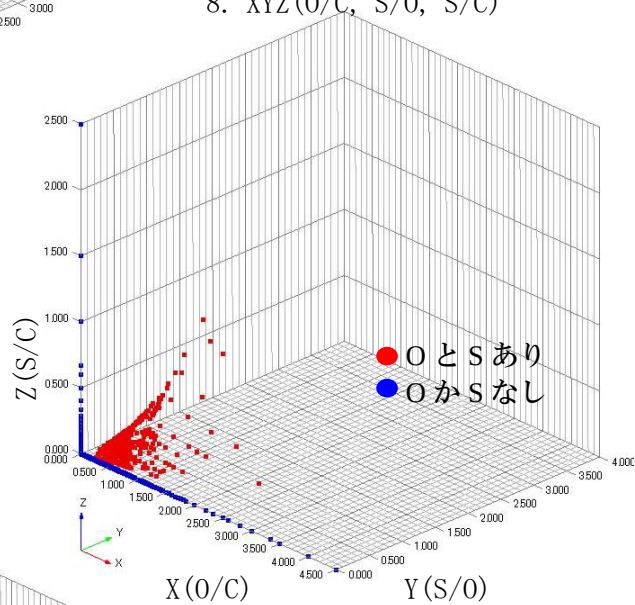


KEGG compound

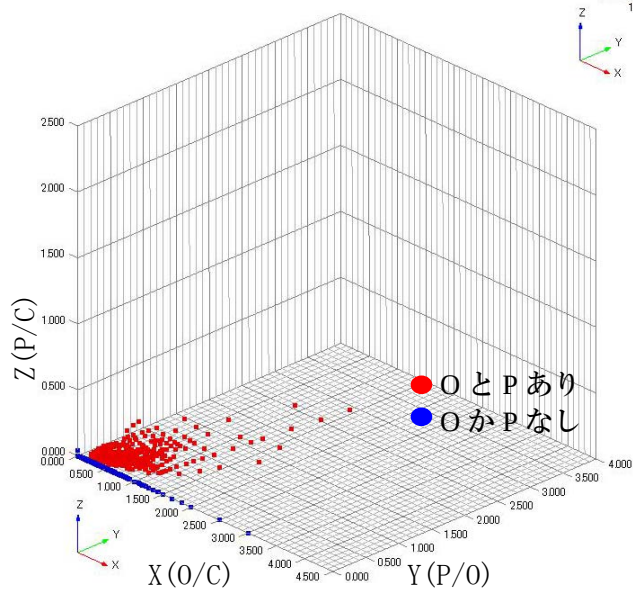
7. XYZ(O/C, N/O, N/C)



8. XYZ(O/C, S/O, S/C)

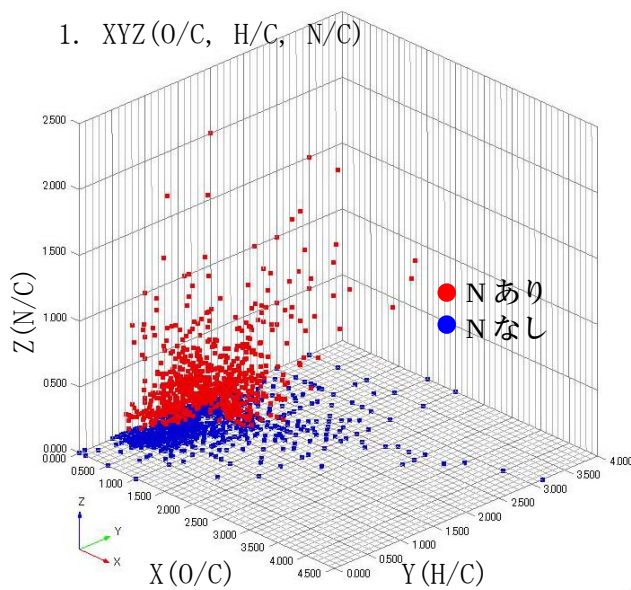


9. XYZ(O/C, P/O, P/C)

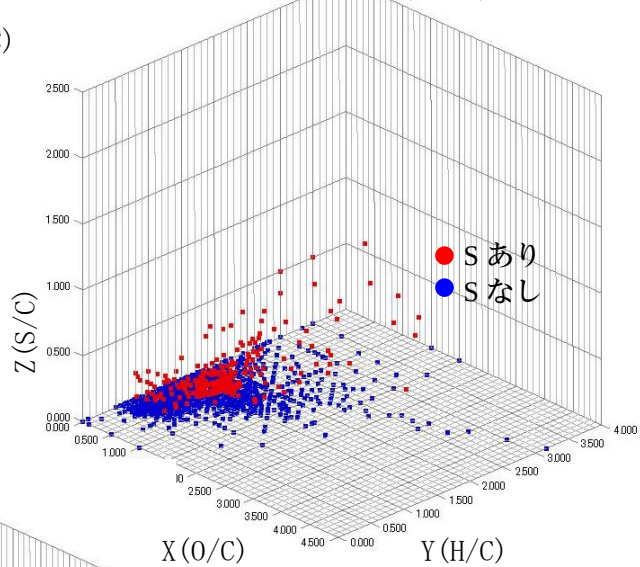


KEGG reaction

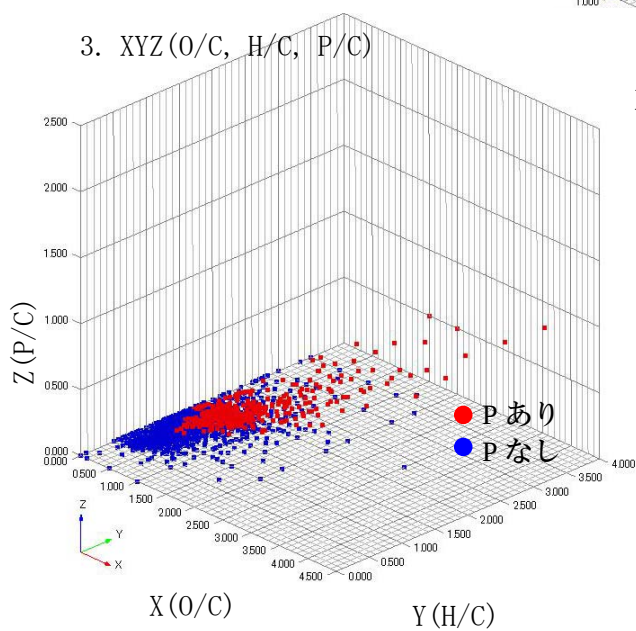
1. XYZ(O/C, H/C, N/C)



2. XYZ(O/C, H/C, S/C)

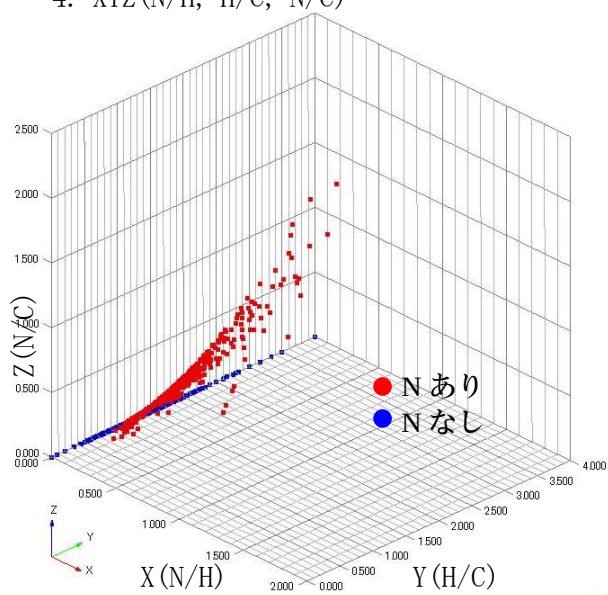


3. XYZ(O/C, H/C, P/C)

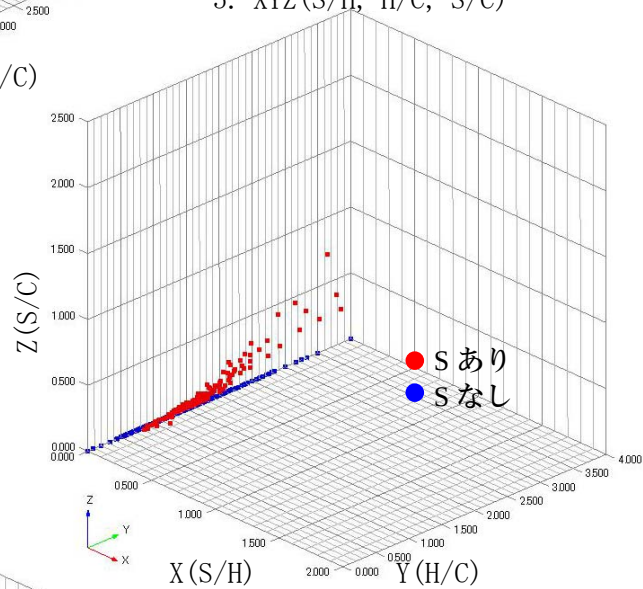


KEGG reaction

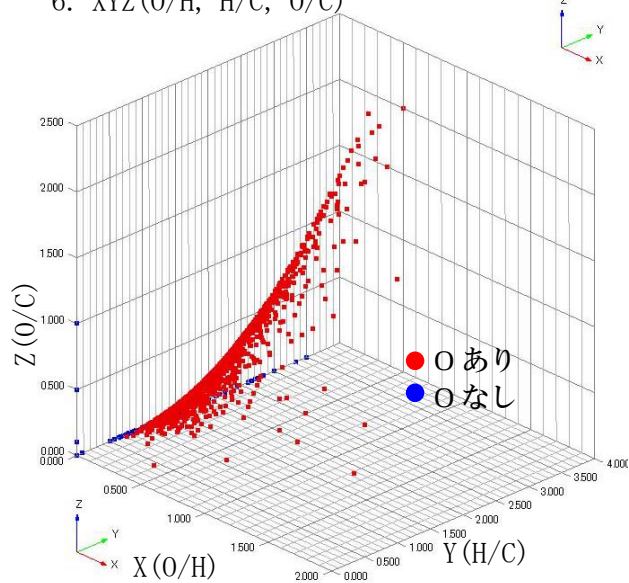
4. XYZ(N/H, H/C, N/C)



5. XYZ(S/H, H/C, S/C)

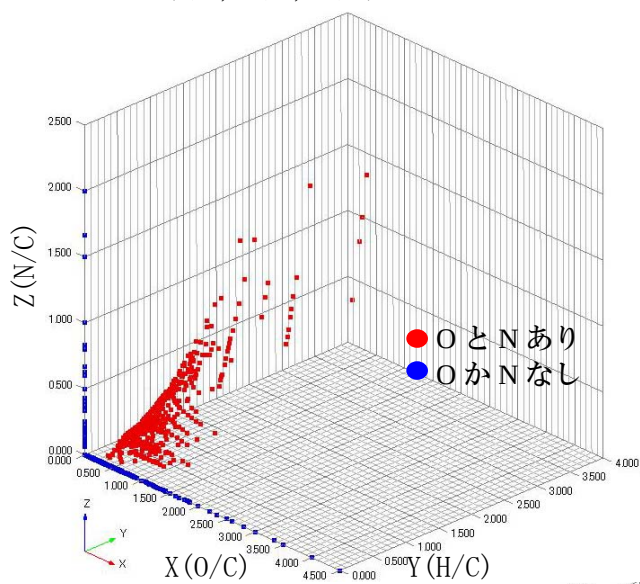


6. XYZ(O/H, H/C, O/C)

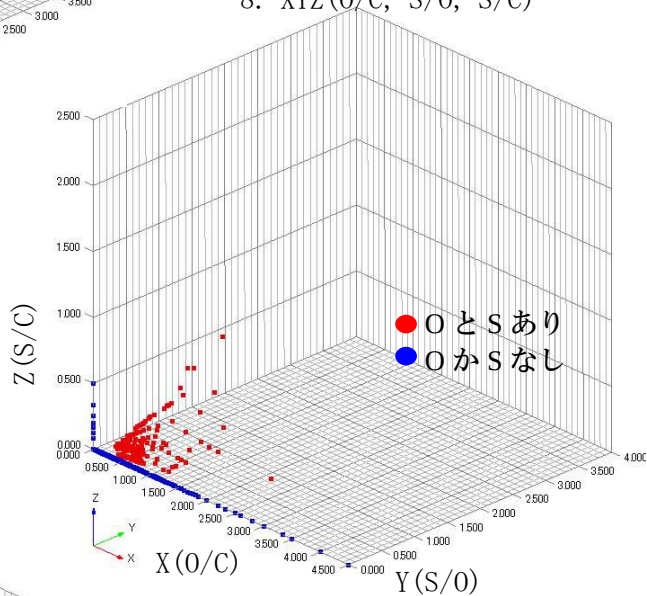


KEGG reaction

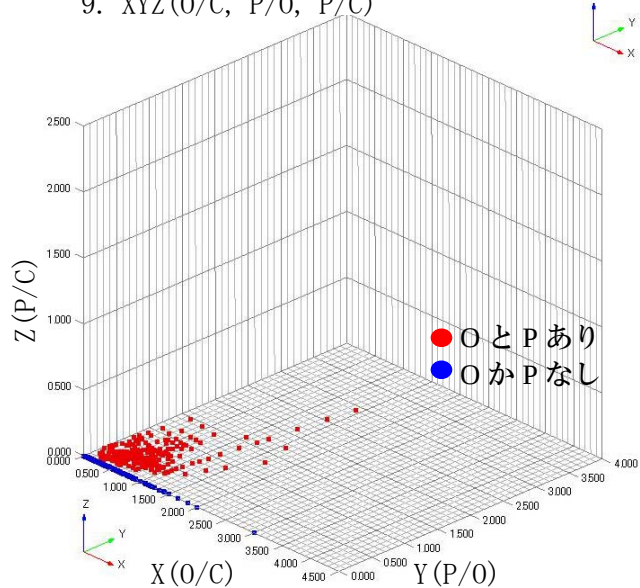
7. XYZ(O/C, N/O, N/C)



8. XYZ(O/C, S/O, S/C)

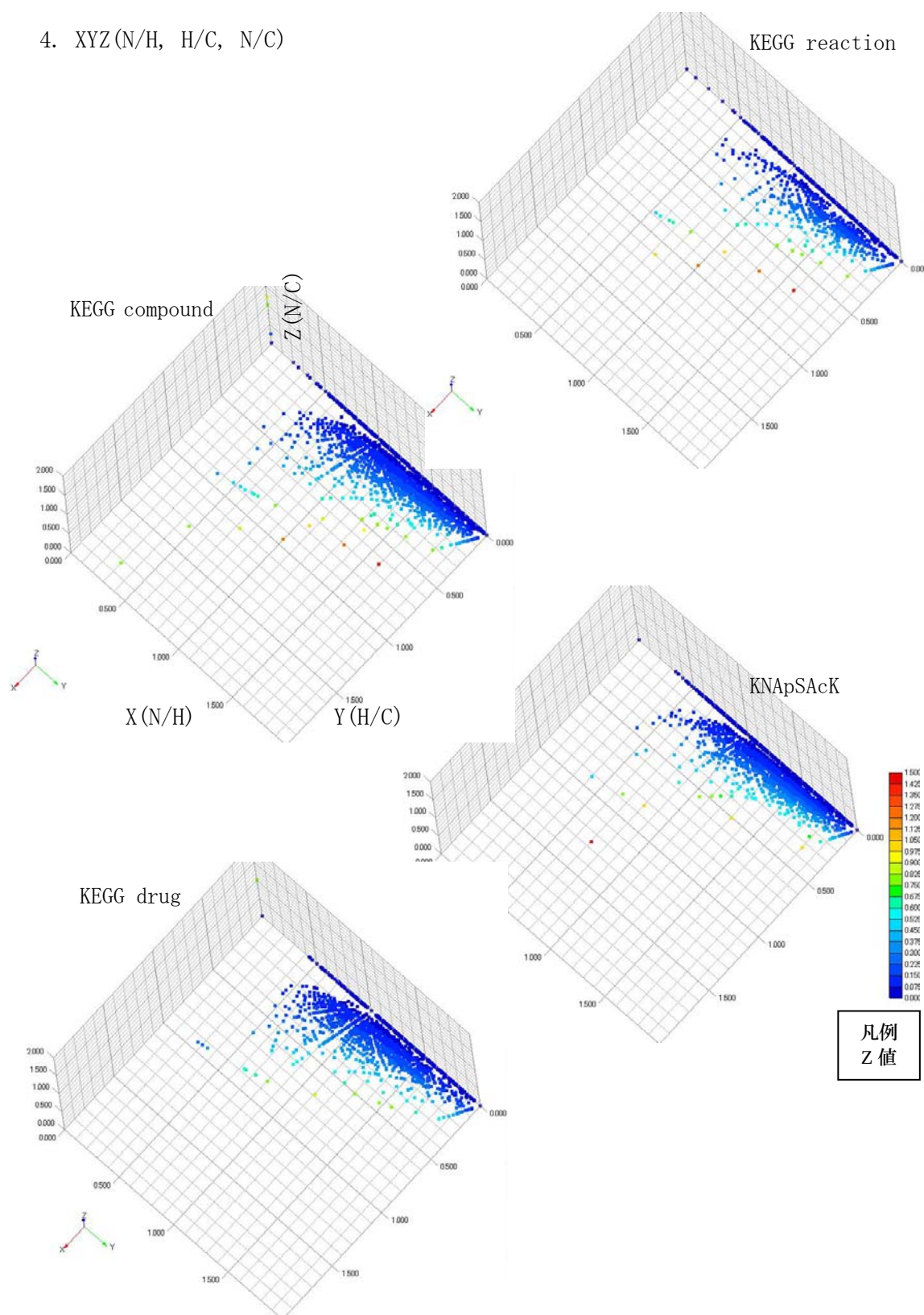


9. XYZ(O/C, P/O, P/C)

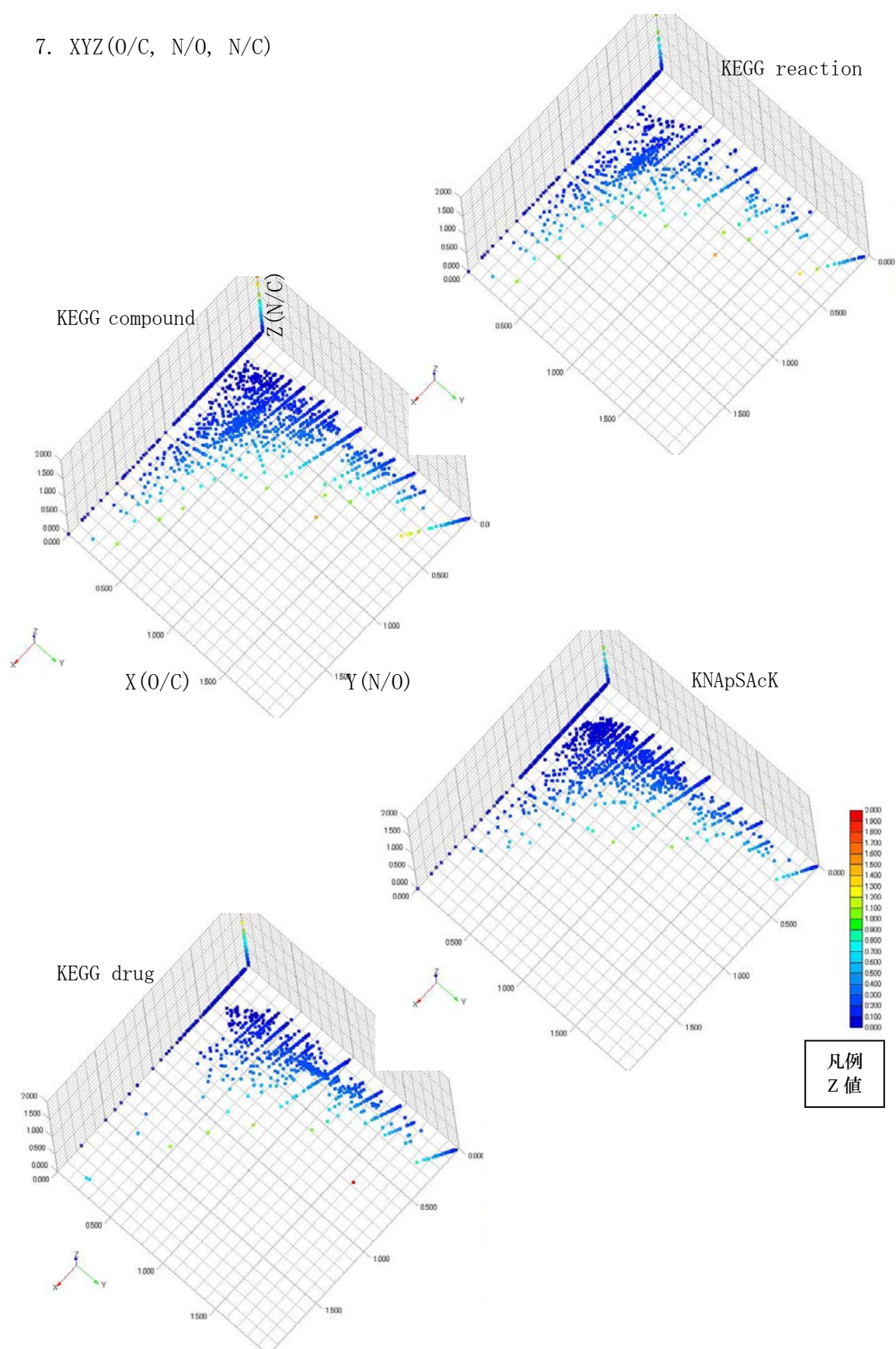


Appendix(2) 構造から見る天然物の全体像

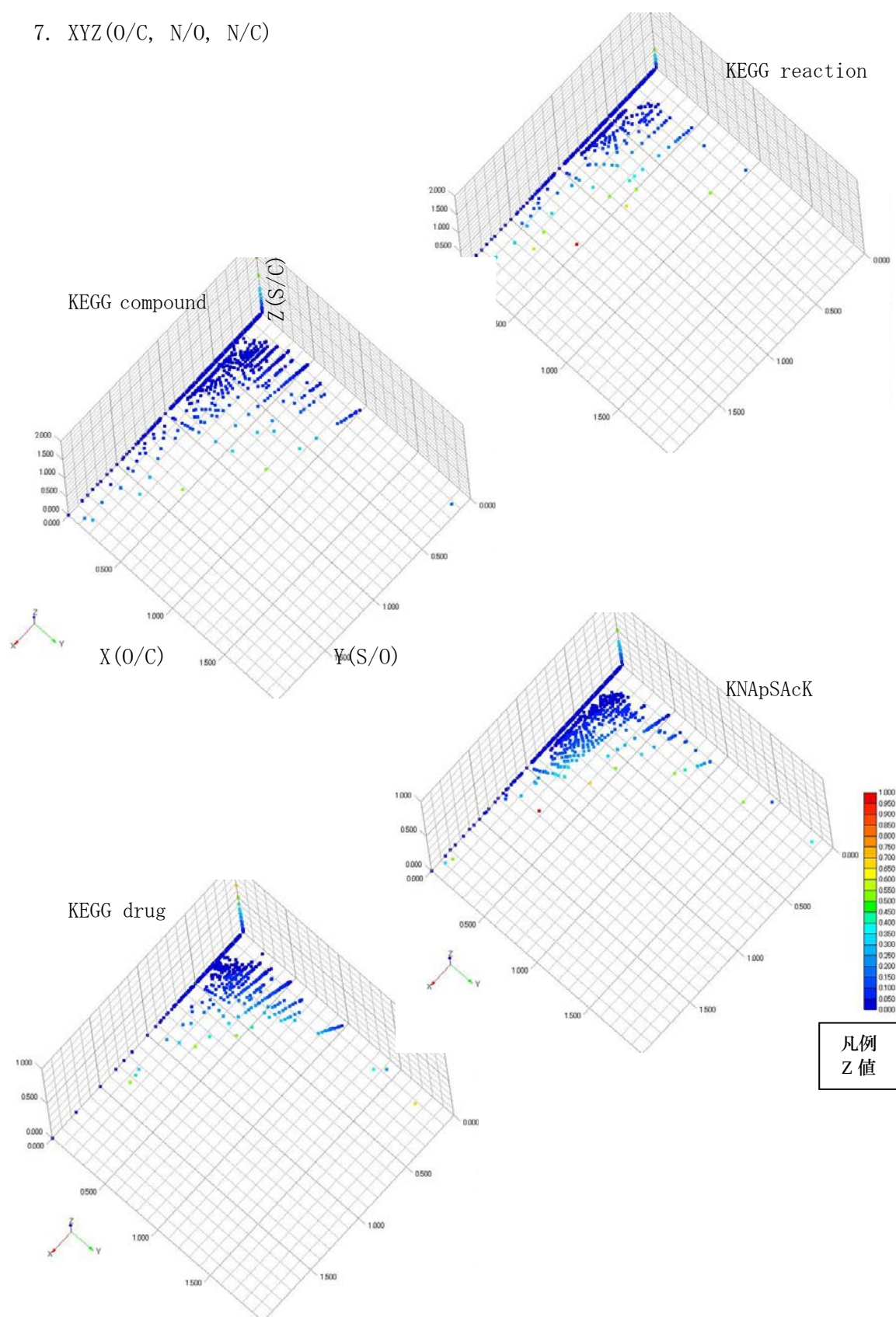
4. XYZ (N/H, H/C, N/C)



7. XYZ (O/C, N/O, N/C)

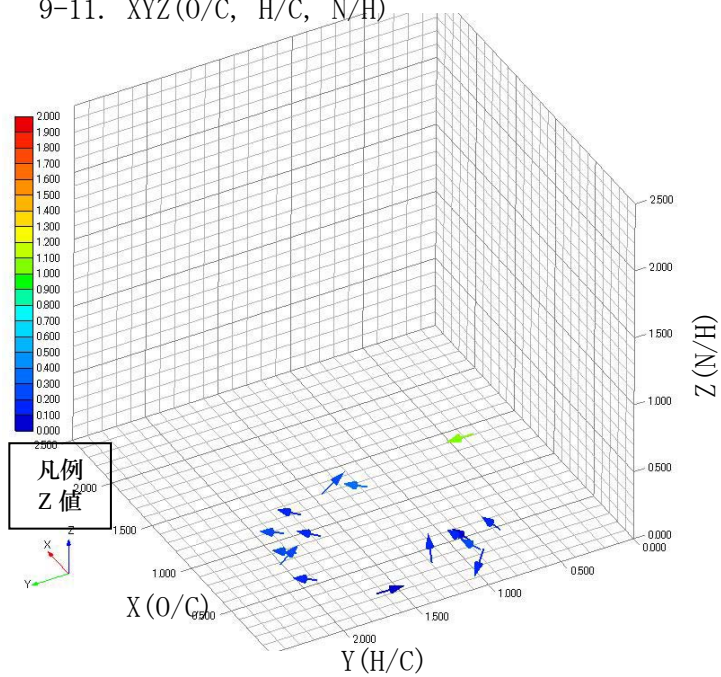


7. XYZ (O/C, N/O, N/C)

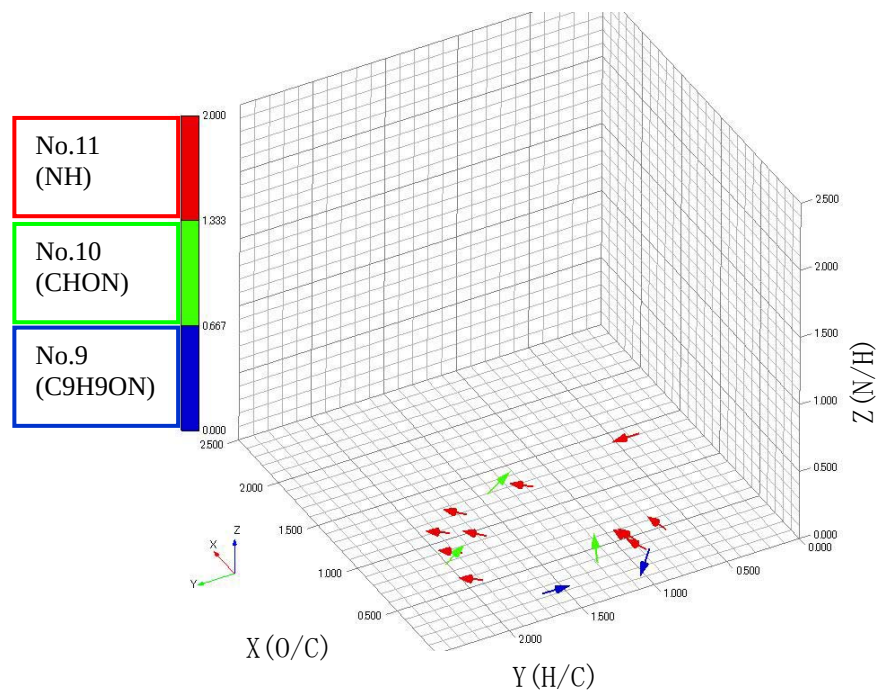


Appendix(3) 代謝反応

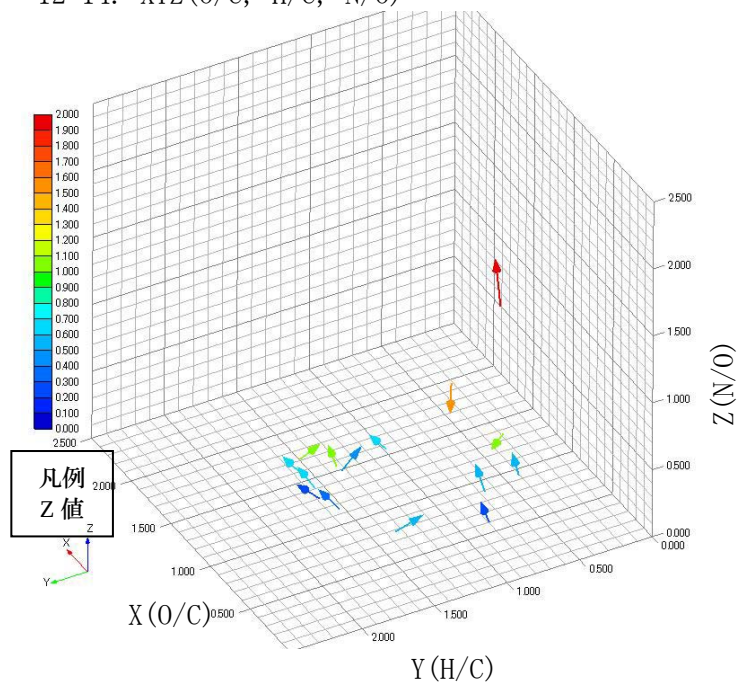
9-11. XYZ(O/C, H/C, N/H)



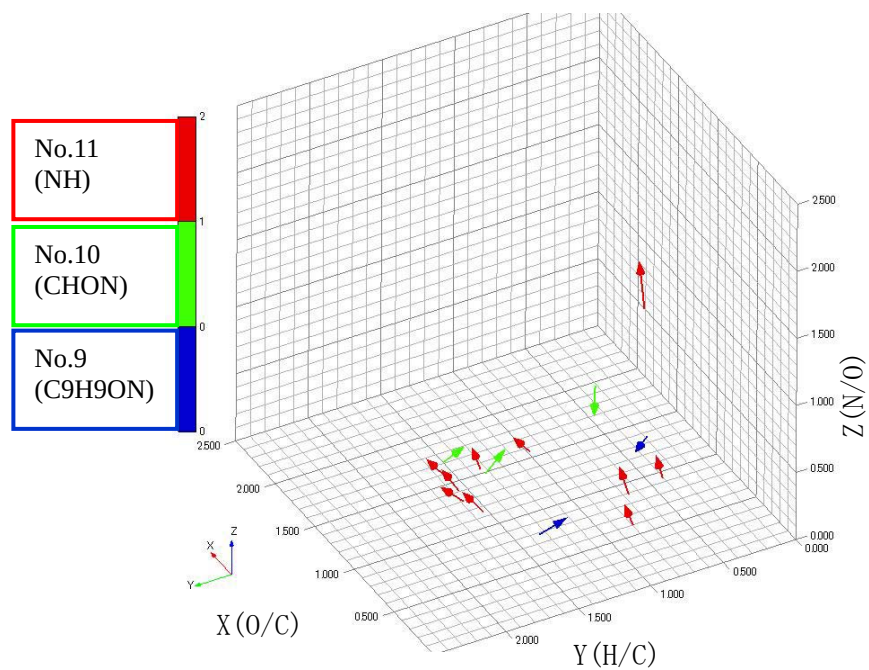
9-11. XYZ(O/C, H/C, N/H)



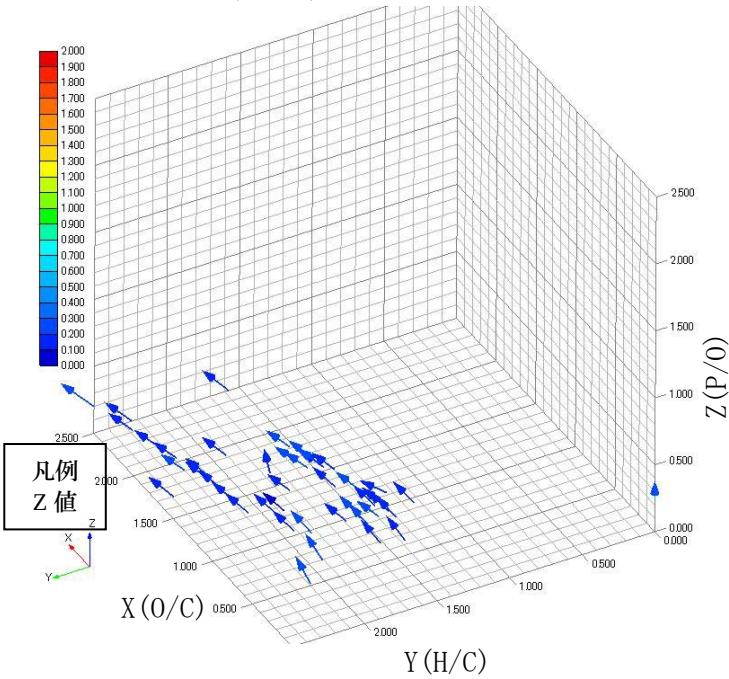
12-14. XYZ(O/C, H/C, N/O)



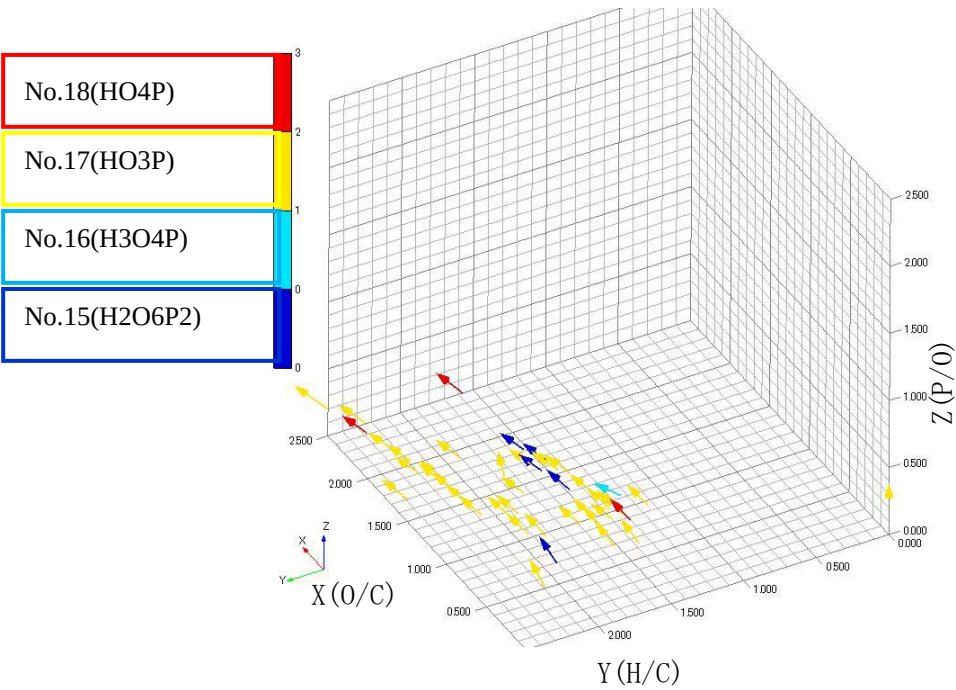
12-14. XYZ(O/C, H/C, N/O)



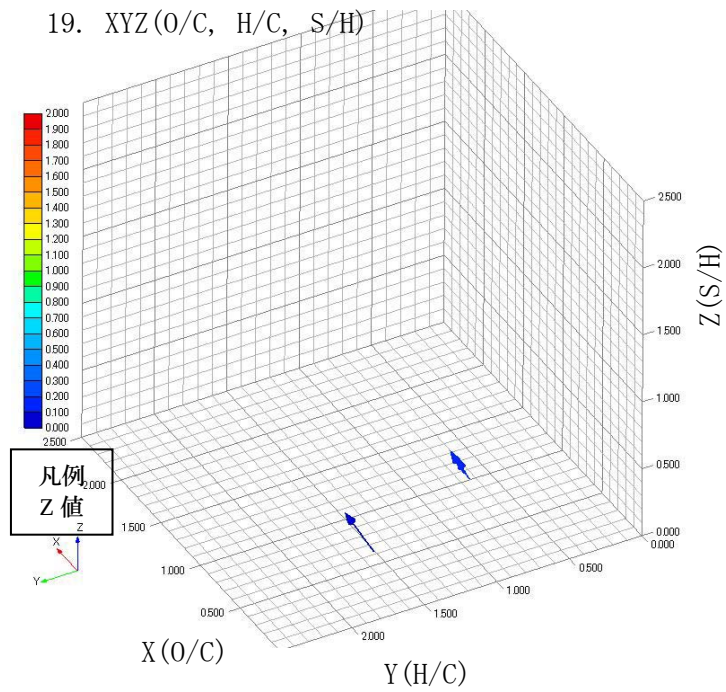
15-18. XYZ(O/C, H/C, P/O)



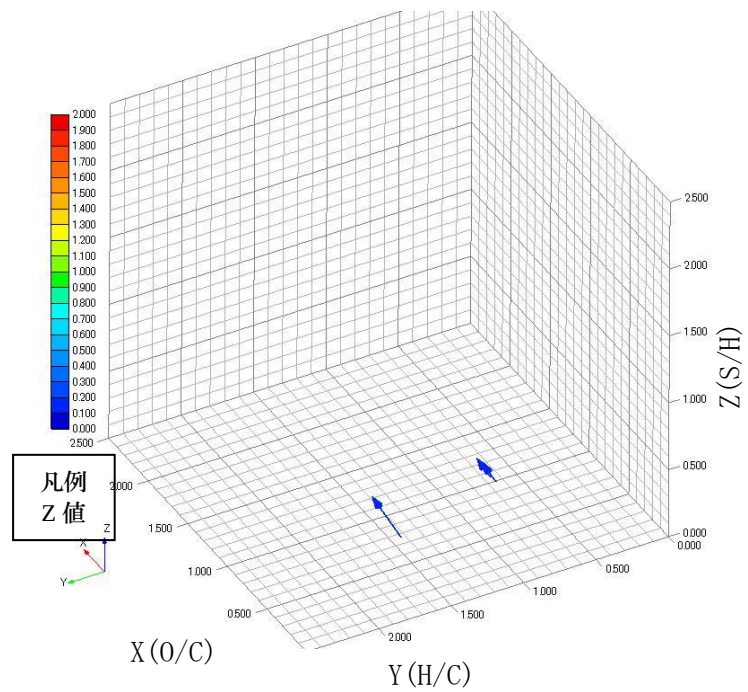
15-18. XYZ(O/C, H/C, P/O)



19. XYZ(O/C, H/C, S/H)



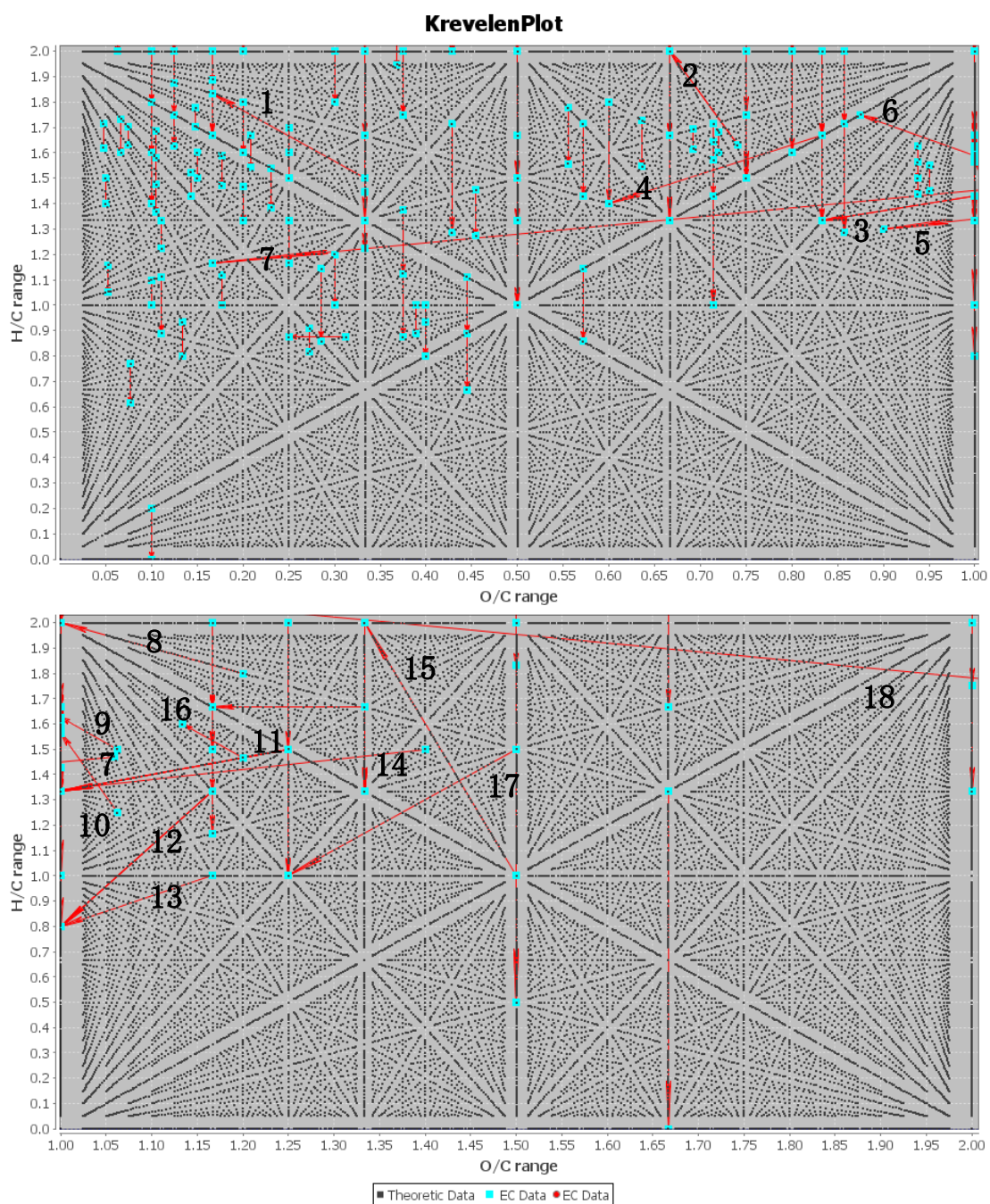
20. XYZ(O/C, H/C, P/O)



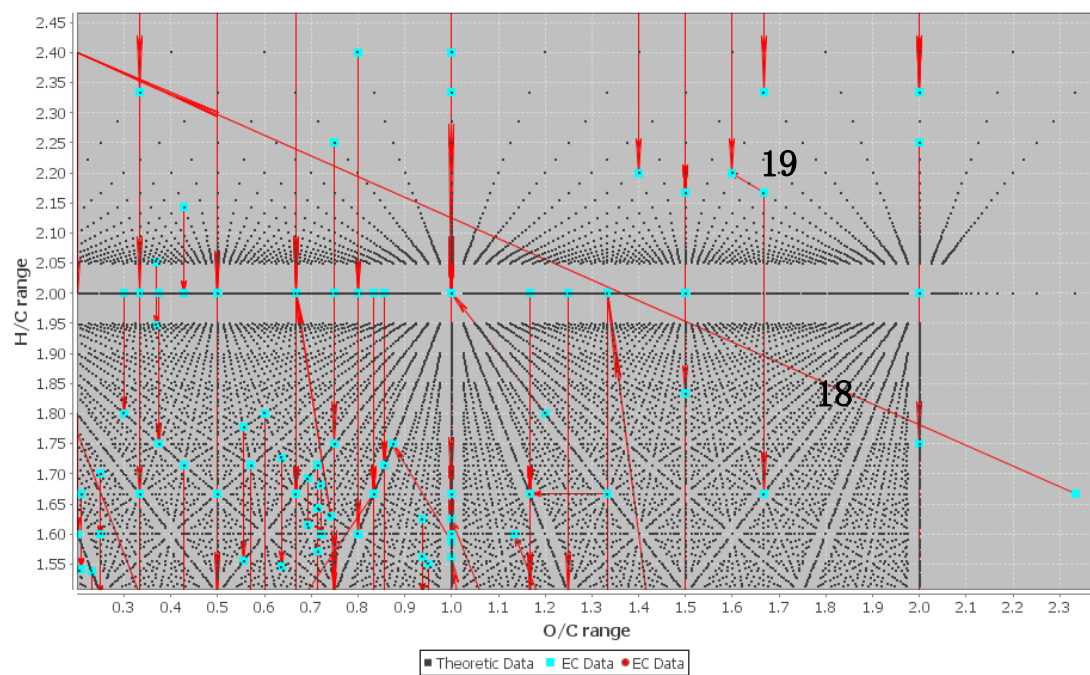
Appendix(4) 酵素反応

	EC	反応（下線付きが反応ペア）	反応前	反応後	反応
1	1. 1. 1. 23	<u>L-histidinol</u> + 2 NAD ⁺ = <u>L-histidine</u> + 2 NADH + 2 H ⁺	C6H11N3O	C6H9N3O2	付加 (H2O)
2	1. 1. 1. 34	<u>(R)-mevalonate</u> + CoA + 2 NADP ⁺ = <u>(S)-3-hydroxy-3-methylglutaryl-CoA</u> + 2 NADPH + 2 H ⁺	C6H12O4	C27H44N7O20 P3S	付加 (CoA)
2	1. 1. 1. 88	<u>(R)-mevalonate</u> + CoA + 2 NAD ⁺ = <u>3-hydroxy-3-methylglutaryl-CoA</u> + 2 NADH + 2 H ⁺	C6H12O4	C27H44N7O20 P3S	付加 (CoA)
3	1. 1. 1. 87	<u>(1R, 2S)-1-hydroxybutane-1, 2, 4-tricarboxylate</u> + NAD ⁺ = <u>2-oxoadipate</u> + CO ₂ + NADH + H ⁺	C7H10O7	C6H8O5	脱離 (CH2O2)
3	1. 1. 1. 286	<u>isocitrate</u> + NAD ⁺ = <u>2-oxoglutarate</u> + CO ₂ + NADH	C6H8O7	C5H4O5	脱離 (CH4O2)
4	1. 1. 1. 84	<u>(R)-3, 3-dimethylmalate</u> + NAD ⁺ = <u>3-methyl-2-oxobutanoate</u> + CO ₂ + NADH	C6H10O5	C5H7O3	脱離 (CH3O2)
5	1. 1. 1. 205	<u>inosine 5'-phosphate</u> + NAD ⁺ + H ₂ O = <u>xanthosine 5'-phosphate</u> + NADH + H ⁺	C10H15N4O14 P3	C10H13N4O9P	脱離 (H2O5P2)
6	1. 1. 1. 240	<u>N-acetyl-D-glucosamine</u> + NAD ⁺ = <u>N-acetyl-D-glucosamine</u> + NADH + H ⁺	C17H27N3O17 P2	C8H14N07	脱離 (C9H13N2O10P2)
7	1. 1. 1. 247	<u>codeine</u> + NADP ⁺ = <u>codeinone</u> + NADPH + H ⁺	C18H21N03	C17H25N3O18 P2	付加 (H4N2O15P2) 脱離 (C)
8	1. 1. 1. 115	<u>D-ribose</u> + NADP ⁺ + H ₂ O = <u>D-ribonate</u> + NADPH + H ⁺	C5H10O5	C5H9O6	脱離 (H) 付加 (O)
9	1. 1. 1. 186	<u>dTDP-D-galactose</u> + 2 NADP ⁺ + H ₂ O = <u>dTDP-D-galacturonate</u> + 2 NADPH + 2 H ⁺	C16H26N2O16 P2	C16H24N2O17 P2	脱 離 (H2) 付加 (O)

10	1. 1. 1. 13 2	<u>GDP-D-mannose</u> + 2 NAD ⁺ + H ₂ O = <u>GDP-D-mannuronate</u> + 2 NADH + 2 H ⁺	C16H25N5016 P2	C16H20N5017 P2	脱 離 (H5) 付加 (O)
11	1. 1. 1. 38	<u>(S)-malate</u> + NAD ⁺ = <u>pyruvate</u> + CO ₂ + NADH	C4H605	C3H403	脱離 (CH2O2)
11	1. 1. 1. 39	<u>(S)-malate</u> + NAD ⁺ = <u>pyruvate</u> + CO ₂ + NADH	C4H605	C3H403	脱離 (CH2O2)
11	1. 1. 1. 40	<u>(S)-malate</u> + NADP ⁺ = <u>pyruvate</u> + CO ₂ + NADPH	C4H605	C3H403	脱離 (CH2O2)
11	1. 1. 1. 83	<u>(R)-malate</u> + NAD ⁺ = <u>pyruvate</u> + CO ₂ + NADH	C4H605	C3H403	脱離 (CH2O2)
12	1. 1. 1. 42	<u>isocitrate</u> + NADP ⁺ = <u>2-oxoglutarate</u> + CO ₂ + NADPH	C6H807	C5H405	脱離 (CH4O2)
13	1. 1. 1. 42	<u>oxalosuccinate</u> + NADP ⁺ = <u>2-oxoglutarate</u> + CO ₂ + NADPH	C6H607	C5H405	脱離 (CH2O2)
14	1. 1. 1. 20 5	<u>inosine 5'-phosphate</u> + NAD ⁺ + H ₂ O = <u>xanthosine 5'-phosphate</u> + NADH + H ⁺	C10H15N4014 P3	C10H13N409P	脱離 (H2O5P2)
15	1. 1. 1. 92	<u>D-glycerate</u> + NAD(P) ⁺ + CO ₂ = <u>2-hydroxy-3-oxosuccinate</u> + NAD(P)H + 2 H ⁺	C3H604	C4H406	付加 (CO2) 脱 離 (H2)
16	1. 1. 1. 22	<u>UDP-glucose</u> + 2 NAD ⁺ + H ₂ O = <u>UDP-glucuronate</u> + 2 NADH + 2 H ⁺	C15H24N2017 P2	C15H22N2018 P2	付加 (O) 脱 離 (H2)
17	1. 1. 1. 93	<u>tartrate</u> + NAD ⁺ = <u>oxaloglycolate</u> + NADH + H ⁺	C4H606	C4H405	付加 (H2O)
18	1. 1. 1. 95	<u>3-phospho-D-glycerate</u> + NAD ⁺ = <u>3-phosphonooxypyruvate</u> + NADH + H ⁺	C5H120	C3H507P	付加 (O6P) 脱離 (C2H7)
18	1. 1. 1. 95	<u>2-hydroxyglutarate</u> + NAD ⁺ = <u>2-oxoglutarate</u> + NADH + H ⁺	C5H805	C5H405	脱離 (H4)
19	1. 1. 1. 44	<u>6-phospho-D-gluconate</u> + NADP ⁺ = <u>D-ribulose 5-phosphate</u> + CO ₂ + NADPH	C6H13010P	C5H1108P	脱離 (CH2O2)



KrevelenPlot



Appendix(5) 生合成反応

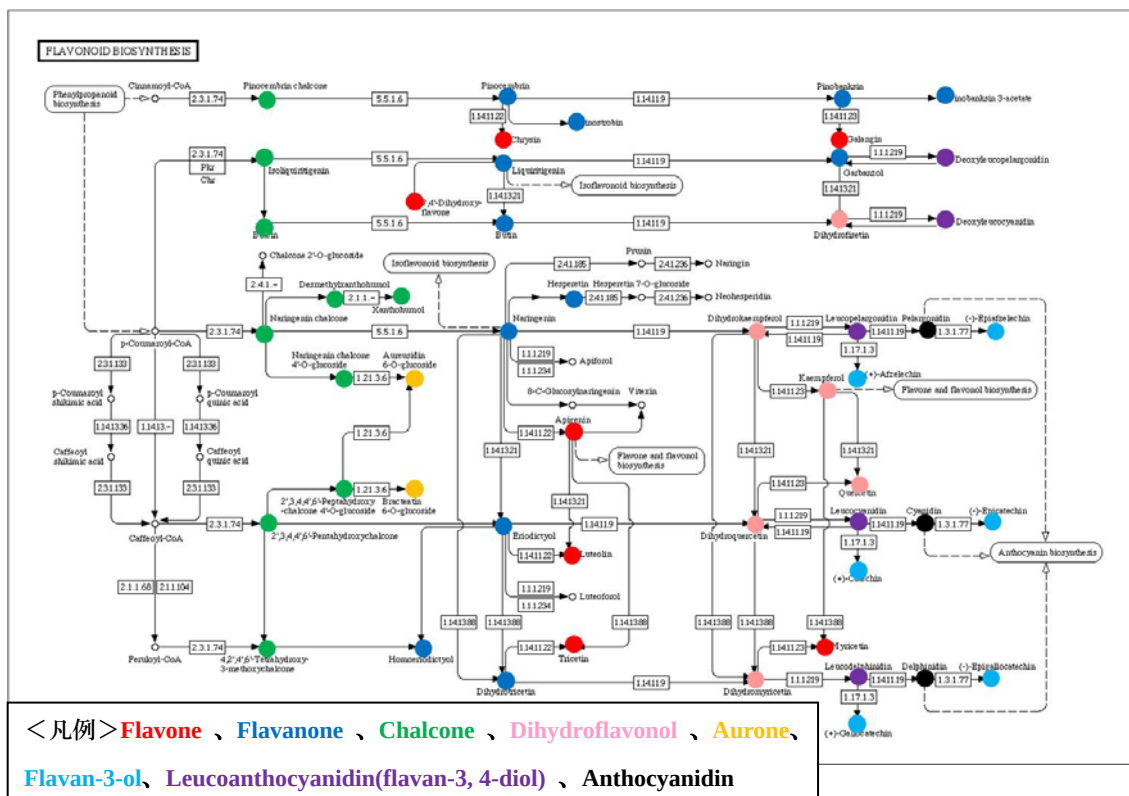
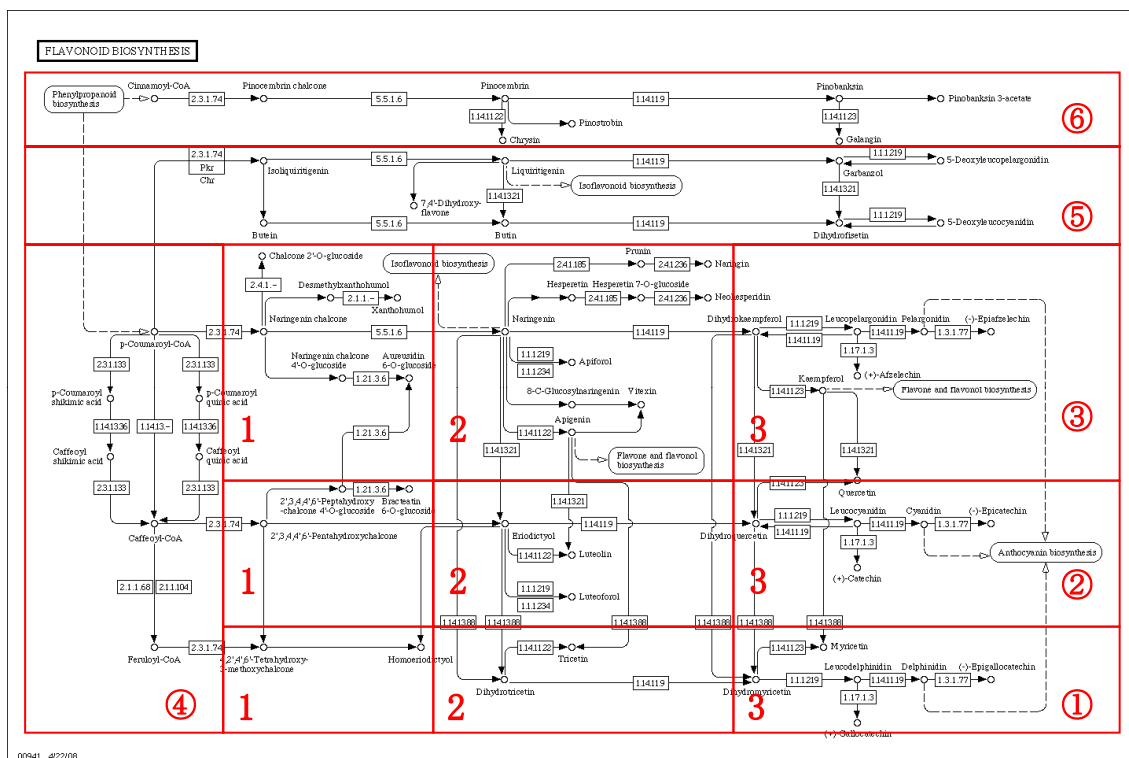


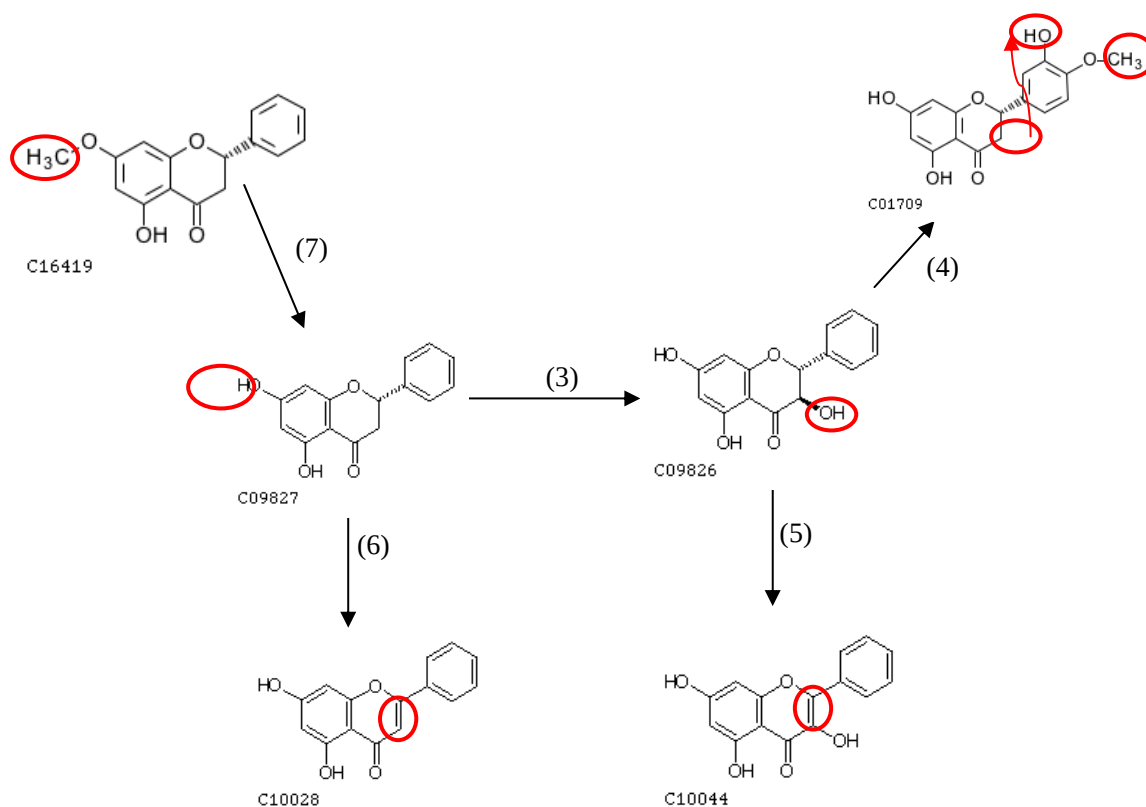
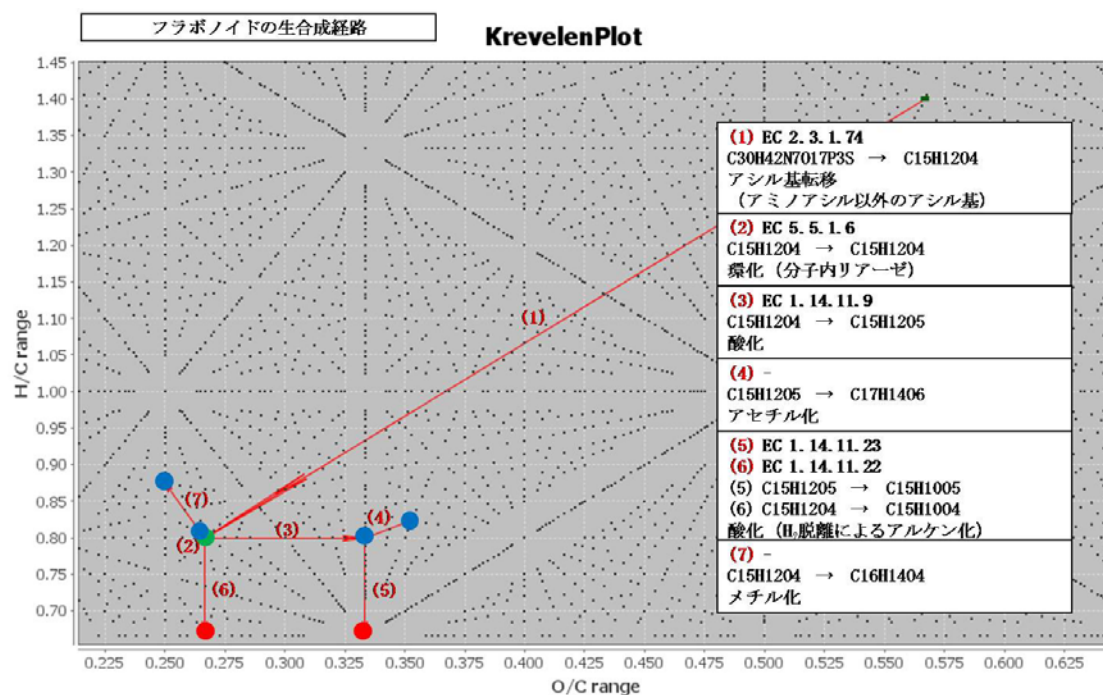
表 0-3-3-1 フラボノイドの生合成反応

No		反応前	反応後	EC	反応種類
①	1	C30H42N7017P3S	C15H1204	2. 3. 1. 74	アシル基転移
	2	C15H1204	C15H1204	5. 5. 1. 6	環化（分子内リアーゼ）
	3	C15H1204	C15H1205	1. 14. 11. 9	酸化
	4	C15H1205	C17H1406	—	アセチル化
	5	C15H1205	C15H1005	1. 14. 11. 23	酸化（H ₂ 脱離によるアルケン化）
	6	C15H1204	C15H1004	1. 14. 11. 22	酸化（H ₂ 脱離によるアルケン化）
	7	C15H1204	C16H1404	—	メチル化
②	1	C30H42N7018P3S	C15H1204	2. 3. 1. 74	アシル基転移
	2	C15H1204	C15H1204	5. 5. 1. 6	環化（分子内リアーゼ）
	3	C15H1204	C15H1205	1. 14. 11. 9	酸化
	4	C15H1205	C15H1405	1. 1. 1. 219	還元
	5	C15H1205	C15H1206	1. 14. 13. 21	酸化
	6	C15H1204	C15H1205	—	酸化
	7	C15H1205	C15H1205	5. 5. 1. 6	環化（分子内リアーゼ）
	8	C15H1204	C15H1004	—	酸化（H ₂ 脱離によるアルケン化）
	9	C15H1204	C15H1205	1. 14. 13. 21	酸化
	10	C15H1205	C15H1206	1. 14. 11. 9	酸化
	11	C15H1205	C15H1206	1. 14. 13. 21	酸化
	12	C15H1206	C15H1406	1. 1. 1. 219	還元
	13	C15H1405	C15H1205	—	酸化
	14	C15H1406	C15H1206	—	酸化
③	1	C16H1607	C30H42N7018P3S	2. 3. 1. 133	アシル基転移
	2	C16H1608	C16H1607	1. 14. 13. 36	アシル基転移
	3	C30H42N7019P3S	C16H1608	2. 3. 1. 133	酸化
	4	C30H42N7019P3S	C30H42N7018P3S	1. 14. 13. —	酸化
	5	C16H1808	C30H42N7018P3S	2. 3. 1. 133	アシル基付加
	6	C16H1809	C16H1808	1. 14. 13. 36	酸化
	7	C30H42N7019P3S	C16H1809	2. 3. 1. 133	アシル基付加
	8	C31H44N7019P3S	C30H42N7019P3S	2. 1. 1. 68 2. 1. 1. 104	メチル基転移（メチル化）
④-1	1	C15H1205	C30H42N7018P3S	2. 3. 1. 74	アシル基転移
	2	C15H1205	C21H22010	2. 4. 1. —	6 炭糖残基の転移

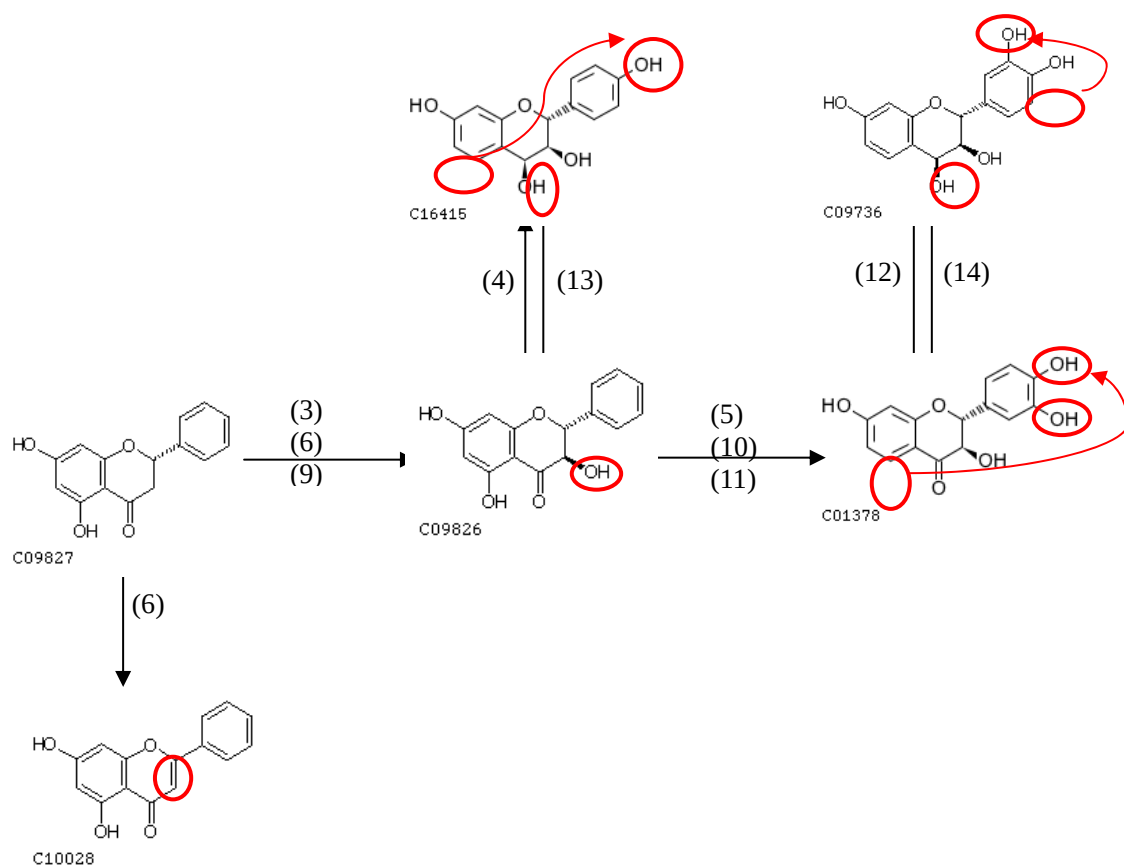
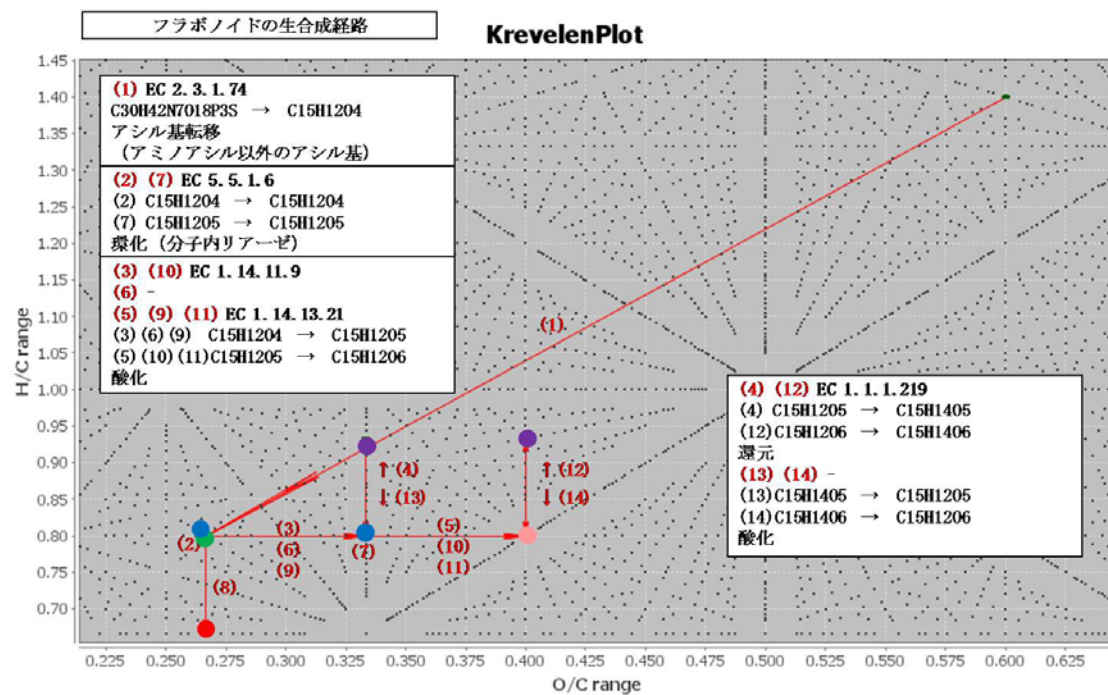
	3	C15H1205	C20H2005	—	アルケンの付加
	4	C20H2005	C21H2205	2. 1. 1. -	メチル基転移 (メチル化)
	5	C15H1205	C15H1205	5. 5. 1. 6	環化 (分子内リアーゼ)
	6	C15H1205	C21H22010	—	6 炭糖残基の転移
	7	C21H22010	C21H20011	1. 21. 3. 6	酸化および脱水
④-2	1	C15H1205	C21H22010	2. 4. 1. 185	6 炭糖残基の転移
	2	C21H22010	C27H32014	2. 4. 1. 236	6 炭糖残基の転移
	3	C15H1205	C16H1406	—	メチル化および酸化
	4	C16H1406	C22H24011	2. 4. 1. 185	6 炭糖残基の転移
	5	C22H24011	C28H34015	2. 4. 1. 236	6 炭糖残基の転移
	6	C15H1205	C15H1405	1. 1. 1. 219	還元
	7	C15H1205	C21H22010	—	6 炭糖残基の転移
	8	C21H22010	C21H20010	—	酸化 (H ₂ 脱離によるアルケン化)
	9	C15H1205	C15H1005	1. 14. 11. 22	酸化 (H ₂ 脱離によるアルケン化)
	10	C15H1205	C15H1206	1. 14. 13. 21	酸化
	11	C15H1205	C15H1207	1. 14. 13. 88	酸化 (2 分子)
	12	C15H1005	C15H1006	1. 14. 13. 21	酸化
	13	C15H1005	C15H1007	1. 14. 13. 88	酸化 (2 分子)
④-3	1	C15H1205	C15H1206	1. 14. 11. 9	酸化
	2	C15H1206	C15H1406	1. 1. 1. 219	還元 (ケトン→アルコール)
	3	C15H1406	C15H1206	1. 14. 11. 19	酸化 (アルコール→ケトン)
	4	C15H1406	C15H1105	1. 14. 11. 19	還元 (脱水+シクロアルカン→ベンゼン)
	5	C15H1406	C15H1405	1. 17. 1. 3	還元 (脱水)
	6	C15H1105	C15H1405	1. 3. 1. 77	還元 (二重結合の開裂)
	7	C15H1206	C15H1006	1. 14. 11. 23	酸化 (二重結合)
	8	C15H1206	C15H1207	1. 14. 13. 21	酸化
	9	C15H1206	C15H1208	1. 14. 13. 88	酸化 (2 分子)
	10	C15H1006	C15H1007	1. 14. 13. 21	酸化
⑤-1	1	C30H42N7019P3S	C15H1206	2. 3. 1. 74	アシル基転移
	2	C15H1206	C21H22011	—	6 炭糖残基の転移
	3	C21H22011	C21H20011	1. 21. 3. 6	酸化および脱水
	4	C21H22011	C21H20012	1. 21. 3. 6	酸化 (2 分子) および脱水
	5	C15H1206	C15H1206	—	環化
	6	C15H1206	C16H1406	—	メチル化

⑤-2	1	C15H1206	C15H1207	1. 14. 11. 9	酸化
	2	C15H1206	C15H1006	1. 14. 11. 22	酸化（二重結合）
	3	C15H1206	C15H1406	1. 1. 1. 219 1. 1. 1. 234	還元（ケトン→アルコール）
	4	C15H1206	C15H1207	1. 14. 13. 88	酸化
	5	C15H1206	C16H1406	—	メチル化
⑤-3	1	C15H1207	C15H1407	1. 1. 1. 219	還元（ケトン→アルコール）
	2	C15H1407	C15H1207	1. 14. 11. 19	還元（ベンゼン→シクロアルカン）
	3	C15H1407	C15H1106	1. 14. 11. 19	還元（脱水+シクロアルカン→ベンゼン）
	4	C15H1407	C15H1406	1. 17. 1. 3	還元（脱水）
	5	C15H1106	C15H1406	1. 3. 1. 77	酸化（アルコール→ケトン）
	6	C15H1207	C15H1208	1. 14. 13. 88	酸化
	7	C15H1207	C15H1007	1. 14. 11. 23	酸化（二重結合）
⑥-1	1	C31H44N7019P3S	C16H1406	2. 3. 1. 74	アシル基転移
	2	C16H1406	C16H1406	—	環化
⑥-2	1	C15H1207	C15H1007	1. 14. 11. 22	酸化（二重結合）
	2	C15H1207	C15H1208	1. 14. 11. 9	還元（ケトン→アルコール）
⑥-3	1	C15H1208	C15H1008	1. 14. 11. 23	酸化（二重結合）
	2	C15H1208	C15H1408	1. 1. 1. 219	還元（ケトン→アルコール）
	3	C15H1408	C15H1107	1. 14. 11. 19	酸化（二重結合）
	4	C15H1107	C15H1407	1. 3. 1. 77	還元（ベンゼン→シクロアルカン）
	5	C15H1408	C15H1407	1. 17. 1. 3	還元（脱水）

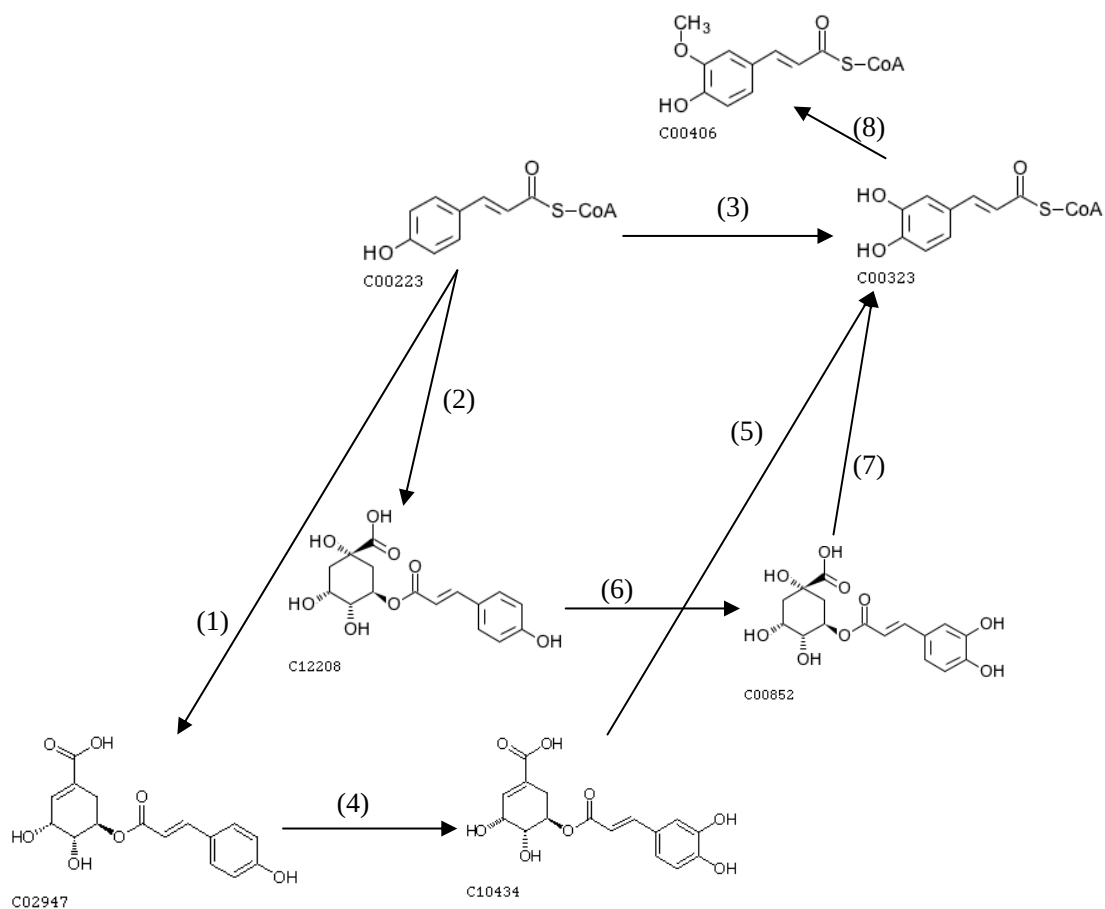
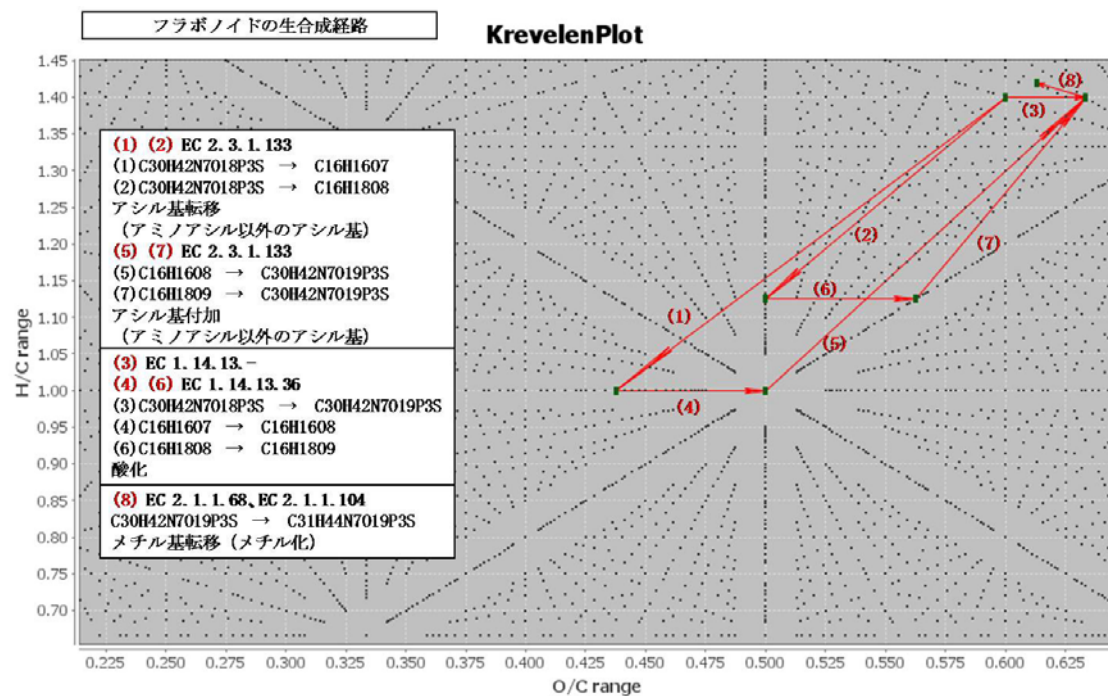
フラボノイドの生合成経路①



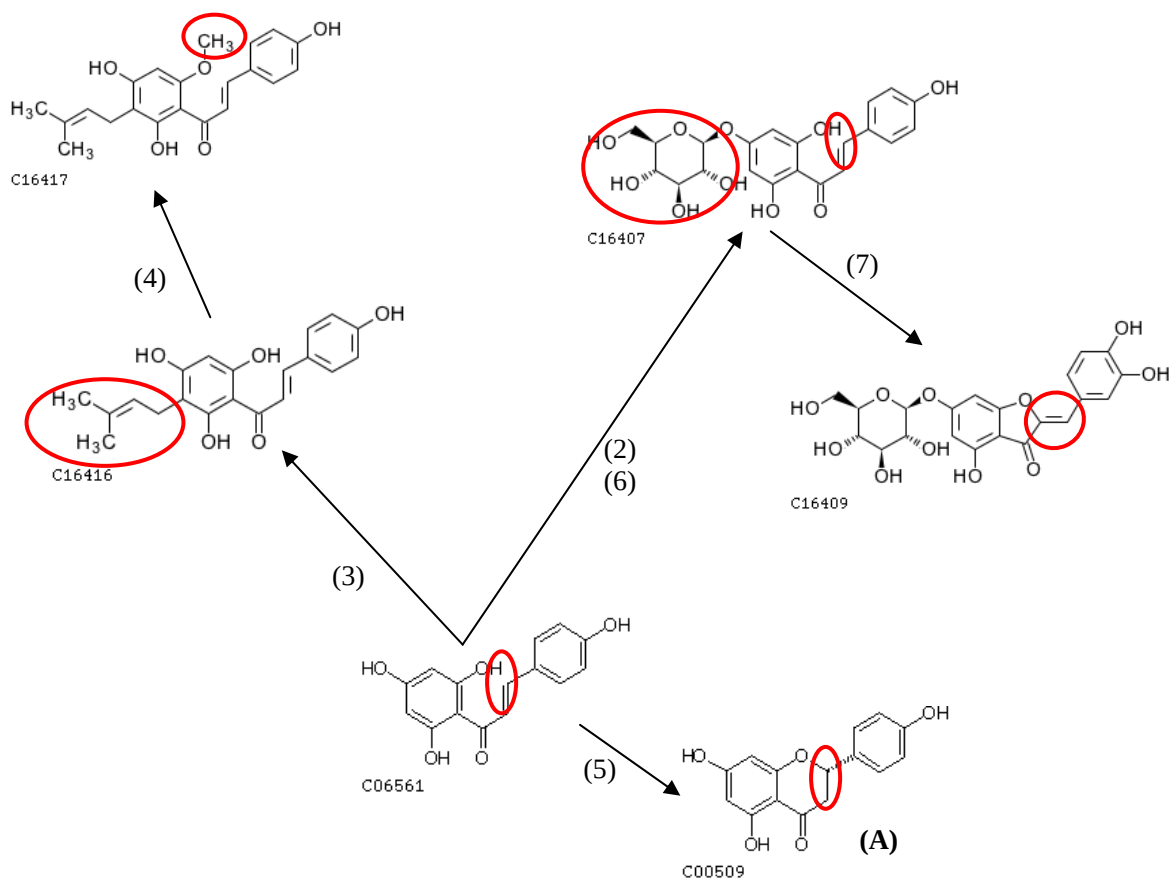
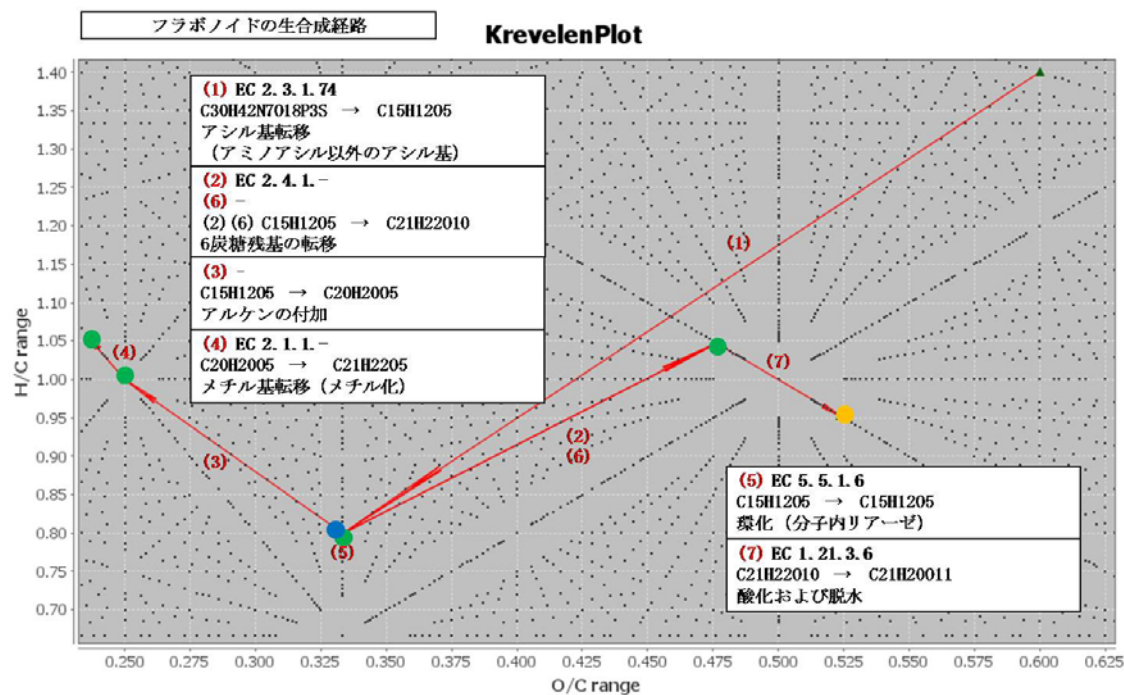
フラボノイドの生合成経路②



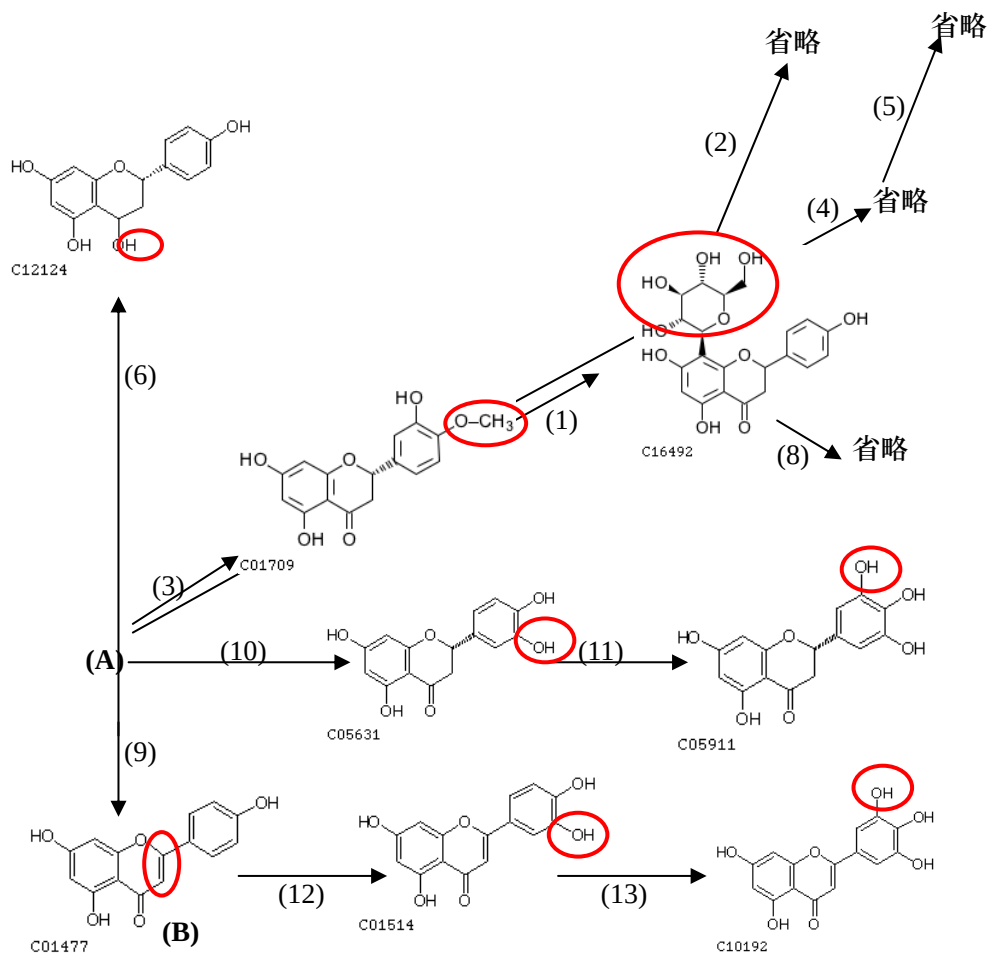
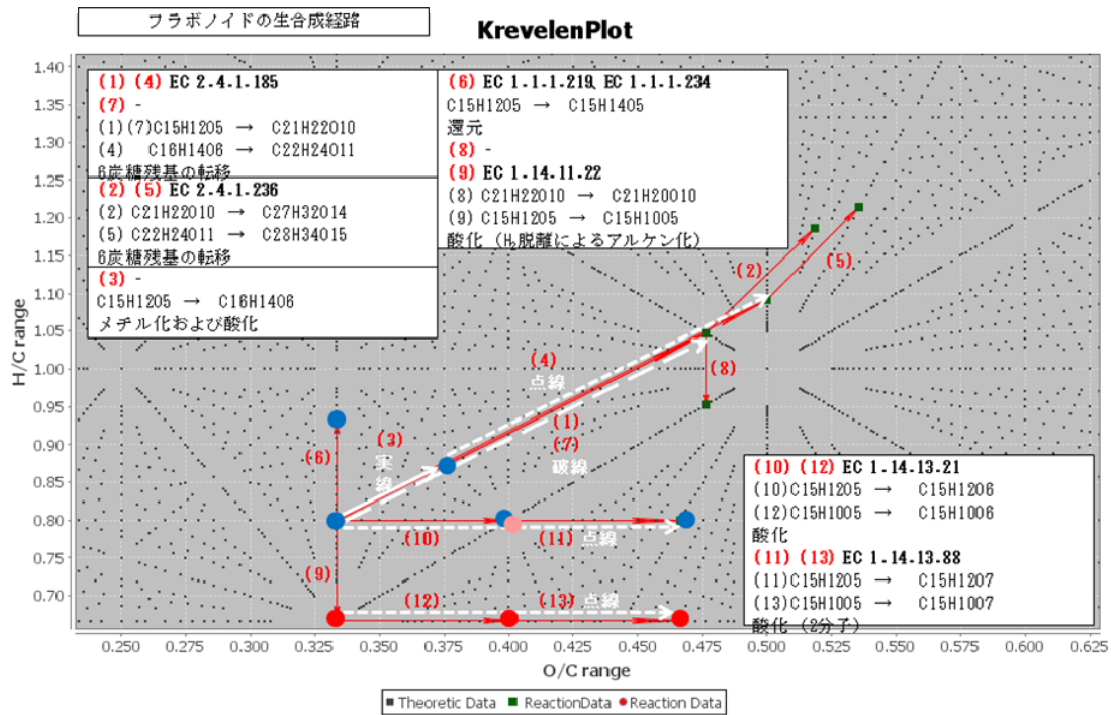
フラボノイドの生合成経路③



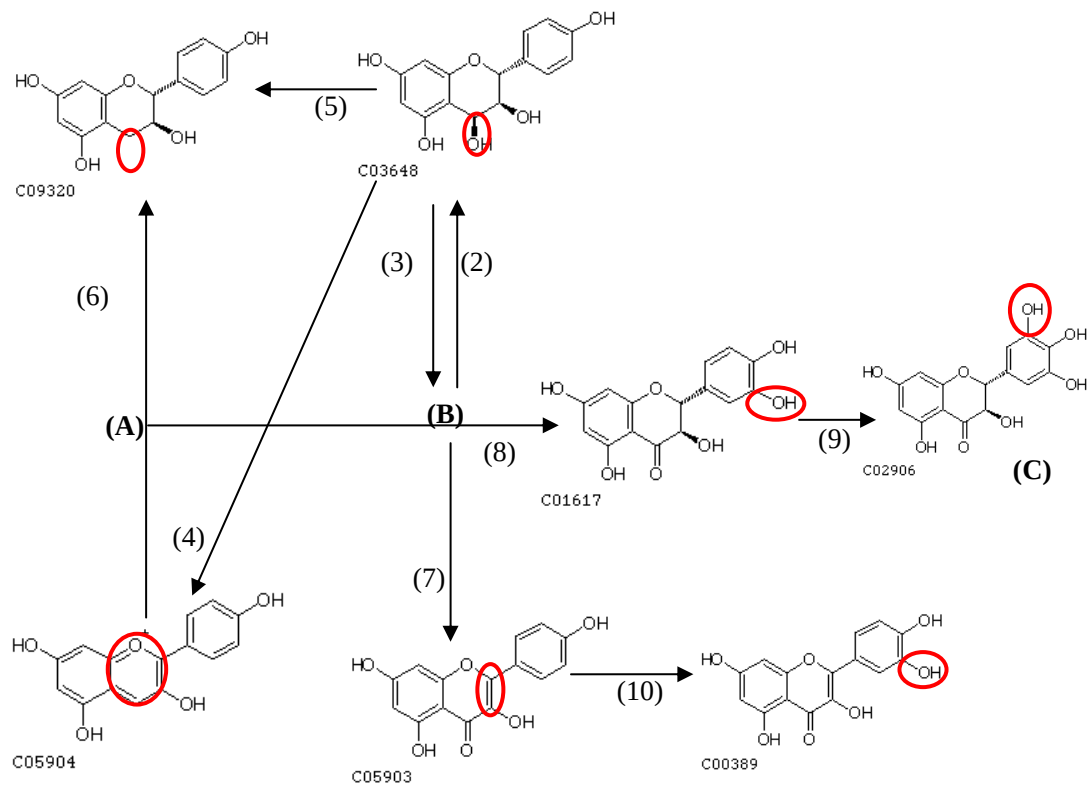
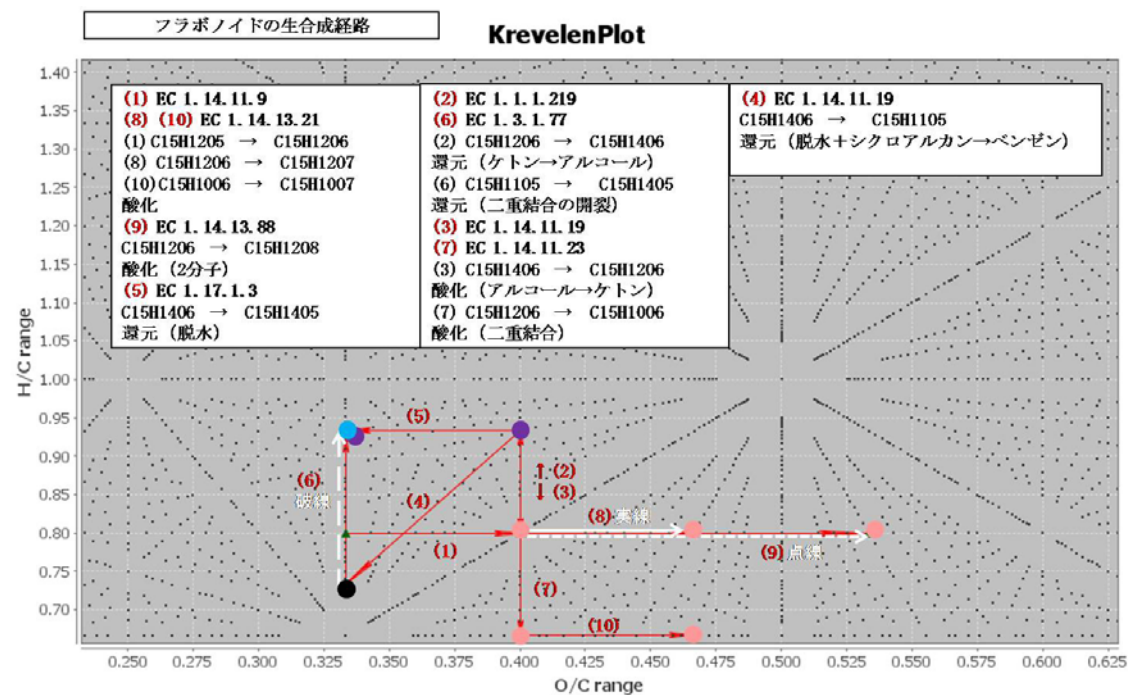
フラボノイドの生合成経路④-1



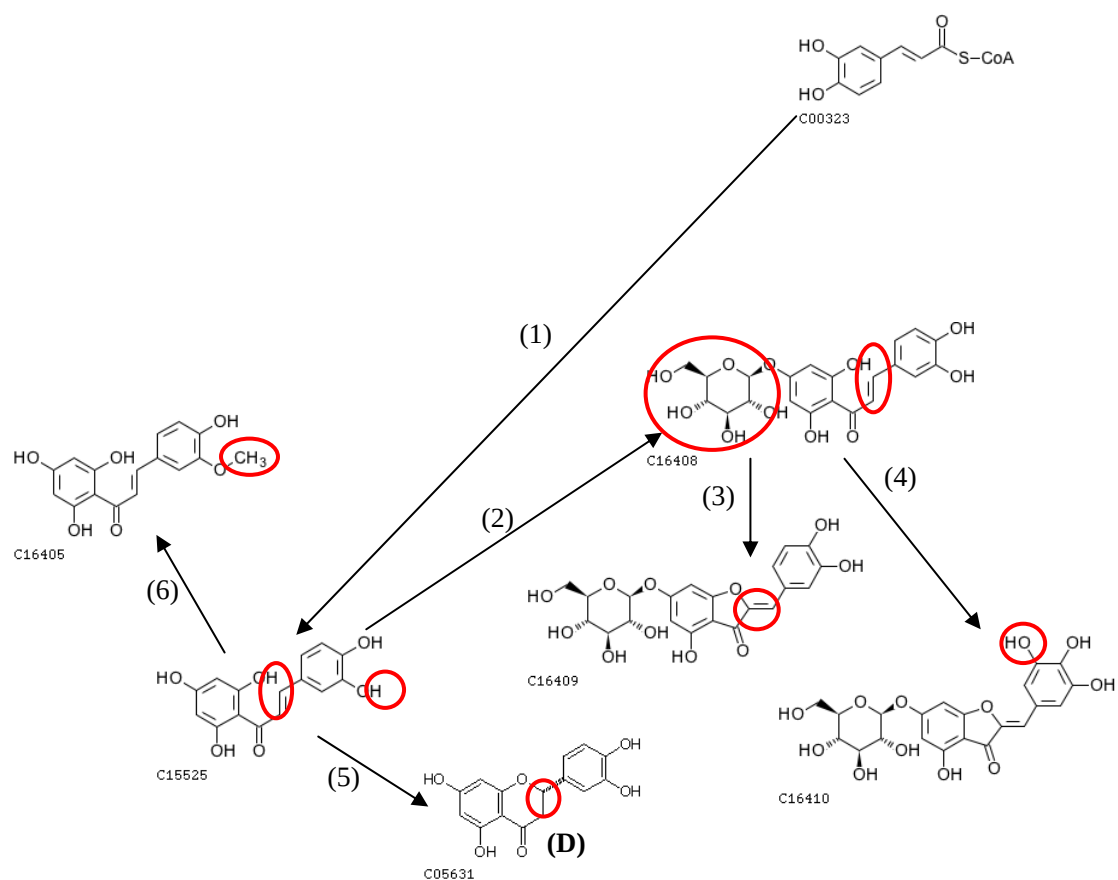
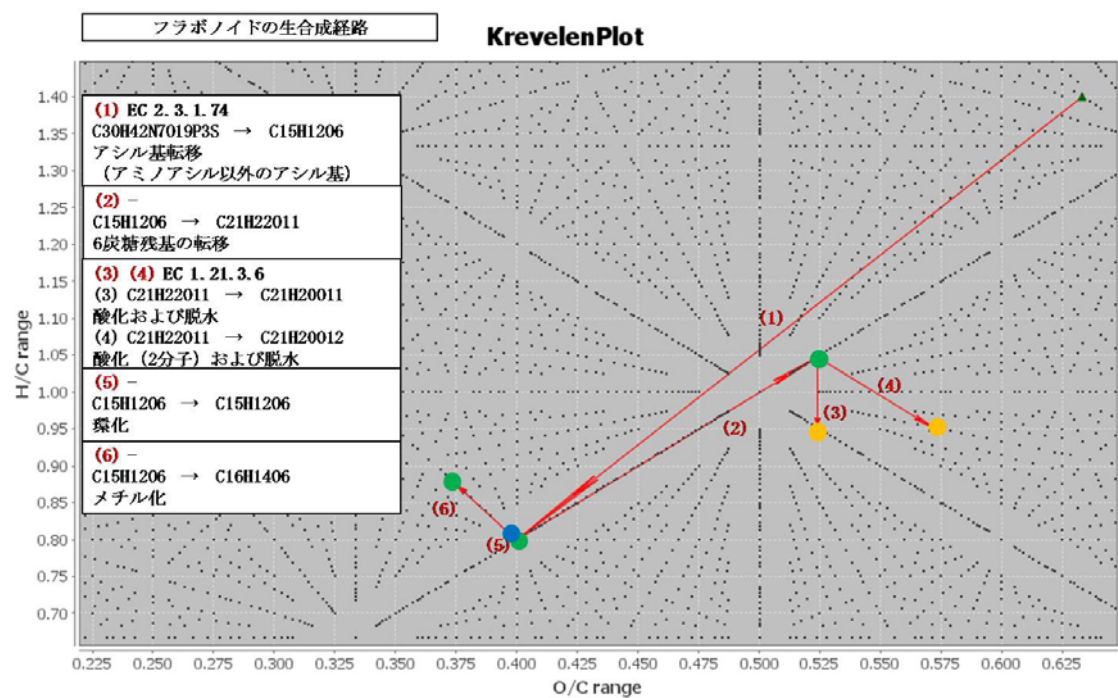
フラボノイドの生合成経路④-2



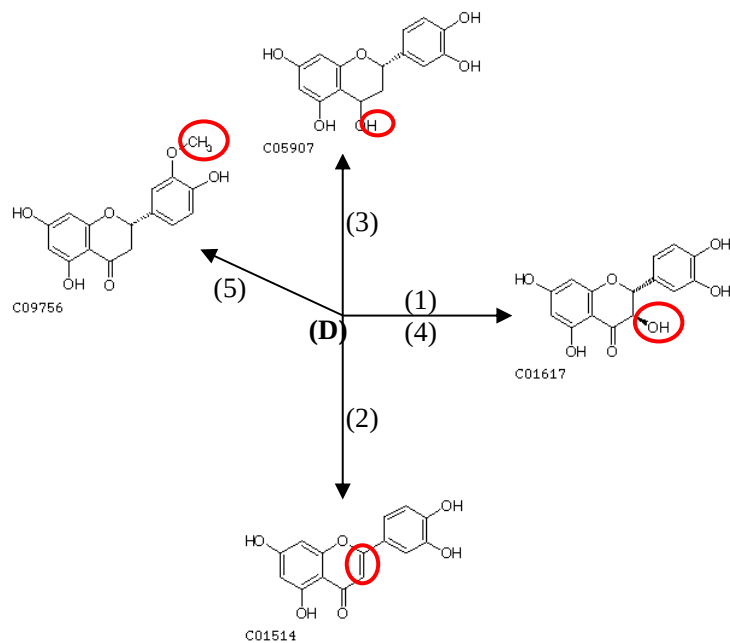
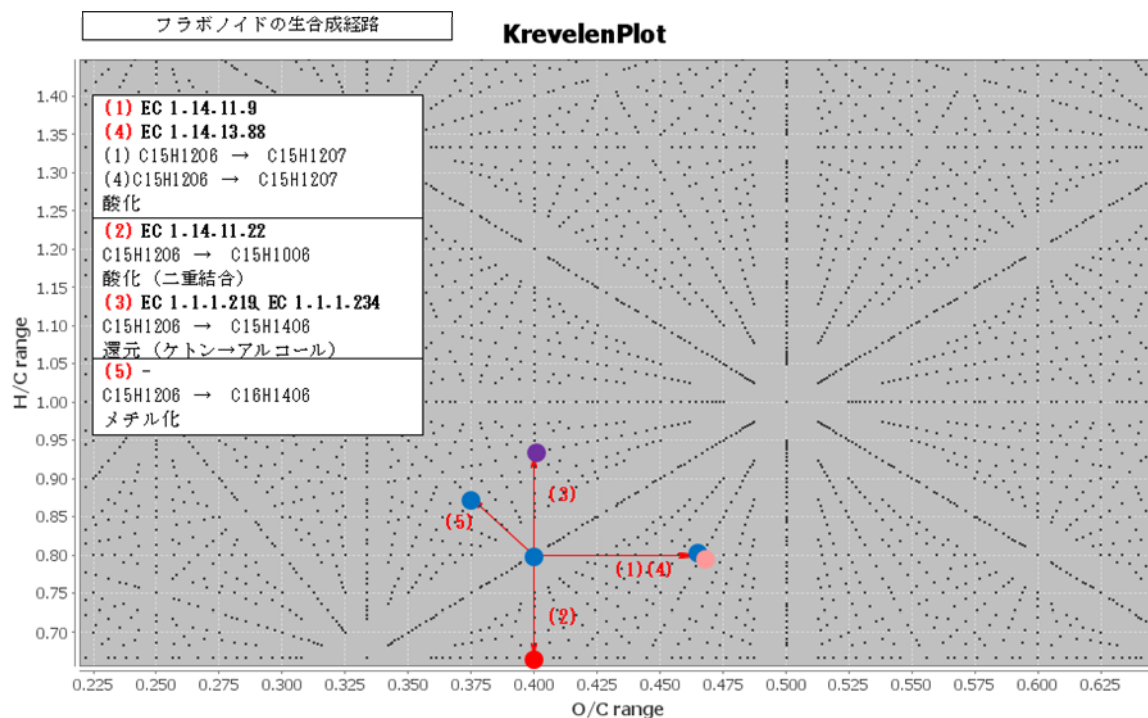
フラボノイドの生合成経路④-3



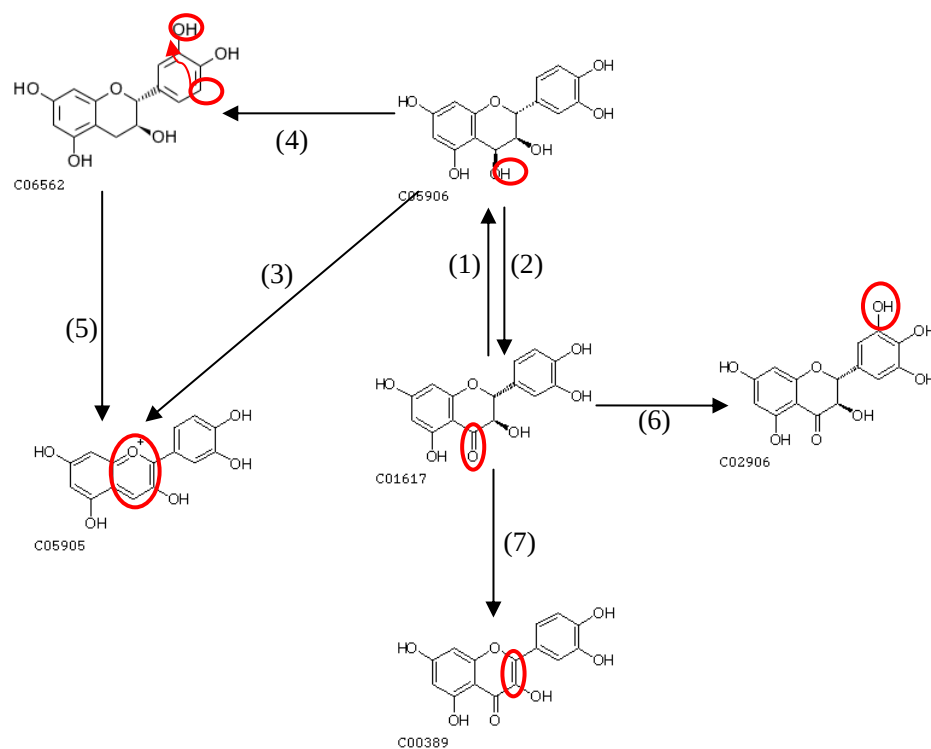
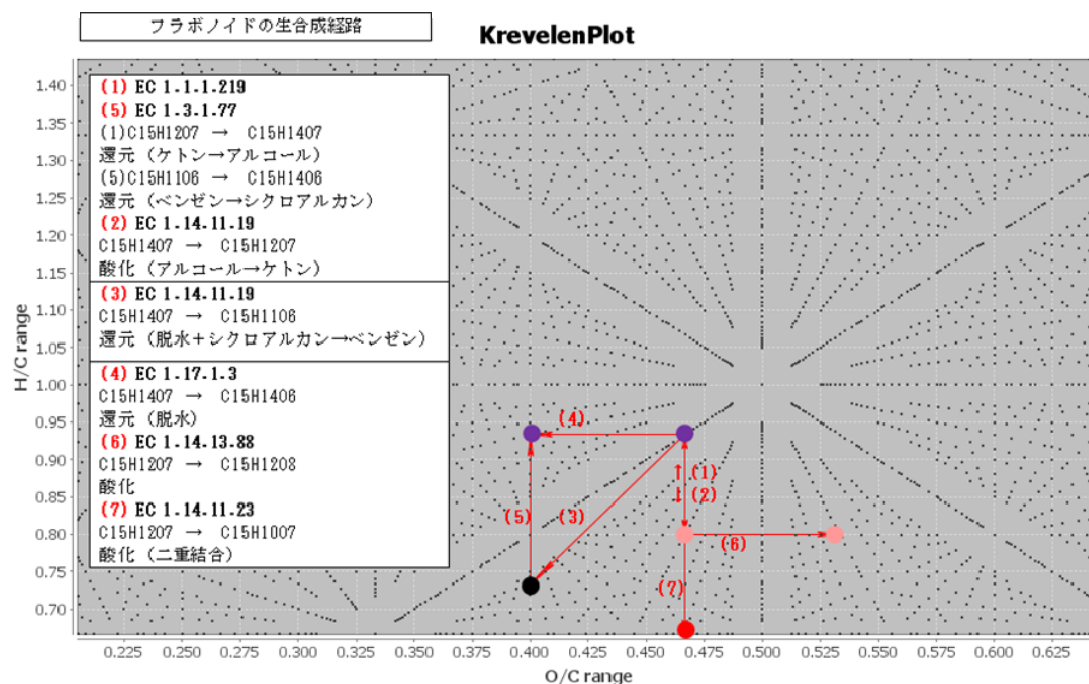
フラボノイドの生合成経路⑤-1



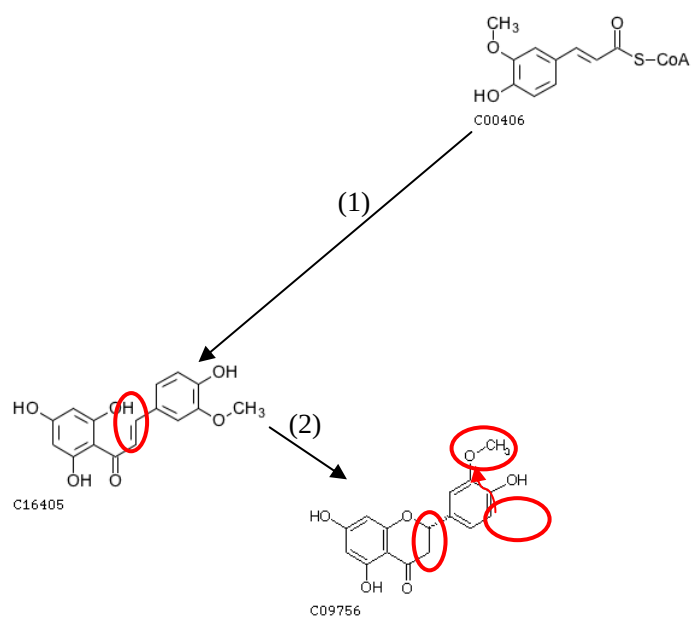
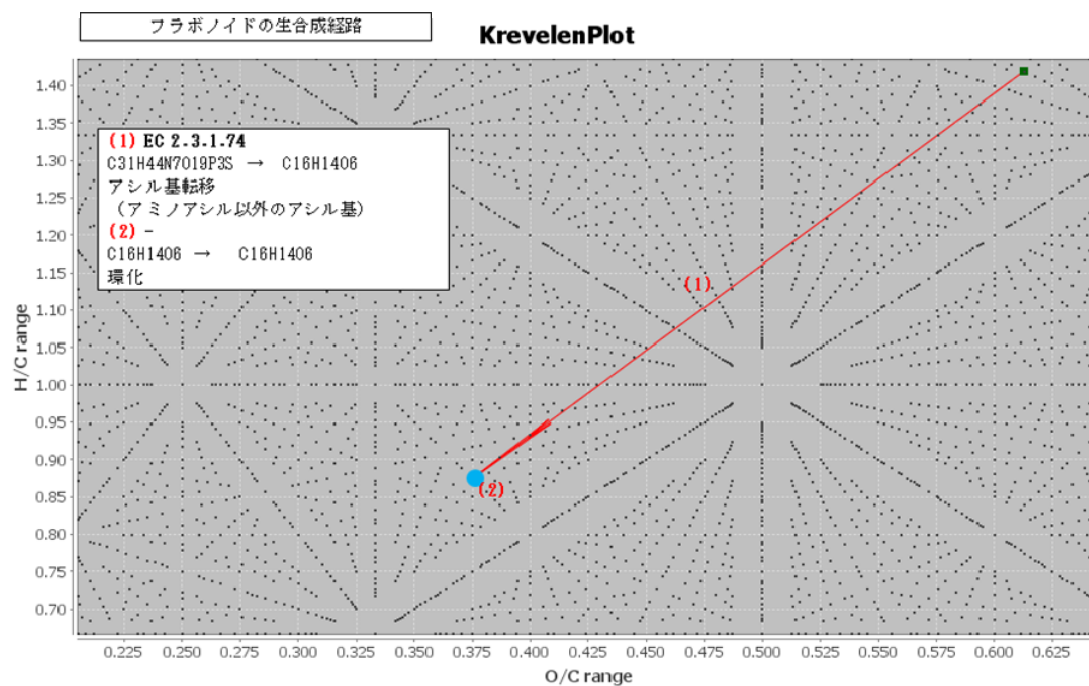
フラボノイドの生合成経路⑤-2



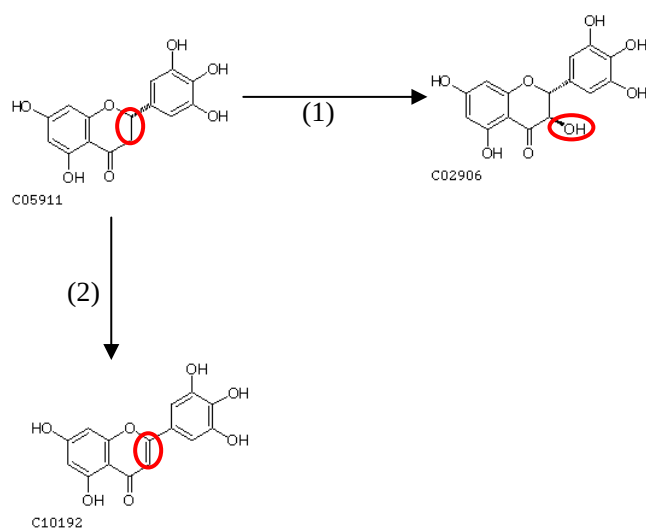
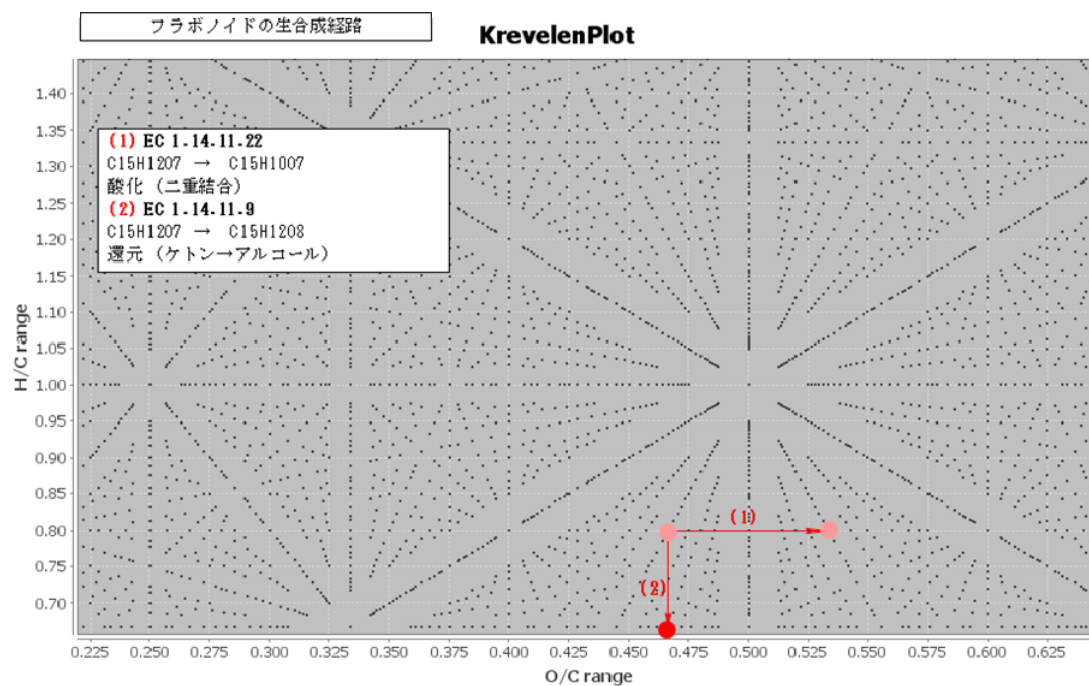
フラボノイドの生合成経路⑤-3



フラボノイドの生合成経路⑥-1



フラボノイドの生合成経路⑥-2



フラボノイドの生合成経路⑥-3

