

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5734586号
(P5734586)

(45) 発行日 平成27年6月17日(2015.6.17)

(24) 登録日 平成27年4月24日(2015.4.24)

(51) Int.Cl. F I
GO 1 N 33/00 (2006.01) GO 1 N 33/00 D

請求項の数 5 (全 18 頁)

(21) 出願番号 特願2010-158035 (P2010-158035)
(22) 出願日 平成22年7月12日(2010.7.12)
(65) 公開番号 特開2012-21806 (P2012-21806A)
(43) 公開日 平成24年2月2日(2012.2.2)
審査請求日 平成25年7月5日(2013.7.5)(73) 特許権者 000173924
公益財団法人野口研究所
東京都板橋区加賀一丁目8番1号
(74) 復代理人 100115624
弁理士 濱中 淳宏
(74) 復代理人 100162950
弁理士 久下 範子
(74) 代理人 110001243
特許業務法人 谷・阿部特許事務所
(72) 発明者 山田 一作
東京都板橋区加賀1丁目8番1号 財団法人野口研究所内

審査官 三木 隆

最終頁に続く

(54) 【発明の名称】 糖鎖構造認識用解析方法、糖鎖構造認識用解析装置およびプログラム

(57) 【特許請求の範囲】

【請求項1】

椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析方法であって、

解析対象の化学構造式データを受信するステップと、
前記受信した解析対象の化学構造式データから単糖の部分構造を抽出するステップと、
前記抽出した単糖の部分構造について構成原子の位置情報を取得するステップと、
前記取得した位置情報に基づいて、単糖の環の内角、および前記内角をなす角とするモードベクトルの外積を算出するステップと、

前記算出したモードベクトルの外積と前記算出した内角に基づいて、前記モードベクトルの外積および前記内角とモードとを関連付けるモードテーブルを使用して、単糖の回転構造のモードを判定するステップと、

前記取得した位置情報に基づいて、向き判定対象原子と、前記向き判定対象原子に結合する第1の環構成原子と、前記第1の環構成原子に隣接する第2の環構成原子とで構成される角をなす角とする向きベクトルの外積を算出するステップと、

前記向きベクトルの外積、前記内角、前記取得した位置情報、および前記判定したモードに基づいて、前記向き判定対象原子の向きを判定するステップと、

前記判定した回転構造および前記判定した向きに基づいて、単糖の略号と化学構造とを関連付ける略号テーブルを使用して、前記解析対象である単糖の化学構造を対応する略号に変換するステップと、

10

20

前記変換した略号を出力するステップと
を含むことを特徴とする糖鎖構造認識用解析方法。

【請求項 2】

環構成原子間の結合情報を取得するステップをさらに含み、
前記モードを判定するステップは、前記取得した結合情報に基づいてモードを判定し、
前記向きを判定するステップは、前記取得した結合情報に基づいて向きを判定すること
を特徴とする請求項 1 に記載の糖鎖構造認識用解析方法。

【請求項 3】

椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析装置であって、
解析対象の化学構造式データを受信する化学構造式データ受信部と、
前記受信した解析対象の化学構造式データから単糖の部分構造を抽出する部分構造検索部と、
前記抽出した単糖の部分構造について構成原子の位置情報を取得する位置情報取得部と、

前記取得した位置情報に基づいて単糖の環の内角、前記内角をなす角とするモードベクトルの外積、および向き判定対象原子と、前記向き判定対象原子に結合する第 1 の環構成原子と、前記第 1 の環構成原子に隣接する第 2 の環構成原子とで構成される角をなす角とする向きベクトルの外積を算出する算出部と、

前記モードベクトルの外積および前記内角とモードとを関連付けるモードテーブルと、
単糖の略号と化学構造とを関連付ける略号テーブルとを管理する記憶部と、

前記算出したモードベクトルの外積と前記算出した内角に基づいて、前記モードテーブルを使用して単糖の回転構造のモードを判定するモード判定部と、

前記向きベクトルの外積、前記内角、前記取得した位置情報、および前記判定したモードに基づいて、前記向き判定対象原子の向きを判定する向き判定部と、

前記判定した回転構造および前記判定した向きに基づいて、前記略号テーブルを使用して前記解析対象である単糖の化学構造を対応する略号に変換するデータ変換部と、

前記変換した略号を出力する出力部と
を備えたことを特徴とする糖鎖構造認識用解析装置。

【請求項 4】

環構成原子間の結合情報を取得する結合情報取得部をさらに含み、
前記モード判定部は、前記取得した結合情報に基づいてモードを判定し、
前記向き判定部は、前記取得した結合情報に基づいて向きを判定することを特徴とする
請求項 3 に記載の糖鎖構造認識用解析装置。

【請求項 5】

請求項 1 または 2 に記載の糖鎖構造認識用解析方法を、コンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、糖鎖の構造を認識するための解析方法、解析装置およびプログラムに関し、
具体的には、椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析方法、
解析装置およびプログラムに関する。

【背景技術】

【0002】

糖鎖は、遺伝情報を担う核酸、生体機能分子を構成するタンパク質に続く、第三の生命鎖として位置づけられ、その重要な特性は構造の多様性にある。核酸は 4 種類の塩基、タンパク質は 20 種類のアミノ酸を構成要素とし、一列に並ぶのに対して、糖鎖は、グルコース、ガラクトースなどの単糖を構成要素とする鎖状物質であり、単糖が保有する複数の水酸基を結合に活用し得るため、図 1 に示すように、複雑な構造を作り出すことができる

10

20

30

40

50

。

【0003】

糖鎖の構造を2次元上に表記する場合、構成要素である単糖単位でその種別を把握して、把握した単糖を組み合わせることにより糖鎖を識別し、表現することが慣例となっている。しかし、単糖の環状形を描く方法として、図2に示すように、Fischer投影、Haworth投影、Mills表示、配座を考慮した表示形式(以下、配座表示)、Glcなどの簡略化された文字列略号など、複数の表記方法が存在し、画一的ではない。さらに、例えば、D-グルコースを配座表示により描いた場合、立体構造の描画における観察方向の違いおよび配座によって、図3に示すように、同一の構造であるにも関わらず、24通りに描かれる。

10

【0004】

一方、単糖の構造を特定するためには、3次元構造における6員環炭素に結合する水酸基の結合向きを判定し、D/L異性体、L体/D体異性体を区別する必要がある。さらに、椅子型配座の糖については、 1C_4 と 4C_1 の2通りの異性体を区別する必要がある。したがって、ある表記方法で描かれた糖鎖を構成する単糖の構造を特定するためには、水酸基の向きを判定し、糖鎖における部分構造の母体となる単糖の種類を識別する処理、および図3に示される各化学構造式が同一の化合物であることを認識する処理が必要となる。また、糖鎖については、糖鎖の隣り合う単糖同士が、いずれの水酸基を用いてどのような向きで結合しているのかを判定する処理が必要となる。

【0005】

化学情報に関するデータベースは、現代の化学・創薬研究において欠くことのできない重要なツールとなっているばかりでなく、特許情報や試薬管理などにおいても必要不可欠なツールとなっている。コンピュータが化学情報をデータとして取り扱うことを可能にする、化学志向のアプリケーションプログラムとしては、ISIS Draw、ChemDraw(登録商標)、およびACD/ChemSketchなどが知られている。また、化学情報の入力を支援する機能として、化学略号やテンプレートを指定することにより、該当する化学構造式の入力を可能とするシステムが提案されている。

20

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特表2003-531419号公報

【特許文献2】特表2003-502773号公報

【非特許文献】

【0007】

【非特許文献1】M. Arita, T. Tokimatsu, ' 'Detection of monosaccharide types from coordinates' ', Proceedings of the 18th International Conference on Genome Informatics (Genome Informatics Series Vol. 19) pp3-14, 2007.

【発明の概要】

【発明が解決しようとする課題】

【0008】

上述のように、化学情報のシステム化が進んでおり、研究開発を支援する有益な情報を提供することが可能となっている。また、システムを利用するにあたって、化学構造式を入力するコストは削減されてきている。しかしながら、入力した化学構造式が意図した化学構造式となっているか否か確認するという点においては十分ではなかった。例えば、データ作成者が化学構造式を入力する場合、すでに入力されている類似の化学構造式を修正して新規の構造式を作成する場合がある。この際、修正漏れや変更不備により、意図している構造とは異なる化学構造式を作成してしまう可能性が考えられるが、上述したように複雑な化学構造を有する物質に関して、正確に確認することは困難である。

40

【0009】

化学構造式を入力する段階で誤った場合、以降のデータ検索において活用されない、誤

50

った情報を提供する、誤った化学構造式を利用することにより二次的な誤入力を招くなど、莫大な損失を被る問題に発展する。

【0010】

本発明は、このような問題に鑑みてなされたもので、その目的とするところは、入力された化学構造式が正確か否かを判定するために、簡略化された記号・略号を用いて化学構造式を表示することにより、ユーザの構造認識を支援する解析方法、解析装置およびプログラムを提供することにある。

【課題を解決するための手段】

【0011】

発明者は、椅子型配座の単糖の構造認識において、構成原子に関する所与の計算値を用いることによって、従来不可能であった / 異性体、および 1C_4 / 4C_1 異性体を区別できることを見いだした。そして、この知見により、椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析方法、解析装置、およびプログラムを完成させた。

10

【0012】

請求項1に記載の発明は、椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析方法であって、解析対象の化学構造式データを受信するステップと、前記受信した解析対象の化学構造式データから単糖の部分構造を抽出するステップと、前記抽出した単糖の部分構造について構成原子の位置情報を取得するステップと、前記取得した位置情報に基づいて、単糖の環の内角、および前記内角をなす角とするモードベクトルの外積を算出するステップと、前記算出したモードベクトルの外積と前記算出した内角に基づいて、前記モードベクトルの外積および前記内角とモードとを関連付けるモードテーブルを使用して、単糖の回転構造のモードを判定するステップと、前記取得した位置情報に基づいて、向き判定対象原子と、前記向き判定対象原子に結合する第1の環構成原子と、前記第1の環構成原子に隣接する第2の環構成原子とで構成される角をなす角とする向きベクトルの外積を算出するステップと、前記向きベクトルの外積、前記内角、前記取得した位置情報、および前記判定したモードに基づいて、前記向き判定対象原子の向きを判定するステップと、前記判定した回転構造および前記判定した向きに基づいて、単糖の略号と化学構造とを関連付ける略号テーブルを使用して、前記解析対象である単糖の化学構造を対応する略号に変換するステップと、前記変換した略号を出力するステップとを含むことを特徴とする。

20

30

【0013】

請求項2に記載の発明は、請求項1に記載の椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析方法であって、環構成原子間の結合情報を取得するステップをさらに含み、前記モードを判定するステップは、前記取得した結合情報に基づいてモードを判定し、前記向きを判定するステップは、前記取得した結合情報に基づいて向きを判定することを特徴とする。

【0014】

請求項3に記載の発明は、椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析装置であって、解析対象の化学構造式データを受信する化学構造式データ受信部と、前記受信した解析対象の化学構造式データから単糖の部分構造を抽出する部分構造検索部と、前記抽出した単糖の部分構造について構成原子の位置情報を取得する位置情報取得部と、前記取得した位置情報に基づいて単糖の環の内角、前記内角をなす角とするモードベクトルの外積、および向き判定対象原子と、前記向き判定対象原子に結合する第1の環構成原子と、前記第1の環構成原子に隣接する第2の環構成原子とで構成される角をなす角とする向きベクトルの外積を算出する算出部と、前記モードベクトルの外積および前記内角とモードとを関連付けるモードテーブルと、単糖の略号と化学構造とを関連付ける略号テーブルとを管理する記憶部と、前記算出したモードベクトルの外積と前記算出した内角に基づいて、前記モードテーブルを使用して単糖の回転構造のモードを判定するモード判定部と、前記向きベクトルの外積、前記内角、前記取得した位置情報、および前記判定したモードに基づいて、前記向き判定対象原子の向きを判定する向き判定部と、前

40

50

記判定した回転構造および前記判定した向きに基づいて、前記略号テーブルを使用して前記解析対象である単糖の化学構造を対応する略号に変換するデータ変換部と、前記変換した略号を出力する出力部とを備えたことを特徴とする。

【0015】

請求項4に記載の発明は、請求項3に記載の椅子型配座の単糖を構成要素とする糖鎖の化学構造を認識するための解析装置であって、環構成原子間の結合情報を取得する結合情報取得部をさらに含み、前記モード判定部は、前記取得した結合情報に基づいてモードを判定し、前記向き判定部は、前記取得した結合情報に基づいて向きを判定することを特徴とする。

【0016】

請求項5に記載の発明は、プログラムであって、請求項1または2に記載の糖鎖構造認識用解析方法を、コンピュータに実行させることを特徴とする。

【発明の効果】

【0017】

本発明により、化学構造式のデータ作成者は、作成した化学構造式が目的とする化学構造式と一致しているかを視覚的・直感的に判断することが可能となる。さらに、作成した化学構造式が目的とする化学構造式と一致していない場合、意図しない化学構造式を目的とする化学構造式に容易に修正することが可能となる。

【図面の簡単な説明】

【0018】

【図1】糖鎖の分岐構造を示す図である。

【図2】単糖の環状形を描く表記方法を示す図である。

【図3】 -D-グルコースの回転構造を示す図である。

【図4】本発明の一実施形態にかかる化学構造認識支援システムを示す構成図である。

【図5】本発明の一実施形態にかかる化学構造式解析サーバのモジュール構成図である。

【図6】本発明の一実施形態にかかる単糖を構成要素とする糖鎖の化学構造を認識するための全体の処理を示すフローチャートである。

【図7】本発明の一実施形態にかかる解析対象の化学構造式を示す図である。

【図8】本発明の一実施形態にかかるナンバリングした単糖を示す図である。

【図9】本発明の一実施形態にかかる配座表示で表記された単糖の回転構造を判定する処理を示すフローチャートである。

【図10】本発明の一実施形態にかかる配座表示で表記された単糖の水酸基の向きを判定する処理を示すフローチャートである。

【図11】本発明の一実施形態にかかる単糖の水酸基の向きを判定する処理において、6員環構成原子へ割り当てられる情報を示す図である。

【図12】本発明の一実施形態にかかる配座表示で表記された単糖の水酸基の向きを判定する処理を示すフローチャートである。

【図13】本発明の一実施形態にかかる配座表示で表記された単糖の水酸基の向きを判定する処理を示すフローチャートである。

【図14】本発明の一実施形態にかかる単糖の水酸基の向きを判定する処理において、判定に用いる領域情報を示す図である。

【図15】本発明の一実施形態にかかるモードDBに格納された情報の一例を示す図である。

【図16】本発明の一実施形態にかかる略号DBに格納された情報の一例を示す図である。

【図17】本発明の一実施形態にかかる化学構造式を解析した結果画面の一例を示す図である。

【図18】本発明の一実施形態にかかる化学構造式を解析した結果画面の変換例を示す図である。

【図19】本発明の一実施形態にかかる化学構造式を解析した結果画面の変換例を示す図

10

20

30

40

50

である。

【図20】本発明の一実施形態にかかる化学構造式を解析した結果画面の変換例を示す図である。

【発明を実施するための形態】

【0019】

図4は、本発明の一実施形態にかかる化学構造認識支援システムを示す構成図である。化学構造認識支援システムを実装する、化学構造式を入力するデータ作成者が使用するクライアントコンピュータ401と、化学構造式解析サーバ403とが、ネットワーク402を介して通信を行うよう構成されている。クライアントコンピュータ401は、液晶ディスプレイなどの表示手段、およびマウス、キーボードなどの入力手段を備える。本実施形態のネットワーク402は、本技術分野で知られたインターネット通信網を使用することができるが、これに限られず、専用の、あるいは汎用のネットワークを使用することができる。

10

【0020】

図5は、本発明の一実施形態にかかる化学構造式解析サーバのモジュール構成図である。化学構造式解析サーバ403は、解析対象の化学構造式データをクライアントコンピュータ401から受信する化学構造式データ受信部501、解析対象の化学構造式データから特定の部分構造を抽出する部分構造検索部502、構成原子の位置情報を取得する位置情報取得部503、取得した位置情報に基づいて、特定の角度、外積を計算する算出部504、単糖の回転構造を表すモードを判定するモード判定部505、単糖の水酸基の向きを判定する向き判定部506、解析結果に基づいて単糖を略号に変換するデータ変換部507、関連データを管理する記憶部508、およびクライアントコンピュータに解析結果を表示する画面を提供する解析結果出力部509を備える。

20

【0021】

記憶部508は、単糖の回転構造を表すモードに関する情報を格納するモードDB、単糖の略号に関する情報を格納する略号DBを管理する。以上、本実施形態のモジュール構成を説明したが、これは単なる例示であり、各モジュールをさらに機能毎に分解し、あるいは各モジュールの機能を統合した新たなモジュールを想定して実装することができる。

【0022】

(単糖を構成要素とする糖鎖の化学構造を認識するための全体の処理)

30

図6は、本発明の一実施形態にかかる単糖を構成要素とする糖鎖の化学構造を認識するための全体の処理を示すフローチャートである。処理S601において、化学構造式受信部501は、解析対象の化学構造式データとして本実施形態の一例である、例えば、図7に示すような糖鎖の化学構造式データを受信したものとする。この化学構造式データは、クライアントコンピュータ401を介してユーザから、または外部のシステムなどから受信することができる。

【0023】

処理S602において、部分構造検索部502は、受信した化学構造式データから6員環を抽出する。本実施形態では、10個の6員環が抽出される。処理S603以降については、処理S602で抽出された全ての6員環に対して、繰り返し処理が行われる。

40

【0024】

処理S603において、位置情報取得部503は、6員環の構成原子、および構成原子に結合する酸素原子(O原子)と炭素原子(C原子)について、順位規則に基づいてナンバリングを行う。図8に、本発明の一実施形態にかかるナンバリングした単糖を示す。さらに位置情報取得部503は、ナンバリングした原子の位置情報を取得する。

【0025】

処理S604において、算出部504は、6員環の内角をなす角とするベクトルの外積 $c1k \sim c5k$ 、および $o5k$ を算出する。具体的には、 $C1 \ O5$ と $C1 \ C2$ の外積($c1k$)、 $C2 \ C1$ と $C2 \ C3$ の外積($c2k$)、 $C3 \ C2$ と $C3 \ C4$ の外積($c3k$)、 $C4 \ C3$ と $C4 \ C5$ の外積($c4k$)、 $C5 \ C4$ と $C5 \ O5$ の外積($c5k$)および $O5 \ C5$ と $O5 \ C1$ の外積($o5k$)を算出する。

50

【 0 0 2 6 】

処理 S 6 0 5 において、算出部 5 0 4 は、以下の数式 1 で示される、処理 S 6 0 4 で算出した外積の SIGN 関数の和の絶対値を算出する。

【 0 0 2 7 】

【 数 1 】

$$| \text{SIGN}(c1k) + \text{SIGN}(c2k) + \text{SIGN}(c3k) + \text{SIGN}(c4k) + \text{SIGN}(c5k) + \text{SIGN}(o5k) |$$

【 0 0 2 8 】

処理 S 6 0 6 において、算出される絶対値が 6 でない場合、解析対象の単糖が配座表示で表記されていると判定され、処理 S 6 0 7 に進む。算出される絶対値が 6 である場合、解析対象の単糖が Mills 表示で表記されていると判定され、処理 S 6 0 8 に進む。本実施形態では、算出される絶対値は 2 なので、処理 S 6 0 7 に進む。

10

【 0 0 2 9 】

後述するように、処理 S 6 0 7 または処理 S 6 0 8 において、各表記方法に沿った解析処理を行い、処理 S 6 0 9 において、データ変換部 5 0 7 は、解析結果に基づいて化学構造式を変換する。

【 0 0 3 0 】

処理 S 6 1 0 において、処理 S 6 0 2 で抽出したすべての 6 員環について解析が終了したかどうか判定する。すべての 6 員環について解析が終了していない場合、処理 S 6 0 3 に戻って、処理 S 6 0 3 ~ 処理 S 6 1 0 を繰り返す。

20

【 0 0 3 1 】

すべての 6 員環について解析が終了している場合、処理 S 6 1 1 において、解析結果出力部 5 0 9 は、クライアントコンピュータに解析結果を表示する画面を提供し、終了する。

【 0 0 3 2 】

(単糖の回転構造を判定する処理)

図 9 は、本発明の一実施形態にかかる配座表示で表記された単糖の回転構造を判定する処理を示すフローチャートである。図 6 の処理 S 6 0 6 に続いて、図 8 に示す単糖の回転構造を判定するものとする。

【 0 0 3 3 】

処理 S 9 0 1 において、モード判定部 5 0 5 は、図 6 の処理 S 6 0 4 で算出した外積に基づいて、モード DB を使用して、基準角を特定する。モード DB は、図 1 5 に示すように、少なくとも外積の符号、基準角、鋭角フラグ、 4C_1 のモード、および 1C_4 のモードに関する属性を有する。基準角は、6 員環の内角のうちの 1 つであり、外積の符号の組み合わせにより特定することができる。例えば、本実施形態では、外積 c1k、外積 c3k、外積 c4k および外積 o5k の符号がマイナスであり、外積 c2k および外積 c5k の符号がプラスであることから、基準角は C1 となる。また、鋭角フラグは、基準角が鋭角であるか否かを示す。

30

【 0 0 3 4 】

処理 S 9 0 2 において、算出部 5 0 4 は基準角の角度を算出し、モード判定部 5 0 5 は、基準角が鋭角であるか否かに基づいて、モード DB のレコードを一意に特定する。図 1 5 に示すように、モード DB のレコードは、6 員環の内角をなす角とするベクトルの外積の符号および鋭角フラグの組み合わせにより一意に特定され、モードを 1C_4 のモードのうちの 1 つと 4C_1 のモードのうちの 1 つのいずれかに絞り込むことができる。例えば、本実施形態では、外積の符号により判定した基準角 C1 が鋭角であるため、モード判定部 5 0 5 は ID が 1 であるレコードを特定し、モードは ${}^4C_1+0$ または ${}^1C_4+180$ のいずれかに絞り込まれる。

40

【 0 0 3 5 】

処理 S 9 0 3 において、算出部 5 0 4 は Z 軸回転角を算出し、モード判定部 5 0 5 は、Z 軸回転角に基づいて、モードを判定する。Z 軸回転角は、処理 S 9 0 2 において特定したレコードの 4C_1 のモードに対する標準構造 (図 3 に示す) から Z 軸方向に回転した角度であ

50

る。具体的には、モード判定部505は、Z軸回転角が90度以下または270度以上である場合は 4C_1 のモードであると判定し、その他の場合は 1C_4 のモードであると判定し、暫定的に1つのモードに特定する。本実施形態では、特定したレコードの ${}^4C_1+0$ に対する標準構造からの回転が0度であるため、モードは暫定的に ${}^4C_1+0$ に特定される。

【0036】

さらに、処理S904において、モード判定部505は、くさびを考慮して最終的なモードを確定するために、6員環構成原子のうち、標準構造となるようにZ軸方向に回転した後y軸上で上に位置する3つの構成原子および下側の鋭角の頂点である構成原子を結ぶ3本の結合のいずれかがくさびで表記されているかどうか判定する。本実施形態では、y軸上で上に位置する3つの構成原子(C4原子、C5原子、O5原子)、および下側の鋭角の頂点である構成原子(C1原子)を結ぶ3本の結合はくさびで表記されていないので、単糖のモードは ${}^4C_1+0$ に確定し、処理は終了する。

【0037】

処理S904において、くさびで表記されていると判定されると、処理S905において、モード判定部505は、処理S903で判定したモードの4と1、および \pm を入れ替える。例えば、処理S903で ${}^4C_1+60$ と判定された単糖の対応する結合が、くさびで表記されている場合、処理S905によってモードが ${}^1C_4-60$ に確定する。

【0038】

(単糖の水酸基の向きを判定する処理)

続いて、単糖の水酸基の向きを判定する処理を、図10～図14を参照して説明する。処理S1001において、向き判定部506は、6員環構成原子について、O1原子～O4原子、またはC6原子が結合しているかどうか判定する。結合している場合、処理S1003以降の処理によって、水酸基の向きを判定する。いずれかの処理において水酸基の向きが判定されると、処理S1002に戻る(図示せず)。

【0039】

O1原子～O4原子、またはC6原子が結合していない場合、処理S1002において、向き判定部506は、すべての6員環構成原子の解析が終了したかどうか判定する。すべての6員環構成原子の解析が終了している場合、単糖の水酸基の向きを判定する処理を終了する。解析が終了していない場合、再び処理S1001に戻り、未解析の6員環構成原子について解析を行う。

【0040】

処理S1003において、算出部504は、図11に示すように、O1原子～O4原子またはC6原子のいずれかと接続しているC原子をC_connectとし、C_connectと隣接する2つの6員環構成原子をそれぞれC_neighbor1、C_neighbor2とした場合に、C_connect C_neighbor1とC_connect 0(またはC6)の外積o_1、およびC_connect C_neighbor2とC_connect 0(またはC6)の外積o_2を算出する。

【0041】

処理S1004において、向き判定部506は、C_neighbor1-C_connect-C_neighbor2で構成される内角ckが90度未満かどうか判定する。内角ckが90度未満である場合、処理S1005に進み、内角ckが90度以上である場合は図12に続く。本実施形態では、c1kおよびc4kの場合、内角ckが90度未満であると判定され、処理S1005に進む。

【0042】

処理S1005において、向き判定部506は、処理S1003で算出した外積o_1およびo_2の符号が共に正であるか、共に負であるか、またはその他であるかどうか判定する。外積の符号が共に正である場合、処理S1006に進み、向き判定部506は、C_connectの位置に基づいて最終的な水酸基の向きを判定する。外積の符号が共に負である場合、処理S1007に進み、向き判定部506は、C_connectの位置に基づいて最終的な水酸基の向きを判定する。

【0043】

ここで、C_connectの位置とは、標準構造となるようにZ軸方向に回転した際に、処理S

10

20

30

40

50

1 0 0 4で90度未満であると判定される2つの内角の頂点である6員環構成原子のうちC_connectが、他方に対してx軸上で右であるか、左であるかを表す。本実施形態では、c1kの頂点であるC1原子がx軸上で右であり、c4kの頂点であるC4原子がx軸上で左である。

【0044】

C1原子に結合する水酸基については、処理S1005で外積の符号が共に正であると判定され、処理S1006でC_connectが右に位置すると判定され、上向きであることがわかる。C4原子に結合する水酸基については、処理S1005で外積の符号が共に正であると判定され、処理S1006でC_connectが左に位置すると判定され、下向きであることがわかる。

【0045】

なお、外積の符号がその他である場合、水酸基の向きは上向きと下向きの中間であると判定される。

【0046】

ここで、図10、図12、図13を通じて判定される水酸基の向きは、「モードが 4C_1 のモードであり、かつ、処理S904で判定した結合がくさびで表記されていない単糖」または「モードが 1C_4 のモードであり、かつ、処理S904で判定した結合がくさびで表記されている単糖」に対する結果を示すものである。上述した条件を満たさない場合、水酸基の向きは上下が逆になる。

【0047】

図10の処理S1004で内角ckが90度以上であると判定されると図12に進み、処理S1201において、向き判定部506は、C_neighbor1-C_connect-C_neighbor2で構成される内角ckが180度未満かどうか判定する。内角ckが180度未満である場合、処理S1202に進み、内角ckが180度以上である場合は図13に続く。本実施形態では、c3kおよびo5kの場合、内角ckが180度未満であると判定され、処理S1202に進む。

【0048】

処理S1202において、算出部504は、図11に示すように、C_connectの対面に位置する6員環構成原子をC_diagonalとした場合に、C_connect C_neighbor1とC_connect C_neighbor2の内、傾きがC_connect C_diagonalに近い方と、C_connect 0(またはC6)の外積を算出する。ここで、傾きが近いとは、x軸に対する傾きの差が少ないことを表す。

【0049】

処理S1203において、向き判定部506は、算出した外積のSignが1であるか、-1であるか、または0であるかどうか判定する。Signが1である場合、処理S1204に進み、向き判定部506は、C_connectの位置に基づいて最終的な水酸基の向きを判定する。Signが-1である場合、処理S1205に進み、向き判定部506は、C_connectの位置に基づいて最終的な水酸基の向きを判定する。

【0050】

ここで、C_connectの位置とは、標準構造となるようにZ軸方向に回転した際に、処理S1201で180度未満であると判定される2つの内角の頂点である6員環構成原子のうちC_connectが、他方に対してx軸上で右であるか、左であるかを表す。本実施形態では、c3kの頂点であるC3原子がx軸上で左である。C3原子に結合する水酸基については、処理S1203でSignが-1であると判定され、処理S1205でC_connect左に位置すると判定され、上向きであることがわかる。

【0051】

なお、Signが0である場合、水酸基の向きは上向きと下向きの中間であると判定される。

【0052】

図12の処理S1201で内角ckが180度以上であると判定されると図13に進み、処理S1301において、向き判定部506は、C_connect 0(またはC6)の位置が領域

10

20

30

40

50

1 であるか、領域 2 であるか、またはその他であるかどうか判定する。ここで、領域 1 は、図 1 3 に示すように、6 員環の外側の領域を表す。領域 2 は、図 1 3 に示すように、C_neighbor1-C_connect 結合の延長線と C_neighbor2-C_connect 結合の延長線で構成される、領域 1 に対して対称の領域を表す。

【 0 0 5 3 】

C_connect 0 (または C6) の位置が領域 1 である場合、処理 S 1 3 0 2 に進み、向き判定部 5 0 6 は、C_connect の位置に基づいて最終的な水酸基の向きを判定する。C_connect

0 (または C6) の位置が領域 2 である場合、処理 S 1 3 0 3 に進み、向き判定部 5 0 6 は、C_connect の位置に基づいて最終的な水酸基の向きを判定する。

【 0 0 5 4 】

ここで、C_connect の位置とは、標準構造となるように Z 軸方向に回転した際に、図 1 2 の処理 S 1 2 0 1 で 1 8 0 度以上であると判定される 2 つの内角の頂点である 6 員環構成原子のうち C_connect が、他方に対して y 軸上で上であるか、下であるかを表す。本実施形態では、c2k の頂点である C2 原子が y 軸上で下であり、c5k の頂点である C5 原子が y 軸上で上である。C2 原子に結合する水酸基については、処理 S 1 3 0 1 で C_connect 0 の位置が領域 1 であると判定され、処理 S 1 3 0 2 で C_connect が下に位置すると判定され、下向きであることがわかる。C5 原子に結合する C6 原子については、処理 S 1 3 0 1 で C_connect C6 の位置が領域 1 であると判定され、処理 S 1 3 0 2 で C_connect が上に位置すると判定され、上向きであることがわかる。

【 0 0 5 5 】

なお、C_connect 0 (または C6) の位置がその他である場合、水酸基の向きは上向きと下向きの中間であると判定される。

【 0 0 5 6 】

(単糖の化学構造式を略号に変換する処理)

図 9、図 1 0、図 1 2、図 1 3 の処理を通じて、単糖のモードおよび水酸基の向きを判定することができる。単糖の化学構造式の略号を特定するために、データ変換部 5 0 7 は、変換のために必要な情報として、(1) C1 原子に結合する C0 原子が存在するか否か、(2) C6 原子が存在するか否か、(3) 置換基が存在するか否か、および置換基の種類などを化学構造式データから取得する。これらの情報は、モードおよび水酸基の向きの判定とは異なり、化学構造式を構成する特定の構成原子の情報に基づいて容易に判定することができるため、ここでは詳述しない。

【 0 0 5 7 】

変換に必要な情報を取得した後、データ変換部 5 0 7 は、C0 原子および C6 原子の存在の有無、並びに C2 原子 ~ C5 原子に結合する基の向きを示すシンボルデータを生成する。本実施形態では、C0 原子および C6 原子について、構成原子が存在しない場合に「 / 」、構成原子が存在する場合に「 | 」を用い、また、C2 原子 ~ C5 原子に結合する基について、向きが上向きの場合に「 + 」、下向きの場合に「 - 」、構成原子に結合する基が存在しない場合に「 * 」を用いることとする。例えば、図 8 に示す単糖の場合、「 C0 / C2 - C3 + C4 - C5 + C6 | 」が生成される。

【 0 0 5 8 】

データ変換部 5 0 7 は、記憶部 5 0 8 を介して略号 D B にアクセスを行い、作成したシンボルデータに対応する単糖の略号を取得する。略号 D B は、図 1 6 に示すように、少なくとも単糖の略号および略号に対応するシンボルに関する属性を有する。作成したシンボルデータに該当するレコードが存在しない場合、糖として認識し得る化学構造ではないことを表す。次に、データ変換部 5 0 7 は、C1 原子に結合する水酸基の向きによって / を判定し、さらに、置換基などを考慮して最終的な略号を特定する。

【 0 0 5 9 】

(解析結果を出力する処理)

すべての単糖について解析が終了した後、解析結果出力部 5 0 9 は、図 1 7 に示すような解析結果を、クライアントコンピュータ 4 0 1 に送信して表示させる。本実施形態では

10

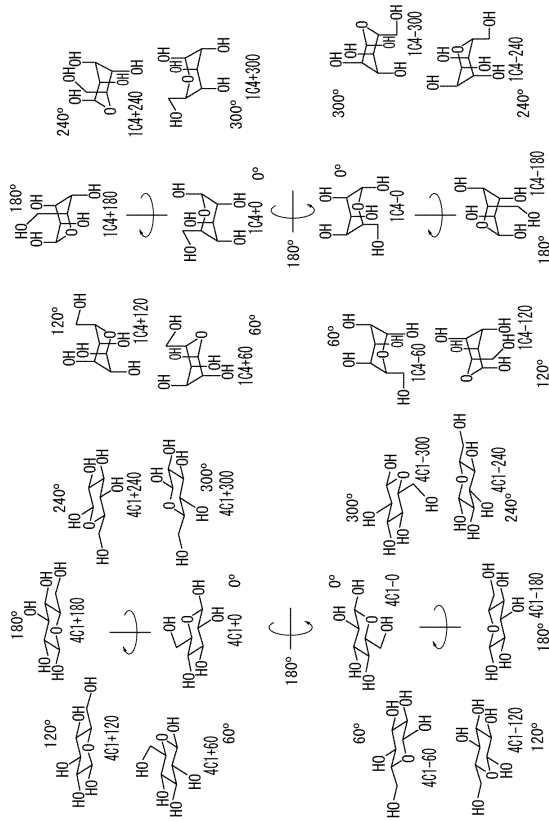
20

30

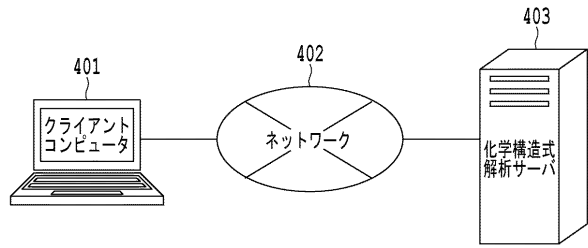
40

50

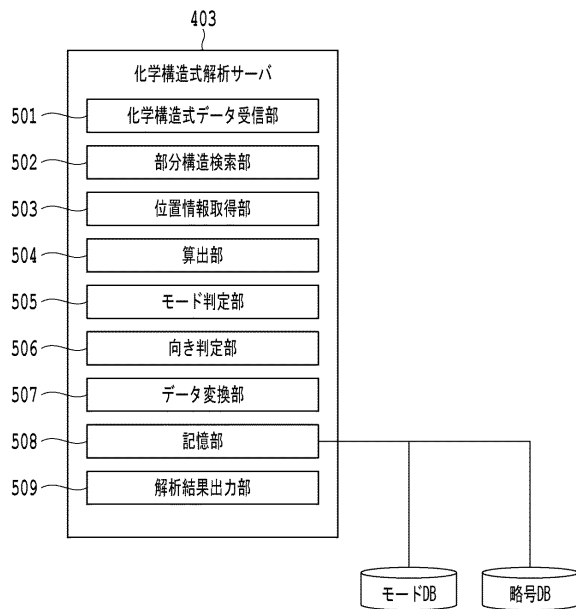
【図3】



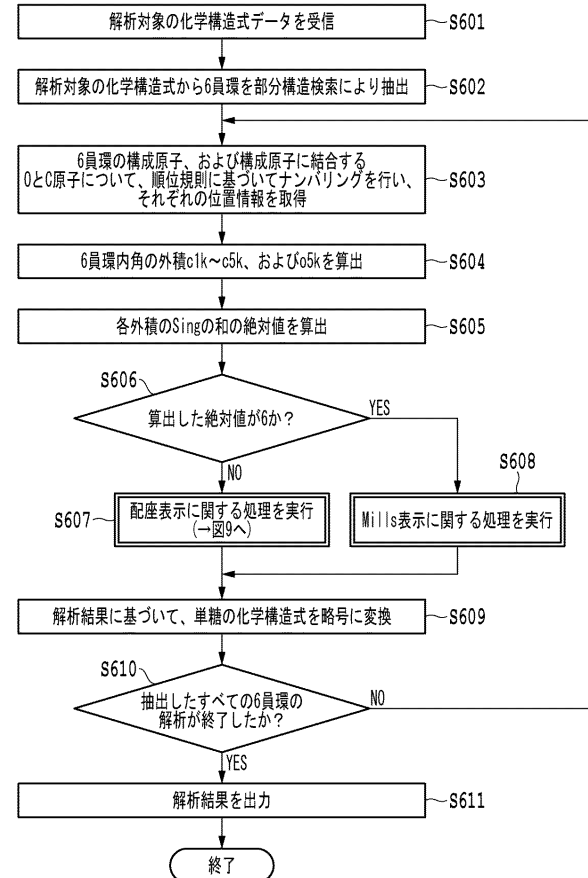
【図4】



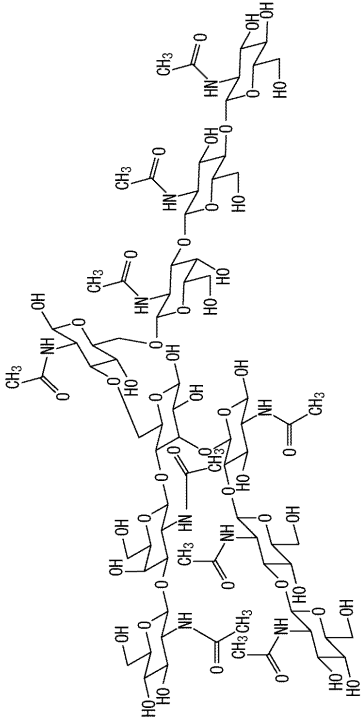
【図5】



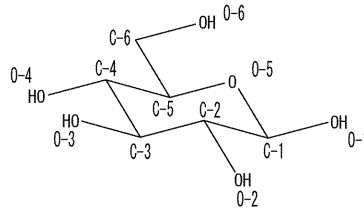
【図6】



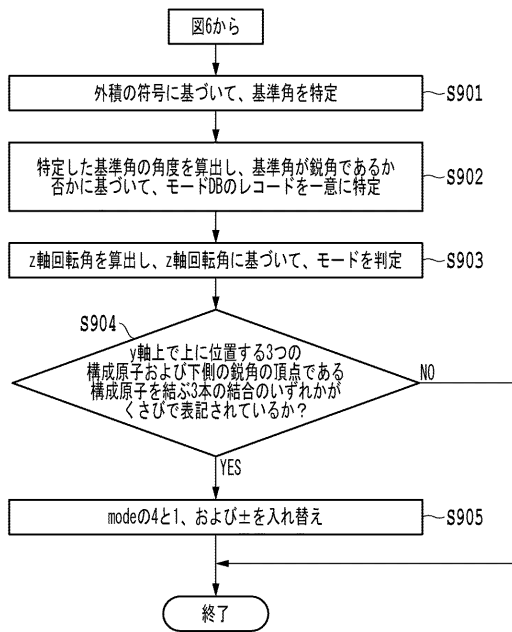
【図7】



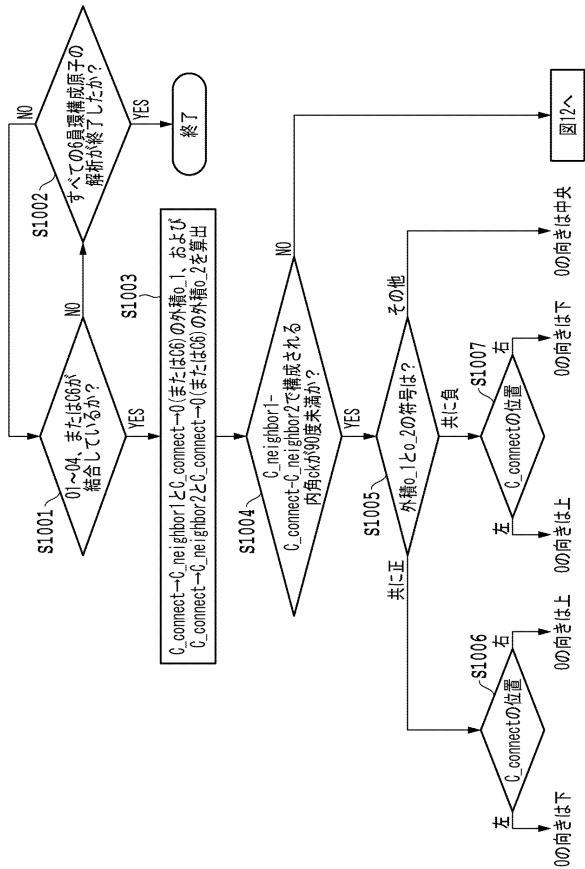
【図8】



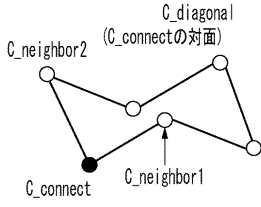
【図9】



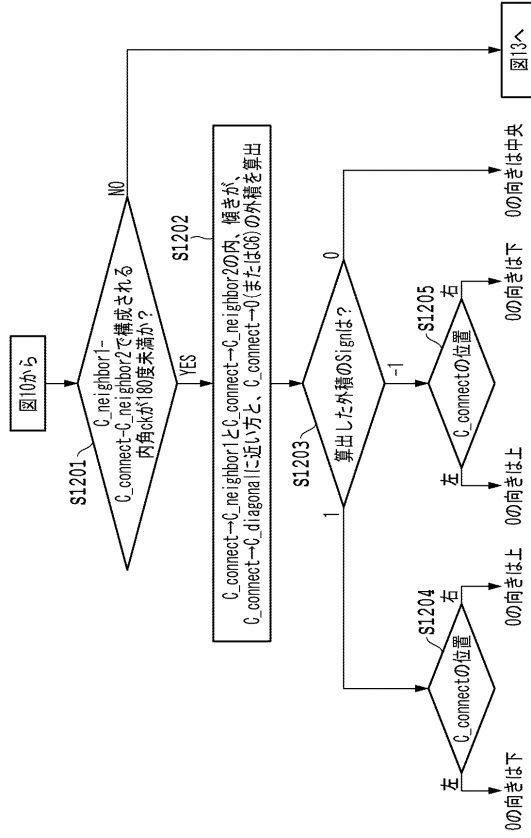
【図10】



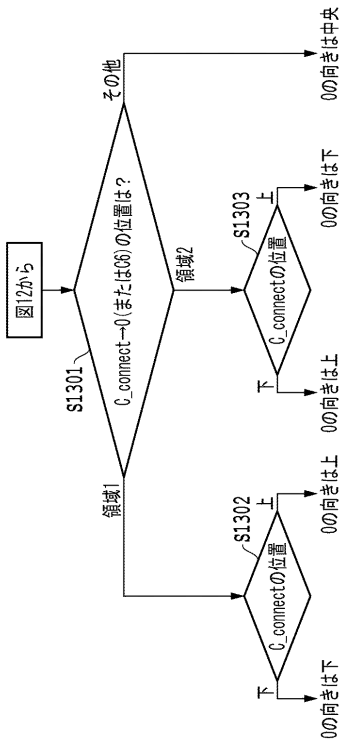
【図11】



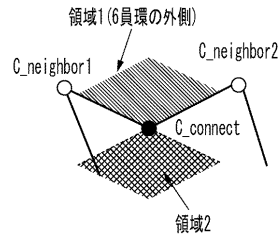
【図12】



【図13】



【図14】



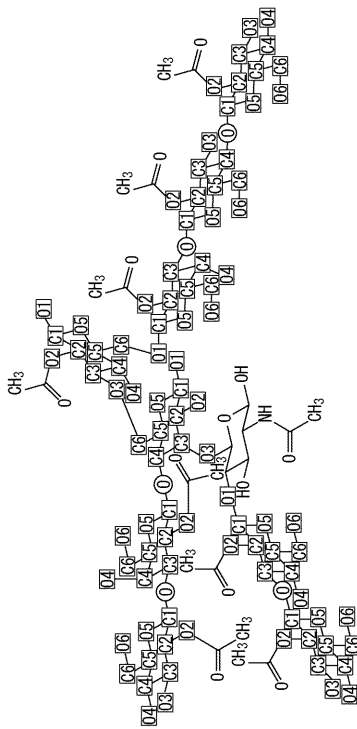
【 図 1 5 】

ID	外積の符号					基準角	鋭角フラグ	モード	
	c1k	c2k	c3k	c4k	c5k			4 ₀	1 ₀
1	-1	1	-1	-1	1	C1	YES	4 ₀ +0	1 ₀ +180
2	-1	1	-1	-1	1	C1	NO	4 ₀ +60	1 ₀ +240
3	-1	-1	1	-1	1	C1	NO	4 ₀ +120	1 ₀ +300
4	-1	-1	1	-1	1	C1	YES	4 ₀ +180	1 ₀ +0
5	1	-1	-1	1	-1	C2	NO	4 ₀ +240	1 ₀ +60
6	1	-1	-1	1	-1	C2	YES	4 ₀ +300	1 ₀ +120
7	-1	1	1	-1	1	C2	YES	4 ₀ -60	1 ₀ +240
8	-1	1	1	-1	1	C2	NO	4 ₀ -120	1 ₀ +300
9	1	1	-1	1	1	C1	YES	4 ₀ -180	1 ₀ +0
10	1	1	-1	1	1	C1	NO	4 ₀ -240	1 ₀ +60
11	1	-1	1	1	1	C1	NO	4 ₀ -300	1 ₀ +120
12	1	-1	1	1	1	C1	YES	4 ₀ -0	1 ₀ +180

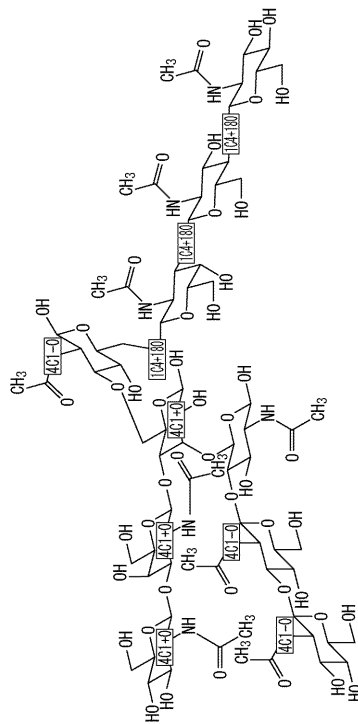
【 図 1 6 】

ID	略号	シンボル
1	D-Ara	C0/C2+C3-C4-C5*C6/
2	D-Lyx	C0/C2+C3+C4-C5*C6/
3	D-Rib	C0/C2-C3-C4-C5*C6/
4	D-Xyl	C0/C2-C3+C4-C5*C6/
5	D-Fuc	C0/C2-C3+C4+C5*C6/
6	D-Qui	C0/C2-C3+C4-C5+C6/
7	D-Rha	C0/C2+C3+C4-C5+C6/
8	D-Glc	C0/C2-C3+C4-C5+C6
9	D-Man	C0/C2+C3+C4-C5+C6
10	D-Gal	C0/C2-C3+C4+C5+C6
11	D-Ido	C0/C2+C3-C4+C5+C6
12	D-AlI	C0/C2-C3-C4-C5+C6
13	D-Alt	C0/C2+C3-C4-C5+C6
14	D-Gul	C0/C2-C3-C4+C5+C6
15	D-Tal	C0/C2+C3+C4+C5+C6
16	D-Fru	C0/C2+C3-C4-C5*C6/
17	D-Tag	C0/C2+C3+C4-C5*C6/
18	D-Sor	C0/C2-C3+C4-C5*C6/
19	D-Psi	C0/C2-C3-C4-C5*C6/
20	D-KDN	C0/C2*C3-C4+C5-C6
21	D-KDO	C0/C2*C3+C4+C5+C6
22	L-Ara	C0/C2-C3+C4+C5*C6/
23	L-Lyx	C0/C2-C3+C4+C5*C6/
24	L-Rib	C0/C2+C3+C4+C5*C6/
25	L-Xyl	C0/C2+C3-C4+C5*C6/
26	L-Fuc	C0/C2+C3-C4-C5-C6/
27	L-Qui	C0/C2+C3-C4+C5-C6/
28	L-Rha	C0/C2-C3-C4+C5-C6/
29	L-Glc	C0/C2+C3-C4+C5-C6
30	L-Man	C0/C2-C3-C4+C5-C6
31	L-Gal	C0/C2+C3-C4-C5-C6
32	L-Ido	C0/C2-C3+C4-C5-C6
33	L-AlI	C0/C2+C3+C4+C5-C6
34	L-Alt	C0/C2-C3+C4+C5-C6
35	L-Gul	C0/C2+C3+C4-C5-C6
36	L-Tal	C0/C2-C3-C4-C5-C6
37	L-Fru	C0/C2-C3+C4+C5*C6/
38	L-Tag	C0/C2-C3-C4+C5*C6/
39	L-Sor	C0/C2+C3-C4+C5*C6/
40	L-Psi	C0/C2+C3+C4+C5*C6/
41	L-KDN	C0/C2*C3+C4-C5+C6
42	L-KDO	C0/C2*C3-C4-C5-C6

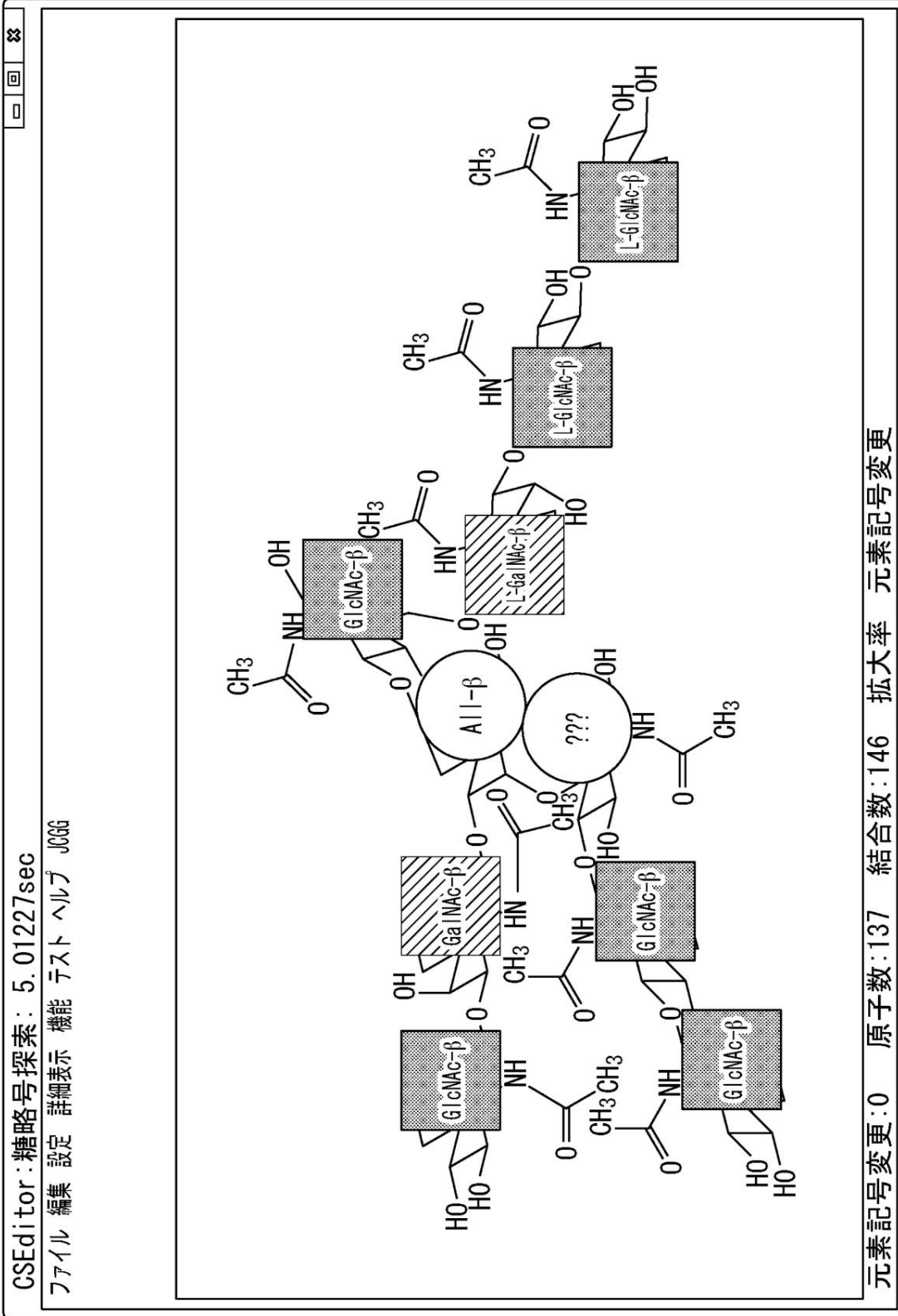
【 図 1 8 】



【 図 1 9 】

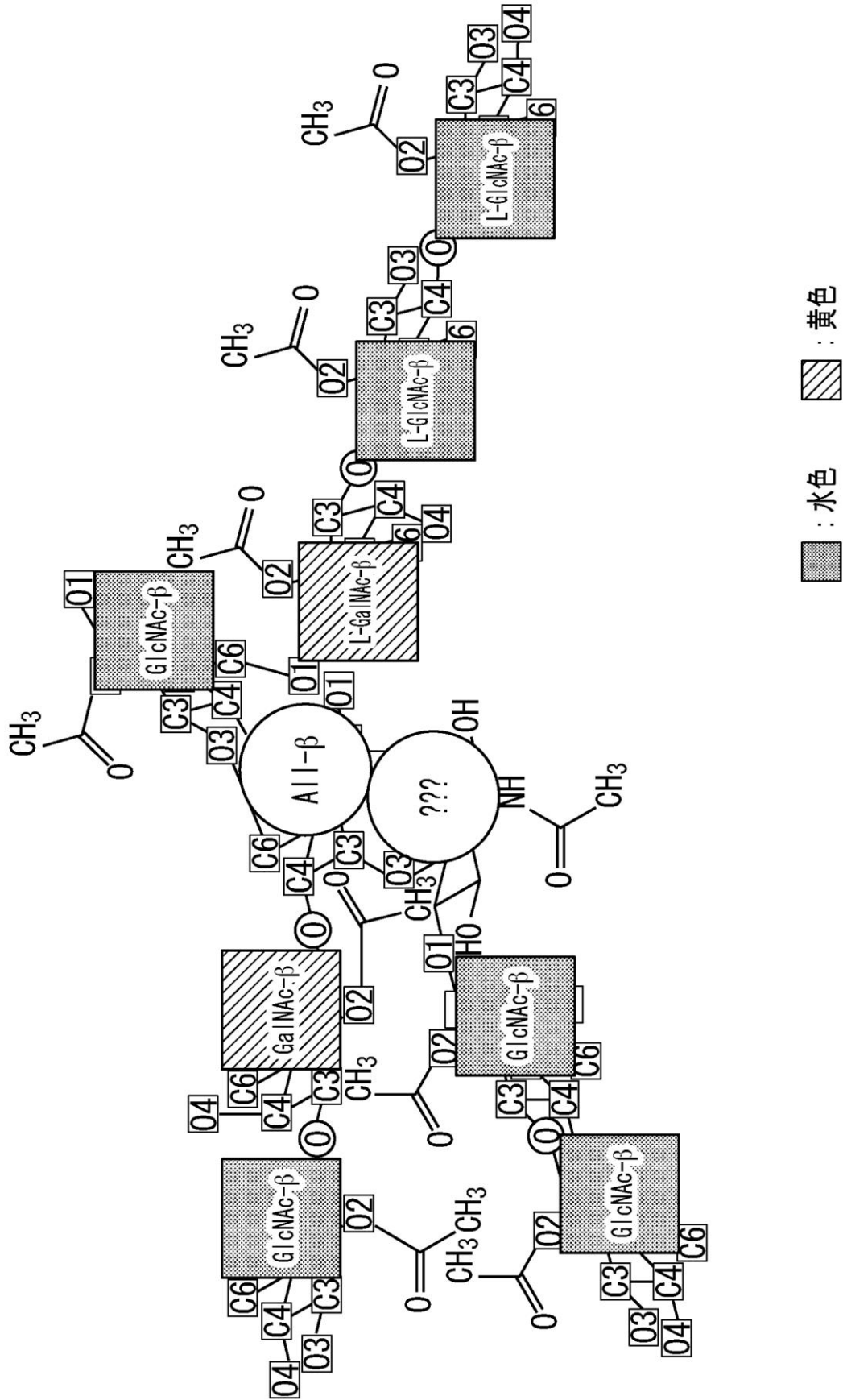


【 図 17 】



: 水色
 : 黄色
 : 黄色

【 図 20 】



フロントページの続き

- (56)参考文献 国際公開第2004/006149(WO, A1)
特開2001-114708(JP, A)
特開平06-249841(JP, A)
特表平05-501324(JP, A)
特開2000-251081(JP, A)
特開平10-031751(JP, A)
特開平09-128562(JP, A)
特表2003-502773(JP, A)
特表2003-531419(JP, A)
特開昭62-169290(JP, A)
特開2005-293296(JP, A)
国際公開第2006/112343(WO, A1)
特開2002-131255(JP, A)
特開2005-265697(JP, A)
日本コンピュータ化学会年会講演予稿集, 2007年, Vol.2007 春季 Page.33
情報化学討論会講演要旨集, 2008年, Vol.31st Page.59-62
野口研究所時報, 2012年 9月30日, No.55 Page.32-34
日本糖質学会年会要旨集, 2012年 8月31日, Vol.31st Page.87
Nucleic Acids Res, 2004年, Vol.32 No.Web Server Issue, Page.W267-W272

(58)調査した分野(Int.Cl., DB名)

G01N 33/00
CAplus/MEDLINE/BIOSIS(STN)
JSTPlus/JMEDPlus/JST7580(JDreamIII)