

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7626466号
(P7626466)

(45)発行日 令和7年2月4日(2025.2.4)

(24)登録日 令和7年1月27日(2025.1.27)

(51)国際特許分類

G 1 6 B	5/00 (2019.01)	F I	G 1 6 B	5/00
G 0 6 N	20/00 (2019.01)		G 0 6 N	20/00

請求項の数 13 (全26頁)

(21)出願番号	特願2022-108668(P2022-108668)	(73)特許権者	519115853 アヘッド・バイオコンピューティング株 式会社 神奈川県川崎市川崎区駅前本町11番地 2 川崎フロンティアビル4階
(22)出願日	令和4年7月5日(2022.7.5)	(74)代理人	110001519 弁理士法人太陽国際特許事務所
(62)分割の表示	特願2021-23750(P2021-23750)の 分割	(72)発明者	秋山 泰 東京都目黒区大岡山2丁目12番1号 国立大学法人東京工業大学内
原出願日	令和3年2月17日(2021.2.17)	(72)発明者	大上 雅史 東京都目黒区大岡山2丁目12番1号 国立大学法人東京工業大学内
(65)公開番号	特開2022-137148(P2022-137148 A)	(72)発明者	柳澤 溪甫 東京都目黒区大岡山2丁目12番1号 最終頁に続く
(43)公開日	令和4年9月21日(2022.9.21)		
審査請求日	令和6年2月19日(2024.2.19)		
(31)優先権主張番号	特願2020-189856(P2020-189856)		
(32)優先日	令和2年11月13日(2020.11.13)		
(33)優先権主張国・地域又は機関	日本国(JP)		

(54)【発明の名称】 情報処理装置、情報処理方法、情報処理プログラム、及び情報処理システム

(57)【特許請求の範囲】**【請求項1】**

端末から送信された要求信号を受信する受信部と、
 前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成する予測部と、
 前記予測部により生成された前記予測情報を前記端末へ送信する送信部と、

前記ペプチドを構成する要素の変更箇所の候補を含む設計支援情報を生成する設計支援部
と、

を含み、

前記設計支援部は、前記ペプチドを構成する複数の残基の各々について、前記予測情報を生成するための学習済みモデルのパラメータを解析することにより、前記残基が前記予測情報に影響を与える度合いを表す残基影響度スコアを計算し、

前記複数の残基の各々に対して計算された前記残基影響度スコアのうち所定閾値以上の残基影響度スコアの前記残基を特定し、

前記ペプチドの構造のうち、前記特定された前記残基を交換候補の残基として設定し、前記交換候補の残基を変更箇所の候補とし、

前記送信部は、前記設計支援情報を前記端末へ送信する、

情報処理装置。

【請求項2】

前記要求信号には、予測情報を生成するための予測モデルの選択情報が含まれており、
 前記予測部は、複数の前記予測モデルから前記選択情報に応じた予測モデルを選択し、

選択した前記予測モデルに対して、前記ペプチドを表すペプチド情報を入力することにより、前記ペプチドの前記予測情報を生成する、
請求項 1 に記載の情報処理装置。

【請求項 3】

前記要求信号にはユーザ ID が含まれており、
複数の予測モデルの各々は、ユーザ毎に予め用意されている予測モデルであり、
前記予測部は、前記複数の予測モデルのうちの前記ユーザ ID に対応する予測モデルに
対して、前記ペプチドを表すペプチド情報を入力することにより、前記ペプチドの前記予
測情報を生成する、
請求項 2 に記載の情報処理装置。

10

【請求項 4】

複数の予測モデルの各々は、ユーザ毎に予め用意された学習用データによって予め学習
されている学習済みモデルである、
請求項 3 に記載の情報処理装置。

【請求項 5】

前記要求信号にはユーザ ID が含まれており、
前記ユーザ ID に応じて、ユーザ毎に予め用意されている複数のデータベースのうちの
前記ユーザ ID に対応するデータベースに対して、前記予測情報を格納する処理部を更に
含む、
請求項 1 ~ 請求項 4 の何れか 1 項に記載の情報処理装置。

20

【請求項 6】

前記要求信号にはユーザ ID が含まれており、
前記ユーザ ID に応じて、ユーザ毎に予め用意されている複数のデータベースのうちの
前記ユーザ ID に対応するデータベースに対して、前記端末から送信されたデータを格納
する処理部を更に含む、
請求項 1 ~ 請求項 5 の何れか 1 項に記載の情報処理装置。

【請求項 7】

前記処理部は、前記端末から送信された、ペプチドの体内動態に関するデータを前記デ
ータベースへ格納する、
請求項 6 に記載の情報処理装置。

30

【請求項 8】

前記送信部は、前記予測部により予測された前記予測情報を、予測対象の前記ペプチド
とを対応付けて、前記端末へ送信する、
請求項 1 ~ 請求項 7 の何れか 1 項に記載の情報処理装置。

【請求項 9】

前記設計支援部は、前記ペプチドを構成する複数の残基のうちの少なくとも 1 つの残基
を予め定めた別の残基に置換した候補ペプチドを複数生成し、
前記予測部は、残基を置換する前の前記ペプチドの前記予測情報を、前記複数の前記候
補ペプチドの各々の前記予測情報を生成し、

前記設計支援部は、複数の前記候補ペプチドの各々について、残基を置換する前の前記
ペプチドの前記予測情報と、前記候補ペプチドの前記予測情報との間の差分を計算し、前
記差分が所定閾値以上である前記候補ペプチドを特定し、特定した前記候補ペプチドのう
ちの置換後の前記残基の箇所を特定し、

前記ペプチドの構造のうち、前記特定された箇所の残基を交換候補の残基として設定し
、前記交換候補の残基を変更箇所の候補とする、
請求項 1 ~ 請求項 8 の何れか 1 項に記載の情報処理装置。

【請求項 10】

前記端末から送信された、前記ペプチドの前記予測情報の教師データである、前記ペ
プチドの体内動態に関する実験データ、前記実験データを得た際の実験手法を表す情報、ユ
ーザが利用する予測モデル、及び前記ユーザにより生成された前記ペプチドの体内動態に

50

関する予測情報の少なくとも 1 つのデータが、前記情報処理装置が管理する記憶部に格納された場合、前記端末に付与されているユーザ ID に対して、前記データの提供に対する報酬を表すポイントを付与する報酬付与部を更に備える、

請求項 1 ~ 請求項 9 の何れか 1 項に記載の情報処理装置。

【請求項 1 1】

端末から送信された要求信号を受信し、

前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成し、

生成された前記予測情報を前記端末へ送信する、

処理をコンピュータが実行する情報処理方法であって、

前記ペプチドを構成する要素の変更箇所の候補を含む設計支援情報を生成し、

前記ペプチドを構成する複数の残基の各々について、前記予測情報を生成するための学習済みモデルのパラメータを解析することにより、前記残基が前記予測情報に影響を与えて
いる度合いを表す残基影響度スコアを計算し、

前記複数の残基の各々に対して計算された前記残基影響度スコアのうち所定閾値以上の残基影響度スコアの前記残基を特定し、

前記ペプチドの構造のうち、前記特定された前記残基を交換候補の残基として設定し、前記交換候補の残基を変更箇所の候補とし、

前記設計支援情報を前記端末へ送信する、

情報処理方法。

【請求項 1 2】

端末から送信された要求信号を受信し、

前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成し、

生成された前記予測情報を前記端末へ送信する、

処理をコンピュータに実行させるための情報処理プログラムであって、

前記ペプチドを構成する要素の変更箇所の候補を含む設計支援情報を生成し、

前記ペプチドを構成する複数の残基の各々について、前記予測情報を生成するための学習済みモデルのパラメータを解析することにより、前記残基が前記予測情報に影響を与えて
いる度合いを表す残基影響度スコアを計算し、

前記複数の残基の各々に対して計算された前記残基影響度スコアのうち所定閾値以上の残基影響度スコアの前記残基を特定し、

前記ペプチドの構造のうち、前記特定された前記残基を交換候補の残基として設定し、前記交換候補の残基を変更箇所の候補とし、

前記設計支援情報を前記端末へ送信する、

情報処理プログラム。

【請求項 1 3】

端末と請求項 1 ~ 請求項 1 0 の何れか 1 項に記載の情報処理装置とを含む情報処理システムであって、

前記端末は、要求信号を前記情報処理装置へ送信し、

前記情報処理装置は、前記端末から送信された前記要求信号を受信し、前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成し、生成された前記予測情報を前記端末へ送信し、

前記端末は、前記情報処理装置から送信された前記予測情報を表示部に表示させ、

前記情報処理装置は、ペプチドを構成する要素の変更箇所の候補を含む設計支援情報を生成し、前記設計支援情報を前記端末へ送信し、

前記端末は、前記設計支援情報を前記表示部に更に表示させる、

情報処理システム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示は、情報処理装置、情報処理方法、情報処理プログラム、及び情報処理システム

10

20

30

40

50

に関する。

【背景技術】

【0002】

特許文献1には、生体高分子の構造解析において、次元集合に含まれるすべての指標次元それぞれを座標軸とする多次元空間において複数の構造のクラスタリングを行い、いずれのクラスタにも含まれないはずれ値となった構造を初期構造とする分子動力学シミュレーションを行うことが開示されている（例えば、請求項4）。

【0003】

特許文献2には、タンパク質の立体構造を予測するタンパク質立体構造予測プログラムが開示されている。このタンパク質立体構造予測プログラムを実行するコンピュータは、タンパク質のアミノ酸配列を読み込み、2次構造情報を予測する。次に、このコンピュータは、2次構造情報に基づいてターンを形成するアミノ酸の個数を算出し、算出されたアミノ酸の個数と2次構造情報とによって、存在確率の高いターンのターン構造情報を取得してターンを予測再現し、タンパク質の立体構造を予測する。

10

【0004】

また、特許文献3には、新生抗原のセットのそれぞれのペプチド配列を、機械学習させた提示モデルに入力することにより、患者についての新生抗原のセットについての数値的提示尤度のセットをそれぞれの患者について生成する方法が開示されている（例えば、請求項1）。

【0005】

また、特許文献4には、コンピュータのプロセッサを使用してペプチドの数値ベクトルをディープラーニング提示モデルに入力して、新生抗原のセットについての提示尤度のセットを生成する方法が開示されている（例えば、請求項1）。

20

【先行技術文献】

【特許文献】

【0006】

【文献】特開2017-37378号公報

【文献】国際公開第2003/054743号

【文献】特表2020-523010号公報

【文献】特表2020-519246号公報

30

【発明の概要】

【発明が解決しようとする課題】

【0007】

ところで、近年、中分子医薬の一つとしてペプチド医薬が注目されている。しかし、ペプチドの体内動態については不明な点が多い。

【0008】

上記特許文献1～4に開示されている技術は、生体高分子の分子動力学シミュレーションを実行する技術、コンピュータによってタンパク質の立体構造を予測する技術、及び新生抗原に対して有効なペプチドを予測する技術であり、ペプチドの体内動態を予測するものではない。このため、上記特許文献1～4の技術は、ペプチドの体内動態を予測することはできない、という課題がある。

40

【0009】

本開示は上記事実を考慮して、ペプチドの体内動態を予測することを目的とする。

【課題を解決するための手段】

【0010】

本開示の第1の態様は、端末から送信された要求信号を受信する受信部と、前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成する予測部と、前記予測部により生成された前記予測情報を前記端末へ送信する送信部と、を含む情報処理装置である。

【0011】

本開示の第2の態様は、端末から送信された要求信号を受信し、前記要求信号に応答し

50

て、ペプチドの体内動態に関する予測情報を生成し、生成された前記予測情報を前記端末へ送信する、処理をコンピュータが実行する情報処理方法である。

【0012】

本開示の第3の態様は、端末から送信された要求信号を受信し、前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成し、生成された前記予測情報を前記端末へ送信する、処理をコンピュータに実行させるための情報処理プログラムである。

【0013】

本開示の第4の態様は、端末と情報処理装置とを含む情報処理システムであって、前記端末は、要求信号を前記情報処理装置へ送信し、前記情報処理装置は、前記端末から送信された前記要求信号を受信し、前記要求信号に応答して、ペプチドの体内動態に関する予測情報を生成し、生成された前記予測情報を前記端末へ送信し、前記端末は、前記情報処理装置から送信された前記予測情報を表示部に表示させる、情報処理システムである。
10

【発明の効果】

【0014】

本開示によれば、ペプチドの体内動態を予測することができる、という効果が得られる。

【図面の簡単な説明】

【0015】

【図1】第1実施形態に係る情報処理システムを示すブロック図である。

【図2】共通データベース142に格納されるデータの一例を説明するための図である。

【図3】ユーザ用データベース144に格納されるデータの一例を説明するための図である。
20

【図4】本実施形態の予測モデルを説明するための図である。

【図5】情報処理システムの各機器を実現するコンピュータを示す図である。

【図6】情報処理システムにおいて実行されるシーケンスである。

【図7】端末の表示部に表示される画面の一例を示す図である。

【図8】情報処理システムにおいて実行されるシーケンスである。

【図9】設計支援情報の一例を示す図である。

【図10】第2実施形態に係る情報処理システムを示すブロック図である。

【図11】第2実施形態のスコアを説明するための図である。

【図12】候補ペプチドを説明するための図である。
30

【図13】設計支援情報の一例を示す図である。

【図14】第4実施形態に係る情報処理システムを示すブロック図である。

【図15】報酬記憶部に格納されているテーブルの一例を示す図である。

【発明を実施するための形態】

【0016】

以下、図面を参照して本発明の実施形態を詳細に説明する。

【0017】

[第1実施形態]

【0018】

<情報処理システムの構成>

【0019】

図1は、第1実施形態に係る情報処理システム10の構成の一例を示すブロック図である。情報処理システム10は、機能的には、図1に示されるように、複数の端末12A, 12B, 12Cと、情報処理装置の一例であるサーバ14とを備えている。複数の端末12A, 12B, 12Cとサーバ14とは、例えば、インターネット等のネットワーク16を介して接続される。なお、以下では、1つの端末を指し示す場合には、単に、端末12と称する。なお、図1では、複数の端末12A, 12B, 12Cが例示されているが、情報処理システム10は、より多い又はより少ない端末を備えていても良い。
40

【0020】

(端末)

10

20

30

40

50

【 0 0 2 1 】

複数の端末 12A, 12B, 12C の各々は、複数の異なるユーザによって操作される。ユーザは、自身が分析したいペプチドを表すペプチド情報を、自身が操作する端末 12 へ入力する。そして、ユーザは、そのペプチド情報を後述するサーバ 14 へ送信するよう に端末 12 を操作する。端末 12 は、ユーザから入力されたペプチド情報を、ネットワーク 16 を介してサーバ 14 へ送信する。なお、ペプチド情報には、ペプチドの構造式、ペプチドの S M I L E S 表記、ペプチドの一次構造、ペプチドの二次構造、ペプチドの三次構造、及びペプチドの四次構造のうちの少なくとも 1 つの情報が含まれている。

【 0 0 2 2 】

(サーバ)

10

【 0 0 2 3 】

サーバ 14 は、図 1 に示されるように、受信部 140 と、共通データベース 142 と、複数のユーザ用データベース 144A, 144B, 144C と、共通予測モデル記憶部 145 と、複数のユーザ用予測モデル記憶部 146A, 146B, 146C と、予測部 148 と、処理部 150 と、送信部 152 とを備えている。なお、以下では、1 つのユーザ用データベースを指し示す場合には、単に、ユーザ用データベース 144 と称する。また、1 つのユーザ用予測モデル記憶部を指し示す場合には、単に、ユーザ用予測モデル記憶部 146 と称する。なお、図 1 では、複数のユーザ用データベース 144A, 144B, 144C が例示されているが、サーバ 14 は、より多い又はより少ないユーザ用データベースを備えていても良い。また、同様に、図 1 では、複数のユーザ用予測モデル記憶部 146A, 146B, 146C が例示されているが、サーバ 14 は、より多い又はより少ないユーザ用予測モデル記憶部を備えていても良い。

20

【 0 0 2 4 】

受信部 140 は、端末 12 から送信された、端末を識別するためのユーザ ID と、ペプチド情報と、後述する予測モデルの選択情報を含む要求信号を受信する。

【 0 0 2 5 】

共通データベース 142 には、サーバ 14 を管理する管理者のデータと、端末 12 を操作するユーザのデータとが格納される。図 2 に、共通データベース 142 に格納されるデータの一例を示す。図 2 に示されるように、各データにはデータ ID が付与されている。このデータ ID は、データを識別するための識別情報である。また、データ ID は、データが管理者由来であるのかユーザ由来であるのかを識別するための情報である。

30

【 0 0 2 6 】

図 2 に示されるように、データ ID 「00001」～「00004」等に対応するデータは管理者由来のデータであり、ペプチド情報と、ペプチド情報の特徴情報と、ペプチドの膜透過性の実験値と、ペプチドの膜透過性の実験値を得た際の実験手法と、ペプチドの体内持続性の実験値と、ペプチドの体内持続性の実験値を得た際の実験手法と、注釈と、後述する予測部 148 により生成された予測情報と、が対応付けられて格納されている。なお、ペプチドの体内持続性の値は、例えば、ペプチドの血漿タンパク質結合率(Plasma Protein Binding)である。注釈は、そのデータがどの様にして得られたのかを表す情報等である。なお、注釈の欄には、当該データの信頼度を表すスコアが格納されていてもよい。データ ID 「00001」～「00004」等に対応するデータは、サーバ 14 の管理者によって予め共通データベース 142 に格納される。

40

【 0 0 2 7 】

一方、図 2 に示されるように、データ ID 「A_00001」、「A_00002」、「B_00001」等に対応するデータはユーザ由来のデータであり、ペプチド情報は対応付けられておらず、ペプチド情報の特徴情報と、ペプチドの膜透過性の実験値と、ペプチドの膜透過性の実験値を得た際の実験手法と、ペプチドの体内持続性の実験値と、ペプチドの体内持続性の実験値を得た際の実験手法と、注釈と、後述する予測部 148 により生成された予測情報と、が対応付けられて格納されている。データ ID 「A_00001」、「A_00002」、「B_00001」等に対応するデータは、ユーザによる承諾を経て共通データベース 142 に

50

格納される。

【0028】

複数のユーザ用データベース 144A, 144B, 144C の各々には、ユーザ由來のデータが格納される。図 3 に、ユーザ用データベース 144 に格納されるデータの一例を示す。図 3 に示されるように、ユーザ用データベース 144 には、ペプチド情報と、ペプチドの特徴情報と、ペプチドの膜透過性の実験値と、ペプチドの膜透過性の実験値を得た際の実験手法と、ペプチドの体内持続性の実験値と、ペプチドの体内持続性の実験値を得た際の実験手法と、注釈と、後述する予測部 148 により生成された予測情報と、が対応付けられて格納されている。なお、ユーザ用データベース 144 に格納される各データについては、その一部又は全てが欠損していてもよい。

10

【0029】

上述したように、共通データベース 142 にはユーザ由來のデータのうちのペプチド情報は格納されないが、ユーザ用データベース 144 にはペプチド情報が格納される。ペプチド情報は秘匿性が高い情報である。このため、ユーザは、他のユーザ及び管理者のデータも格納される共通データベース 142 に、自らのペプチド情報が格納されることに関しては抵抗があると考えられる。

【0030】

そこで、本実施形態では、共通データベース 142 とは異なるユーザ用データベース 144 を設け、ユーザ用データベース 144 にのみユーザのペプチド情報を格納する。一方、共通データベース 142 には、ユーザのペプチド情報は格納されない。

20

【0031】

なお、ペプチド情報から得られる特徴情報は、ペプチドのさまざまな局所又は全体の構成に注目して、可逆的又は不可逆的な演算処理により計算された多数の数値を集めて構成されるベクトル形式の情報であり、ペプチド情報に比べれば秘匿性は低い。このため、特徴情報に関しては共通データベース 142 に格納してもよいと考えるユーザも存在すると考えられる。

【0032】

そこで、本実施形態では、特徴情報に関しては共通データベース 142 に格納する。この共通データベース 142 に格納された情報は、ペプチドの膜透過性又は体内持続性を予測するための予測モデルの生成等に用いられる。予測モデルについては後述する。また、共通データベース 142 に格納された情報は、サーバ 14 を管理する管理者によって利用可能な情報として取り扱われる。なお、共通データベース 142 に格納された情報は、管理者のみならず、全てのユーザから利用可能な情報としても取り扱われる。なお、共通データベース 142 に格納された情報のうちの一部については、ユーザは利用することができない情報として設定されていてもよい。

30

【0033】

共通予測モデル記憶部 145 には、予測モデルが格納される。

【0034】

複数のユーザ用予測モデル記憶部 146A, 146B, 146C の各々には、ユーザ用の予測モデルが格納される。

40

【0035】

図 4 に、本実施形態の予測モデルの一例を示す。図 4 には、予測モデルの一例である、膜透過性予測用学習済みモデル M1 と、体内持続性予測用学習済みモデル M2 と、ペプチドの体内動態に関する分子動力学シミュレーションを実施するためのシミュレーションモデル M3 が示されている。図 4 に示されるように、ペプチド情報から得られる特徴情報が、膜透過性予測用学習済みモデル M1 に入力されると、そのペプチドに対応する膜透過性の予測値がモデルから出力される。また、ペプチド情報から得られる特徴情報が、体内持続性予測用学習済みモデル M2 に入力されると、そのペプチドに対応する体内持続性の予測値がモデルから出力される。また、ペプチド情報がシミュレーションモデル M3 に入力されると、そのペプチドの体内動態に関する分子動力学シミュレーションのトラジェクト

50

リーデータがシミュレーションモデル M 3 から出力される。なお、シミュレーションモデル M 3 から出力されるデータはトراجيكتريーデータであるが、トراجيكتريーデータに対して統計的分析等を行うことにより、ペプチドの膜透過性又は体内持続性の予測値が算出され、それらの値がシミュレーションモデル M 3 から出力される場合もある。

【 0 0 3 6 】

なお、これらの学習済みモデルは、共通データベース 1 4 2 又はユーザ用データベース 1 4 4 に格納されているデータの全て又は一部を学習用データとして生成される。具体的には、特徴情報に対応付けられた実験値が教師データとして用いられ、教師あり機械学習によって学習済みモデルが生成される。これらの学習済みモデルは、例えば、ニューラルネットワーク（例えば、深層学習により学習されるディープニューラルネットワークも含む）、及びサポートベクターマシン等によって実現される。なお、学習済みモデルは、これらのような機械学習モデルに限らず、ロジスティック回帰等の手法によって実現されてもよい。

10

【 0 0 3 7 】

なお、共通予測モデル記憶部 1 4 5 には、共通データベース 1 4 2 に格納されているデータを学習用データとして生成された学習済みモデルと、シミュレーションモデルとが予測モデルとして格納される。一方、複数のユーザ用予測モデル記憶部 1 4 6 A , 1 4 6 B , 1 4 6 C の各々には、ユーザ用データベース 1 4 4 に格納されているデータを学習用データとして生成された学習済みモデルと、ユーザ毎に用意されたシミュレーションモデルとが予測モデルとして格納される。

20

【 0 0 3 8 】

ペプチドの挙動を予測する際、あるユーザ A はある種類のペプチドの挙動を解析したいと考え、別のユーザ B は別の種類のペプチドの挙動を解析したい、と考えることが想定される。また、ユーザ毎に、ユーザ用データベース 1 4 4 に格納されるデータは異なるため、予測モデルの最適な構成方法が異なってくることが考えられる。そこで、本実施形態では、ペプチドの体内動態を予測する予測モデルをユーザ毎に用意する。

【 0 0 3 9 】

具体的には、ユーザ用予測モデル記憶部 1 4 6 A には、ユーザ A 用の予測モデルが格納されている。また、ユーザ用予測モデル記憶部 1 4 6 B には、ユーザ B 用の予測モデルが格納されている。ユーザ用予測モデル記憶部 1 4 6 C には、ユーザ C 用の予測モデルが格納されている。これにより、ユーザが解析したいペプチドに応じて予測モデルが用意され、ユーザはその予測モデルを利用してペプチドの予測情報を得ることができる。

30

【 0 0 4 0 】

一方、自らのデータから生成された固有の予測モデルではなく、一般の予測モデルを用いてペプチドの体内動態を予測したいといったユーザの存在も考えられる。このため、本実施形態では、共通予測モデル記憶部 1 4 5 に、共通データベース 1 4 2 に格納されているデータを学習用データとして生成された学習済みモデルと、標準的なシミュレーションモデルとが格納される。

【 0 0 4 1 】

このため、予測部 1 4 8 は、端末 1 2 から送信された要求信号に含まれる予測モデルの選択情報が、ユーザ用予測モデル記憶部 1 4 6 に格納されている予測モデルを表している場合には、受信部 1 4 0 により受信されたユーザ ID に基づいて、当該ユーザ ID に対応するユーザ用予測モデル記憶部 1 4 6 から、予測モデルを読み出す。

40

【 0 0 4 2 】

一方、予測部 1 4 8 は、端末 1 2 から送信された要求信号に含まれる予測モデルの選択情報が、共通予測モデル記憶部 1 4 5 に格納されている予測モデルを表している場合には、共通予測モデル記憶部 1 4 5 から予測モデルを読み出す。

【 0 0 4 3 】

そして、予測部 1 4 8 は、受信部 1 4 0 により受信されたペプチド情報を既知の手法により特徴情報に変換する。次に、予測部 1 4 8 は、ペプチド情報及び特徴情報の少なくと

50

も一方を読み出した予測モデルへ入力することにより、ペプチド情報に対応する予測情報を生成する。

【0044】

例えば、予測モデルが膜透過性予測用学習済みモデルM1である場合には、膜透過性の予測値が予測情報として生成される。予測モデルが体内持続性予測用学習済みモデルM2である場合には、体内持続性の予測値が予測情報として生成される。また、予測モデルがシミュレーションモデルM3である場合には、予測部148は、受信部140により受信されたペプチド情報をシミュレーションモデルM3へ入力することにより、既知の分子動力学シミュレーション手法によって、体内におけるペプチドの挙動をシミュレーションする。これにより、ペプチドの体内動態に関する予測情報が生成される。

10

【0045】

なお、共通予測モデル記憶部145及びユーザ用予測モデル記憶部146には、同種の予測情報を生成する予測モデルが複数格納されている場合もある。例えば、ユーザ用予測モデル記憶部146には、複数の膜透過性予測用学習済みモデルが格納されている場合もある。また、ユーザ用予測モデル記憶部146には、複数のシミュレーションモデルが格納されている場合もある。

【0046】

例えば、ユーザは、あるペプチド情報について、ユーザ用予測モデル記憶部146に格納されている自らの膜透過性予測用学習済みモデルX及び膜透過性予測用学習済みモデルYと、共通予測モデル記憶部145に格納されている膜透過性予測用学習済みモデルZ及び膜透過性予測用学習済みモデルWの各々を用いて、ペプチドの予測情報を生成したいといった場合も想定される。なお、体内持続性予測用学習済みモデル及びシミュレーションモデルについても同様のことが想定される。

20

【0047】

このような場合、ユーザは、上記の同種の複数の予測モデルを用いて、1つのペプチド情報に対して複数の予測情報を生成することも可能である。この場合には、例えば、ユーザは、同種の複数の予測モデルによって生成された予測情報の各々を確認し、どの予測情報が妥当であるのかといった検討も可能である。または、ユーザは、同種の複数の予測モデルによって生成された予測情報の各々を平均するなどして、妥当な予測情報を得ることも可能である。なお、この場合には、予測対象のペプチド情報から特徴情報を生成する処理は一度でよく、生成された特徴情報を複数の予測モデルへ入力することにより、複数の予測情報が生成される。

30

【0048】

なお、共通予測モデル記憶部145には、予測情報を生成する際の各種パラメータも格納されている。これらの各種パラメータは、予測モデルを用いて予測情報を生成する際に用いられる。また、複数のユーザ用予測モデル記憶部146A, 146B, 146Cの各々にも、予測情報を生成する際の各種パラメータが格納されており、これらの各種パラメータはユーザ毎に異なるものとなる。予測情報を生成する際の各種パラメータがユーザ毎に異なることにより、ユーザの要望に応じた適切な予測情報を生成することができる。なお、共通予測モデル記憶部145又はユーザ用予測モデル記憶部146には、同種の各種パラメータが複数格納されている場合もある。ユーザは、ペプチドの予測情報を得る際に、これらの各種パラメータから適切と思われるパラメータを選択して、サーバ14にペプチドの予測情報を生成させる。

40

【0049】

処理部150は、受信部140により受信されたペプチド情報と、予測部148により得られた特徴情報と予測情報を対応付けて、ユーザIDに対応するユーザ用データベース144へ格納する。

【0050】

送信部152は、予測部148により生成された予測情報を、受信部140により受信したユーザIDに対応する端末12へ送信する。

50

【 0 0 5 1 】

端末 1 2 は、サーバ 1 4 から送信された予測情報を受信し、その予測情報を表示部（図示省略）へ表示させる。

【 0 0 5 2 】

端末 1 2 及びサーバ 1 4 は、例えば、図 5 に示すようなコンピュータ 5 0 によって実現することができる。端末 1 2 及びサーバ 1 4 を実現するコンピュータ 5 0 は、C P U 5 1 、一時記憶領域としてのメモリ 5 2 、及び不揮発性の記憶部 5 3 を備える。また、コンピュータは、入出力装置等（図示省略）が接続される入出力 interface (I / F) 5 4 、及び記録媒体 5 9 に対するデータの読み込み及び書き込みを制御する read/write (R / W) 部 5 5 を備える。また、コンピュータは、インターネット等のネットワークに接続されるネットワーク I / F 5 6 を備える。C P U 5 1 、メモリ 5 2 、記憶部 5 3 、入出力 I / F 5 4 、R / W 部 5 5 、及びネットワーク I / F 5 6 は、バス 5 7 を介して互いに接続される。10

【 0 0 5 3 】

記憶部 5 3 は、Hard Disk Drive (H D D) 、Solid State Drive (S S D) 、フラッシュメモリ等によって実現できる。記憶媒体としての記憶部 5 3 には、コンピュータを機能させるためのプログラムが記憶されている。C P U 5 1 は、プログラムを記憶部 5 3 から読み出してメモリ 5 2 に展開し、プログラムが有するプロセスを順次実行する。

【 0 0 5 4 】

次に、実施形態の情報処理システム 1 0 の作用について説明する。20

【 0 0 5 5 】

端末 1 2 を操作するユーザが、予測対象のペプチド情報を端末 1 2 に入力し、そのペプチド情報をサーバ 1 4 へ送信するような操作を実行すると、図 6 に示されるようなシーケンスが実行される。具体的には、端末 1 2 を操作するユーザが、予測対象のペプチド情報を端末 1 2 に入力し、そのペプチド情報をサーバ 1 4 へ送信するとともに、ペプチドの挙動を学習済みモデルによって予測する旨の指示信号をサーバ 1 4 へ送信した場合には、図 6 に示されるようなシーケンスが実行される。なお、図 6 のシーケンスでは、1 つの予測モデルが選択され当該予測モデルによって予測情報が生成される場合を例に説明する。なお、前述したように複数の予測モデルの各々によって予測情報が生成される場合もある。この場合には、複数の予測モデルを指定する情報が、後述する予測モデルの選択情報に含まれる。30

【 0 0 5 6 】

ステップ S 1 0 0 において、端末 1 2 は、ユーザによって入力されたペプチド情報と自身のユーザ I D を含む要求信号をサーバ 1 4 へ送信する。なお、要求信号には、ユーザ用予測モデル記憶部 1 4 6 に格納されているユーザ自らの予測モデルを用いて予測情報を生成するのか、又は、共通予測モデル記憶部 1 4 5 に格納されている予測モデルを用いて予測情報を生成するのかを表す予測モデルの選択情報も含まれている。

【 0 0 5 7 】

ステップ S 1 0 2 において、サーバ 1 4 の受信部 1 4 0 は、上記ステップ S 1 0 0 で端末 1 2 から送信された要求信号を受信する。40

【 0 0 5 8 】

ステップ S 1 0 4 において、サーバ 1 4 の予測部 1 4 8 は、上記ステップ S 1 0 2 で受信した要求信号に含まれるユーザ I D から、端末 1 2 のユーザを特定する。

【 0 0 5 9 】

ステップ S 1 0 6 において、サーバ 1 4 の予測部 1 4 8 は、上記ステップ S 1 0 2 で受信した要求信号に含まれる予測モデルの選択情報に基づいて、予測情報の生成に用いる予測モデルが何れであるのかを判定する。選択情報がユーザ自らの予測モデルを利用するこことを表している場合には、サーバ 1 4 の予測部 1 4 8 は、ユーザ I D に対応するユーザ用予測モデル記憶部 1 4 6 から予測モデルを読み出す。一方、選択情報が共通予測モデル記憶部 1 4 5 の予測モデルを利用することを表している場合には、サーバ 1 4 の予測部 1 450

8は、共通予測モデル記憶部145から予測モデルを読み出す。

【0060】

ステップS108において、サーバ14の予測部148は、上記ステップS102で受信した要求信号に含まれるペプチド情報を特徴情報へ変換する。

【0061】

ステップS110において、サーバ14の予測部148は、上記ステップS106で読み出した予測モデルに対して、上記ステップS108で得られた特徴情報を入力することにより、ペプチドの予測情報を生成する。

【0062】

ステップS112において、サーバ14の送信部152は、上記ステップS110で得られた予測情報を端末12へ送信する。なお、送信部152は、上記ステップS110で得られた予測情報と、予測対象のペプチド情報（例えば、ペプチドの構造式等）とを対応付けて、端末12へ送信するようにしてよい。

10

【0063】

ステップS114において、端末12は上記ステップS112で送信された予測情報を受信する。

【0064】

ステップS116において、端末12は、上記ステップS114で受信した予測情報を表示部（図示省略）に表示させる。

20

【0065】

この場合には、図7に示されるように、端末12の表示部（図示省略）には、ペプチドの構造式を表す情報、ペプチドのSMILE表記、及び予測情報が表示される。図7は、ペプチドの予測情報と、予測対象であるペプチドの構造式とが対応付けられた情報の一例である。

【0066】

なお、予測対象のペプチドが1つではなく複数あってもよい。この場合には、複数のペプチドの各々について、ペプチド情報（例えば、ペプチドの構造式を表す情報）、ペプチドのSMILE表記、及び予測情報が端末12の表示部（図示省略）に表示される。なお、この場合には、上記ステップS106～ステップS110の処理が複数のペプチド情報の各々について繰り返され、複数のペプチド情報の各々についての予測情報が端末12の表示部（図示省略）に表示される。また、上記ステップS106で読み出される予測モデルが複数である場合には、一つのペプチド情報について複数の予測モデルにより予測情報が生成される。

30

【0067】

これにより、ユーザは自らの端末12を操作するのみで、ペプチドの体内動態に関する予測情報を得ることができる。

【0068】

ステップS118において、サーバ14の処理部150は、上記ステップS102で受信した要求信号に含まれるペプチド情報と、上記ステップS108で得られた特徴情報と、上記ステップS110で生成された予測情報を対応付けて、ユーザIDに対応するユーザ用データベース144へ格納する。

40

【0069】

ステップS120において、サーバ14の処理部150は、上記ステップS108で得られた特徴情報と上記ステップS110で生成された予測情報を共通データベース142に格納する。

【0070】

なお、端末12を操作するユーザが、予測対象のペプチド情報を端末12に入力し、そのペプチド情報をサーバ14へ送信するとともに、ペプチドの挙動を分子動力学シミュレーションによって予測する旨の指示信号をサーバ14へ送信した場合には、図8に示されるようなシーケンスが実行される。

50

【0071】

図8に示すステップS100～ステップS106は、図6と同様に実行される。

【0072】

ステップS210において、サーバ14の予測部148は、ステップS102で受信されたペプチド情報を、予測モデルとしてのシミュレーションモデルへ入力することにより、既知の分子動力学シミュレーション手法によって、体内におけるペプチドの挙動をシミュレーションする。これにより、ペプチドの体内動態に関する予測情報が生成される。

【0073】

ステップS212において、サーバ14の予測部148は、上記ステップS210で生成された予測情報を端末12へ送信する。

10

【0074】

ステップS214において、端末12は、上記ステップS212でサーバ14から送信された予測情報を受信する。

【0075】

ステップS216において、端末12は、上記ステップS214で受信した予測情報を表示部(図示省略)へ表示させる。

【0076】

ステップS218において、サーバ14の処理部150は、上記ステップS102で受信した要求信号に含まれるペプチド情報と、上記ステップS210で生成された予測情報とを対応付けて、ユーザIDに対応するユーザ用データベース144へ格納する。

20

【0077】

なお、この場合のペプチドの予測情報は、例えば、ペプチド分子の運動を時系列的に模擬したトラジェクトリーデータ及びトラジェクトリーデータに対して統計的分析等を行うことにより得られた膜透過性又は体内持続性の予測値等の少なくとも一つの情報である。なお、このトラジェクトリーデータに基づき、ペプチドの膜透過性又は体内持続性等に関わる挙動が動画像によって可視化されてもよい。

【0078】

以上詳細に説明したように、情報処理システムのサーバは、端末から送信された要求信号に応答して、ペプチドの体内動態に関する予測情報を生成する。そして、サーバは、予測情報を端末へ送信する。これにより、ペプチドの体内動態を予測することができる。

30

【0079】

なお、本開示は、上述した実施形態に限定されるものではなく、この開示の要旨を逸脱しない範囲内で様々な変形や応用が可能である。

【0080】

例えば、上記実施形態のサーバ14は、ペプチドの体内動態の予測情報として、膜透過性及び体内持続性を予測する場合を例に説明したが、これに限定されるものではない。ペプチドの体内動態であれば、どのようなものであってもよい。例えば、ペプチドの体内動態として、血液脳関門(Blood-Brain Barrier)透過性又は体内環境におけるペプチドの溶解度等の物性パラメータを予測するようにしてもよい。

【0081】

また、上記実施形態のサーバ14は、ペプチドの予測情報のみを生成する場合を例に説明したが、これに限定されるものではない。例えば、サーバ14は、ペプチドの体内動態を改良する目的で、ペプチドを構成する要素のうちの変更箇所の候補を表す設計支援情報を生成する設計支援部を更に備えるようにしてもよい。例えば、ペプチドを構成する要素としてはペプチドの残基が挙げられる。この場合には、サーバ14は、例えば、変更箇所の候補Rを含む設計支援情報Sを生成する。そして、端末12の表示部には、図9に示されるような設計支援情報Sが表示される。これにより、ユーザのペプチドの設計が支援される。なお、サーバ14は、ペプチドの予測情報に代えてペプチドの設計支援情報のみを生成し、設計支援情報を端末12へ送信するようにしてもよい。

【0082】

40

50

また、上記実施形態では、特徴情報及び予測情報が共通データベース 142 に格納され、ペプチド情報、特徴情報、及び予測情報がユーザ用データベース 144 に格納される場合を例に説明したが、これらに関しては種々の変形例が想定される。

【0083】

例えば、上記実施形態では、ユーザが予測対象としてサーバ 14 へ送信したペプチド情報の特徴情報及び予測情報は、共通データベース 142 に全て格納される場合を例に説明したが、これに限定されるものではない。

【0084】

例えば、サーバ 14 は、ペプチド情報の特徴情報及び予測情報のうちユーザの承諾が得られた特徴情報及び予測情報のみを共通データベース 142 に格納するようにもよい。この場合には、サーバ 14 は、例えば、図 6 のステップ S120 において特徴情報及び予測情報を共通データベース 142 に格納する際に、端末 12 に対して特徴情報を共通データベース 142 に格納して良いか否かを表す確認信号を出力する。端末 12 は、サーバ 14 から送信された確認信号を受信すると、自らの表示部（図示省略）に、予測対象のペプチド情報の特徴情報を共通データベース 142 に格納しても良いか否かの確認を取る画面を表示させる。なお、この際には、例えば、ユーザが共通データベース 142 に特徴情報を格納することに同意した場合には、その特徴情報又は予測情報は他のユーザのペプチドの予測情報の生成の際にも用いられ、当該ユーザに対しては特徴情報又は予測情報の提供に対する報酬としてポイント等が付与される旨の表示が端末 12 の表示部（図示省略）に表示される。そして、例えば、ユーザが、特徴情報又は予測情報を共通データベース 142 に格納しても良い旨の操作を端末 12 に対して入力した場合には、端末 12 は、その旨の指示信号をサーバ 14 へ送信する。一方、ユーザが、特徴情報又は特徴情報を共通データベース 142 に格納することには同意しない旨の操作を端末 12 に対して入力した場合には、端末 12 は、その旨の指示信号をサーバ 14 へ送信する。

10

20

30

【0085】

サーバ 14 は、端末 12 から送信された指示信号に応じて、特徴情報又は予測情報を共通データベース 142 に格納するか否かを判定する。そして、サーバ 14 は、端末 12 から送信された指示信号が、共通データベース 142 への特徴情報又は予測情報の格納に同意する旨を表す場合には、当該特徴情報又は当該予測情報を共通データベース 142 に格納する。一方、サーバ 14 は、端末 12 から送信された指示信号が、共通データベース 142 への特徴情報又は予測情報の格納に同意しない旨を表す場合には、特徴情報又は予測情報を共通データベース 142 に格納せずに処理を終了する。このように、ユーザの意向に応じてデータを登録することにより、ユーザはサーバ 14 の利用がしやすくなる。なお、前述したように、ユーザから特徴情報又は予測情報が提供され、特徴情報又は予測情報が共通データベース 142 へ格納された場合には、ユーザに対してポイントが付与される。ユーザに対して付与されたポイントは、例えば、サーバ 14 を利用する際の利用料の割引等に用いられる。

【0086】

また、ユーザはペプチドの予測情報を得た後に、そのペプチドが体内において実際にどのような挙動を示すのかを検証するために実験を行う場合がある。ユーザは、その実験値を確認するとともに、サーバ 14 のユーザ用データベース 144 及び共通データベース 142 に格納することも可能である。この場合、ユーザが、上記の特徴情報及び予測情報の場合と同様に、実験値と実験手法の情報とを共通データベース 142 に格納することに同意し、ユーザによる実験によって得られた各種情報が共通データベース 142 に格納された場合には当該ユーザに対しては実験によって得られた各種情報の提供に対する報酬としてポイント等が付与される。これにより、ユーザの実験によって得られた実験値は、他のユーザ又はサーバ 14 の管理者による利用が可能な情報となり、例えば、ユーザ又は管理者が予測モデルを機械学習させる際の教師データとして活用が可能となる。例えば、図 2 に示されている、c2_A、d2_A、e2_A、及びf2_A は、ユーザから提供された実験情報の一例であり、このように実験情報が共通データベース 142 に格納された場合には、ユー

40

50

ザに対してポイントが付与される。

【 0 0 8 7 】

なお、サーバ14の各種の記憶部にデータを格納する際には、ユーザは、サーバ14には一切データを格納しない、ユーザ用データベース144にのみデータを格納する、又は共通データベース142にもデータを格納する、といった何れかの形態を選択可能である。

【 0 0 8 8 】

また、ユーザに対してポイントが付与される際には、サーバ14の管理者は、ユーザの行動履歴又は図2の注釈等に格納される信頼度等に基づき、ユーザに対してポイントを付与するようにしてもよい。なお、注釈等に格納される信頼度は、ユーザが提供したデータ自体の信頼度及びユーザに対する信頼度の何れであってもよい。この場合には、例えば、高品質なデータを提供してくれるユーザ又は信頼度が高いユーザに対しては多くのポイントが付与される。

10

【 0 0 8 9 】

また、上記実施形態では、複数のユーザ用予測モデル記憶部146A, 146B, 146Cの各々に格納されている学習済みモデルは、ユーザ用データベース144に格納されているデータのみを学習用データとして生成された学習済みモデルである場合を例に説明したが、これに限定されるものではない。例えば、複数のユーザ用予測モデル記憶部146A, 146B, 146Cの各々に格納されている学習済みモデルは、ユーザ用データベース144に格納されているデータに加えて、共通データベース142に格納されているデータを学習用データとして生成されていてもよい。具体的には、ユーザ用データベース144に格納されているデータに加えて、共通データベース142に格納されているデータの一部又は全部を学習用データとして学習済みモデルを生成するようにしてもよい。これにより、予測精度の高い学習済みモデルが生成される。なお、この場合には、ユーザ用データベース144又は共通データベース142には、学習用データとしての各種データが予め登録される。

20

【 0 0 9 0 】

また、ユーザは上記のような各種データをサーバ14の共通データベース142に格納することにより自らのデータを提供するのみではなく、自らが作成した予測モデルをサーバ14に提供するようにしてもよい。この場合には、上記と同様に、予測モデルの提供に際しては当該ユーザに対してポイントが付与される。

30

【 0 0 9 1 】

また、共通データベース142又はユーザ用データベース144に格納されているデータは、端末12へダウンロード可能となっていても良い。また、共通データベース142又はユーザ用データベース144に格納されているデータは、端末12により編集可能な状態となっていてもよい。なお、この場合には、端末12から編集可能なデータは制限がかけられていてもよい。例えば、共通データベース142に格納されているデータに関しては、端末12からは編集ができないように構成されていてもよい。また、端末12又はサーバ14は、共通データベース142又はユーザ用データベース144に格納されているデータを用いて、ペプチドの体内動態に関する各種の情報を生成するようにしてもよい。例えば、端末12又はサーバ14は、ペプチドの特徴量（例えば、脂溶性）を横軸にとり、膜透過性を縦軸にとったグラフを生成するなどして、ペプチドの体内動態に関する各種の情報を生成するようにしてもよい。

40

【 0 0 9 2 】

また、上記実施形態では、サーバ14の予測部148が、端末12から送信された要求信号に含まれるペプチド情報を特徴情報へ変換する場合を例に説明したが、これに限定されるものではない。例えば、端末12から送信される要求信号には特徴情報が含まれていてもよい。

【 0 0 9 3 】

サーバ14により提供されるサービスを利用するユーザの中には、サーバ14に対してペプチドの構造式等のペプチド情報を送信することに躊躇するユーザも存在することが想

50

定される。このため、例えば、ユーザが操作する端末 1 2 又はサーバ 1 4 とは異なるコンピュータによって、ペプチド情報から特徴情報への変換を予め実行し、要求信号にその特徴情報を含ませるようにしてもよい。この場合には、端末 1 2 は、ペプチド情報から変換された特徴情報を含む要求信号をサーバ 1 4 へ送信する。サーバ 1 4 の受信部 1 4 0 は、端末 1 2 から送信された要求信号を受信する。サーバ 1 4 の予測部 1 4 8 は、要求信号に含まれる特徴情報を予測モデルに対して入力することにより、ペプチドの予測情報を生成する。そして、サーバ 1 4 の送信部 1 5 2 は、得られた予測情報を端末 1 2 へ送信する。これにより、ユーザは、ペプチドの構造式等であるペプチド情報をサーバ 1 4 へ送信することなく、ペプチドの予測情報を得ることができる。ただし、この場合には、予測モデルをシミュレーションモデルとする選択は不可となる。シミュレーションモデルによる予測には、ペプチドの構造式等のペプチド情報が必要となるためである。なお、この場合には、例えば、サーバ 1 4 又はサーバ 1 4 とは異なるコンピュータによって利用される、ペプチド情報から特徴情報への変換プログラム等を、端末 1 2 又はサーバ 1 4 とは異なるコンピュータに予め提供しておく等の対応が想定される。

【 0 0 9 4 】

[第 2 実施形態]

【 0 0 9 5 】

次に、第 2 実施形態について説明する。第 2 実施形態の情報処理システムは、ユーザによるペプチドの設計を支援する。なお、第 2 実施形態では、第 1 実施形態をより具体化させた内容について説明する。なお、第 2 実施形態に係る情報処理システムの構成のうちの第 1 実施形態の情報処理システムと同様の部分は、同一符号を付して説明を省略する。

【 0 0 9 6 】

ユーザがペプチドを設計する場合、所望の膜透過性又は体内持続性を持つペプチドの構成を知りたい場合がある。この点、ペプチドの構成のうちどの部分を変更すれば、より好ましい膜透過性又は体内持続性を持つペプチドとなるのかといった情報は、ペプチドを設計するユーザにとっては有用な情報である。

【 0 0 9 7 】

そこで、第 2 実施形態の情報処理システムは、ペプチドを構成する複数の残基のうち予測情報に特に影響を与えていたる残基を特定し、その残基を交換候補の残基としてユーザに対して提示する。これにより、より好ましい膜透過性又は体内持続性を有するペプチドの探索を支援することができる。また、ユーザに対してペプチドの設計に関するナビゲーションサービスを提供することができる。

【 0 0 9 8 】

図 1 0 は、第 2 実施形態に係る情報処理システム 2 1 0 の構成の一例を示すブロック図である。第 2 実施形態の情報処理システム 2 1 0 のサーバ 2 1 4 は、図 1 0 に示されるように、設計支援部 2 5 4 を更に備えている。設計支援部 2 5 4 は、ユーザによるペプチドの設計を支援する情報を端末 1 2 へ送信する。以下、具体的に説明する。

【 0 0 9 9 】

なお、第 2 実施形態の情報処理システム 2 1 0 のサーバ 2 1 4 の予測部 1 4 8 は、第 1 実施形態と同様に、ニューラルネットワーク等の学習済みモデルによってペプチドの予測情報を生成する。また、情報処理装置 2 1 0 の予測部 1 4 8 は、第 1 実施形態と同様に、予測対象のペプチドのペプチド情報から特徴情報として特徴ベクトル x を抽出し、その特徴ベクトル x を学習済みモデルへ入力することにより予測情報を生成する。

【 0 1 0 0 】

なお、特徴ベクトル x は、ペプチドが有する複数の残基の各々から得られる。例えば、特徴ベクトル $x = [x_{11}, x_{12}, \dots, x_{21}, x_{22}, \dots, x_{N1}, x_{N2}, \dots]$ のうちの、 $[x_{11}, x_{12}, \dots]$ はペプチドに含まれる 1 番目の残基から得られるベクトルであり、 $[x_{21}, x_{22}, \dots]$ はペプチドに含まれる 2 番目の残基から得られるベクトルであり、 $[x_{N1}, x_{N2}, \dots]$ はペプチドに含まれる N 番目の残基から得られるベクトルである。

【0101】

まず、設計支援部254は、ペプチドを構成する複数の残基の各々について、ペプチドの予測情報を生成するための学習済みモデルのパラメータを、例えば既知の技術を用いて解析することにより、当該残基が予測情報に影響を与えている度合いを表すスコア（以下、単に「残基影響度スコア」とも称する。）を計算する。

【0102】

具体的には、まず、設計支援部254は、学習済みモデルに入力される特徴ベクトル x のうちの各要素 x_{ij} について、学習済みモデルから出力される予測情報が表す値 y を要素 x_{ij} によって偏微分することにより得られる微分値を計算する。なお、予測情報が表す値 y の要素 x_{ij} による偏微分は次式によって表される。

10

【0103】**【数1】**

$$\frac{\partial y}{\partial x_{ij}}$$

【0104】

20

この微分値は、学習済みモデルのパラメータを解析することにより得られる。また、以下の微分値

【0105】**【数2】**

$$\frac{\partial y}{\partial x_{i1}}, \frac{\partial y}{\partial x_{i2}}, \dots,$$

30

【0106】

の絶対値を、 i 番目の残基の各特徴量が予測情報に影響を与えている度合いを表すスコア（以下、単に「特徴量影響度スコア」とも称する。）とする。設計支援部254は、この特徴量影響度スコアを特徴ベクトル x のうちの各要素 x_{ij} について計算する。

【0107】

図11に、本実施形態の特徴量影響度スコアを説明するための図を示す。図11は、他分野においてはサリエンシーマップとも称される。

【0108】

図11に示されるマップでは、横軸の「1, 2, ...」が残基のIDを表し、縦軸の「 x_{ij} 」が特徴ベクトル x に含まれる特徴量の種類を表す。また、マップの濃淡は、特徴ベクトル x の各要素 x_{ij} が予測に影響を与えている度合いを表す特徴量影響度スコアに相当し、マップの濃淡が濃いほど特徴量影響度スコアが高いことが表されている。

40

【0109】

図11の例では、例えば、ID = 7である残基の特徴ベクトルの要素 x_{71} の特徴量影響度スコアが高い。このため、ID = 7である残基を別の残基等に交換すれば、予測情報が大きく変化するものと予想される。なお、このとき、特徴量影響度スコアに対応する微分値の正負を確認することにより、この残基が存在することによる影響の方向を示すことも可能である。

【0110】

50

次に、例えば、設計支援部 254 は、複数の残基の各々について、特徴ベクトル \times の各要素に対して計算された特徴量影響度スコアの総和を計算するなどして、当該残基の残基影響度スコアを計算する。そして、設計支援部 254 は、複数の残基の各々に対して計算された当該残基の残基影響度スコアのうち所定閾値以上の残基影響度スコアの残基を特定する。

【0111】

なお、特徴量影響度スコアから残基影響度スコアを計算する方法は、上記の手法に限定されるものではなく、例えば、特徴ベクトル \times の各要素に対して計算された特徴量影響度スコアの重み付き平均、最大値、又は最小値等を残基影響度スコアとしてもよい。

【0112】

次に、設計支援部 254 は、ペプチドの構造のうち、特定された残基を交換候補の残基として設定し、交換候補の残基を変更箇所の候補とする。

【0113】

例えば、設計支援部 254 は、残基影響度スコアが所定閾値以上である残基を変更箇所の候補として設定する。そして、設計支援部 254 は、交換候補の残基を提案する設計支援情報を生成する。

【0114】

第2実施形態のサーバ214の送信部152は、設計支援部254により生成された設計支援情報を端末12へ送信する。なお、送信部152は、上記図11に示されるようなマップを設計支援情報として端末12へ送信するようにしてもよい。

10

【0115】

なお、第2実施形態の情報処理システム210の他の構成及び作用については、第1実施形態と同様であるため、説明を省略する。

【0116】

以上説明したように、第2実施形態の情報処理システムのサーバは、ペプチドを構成する複数の残基の各々について、予測情報を生成するための学習済みモデルのパラメータを解析することにより、当該残基が予測情報に影響を与えている度合いを表す残基影響度スコアを計算する。サーバは、複数の残基の各々に対して計算された残基影響度スコアのうち所定閾値以上の残基影響度スコアの残基を特定し、ペプチドの構造のうち、特定された残基を交換候補の残基として設定し、交換候補の残基を変更箇所の候補とする。また、サーバは、変更箇所の候補を提案する設計支援情報を端末へ送信する。端末は、サーバから送信された設計支援情報を表示部（図示省略）に表示させる。これにより、ユーザは、ペプチドを構成する残基のうち何れの残基をどのように変更すれば良いのかに関する手掛けかりを得ることができる。また、ユーザによるペプチドの設計を支援することができる。

20

【0117】

なお、第2実施形態のサーバは、複数の残基の各々に対して計算された残基影響度スコアのうち所定閾値未満の残基影響度スコアの残基を特定し、その情報をユーザに対して提示するようにしてもよい。この場合には、変更しても膜透過性又は体内持続性にあまり影響しない残基が特定されうるため、ユーザは、ペプチドを構成する残基のうち何れの残基を変更すれば良いのかに関する手掛けかりを得ることができる。

30

【0118】

[第3実施形態]

【0119】

次に、第3実施形態について説明する。第3実施形態の情報処理システムは、ペプチドを構成する複数の残基のうちの少なくとも1つの残基を予め定めた別の残基に置換した候補ペプチドを複数生成し、複数の候補ペプチドの各々の予測情報を生成することにより、予測情報に影響を与えている残基を特定し、具体的に残基交換を提案する点が、第1及び第2実施形態と異なる。なお、第3実施形態に係る情報処理システムの構成は、第2実施形態の情報処理システムと同様の構成であるため、同一符号を付して説明を省略する。

【0120】

40

50

第3実施形態のサーバ214の設計支援部254は、ペプチドを構成する複数の残基のうちの少なくとも1つの残基を、アラニンなどの予め定めた別の残基に置換した候補ペプチドを生成する。

【0121】

図12に、候補ペプチドを説明するための図を示す。図12は、あるペプチドS1のうちの残基Re1を別の残基へ置換する場合の例が示されている。例えば、設計支援部254は、図12のペプチドS1のうちの残基Re1を別の残基Re2（例えば、グリシン）へ置換することにより、候補ペプチドS2を生成する。また、設計支援部254は、図12のペプチドS1のうちの残基Re1を別の残基Re3（例えば、アラニン）へ置換することにより、候補ペプチドS3を生成する。なお、別の残基は、例えば、バリン、ロイシン、アルギニン、若しくはアスパラギン酸等の天然アミノ酸、又は任意の人工アミノ酸等であってもよい。

10

【0122】

そして、第3実施形態の予測部148は、第1実施形態と同様の手法を用いて、残基を置換する前のペプチドの予測情報と、複数の候補ペプチドの各々の予測情報を生成する。

【0123】

次に、設計支援部254は、複数の候補ペプチドの各々について、候補ペプチドの予測情報と、残基を置換する前のペプチドの予測情報との差分を計算する。そして、設計支援部254は、差分が所定閾値以上である候補ペプチドを特定し、特定した候補ペプチドのうちの置換後の残基の箇所を特定する。

20

【0124】

残基を置換する前のペプチドの予測情報と、候補ペプチドの予測情報との間の差分が大きい場合、置換前の残基及び置換後の残基の少なくとも一方は予測情報に対して大きな影響を与える残基であることになる。このため、設計支援部254は、差分が大きい候補ペプチドのうちの置換後の残基の箇所を特定し、残基を置換する前のペプチドにおいて当該箇所に存在していた残基を特定する。

【0125】

そして、設計支援部254は、ペプチドの構造を表すペプチド構造情報のうち、特定された箇所の残基を交換候補の残基として設定し、交換候補の残基を変更箇所の候補とする。そして、設計支援部254は、交換候補の残基を別の残基へ交換することを提案する設計支援情報を生成する。

30

【0126】

図13に、設計支援情報の一例を示す。図13に示されるように、例えば、設計支援情報には、残基を変更する前のペプチド情報（例えば、ペプチドの構造式）、交換候補の残基を表す情報（図中では点線部分が交換候補の残基を表す）、及びペプチドの予測情報が含まれる。また、例えば、設計支援情報には、残基を交換した後の候補ペプチドのペプチド情報（例えば、ペプチドの構造式）、候補ペプチドの予測情報、及び候補ペプチドに新たに組み込まれた残基の情報（例えば、残基の構造式）が含まれる。なお、設計支援情報には、交換候補の残基を複数の異なる別の残基へ置換し、交換候補の残基が置換された際の予測情報の変化（または、変化の重み付き平均）を含んでいてもよい。

40

【0127】

第3実施形態のサーバ214の送信部152は、設計支援部254により生成された設計支援情報を端末12へ送信する。

【0128】

なお、第3実施形態の情報処理システム210の他の構成及び作用については、第1又は第2実施形態と同様であるため、説明を省略する。

【0129】

以上説明したように、第3実施形態の情報処理システムのサーバは、ペプチドを構成する複数の残基のうちの少なくとも1つの残基を予め定めた別の残基に置換した候補ペプチドを複数生成し、残基を置換する前のペプチドの予測情報と、複数の候補ペプチドの各々

50

の予測情報を生成する。次に、サーバは、複数の候補ペプチドの各々について、残基を置換する前のペプチドの予測情報と、候補ペプチドの予測情報との間の差分を計算し、差分が所定閾値以上である候補ペプチドを特定し、特定した候補ペプチドのうちの置換後の残基の箇所を特定する。そして、サーバは、ペプチドの構造のうち、特定された箇所の残基を交換候補の残基として設定し、交換候補の残基を変更箇所の候補とする。これにより、ユーザは、ペプチドを構成する残基のうち何れの残基を変更すれば良いのかに関する手掛かりを得ることができる。また、ユーザによるペプチドの設計を支援することができる。

【0130】

[第4実施形態]

【0131】

10

次に、第4実施形態について説明する。第4実施形態の情報処理システムは、ユーザから所定のデータが提供された場合に、当該ユーザに対して報酬を表すポイントを付与する点が、第1～第3実施形態と異なる。なお、第4実施形態に係る情報処理システムの構成は、第1実施形態の情報処理システムと同様の構成であるため、同一符号を付して説明を省略する。

【0132】

ペプチドの体内動態に関する予測情報を精度良く生成するためには、多くのデータが必要となる。例えば、あるペプチドの体内動態に関する所定の実験により得られた膜透過性又は体内持続性を表す実験データは、他のペプチドの体内動態を予測するための有用なデータとなり得る。

20

【0133】

そこで、第4実施形態の情報処理システムは、ユーザから所定のデータが提供された場合に、当該ユーザに対して報酬を表すポイントを付与する。このポイントは、例えば、本情報処理システムを用いて提供されるサービスの利用料の割引等に利用される。

【0134】

図14は、第4実施形態に係る情報処理システム410の構成の一例を示すブロック図である。第4実施形態の情報処理システム410のサーバ414は、図14に示されるように、報酬付与部456と、報酬記憶部458とを更に備えている。

【0135】

報酬付与部456は、ユーザから提供されたデータを、共通データベース142又は共通予測モデル記憶部145へ格納する。例えば、報酬付与部456は、端末12から送信された、ペプチドの予測情報の教師データである、ペプチドの体内動態に関する実験データ、実験データを得た際の実験手法を表す情報、ユーザ用予測モデル記憶部146に保存された予測モデル、ユーザが自らの端末で利用している予測モデル（例えば、ユーザが自らの端末で独自に生成した予測モデル又はユーザが外部から持ち込んできた予測モデル等）、及びユーザにより生成されたペプチドの体内動態に関する予測情報の少なくとも1つのデータを受け付ける。そして、報酬付与部456は、受け付けたデータを、情報処理装置414が管理する記憶部の一例である共通データベース142又は共通予測モデル記憶部145に格納する。

30

【0136】

なお、実験データは、既知の実験手法により得られる予測対象のペプチドの体内動態に関するデータである。この実験データは、例えば、学習済みモデルの生成の際に教師データとして用いられる。また、実験データを得た際の実験手法を表す情報に関しては、どの実験手法を用いたかによって実験データの信頼性は異なるものとなるため、有用な情報である。また、ユーザが独自に生成した学習済みモデル等の予測モデルが提供されれば、その予測モデルを基に新たな予測モデルを生成することも可能であるため、同様に有用である。また、ユーザが独自に生成したペプチドの予測情報も同様に有用である。

40

【0137】

そのため、報酬付与部456は、データが送信された端末12に付与されているユーザIDに対して、当該データの提供に対する報酬を表すポイントを付与する。

50

【 0 1 3 8 】

例えば、報酬付与部 456 は、報酬記憶部 458 に格納されている、ユーザ ID と当該ユーザ ID に付与されているポイントとが対応付けられているテーブルを更新することにより、ユーザ ID に対してポイントを付与する。

【 0 1 3 9 】

図 15 は、報酬記憶部 458 に格納されているテーブルの一例である。図 15 に示される例では、ユーザ ID 「USER_01」と当該ユーザ ID に付与されているポイント「P_USER_01」が示されている。報酬付与部 456 は、ユーザ ID 「USER_01」に対応する端末 12 からデータの提供を受け付けた場合には、ユーザ ID 「USER_01」に付与されているポイント「P_USER_01」に所定のポイントを加算することにより、報酬記憶部 458 に格納されているテーブルを更新する。なお、データ提供に対してどの程度のポイントが付与されるのかに関してのルールは予め設定されている。例えば、報酬付与部 456 は、提供されるデータの種類（例えば、提供されるデータは、実験値を表すデータ、予測モデルを表すデータ、及び予測情報を表すデータの何れであるか）、データの量、データの履歴情報（例えば、提供されるデータが膜透過性の実験値を表すデータである場合、そのデータは PAMPA 試験、Caco2 細胞、MDCK 細胞、及び LLC-PK1 細胞等のうちどの手段で得られたものであるかといった履歴情報）、データのカテゴリ（例えば、提供されるデータは、環状ペプチドに関するデータ及び低分子に関するデータの何れであるかといった属性）、及びデータが提供されたユーザの信頼度等に応じて、ユーザに対してポイントが付与される。なお、ユーザの信頼度は、人手によって設定することも可能であるし、サーバ 414 のようなコンピュータによって設定することも可能である。例えば、サーバ 414 の報酬付与部 456 は、情報処理システム 410 が提供しているサービスの利用頻度及び利用履歴等に応じて、ユーザの信頼度を決定する。例えば、報酬付与部 456 は、利用頻度が高いユーザほど信頼度が高くなるように設定する。または、報酬付与部 456 は、利用履歴に応じて利用年数が長いユーザほど信頼度が高くなるように設定する。また、ユーザからデータの提供を受けた際のポイントについては、例えば、報酬付与部 456 は、提供されたデータを予測モデルへ反映させた場合の性能向上への貢献度合いの推定値を計算し、当該推定値に基づいて、ユーザに付与するポイントを変化させててもよい。例えば、報酬付与部 456 は、ユーザから提供されたデータとサーバ 414 が既に保持しているデータとの間の類似度を計算し、類似度が低いデータを提供したユーザに対しては高いポイントを付与するようにしてもよい。これにより、サーバ 414 が保持していないデータの提供が促される。また、例えば、報酬付与部 456 は、ペプチドの予測に悪影響があると思われるデータの提供がされた場合には、ポイントを付与しないようにしてもよい。

【 0 1 4 0 】

このように、データの提供に対するインセンティブをユーザに対して与えることにより、サーバ 414 の共通データベース 142 又は共通予測モデル記憶部 145 には、より多くのデータが格納され、それらのデータをペプチドの体内動態の予測に役立てることができる。予測モデルの提供に対するポイントの付与についても、上記とほぼ同様である。例えば、予め準備しておいた正解の判っている試験用データを用いて、提供された予測モデルの性能を推定し、優れた性能を発揮すると期待できる予測モデルには高いポイントを付与するようにしてもよい。また、例えば、従来の予測モデルでは成績が悪かったペプチドに対して、予測が改善されているような予測モデルには高いポイントを付与するようにしてもよい。

【 0 1 4 1 】

以上説明したように、第 4 実施形態の情報処理システムのサーバは、端末から送信された、ペプチドの予測情報の教師データである実験データ、実験データを得た際の実験手法を表す情報、ユーザが利用する予測モデル、及びユーザにより生成されたペプチドの体内動態に関する予測情報の少なくとも 1 つのデータが、サーバが管理する記憶部に格納された場合、端末に付与されているユーザ ID に対して、データの提供に対する報酬を表すポイントを付与する。これにより、ペプチドの体内動態の予測に有用なデータをより多く収

10

20

30

40

50

集することができる。

【 0 1 4 2 】

なお、本開示は、上述した実施形態に限定されるものではなく、この開示の要旨を逸脱しない範囲内で様々な変形や応用が可能である。

【 0 1 4 3 】

また、上記では本発明に係るプログラムが記憶部（図示省略）に予め記憶（インストール）されている様子を説明したが、本発明に係るプログラムは、CD-ROM、DVD-ROM及びマイクロSDカード等の記録媒体に記録されている形態で提供することも可能である。

【 0 1 4 4 】

なお、上記実施形態でCPUがソフトウェア（プログラム）を読み込んで実行した処理を、CPU以外の各種のプロセッサが実行してもよい。この場合のプロセッサとしては、FPGA（Field-Programmable Gate Array）等の製造後に回路構成を変更可能なPLD（Programmable Logic Device）、及びASIC（Application Specific Integrated Circuit）等の特定の処理を実行させるために専用に設計された回路構成を有するプロセッサである専用電気回路等が例示される。また、各処理を、これらの各種のプロセッサのうちの1つで実行してもよいし、同種又は異種の2つ以上のプロセッサの組み合わせ（例えば、複数のFPGA、及びCPUとFPGAとの組み合わせ等）で実行してもよい。また、これらの各種のプロセッサのハードウェア的な構造は、より具体的には、半導体素子等の回路素子を組み合わせた電気回路である。

10

【 0 1 4 5 】

また、本実施形態の各処理を、汎用演算処理装置及び記憶装置等を備えたコンピュータ又はサーバ等により構成して、各処理がプログラムによって実行されるものとしてもよい。このプログラムは記憶装置に記憶されており、磁気ディスク、光ディスク、半導体メモリ等の記録媒体に記録することも、ネットワークを通して提供することも可能である。もちろん、その他いかなる構成要素についても、単一のコンピュータやサーバによって実現しなければならないものではなく、ネットワークによって接続された複数のコンピュータに分散して実現してもよい。

20

【 符号の説明 】

【 0 1 4 6 】

30

1 0 情報処理システム

1 2 A , 1 2 B , 1 2 C 端末

1 4 サーバ

1 6 ネットワーク

1 4 0 受信部

1 4 2 共通データベース

1 4 4 A , 1 4 4 B , 1 4 4 C ユーザ用データベース

1 4 5 共通予測モデル記憶部

1 4 6 A , 1 4 6 B , 1 4 6 C ユーザ用予測モデル記憶部

1 4 8 予測部

40

1 5 0 処理部

1 5 2 送信部

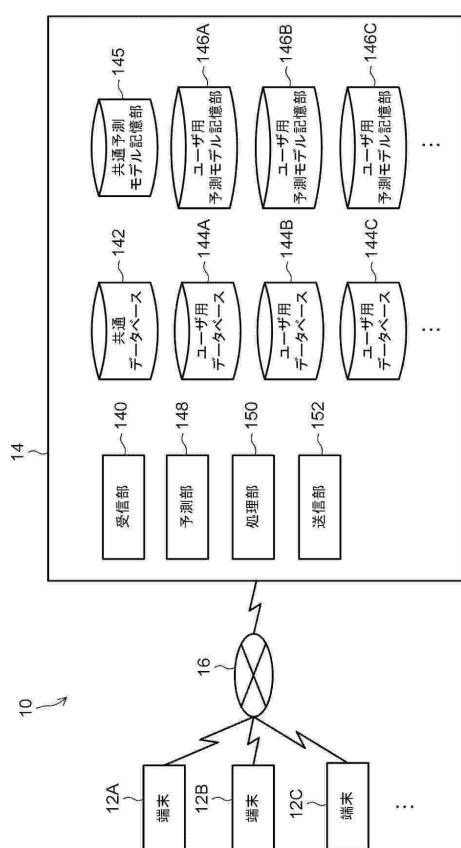
2 5 4 設計支援部

4 5 6 報酬付与部

4 5 8 報酬記憶部

50

【図面】
【図 1】



【図 3】

データID	ペプチド情報	特徴情報	膜透過性実験値	膜透過性実験値	体内持続性実験手法	体内持続性実験手法	注釈	予測情報	...
A_00001	a1_A	b1_A	c1_A	d1_A	e1_A	f1_A	g1_A	h1_A	...
...
...

【図 2】

データID	ペプチド情報	特徴情報	膜透過性実験値	膜透過性実験手法	体内持続性実験手法	注釈	予測情報	...	
00001	a1	b1	c1	d1	e1	f1	g1	h1	...
00002	a2	b2	c2	d2	e2	f2	g2	h2	...
00003	a3	b3	—	—	e3	f3	g3	h3	...
00004	a4	b4	c4	d4	—	—	g4	h4	...
...
A_00001	—	b1_A	—	—	e1_A	f1_A	g1_A	h1_A	...
A_00002	—	b2_A	c2_A	d2_A	e2_A	f2_A	g2_A	h2_A	...
...
B_00001	—	b1_B	c1_B	d1_B	e1_B	f1_B	g1_B	h1_B	...
...

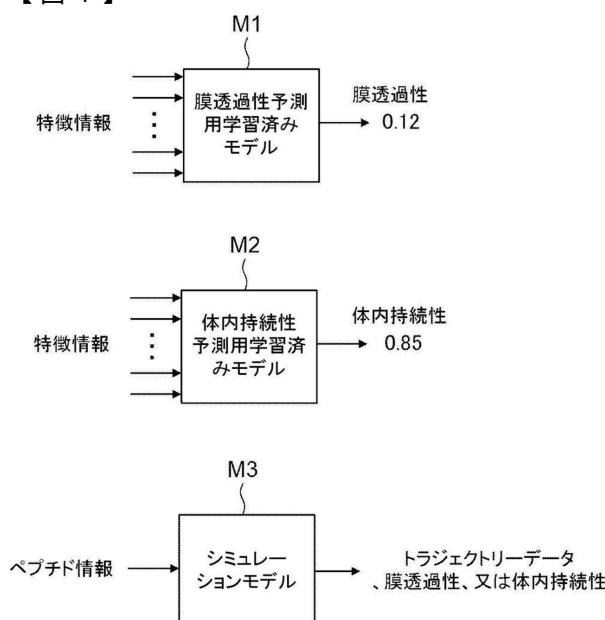
10

20

30

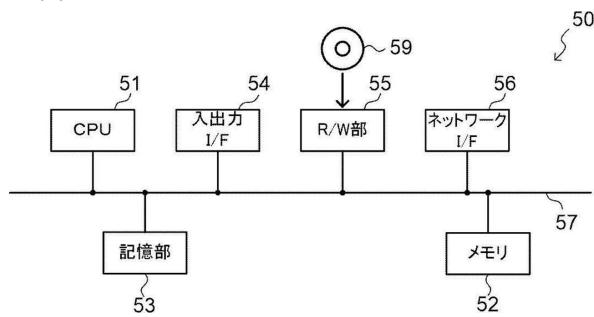
40

【図 4】

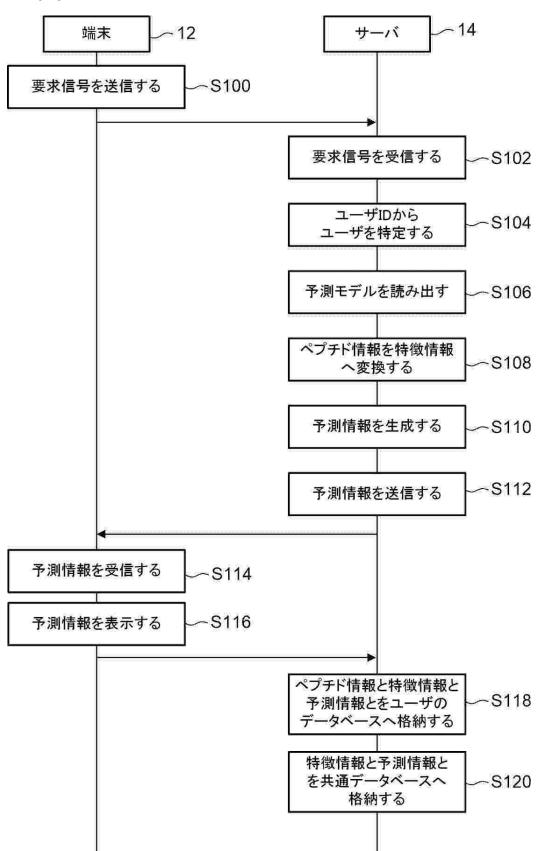


50

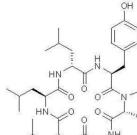
【 図 5 】



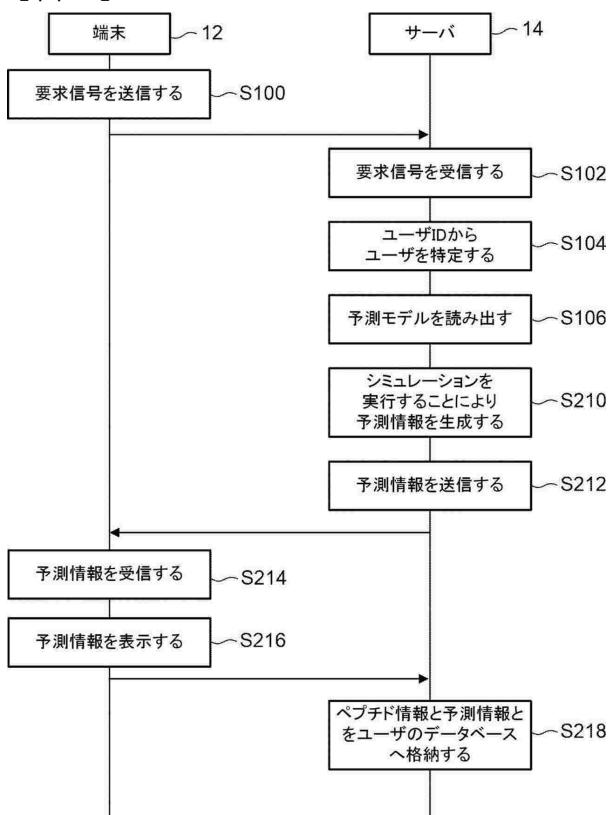
【図6】



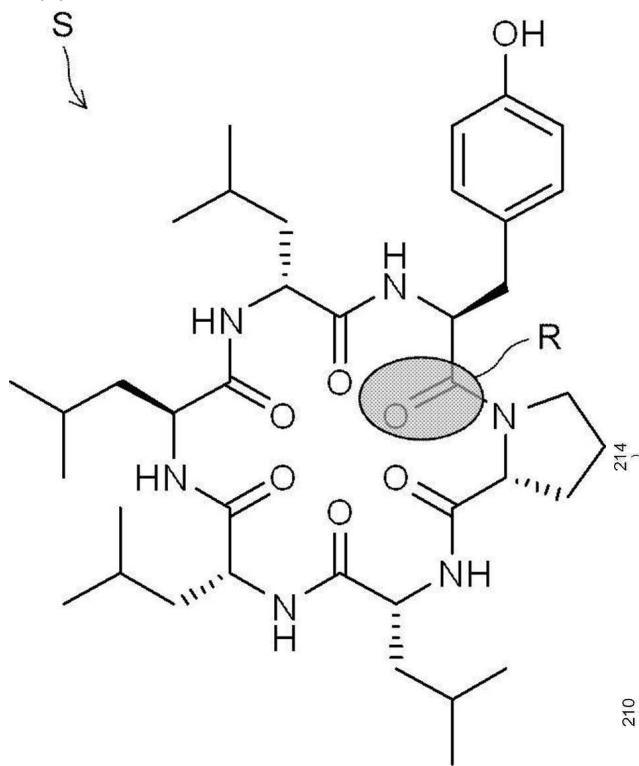
【 四 7 】

ペプチドID	ペプチドの構造式	SMILES表記	予測情報
00001		O=C([C@H](N1)CC2=CC=C(C(C=C2)N3CCC[C@H]3)C(N([C@H](CC(C)C)C(N([C@H](CC(C)C)C(N([C@H](CC(C)C)C(N([C@H](CC(C)C)C1=O)=O)=O)=O)=O)	...
00002
...

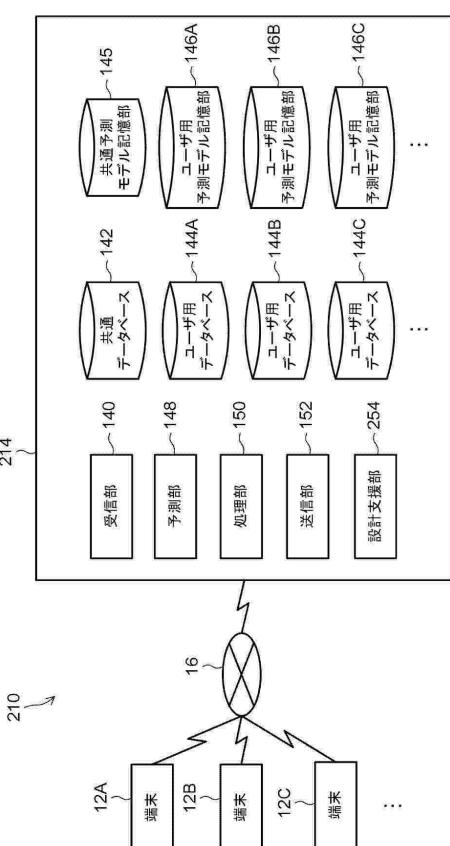
【 図 8 】



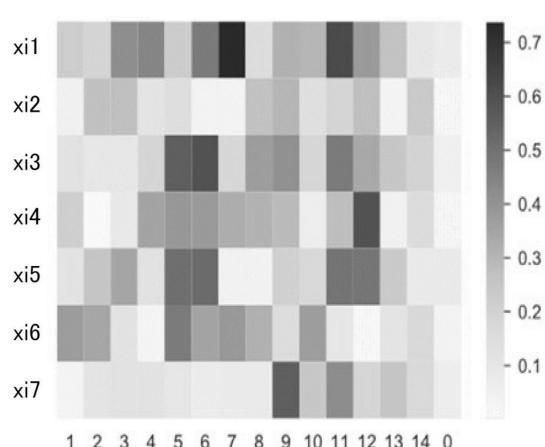
【図 9】



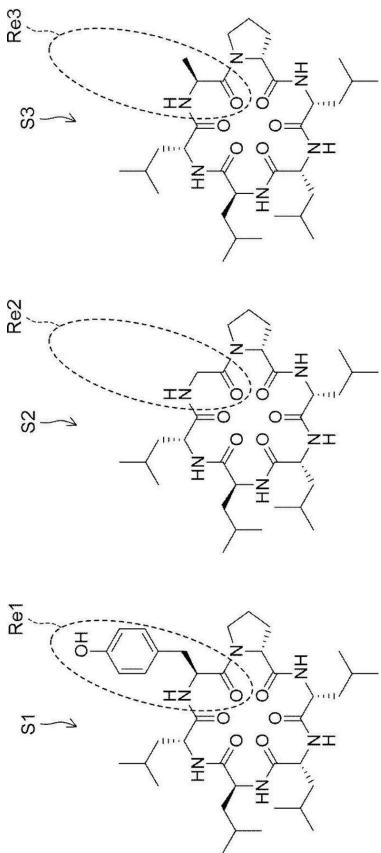
【図 10】



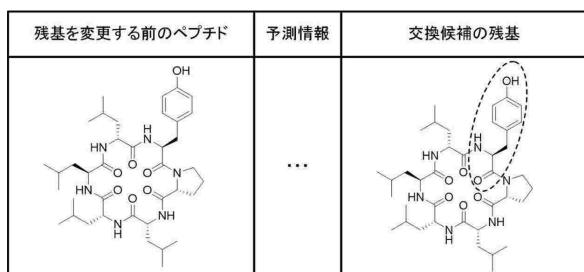
【図 11】



【図 12】

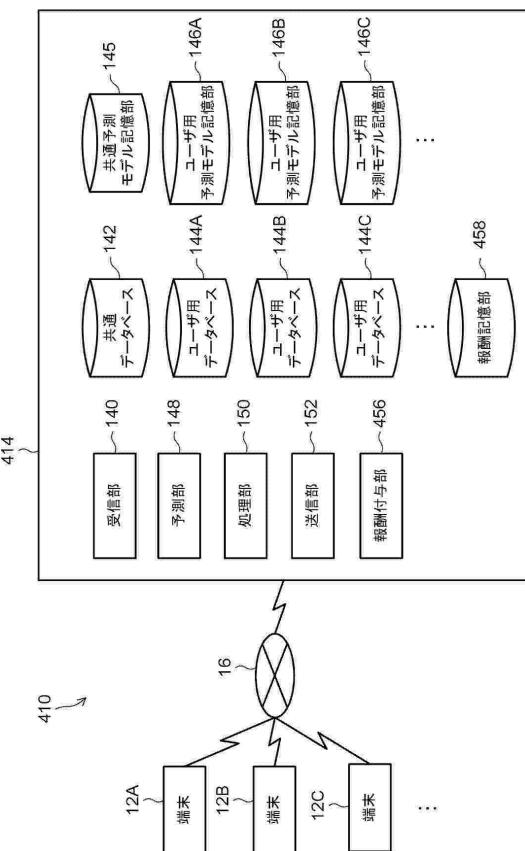


【図 1 3】



候補ID	残基を変更した後の候補ペプチド	予測情報	新たに組み込まれた残基の情報
0001	
...

【図 1 4】



【図 1 5】

ユーザID	ポイント	...
USER_01	P_USER_01	...
...
...

10

20

30

40

50

フロントページの続き

国立大学法人東京工業大学内

(72)発明者 吉川 寧

東京都目黒区大岡山2丁目12番1号 国立大学法人東京工業大学内

審査官 塩田 徳彦

(56)参考文献 特開2019-028879 (JP, A)

米国特許出願公開第2015/0300968 (US, A1)

特開2019-179356 (JP, A)

米国特許出願公開第2010/0121791 (US, A1)

(58)調査した分野 (Int.Cl., DB名)

G 16 B 5 / 00 - 99 / 00

G 06 N 20 / 00