



(12) 发明专利申请

(10) 申请公布号 CN 114496093 A

(43) 申请公布日 2022. 05. 13

(21) 申请号 202111341383.5

(22) 申请日 2021.11.12

(30) 优先权数据

2020-189856 2020.11.13 JP

2021-023750 2021.02.17 JP

(71) 申请人 国立大学法人东京工业大学

地址 日本东京都

(72) 发明人 秋山泰 大上雅史 柳泽溪甫

吉川宁

(74) 专利代理机构 北京中博世达专利商标代理

有限公司 11274

专利代理师 侯志源

(51) Int.Cl.

G16B 40/00 (2019.01)

G16B 50/00 (2019.01)

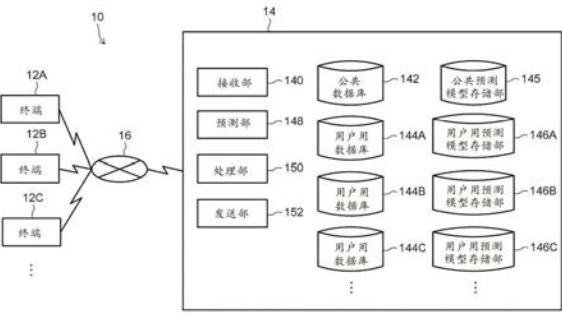
权利要求书2页 说明书15页 附图15页

(54) 发明名称

信息处理装置、方法、系统及记录信息处理程序的介质

(57) 摘要

本发明涉及一种信息处理装置、信息处理方法、记录信息处理程序的记录介质以及信息处理系统。服务器(14)响应于从终端(12)发送的请求信号,生成与肽的药代动力学相关的预测信息。然后,服务器(14)向终端(12)发送预测信息。



1. 一种信息处理装置,所述信息处理装置包括:  
接收部,所述接收部接收从终端发送的请求信号;  
预测部,所述预测部响应于所述请求信号,生成与肽的药代动力学相关的预测信息;以及  
发送部,所述发送部向所述终端发送由所述预测部生成的所述预测信息。
2. 根据权利要求1所述的信息处理装置,其中,  
在所述请求信号中包括用于生成预测信息的预测模型的选择信息,  
所述预测部通过从多个所述预测模型中选择与所述选择信息相应的预测模型,并对所选择的所述预测模型输入表示所述肽的肽信息,从而生成所述肽的所述预测信息。
3. 根据权利要求2所述的信息处理装置,其中,  
在所述请求信号中包括用户ID,  
多个预测模型中的每一个是为每个用户预先准备的预测模型,  
所述预测部通过将表示所述肽的肽信息输入到所述多个预测模型中的与所述用户ID对应的预测模型中,从而生成所述肽的所述预测信息。
4. 根据权利要求3所述的信息处理装置,其中,所述多个预测模型中的每一个是通过为每个用户预先准备的学习用数据来预先学习的学习完成模型。
5. 根据权利要求1至4中任一项所述的信息处理装置,其中,  
在所述请求信号中包括用户ID,  
所述信息处理装置还包括处理部,所述处理部根据所述用户ID,将所述预测信息存储在为每个用户预先准备的多个数据库中的、与所述用户ID对应的数据库中。
6. 根据权利要求1至5中任一项所述的信息处理装置,其中,  
在所述请求信号中包括用户ID,  
所述信息处理装置还包括处理部,所述处理部根据所述用户ID,将从所述终端发送的数据存储在为每个用户预先准备的多个数据库中的、与所述用户ID对应的数据库中。
7. 根据权利要求6所述的信息处理装置,其中,所述处理部将从所述终端发送的与所述肽的药代动力学相关的数据存储在所述数据库中。
8. 根据权利要求1至7中任一项所述的信息处理装置,其中,  
所述发送部将由所述预测部预测的所述预测信息与作为预测对象的所述肽相关联,并发送至所述终端。
9. 根据权利要求1至8中任一项所述的信息处理装置,所述信息处理装置还包括设计支援部,所述设计支援部生成包含构成所述肽的要素的变更位置的候选的设计支援信息,  
所述发送部将所述设计支援信息向所述终端发送。
10. 根据权利要求9所述的信息处理装置,其中,  
所述设计支援部对构成所述肽的多个残基中的每一个,通过解析用于生成所述预测信息的学习完成模型的参数,来计算表示所述残基对所述预测信息产生影响的程度的残基影响度分数;  
在针对所述多个残基中的每一个计算出的所述残基影响度分数中,确定规定阈值以上的残基影响度分数的所述残基;  
将在所述肽的结构中的所述确定的所述残基设定为交换候选的残基,并将所述交换候

选的残基作为变更位置的候选。

11. 根据权利要求9所述的信息处理装置, 其中,

所述设计支援部生成将构成所述肽的多个残基中的至少一个残基取代为预定的其他残基而得到的多个候选肽;

所述预测部生成取代残基之前的所述肽的所述预测信息、以及多个所述候选肽中的每一个的所述预测信息;

所述设计支援部对多个所述候选肽中的每一个, 计算取代残基之前的所述肽的所述预测信息与所述候选肽的所述预测信息之间的差分, 确定所述差分为规定阈值以上的所述候选肽, 并对在所确定的所述候选肽中的取代后的所述残基的位置进行确定;

将在所述肽的结构中的所述确定的位置的残基设定为交换候选的残基, 并将所述交换候选的残基作为变更位置的候选。

12. 根据权利要求1至11中任一项所述的信息处理装置, 所述信息处理装置还包括报酬赋予部,

在从所述终端发送的作为所述肽的所述预测信息的监督数据的与所述肽的药代动力学相关的实验数据、表示获得所述实验数据时的实验方法的信息、用户使用的预测模型、以及由所述用户生成的与所述肽的药代动力学相关的预测信息中的至少一个数据被存储在由所述信息处理装置管理的存储部中的情况下, 所述报酬赋予部针对被赋予给所述终端的用户ID、赋予表示针对提供所述数据的报酬的积分。

13. 一种信息处理方法, 所述信息处理方法由计算机执行如下处理:

接收从终端发送的请求信号,

响应于所述请求信号, 生成与肽的药代动力学相关的预测信息,

向所述终端发送所生成的所述预测信息。

14. 一种信息处理程序, 所述信息处理程序用于使计算机执行如下处理:

接收从终端发送的请求信号,

响应于所述请求信号, 生成与肽的药代动力学相关的预测信息,

向所述终端发送所生成的所述预测信息。

15. 一种信息处理系统, 所述信息处理系统包括终端和信息处理装置,

所述终端向所述信息处理装置发送请求信号,

所述信息处理装置接收从所述终端发送的所述请求信号, 响应于所述请求信号, 生成与肽的药代动力学相关的预测信息, 向所述终端发送所生成的所述预测信息,

所述终端使显示部显示从所述信息处理装置发送的所述预测信息。

16. 根据权利要求15所述的信息处理系统, 其中,

所述信息处理装置生成包含构成肽的要素的变更位置的候选的设计支援信息, 并将所述设计支援信息向所述终端发送,

所述终端还在所述显示部上显示所述设计支援信息。

## 信息处理装置、方法、系统及记录信息处理程序的介质

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年11月13日提交日本特许厅的、申请号为2020-189856以及2021年2月17日递交日本特许厅的、申请号为2021-023750的日本专利申请的优先权和权益,其全部内容通过引用结合在申请中。

### 技术领域

[0003] 本公开涉及一种信息处理装置、信息处理方法、记录信息处理程序的介质及信息处理系统。

### 背景技术

[0004] 在日本特开2017-37378号公报中公开了:在生物体高分子的结构解析中,在将包含在维度集合中的所有索引维度中的每一个作为坐标轴的多维空间中、进行多个结构的聚类,并将不包含在任何簇中的离群值的结构作为初始结构来进行分子动力学模拟(例如权利要求4)。

[0005] 在国际公开第2003/054743号中公开了一种预测蛋白质立体结构的蛋白质立体结构预测程序。执行该蛋白质立体结构预测程序的计算机读取蛋白质的氨基酸序列,并预测二级结构信息。接着,该计算机基于二级结构信息计算形成转角的氨基酸的个数,并通过计算出的氨基酸的个数和二级结构信息获得具有存在概率高的转角结构信息,预测并再现转角,从而预测蛋白质的立体结构。

[0006] 另外,在日本特表2020-523010号公报中公开了一种如下的方法:通过将新生抗原集合中的每一个新生抗原的肽序列输入到进行了机械学习的呈递模型中,来对各患者生成对于患者的新生抗原集合的数值呈递可能性集合(例如权利要求1)。

[0007] 此外,在日本特表2020-519246号公报中公开了一种如下的方法:使用计算机处理器将肽的数值矢量输入到深度学习呈递模型中,来生成对于新生抗原集合的呈递可能性集合(例如权利要求1)。

### 发明内容

[0008] 发明要解决的课题

[0009] 近年来,作为中分子药物之一的肽药物受到关注。但是,关于肽的药代动力学还有很多不清楚的地方。

[0010] 在上述日本特开2017-37378号公报、国际公开第2003/054743号、日本特表2020-523010号公报、或日本特表2020-519246号公报中公开的技术是执行生物高分子的分子动力学模拟的技术、通过计算机预测蛋白质立体结构的技术、以及预测对新生抗原有效的肽的技术,并不是预测肽的药代动力学的技术。因此,上述文献的技术存在无法预测肽的药代动力学的问题。

[0011] 本公开考虑到上述事实,其目的在于预测肽的药代动力学。

[0012] 解决课题的手段

[0013] 本公开的第一方面是一种信息处理装置,该信息处理装置包括:接收部,该接收部接收从终端发送的请求信号;预测部,该预测部响应于所述请求信号,生成与肽的药代动力学相关的预测信息;以及发送部,该发送部向所述终端发送由所述预测部生成的所述预测信息。

[0014] 本公开的第二方面是一种信息处理方法,该信息处理方法由计算机执行如下处理:接收从终端发送的请求信号,响应于所述请求信号,生成与肽的药代动力学相关的预测信息,向所述终端发送所生成的所述预测信息。

[0015] 本公开的第三方面是一种记录有信息处理程序的记录介质,该信息处理程序用于使计算机执行如下处理:接收从终端发送的请求信号,响应于所述请求信号,生成与肽的药代动力学相关的预测信息,向所述终端发送所生成的所述预测信息。

[0016] 本公开的第四方面是一种信息处理系统,该信息处理系统包括终端和信息处理装置,所述终端向所述信息处理装置发送请求信号,所述信息处理装置接收从所述终端发送的所述请求信号,响应于所述请求信号,生成与肽的药代动力学相关的预测信息,向所述终端发送所生成的所述预测信息,所述终端使显示部显示从所述信息处理装置发送的所述预测信息。

[0017] 发明效果

[0018] 根据本公开,能够得到可预测肽的药代动力学这样的效果。

## 附图说明

[0019] 图1是示出了第一实施方式所涉及的信息处理系统的框图。

[0020] 图2是用于说明公共数据库142中存储的数据的一个示例的图。

[0021] 图3是用于说明在用户用数据库144中存储的数据的一个示例的图。

[0022] 图4是用于说明本实施方式的预测模型的图。

[0023] 图5是实现信息处理系统的各设备的计算机的示意图。

[0024] 图6是在信息处理系统中执行的序列。

[0025] 图7是在终端的显示部上显示的画面的一个示例的示意图。

[0026] 图8是在信息处理系统中执行的序列。

[0027] 图9是设计支援信息的一个示例的示意图。

[0028] 图10是示出了第二实施方式所涉及的信息处理系统的框图。

[0029] 图11是用于说明第二实施方式的分数的图。

[0030] 图12是用于说明候选肽的图。

[0031] 图13是设计支援信息的一个示例的示意图。

[0032] 图14是示出了第四实施方式所涉及的信息处理系统的框图。

[0033] 图15是存储在报酬存储部中的表的一个示例的示意图。

## 具体实施方式

[0034] 下面将参考附图,对本发明的实施方式进行详细地说明。

[0035] [第一实施方式]

[0036] <信息处理系统的结构>

[0037] 图1是示出了第一实施方式所涉及的信息处理系统10的结构的一个示例的框图。如图1所示,信息处理系统10在功能上包括多个终端12A、12B、12C以及作为信息处理装置的一个示例的服务器14。多个终端12A、12B、12C与服务器14例如经由互联网等网络16连接。另外,以下在指示一个终端的情况下,简称为终端12。另外,在图1中,例示了多个终端12A、12B、12C,但信息处理系统10也可以包括更多或更少的终端。

[0038] (终端)

[0039] 多个终端12A、12B、12C中的每一个由多个不同的用户操作。用户将表示自己想分析的肽的肽信息输入到自己操作的终端12。然后,用户操作终端12以将该肽信息发送给后述的服务器14。终端12将从用户输入的肽信息经由网络16发送给服务器14。另外,肽信息包括肽的结构式、肽的SMILES(Simplified molecular input line entry system,简化分子线性输入规范)表示、肽的一级结构、肽的二级结构、肽的三级结构及肽的四级结构中的至少一种信息。

[0040] (服务器)

[0041] 如图1所示,服务器14包括接收部140、公共数据库142、多个用户用数据库144A、144B、144C、公共预测模型存储部145、多个用户用预测模型存储部146A、146B、146C、预测部148、处理部150、以及发送部152。另外,以下,在指示一个用户用数据库的情况下,将其简称为用户用数据库144。此外,在指示一个用户用预测模型存储部的情况下,将其简称为用户用预测模型存储部146。另外,在图1中,例示了多个用户用数据库144A、144B、144C,但服务器14也可以包括更多或更少的用户用数据库。此外,同样地,在图1中例示了多个用户用预测模型存储部146A、146B、146C,但服务器14也可以包括更多或更少的用户用预测模型存储部。

[0042] 接收部140接收从终端12发送的请求信号,该请求信号包括用于识别终端的用户ID、肽信息、以及后述的预测模型的选择信息。

[0043] 在公共数据库142中,存储有管理服务器14的管理员的数据、以及操作终端12的用户的数据。图2示出了公共数据库142中存储的数据的一个示例。如图2所示,各数据被赋予了数据ID。该数据ID是用于识别数据的识别信息。此外,数据ID也是用于识别数据是来自管理员还是用户的信息。

[0044] 如图2所示,与数据ID“00001”~“00004”等对应的数据是来自管理员的数据,并且,肽信息、肽信息的特征信息、肽的膜透过性的实验值、获得肽的膜透过性的实验值时的实验方法、肽的体内持续性的实验值、获得肽的体内持续性的实验值时的实验方法、注释、由后述的预测部148生成的预测信息被相关联地存储。另外,肽的体内持续性的值例如为肽的血浆蛋白结合率(Plasma Protein Binding)。注释是表示如何获得该数据的信息等。另外,在注释栏中,也可以存储表示该数据的可靠度的分数。与数据ID“00001”~“00004”等对应的数据由服务器14的管理员预先存储在公共数据库142中。

[0045] 另一方面,如图2所示,与数据ID“A\_00001”、“A\_00002”、“B\_00001”等对应的数据是来自用户的数据,并且肽信息没有被相关联、而肽信息的特征信息、肽的膜透过性的实验值、获得肽的膜透过性的实验值时的实验方法、肽的体内持续性的实验值、获得肽的体内持续性的实验值时的实验方法、注释、由后述的预测部148生成的预测信息被相关联地存储。

与数据ID“A\_00001”、“A\_00002”、“B\_00001”等对应的数据,经由用户的承诺被存储在公共数据库142中。

[0046] 在多个用户用数据库144A、144B、144C的每一个中,存储有来自用户的数据。图3示出了用户用数据库144中存储的数据的一个示例。如图3所示,在用户用数据库144中,肽信息、肽的特征信息、肽的膜透过性的实验值、获得肽的膜透过性的实验值时的实验方法、肽的体内持续性的实验值、获得肽的体内持续性的实验值时的实验方法、注释、由后述的预测部148生成的预测信息被相关联地存储。另外,关于存储在用户用数据库144中的各数据,也可以欠缺其一部分或全部。

[0047] 如上所述,在公共数据库142中不存储来自用户的数据中的肽信息,但在用户用数据库144中存储肽信息。肽信息是隐私性高的信息。因此,可以认为用户对将自己的肽信息存储在还存储有其他用户和管理员的数据的公共数据库142中具有抵抗。

[0048] 因此,在本实施方式中,设置与公共数据库142不同的用户用数据库144,仅在用户用数据库144中存储用户的肽信息。另一方面,在公共数据库142中不存储用户的肽信息。

[0049] 另外,从肽信息获得的特征信息,是关注肽的各种局部或整体结构、并汇集通过可逆或不可逆的运算处理计算出的大量数值而构成的矢量形式的信息,与肽信息相比隐私性较低。因此,关于特征信息,考虑到也存在认为也可以存储在公共数据库142中的用户。

[0050] 因此,在本实施方式中,关于特征信息,存储在公共数据库142中。存储在该公共数据库142中的信息用于生成预测模型等,该预测模型用于预测肽的膜透过性或体内持续性。后面将描述预测模型。此外,存储在公共数据库142中的信息被作为可由管理服务器14的管理员使用的信息来处理。另外,存储在公共数据库142中的信息也作为不仅可由管理者、还可由所有用户使用的信息来处理。另外,存储在公共数据库142中的信息中的一部分也可以作为用户无法使用的信息而设定。

[0051] 公共预测模型存储部145存储有预测模型。

[0052] 在多个用户用预测模型存储部146A、146B、146C的每一个中分别存储有用户用的预测模型。

[0053] 图4示出了本实施方式的预测模型的一个示例。图4示出了作为预测模型的一个示例的、膜透过性预测用学习完成模型M1、体内持续性预测用学习完成模型M2、以及用于实施与肽的药代动力学相关的分子动力学模拟的模拟模型M3。如图4所示,当从肽信息获得的特征信息被输入到膜透过性预测用学习完成模型M1中时,从模型输出与该肽对应的膜透过性的预测值。此外,当从肽信息获得的特征信息被输入到体内持续性预测用学习完成模型M2中时,从模型输出与该肽对应的体内持续性的预测值。并且,当肽信息被输入到模拟模型M3中时,从模拟模型M3输出与该肽的药代动力学相关的分子动力学模拟的轨迹数据。另外,从模拟模型M3输出的数据为轨迹数据,但也存在如下的情况:通过对轨迹数据进行统计分析等,来计算肽的膜透过性或体内持续性的预测值,并从模拟模型M3输出这些值。

[0054] 另外,这些学习完成模型是将存储在公共数据库142或用户用数据库144中的数据的全部或一部分作为学习用数据而生成的。具体地,与特征信息相关联的实验值被用作监督数据,并且通过有监督机器学习生成学习完成模型。这些学习完成模型例如通过神经网络(例如,还包括通过深度学习来学习的深度神经网络)、以及支持向量机等来实现。另外,学习完成模型不限于上述这些机器学习模型,也可以通过逻辑回归等方法来实现。

[0055] 此外,在公共预测模型存储部145中,将存储在公共数据库142中的数据作为学习用数据而生成的学习完成模型、以及模拟模型作为预测模型被存储。另一方面,在多个用户用预测模型存储部146A、146B、146C的每一个中,将存储在用户用数据库144中的数据作为学习用数据而生成的学习完成模型、以及为每个用户准备的模拟模型作为预测模型被存储。

[0056] 在预测肽的行为时,可以假定,某个用户A想要解析某种肽的行为,而另一个用户B想要解析另一种肽的行为。此外,由于存储在用户用数据库144中的数据对于每个用户是不同的,因此可以考虑预测模型的最佳构成方法是不同的。因此,在本实施方式中,为每个用户准备预测肽的药代动力学的预测模型。

[0057] 具体地,在用户用预测模型存储部146A中存储有用户A用的预测模型。此外,在用户用预测模型存储部146B中存储有用户B用的预测模型。在用户用预测模型存储部146C中存储有用户C用的预测模型。由此,根据用户想解析的肽来准备预测模型,用户可以使用该预测模型来获得肽的预测信息。

[0058] 另一方面,还可以认为存在这样的用户,即,希望使用一般的预测模型而不是使用从自己的数据生成的固有的预测模型、来预测肽的药代动力学。因此,在本实施方式中,在公共预测模型存储部145中存储有将存储在公共数据库142中的数据作为学习用数据而生成的学习完成模型、以及标准的模拟模型。

[0059] 因此,在从终端12发送的请求信号中包含的预测模型的选择信息、表示用户用预测模型存储部146中存储的预测模型的情况下,预测部148基于由接收部140接收到的用户ID,从与该用户ID对应的用户用预测模型存储部146读取预测模型。

[0060] 另一方面,在从终端12发送的请求信号中包含的预测模型的选择信息、表示公共预测模型存储部145中存储的预测模型的情况下,预测部148从公共预测模型存储部145读取预测模型。

[0061] 然后,预测部148通过已知的方法将由接收部140接收到的肽信息转换为特征信息。接着,预测部148通过将肽信息和特征信息中的至少一个输入到读取出的预测模型中,生成与肽信息对应的预测信息。

[0062] 例如,在预测模型是膜透过性预测用学习完成模型M1的情况下,膜透过性的预测值作为预测信息而生成。在预测模型是体内持续性预测用学习完成模型M2的情况下,体内持续性的预测值作为预测信息而生成。此外,在预测模型是模拟模型M3的情况下,预测部148通过将由接收部140接收到的肽信息输入到模拟模型M3,从而使用已知的分子动力学模拟方法来模拟体内的肽的行为。由此,生成与肽的药代动力学相关的预测信息。

[0063] 另外,公共预测模型存储部145和用户用预测模型存储部146也可以存储生成相同类型的预测信息的多个预测模型。例如,用户用预测模型存储部146还可以存储有多个膜透过性预测用学习完成模型。另外,用户用预测模型存储部146还可以存储有多个模拟模型。

[0064] 例如,还可以假定这样的情况,即,对于某一肽信息,用户希望使用存储在用户用预测模型存储部146中的自己的膜透过性预测用学习完成模型X和膜透过性预测用学习完成模型Y、以及存储在公共预测模型存储部145中的膜透过性预测用学习完成模型Z和膜透过性预测用学习完成模型W中的每一个,来生成肽的预测信息。另外,对于体内持续性预测用学习完成模型以及模拟模型也假定同样的情况。



[0065] 在这种情况下,用户也可以使用上述相同类型的多个预测模型,从而对一个肽信息生成多个预测信息。在这种情况下,例如,用户可以确认由相同类型的多个预测模型生成的预测信息中的每一个,并研究哪个预测信息是合适的。或者,用户也可以对由相同类型的多个预测模型生成的预测信息中的每一个进行平均等,从而得到合适的预测信息。另外,在这种情况下,从预测对象的肽信息生成特征信息的处理只进行一次即可,通过将所生成的特征信息输入到多个预测模型,从而生成多个预测信息。

[0066] 另外,公共预测模型存储部145还存储有生成预测信息时的各种参数。这些各种参数在使用预测模型生成预测信息时使用。此外,在多个用户用预测模型存储部146A、146B、146C的每一个中,也存储有生成预测信息时的各种参数,这些各种参数针对每个用户而不同。由于使生成预测信息时的各种参数针对每个用户是不同的,能够根据用户的期望生成合适的预测信息。另外,公共预测模型存储部145或用户用预测模型存储部146也可以存储相同类型的多个各种参数。用户在获得肽的预测信息时,从这些各种参数中适当选择认为合适的参数,从而使服务器14生成肽的预测信息。

[0067] 处理部150将由接收部140接收到的肽信息、与由预测部148获得的特征信息及预测信息相关联地存储在与用户ID对应的用户用数据库144中。

[0068] 发送部152,将由预测部148生成的预测信息,发送给与由接收部140接收到的用户ID对应的终端12。

[0069] 终端12接收从服务器14发送的预测信息,并使显示部(省略图示)显示该预测信息。

[0070] 终端12以及服务器14例如能够通过图5所示的计算机50来实现。实现终端12以及服务器14的计算机50包括CPU 51、作为临时存储区域的存储器52、以及非易失性的存储部53。此外,计算机包括连接有输入输出装置等(省略图示)的输入输出界面(interface) (I/F) 54、以及控制对记录介质59读取和写入数据的读/写(read/write) (R/W) 部55。此外,计算机包括连接到互联网等的网络I/F 56。CPU 51、存储器52、存储部53、输入输出I/F 54、R/W 部55以及网络I/F 56经由总线57相互连接。

[0071] 存储部53能够通过硬盘驱动器(Hard Disk Drive,HDD)、固态硬盘(Solid State Drive,SSD)、闪存等来实现。在作为存储介质的存储部53中存储有用于使计算机发挥功能的程序。CPU 51从存储部53读取程序并扩展到存储器52中,依次执行程序所具有的进程(process)。

[0072] 接着,将对实施方式的信息处理系统10的作用进行说明。

[0073] 当操作终端12的用户执行将预测对象的肽信息输入到终端12、并将该肽信息发送给服务器14的操作时,执行图6所示的序列。具体地,在操作终端12的用户执行将预测对象的肽信息输入到终端12、并将该肽信息发送给服务器14、并且将通过学习完成模型预测肽的行为的意思的指示信号发送给服务器14的情况下,执行图6所示的序列。另外,在图6的序列中,将以选择一个预测模型并且使用该预测模型生成预测信息的情况为例进行说明。此外,如上所述,也存在通过多个预测模型的每一个生成预测信息的情况。在这种情况下,指定多个预测模型的信息包含在后述的预测模型的选择信息中。

[0074] 在步骤S100,终端12向服务器14发送请求信号,该请求信号包括由用户输入的肽信息和自己的用户ID。另外,在请求信号中还包含预测模型的选择信息,该预测模型的选择

信息表示是使用存储在用户用预测模型存储部146中的用户自己的预测模型来生成预测信息、还是使用存储在公共预测模型存储部145中的预测模型来生成预测信息。

[0075] 在步骤S102中,服务器14的接收部140接收在上述步骤S100中从终端12发送的请求信号。

[0076] 在步骤S104中,服务器14的预测部148根据在上述步骤S102中接收到的请求信号所包含的用户ID,确定终端12的用户。

[0077] 在步骤S106中,服务器14的预测部148基于在步骤S102中接收到的请求信号中包含的预测模型的选择信息,来判断用于生成预测信息的预测模型是哪一个。在选择信息表示使用用户自己的预测模型的情况下,服务器14的预测部148从与用户ID对应的用户用预测模型存储部146读取预测模型。另一方面,在选择信息表示使用公共预测模型存储部145的预测模型的情况下,服务器14的预测部148从公共预测模型存储部145读取预测模型。

[0078] 在步骤S108中,服务器14的预测部148将在上述步骤S102中接收到的请求信号中所包含的肽信息转换为特征信息。

[0079] 在步骤S110,服务器14的预测部148通过将在上述步骤S108中获得的特征信息输入到在上述步骤S106中读取到的预测模型中,来生成肽的预测信息。

[0080] 在步骤S112中,服务器14的发送部152将在上述步骤S110中获得的预测信息发送给终端12。另外,发送部152也可以将在上述步骤S110中获得的预测信息与预测对象的肽信息(例如,肽的结构式等)相关联,并发送给终端12。

[0081] 在步骤S114中,终端12接收在上述步骤S112中发送的预测信息。

[0082] 在步骤S116中,终端12使显示部(省略图示)显示在上述步骤S114中接收到的预测信息。

[0083] 在这种情况下,如图7所示,在终端12的显示部(省略图示)上显示表示肽的结构式的信息、肽的SMILES表示和预测信息。图7是肽的预测信息与作为预测对象的肽的结构式相关联的信息的一个示例。

[0084] 另外,作为预测对象的肽也可以不是一个而是多个。在这种情况下,对于多个肽中的每一个,肽信息(例如,表示肽的结构式的信息)、肽的SMILES表示以及预测信息显示在终端12的显示部(省略图示)上。另外,在这种情况下,对于多条肽信息中的每一条重复步骤S106~S110的处理,并且在终端12的显示部(省略图示)上显示多条肽信息中的每一条的预测信息。此外,当在步骤S106中读取到的预测模型为多个时,针对一个肽信息通过多个预测模型生成预测信息。

[0085] 由此,用户可以仅通过操作自己的终端12,来获得与肽的药代动力学相关的预测信息。

[0086] 在步骤S118中,服务器14的处理部150将在上述步骤S102中接收到的请求信号所包含的肽信息、在上述步骤S108中获得的特征信息、在上述步骤S110中生成的预测信息相关联地保存到与用户ID对应的用户用数据库144中。

[0087] 在步骤S120中,服务器14的处理部150将在上述步骤S108中获得的特征信息和在上述步骤S110中生成的预测信息存储在公共数据库142中。

[0088] 另外,在操作终端12的用户执行将预测对象的肽信息输入到终端12、并将该肽信息发送到服务器14、并且将通过分子动力学模拟预测肽的行为的意思的指示信号发送给服

务器14的情况下,执行图8所示的序列。

[0089] 图8所示的步骤S100~步骤S106以与图6相同的方式执行。

[0090] 在步骤S210中,服务器14的预测部148通过将在步骤S102中接收到的肽信息输入到作为预测模型的模拟模型中,从而使用已知的分子动力学模拟方法来模拟体内的肽的行为。由此,生成与肽的药代动力学相关的预测信息。

[0091] 在步骤S212中,服务器14的预测部148将在上述步骤S210中生成的预测信息发送给终端12。

[0092] 在步骤S214中,终端12接收在上述步骤S212中从服务器14发送的预测信息。

[0093] 在步骤S216中,终端12使在上述步骤S214中接收到的预测信息显示在显示部(省略图示)上。

[0094] 在步骤S218中,服务器14的处理部150将在上述步骤S102中接收到的请求信号所包含的肽信息、与在上述步骤S210中生成的预测信息相关联地保存到与用户ID对应的用户用数据库144中。

[0095] 另外,这种情况下的肽的预测信息,例如为按时间序列模拟肽分子的运动得到的轨迹数据、及通过对轨迹数据进行统计分析等而得到的膜透过性或体内持续性的预测值等的至少一个信息。此外,基于该轨迹数据,可以通过动态图像使与肽的膜透过性或体内持续性等相关的行为可视化。

[0096] 如以上详细说明所述,信息处理系统的服务器响应于从终端发送的请求信号,生成与肽的体内动态相关的预测信息。然后,服务器向终端发送预测信息。由此,能够预测肽的药代动力学。

[0097] 另外,本公开不限于上述实施方式,在不偏离本公开的要旨的范围内可以进行各种变形和应用。

[0098] 例如,上述实施方式的服务器14以预测膜透过性和体内持续性作为肽的药代动力学的预测信息的情况为例进行了说明,但不限于此。只要是肽的药代动力学,可以是任何信息,例如,作为肽的药代动力学,可以预测血脑屏障(Blood-Brain Barrier)渗透性或体内环境中的肽的溶解度等的物性参数。

[0099] 此外,上述实施方式的服务器14以仅生成肽的预测信息的情况为例进行了说明,但不限于此。例如,以改善肽的药代动力学为目的,服务器14还可以具备设计支援部,该设计支援部生成表示构成肽的要素中的变更位置的候选的设计支援信息。例如,作为构成肽的要素,可以举出肽的残基。在这种情况下,服务器14例如生成包含变更位置的候选R的设计支援信息S。然后,在终端12的显示部上显示图9所示的设计支援信息S。由此,用户设计肽得以支援。另外,服务器14也可以只生成肽的设计支援信息来代替肽的预测信息,并将设计支援信息发送给终端12。

[0100] 此外,在上述实施方式中,以特征信息和预测信息存储在公共数据库142中且肽信息、特征信息及预测信息存储在用户用数据库144中的情况为例进行了说明,但对于它们可以假定各种变形例。

[0101] 例如,在上述实施方式中,以由用户向服务器14发送以作为预测对象的肽信息的特征信息及预测信息全部存储在公共数据库142中的情况为例进行了说明,但不限定于此。

[0102] 例如,服务器14可以仅将肽信息的特征信息和预测信息中的获得用户同意的特征

信息和预测信息存储在公共数据库142中。在这种情况下,例如,服务器14当在图6的步骤S120中将特征信息和预测信息存储在公共数据库142中时,向终端12输出表示是否可以将特征信息存储在公共数据库142中的确认信号。当终端12接收从服务器14发送的确认信号时,终端12使自己的显示部(省略图示)显示用于确认是否可以将预测对象的肽信息的特征信息存储在公共数据库142中的画面。另外,此时,例如,在用户同意在公共数据库142中存储特征信息的情况下,该特征信息或预测信息也在其他用户生成肽的预测信息时被使用,并在终端12的显示部(省略图示)中显示如下的意思的显示:对该用户赋予积分等,以作为提供特征信息或预测信息的报酬。并且,例如,在用户向终端12输入了可以将特征信息或预测信息存储在公共数据库142中的意思的操作情况下,终端12向服务器14发送表示该意思的指示信号。另一方面,在用户向终端12输入了不同意将特征信息或特征信息存储在公共数据库142中的意思的操作的情况下,终端12向服务器14发送表示该意思的指示信号。

[0103] 服务器14根据从终端12发送的指示信号,判断是否将特征信息或预测信息存储在公共数据库142中。然后,在从终端12发送的指示信号表示同意向公共数据库142存储特征信息或预测信息的意思的情况下,服务器14将该特征信息或该预测信息存储在公共数据库142中。另一方面,在从终端12发送的指示信号表示不同意向公共数据库142存储特征信息或预测信息的意思的情况下,服务器14结束处理而不将特征信息或预测信息存储在公共数据库142中。这样,通过根据用户的意向来登记数据,对于用户,服务器14的使用变得容易。另外,如上所述,在用户提供特征信息或预测信息并且特征信息或预测信息被存储在公共数据库142中的情况下,向用户赋予积分。赋予用户的积分例如用于在使用服务器14时的使用费的折扣等。

[0104] 此外,在获得肽的预测信息之后,用户可以进行实验以验证该肽在体内实际上表现出怎样的行为。用户可以确认该实验值,并且存储在服务器14的用户用数据库144和公共数据库142中。在这种情况下,如果用户同意以与上述特征信息和预测信息的情况相同的方式、将实验值和实验方法的信息存储在公共数据库142中并且当通过用户的实验获得的各种信息被存储在公共数据库142中,则对该用户赋予积分等以作为提供通过实验获得的各种信息的报酬。由此,通过用户的实验获得的实验值成为其他用户或服务器14的管理员能够利用的信息,例如,能够作为用户或管理员使预测模型进行机器学习时的监督数据而进行活用。例如,图2所示的c2\_A、d2\_A、e2\_A和f2\_A是用户提供的实验信息的一个示例,并且在如此将实验信息存储在公共数据库142中的情况下,对用户赋予积分。

[0105] 另外,在将数据存储在服务器14的各种存储部中时,用户可以选择服务器14不存储任何数据、仅用户用数据库144存储数据、或者公共数据库142也存储数据中的任意形式。

[0106] 此外,在对用户赋予积分时,服务器14的管理员也可以基于在用户的行为历史或者图2的注释等中存储的可靠度等,来对用户赋予积分。另外,存储在注释等中的可靠度可以是用户提供的数据本身的可靠度和用户的可靠度中的任意一个。在这种情况下,例如,对提供了高质量数据的用户或可靠度高的用户赋予大量积分。

[0107] 此外,在上述实施方式中,关于存储在多个用户用预测模型存储部146A、146B、146C的每一个中的学习完成模型,以仅将存储在用户用数据库144的数据作为学习用数据而生成的学习完成模型的情况为例进行了说明,但不限于此。例如,存储在多个用户用预测模型存储部146A、146B、146C的每一个中的学习完成模型,除了存储在用户用数据库144中

的数据以外,也可以将存储在公共数据库142中的数据作为学习用数据来生成。具体地,除了存储在用户用数据库144中的数据之外,也可以使存储在公共数据库142中的数据的一部分或全部作为学习用数据来生成学习完成模型。由此,生成预测精度高的学习完成模型。另外,在这种情况下,在用户用数据库144或公共数据库142中预先登记有作为学习用数据的各种数据。

[0108] 此外,用户不仅可以通过将上述各种数据存储在服务器14的公共数据库142中来提供自己的数据,还可以将自己制作的预测模型提供给服务器14。在这种情况下,与上述相同,在提供预测模型时,对该用户赋予积分。

[0109] 此外,存储在公共数据库142或用户用数据库144中的数据也可以能够下载到终端12中。并且,存储在公共数据库142或用户用数据库144中的数据也可以处于能够由终端12编辑的状态。另外,在这种情况下,能够由终端12编辑的数据也可以受到限制。例如,关于存储在公共数据库142中的数据,也可以构成为无法由终端12进行编辑。此外,终端12或服务器14也可以使用存储在公共数据库142或用户用数据库144中的数据,生成与肽的药代动力学相关的各种信息。例如,终端12或服务器14也可以生成以肽的特征量(例如脂溶性)为横轴、以膜透过性为纵轴的图表等,从而生成与肽的药代动力学相关的各种信息。

[0110] 此外,在上述实施方式中,以服务器14的预测部148将从终端12发送的请求信号所包含的肽信息转换为特征信息的情况为例进行了说明,但不限于此。例如,在从终端12发送的请求信号中也可以包含特征信息。

[0111] 在使用由服务器14提供的服务的用户中,假定还存在对向服务器14发送肽的结构式等肽信息犹豫不决的用户。因此,例如,也可以通过与用户操作的终端12或服务器14不同的计算机,预先执行从肽信息向特征信息的转换,使请求信号包含该特征信息。在这种情况下,终端12向服务器14发送包含从肽信息转换得到的特征信息的请求信号。服务器14的接收部140接收从终端12发送的请求信号。服务器14的预测部148通过将请求信号中包含的特征信息输入到预测模型,来生成肽的预测信息。然后,服务器14的发送部152将获得的预测信息发送给终端12。由此,用户可以获得肽的预测信息,而无需将肽的结构式等肽信息发送给服务器14。但是,在这种情况下,无法进行将预测模型作为模拟模型的选择。这是因为,在基于模拟模型的预测中需要肽的结构式等肽信息。另外,在这种情况下,例如,假定如下的应对:将由服务器14或与服务器14不同的计算机使用的、从肽信息向特征信息的转换程序等预先提供给终端12或与服务器14不同的计算机等。

[0112] [第二实施方式]

[0113] 接着,对第二实施方式进行说明。第二实施方式的信息处理系统支援由用户进行的肽的设计。另外,在第二实施方式中,对使第一实施方式更具体化的内容进行说明。另外,对于在第二实施方式所涉及的信息处理系统的结构中的与第一实施方式的信息处理系统相同的部分,标记相同的附图标记并省略其说明。

[0114] 当用户设计肽时,可能想要知道具有期望的膜透过性或体内持续性的肽的结构。在这点上,若变更肽的结构中的哪个部分、则成为具有更优选的膜透过性或体内持续性的肽的这一信息,对设计肽的用户是有用的信息。

[0115] 因此,第二实施方式的信息处理系统确定构成肽的多个残基中的对预测信息有特别影响的残基,将该残基作为交换候选的残基呈递给用户。由此,能够支援搜索具有更优选

的膜透过性或体内持续性的肽。此外,还可以为用户提供与肽的设计相关的导航服务。

[0116] 图10是示出了第二实施方式所涉及的信息处理系统210的结构的一个示例的框图。如图10所示,第二实施方式的信息处理系统210的服务器214还包括设计支援部254。设计支援部254将支援由用户设计肽的信息发送给终端12。以下,将具体地进行说明。

[0117] 另外,与第一实施方式相同地,第二实施方式的信息处理系统210的服务器214的预测部148,通过神经网络等学习完成模型来生成肽的预测信息。此外,与第一实施方式相同地,信息处理装置210的预测部148从作为预测对象的肽的肽信息中抽取特征矢量 $x$ 作为特征信息,并将该特征矢量 $x$ 输入到学习完成模型中,从而生成预测信息。

[0118] 另外,从肽所具有的多个残基中的每一个获得特征矢量 $x$ 。例如,特征矢量 $x = [x_{11}, x_{12}, \dots, x_{21}, x_{22}, \dots, x_{N1}, x_{N2}, \dots]$ 中的 $[x_{11}, x_{12}, \dots]$ 是从肽中包含的第一个残基获得的矢量, $[x_{21}, x_{22}, \dots]$ 是从肽中包含的第二个残基获得的矢量,而 $[x_{N1}, x_{N2}, \dots]$ 是从肽中包含的第 $N$ 个残基获得的矢量。

[0119] 首先,设计支援部254对构成肽的多个残基的每一个,通过使用例如已知的技术解析用于生成肽的预测信息的学习完成模型的参数,来计算表示该残基对预测信息的影响程度的分数(以下简称为“残基影响度分数”)。

[0120] 具体地,首先,设计支援部254对输入到学习完成模型的特征矢量 $x$ 中的各要素 $x_{ij}$ ,计算通过要素 $x_{ij}$ 对从学习完成模型输出的预测信息所表示的值 $y$ 进行偏微分而得到的微分值。另外,预测信息所表示的值 $y$ 的要素 $x_{ij}$ 的偏微分由以下算式表示。

[0121] [数式1]

$$[0122] \quad \frac{\partial y}{\partial x_{ij}}$$

[0123] 该微分值通过解析学习完成模型的参数而获得。此外,将以下微分值的绝对值设为表示第 $i$ 个残基的各特征量对预测信息的影响程度的分数(以下简称为“特征量影响度分数”)。

[0124] [数式2]

$$[0125] \quad \frac{\partial y}{\partial x_{i1}}, \frac{\partial y}{\partial x_{i2}}, \dots,$$

[0126] 设计支援部254对特征矢量 $x$ 中的各要素 $x_{ij}$ 计算该特征量影响度分数。

[0127] 图11是用于说明本实施方式的特征量影响度分数的示意图。图11在其他领域也称为显著图。

[0128] 在图11所示的显著图中,横轴上的“1、2、……”表示残基的ID,纵轴上的“ $x_{ij}$ ”表示特征矢量 $x$ 中包含的特征量的种类。此外,显著图的浓淡相当于表示特征矢量 $x$ 的各要素 $x_{ij}$ 对预测的影响程度的特征量影响度分数,显著图的浓淡越浓,表示特征量影响度分数越高。

[0129] 在图11的示例中,例如,ID=7的残基的特征矢量的元素 $x_{71}$ 的特征量影响度分数高。因此,可以预想到若将ID=7的残基交换为其他残基等,则预测信息将发生较大变化。另外,此时,通过确认与特征量影响度分数对应的微分值的正负,也能够示出该残基存在所带来的影响的方向。

[0130] 接着,例如,设计支援部254对多个残基的每一个,通过计算针对特征矢量x的各要素计算出的特征量影响度分数的总和等,来计算该残基的残基影响度分数。并且,设计支援部254在针对多个残基的每一个计算出的该残基的残基影响度分数中,确定具有规定阈值以上的残基影响度分数的残基。

[0131] 另外,根据特征量影响度分数计算残基影响度分数的方法不限于上述方法,例如,也可以将针对特征矢量x的各要素计算出的特征量影响度分数的加权平均、最大值、或最小值等设置为残基影响度分数。

[0132] 接着,设计支援部254将肽的结构中的被确定的残基设定为交换候选的残基,并将交换候选的残基作为变更位置的候选。

[0133] 例如,设计支援部254将残基影响度分数为规定阈值以上的残基设定为变更位置的候选。然后,设计支援部254生成用于建议交换候选的残基的设计支援信息。

[0134] 第二实施方式的服务器214的发送部152将由设计支援部254生成的设计支援信息发送给终端12。另外,发送部152也可以将上述图11所示的显著图作为设计支援信息发送给终端12。

[0135] 另外,第二实施方式的信息处理系统210的其他结构和作用与第一实施方式相同,因此省略对其的说明。

[0136] 如上所述,第二实施方式的信息处理系统的服务器,对构成肽的多个残基的每一个,通过解析用于生成预测信息的学习完成模型的参数,来计算表示该残基对预测信息的影响程度的残基影响度分数。服务器在针对多个残基中的每一个计算出的残基影响度分数中、确定规定阈值以上的残基影响度分数的残基,并将在肽的结构中的被确定的残基设定为交换候选的残基,并将交换候选的残基作为变更位置的候选。此外,服务器向终端发送用于建议变更位置的候选的设计支援信息。终端在显示部(省略图示)上显示从服务器发送的设计支援信息。由此,用户能够获得关于如何变更构成肽的残基中的哪个残基即可的线索。此外,能够支援用户设计肽。

[0137] 另外,第二实施方式的服务器也可以在针对多个残基的每一个计算出的残基影响度分数中、确定不足规定阈值的残基影响度分数的残基,并向用户呈递该信息。在这种情况下,由于可以确定即使变更也不太影响膜透过性或体内持续性的残基,因此用户可以获得关于变更构成肽的残基中的哪个残基即可的线索。

[0138] [第三实施方式]

[0139] 接着,对第三实施方式进行说明。第三实施方式的信息处理系统与第一实施方式及第二实施方式的不同点在于,通过生成多个将构成肽的多个残基中的至少一个残基取代为预定的其他残基而得到的候选肽,并生成多个候选肽的每一个的预测信息,从而确定影响预测信息的残基,具体地,建议残基交换。另外,第三实施方式所涉及的信息处理系统的结构由于是与第二实施方式的信息处理系统相同的结构,因此标记相同的附图标记并省略其说明。

[0140] 第三实施方案的服务器214的设计支援部254生成将构成肽的多个残基中的至少一个残基取代为丙氨酸等预定的其他残基而得到的候选肽。

[0141] 图12是用于说明候选肽的示意图。图12示出了将某种肽S1中的残基Re1取代为其他残基的情况的示例。例如,设计支援部254通过将图12的肽S1中的残基Re1取代为其他残

基Re2 (例如, 甘氨酸), 来生成候选肽S2。此外, 设计支援部254通过将图12的肽S1中的残基Re1取代为其他残基Re3 (例如, 丙氨酸), 来生成候选肽S3。另外, 其他残基例如可以是缬氨酸、亮氨酸、精氨酸或天冬氨酸等天然氨基酸、或任意的人工氨基酸等。

[0142] 并且, 第三实施方式的预测部148使用与第一实施方式相同的方法, 生成取代残基之前的肽的预测信息和多个候选肽的每一个的预测信息。

[0143] 接着, 设计支援部254对多个候选肽的每一个, 计算候选肽的预测信息与取代残基之前的肽的预测信息的差分。然后, 设计支援部254确定差分为规定阈值以上的候选肽, 并对在确定的候选肽中的取代后的残基的位置进行确定。

[0144] 在取代残基之前的肽的预测信息与候选肽的预测信息之间的差分较大的情况下, 取代前的残基和取代后的残基中的至少一方是对预测信息有较大影响的残基。因此, 设计支援部254确定差分较大的候选肽中的取代后的残基的位置, 并在取代残基之前的肽中确定存在于该位置的残基。

[0145] 并且, 设计支援部254将表示肽的结构的信息中的被确定的位置的残基设定为交换候选的残基, 并将交换候选的残基作为变更位置的候选。然后, 设计支援部254生成用于建议将交换候选的残基转换为其他残基的设计支援信息。

[0146] 图13示出了设计支援信息的一个示例。如图13所示, 例如, 设计支援信息可以包括变更残基之前的肽信息 (例如, 肽的结构式)、表示交换候选的残基的信息 (图中虚线部分表示交换候选的残基)、以及肽的预测信息。此外, 例如, 设计支援信息可以包含在交换残基之后的候选肽的肽信息 (例如, 肽的结构式)、候选肽的预测信息、以及新并入候选肽的残基的信息 (例如, 残基的结构式)。另外, 设计支援信息中也可以包含将交换候选的残基取代为多个不同的其他残基并交换候选的残基被取代后的预测信息的变化 (或变化的加权平均)。

[0147] 第三实施方式的服务器214的发送部152将由设计支援部254生成的设计支援信息发送给终端12。

[0148] 另外, 第三实施方式的信息处理系统210的其他结构和作用与第一实施方式或第二实施方式相同, 因此省略对其的说明。

[0149] 如以上说明, 第三实施方式的信息处理系统的服务器生成将构成肽的多个残基中的至少一个残基取代为规定的其他残基而得到的多个候选肽, 并生成取代残基之前的肽的预测信息以及多个候选肽的每一个的预测信息。接着, 服务器对多个候选肽的每一个, 计算取代残基之前的肽的预测信息与候选肽的预测信息之间的差分, 确定差分为规定阈值以上的候选肽, 并对在确定的候选肽中的取代后的残基的位置进行确定。然后, 服务器将肽的结构中的被确定的位置的残基设定为交换候选的残基, 并将交换候选的残基作为变更位置的候选。由此, 用户能够得到关于变更构成肽的残基中的哪个残基即可的线索。此外, 能够支援用户设计肽。

[0150] [第四实施方式]

[0151] 接着, 对第四实施方式进行说明。第四实施方式的信息处理系统与第一实施方式~第三实施方式的不同点在于, 在用户提供规定数据的情况下, 对该用户赋予表示报酬的积分。另外, 第四实施方式所涉及的信息处理系统的结构由于是与第一实施方式的信息处理系统相同的结构, 因此标记相同的附图标记并省略其说明。

[0152] 为了高精度地生成与肽的药代动力学相关的预测信息, 需要大量的数据。例如, 通



过与某种肽的药代动力学相关的规定实验获得的表示膜透过性或体内持续性的实验数据，可以是用于预测其他肽的药代动力学的有用数据。

[0153] 因此，在第四实施方式的信息处理系统中，在用户提供规定数据的情况下，向该用户赋予表示报酬的积分。该积分例如用于使用本信息处理系统提供的服务的使用费的折扣等。

[0154] 图14是示出了第四实施方式所涉及的信息处理系统410的结构的一个示例的框图。如图14所示，第四实施方式的信息处理系统410的服务器414还包括报酬赋予部456和报酬存储部458。

[0155] 报酬赋予部456将用户提供的数据存储到公共数据库142或公共预测模型存储部145中。例如，报酬赋予部456接受从终端12发送的作为肽的预测信息的监督数据的与肽的药代动力学相关的实验数据、表示获得实验数据时的实验方法的信息、保存于用户用预测模型存储部146的预测模型、用户在自己的终端中使用的预测模型（例如，用户在自己的终端独自生成的预测模型或者用户可从外部引入的预测模型等）、以及由用户生成的与肽的药代动力学相关的预测信息中的至少一个数据。然后，报酬赋予部456将接受的数据存储在信息处理装置414管理的存储部的一个示例即公共数据库142或公共预测模型存储部145中。

[0156] 另外，实验数据是通过已知的实验方法获得的与作为预测对象的肽的药代动力学相关的数据。该实验数据例如在生成学习完成模型时作为监督数据使用。此外，关于表示获得实验数据时的实验方法的信息，由于实验数据的可靠性根据使用哪种实验方法而不同，因此是有用的信息。此外，如果提供了由用户独立生成的学习完成模型等预测模型，则也可以基于该预测模型来生成新的预测模型，这同样是有用的。此外，由用户独自生成的肽的预测信息也同样有用。

[0157] 因此，报酬赋予部456对被赋予给发送了数据的终端12的用户ID，赋予表示针对提供该数据的报酬的积分。

[0158] 例如，报酬赋予部456通过更新存储在报酬存储部458中的、用户ID与赋予该用户ID的积分相关联的表，从而对用户ID赋予积分。

[0159] 图15是存储在报酬存储部458中的表的一个示例。在图15所示的示例中，图示了用户ID“USER\_01”和赋予该用户ID的积分“P\_USER\_01”。在报酬赋予部456在从与用户ID“USER\_01”对应的终端12接受了数据的提供的情况下，通过对赋予给用户ID“USER\_01”的积分“P\_USER\_01”加上规定的积分，来更新存储在报酬存储部458中的表。另外，预先设定关于对数据提供赋予何种程度的积分的规则。例如，报酬赋予部456根据提供的数据的种类（例如，提供的数据是否是表示实验值的数据、表示预测模型的数据、以及表示预测信息的数据中的哪一个）、数据的量、数据的历史信息（例如，在提供的数据是表示膜透过性的实验值的数据的情况下，表示该数据是通过PAMPA试验、Caco2细胞、MDCK细胞以及LLC-PK1细胞等中的哪种手段获得的这样的履历信息）、数据的类别（例如，提供的数据是与环状肽相关的数据以及与低分子相关的数据中的哪一种的这样的属性）、以及提供数据的用户的可靠度等，对用户赋予积分。另外，用户的可靠度可以手动设定、或者也可以由诸如服务器414的计算机设定。例如，服务器414的报酬赋予部456根据信息处理系统410所提供的服务的使用频率以及使用历史等，决定用户的可靠度。例如，报酬赋予部456设定为使用频率越高的用户、可

靠度越高。或者,报酬赋予部456根据使用历史设定为使用年数越长的用户、可靠度越高。此外,关于从用户接收数据的提供后的积分,例如,报酬赋予部456可以计算向预测模型反映所提供的数据情况下的对性能提高的贡献程度的估计值,并且基于该估计值改变赋予用户的积分。例如,报酬赋予部456也可以计算从用户提供的数据与服务器414已保持的数据之间的相似度,对提供了相似度低的数据的用户赋予较高的积分。由此,促进提供服务器414未持有的数据。此外,例如,在提供被认为对肽的预测具有不利影响的数据的情况下,报酬赋予部456也可以不赋予积分。

[0160] 这样,通过向用户给予针对提供数据的奖励,在服务器414的公共数据库142或公共预测模型存储部145中能够存储更多的数据,并且能够使这些数据对预测肽的药代动力学有用。对于针对提供预测模型的积分的赋予也与上述大致相同。例如,也可以使用预先准备好的已知正确的试验用数据、来估计所提供的预测模型的性能,并对能够期待发挥优良性能的预测模型赋予较高积分。此外,例如,对于在现有的预测模型中成绩较差的肽,对预测得到改善这样的预测模型也可以赋予较高的积分。

[0161] 如上所述,在从终端发送的作为肽的预测信息的监督数据的实验数据、表示获得实验数据后的实验方法的信息、用户使用的预测模型、以及由用户生成的与肽的药代动力学相关的预测信息中的至少一个数据被存储在由服务器管理的存储部中的情况下,第四实施方式的信息处理系统的服务器,对被赋予给终端的用户ID、赋予表示针对提供数据的报酬的积分。由此,能够收集更多的对预测肽的药代动力学有用的数据。

[0162] 另外,本公开不限于上述各实施方式,在不偏离该公开的要旨的范围内可以进行各种变形和应用。

[0163] 此外,在上述内容中,对本发明所涉及的程序预先存储(安装)在存储部(省略图示)中的方式进行了说明,但本发明所涉及的程序也能够以记录在CD-ROM、DVD-ROM以及微型SD卡等记录介质中的方式来提供。

[0164] 另外,在上述实施方式中由中央处理器(Central Processing Unit,CPU)读取软件(程序)并执行的处理,也可以由CPU以外的各种处理器执行。作为该情况下的处理器,示例出在现场可编程逻辑门阵列(Field-Programmable Gate Array,FPGA)等的制造后能够变更电路结构的可编程逻辑器件(Programmable Logic Device,PLD)及专用集成电路(Application Specific Integrated Circuit,ASIC)等、具有为了执行确定的处理而专门设计的电路结构的处理器即专用电路等。另外,既可以通过这些各种处理器中的一个来执行各处理,也可以通过同种或异种的两个以上的处理器的组合(例如,多个FPGA、以及CPU与FPGA的组合等)来执行各处理。此外,这些各种处理器的硬件结构,更具体地,是组合了半导体元件等电路元件的电路。

[0165] 此外,也可以由具备通用运算处理装置及存储装置等的计算机或服务器等构成本实施方式的各处理,并通过程序执行各处理。该程序存储在存储装置中,并且可以记录在诸如磁盘、光盘或半导体存储器等记录介质中,或者可以通过网络提供。当然,任何其他构成元件都不需要由单个计算机或服务器来实现,而是可以通过分布到经由网络连接的多个计算机来实现。

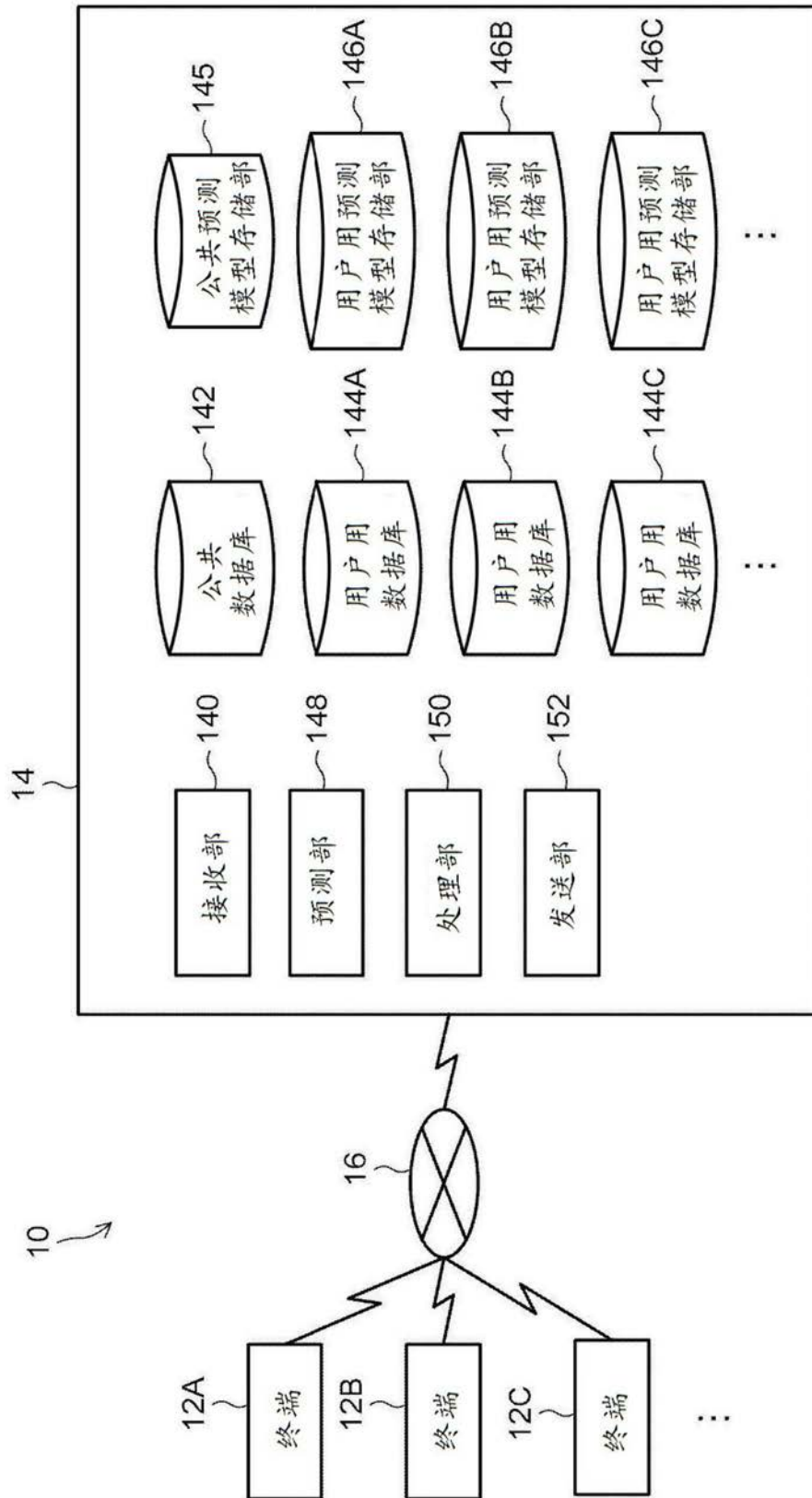


图1

数据 ID	肽信息	特征信息	膜透过性 实验值	膜透过性 实验方法	体内持续性 实验值	体内持续性 实验方法	注释	预测信息	...
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	...
...	...	...	...	...	...	...	...	...	...

图2

数据ID	肽信息	特征信息	膜透过性 实验值	膜透过性 实验方法	体内持续性 实验值	体内持续性 实验方法	注释	预测信息	...
A_00001	a1_A	b1_A	c1_A	d1_A	e1_A	f1_A	g1_A	h1_A	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...

图3

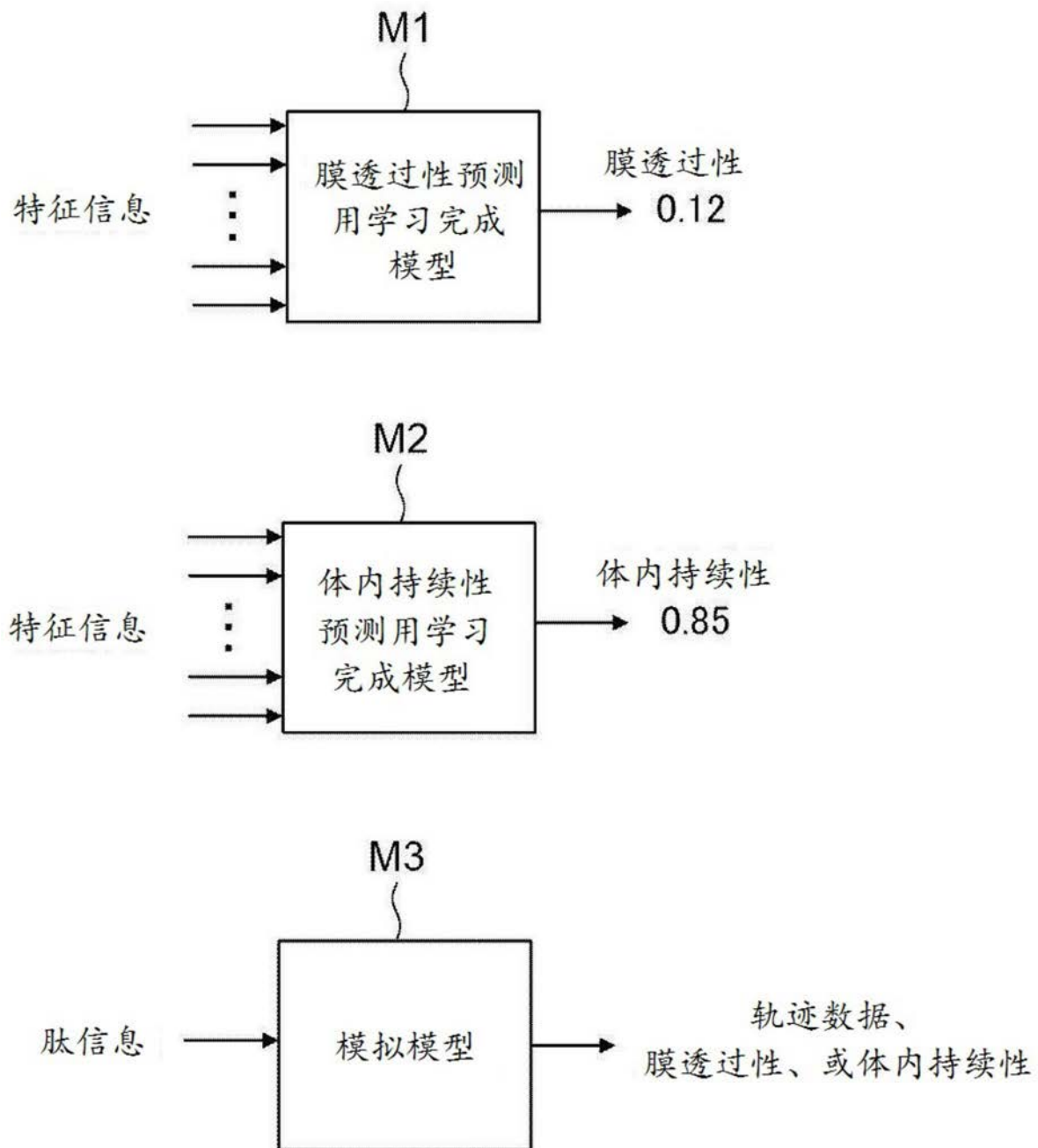


图4

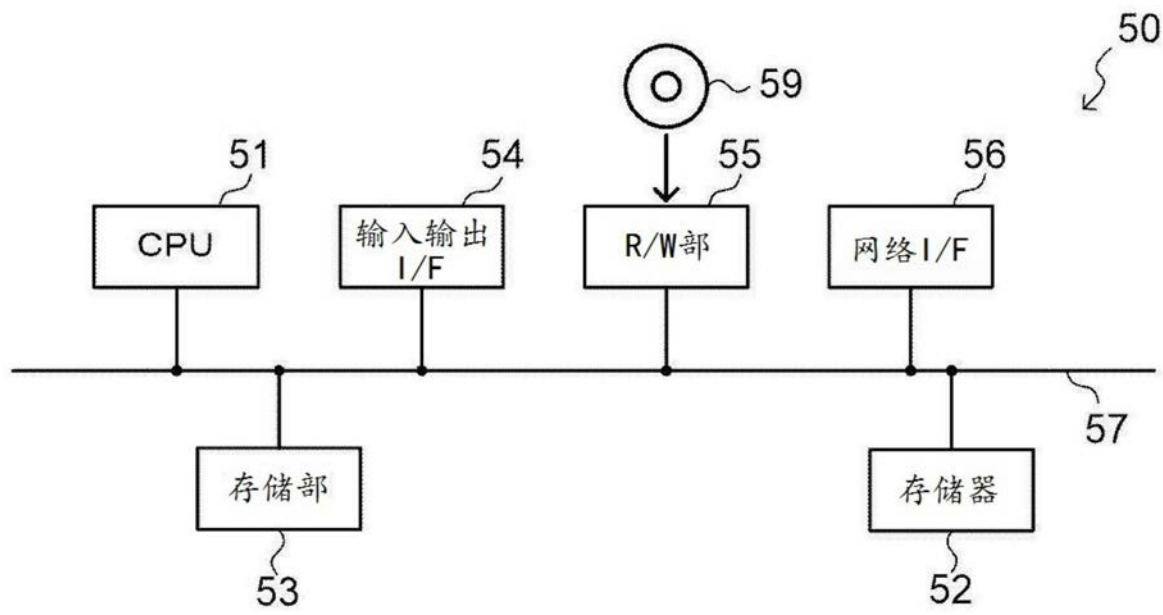


图5

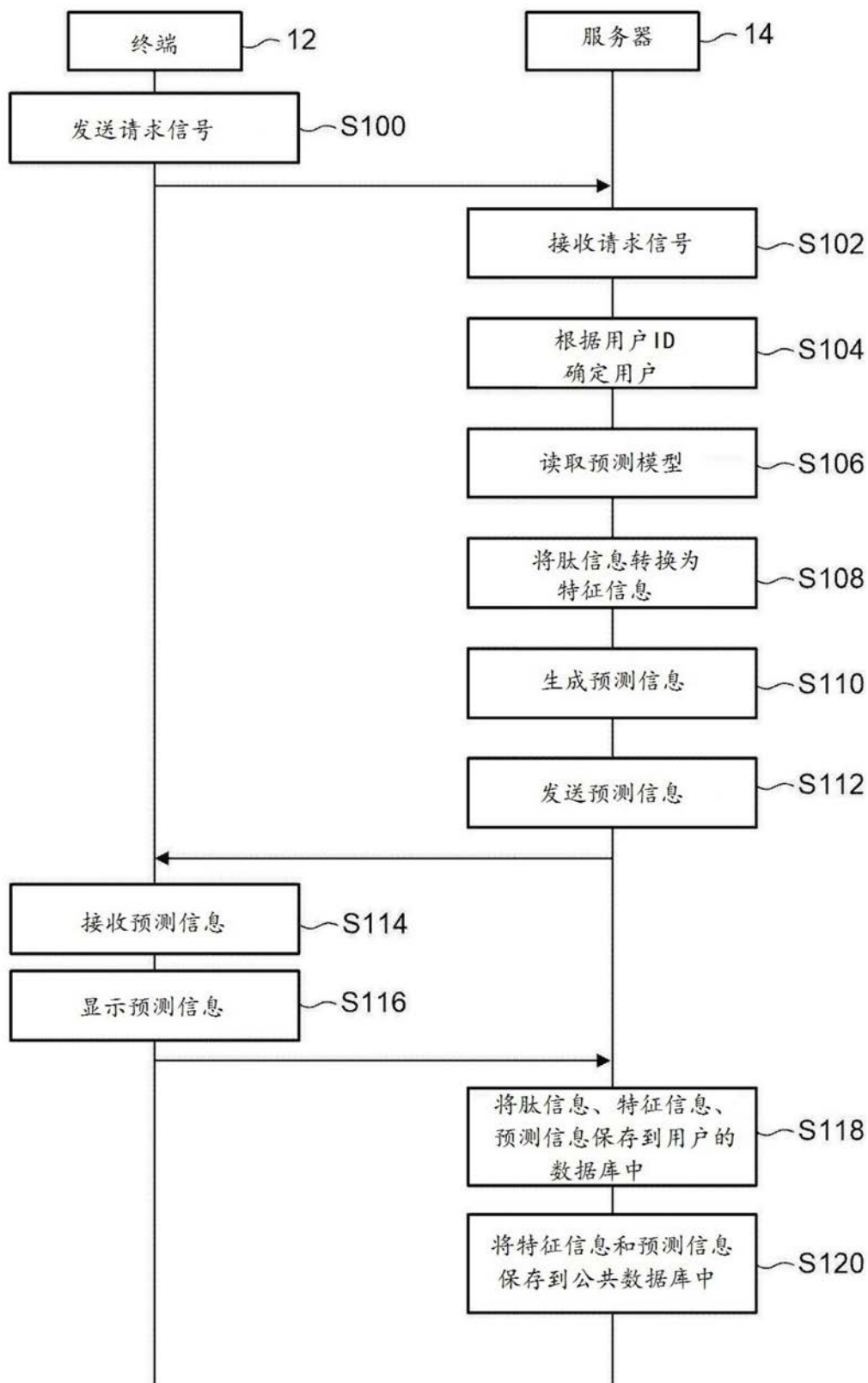


图6



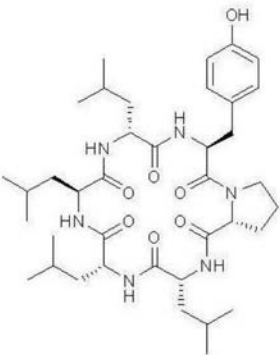
肽ID	肽的结构式	SMILES表示	预测信息
00001		<chem>O=C([C@@H](N1)CC2=CC=C(C=C2)O)N3CCC[C@@H]3C(N[C@H](CC(C)C)C(N[C@H](CC(C)C)C(N[C@@H](CC(C)C)C(N[C@H](CC(C)C)C(N[C@H](CC(C)C)C1=O)=O)=O)=O)=O</chem>	...
00002	...	...	...
...	...	...	...

图7

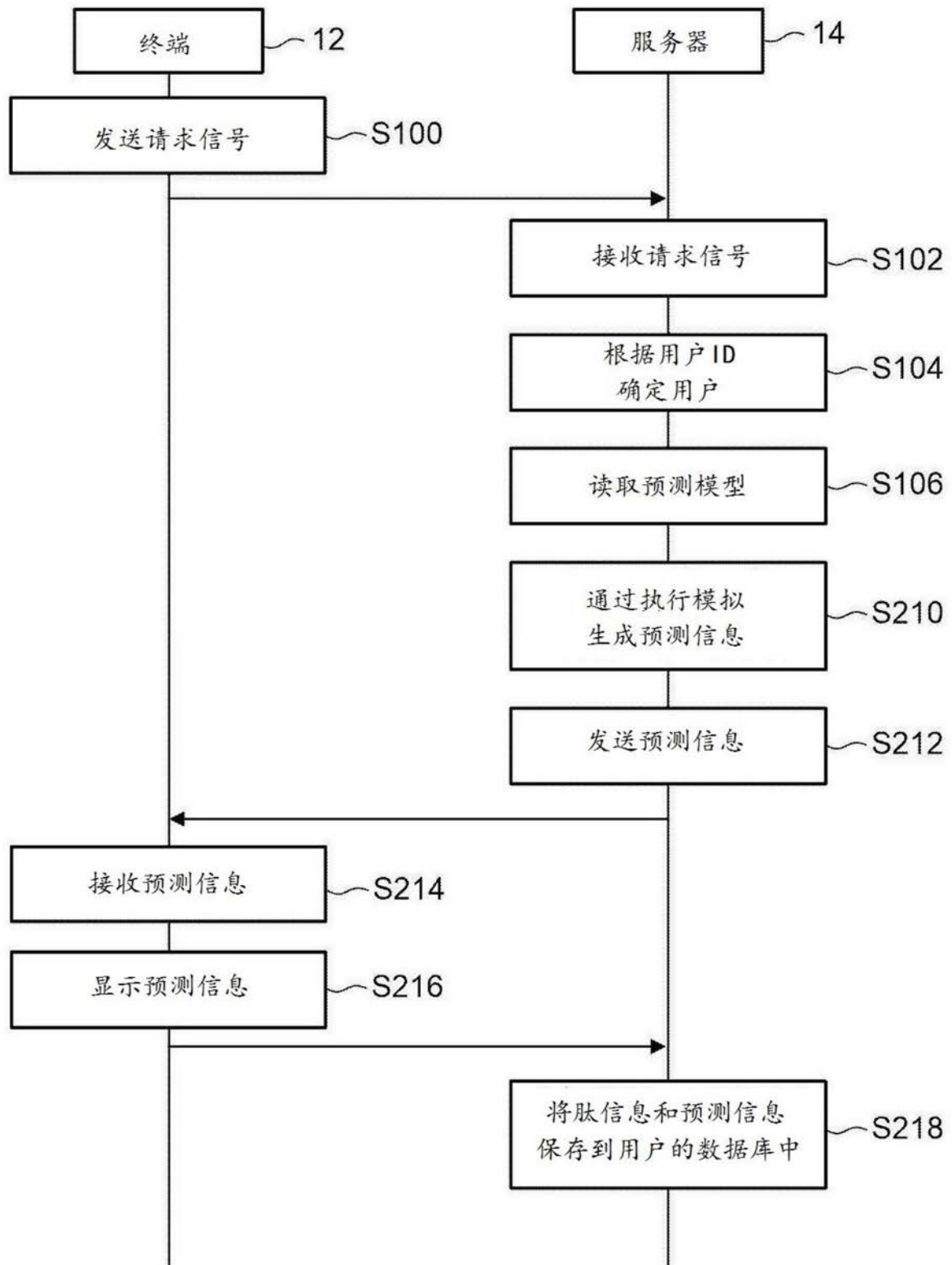


图8

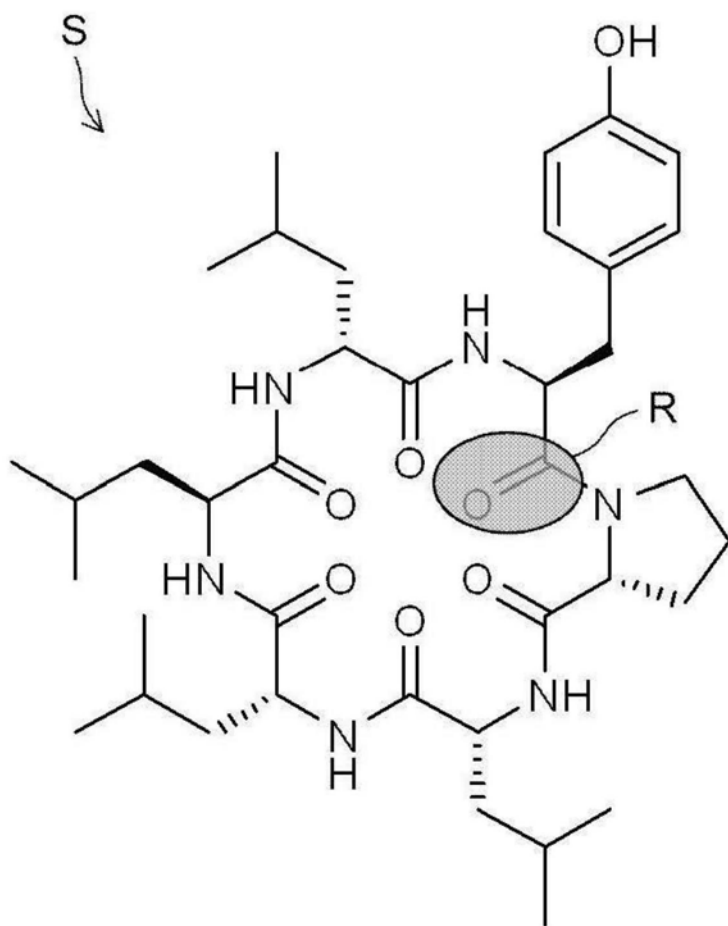


图9

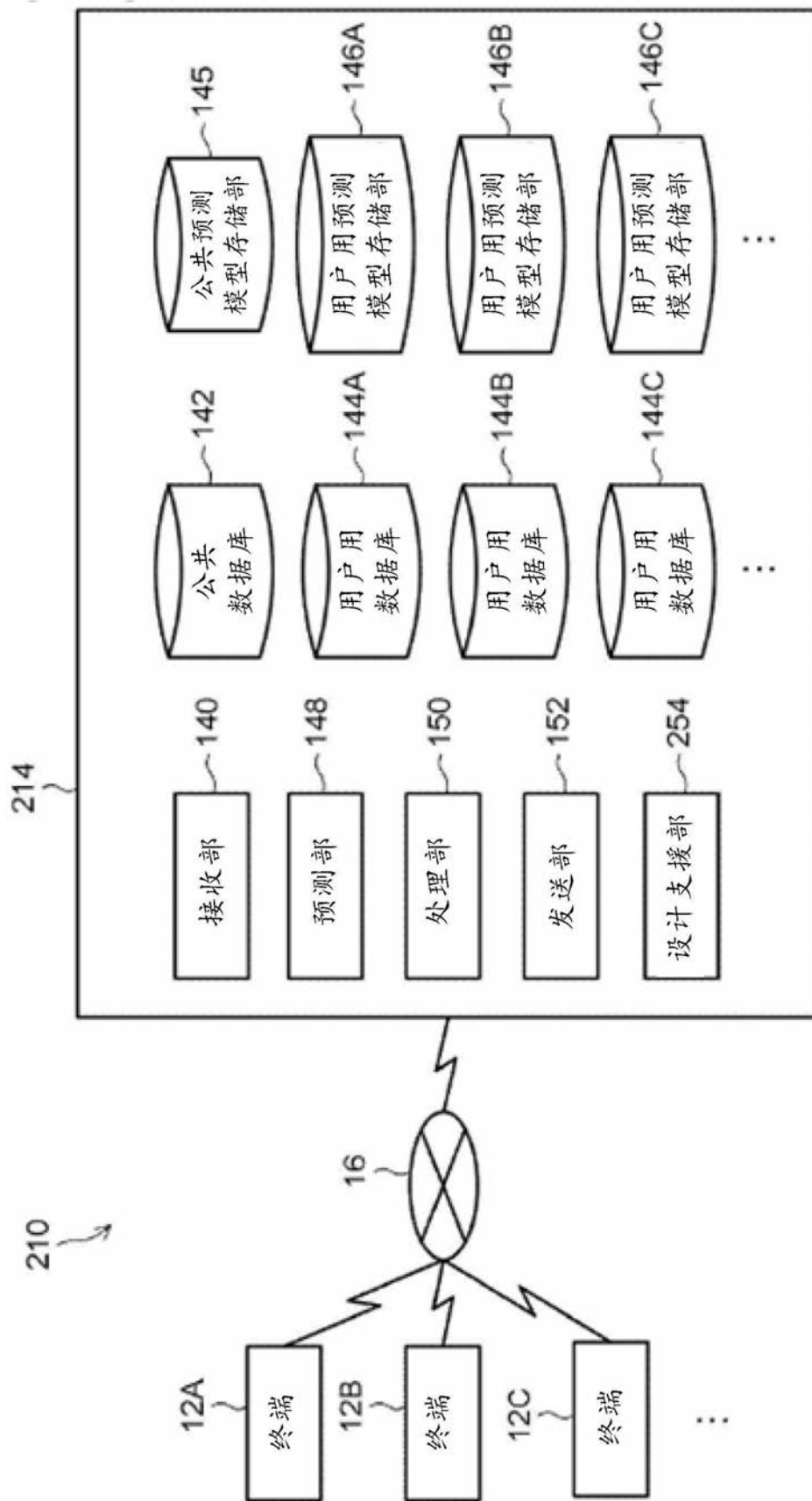


图10

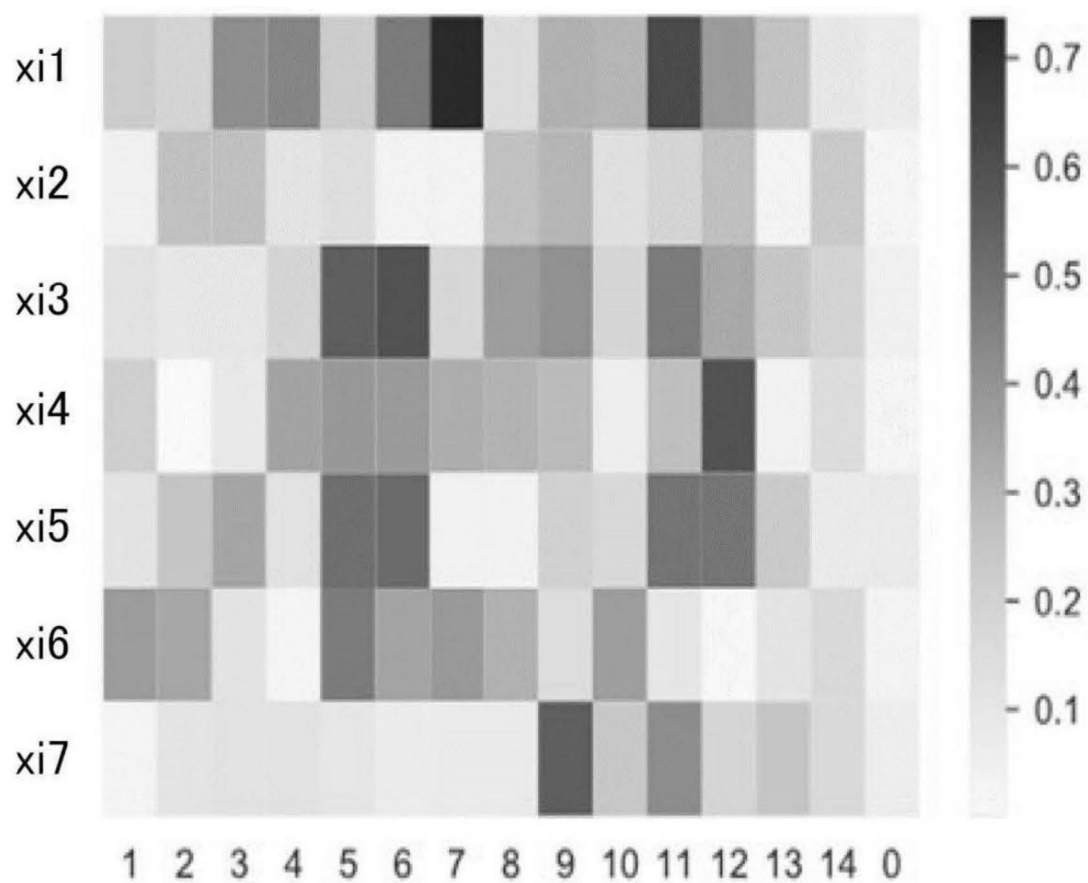


图11

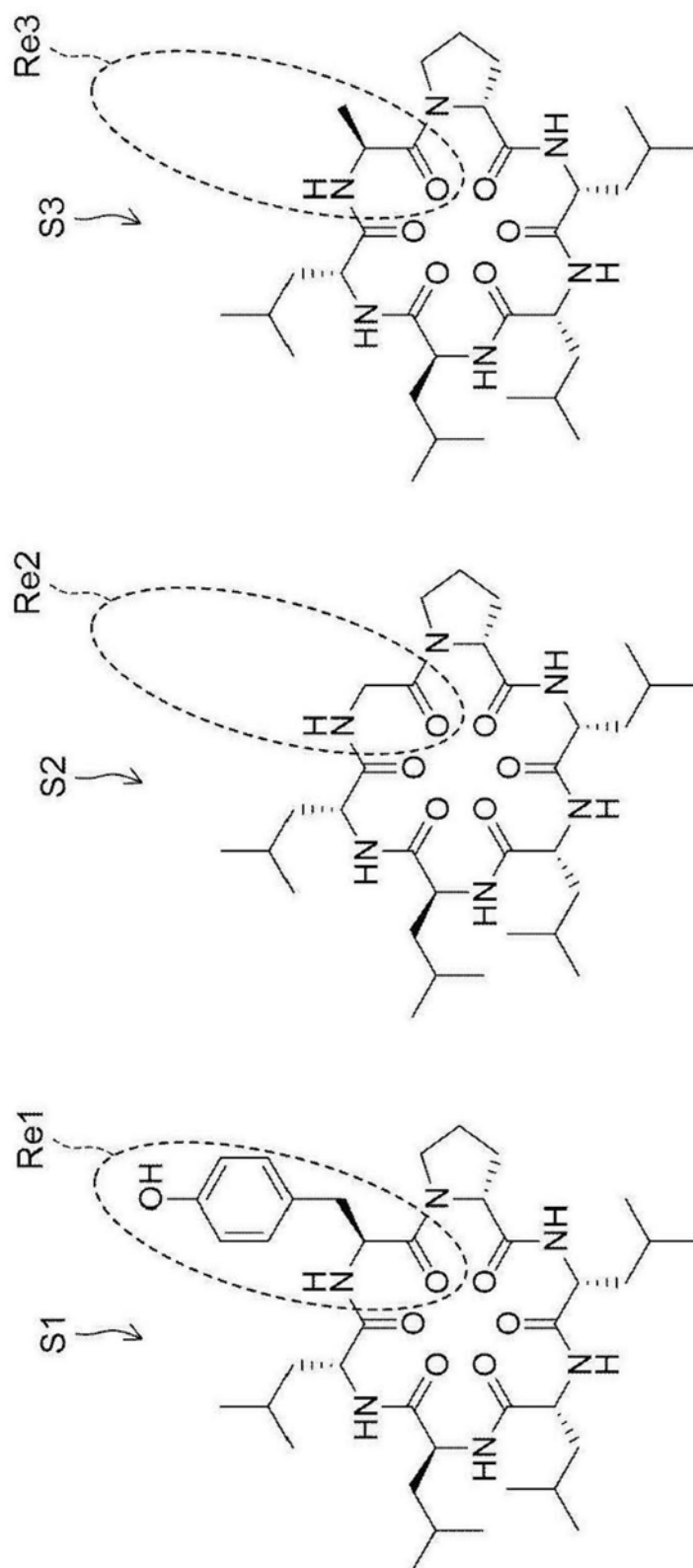
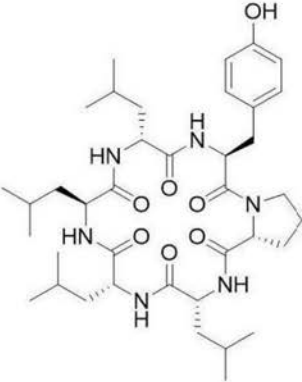
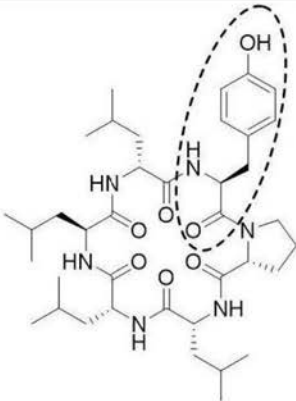


图12

变更残基之前的肽	预测信息	交换候选的残基
	...	

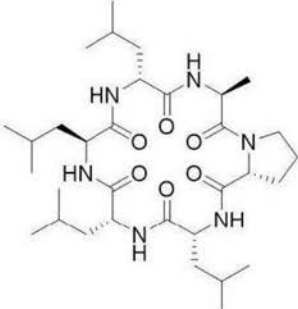
候选ID	变更残基之后的候选肽	预测信息	新并入的残基的信息
0001		...	...
...	...	...	...

图13

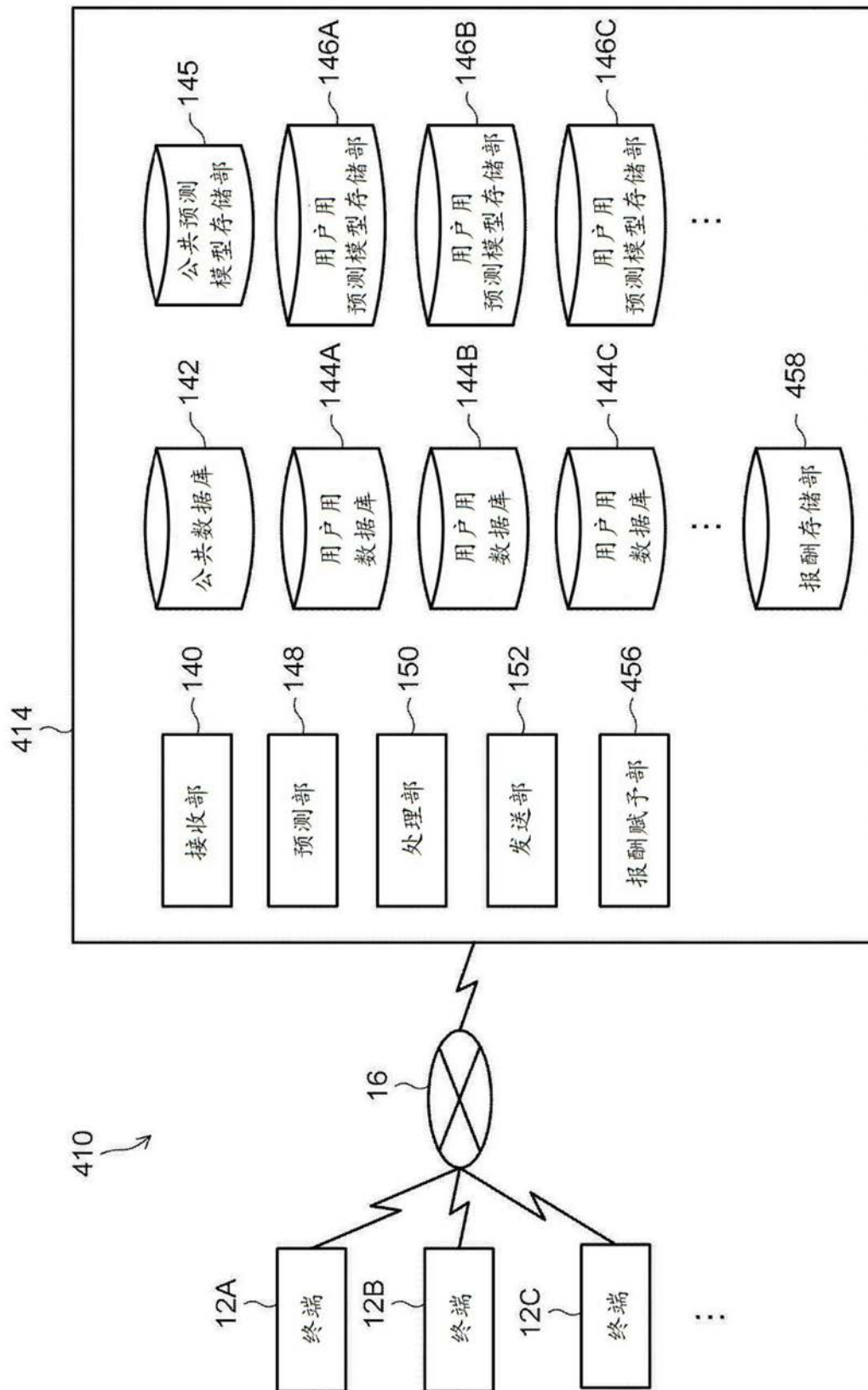


图14



用户 ID	积分	...
USER_01	P_USER_01	...
...	...	...
...	...	...

图15