

(19)



**Евразийское
патентное
ведомство**

(21) **202390440** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.04.03

(51) Int. Cl. **G16C 20/30** (2019.01)
G16C 60/00 (2019.01)

(22) Дата подачи заявки
2021.07.16

**(54) ВЫБОР СМОЛ ДЛЯ ИСПОЛЬЗОВАНИЯ В ПРОЦЕССАХ ОЧИСТКИ ПУТЕМ
ХРОМАТОГРАФИИ**

(31) **63/058,050**

(32) **2020.07.29**

(33) **US**

(86) **PCT/US2021/041973**

(87) **WO 2022/026211 2022.02.03**

(71) Заявитель:
ЭМДЖЕН ИНК. (US)

(72) Изобретатель:

**Ногюэрес Гектор, Коуфаль Мира,
Сото-Ортега Дебора (US)**

(74) Представитель:

Медведев В.Н. (RU)

(57) Способ выбора исходных материалов для использования в процессе очистки путем колоночной хроматографии, в котором для каждой из одной или нескольких кандидатных смол получают соответственный набор значений атрибута смолы, содержащий по меньшей мере одно аналитическое измерение кандидатной смолы. Способ также включает для каждой кандидатной смолы прогнозирование соответственного значения показателя эффективности для процесса очистки путем колоночной хроматографии путем применения соответственного набора значений атрибута смолы и, возможно, других значений параметра (например, параметров собранного фильтрата и/или процесса очистки) в качестве входных данных к многомерной статистической модели. Способ дополнительно включает выбор смолы из одной или нескольких кандидатных смол на основе, по меньшей мере, частично прогнозированного соответственного значения (прогнозированных соответственных значений) показателя эффективности и выполнение процесса очистки путем колоночной хроматографии с использованием выбранной смолы в качестве неподвижной фазы.

A1

202390440

202390440

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-577132EA/022

ВЫБОР СМОЛ ДЛЯ ИСПОЛЬЗОВАНИЯ В ПРОЦЕССАХ ОЧИСТКИ ПУТЕМ ХРОМАТОГРАФИИ

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

[0001] Настоящее изобретение в целом относится к производству биофармацевтических продуктов, а более конкретно к методам облегчения выбора (например, отбора) смол для процессов очистки путем колоночной хроматографии таким образом, чтобы учитывать изменчивость (например, изменчивость от партии к партии) в атрибутах смолы.

УРОВЕНЬ ТЕХНИКИ

[0002] В биофармацевтической промышленности крупные, сложные белковые молекулы, известные как биологические или терапевтические белки, получают из живых систем. На высоком уровне процесс производства терапевтических белков включает следующие стадии: (1) стадия выбора клетки-хозяина, на которой производится основная линия клеток, содержащая ген, который производит желаемый белок (например, с использованием клеток яичника китайского хомячка (СНО)); (2) стадия культивирования клеток, на которой определенные культуральные среды используются для выращивания большого количества клеток, производящих белок, в биореакторах; (3) стадия очистки, на которой происходит восстановление и очистка продукта, полученного на предыдущем этапе, для выделения белка; и (4) стадия приготовления состава, его наполнения и фасовки, на которой белок готовится для использования врачами или пациентами.

[0003] На фиг. 2 изображен стандартный процесс 10 очистки для производства терапевтических белков (т. е. третья стадия, перечисленная выше). На первом этапе 12 терапевтический белок (также называемый «целевой молекулой») собирают из клеточной культуры (например, из биореакторов). Собранный материал также содержит нежелательные побочные продукты, такие как белки клетки-хозяина (НСР), которые были секретированы вместе с целевой молекулой и/или высвобождены в технологический поток во время этапов сбора, а также другие нежелательные вещества (например, деградированные или агрегированные белки). На этапе 14 собранный материал подвергается очистке путем колоночной хроматографии, обычно с использованием нескольких хроматографических колонок. Очистка на этапе 14 играет важную роль в удалении НСР и других примесей из собранного материала. В качестве примера, этап 14 может включать очистку через четыре хроматографические колонки, при этом первая колонка предназначена для выбора целевой молекулы и снижения примесей клетки-хозяина, вторая колонка - для снижения примесей, связанных с продуктом и процессом, третья колонка - для последующего концентрирования желаемого продукта, а четвертая колонка - для дальнейшего повышения чистоты продукта. Затем на этапах 16 и 18 очищенное вещество подвергается вирусной фильтрации и конечной фильтрации соответственно. Очищенное/фильтрованное вещество затем обычно соединяют со

вспомогательным веществом для получения стерильного раствора, который можно вводить или вливать, и помещают в целевой буфер, получая состав, который помещают в контейнеры (например, флаконы или шприцы) для маркировки, длительного хранения и отправки. Хотя пример процесса 10 очистки приведен для иллюстрации, следует понимать, что методы, описанные в настоящем документе, могут быть легко применены к другим процессам очистки, включающим колоночную хроматографию.

[0004] В общем, «хроматография» (например, как выполняется на этапе 14) относится к процессу разделения, в котором молекулы распределяются между двумя фазами: (1) неподвижной фазой, которая часто представляет собой смолу; и (2) подвижной фазой, которая в случае разделения белков представляет собой растворитель, такой как вода или хлороформ. Молекулы, которые сильнее притягиваются к неподвижной фазе, медленнее проходят через систему по сравнению с теми, которые сильнее притягиваются к подвижной фазе. Для очистки в коммерческом производстве хроматография обычно проводится в виде колоночной хроматографии из-за соображений масштаба. В ходе обычной хроматографической операции в колонку вводится объем образца. Затем через колонку прокачивается элюент, в результате чего молекулы разделяются на основе их относительного сродства со стационарной смолой и элюентом. Разные молекулы будут выходить из колонки в разное время и после прохождения через колонку разных объемов элюента. Соответственно, терапевтические белки могут быть отделены от других веществ, которые выходят из колонки в более раннее или более позднее время, чем терапевтические белки. Эта информация фиксируется на хроматограмме, которая представляет собой график, например, график поглощения ультрафиолетового излучения, концентрации, выходящей из колонки, в зависимости от времени.

[0005] Приведем несколько примеров: хроматография гидрофобного взаимодействия может использоваться для разделения белков на основе различий в гидрофобности, аффинная хроматография может использоваться для разделения молекул на основе различий в сродстве с целевым лигандом, присоединенным к хроматографической смоле, и ионообменная хроматография может использоваться для разделения молекул на основе различий в молекулярном заряде. В качестве более конкретного примера, катионообменная хроматография (CEX) представляет собой ионообменную хроматографию, используемую, когда интересующая молекула положительно заряжена. Другие распространенные типы хроматографии включают размерно-эксклюзионную хроматографию (SEC), в которой молекулы в растворе разделяются по размеру и/или молекулярной массе, и хроматографию с применением белка А.

[0006] В целом, для обеспечения надежного коммерческого производственного процесса важно охарактеризовать биологические процессы, в частности, определить, как изменчивость свойств исходных материалов влияет на эффективность процесса и качество продукции. Однако по различным причинам производители/поставщики, как правило, не

оценивают, как исходные материалы, произведенные в пределах сертификата анализа (или «CoA»), из-за изменчивости между производственными партиями могут влиять на согласованность процессов и качество продукции. Вместо этого влияние изменчивости исходного материала оценивается, если вообще оценивается, на целевых диапазонах с помощью различных анализов, основанных на оценке риска. Основанные на риске подходы такого рода могут быть недостаточными для определенных видов исходных материалов, таких как смола, используемая в качестве неподвижной фазы в процессах очистки путем колоночной хроматографии (например, процессах очистки, обсужденных выше со ссылкой на этап 14 по фиг. 2). В частности, изменчивость этих исходных материалов от партии к партии может привести к дорогостоящим сбоям, таким как неприемлемые показатели эффективности очистки (например, чрезмерные уровни НСР), которые требуют отбраковки или дополнительной обработки серий лекарственных средств.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[0007] Варианты осуществления, описанные в настоящем документе, относятся к системам и способам, которые облегчают выбор смолы для неподвижной фазы процесса очистки путем колоночной хроматографии при производстве терапевтического белка, такого как, например, моноклональное антитело («mAb») или биспецифическое или другое полиспецифическое антитело. В этих вариантах осуществления многомерная статистическая модель позволяет выбрать смолы (например, выбрать конкретные партии смол), которые не снижают (или чрезмерно не снижают) эффективность процесса очистки, путем учета изменчивости (например, изменчивости от партии к партии) при производстве смол. Многомерная статистическая модель прогнозирует показатель эффективности, такой как уровень НСР и/или одной или нескольких других примесей, для процесса очистки путем колоночной хроматографии (например, СЕХ, SEC, с применением белка А или другого подходящего процесса хроматографии, использующего смолу в качестве неподвижной фазы) на основе различных атрибутов смолы и, возможно, одного или нескольких других типов входных данных (например, коэффициентов загружаемого материала собранного фильтрата и/или параметров процесса хроматографии). Атрибуты смолы могут быть предоставлены производителем, например, в сертификате анализа (CoA).

[0008] «Выбор» смолы может означать выбор одной или нескольких смол из множества кандидатных смол или подтверждение того, является ли одна кандидатная смола приемлемой для использования (т. е. «отбор» кандидатной смолы перед использованием в промышленном масштабе). Например, партии смолы, полученные от поставщика, могут быть отобраны с использованием многомерной статистической модели и данных CoA, предоставленных производителем, чтобы определить, какие партии приемлемы, а какие должны быть отклонены/заменены (или потребуют дополнительных этапов очистки для соответствия требованиям и т. д.). В качестве другого примера, конкретные партии смолы могут быть выбраны (например, заказаны) в первом случае на

основе того, какие партии обеспечивают наибольший допуск относительно порогов приемлемости (например, путем выбора партий смолы, для которых многомерная статистическая модель прогнозирует самые низкие уровни НСР). В качестве других примеров, смолы разных производителей и/или разные типы или составы смол могут быть выбраны путем применения разных, соответствующих наборов значений атрибута смолы к многомерной статистической модели.

[0009] С помощью таких методов количество лекарственного вещества, которое необходимо отклонить/забраковать, и/или количество времени и других ресурсов, необходимых для обеспечения приемлемой эффективности очистки, может быть существенно снижено. Например, время, необходимое для обеспечения приемлемой эффективности очистки, может быть сокращено с десятков или даже сотен часов до приблизительно порядка одного или двух часов.

[0010] Кроме того, некоторые варианты осуществления, описанные в настоящем документе, определяют, какие атрибуты смолы оказывают наибольшее влияние на эффективность (например, по результатам измерения снижения уровня НСР) процесса очистки путем колоночной хроматографии. Эти атрибуты смолы могут быть определены с помощью мелкомасштабной модели процесса колоночной хроматографии промышленного масштаба. Используемый в настоящем документе в качестве описания конкретного процесса термин «промышленный масштаб» или «коммерческий масштаб» указывает на то, что процесс используется в ходе производства или тестирования или в ходе выявления, получения и/или отбора конкретных поставок/материалов, которые будут использоваться при производстве или тестировании партии или серии лекарственного препарата, предназначенного для продажи и/или распространения среди клиентов (например, пациентов, аптек и т. д.), возможно, подлежащего одному или нескольким последующим этапам отбора (например, визуальный осмотр флакона или шприца, наполненного произведенным продуктом в виде лекарственного средства и т. д.). Коммерческий масштаб может означать использование биореакторов по меньшей мере 500 л, 1000 л, 2000 л или более. И наоборот, «мелкомасштабный» или «маломасштабный» указывает на то, что процесс не имеет промышленного масштаба (т. е. выполняется «автономно»). Например, лабораторная хроматографическая станция для отбора партий смолы является «мелкомасштабной», а не «промышленной», если станция не используется для отбора партий смол специально для использования в коммерческом производстве лекарственных средств, независимо от физического размера лабораторной станции по сравнению с хроматографической станцией промышленного масштаба. После определения наиболее важных атрибутов смолы определяются конкретные значения или диапазоны значений для этих атрибутов, которые существенно улучшают эффективность очистки. Затем результаты могут быть предоставлены производителю смолы, который может использовать их для внесения надлежащих изменений в процесс производства смолы. Более того, благодаря способности мелкомасштабной модели точно воспроизводить эффективность промышленного масштаба, данные, полученные в ходе

прогонов мелкомасштабной модели (например, значения атрибута смолы, параметры процесса очистки, результирующие уровни НСР и т. д.), могут быть использованы для увеличения размера набора обучающих данных для многомерной статистической модели, тем самым повышая точность модели. Кроме того, многомерная статистическая модель может дать метрические показатели, которые проливают дополнительный свет на то, какие атрибуты смолы (и/или другие факторы) являются более прогнозируемыми для эффективности очистки.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0011] Специалисту в данной области техники будет понятно, что фигуры, описанные в настоящем документе, включены для целей иллюстрации и не ограничивают настоящее изобретение. Графические материалы не обязательно изображены в масштабе, и вместо этого акцент делается на иллюстрацию принципов настоящего изобретения. Следует понимать, что в некоторых случаях различные аспекты описанных вариантов реализации могут быть укрупнены или увеличены для способствования пониманию описанных вариантов реализации. На графических материалах подобные ссылочные позиции на различных графических материалах в целом выполняют одинаковые функции и/или являются структурно одинаковыми компонентами.

[0012] На фиг. 1 представлена упрощенная структурная схема иллюстративной системы, которая может реализовывать методы, описанные в настоящем документе.

[0013] На фиг. 2 изображен процесс очистки для производства терапевтических белков, известный из уровня техники.

[0014] На фиг. 3 изображена измеренная изменчивость эффективности очистки в промышленном масштабе от партии к партии в разных партиях смолы.

[0015] На фиг. 4 изображен пример зависимости эффективности очистки от высоты колонки в мелкомасштабной модели системы колоночной хроматографии промышленного масштаба.

[0016] На фиг. 5 изображена диаграмма сравнения эффективности мелкомасштабной модели и эффективности системы колоночной хроматографии промышленного масштаба.

[0017] На фиг. 6 изображена эффективность очистки, достигнутая у разных партий смолы, согласно мелкомасштабной модели системы колоночной хроматографии промышленного масштаба.

[0018] На фиг. 7 изображена общая изменчивость уровней НСР после очистки, обусловленная изменчивостью смолы и эксплуатационной изменчивостью.

[0019] На фиг. 8 изображен график фактических и прогнозированных данных, который сравнивает эффективность модели с экспериментальным определением.

[0020] На фиг. 9А и 9В изображено влияние определенных атрибутов смолы и производственных рабочих параметров смолы на эффективность очистки согласно мелкомасштабной модели системы колоночной хроматографии промышленного масштаба.

[0021] На фиг. 10 изображено влияние определенных атрибутов смолы и производственных рабочих параметров смолы на эффективность очистки согласно мелкомасштабной модели.

[0022] На фиг. 11 изображен иллюстративный процесс производства смолы, в котором обратная связь обеспечивается на основе результатов мелкомасштабного моделирования.

[0023] На фиг. 12 изображены примеры модификаций процесса производства смолы, которые могут улучшить эффективность очистки.

[0024] На фиг. 13 изображена диаграмма, на которой сравнивается фактическая эффективность очистки с доверительными пределами, прогнозированными многомерной статистической моделью.

[0025] На фиг. 14 представлена блок-схема, в которой изображен пример способа выбора исходных материалов для использования в процессе очистки путем колоночной хроматографии.

ПОДРОБНОЕ ОПИСАНИЕ

[0026] Различные концепции, представленные выше и более подробно обсужденные ниже, могут быть реализованы многочисленными способами, и описанные концепции не ограничены каким-либо конкретным способом реализации. Примеры вариантов реализации представлены для иллюстративных целей.

[0027] На фиг. 1 представлена упрощенная структурная схема иллюстративной системы 100, которая может реализовывать методы, описанные в настоящем документе. Система 100 содержит вычислительную систему 102, соединенную с возможностью связи с обучающим сервером 104 и сервером 106 поставщика по сети 108. В целом, вычислительная система 102 и/или обучающий сервер 104 выполнены с возможностью обучения многомерной статистической модели 110 (также упоминаемой в настоящем документе как просто «модель 110») с использованием обучающих данных в базе 112 обучающих данных и использования обученной модели 110 для прогнозирования эффективности очистки (например, уровня нежелательного белка клетки-хозяина или «НСР») для процесса колоночной хроматографии, который может быть использован в производстве терапевтических белков. Должно быть понятно, что «процесс колоночной хроматографии» может представлять собой процесс, который будет реализован в будущем, процесс, который был реализован в прошлом, процесс, который реализуется в настоящее время, или процесс, который строго гипотетически никогда не был и никогда не будет реализован. То есть соответствующий реальный процесс может существовать или может не существовать. Процесс очистки путем колоночной хроматографии может включать по меньшей мере один из процесса СЕХ, процесса SEC, процесса хроматографии с применением белка А, любого обращенно-фазового процесса хроматографии или любого другого подходящего процесса хроматографии.

[0028] Как более подробно обсуждено где-то в настоящем документе, модель 110 прогнозирует эффективность очистки на основе по меньшей мере частично значений

атрибута для исходных материалов (в частности, смолы), которые будут использоваться в качестве неподвижной фазы в (реальном или гипотетическом) процессе очистки путем колоночной хроматографии. В некоторых вариантах осуществления модель 110 представляет собой модель проекции на латентные структуры (PLS). Как более подробно обсуждено где-то в настоящем документе, модель PLS может обеспечить высокий уровень точности при работе со значениями атрибута смолы и, возможно, другими входными параметрами для прогнозирования показателя эффективности (например, концентрации НСР) в коммерческом масштабе. Хотя для прогнозирования эффективности колоночной хроматографии были предложены многомерные статистические модели, настоящие варианты осуществления могут обеспечить существенно более надежное прогнозирование за счет учета изменчивости (например, изменчивости от партии к партии) в значениях атрибута смолы. В других вариантах осуществления модель 110 представляет собой другой подходящий тип многомерной статистической (например, регрессионной) модели. В некоторых вариантах осуществления, например, модель 110 может представлять собой или содержать регрессионную модель дерева (или «решений», или «ID»), модель эластичной сети, модель лассо, модель гребня, модель машины опорных векторов (SVM) и т. д. Более того, в некоторых вариантах осуществления модель 110 может содержать разные модели, обученные для прогнозирования разных показателей эффективности. В некоторых вариантах осуществления, например, модель 110 конкретно содержит модель PLS для прогнозирования уровней НСР, модель дерева решений для прогнозирования уровня агрегированных белков и/или фрагментов белков и так далее. Кроме того, в некоторых вариантах осуществления модель 110 может содержать более одной модели любого заданного типа (например, две или более моделей одного типа, которые обучены на разных наборах исторических данных, используют разные наборы признаков и/или имеют разные гиперпараметры).

[0029] Значения атрибута, с которыми работает модель 110, для любого данного прогона/прогнозирования может соответствовать определенной одной из N партий 114 смолы, например, где N - это любое целое число больше нуля. Значения атрибута смолы могут включать, например, параметры из сертификата анализа («CoA»), такие как любой один или несколько из следующего неисключительного списка: диаметр пор; объем пор; %20-30 мкм, несвязанные; коэффициент емкости; % по количеству 2-10 мкм в среднем 3-связанные; % по объему 20-30 мкм в среднем 3-связанные; средний размер частиц; время удерживания рибонуклеазы; время удерживания инсулина; время удерживания лизоцима; время удерживания миоглобина; время удерживания овальбумина; время удерживания окситоцина; время удерживания брадикинина; время удерживания ангиотензина II (angioII); время удерживания нейротензина (neuro); и/или время удерживания ангиотензина I (angioI).

[0030] Аналитические измерения конкретной одной из партий 114 смолы (например, измерения любого одного или нескольких типов значений атрибута смолы, отмеченных выше) могут проводиться поставщиком (например, производителем). В

качестве альтернативы, аналитические измерения могут быть выполнены производителем лекарственного средства (например, производителем лекарственного средства, связанным с вычислительной системой 102) и/или другим субъектом (например, подрядчиком производителя смолы или производителя лекарственного средства).

[0031] В последующем описании значения атрибута для разных смол могут быть значениями атрибута, которые соответствуют разным партиям смолы (например, разным партиям 114 смолы). Однако следует понимать, что значения атрибута для разных смол могут вместо этого соответствовать разным поднаборам одной партии смолы, разным типам смол (например, смолам, произведенным по разным рецептам или составам), смолам, предоставленным разными производителями, и так далее.

[0032] Некоторые или все значения атрибута смолы могут быть указаны в CoA, который производитель (или другой поставщик и т. д.) предоставляет субъекту, владеющему вычислительной системой 102 и/или обслуживающему ее (например, производителю лекарственного средства). На фиг. 1 показан вариант осуществления, в котором значения CoA хранятся в памяти сервера 106 поставщика в виде данных 116 CoA смолы, которые сервер 106 может затем электронно отправить в вычислительную систему 102 и/или обучающий сервер 104 по сети 108 (например, через HTTP, FTP, электронную почту и т. д.). Однако в других вариантах осуществления система 100 исключает сервер 106, а поставщик или другой субъект предоставляет данные 116 CoA смолы (или другую форму информации, определяющую значения атрибута смолы) вычислительной системе 102 и/или обучающему серверу 104 другими способами, например, на бумаге, прилагаемой к физической поставке, которая может включать QR-код или подобное для доступа к данным CoA смолы с сервера, или на любом виде машиночитаемого носителя и т. д.

[0033] В дополнении к значениям атрибута смолы многомерная статистическая модель 110 может работать с одним или несколькими другими типами входных данных. Например, входные данные для модели 110 могут также содержать один или несколько рабочих параметров процесса очистки (также именуемых в настоящем документе просто как «параметры процесса очистки»), один или несколько параметров эффективности процесса собранного фильтрата (также именуемых в настоящем документе просто как «параметры собранного фильтрата») и/или один или несколько других типов числовых и/или категориальных параметров (например, параметр, указывающий на модальность желаемого терапевтического белка, например, моноклональный или биспецифический и т. д.). Параметры процесса очистки могут содержать, например, HETP колонки, асимметрию колонки и/или другие подходящие параметры. Параметры собранного фильтрата могут содержать, например, конечную жизнеспособность производственных биореакторов, общую площадь индивидуальной RP-HPLC для DFM, основную площадь индивидуальной RP-HPLC для DFM, площадь примесей индивидуальной RP-HPLC для DFM, индивидуальную PI для DFM, титр индивидуальной PI для DFM и/или другие подходящие параметры, где DFM=диафильтрованная среда, HETP=эквивалент высоты теоретической

тарелки, PI=изоформа продукта, и RP-HPLC=обращенно-фазовая высокоэффективная жидкостная хроматография.

[0034] Вычислительная система 102 также может быть в целом выполнена так, чтобы позволить одному или нескольким пользователям, которые могут быть локальными или удаленно распределенными, использовать возможности прогнозирования вычислительной системы 102 и предоставлять различные интерактивные возможности пользователю (пользователям), как обсуждено где-то в настоящем документе.

[0035] Сеть 108 может представлять собой одну сеть связи или может содержать несколько сетей связи одного или нескольких типов (например, одну или несколько проводных и/или беспроводных локальных сетей (LAN) и/или одну или несколько проводных и/или беспроводных глобальных сетей (WAN), таких как Интернет). В различных вариантах осуществления обучающий сервер 104 может обучить и/или использовать многомерную статистическую модель 110 как «облачную» службу (например, веб-службы Amazon), или обучающий сервер 104 может быть локальным сервером. Однако в изображенном варианте осуществления модель 110 обучается на сервере 104, а затем передается в вычислительную систему 102 по сети 108 по мере необходимости. В других вариантах осуществления модель 110 обучается на вычислительной системе 102, а затем загружается на обучающий сервер 104 для последующего доступа. В еще одних вариантах осуществления вычислительная система 102 обучает и поддерживает/сохраняет многомерную статистическую модель 110, в этом случае система 100 может не содержать обучающий сервер 104 (и, возможно, сеть 108), или сервер 104 может быть частью вычислительной системы 102.

[0036] Вычислительная система 102 может содержать один или несколько компьютеров общего назначения, специально запрограммированных на выполнение операций, обсужденных в настоящем документе, и/или может содержать один или несколько вычислительных устройств специального значения. Как видно на фиг. 1, вычислительная система 102 содержит процессорный узел 120, сетевой интерфейс 122, дисплей 124, устройство 126 пользовательского ввода и узел 128 памяти. В вариантах осуществления, где вычислительная система 102 содержит два или более компьютеров (расположенных совместно или удаленных друг от друга), операции, описанные в настоящем документе, относящиеся к по меньшей мере процессорному узлу 120, сетевому интерфейсу 122 и/или узлу 128 памяти, могут быть распределены между несколькими процессорными узлами, несколькими сетевыми интерфейсами и/или несколькими узлами памяти соответственно. Более того, дисплей 124 и устройство 126 пользовательского ввода, хотя и упоминаются в настоящем документе в единственном числе, могут включать в себя несколько дисплеев и несколько устройств пользовательского ввода соответственно. Например, дисплей 124 может включать в себя по меньшей мере один дисплей на каждом из ряда удаленных, специфических для пользователя клиентских устройств, а устройство 126 пользовательского ввода может включать в себя по меньшей мере одно устройство пользовательского ввода для каждого из этих клиентских устройств.

[0037] Процессорный узел 120 содержит один или несколько процессоров, каждый из которых может представлять собой программируемый микропроцессор, который исполняет программные инструкции, хранящиеся в узле 128 памяти, с целью исполнения некоторых или всех функций вычислительной системы 102, как описано в настоящем документе. Процессорный узел 120 может содержать один или несколько центральных процессорных узлов (CPU) и/или один или несколько графических процессорных узлов (GPU), например. Альтернативно или в дополнение некоторые из процессоров в процессорном узле 120 могут представлять собой процессоры других типов (например, специализированные интегральные микросхемы (ASIC), вентильные матрицы с эксплуатационным программированием (FPGA) и т. д.), и некоторые функциональные возможности вычислительной системы 102, как описано в настоящем документе, альтернативно могут быть реализованы в аппаратном обеспечении.

[0038] Сетевой интерфейс 122 может содержать любое подходящее аппаратное обеспечение (например, клиентское приемопередающее аппаратное обеспечение), программно-аппаратное обеспечение и/или программное обеспечение, выполненное с возможностью установления связи с обучающим сервером 104 по сети 108 с использованием одного или нескольких проводных и/или беспроводных протоколов связи. Например, сетевой интерфейс 122 может представлять собой или содержать интерфейс WiFi или Ethernet, позволяющий вычислительной системе 102 устанавливать связь с обучающим сервером 104 через Интернет или интранет и т. д.

[0039] В дисплее 124 может использоваться любая подходящая технология (например, LED, OLED, LCD и т. д.) для представления информации пользователю, а устройство 126 пользовательского ввода может представлять собой клавиатуру или другое подходящее устройство ввода. В некоторых вариантах осуществления дисплей 124 и устройство 126 пользовательского ввода объединены в одном устройстве (например, сенсорном дисплее). В целом, дисплей 124 и устройство 126 пользовательского ввода могут быть объединены, чтобы позволить пользователю взаимодействовать с графическими пользовательскими интерфейсами (GUI), предоставляемыми вычислительной системой 102. Однако вычислительная система 102 может не содержать дисплей 124 и/или устройство 126 пользовательского ввода, например, в определенных вариантах осуществления, где вычислительная система 102 взаимодействует с другими вычислительными устройствами или системами (например, клиентскими устройствами третьих лиц) для обеспечения взаимодействия пользователей этих устройств или систем.

[0040] Узел 128 памяти может содержать одно или несколько энергозависимых и/или энергонезависимых запоминающих устройств. Могут содержаться запоминающие устройства любого подходящего типа или типов, такие как постоянное запоминающее устройство (ROM), оперативное запоминающее устройство (RAM), флеш-память, твердотельный накопитель (SSD), накопитель на жестком диске (HDD) и так далее. В совокупности узел 128 памяти может хранить одно или несколько программных приложений, данные, полученные/используемые этими приложениями, и данные,

выводимые/генерируемые этими приложениями. Эти приложения включают приложение 130 для выбора смолы, которое при исполнении процессорным узлом 120 прогнозирует и представляет эффективность виртуального (*in silico*) процесса колоночной хроматографии для очистки при производстве терапевтического белка. В некоторых вариантах осуществления различные «узлы» приложения 130 для выбора смолы, обсужденные в настоящем документе, могут быть распределены по разным программным приложениям, и/или функциональные возможности любого такого узла могут быть разделены между двумя или более программными приложениями.

[0041] В примере системы 100 приложение 130 для выбора смолы содержит узел 132 сбора данных, узел 134 прогнозирования и узел 136 визуализации. В общем, узел 132 сбора данных получает (например, извлекает) параметры, которые узел 134 прогнозирования применяет в качестве входных данных к локальной многомерной статистической модели 138, чтобы прогнозировать показатель эффективности. В изображенном варианте осуществления модель 138 является локальной копией модели 110, обученной обучающим сервером 104, и может храниться в RAM или ROM узла 128 памяти, например. Однако, как отмечено выше, обучающий сервер 104 может использовать/прогонять многомерную статистическую модель 110 в некоторых вариантах осуществления, в случае чего нет необходимости иметь локальную копию в узле 128 памяти, или многомерная статистическая модель 110 может изначально находиться в постоянной памяти узла 128 памяти, а не извлекаться из обучающего сервера 104 по мере необходимости. Узел 132 сбора данных может получать значения атрибута смолы (например, данные 116 CoA смолы) от сервера 106 поставщика по сети 108 и может получать другие параметры, с которыми работает локальная многомерная статистическая модель 138, от пользователя, вводящего параметры/значения через GUI (например, представленный на дисплее 124), который генерируется или заполняется узлом 136 визуализации, и/или в виде одного или нескольких файлов или другой передачи данных (например, используя пути к файлам, указанные пользователем через такой GUI), например.

[0042] Узел 136 визуализации может также генерировать и/или заполнять GUI для просмотра и/или взаимодействия с прогнозированными результатами смоделированного процесса (например, значениями показателя эффективности, выведенными узлом 134 прогнозирования с использованием такой модели, как локальная многомерная статистическая модель 138), например. Например, узел 136 визуализации может предписывать GUI отобразить прогнозируемую концентрацию НСР (или концентрацию другого типа примесей, или общую концентрацию примесей и т. д.) для данного набора значений атрибута смолы, которые соответствуют конкретной одной из партий 114 смолы, а также любые другие параметры, используемые в качестве входных данных для локальной многомерной статистической модели 138 (например, значения различных параметров процесса и/или параметров собранного фильтрата).

[0043] Работа системы 100 согласно одному варианту осуществления теперь будет

описана более подробно. Первоначально обучающий сервер 104 обучает многомерную статистическую модель 110, используя исторические данные, хранящиеся в базе 112 обучающих данных. База 112 обучающих данных может содержать одну базу данных, хранящуюся в одном запоминающем устройстве (например, HDD, SSD и т. д.), или может содержать несколько баз данных, хранящихся в одном или нескольких запоминающих устройствах. В некоторых вариантах осуществления и как более подробно обсуждено в настоящем документе, различные методы (например, мелкомасштабное моделирование) могут быть использованы для определения того, какие признаки (например, значения атрибута смолы, параметры процесса очистки и т. д.) являются наиболее прогнозируемыми для конкретного показателя эффективности, и/или им можно обучить или переобучить с использованием набора признаков, который содержит только признаки, наиболее прогнозируемые для конкретного показателя эффективности. Хотя многомерная статистическая модель 110 может содержать несколько отдельных моделей, для простоты объяснения в описании в настоящем документе многомерная статистическая модель 110 упоминается в единственном числе, и подразумевается, что методы, описанные в настоящем документе, могут быть применены к нескольким моделям.

[0044] База 112 обучающих данных сохраняет набор обучающих данных для обучения многомерной статистической модели 110 (например, входные данные/данные признаков и соответствующие маркировки). Для обучения модели, прогнозирующей концентрацию НСР, например, база 112 обучающих данных может содержать многочисленные наборы входных данных/признаков, каждый из которых содержит исторические значения атрибута смолы (и, возможно, параметры процесса очистки и/или параметры собранного фильтрата и т. д.), а также известную (например, измеренную) концентрацию НСР, соответствующую каждому набору признаков. В некоторых вариантах осуществления все признаки и маркировки являются числовыми, при этом нечисловые классификации или категории преобразовываются в числовые значения (например, допустимые значения [моноклональный, биспецифический формат 1, биспецифический формат 2, биспецифический формат 1 или 2] признака/входных данных модальности преобразуются в значения [00, 10, 01, 11]).

[0045] В некоторых вариантах осуществления обучающий сервер 104 использует дополнительные маркированные наборы данных в базе 112 обучающих данных для подтверждения/проверки обученной многомерной статистической модели 110 (например, для подтверждения того, что многомерная статистическая модель 110 обеспечивает по меньшей мере некоторую минимальную приемлемую точность). В некоторых вариантах осуществления обучающий сервер 104 также обновляет/уточняет многомерную статистическую модель 110 на постоянной основе. Например, после того как многомерная статистическая модель 110 первоначально обучена для обеспечения достаточного уровня точности и введена в эксплуатацию для процесса промышленного масштаба (например, для отбора партий смолы), дополнительные измерения показателя эффективности (и соответствующих входных данных/признаков) в коммерческом масштабе могут быть

использованы для дальнейшего повышения точности прогнозирования многомерной статистической модели 110.

[0046] Приложение 130 для выбора смолы может в какой-то более поздний момент получить от обучающего сервера 104 по сети 108 и через сетевой интерфейс 122 копию многомерной статистической модели 110. После получения модели вычислительная система 102 сохраняет локальную копию в виде локальной многомерной статистической модели 138. В других вариантах осуществления, как отмечено выше, модель не извлекается, а вместо этого входные данные/данные признаков отправляют на обучающий сервер 104 (или другой сервер), что необходимо для использования многомерной статистической модели 110, или многомерная статистическая модель 110 может находиться только в вычислительной системе 102.

[0047] В соответствии с набором признаков, для которого разработана/обучена локальная многомерная статистическая модель 138, узел 132 сбора данных собирает необходимые данные. Например, узел 132 сбора данных может получать данные 116 CoA смолы от сервера 106 поставщика или посредством ввода информации пользователем (например, на GUI, представленном на дисплее 124). Узел 132 сбора данных также собирает любые другие параметры, используемые в качестве входных данных модели, такие как введенные пользователем параметры процесса очистки и/или параметры собранного фильтрата, например. После того как узел 132 сбора данных собрал входные данные модели для конкретной кандидатной смолы (например, одной из партий 114 смолы), узел 134 прогнозирования предписывает локальной многомерной статистической модели 138 работать с этими входными данными/признаками для прогнозирования желаемого показателя эффективности для процесса колоночной хроматографии. В некоторых вариантах осуществления многомерная статистическая модель 138 прогнозирует уровни НСР (например, концентрацию НСР), и узел 132 сбора данных собирает значения атрибута смолы (например, значения любого одного или нескольких иллюстративных параметров CoA, перечисленных выше), значения параметра процесса очистки (например, значения любого одного или нескольких иллюстративных параметров процесса очистки, перечисленных выше) и параметры собранного фильтрата (например, значения любого одного или нескольких иллюстративных параметров собранного фильтрата, перечисленных выше) для использования в качестве входных данных для локальной многомерной статистической модели 138.

[0048] Узел 136 визуализации может затем предписывать GUI, изображенному на дисплее 124, представить прогнозированный показатель эффективности и/или другую информацию на основе прогнозированного показателя эффективности (например, список/классификацию прогнозированных показателей эффективности для разных партий смолы, двоичное указание того, является ли прогнозированный показатель эффективности «приемлемым» по сравнению с заранее определенным порогом, и т. д.). Узел 136 визуализации может также предписывать GUI представить доверительные метрические показатели, связанные с прогнозированными показателями эффективности (например,

доверительные метрические показатели, созданные локальной многомерной статистической моделью 138). Например, GUI может отображать диапазон уровней НСР, которые соответствуют по меньшей мере уровню достоверности 90% (или 80%, 95% и т. д.). При повторении для нескольких партий 114 смолы пользователь может выбрать, какие партии смолы (или типы смолы и т. д.) являются приемлемыми или неприемлемыми (например, путем сравнения разных прогнозов друг с другом или путем сравнения каждого прогноза с порогом приемлемости и т. д.). В общем, для любой данной одной из партий 114 смолы пользователь может использовать отображаемый прогноз и/или результат (возможно, в сочетании с другой информацией), чтобы определить, является ли партия приемлемой. Если партия является приемлемой, пользователь может выбрать эту партию (например, указать одобрение партии с помощью GUI или других средств) для использования в реальном процессе колоночной хроматографии. С этой целью иллюстративная система 100 содержит систему 140 колоночной хроматографии для решения реальных задач, которая выполнена с возможностью выполнения процесса колоночной хроматографии (например, CEX, SEC, с применением белка А или другого типа колоночной хроматографии), а выбранная/принятая смола может быть использована в качестве неподвижной фазы для этого процесса. Система 140 колоночной хроматографии может представлять собой систему колоночной хроматографии промышленного масштаба, используемую, например, при коммерческом производстве терапевтического белка. В зависимости от вариантов осуществления, система 140 колоночной хроматографии может содержать одну или несколько колонок, а выбранная смола может использоваться для одной, некоторых или всех из этих колонок.

[0049] Процесс прогнозирования/визуализации может быть выполнен только один раз (например, при отборе одной полученной партии смолы для определения пригодности партии) или несколько раз (например, при выборе того, какая из нескольких полученных партий 114 смолы должна быть сохранена или заказана в первую очередь и т. д.). Независимо от того, является ли целью отбор одной партии смолы или выбор из нескольких партий кандидатной смолы, этот процесс прогнозирования и визуализации эффективности очистки может существенно сократить количество лекарственного вещества, которое должно быть отклонено/забраковано из-за низкой эффективности очистки, и/или существенно сократить количество времени, необходимого для обеспечения приемлемой эффективности очистки. Например, время, необходимое для отбора (т. е. обеспечения приемлемой эффективности для) одной или нескольких партий смолы, может быть сокращено с десятков или даже сотен часов до приблизительно порядка одного или двух часов.

[0050] В других вариантах осуществления или сценариях процесс можно повторять столько раз, сколько необходимо для чисто гипотетических значений атрибута смолы, например, чтобы идентифицировать критические атрибуты смолы, идентифицировать оптимальные значения атрибута смолы или диапазоны значений или идентифицировать способы, которыми разные атрибуты смолы взаимодействуют с другими параметрами

(например, для оценки корреляций собранного фильтрата и/или параметров процесса хроматографии с конкретными свойствами смолы), и так далее. Результаты этих виртуальных экспериментов могут быть переданы производителям по мере необходимости (например, чтобы производитель мог соответствующим образом изменить состав/рецепт своей смолы).

[0051] Также можно использовать различные методы, отдельно или в сочетании с (например, в качестве предшественника) методами статистического моделирования, описанными в настоящем документе, чтобы лучше понять, как изменчивость свойств смолы может повлиять на процесс очистки путем колоночной хроматографии. Например, в то время как обычные анализы, основанные на оценке риска, могут гарантировать, что большинство партий смолы обеспечивают приемлемую эффективность очистки, тем не менее, могут возникать неожиданные остатки. На фиг. 3 изображен пример двух таких случаев. На диаграмме 300 по фиг. 3 показана измеренная изменчивость эффективности очистки в промышленном масштабе от партии к партии (в частности, уровень/концентрация НСР) по ряду разных партий смолы за конкретный период времени. Как может быть видно из диаграммы 300, большинство партий смолы привели к кажущейся безопасной степени очистки относительно допустимого предела/порога для полученного уровня НСР. Однако изменения атрибутов произведенной смолы со временем привели к тому, что предел приемлемости был приближен и, в конечном итоге, превышен как в первом событии 302, так и в более позднем втором событии 304.

[0052] Создание и внедрение мелкомасштабной модели колоночной хроматографии, разработанной для воспроизведения в важных аспектах (но в меньших размерах/количествах) системы колоночной хроматографии промышленного масштаба, что дало в результате уровни НСР, показанные на фиг. 3, привело к успешному воссозданию отклонения, показанного как первое событие 302 по фиг. 3. Изменяя высоту колонки со смолой (также называемую «высотой слоя» или «высотой набивочного слоя») с использованием этой мелкомасштабной модели, было определено, что изменчивость высоты колонки влияет на эффективность хроматографии с обычно квадратичным откликом, как видно на диаграмме 400 по фиг. 4. На фиг. 4 параметры процесса для мелкомасштабной модели были установлены на основе предполагаемой высоты колонки, а затем скорректированы на основе фактической/наблюдаемой высоты колонки. Первопричина первого события 302 была связана с высотой колонки и ее влиянием на рабочие параметры скорости потока, объема промывки и наклона градиента при элюировании. Затем на основе полученных данных были предприняты корректирующие действия, включая (1) сужение нормального диапазона эксплуатации (NOR) процесса для уменьшения изменчивости высоты слоя и (2) установление автоматических заданных значений для расхода, объема промывки и наклона градиента на основе фактически измеренной высоты слоя (а не теоретического значения).

[0053] Мелкомасштабная модель была также использована для характеристики различных других аспектов процесса колоночной хроматографии промышленного

масштаба с целью оптимизации рабочих параметров (например, градиента, температуры, концентраций буфера, градиента в начале, градиента в конце и т. д.) без влияния на другие атрибуты продукта. Кроме того, были определены более узкие диапазоны для рабочих параметров с учетом допусков на управление оборудованием и моделированием рисков.

[0054] На фиг. 5 представлены диаграммы 500 и 520 сравнения эффективности очистки для мелкомасштабной модели и для промышленного масштаба с учетом процентного содержания на этапе восстановления (во второй колонке четырехколоночной системы очистки) и уровня НСР (в третьей колонке четырехколоночной системы очистки) соответственно. Как может быть видно на фиг. 5, эффективность мелкомасштабной модели (SSM) близко приближена к эффективности промышленного масштаба (CS). Соответственно, мелкомасштабная модель может быть использована для отбора смолы перед использованием в промышленном масштабе. Таким образом, можно поддерживать хорошую эффективность для очистки в коммерческом масштабе, несмотря на изменчивость смолы от партии к партии (например, партии смолы с плохими характеристиками могут быть отбракованы, или возвращены, или использованы в сочетании с дополнительными этапами очистки и т. д.). Однако отбор с помощью мелкомасштабной модели системы колоночной хроматографии (например, мелкомасштабной модели системы 140 колоночной хроматографии по фиг. 1) может потребовать десятки или даже сотни часов по сравнению с приблизительно потенциально порядком одного или двух часов при использовании многомерной статистической модели, описанной в настоящем документе (например, многомерной статистической модели 110 или локальной многомерной статистической модели 138 по фиг. 1). Таким образом, может оказаться предпочтительным использовать мелкомасштабное моделирование для получения определенных знаний об атрибутах смолы и их взаимодействии с различными рабочими параметрами хроматографии, а вместо него использовать многомерную статистическую модель для отбора партий смолы.

[0055] На фиг. 6 изображена диаграмма 600, показывающая экспериментальную эффективность очистки (уровни НСР) для разных партий смолы с использованием мелкомасштабной модели системы колоночной хроматографии промышленного масштаба. Как может быть видно, эта мелкомасштабная модель определила две из 18 партий смолы как неудовлетворительные (т. е. превышающие допустимый порог/предел НСР). Таким образом, две партии смолы с низкими характеристиками можно отсеять, а не использовать эти партии в коммерческом масштабе только для того, чтобы получить неприемлемые уровни чистоты.

[0056] Снова обращаясь к фиг. 3, после первого события 302 для системы хроматографии промышленного масштаба было выполнено несколько корректирующих модификаций на основе результатов, полученных от мелкомасштабной модели. Например, вышеупомянутые корректировки, касающиеся высоты слоя, расхода, объема промывки и наклона градиента, были реализованы в коммерческом масштабе. Тем не менее, годы спустя наблюдалось другое событие в коммерческом масштабе (т. е. второе

событие 304 по фиг. 3), при этом результирующий уровень НСР снова превысил допустимый предел, несмотря на то что мелкомасштабная модель определила партию смолы как приемлемую.

[0057] Исследование второго события 304 показало, что в дополнение к атрибутам смолы для конкретной партии технологический собранный фильтрат оказал значительное влияние на результирующие уровни НСР. Это было подтверждено характеризующими данными, показывающими изменчивость уровней НСР между разными партиями собранного фильтрата. На фиг. 7 изображена диаграмма 700, показывающая комбинированную/общую изменчивость уровней НСР в результате как изменчивости смолы от партии к партии, так и эксплуатационной изменчивости, где «эксплуатационная изменчивость» относится к комбинации (1) изменчивости параметров процесса очистки, которые были определены во время первого события 302, и (2) изменчивости от партии к партии собранного фильтрата. Параметры процесса очистки, которые могут изменяться, могут содержать, например, НЕТР колонки 1, асимметрию колонки 1 и/или другие параметры. Параметры собранного фильтрата, которые могут изменяться, могут содержать, например, конечную жизнеспособность производственного биореактора, общую площадь индивидуальной RP-HPLC для DFM, основную площадь индивидуальной RP-HPLC для DFM, площадь примесей индивидуальной RP-HPLC для DFM, индивидуальную PI для DFM, титр индивидуальной PI для DFM и/или другие параметры. Как может быть видно на диаграмме 700, при рассмотрении изолированно (т. е. без эксплуатационной изменчивости) возникающая изменчивость смолы от партии к партии может указывать на приемлемый допуск НСР относительно порога приемлемости даже в тех случаях, когда фактически порог может быть превышен.

[0058] Чтобы лучше понять, как изменчивость партии смолы влияет на изменчивость НСР, было проведено совместное исследование с производителем смолы. Совместное исследование было сфокусировано на понимание того, какие производственные параметры/атрибуты смолы существенно влияют на очистку НСР на стадии хроматографии, чтобы контролировать эти атрибуты и обеспечить постоянство эффективности очистки. В совместном исследовании загружаемый хроматографический материал (представитель материала, используемого в коммерческом масштабе) использовался для мелкомасштабного моделирования, а очистка проводилась с использованием хроматографических колонок Omnifit® Benchmark с внутренним диаметром 1,0 см. Прогоны мелкомасштабной модели были выполнены с помощью систем GE® Healthcare ÄKTA Explorer® 100 с программным обеспечением UNICORN® 5.0. Все используемые растворы и буферы были получены с использованием исходных материалов от квалифицированных поставщиков и в соответствии с мелкомасштабными рецептами. Уровни/результаты НСР оценивались с помощью ферментного иммуносорбентного анализа (ELISA). Построение экспериментального дизайна и анализ моделей для экспериментов проводились с использованием программного обеспечения JMP® для статистических открытий от SAS®.

[0059] В совместном исследовании для трех атрибутов смолы был осуществлен способ полного факторного планирования эксперимента (DOE), включая прогоны в центральной точке для проверки несоответствия из-за нелинейных эффектов: уровень лиганда А, уровень лиганда В и уровень блокирующей концевой группы. Выбор этих атрибутов был основан на научном обосновании, связанном с известным механизмом связывания белка и взаимодействием в смешанном режиме. Для исследования три атрибута были изменены в пределах нормального диапазона эксплуатации (NOR), что привело к девяти разным перестановкам, которые были использованы производителем для создания разных образцов смолы. После производства образцы смолы использовались в качестве неподвижной фазы для мелкомасштабной модели. Затем результаты НСР использовались для оценки мелкомасштабной модели на основе скорректированного графика фактических и прогнозированных данных коэффициента детерминации (« R^2 adj»). После этого мелкомасштабная модель использовалась для определения основных эффектов, факторов и взаимодействий, которые способствовали получению результатов НСР.

[0060] На основе результатов DOE был проведен пошаговый регрессионный анализ Fit Model для оценки влияния одного фактора и для моделирования многофакторных взаимодействий на переменную отклика (т. е. уровень НСР). Как показано на диаграмме 800 по фиг. 8, в результате регрессионного анализа был получен график фактических и прогнозированных данных со скорректированным коэффициентом детерминации (R^2 Adj), составляющим 0,84, что подтверждает высокий уровень прогнозируемости, предлагаемый моделью. Статистика R^2 Adj является модифицированной версией коэффициента детерминации (R^2) и сравнивает описательную силу регрессионных моделей, включающих разное количество предикторов. Каждый добавленный в модель предиктор увеличивает R^2 для этой модели. Таким образом, в то время как модель с большим количеством условий может показаться лучше подходящей просто потому, что в ней больше условий, R^2 Adj компенсирует добавление переменных. R^2 Adj увеличивается только в том случае, если новое условие улучшает модель сверх того, что могло бы получиться случайно, и уменьшается только в том случае, если новое условие ухудшает модель сверх того, что могло бы получиться случайно.

[0061] Из трех оцениваемых факторов (лиганд А, лиганд В и уровни блокирующей концевой группы) статистически значимое влияние на уровень/чистоту НСР (здесь определяется как р-значение, составляющее менее 0,05) было получено только для уровня лиганда А (р-значение 0,0020) и для взаимодействия между лигандом А и уровнями блокирующей концевой группы (р-значение 0,0277), причем уровень лиганда А оказался наиболее значимым. Связь между уровнем НСР и каждым из лиганда А, лиганда В и уровнем блокирующей концевой группы показана на диаграммах 900 по фиг. 9А. На диаграммах 900 большой наклон соответствует большему влиянию на уровни/чистоту НСР. На фиг. 9В изображены диаграммы 950, показывающие взаимодействие между уровнем лиганда А, уровнем блокирующей концевой группы и уровнем/чистотой НСР.

Как видно из диаграмм 950, наилучшая эффективность (самый низкий уровень НСР) получается при сочетании относительно высокого уровня лиганда А и относительно высокого уровня блокирующей концевой группы.

[0062] На фиг. 10 изображена диаграмма 1000 с более подробными результатами. На диаграмме 1000 показаны результаты НСР для девяти пробных прогонов, обозначенных «1» - «9», причем каждый прогон соответствует разной перестановке относительно высоких уровней («+»), относительно низких уровней («-») или промежуточных уровней («0») для трех атрибутов смолы, как показано ниже в таблице 1.

ПРОГОН	Уровень лиганда А	Уровень лиганда В	Уровень блокирующей концевой группы
1	-	-	+
2	+	+	-
3	-	+	-
4	-	-	-
5	+	-	-
6	0	0	0
7	-	+	+
8	+	+	+
9	+	-	+

Таблица 1

Как видно из диаграммы 1000, прогоны под номерами 8 и 9 обеспечивают наилучшие результаты НСР из всех девяти прогонов, при этом уровни лиганда В показывают лишь очень незначительное влияние на эффективность.

[0063] Производитель смолы может использовать эту информацию для улучшения процесса производства смолы. Пример одного такого процесса 1100 показан на фиг. 11. В процессе 1100 производство смолы начинается с синтеза 1110 основы смолы, затем следуют этапы определения размеров 1120 и связывания 1130, а затем конечное тестирование 1140. На основании значений атрибута смолы, определенных в ходе совместного исследования, на этапе связывания 1130 могут быть внесены изменения. Тогда как на фиг. 11 показаны входные данные относительно лиганда А, лиганда В и уровней блокирующей концевой группы, может быть достаточно скорректировать процесс 1100 производства смолы только на основе желаемых уровней лиганда А (учитывая относительно небольшое влияние уровней блокирующей концевой группы и лиганда В) или только на основе желаемой комбинации уровней лиганда А и блокирующей концевой группы (учитывая особенно небольшое влияние уровней лиганда В).

[0064] На фиг. 12 изображены примеры модификаций 1200 процесса производства

смолы (например, процесс 1100) на основе результатов мелкомасштабной модели (показаны на фиг. 9 и 10 и в таблице 1), что улучшает эффективность очистки. Как видно на фиг. 12, текущее состояние 1210 представляет процесс производства смолы до входных данных/обратной связи, предусмотренных для производителя смолы на этапе 1130 по фиг. 11. Основа смолы, которая содержит отрицательно заряженные силанольные группы, связана с лигандами, включая лиганд А и лиганд В. Первая стадия 1220 модификации включает увеличение концентрации/количества как лиганда А, так и лиганда В, чтобы увеличить углеродную нагрузку и поверхностное покрытие на основе. По желанию лиганд В можно не увеличивать из-за его относительно небольшого влияния на эффективность. Вторая стадия 1230 модификации включает дальнейшее увеличение концентраций лиганда А и лиганда В, а также увеличение уровня блокирующей концевой группы, чтобы увеличить углеродную нагрузку и поверхностное покрытие на основе, и чтобы уменьшить доступные силанольные группы для влияния на взаимодействие в смешанном режиме.

[0065] Данные, полученные из результатов совместного исследования DOE, из анализа вклада собранного фильтрата, а также из исторических уровней/результатов НСР исторической модели промышленного масштаба и мелкого масштаба, были использованы для создания многомерной статистической прогностической модели с целью обеспечения более эффективного и надежного процесса отбора смол перед использованием смол в процессе очистки путем колоночной хроматографии промышленного масштаба. Полученная в результате модель может быть использована, например, в качестве многомерной статистической модели 110 и/или локальной многомерной статистической модели 138 по фиг. 1. В данном случае многомерная статистическая модель была моделью проекции на латентные структуры (PLS), целью которой было установление корреляции между параметрами процесса и реакциями процесса, а также определение параметров, которые в наибольшей степени влияют на качество реакций процесса. Хотя в данном случае была построена и протестирована модель PLS, понятно, что в других вариантах осуществления вместо нее могут использоваться другие типы моделей (например, эластичная сеть, дерево решений и т. д.).

[0066] Данные, использованные для обучения модели PLS, были получены из собранного фильтрата клеточной культуры, параметров процесса очистки и значений атрибута смолы (т. е. значений атрибута смолы, указанных в CoA для различных партий смолы). Данные отражали ряд «наблюдений», которые были разделены на обучающие поднаборы и поднаборы для проверки/подтверждения. Набор данных для подтверждения содержал серии лекарственных веществ, выбранные случайным образом в диапазоне результатов/уровней НСР промышленного масштаба. Набор обучающих данных содержал оставшиеся серии лекарственного вещества, а также данные, полученные в ходе прогонов мелкомасштабной модели. В любом данном прогоне серии лекарственного вещества или мелкомасштабной модели могла использоваться другая смесь загружаемого материала собранного фильтрата и/или партий смолы. Таким образом, каждые обучающие входные данные (или «переменная x») были расширены до трех единиц входных

данных/переменных x (например, минимальной, максимальной и средневзвешенной) для лучшего улавливания потенциальных вкладов в модель PLS и прогнозирования выходных данных («переменной y », здесь уровень НСР) на оцениваемом этапе хроматографии.

[0067] Затем наборы обучающих данных и наборы данных для подтверждения были обработаны для создания и проверки первой итерации модели PLS. Это было сделано с помощью инструментов SIMCA 14.1 от Umetrics®, хотя более новые версии могут быть использованы при обновлении модели более свежими данными (т. е. для расширения обучающего набора и, следовательно, диапазона прогнозируемости). Затем была определена и проанализирована способность прогнозирования каждой единицы входных данных/переменной x . Для этого были использованы инструменты SIMCA 14.1 для построения графика, показывающего важность переменных для проекции (или «VIP») каждой переменной x . Переменные x с более высокими значениями VIP имеют больший вклад в соответствие и прогнозируемость модели. Из всех минимальных/максимальных/средневзвешенных значений, связанных с каждой переменной x , для следующей итерации модели PLS сохранялось/использовалось только то, которое имело наибольшее значение VIP.

[0068] После создания конечной обученной модели PLS выходные данные/переменные y (прогнозируемые уровни НСР) для набора для подтверждения были прогнозированы и сравнены с фактическими/известными значениями (измеренными уровнями НСР). Пригодность и прогнозируемость конечной модели PLS оценивалась на основе различных типов информации, таких как график остатков (т. е. график остатков, стандартизированных по двойной логарифмической шкале), график перестановок (т. е. график, отражающий вариации в частях набора данных, используемых для обучения и подтверждения, для оценки риска того, что текущая модель PLS хорошо подходит для набора обучающих данных, но плохо прогнозирует выходные данные для новых наблюдений), график VIP (т. е. для обобщения важности переменных, как для объяснения входных данных/переменных x , так и для корреляции с выходными данными/переменной y) и график, отображающий наблюдаемые значения в сравнении с прогнозируемыми значениями выходных данных/переменной y .

[0069] Было создано шесть итераций модели PLS (каждая с двумя главными компонентами) со следующими метрическими показателями эффективности.

Модель	R^2X	R^2Y	Q^2
M1	0,472	0,862	0,831
M2	0,523	0,843	0,815
M3	0,517	0,850	0,820
M4	0,532	0,851	0,822
M5	0,549	0,850	0,824
M6	0,570	0,848	0,822

Таблица 2

R^2 - это мера того, насколько хорошо модель соответствует набору данных (при этом R^2X измеряет соответствие по входным данным, а R^2Y измеряет соответствие по выходным данным/НСР), а Q^2 - это мера того, насколько прогнозируема/точна модель. Целью является максимизация R^2Y , хотя могут также учитываться и другие факторы, например, простота модели (например, количество входных данных).

[0070] Модель, полученная в результате конечной итерации (М6), имела значение R^2Y (0,848), которое немного ниже самого высокого значения R^2Y (0,862), но имела преимущество в том, что была обучена на меньшем количестве переменных x , чем другие модели. В частности, М6 был обучен на входном наборе, состоящем из 17 значений атрибута смолы из СоА (диаметр пор, объем пор, %20-30 мкм, несвязанные, коэффициент емкости, % по количеству 2-10 мкм в среднем 3-связанных, % по объему 20-30 мкм в среднем 3-связанных, средний размер частиц, время удерживания рибонуклеазы, время удерживания инсулина, время удерживания лизоцима, время удерживания миоглобина, время удерживания овальбумина, время удерживания окситоцина, время удерживания брадикинина, время удерживания angioII , время удерживания peugo и время удерживания angioI), шести параметров загружаемого материала собранного фильтрата (конечная жизнеспособность производственного биореактора, общая площадь индивидуальной RP-HPLC для DFM, основная площадь индивидуальной RP-HPLC для DFM, площадь примесей индивидуальной RP-HPLC для DFM, индивидуальная PI для DFM, титр индивидуальной PI для DFM) и двух параметров процесса очистки ниже по потоку (HETP колонки 1, асимметрия колонки 1). График нормальной вероятности остатков показал отсутствие остатков в конечной (М6) модели PLS (все вероятности находятся в пределах плюс-минус четырех стандартных отклонений). Более того, график перестановок показал, что конечная модель PLS является уникальным решением для набора обучающих данных. Более конкретно, график значений R^2Y и Q^2 в зависимости от корреляции между переставленной переменной y и исходной переменной y показал большое и четкое разделение между значениями для исходной модели М6 и значениями для всех перестановок модели М6 (т. е. для исходной модели М6 значения R^2Y и Q^2 составляют 0,848 и 0,822 соответственно, а для перестановок все значения меньше приблизительно 0,3 или меньше приблизительно 0,1 соответственно). График перестановки также показал, что линия регрессии для Q^2 опустилась ниже нуля, что еще больше указывает на то, что модель PLS является уникальным решением для набора данных.

[0071] На фиг. 13 изображена диаграмма 1300, на которой сравниваются фактическая эффективность очистки (в данном примере уровни НСР) с доверительными пределами, прогнозированными моделью М6, для четырех разных партий лекарственного вещества («Партия 1» - «Партия 4»), представленных в наборе данных для подтверждения. Как видно на диаграмме 1300, во всех партиях лекарственного вещества в наборе данных для подтверждения фактические уровни НСР находились в пределах прогнозированного диапазона НСР. Кроме того, абсолютные различия между фактическими и прогнозированными значениями находились в пределах среднеквадратичной ошибки от

перекрестной проверки или «RMSECV». Модель М6 продемонстрировала точность, подобную отбору мелкомасштабной модели, описанному в настоящем документе, но потребовала меньшего использования ресурсов (например, меньше рабочей силы), отсутствия аналитических испытаний и значительного сокращения времени исполнения и обработки результатов (т. е. приблизительно с 80 часов до двух часов для времени исполнения и с 672 часов до двух часов для времени обработки результатов). Еще одно преимущество предиктивной (например, PLS) многомерной статистической модели по сравнению с отбором мелкомасштабной модели заключается в том, что первая позволяет пользователю оценить любую комбинацию значений атрибута загружаемого собранного материала и смолы *in silico*. Как и в случае прогнозирования НСР с помощью мелкомасштабной модели, прогнозы с помощью многомерной статистической модели предполагают, что последующая часть процесса производства лекарственных веществ выполняется в пределах нормального диапазона эксплуатации (NOR).

[0072] На фиг. 14 представлена блок-схема, в которой изображен иллюстративный способ 1400 выбора исходных материалов для использования в процессе очистки путем хроматографии, например, для производства терапевтических белков. Способ 1400 может быть реализован частично (например, блоки с 1410 по 1420 или 1430) с помощью процессорного узла 120 вычислительной системы 102 (при исполнении программных инструкций приложения 130 для выбора смолы, хранящегося в узле 128 памяти) или с помощью одного или нескольких процессоров обучающего сервера 104 (например, в реализации облачного сервиса), например.

[0073] В блоке 1410 для каждой из одной или нескольких кандидатных смол получают соответственный набор значений атрибута смолы, при этом каждый набор содержит по меньшей мере одно аналитическое измерение кандидатной смолы. Если способ 1400 используется для отбора отдельных партий смолы, например, блок 1410 может включать получение одного набора значений атрибута смолы для одной партии смолы. В качестве другого примера, если способ 1400 используется для выбора среди разных продуктов смолы или разных партий смолы, предложенных производителем, блок 1410 может включать получение нескольких наборов значений атрибута смолы, соответствующих разным продуктам или партиям смолы. Некоторые или все значения атрибута смолы могут быть получены (прямо или косвенно) от производителя или поставщика кандидатной смолы (кандидатных смол), например, в виде CoA или другом формате. Например, производитель или поставщик смолы может провести любое аналитическое измерение (любые аналитические измерения), необходимое (необходимые) для получения данных CoA, а затем физически или в электронном виде предоставить CoA субъекту (например, производителю лекарственного средства), который выполняет способ 1400. В качестве примеров, значения атрибута смолы могут содержать один или несколько из диаметра пор, объема пор, %20-30 мкм, несвязанные, коэффициента емкости, % по количеству 2-10 мкм в среднем 3-связанных, % по объему 20-30 мкм в среднем 3-связанных, среднего размера частиц, времени удерживания рибонуклеазы,

времени удерживания инсулина, времени удерживания лизоцима, времени удерживания миоглобина, времени удерживания овальбумина, времени удерживания окситоцина, времени удерживания брадикинина, времени удерживания angioII , времени удерживания neuro и/или времени удерживания angioI . В некоторых вариантах осуществления блок 1410 включает прием данных, введенных пользователем вручную (например, пользователь вводит данные из CoA).

[0074] В блоке 1420 для каждой из одной или нескольких кандидатных смол прогнозируют соответствующее значение показателя эффективности (для процесса очистки путем колоночной хроматографии) путем применения соответствующего набора значений атрибута смолы в качестве входных данных к многомерной статистической модели (например, многомерной статистической модели 110 или локальной многомерной статистической модели 138 по фиг. 1). Следует понимать, что «применение» и «прогнозирование» блока 1420 может включать непосредственную работу модели (например, в качестве локальной копии, такой как локальная многомерная статистическая модель 138) или может включать инициирование работы модели (например, путем отправки входных значений на обучающий сервер 104 по сети 108 и запроса прогноза/результата и т. д.).

[0075] Модель может представлять собой модель проекции на латентные пространства (PLS), например, с любым подходящим количеством главных компонентов (например, двумя главными компонентами). Альтернативно модель может быть любым другим подходящим типом многомерной статистической модели (например, эластичная сеть, дерево решений и т. д.). Показатель эффективности может представлять собой уровень НСР (например, концентрацию), полученный, например, в результате процесса очистки путем колоночной хроматографии. Альтернативно показатель эффективности может быть уровнем другого типа примеси (например, агрегированные белки, белковые фрагменты и т. д.), общим уровнем всех типов примесей или любым другим подходящим показателем чистоты материала, полученным в результате от процесса очистки путем колоночной хроматографии. В некоторых вариантах осуществления каждое соответствующее «значение» представляет собой диапазон значений. Например, модель может вывести диапазон значений, для которых превышен некоторый минимальный уровень достоверности (например, 90%, 95% и т. д.).

[0076] В некоторых вариантах осуществления блок 1420 также включает применение одного или нескольких других типов входных данных к многомерной статистической модели вместе со значениями атрибута смолы. Например, блок 1420 может включать применение одного или нескольких значений параметров собранного фильтрата (например, конечная жизнеспособность производственного биореактора, общая площадь индивидуальной RP-HPLC для DFM, основная площадь индивидуальной RP-HPLC для DFM, площадь примесей индивидуальной RP-HPLC для DFM, индивидуальная PI для DFM, титр индивидуальной PI для DFM и/или другие параметры) и/или одного или нескольких значений параметра процесса хроматографии/очистки (например, HETP

колонки 1, асимметрия колонки 1 и/или другие параметры) в качестве входных данных к многомерной статистической модели.

[0077] В блоке 1430 смолу из одной или нескольких кандидатных смол выбирают на основе по меньшей мере частично прогнозированного соответственного значения (прогнозированных соответственных значений) показателя эффективности. «Выбор» может быть подтверждением или утверждением конкретной партии смолы, например, или обозначением конкретного типа или партии смолы как приемлемого и т. д. В некоторых вариантах осуществления блок 1430 выполняется автоматически программным обеспечением (например, процессорным узлом 120 вычислительной системы 102 при исполнении программных инструкций приложения 130 для выбора смолы или одним или несколькими процессорами обучающего сервера 104 и т. д.). Альтернативно блок 1430 может быть полностью или частично выполнен одним или несколькими пользователями с учетом прогнозированного значения (прогнозированных значений). Для этого блок 1430 может включать предписание пользовательскому интерфейсу (например, GUI, созданному или заполненному узлом 136 визуализации и представленному на дисплее 124 по фиг. 1) представить значение (значения) прогнозированного показателя эффективности и/или результат на основе прогнозированного значения (прогнозированных значений) (например, указание на то, удовлетворяют ли одно или несколько значений прогнозированного показателя эффективности одному или нескольким критериям приемлемости, например, при превышении некоторого порогового значения), чтобы облегчить пользователю выбор одной или нескольких конкретных кандидатных смол (например, партий смол).

[0078] Независимо от того, реализуется ли блок 1430 автоматически, вручную или с помощью некоторой их комбинации, блок 1430 может включать сравнение значения (значений) прогнозированного показателя эффективности с заранее определенным порогом приемлемости. Например, блок 1430 может включать выбор смолы, только если соответствующий показатель эффективности (например, уровень НСР) ниже порога приемлемости. Альтернативно, если многомерная статистическая модель выводит диапазон значений (например, диапазон, для которого превышен порог достоверности), блок 1430 может включать сравнение диапазона (диапазонов) значений показателя эффективности с заранее определенным порогом приемлемости. Например, блок 1430 может включать выбор смолы, только если все значения в соответствующем диапазоне (например, соответствующем диапазоне уровней НСР) ниже порога приемлемости.

[0079] В блоке 1440 процесс очистки путем колоночной хроматографии выполняют с использованием смолы, которая была выбрана (например подтверждена/утверждена) в блоке 1430 в качестве неподвижной фазы в системе колоночной хроматографии. В некоторых вариантах осуществления система колоночной хроматографии представляет собой систему промышленного масштаба. Блок 1440 может быть выполнен с помощью системы 140 колоночной хроматографии по фиг. 1, например, возможно, с помощью ручного управления. В многоколоночных системах выбранная смола может

использоваться в качестве неподвижной фазы в одной, некоторых или всех из колонок.

[0080] В некоторых вариантах осуществления способ 1400 включает один или несколько других блоков, не видимых на фиг. 14. Если, например, процесс/система хроматографической очистки из блока 1440 имеет промышленный масштаб, способ 1400 может включать дополнительный блок перед блоком 1410, в котором многомерная статистическая модель обучается с использованием исторических данных хроматографической очистки промышленного масштаба (и, возможно, также мелкомасштабных) (т. е. исторических значений входных данных/признаков и соответствующих показателей эффективности/маркировок).

[0081] Хотя системы, способы, устройства и их компоненты были описаны в рамках примерных вариантов осуществления, они не ограничиваются ими. Подробное описание следует интерпретировать как приведенное только в качестве примера, и в нем не описан каждый возможный вариант осуществления настоящего изобретения, поскольку описание каждого возможного варианта осуществления было бы непрактичным, если не невозможным. Могут быть реализованы многочисленные альтернативные варианты осуществления с использованием либо современной технологии, либо технологии, разработанной после даты подачи настоящего патента, которые, тем не менее, находятся в пределах объема формулы изобретения, определяющей настоящее изобретение.

[0082] Специалисты в данной области техники поймут, что широкое разнообразие модификаций, изменений и комбинаций может быть выполнено применительно к описанным выше вариантам осуществления без отхода от объема настоящего изобретения, и что такие модификации, изменения и комбинации необходимо рассматривать как попадающие в объем идеи настоящего изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ выбора исходных материалов для использования в процессе очистки путем колоночной хроматографии, причем способ включает:

для каждой из одной или нескольких кандидатных смол получение с помощью одного или нескольких процессоров вычислительной системы соответственного набора значений атрибута смолы, причем соответственный набор значений атрибута смолы содержит по меньшей мере одно аналитическое измерение кандидатной смолы;

для каждой из одной или нескольких кандидатных смол прогнозирование с помощью одного или нескольких процессоров, применяющих соответственный набор значений атрибута смолы в качестве входных данных к многомерной статистической модели, соответственного значения показателя эффективности для процесса очистки путем колоночной хроматографии;

выбор смолы из одной или нескольких кандидатных смол на основе по меньшей мере частично одного или нескольких прогнозированных соответственных значений показателя эффективности; и

выполнение процесса очистки путем колоночной хроматографии с использованием выбранной смолы в качестве неподвижной фазы.

2. Способ по п. 1, отличающийся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра собранного фильтрата в качестве входных данных к многомерной статистической модели.

3. Способ по п. 1, отличающийся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

4. Способ по п. 1, отличающийся тем, что прогнозирование соответственного значения показателя эффективности включает применение (i) соответственного набора значений атрибута смолы, (ii) одного или нескольких значений параметра собранного фильтрата и (iii) одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

5. Способ по любому из пп. 1-4, отличающийся тем, что прогнозирование соответственного значения показателя эффективности включает прогнозирование уровня белка клетки-хозяина, полученного в результате процесса очистки путем колоночной хроматографии.

6. Способ по любому из пп. 1-5, отличающийся тем, что выбор смолы включает сравнение одного или нескольких соответственных значений показателя эффективности с заранее определенным порогом приемлемости.

7. Способ по любому из пп. 1-6, отличающийся тем, что:

прогнозирование соответственного значения показателя эффективности включает прогнозирование соответственного диапазона значений показателя эффективности; и

выбор смолы включает сравнение каждого из одного или нескольких соответственных диапазонов значений с заранее определенным порогом приемлемости.

8. Способ по любому из пп. 1-7, отличающийся тем, что прогнозирование соответственного значения показателя эффективности включает применение соответственного набора характеристик смолы в качестве входных данных к регрессионной модели проекции на латентные структуры (PLS).

9. Способ по любому из пп. 1-8, отличающийся тем, что для каждой из одной или нескольких кандидатных смол получение соответственного набора значений атрибута смолы включает получение значений атрибута смолы, предоставленных производителем или поставщиком кандидатной смолы.

10. Способ по любому из пп. 1-9, отличающийся тем, что:
одна или несколько кандидатных смол включают в себя множество кандидатных смол, которые соответствуют разным производственным партиям одного типа смолы.

11. Способ по любому из пп. 1-10, отличающийся тем, что:
процесс очистки путем колоночной хроматографии представляет собой процесс очистки путем хроматографии промышленного масштаба; и

способ дополнительно включает обучение многомерной статистической модели с использованием исторических данных процесса очистки путем хроматографии мелкого масштаба и промышленного масштаба.

12. Система, содержащая:
вычислительную систему, которая содержит один или несколько процессоров и одно или несколько запоминающих устройств, причем одно или несколько запоминающих устройств хранят инструкции, которые при исполнении одним или несколькими процессорами предписывают вычислительной системе для каждой из одной или нескольких кандидатных смол

получить соответственный набор значений атрибута смолы, причем соответственный набор значений атрибута смолы содержит по меньшей мере одно аналитическое измерение кандидатной смолы,

спрогнозировать путем применения соответственного набора значений атрибута смолы в качестве входных данных к многомерной статистической модели соответственного значения показателя эффективности для процесса очистки путем колоночной хроматографии и

отобразить соответственное значение показателя эффективности или результат на основе соответственного значения показателя эффективности пользователю для облегчения выбора смолы из одной или нескольких кандидатных смол на основе по меньшей мере частично одного или нескольких прогнозированных соответственных значений показателя эффективности; и

систему колоночной хроматографии, выполненную с возможностью выполнения процесса очистки путем колоночной хроматографии с использованием выбранной смолы в качестве неподвижной фазы.

13. Система по п. 12, отличающаяся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра собранного фильтрата в качестве входных данных к многомерной статистической модели.

14. Система по п. 12, отличающаяся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

15. Система по п. 12, отличающаяся тем, что прогнозирование соответственного значения показателя эффективности включает применение (i) соответственного набора значений атрибута смолы, (ii) одного или нескольких значений параметра собранного фильтрата и (iii) одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

16. Система по любому из пп. 12-15, отличающаяся тем, что соответственное значение показателя эффективности представляет собой уровень белка клетки-хозяина, полученного в результате процесса очистки путем колоночной хроматографии.

17. Система по любому из пп. 12-16, отличающаяся тем, что прогнозирование соответственного значения показателя эффективности включает прогнозирование соответственного диапазона значений показателя эффективности.

18. Система по любому из пп. 12-17, отличающаяся тем, что многомерная статистическая модель содержит регрессионную модель проекции на латентные структуры (PLS).

19. Система по любому из пп. 12-18, отличающаяся тем, что для каждой из одной или нескольких кандидатных смол получение соответственного набора значений атрибута смолы включает получение значений атрибута смолы, предоставленных производителем или поставщиком кандидатной смолы.

20. Система по любому из пп. 12-19, отличающаяся тем, что одна или несколько кандидатных смол включают в себя множество кандидатных смол, которые соответствуют разным производственным партиям одного типа смолы.

21. Энергонезависимый машиночитаемый носитель, хранящий инструкции, которые при исполнении одним или несколькими процессорами или вычислительной системой предписывают вычислительной системе для каждой из одной или нескольких кандидатных смол:

получить соответственный набор значений атрибута смолы, причем соответственный набор значений атрибута смолы содержит по меньшей мере одно аналитическое измерение кандидатной смолы;

спрогнозировать путем применения соответственного набора значений атрибута смолы в качестве входных данных к многомерной статистической модели соответственного значения показателя эффективности для процесса очистки путем колоночной хроматографии; и

отобразить соответственное значение показателя эффективности или результат на основе соответственного значения показателя эффективности пользователю для облегчения выбора смолы из одной или нескольких кандидатных смол для использования в качестве неподвижной фазы в процессе очистки путем колоночной хроматографии на основе по меньшей мере частично одного или нескольких прогнозируемых соответственных значений показателя эффективности.

22. Энергонезависимый машиночитаемый носитель по п. 21, отличающийся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра собранного фильтрата в качестве входных данных к многомерной статистической модели.

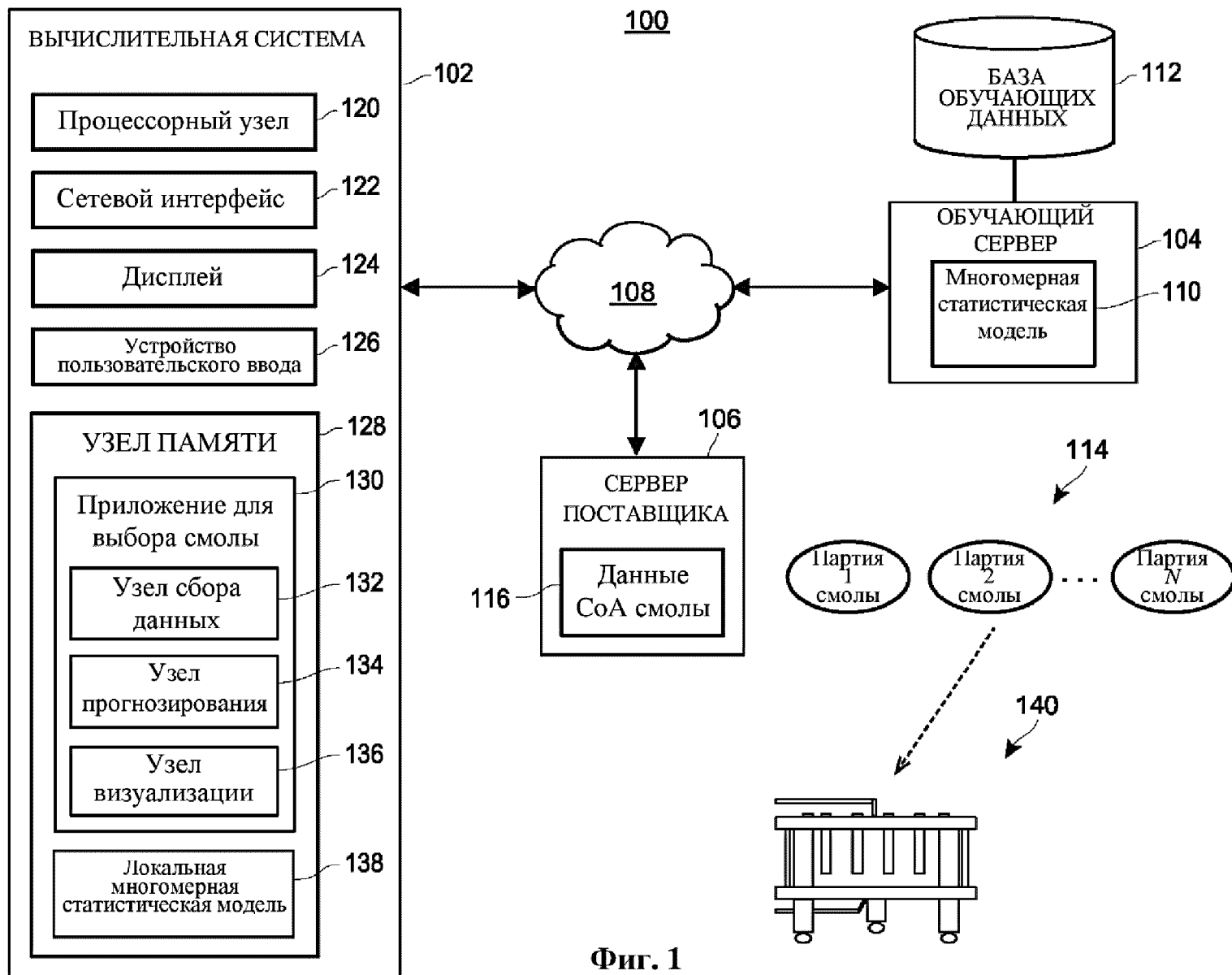
23. Энергонезависимый машиночитаемый носитель по п. 21, отличающийся тем, что прогнозирование соответственного значения показателя эффективности дополнительно включает применение одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

24. Энергонезависимый машиночитаемый носитель по п. 21, отличающийся тем, что прогнозирование соответственного значения показателя эффективности включает применение (i) соответственного набора значений атрибута смолы, (ii) одного или нескольких значений параметра собранного фильтрата и (iii) одного или нескольких значений параметра процесса очистки в качестве входных данных к многомерной статистической модели.

25. Энергонезависимый машиночитаемый носитель по любому из пп. 21-24, отличающийся тем, что соответственное значение показателя эффективности представляет собой уровень белка клетки-хозяина, полученного в результате процесса очистки путем колоночной хроматографии.

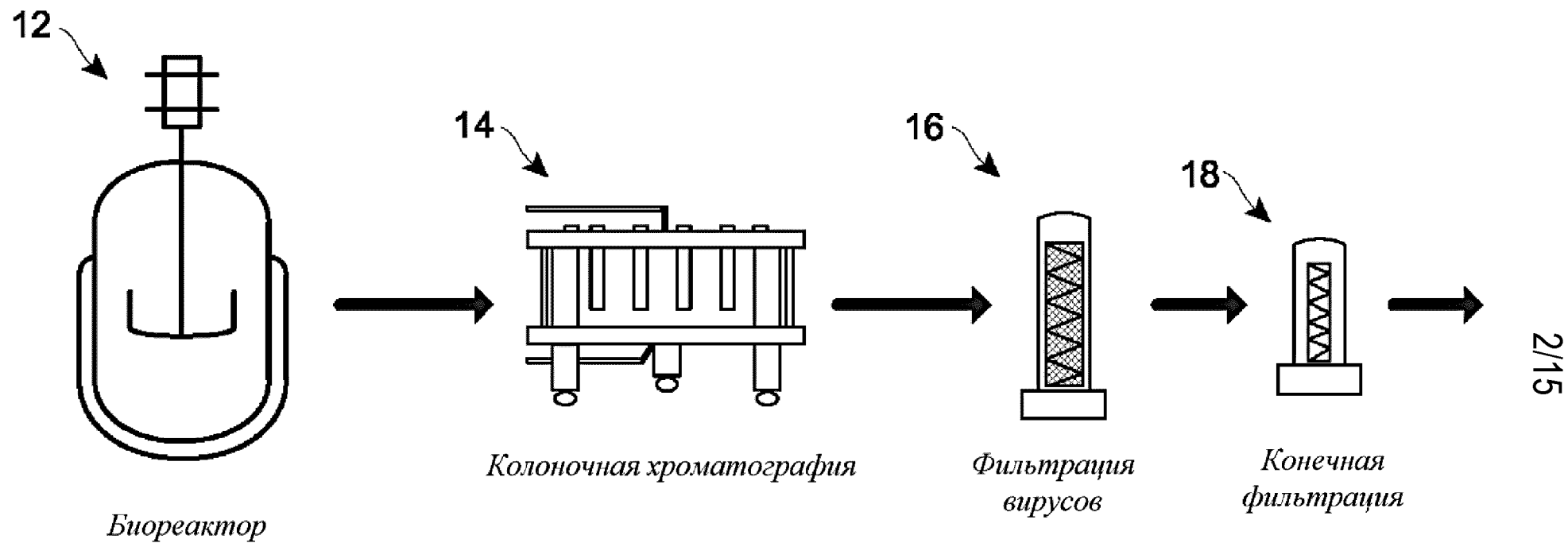
26. Энергонезависимый машиночитаемый носитель по любому из пп. 21-25, отличающийся тем, что многомерная статистическая модель содержит регрессионную модель проекции на латентные структуры (PLS).

По доверенности

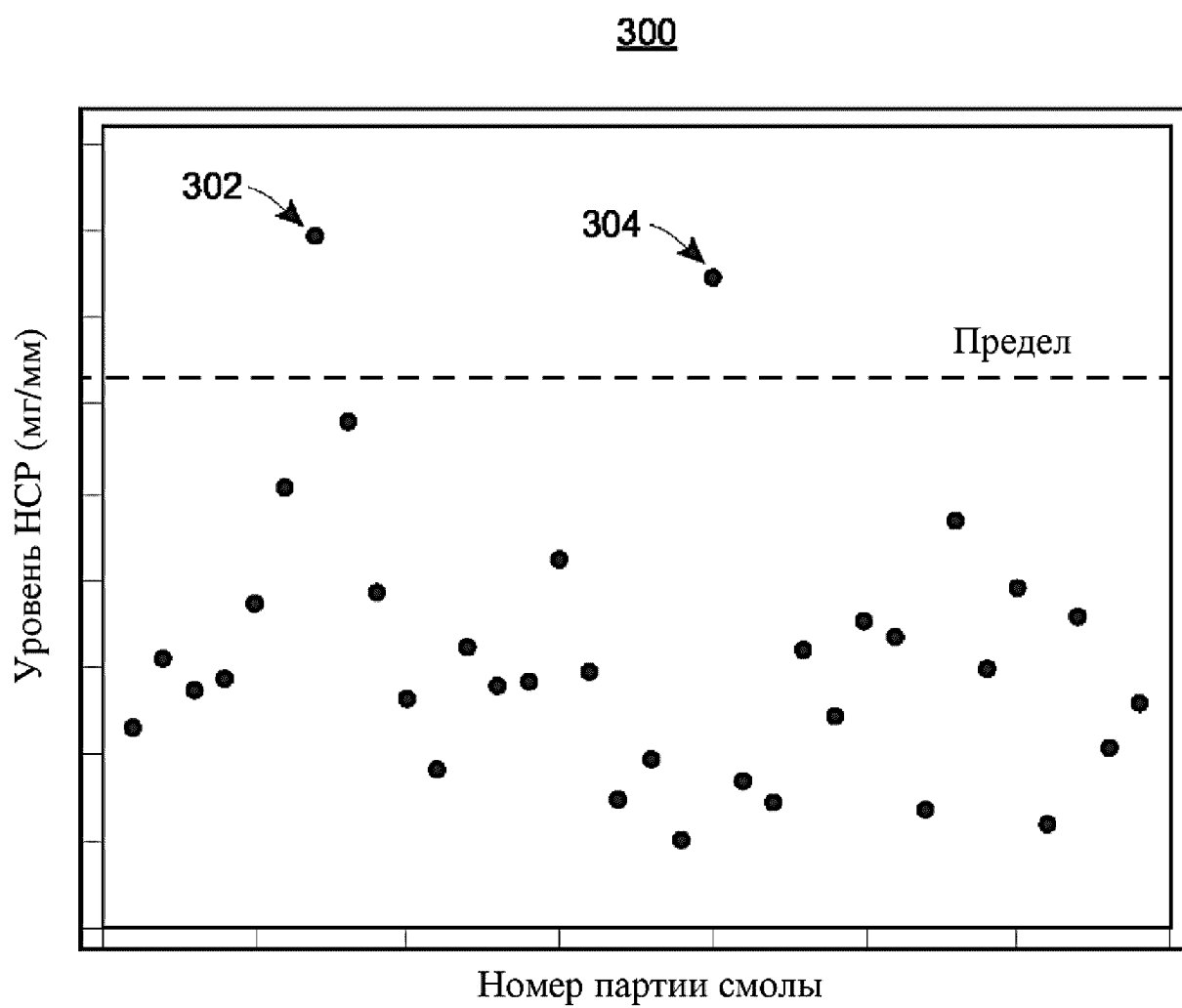


Фиг. 1

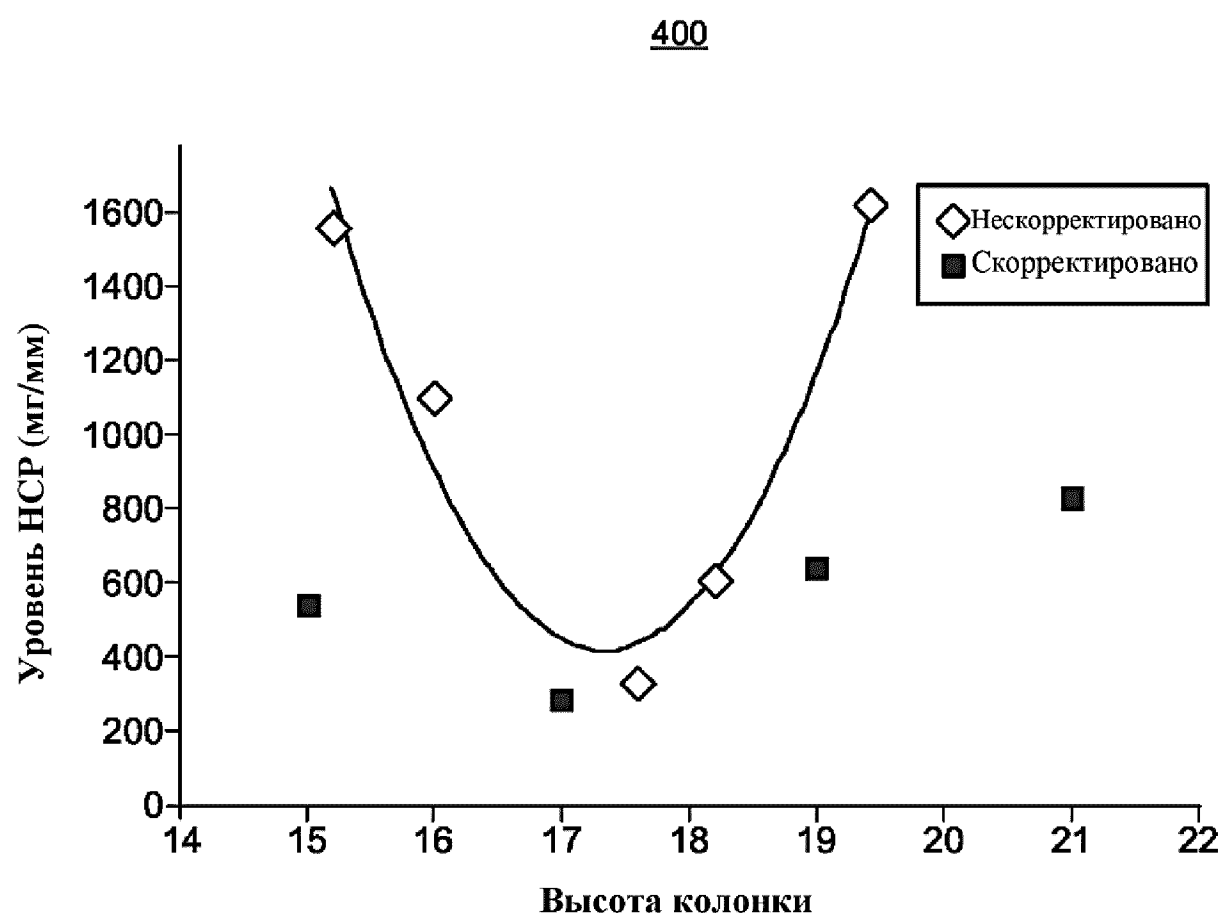
10



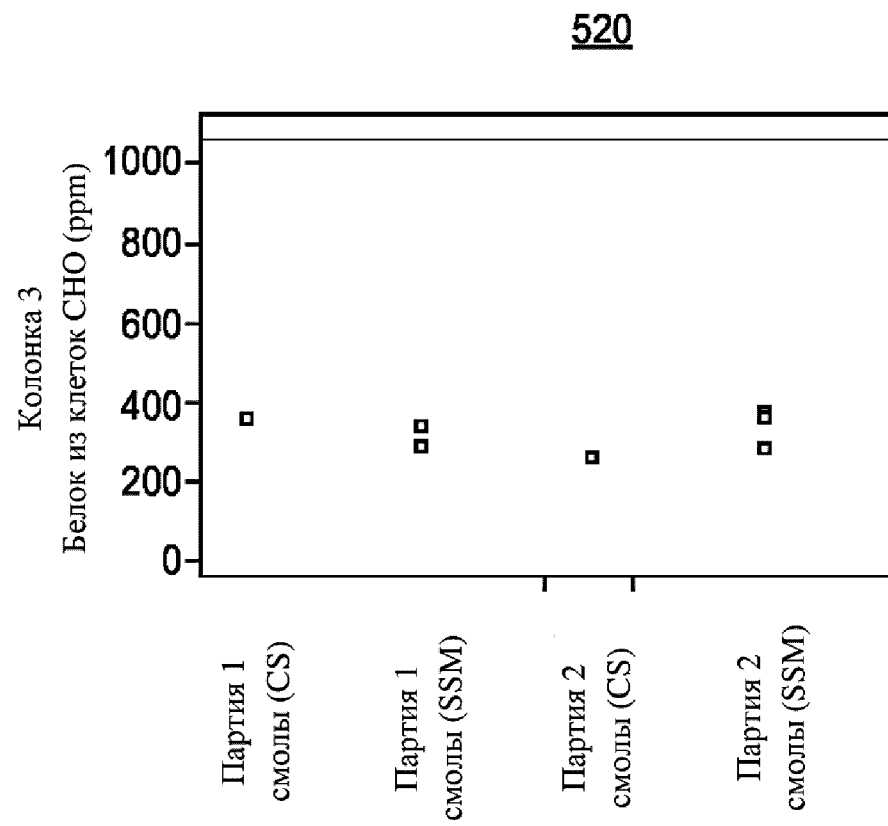
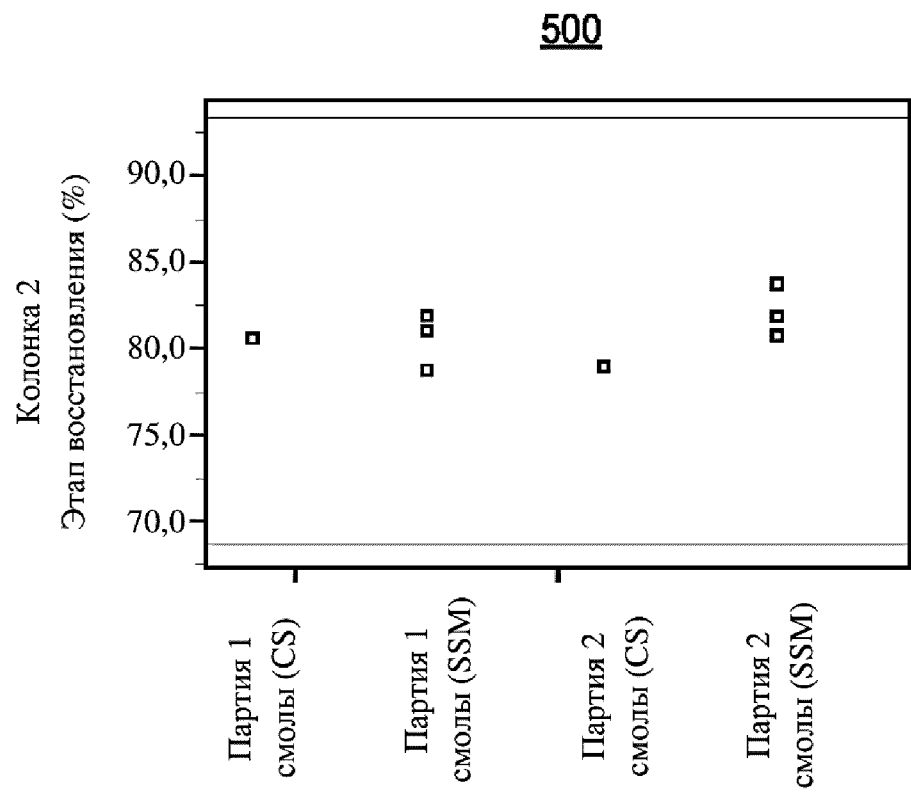
Фиг. 2
(Предыдущий
уровень техники)



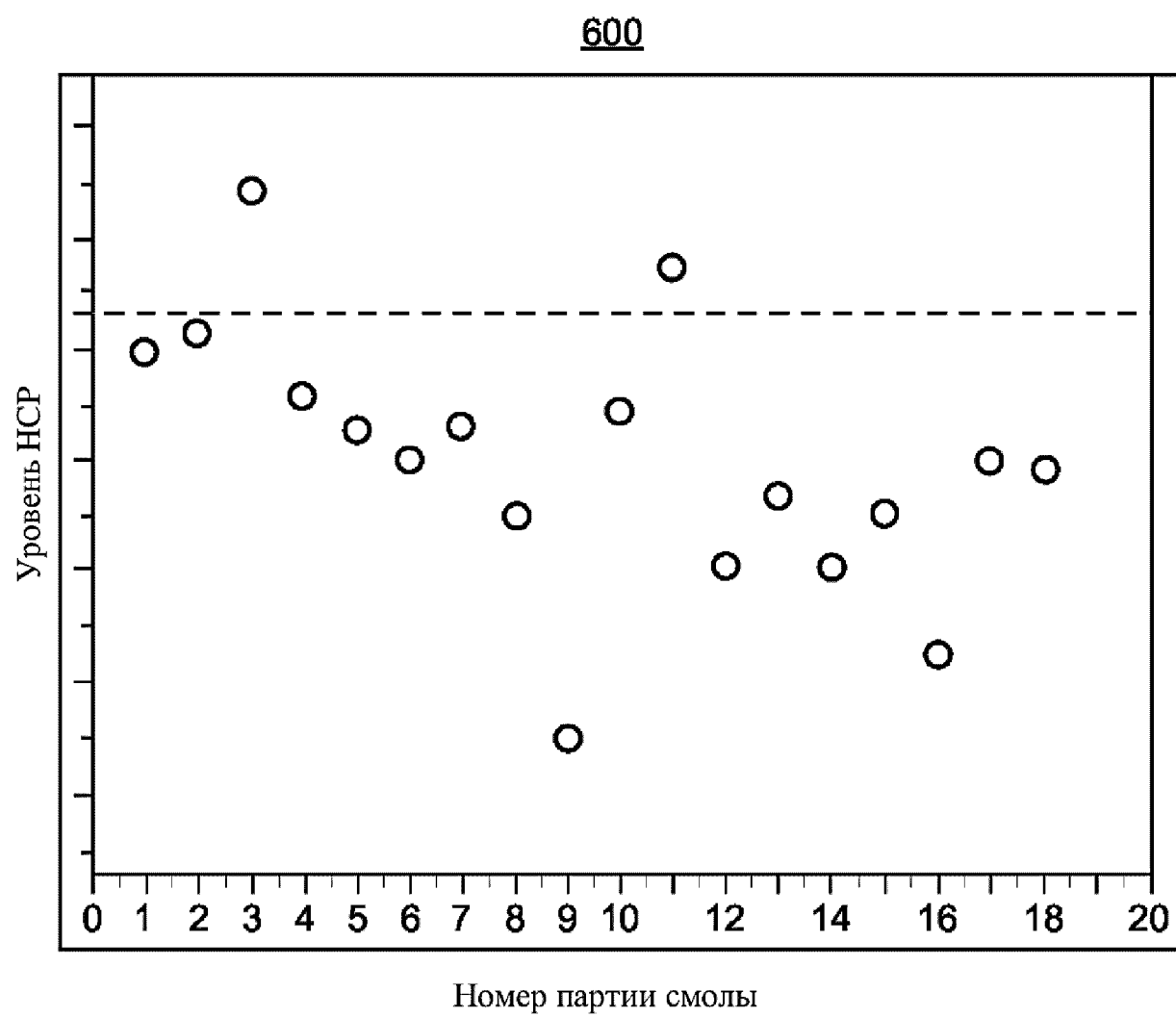
Фиг. 3



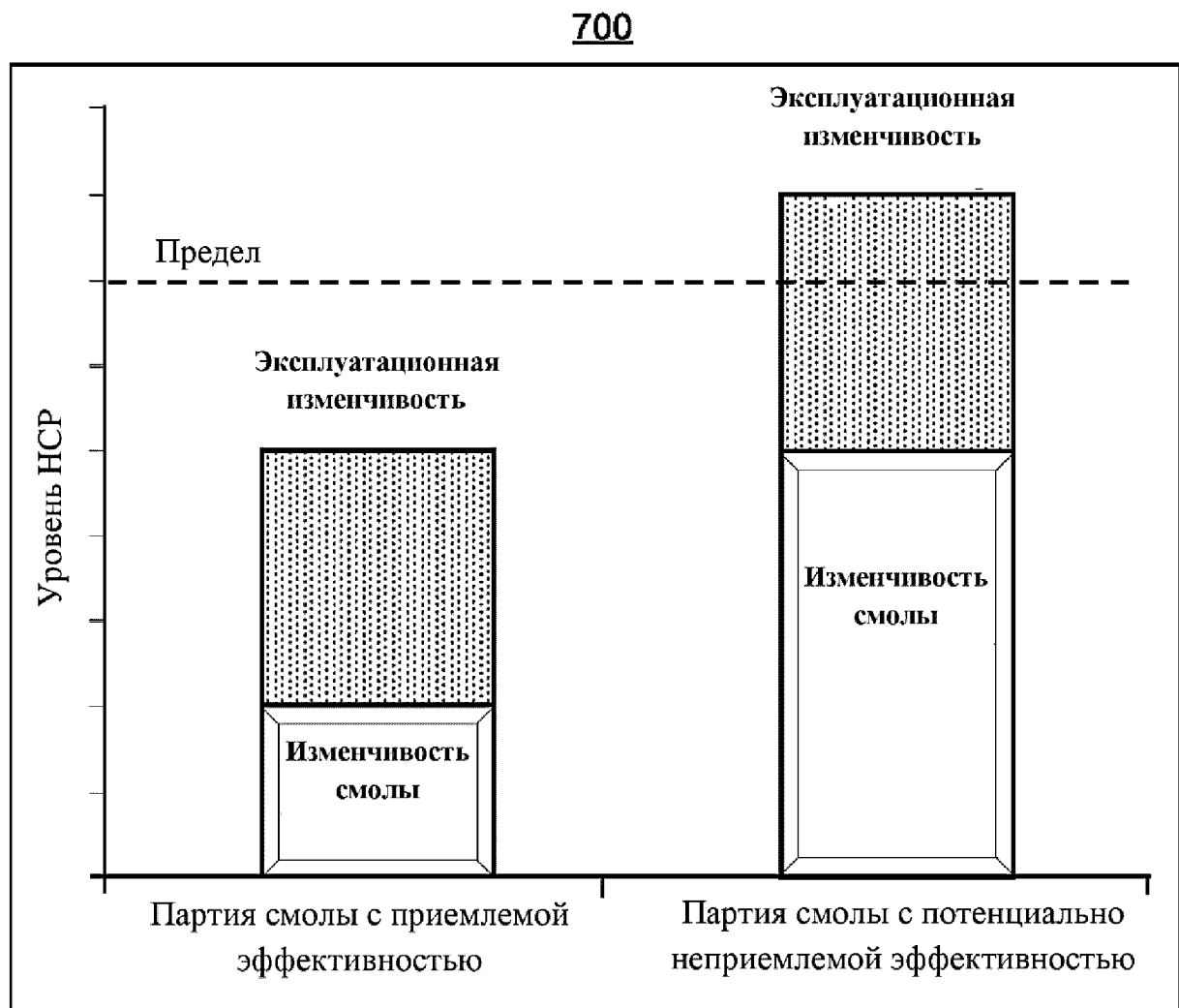
Фиг. 4



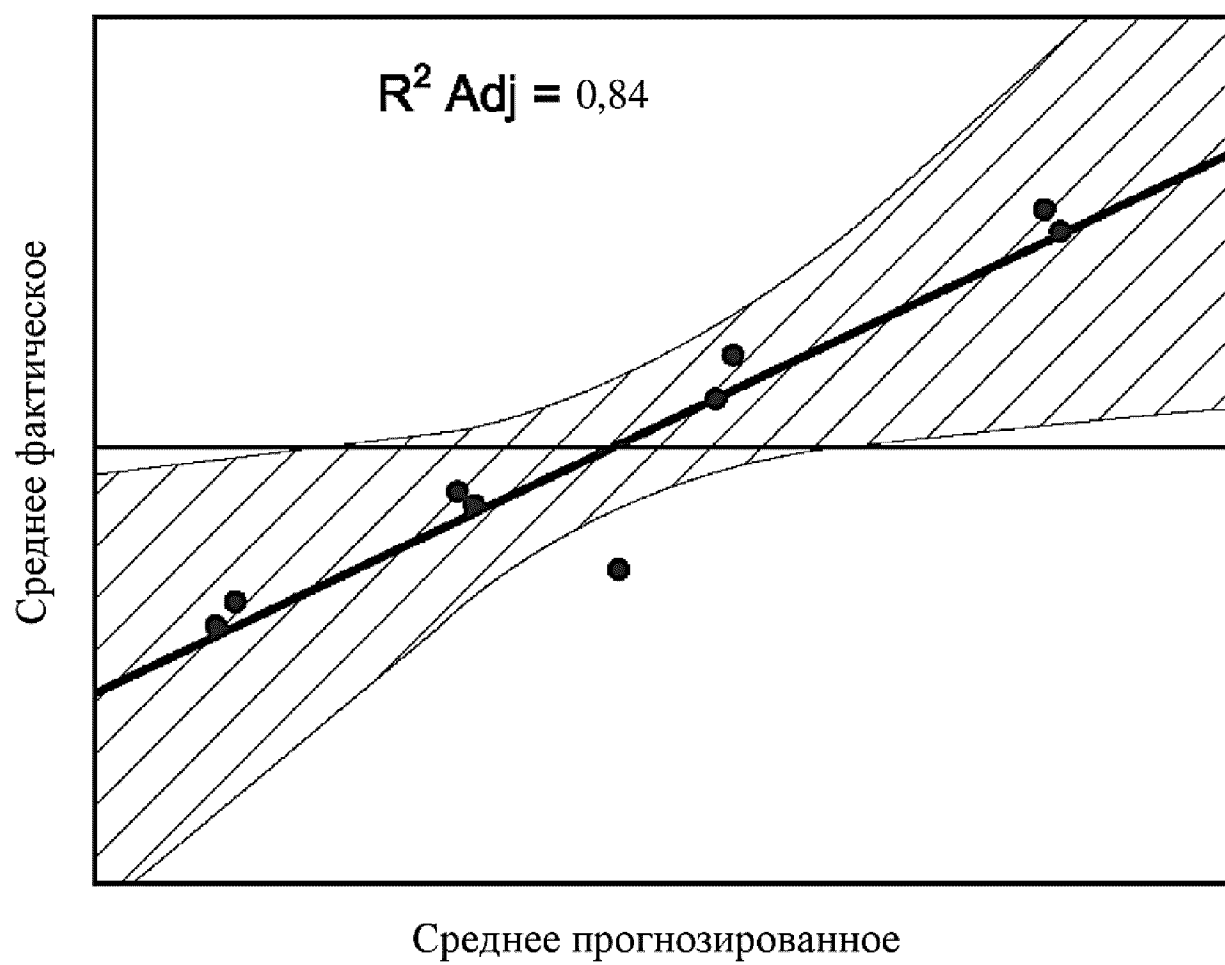
Фиг. 5



Фиг. 6

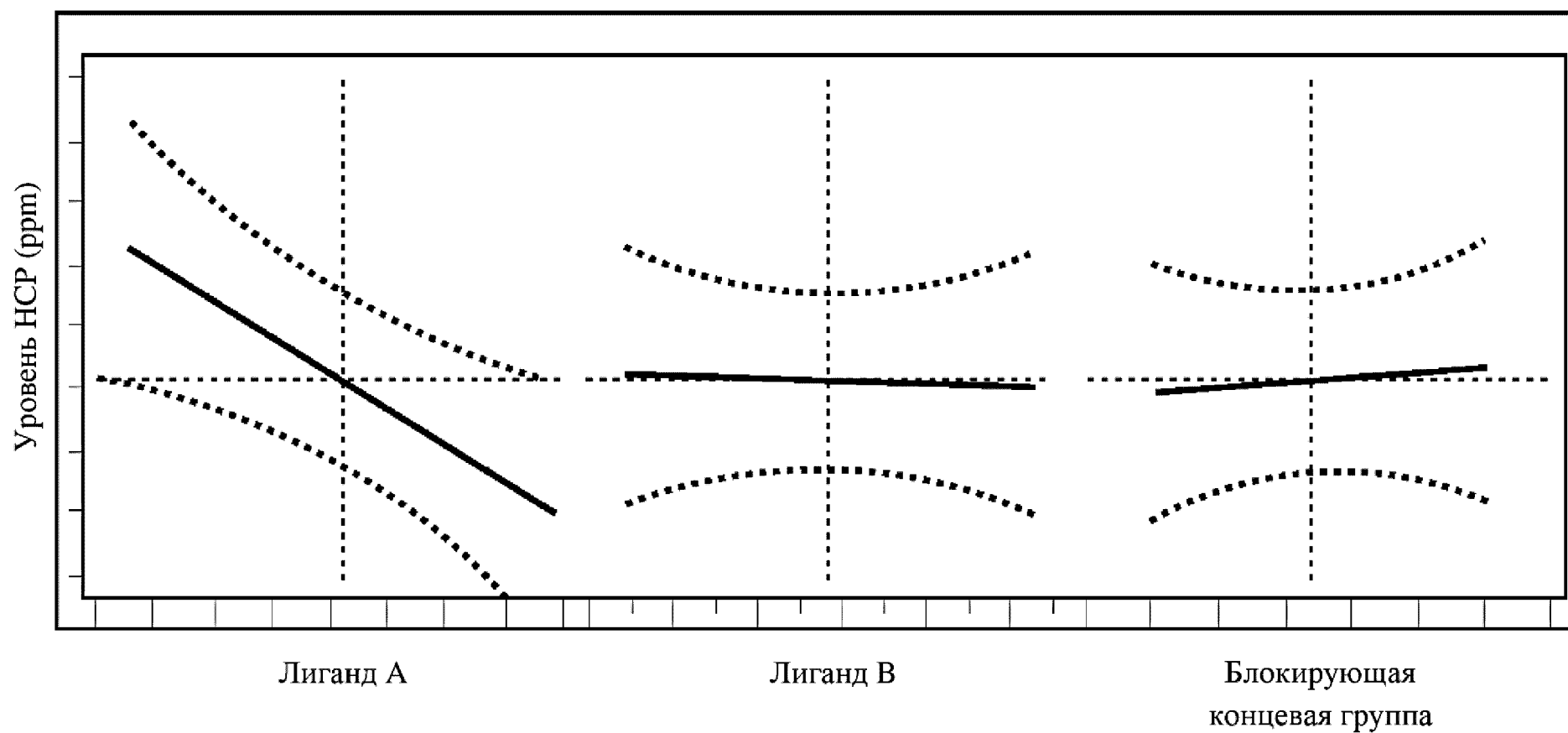


Фиг. 7

800

Фиг. 8

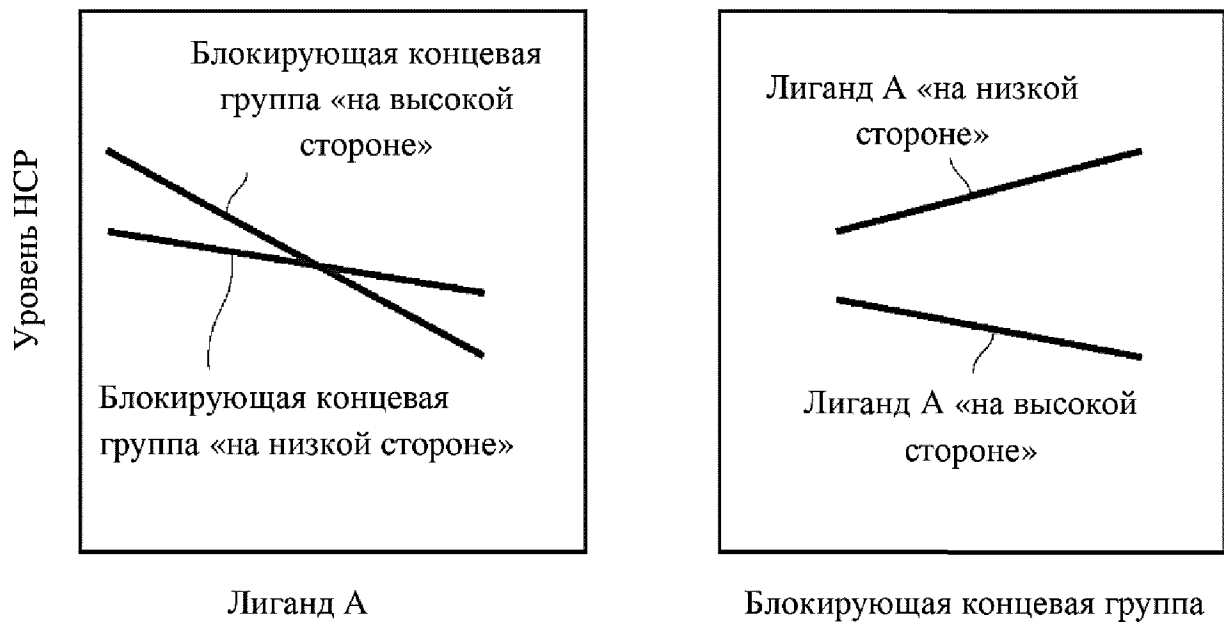
900



9/15

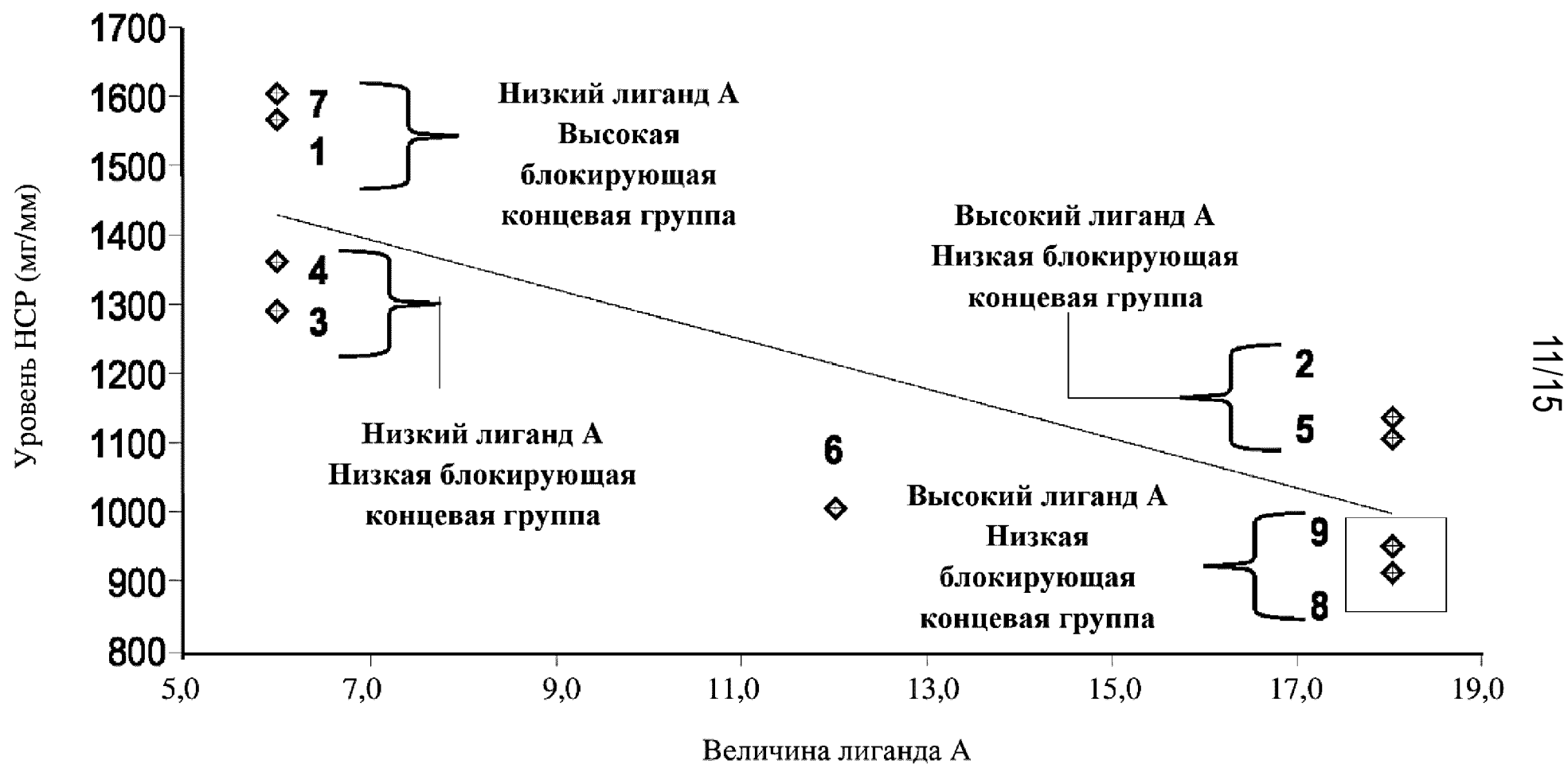
Фиг. 9А

950

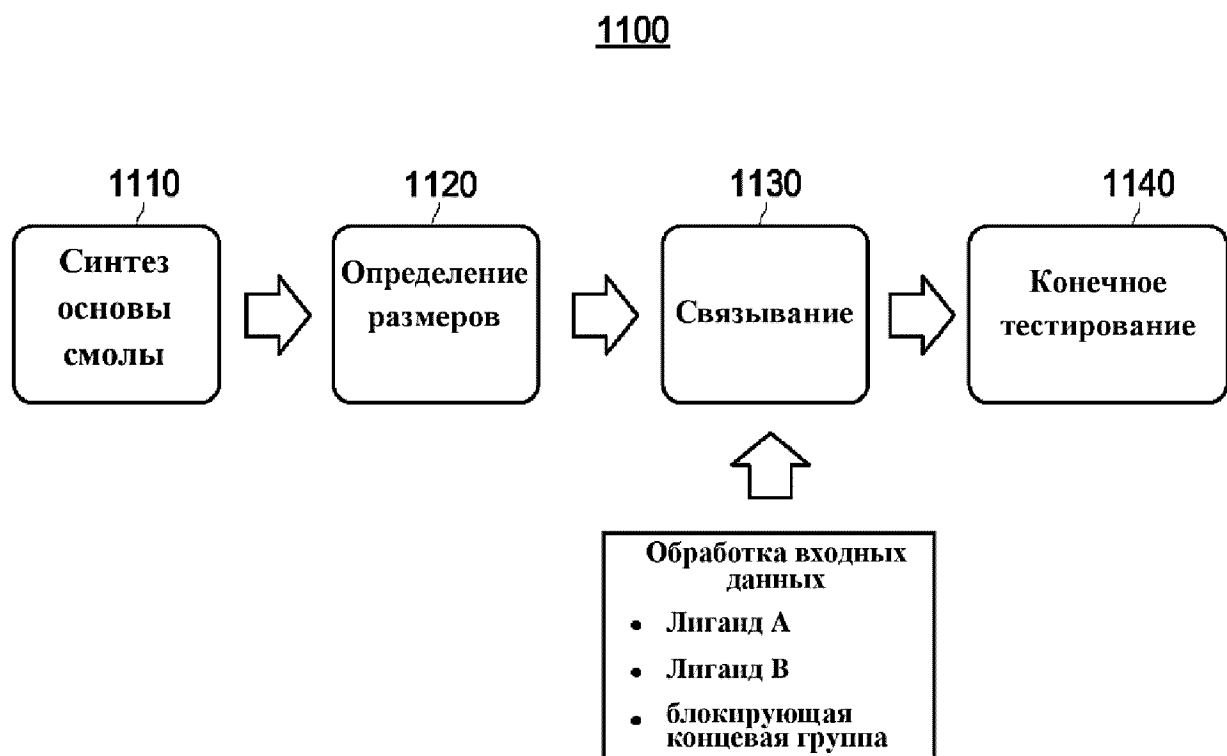


Фиг. 9В

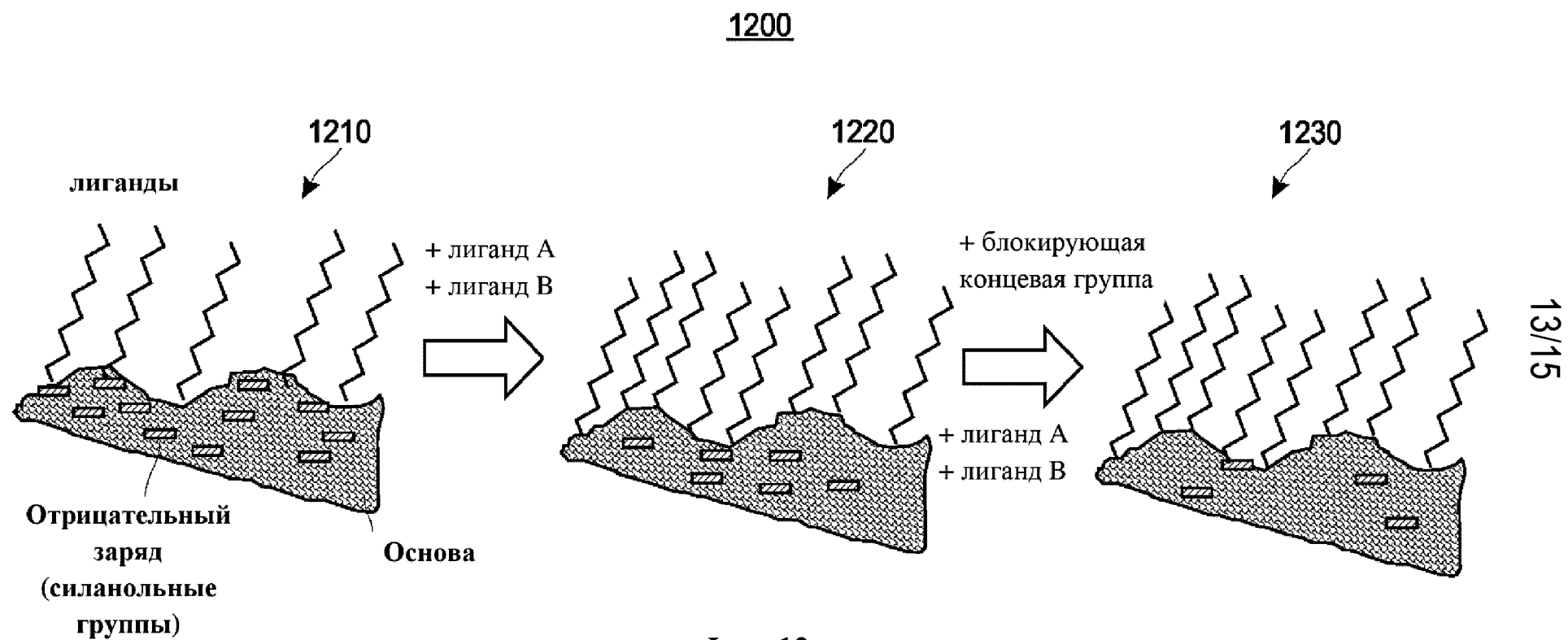
1000



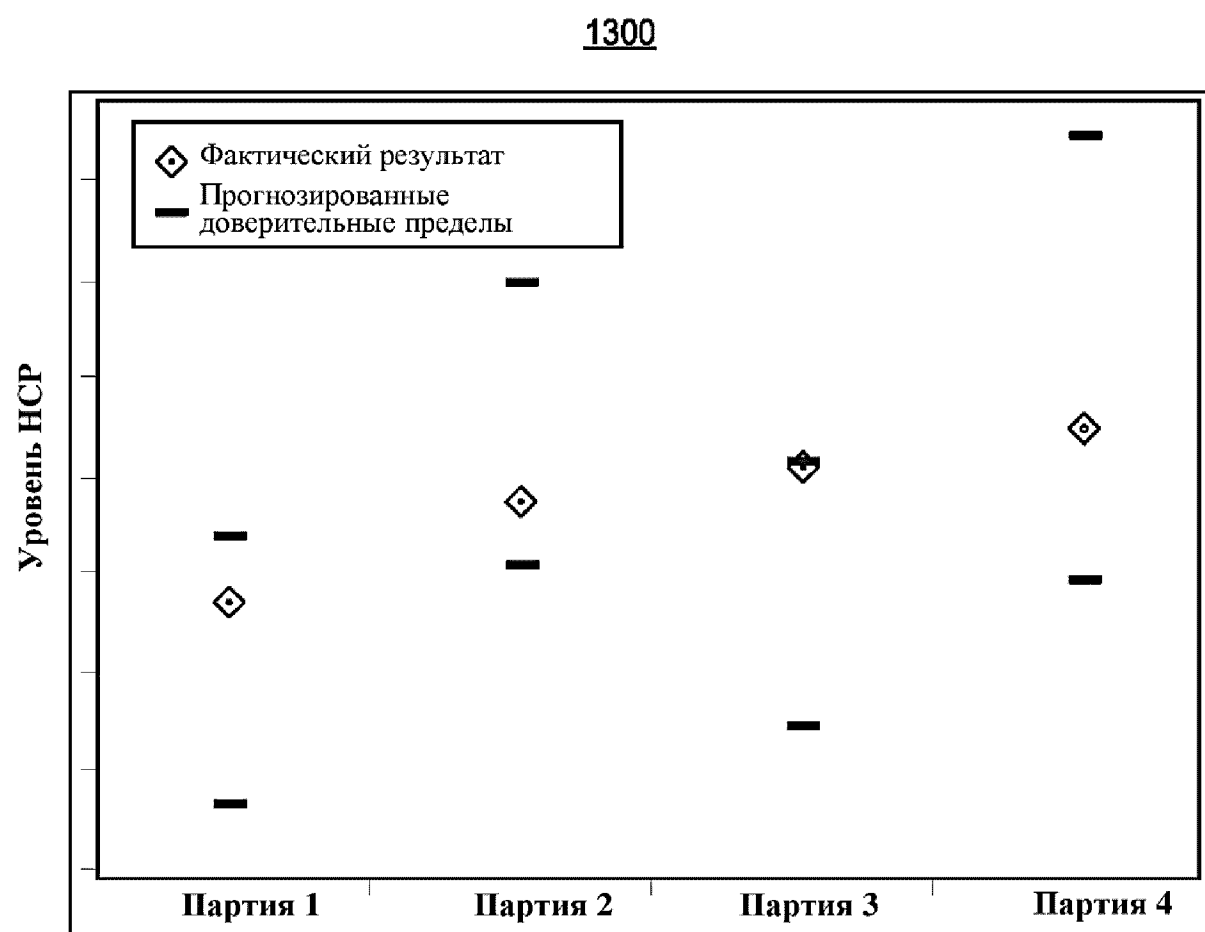
Фиг. 10



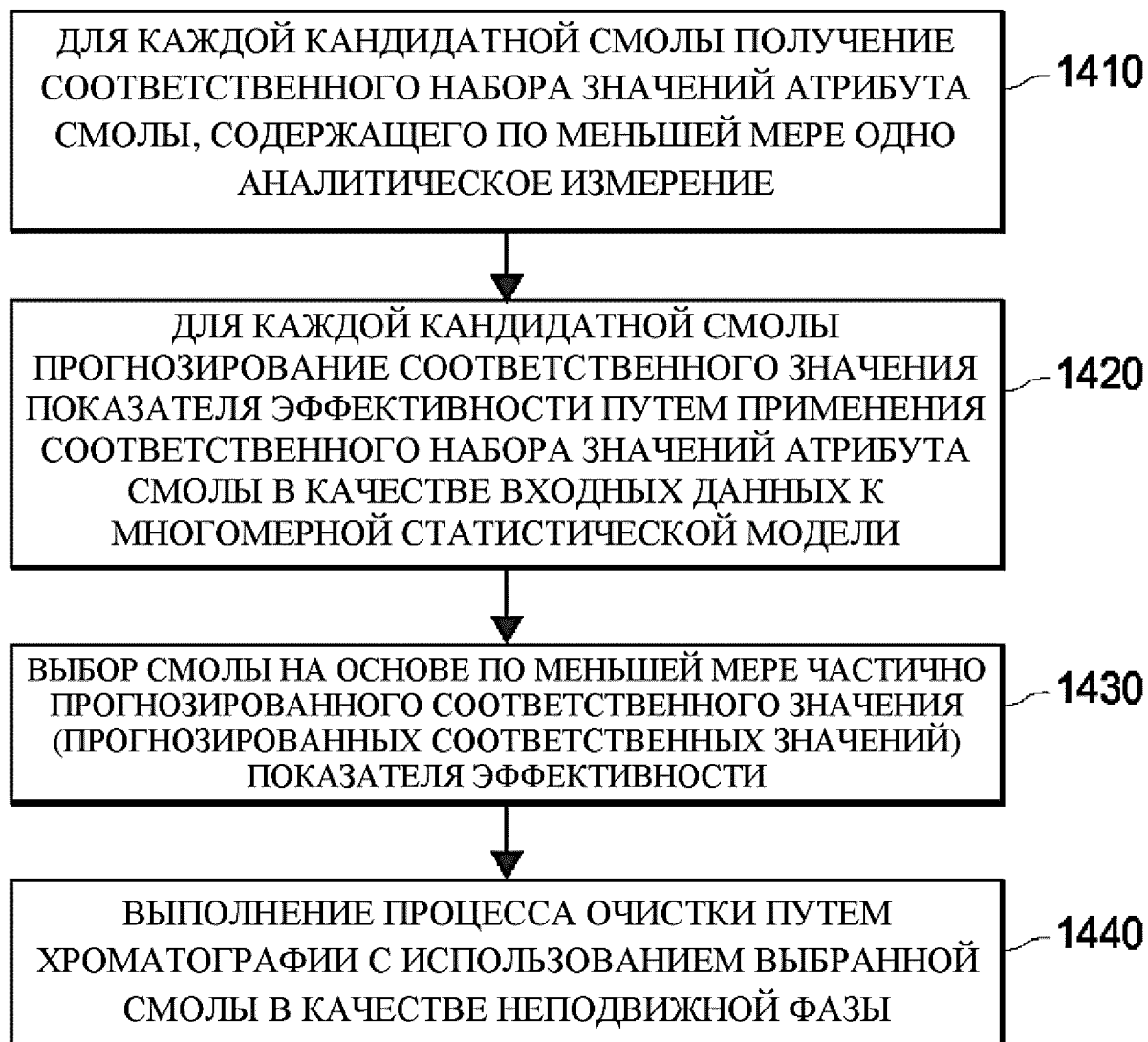
Фиг. 11



Фиг. 12



Фиг. 13

1400**Фиг. 14**