(19) Евразийское патентное ведомство

(21) 202291449 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.10.05

(22) Дата подачи заявки
2020.11.13

(51) Int. Cl. ***C07K 14/00*** (2006.01)
A61K 38/17 (2006.01)
A61K 49/00 (2006.01)
C07K 16/28 (2006.01)

(54) ПОЛИПЕПТИДЫ XTEN СО ШТРИХ-КОДОМ, И КОМПОЗИЦИИ НА ИХ ОСНОВЕ, И СПОСОБЫ ИХ ПОЛУЧЕНИЯ И ПРИМЕНЕНИЯ

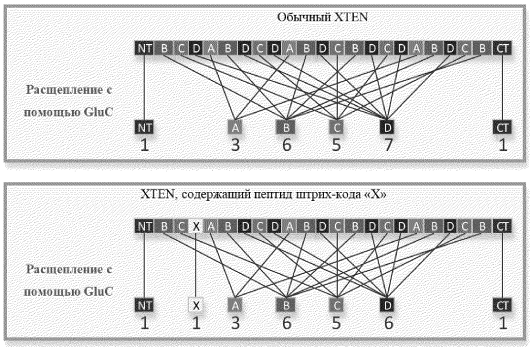
(31) **62/934,980**
(32) **2019.11.13**
(33) **US**
(86) **PCT/US2020/060378**
(87) **WO 2021/097186 2021.05.20**

(71) Заявитель:
АМУНИКС ФАРМАСЬЮТИКАЛЗ, ИНК. (US)

(72) Изобретатель:
Шелленбергер Фолькер, Йохансен Эрик, Хенкенсифкен Анджела (US)

(74) Представитель:
Медведев В.Н. (RU)

(57) В данном документе раскрыты полипептиды, содержащие удлиненный рекомбинантный полипептид (XTEN), состоящий из множества перекрывающихся мотивов последовательностей и одного или более фрагментов со штрих-кодом, высвобождаемых после расщепления протеазами и обнаруживаемых среди всех других протеолитически высвобождаемых фрагментов. Определенные варианты осуществления этих полипептидов дополнительно содержат биологически активный полипептид, где их предпочтительные варианты осуществления содержат сегмент для обеспечения высвобождения, способный к протеолитическому расщеплению, которое расщепляет связь между полипептидом XTEN и биологически активным полипептидом. Также раскрыты способы получения и способы применения указанных полипептидов.



202291449 A1

202291449 A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-574267EA/022

ПОЛИПЕПТИДЫ XTEN СО ШТРИХ-КОДОМ, И КОМПОЗИЦИИ НА ИХ ОСНОВЕ, И СПОСОБЫ ИХ ПОЛУЧЕНИЯ И ПРИМЕНЕНИЯ

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

[1] Настоящая заявка содержит перечень последовательностей, который был представлен в электронном виде в формате ASCII и тем самым включен посредством ссылки во всей своей полноте. Указанная копия в формате ASCII, созданная 6 ноября 2020 г., называется 20-1761-WO_Sequence_Listing_ST25.txt, и ее размер составляет 1494 байта.

ПРЕДПОСЫЛКИ ИЗОБРЕТЕНИЯ

[2] Полипептид может быть получен способом, в результате которого получают смесь полипептидов. Смесь полипептидов часто может включать полноразмерный полипептид вместе с его размерными вариантами (*например*, усеченными). Наличие вариантов, отличающихся по размеру от необходимого полноразмерного продукта, может повлиять на биологическое поведение полипептидного лекарственного вещества, потенциально влияя на безопасность и/или эффективность полипептидного лекарственного вещества. Например, пролекарственные средства на основе белков для терапии рака могут быть построены с нацеленным на опухоль механизмом активации. Более конкретно, полноразмерный терапевтический белок можно получать и вводить в неактивной (нецитотоксической) пролекарственной форме, которая превращается в активное лекарственное средство путем предпочтительного удаления части пролекарственного полипептида в предполагаемом биологическом сайте (*например*, опухоли). Усеченные варианты полноразмерной конструкции могут терять защитные последовательности и становиться цитотоксическими (активными), тем самым «загрязняя» пролекарственную композицию и создавая смесь, содержащую компоненты, которые непреднамеренно активны за пределами предполагаемого биологического сайта. В некоторых случаях такие варианты с более короткой длиной могут представлять больший риск иммуногенности, иметь менее селективную токсичность в отношении опухолевых клеток, или демонстрировать в меньшей степени необходимый фармакокинетический профиль (*например*, приводя к сужению терапевтического окна) по сравнению с полноразмерным белком, или оказывать вредное непредусмотренное воздействие на реципиента за пределами предполагаемого сайта (*например*, в здоровой ткани). В результате обнаружение и количественная оценка структурных вариаций белка могут иметь важное значение для оценки биологических свойств (*например*, клинической безопасности и фармакологической эффективности) биотерапевтических средств и при создании новых биотерапевтических средств (*например*, с повышенной эффективностью и уменьшенными побочными эффектами). Существующие методики и способы идентификации и количественного определения количества «загрязняющих» продуктов усечения могут иметь один или более недостатков, таких как ограниченные чувствительность, простота, эффективность или действенность.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[3] В данном документе раскрыты полипептиды, содержащие удлиненный рекомбинантный полипептид (XTEN), который состоит из множества неперекрывающихся мотивов последовательности. В полипептидах XTEN по настоящему изобретению множество неперекрывающихся мотивов последовательности предусматривает набор неперекрывающихся мотивов последовательности, где каждый из указанных мотивов последовательности повторяется в полипептиде XTEN по меньшей мере дважды; а также уникальный неперекрывающийся мотив последовательности, который встречается в полипептиде XTEN только один раз; где полипептид дополнительно содержит первый фрагмент со штрих-кодом, высвобождаемый из полипептида при расщеплении протеазой. В указанных вариантах осуществления первый фрагмент со штрих-кодом представляет собой часть XTEN, которая включает по меньшей мере часть мотива последовательности, который встречается только один раз в XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой. Кроме того, в вариантах осуществления XTEN по настоящему изобретению, предусмотренных в данном документе, фрагмент со штрих-кодом не включает N-концевую аминокислоту или C-концевую аминокислоту полипептида. Как дополнительно раскрыто в данном документе, полипептиды XTEN по настоящему изобретению характеризуются тем, что содержат по меньшей мере 150 аминокислот, более конкретно 150-3000 аминокислот, в длину. Аминокислоты, входящие в состав полипептидов XTEN по настоящему изобретению, характеризуются тем, что по меньшей мере 90% этих остатков представляют собой глицин (G), аланин (A), серин (S), треонин (T), глутаминовую кислоту (E) или пролин (P), и полипептид XTEN содержит по меньшей мере четыре из этих аминокислот (G, A, S, T, E или P). Кроме того, полипептиды XTEN, предусмотренные в данном документе, содержат неперекрывающиеся мотивы последовательности, которые имеют длину последовательностей от 9 до 14 аминокислот, и в каждом из указанных неперекрывающихся мотивов последовательность аминокислот G, A, S, T, E или P фактически рандомизирована относительно любого другого неперекрывающегося мотива последовательности, входящего в состав полипептида XTEN.

[4] В некоторых вариантах осуществления фрагмент со штрих-кодом не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой кислоте в XTEN. В некоторых вариантах осуществления фрагмент со штрих-кодом содержит глутаминовую кислоту на C-конце. В некоторых вариантах осуществления фрагмент со штрих-кодом содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты. В некоторых вариантах осуществления остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает непосредственно к другому остатку глутаминовой кислоты. В некоторых вариантах осуществления фрагмент со штрих-кодом не включает остаток глутаминовой

кислоты в положении, отличном от С-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 аминокислот до 150 аминокислот либо от N-конца полипептида, либо от С-конца полипептида.

[5] В некоторых вариантах осуществления мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 182-203 и 1715-1722. В некоторых вариантах осуществления мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 186-189. В некоторых вариантах осуществления набор неперекрывающихся мотивов последовательности предусматривает по меньшей мере два, по меньшей мере три или все четыре мотива последовательности SEQ ID NO: 186-189.

[6] В конкретных вариантах осуществления полипептиды, предусмотренные в данном документе, предусматривают полипептид XTEN, раскрытый в данном документе, где фрагмент со штрих-кодом не включает N-концевую аминокислоту или С-концевую аминокислоту полипептида; не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой кислоте в XTEN; содержит глутаминовую кислоту на С-конце; содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты; и расположен на расстоянии от 10 аминокислот до 125 аминокислот либо от N-конца полипептида, либо от С-конца полипептида.

[7] В некоторых из этих конкретных вариантов осуществления остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает непосредственно к другому остатку глутаминовой кислоты. В некоторых из этих конкретных вариантов осуществления фрагмент со штрих-кодом не включает остаток глутаминовой кислоты в положении, отличном от С-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин.

[8] В некоторых вариантах осуществления полипептиды XTEN, предусмотренные в данном документе, содержат множество неперекрывающихся мотивов последовательности, где каждый указанный мотив последовательности повторяется в полипептиде XTEN по меньшей мере дважды и имеет длину от 9 до 14 аминокислот. В некоторых вариантах осуществления мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 182-203 и 1715-1722. В некоторых вариантах осуществления мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 186-189. В некоторых вариантах осуществления набор неперекрывающихся мотивов последовательности предусматривает по меньшей мере два, по меньшей мере три или все четыре мотива последовательности под SEQ ID NO: 186-189. В некоторых вариантах осуществления по меньшей мере 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотных остатков

полипептида XTEN представляют собой комбинацию глицина (G), аланина (A), серина (S), треонина (T), глутаминовой кислоты (E) или пролина (P), где полипептид XTEN содержит по меньшей мере четыре из этих аминокислот (G, A, S, T, E или P). В некоторых вариантах осуществления длина XTEN составляет от 150 до 3000 аминокислот. В некоторых вариантах осуществления длина XTEN составляет от 150 до 1000 аминокислот. В некоторых вариантах осуществления полипептид может быть расщеплен протеазой, которая расщепляет с С-концевой стороны от остатков глутаминовой кислоты, за которыми не следует пролин. В определенных вариантах осуществления протеаза представляет собой протеазу Glu-C.

[9] В некоторых вариантах осуществления полипептидов XTEN, предусмотренных в данном документе, фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от N-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от N-конца белка. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от С-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от С-конца белка. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет по меньшей мере 4 аминокислоты. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет от 4 до 20, от 5 до 15, от 6 до 12 или от 7 до 10 аминокислот. В некоторых вариантах осуществления фрагмент со штрих-кодом идентифицирован в данном документе под SEQ ID NO: 8020-8030 (BAR001-BAR011).

[10] В некоторых вариантах осуществления полипептид дополнительно содержит второй фрагмент со штрих-кодом, где второй фрагмент со штрих-кодом представляет собой часть XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой. В некоторых вариантах осуществления полипептид дополнительно содержит третий фрагмент со штрих-кодом, где третий фрагмент со штрих-кодом представляет собой часть XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

[11] В некоторых вариантах осуществления XTEN характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100% идентичностью последовательности с последовательностью, идентифицированной в данном документе под SEQ ID NO: 8001-8019. В некоторых вариантах осуществления длина XTEN составляет по меньшей мере 200, по меньшей мере 250, по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере

450 или по меньшей мере 500 аминокислот.

[12] В некоторых вариантах осуществления полипептид дополнительно содержит биологически активный полипептид, связанный с полипептидом XTEN (BPXTEN). В некоторых вариантах осуществления полипептид XTEN связан с биологически активным полипептидом на амино-конце или карбоксильном конце XTEN. В любой конфигурации фрагмент со штрих-кодом расположен в области XTEN, которая простирается, при измерении, от амино-конца или карбоксильного конца, связанного с биологически активным полипептидом, на от 5% до 50%, от 7% до 40% или от 10% до 30% длины XTEN.

[13] В некоторых вариантах осуществления полипептид BPXTEN дополнительно содержит один или более эталонных фрагментов, высвобождаемых из полипептида при расщеплении протеазой, где каждый из этих одного или более эталонных фрагментов содержит часть биологически активного полипептида. В некоторых вариантах осуществления один или более эталонных фрагментов представляют собой один эталонный фрагмент, который отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после расщепления полипептида протеазой. В некоторых вариантах осуществления указанный эталонный фрагмент содержит пептид, присутствие которого в смеси полипептидов указывает на его существование или целостность (*m. e.* на то, что белок не подвергся деградации или протеолитическому расщеплению).

[14] В некоторых вариантах осуществления полипептид BPXTEN дополнительно содержит первый сегмент для обеспечения высвобождения (RS1), расположенный между XTEN и биологически активным полипептидом. В некоторых вариантах осуществления RS1 содержит аминокислотную последовательность, характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью, идентифицированной в данном документе под любой из последовательностей из **таблиц 4a-4h**. В некоторых вариантах осуществления биологически активный полипептид идентифицирован в данном документе под любой одной или комбинацией последовательностей из **таблиц 4a-4h и 8a-8b**.

[15] В некоторых вариантах осуществления полипептид BPXTEN предпочтительно имеет конечный период полувыведения, который в по меньшей мере два раза дольше по сравнению с биологически активным полипептидом, не связанным с каким-либо XTEN.

[16] В некоторых вариантах осуществления полипептид BPXTEN предпочтительно является менее иммуногенным по сравнению с биологически активным полипептидом, не связанным с каким-либо XTEN, где иммуногенность можно установить путем измерения уровня продуцирования антител IgG, которые избирательно связываются с биологически активным полипептидом, после введения сравнимых доз человеку или животному.

[17] В некоторых вариантах осуществления полипептид BPXTEN демонстрирует коэффициент кажущейся молекулярной массы в физиологических условиях,

превышающий приблизительно 6.

[18] В некоторых вариантах осуществления полипептид BPXTEN дополнительно содержит второй полипептид XTEN, где второй полипептид XTEN содержит аминокислотную последовательность, имеющую такие же характеристики, которые указаны выше и во всем данном изобретении для первого компонента, представляющего собой XTEN, этих вариантов осуществления BPXTEN, и где первый полипептид XTEN расположен с N-конца биологически активного полипептида, а второй полипептид XTEN расположен с С-конца биологически активного полипептида. В некоторых вариантах осуществления второй полипептид XTEN содержит аминокислотную последовательность, которая отличается от аминокислотной последовательности первого XTEN, содержащего эти варианты осуществления BPXTEN. В некоторых вариантах осуществления аминокислотная последовательность второго полипептида XTEN длиннее, чем аминокислотная последовательность первого полипептида XTEN.

[19] В некоторых вариантах осуществления полипептид BPXTEN дополнительно содержит второй сегмент для обеспечения высвобождения (RS2), расположенный между биологически активным полипептидом и вторым полипептидом XTEN. В некоторых вариантах осуществления RS1 первого полипептида XTEN и RS2 второго полипептида XTEN идентичны по последовательности. В некоторых вариантах осуществления каждый из RS1 первого полипептида XTEN и RS2 второго полипептида XTEN представляет собой субстрат для расщепления несколькими протеазами в одном, или двух, или трех, или более сайтах расщепления в каждой последовательности сегмента для обеспечения высвобождения.

[20] В некоторых из этих вариантов осуществления полипептид BPXTEN содержит добавочный фрагмент со штрих-кодом, который представляет собой часть второго полипептида XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой. В некоторых из этих вариантов осуществления добавочный фрагмент со штрих-кодом не включает С-концевую аминокислоту полипептида. В некоторых из этих вариантов осуществления добавочный фрагмент со штрих-кодом содержит остаток глутаминовой кислоты на С-конце. В некоторых из этих вариантов осуществления добавочный фрагмент со штрих-кодом второго полипептида XTEN расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от С-конца второго компонента, представляющего собой XTEN, полипептида BPXTEN. В некоторых из этих вариантов осуществления добавочный фрагмент со штрих-кодом второго полипептида XTEN расположен в месте, которое находится на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от С-конца второго компонента, представляющего собой XTEN, полипептида BPXTEN. В некоторых из этих вариантов осуществления длина добавочного фрагмента со штрих-кодом составляет от 4 до 20, от 5 до 15, от 6 до 12 или от 7 до 10 аминокислот. В некоторых из этих вариантов осуществления добавочный фрагмент со

штрих-кодом идентифицирован в данном документе под SEQ ID NO: 8020-8030 (BAR001-BAR011).

[21] В некоторых вариантах осуществления второй полипептид XTEN дополнительно содержит набор фрагментов со штрих-кодом, который включает добавочный фрагмент со штрих-кодом и по меньшей мере один дополнительный фрагмент со штрих-кодом, где каждый фрагмент со штрих-кодом из набора фрагментов со штрих-кодом отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида BPXTEN после полного расщепления полипептида протеазой. В некоторых вариантах осуществления второй полипептид XTEN идентифицирован под SEQ ID NO: 8001-8019. В некоторых вариантах осуществления добавочный фрагмент со штрих-кодом не включает остаток глутаминовой кислоты, который непосредственно примыкает к другому остатку глутаминовой кислоты в полипептиде.

[22] В некоторых вариантах осуществления по меньшей мере 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотных остатков второго полипептида XTEN представляют собой комбинацию глицина (G), аланина (A), серина (S), треонина (T), глутаминовой кислоты (E) и пролина (P), где полипептид XTEN содержит по меньшей мере четыре из этих аминокислот (G, A, S, T, E или P). В некоторых вариантах осуществления сумма общего количества аминокислот в первом полипептиде XTEN и общего количества аминокислот во втором полипептиде XTEN составляет по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 500, по меньшей мере 600, по меньшей мере 700 или по меньшей мере 800 аминокислот. В некоторых вариантах осуществления второй полипептид XTEN содержит множество неперекрывающихся мотивов последовательности, где каждый из указанных мотивов последовательности повторяется по меньшей мере два раза в последовательности второго полипептида XTEN и имеет длину от 9 до 14 аминокислот.

[23] В некоторых вариантах осуществления для второго полипептида XTEN мотивы последовательности из множества неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 182-203 и 1715-1722. В некоторых вариантах осуществления мотивы последовательности из множества неперекрывающихся мотивов последовательности идентифицированы в данном документе под SEQ ID NO: 186-189. В некоторых вариантах осуществления для второго полипептида XTEN множество неперекрывающихся мотивов последовательности предусматривают по меньшей мере два, по меньшей мере три или все четыре из следующих мотивов: SEQ ID NO: 186-189. В некоторых вариантах осуществления длина второго полипептида XTEN составляет от 150 до 3000 аминокислот. В некоторых вариантах осуществления длина второго полипептида XTEN составляет от 150 до 1000 аминокислот. В некоторых вариантах осуществления второй полипептид XTEN характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100% идентичностью последовательности с

последовательностью, идентифицированной в данном документе под SEQ ID NO: 8001-8019. В некоторых вариантах осуществления длина второго полипептида XTEN составляет по меньшей мере 200, по меньшей мере 250, по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 450 или по меньшей мере 500 аминокислот.

[24] В конкретных вариантах осуществления полипептиды BPXTEN, предусмотренные в данном документе, содержат первый полипептид XTEN, содержащий последовательность первого RS, расположенную проксимально, но не включающую С-конец полипептида, ковалентно связанную с первым и вторым биологически активными полипептидами, ковалентно связанными тандемно, где второй полипептид XTEN ковалентно связан с С-концом тандемно связанных биологически активных полипептидов, где второй полипептид XTEN содержит последовательность второго RS, расположенную проксимально, но не включающую N-конец второго полипептида XTEN, где последовательности первого и второго RS могут быть одинаковыми или разными. В конкретных вариантах осуществления второй полипептид XTEN содержит аминокислотную последовательность, которая длиннее, чем аминокислотная последовательность первого полипептида XTEN. В определенных вариантах осуществления первый или второй биологически активный белок или оба содержат специфически связывающийся белок, где в определенных вариантах осуществления специфически связывающийся белок специфически связывается с антигеном или агонистом, экспрессируемым в необходимом биологическом сайте. В конкретных вариантах осуществления необходимый биологический сайт представляет собой опухоль, а антиген представляет собой опухолеспецифический антиген. В конкретных вариантах осуществления первый и второй биологически активные полипептиды отличаются друг от друга, в том числе без ограничения разными значениями аффинности специфического связывания.

[25] Кроме того, в данном документе раскрыта нуклеиновая кислота, содержащая полинуклеотид, кодирующий полипептид, такой как любой полипептид XTEN или BPXTEN, раскрытый в данном документе, или обратную последовательность, комплементарную указанному полинуклеотиду.

[26] Также в данном документе раскрыт вектор экспрессии, содержащий любую полинуклеотидную последовательность, раскрытую в данном документе, и регуляторную последовательность, функционально связанную с полинуклеотидной последовательностью, которая регулирует экспрессию или другую биологическую активность указанного полинуклеотида.

[27] В данном документе раскрыта клетка-хозяин, содержащая вектор экспрессии, раскрытый в данном документе. В некоторых вариантах осуществления клетка-хозяин является прокариотической. В некоторых из этих вариантов осуществления клетка-хозяин представляет собой E. coli. В некоторых альтернативных вариантах осуществления клетка-хозяин представляет собой клетку млекопитающего.

[28] Дополнительно, в данном документе раскрыта фармацевтическая композиция, содержащая полипептид, раскрытый в данном документе, и одно или более фармацевтически приемлемых вспомогательных веществ. В некоторых вариантах осуществления фармацевтическая композиция составлена для введения животному и, в частности, человеку, где указанное введение может быть осуществлено при помощи любого терапевтически эффективного пути введения. Фармацевтические композиции, раскрытые в данном документе, могут быть получены и использованы в любом составе, известном в уровне техники и особым образом адаптированном к пути введения, месту и предполагаемому действию на человека или животное.

[29] В данном документе раскрыто применение полипептида, раскрытого в данном документе, и, в частности, полипептида BPXTEN, при получении лекарственного препарата для лечения заболевания, нарушения или состояния у человека или животного. В некоторых вариантах осуществления заболевание, нарушение или состояние могут представлять собой рак.

[30] В данном документе раскрыт способ лечения заболевания у человека или животного, который раскрыт выше и во всем данном изобретении, при этом способ предусматривает введение человеку или животному, нуждающемуся в этом, одной или более терапевтически эффективных доз фармацевтической композиции. В некоторых вариантах осуществления фармацевтическую композицию вводят человеку или животному в виде одной или более терапевтически эффективных доз, вводимых по клинически приемлемому графику ежедневно, еженедельно, ежемесячно или ежегодно и в клинически приемлемой дозе.

[31] В данном документе раскрыта смесь, содержащая множество полипептидов, в частности полипептидов XTEN и BPXTEN, раскрытых в данном документе, различной длины, при этом смесь включает

первый набор полипептидов, где каждый полипептид из первого набора полипептидов содержит фрагмент со штрих-кодом, который высвобождается из полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов; и

второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов;

где каждый из как первого набора полипептидов, так и второго набора полипептидов содержит эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и получается при расщеплении протеазой; и

где отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,7.

[32] В некоторых вариантах осуществления отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8,

0,9, 0,95 или 0,98. В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде из первого набора полипептидов и второго набора полипептидов. В некоторых вариантах осуществления протеаза представляет собой протеазу, которая расщепляет с С-концевой стороны от остатков глутаминовой кислоты. В некоторых вариантах осуществления высвобождение штрих-кода из полипептидов, содержащих первый набор полипептидов, обеспечивается протеазами пепсином, эластазой, термолизином или Glu-C. В некоторых вариантах осуществления высвобождение штрих-кода обеспечивается протеазой Glu-C. В некоторых вариантах осуществления протеаза не является трипсином. В некоторых вариантах осуществления полипептиды с различными значениями длины предусматривают полипептиды, содержащие по меньшей мере один полипептид XTEN, который указан в данном документе.

[33] В некоторых вариантах осуществления первый набор полипептидов включает полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида. В некоторых вариантах осуществления полноразмерный полипептид представляет собой любой раскрытый в данном документе полипептид и, в частности, полипептиды XTEN и BPXTEN. В некоторых вариантах осуществления фрагмент со штрих-кодом не содержит ни N-концевую аминокислоту, ни С-концевую аминокислоту полноразмерного полипептида. В некоторых вариантах осуществления в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-, так и С-концевого усечения полноразмерного полипептида.

[34] В данном документе раскрыт способ оценки в смеси, содержащей полипептиды с различными значениями длины и, в частности, полипептиды XTEN и BPXTEN, раскрытые в данном документе, относительного количества первого набора полипептидов в смеси ко второму набору полипептидов в смеси, где каждый полипептид из первого набора полипептидов имеет общий фрагмент со штрих-кодом, который встречается один раз и только один раз в полипептиде, и каждый полипептид из второго набора полипептидов не имеет фрагмента со штрих-кодом, который является общим для полипептидов из первого набора, где каждый из отдельных полипептидов как из первого набора полипептидов, так и из второго набора полипептидов, содержит эталонный фрагмент, при этом способ предусматривает

приведение смеси в контакт с протеазой для получения множества протеолитических фрагментов, полученных в результате расщепления первого набора полипептидов и второго набора полипептидов, где множество протеолитических фрагментов включает множество эталонных фрагментов и множество фрагментов со штрих-кодом; и

определение отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов с оценкой тем самым относительных количеств первого набора полипептидов ко второму набору полипептидов.

[35] В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде из первого набора полипептидов и второго набора полипептидов.

[36] В некоторых вариантах осуществления протеаза расщепляет полипептиды различной длины с С-концевой стороны от остатков глутаминовой кислоты, за которыми не следует остаток пролина. В некоторых вариантах осуществления протеаза представляет собой протеазу Glu-C. В некоторых вариантах осуществления протеаза не является трипсином. В некоторых вариантах осуществления определение отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов предусматривает количественную оценку фрагментов со штрих-кодом и эталонных фрагментов в смеси после контакта смеси полипептидов с протеазой. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют на основе их соответствующих значений массы. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью масс-спектрометрии. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью жидкостной хроматографии-масс-спектрометрии (LC-MS). В некоторых вариантах осуществления определение отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает изобарическое мечение или мечение стабильным изотопом. В некоторых вариантах осуществления определение отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает введение в смесь одного или обоих из меченного изотопом эталонного фрагмента и меченного изотопом фрагмента со штрих-кодом.

[37] В некоторых из этих вариантов осуществления полипептиды различной длины предусматривают полноразмерный полипептид и его усеченные фрагменты. В некоторых из этих вариантов осуществления в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-, так и С-концевого усечения полноразмерного полипептида. В некоторых из этих вариантов осуществления отношение количества фрагментов со штрих-кодом к эталонным фрагментам составляет более 0,5, 0,6, 0,7, 0,8, 0,9, 0,95, 0,98 или 0,99.

[38] В данном документе раскрыта смесь, содержащая множество полипептидов различной длины, при этом смесь включает первый набор полипептидов, где каждый полипептид из первого набора полипептидов содержит фрагмент со штрих-кодом, который высвобождается из полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов. Указанный вариант осуществления также включает второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов, где как первый набор полипептидов, так и второй набор полипептидов каждый содержит эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и может высвобождаться при расщеплении

протеазой. В указанном варианте осуществления количество эталонных фрагментов, определенное количественно в смеси полипептидов после расщепления протеазой, равно сумме количества первого и второго набора полипептидов в смеси, и количество фрагментов со штрих-кодом, определенное количественно в смеси полипептидов после расщепления протеазой, равно количеству первого набора полипептидов в смеси. В указанном варианте осуществления первый набор полипептидов содержит один эталонный фрагмент, где отношение первого набора полипептидов к полипептидам в смеси, содержащей эталонный фрагмент, составляет более 0,7.

[39] В некоторых вариантах осуществления смесь характеризуется отношением первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, более 0,8, 0,9 или 0,95.

[40] В конкретном варианте осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов. В альтернативных вариантах осуществления эталонный фрагмент встречается дважды в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

[41] В некоторых вариантах осуществления первый набор полипептидов содержит полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида.

[42] В некоторых вариантах осуществления полноразмерный полипептид включает полипептиды, раскрытые в данном документе.

[43] В конкретном варианте осуществления фрагмент со штрих-кодом в смеси не содержит N-концевую аминокислоту и С-концевую аминокислоту полноразмерного полипептида.

[44] В некоторых вариантах осуществления смесь содержит полипептиды с различными значениями длины, которые отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-концевого, так и С-концевого усечения полноразмерного полипептида.

[45] В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов. В альтернативном варианте осуществления количество эталонных фрагментов в первом наборе полипептидов может отличаться от количества эталонных фрагментов во втором наборе полипептидов, но их количество в каждом полипептиде каждого набора должно быть одинаковым.

[46] В одном конкретном варианте осуществления каждый из эталонных фрагментов в полипептидах смеси имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов.

[47] В данном документе раскрыта смесь, содержащая множество полипептидов различной длины, при этом смесь предусматривает первый набор полипептидов, где

каждый полипептид первого набора полипептидов содержит

фрагмент со штрих-кодом, который может высвобождаться из полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, отличающиеся от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов. Смесь дополнительно содержит второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов, где каждый из как первого набора полипептидов, так и второго набора полипептидов содержит эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и может высвобождаться при расщеплении протеазой. Отношение первого набора полипептидов к полипептидам в смеси имеет формулу:

[полипептиды, содержащие штрих-код]/[(полипептиды, содержащие эталонные пептиды) x N],

где N представляет собой количество появлений эталонного пептида, который высвобождается из каждого полипептида в смеси, и где, когда первый набор полипептидов содержит один эталонный фрагмент, отношение

первого набора полипептидов к полипептидам в смеси, включающей эталонный фрагмент, составляет более 0,7.

[48] В конкретном варианте осуществления отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8, 0,9 или 0,95.

[49] В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

[50] В некоторых вариантах осуществления эталонный фрагмент встречается дважды в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

[51] В конкретном варианте осуществления первый набор полипептидов содержит полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида.

[52] В некоторых вариантах осуществления полноразмерный полипептид включает полипептиды, раскрытые в данном документе. В конкретном варианте осуществления фрагмент со штрих-кодом не содержит N-концевую аминокислоту и C-концевую аминокислоту полноразмерного полипептида.

[53] В некоторых вариантах осуществления в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, C-концевого усечения или как N-концевого, так и C-концевого усечения полноразмерного полипептида.

[54] В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора

полипептидов. В дополнительных вариантах осуществления количество эталонных фрагментов в первом наборе полипептидов может отличаться от количества эталонных фрагментов во втором наборе полипептидов, но их количество в каждом полипептиде каждого набора должно быть одинаковым. В некоторых вариантах осуществления эталонные фрагменты в полипептидах смеси имеют последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов.

[55] В данном документе раскрыт способ выявления целостности последовательности полипептидов, содержащих первый набор полипептидов в смеси, раскрытой в данном документе, при этом способ предусматривает стадии расщепления смеси полипептидов протеазой, которая высвобождает фрагмент со штрих-кодом и эталонный фрагмент из первого набора полипептидов и высвобождает эталонный фрагмент из второго набора полипептидов, и определения отношения фрагментов со штрих-кодом из первого набора полипептидов к эталонным фрагментам из первого и второго набора полипептидов. В конкретном варианте осуществления целостность последовательности полипептидов первого набора полипептидов выявляют путем сравнения соотношения фрагментов с ожидаемым соотношением фрагментов на основе количества фрагментов со штрих-кодом и эталонных фрагментов в полипептидах, содержащих первый и второй наборы полипептидов.

[56] Рассматриваемые в данном документе способы легко поддаются качественному и количественному анализам полипептидов, содержащих штрих-коды и/или эталонные фрагменты, например с использованием LC/MS. В одном конкретном варианте осуществления LC/MS является количественной и выявляет изотопно различимое количество фрагментов со штрих-кодом, эталонных фрагментов или того и другого. В примерах таких способов в смесь полипептидов добавляют известное количество «стандартного материала» для обеспечения такого анализа. Например, такой стандартный материал представляет собой материал, который предусматривает меченную изотопами версию указанной смеси множества полипептидов различной длины, подлежащей анализу. Этот меченный изотопами стандарт может быть добавлен к смеси в виде полной последовательности перед расщеплением указанной протеазой. В качестве альтернативы, испытываемый образец смеси полипептидов различной длины и меченного изотопами стандартного материала расщепляют протеазой в отдельных реакциях, а расщепленный протеазой меченный изотопом стандартный материал добавляют к испытываемому образцу перед анализом с помощью LC/MS. Способы по настоящему изобретению дополнительно предусматривают количественный анализ количества фрагментов со штрих-кодом, эталонных фрагментов или того и другого в испытываемом образце путем сравнения с количественным определением выявленных изотопно различимых количеств фрагментов со штрих-кодом, эталонных фрагментов или того и другого.

[57] Вариации и модификации этих вариантов осуществления придут на ум

специалистам в данной области после ознакомления с настоящим изобретением. Вышеупомянутые признаки и аспекты могут быть реализованы в любой комбинации и подкомбинациях (включая множественные зависимые комбинации и подкомбинации) с одним или более другими описанными в данном документе признаками. Различные признаки, описанные или проиллюстрированные выше, включая любые их компоненты, могут быть объединены или интегрированы в другие варианты осуществления. Более того, определенные признаки могут быть опущены или не реализованы.

ВКЛЮЧЕНИЕ ПОСРЕДСТВОМ ССЫЛКИ

[58] Все публикации, патенты и патентные заявки, упоминаемые в настоящем описании, включены в данный документ посредством ссылки в той же степени, как если бы было указано, что каждые отдельные публикация, патент или патентная заявка специально и отдельно включены посредством ссылки.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[59] Различные признаки настоящего изобретения конкретно изложены в прилагаемой формуле изобретения. Лучшее понимание признаков и преимуществ настоящего раскрытия может быть достигнуто при обращении к следующему подробному описанию, в котором изложены иллюстративные варианты осуществления, в которых используют принципы настоящего изобретения, и прилагаемым графическим материалам, в которых представлено следующее.

[60] На фигуре 1 показана смесь икстенилированных полипептидов активируемых протеазой привлекающих Т-клетки активаторов («XPAT»), содержащая полипептиды XTEN с различными значениями длины. Полноразмерный XPAT (вверху) содержит полипептид XTEN длиной 288 аминокислот на N-конце и полипептид XTEN длиной 864 аминокислоты на С-конце. Различные усечения могут происходить в XPAT в одном или обоих из N- и С-концевых полипептидов XTEN, например, во время ферментации, очистки или других стадий получения продукта. В то время как продукты с ограниченными усечениями (усечения рядом с частью полипептида XTEN, удаленные от активируемого протеазой привлекающего Т-клетки активатора, связанного с ним) могут функционировать аналогично полноразмерной конструкции, серьезные усечения (усечения ближе к части полипептида XTEN, расположенные проксимальнее к активируемому протеазой привлекающему Т-клетки активатору, связанному с ним) могут обладать существенно отличающимися от своих полноразмерных аналогов фармакологическими свойствами. Наличие усечений создает проблему для количественной оценки фармакологически эффективных и неэффективных вариантов в продукте XPAT. Как показано на фигуре 1 с использованием полноразмерного XPAT, каждый полипептид XTEN имеет проксимальный конец и дистальный конец, где проксимальный конец расположен по сравнению с дистальным концом ближе к биологически активному полипептиду (*например*, привлекающему Т-клетки активатору, цитокину, моноклональному антителу (mAb), фрагменту антитела или другому белку, который икстенилирован). В зависимости от ориентации связи проксимальные или

дистальные концы полипептида XTEN могут соответствовать N-концу или C-концу полипептида XTEN.

[61] На фигуре 2 показана смесь полипептидов XPAT, содержащих полипептиды XTEN со штрих-кодом с различными значениями длины. В полноразмерном XPAT (вверху) N-концевой полипептид XTEN длиной 288 аминокислот содержит три слитые с возможностью расщепления последовательности штрих-кода: «NA», «NB» и «NC» (от дистального конца к проксимальному концу), и C-концевой полипептид XTEN длиной 864 аминокислоты содержит три слитые с возможностью расщепления последовательности штрих-кода: «CC», «CB» и «CA» (от проксимального конца к дистальному концу). Каждый штрих-код расположен с указанием фармакологически значимой длины соответствующего полипептида XTEN. Например, второстепенные продукты XPAT с N-концевым усечением, лишенные штрих-кода «NA», но имеющие более проксимальные штрих-коды «NB» и «NC», могут проявлять фактически те же фармакологические свойства, что и полноразмерная конструкция. Напротив, основные продукты XPAT с N-концевым усечением, *например* лишенные всех трех штрих-кодов на N-конце, могут заметно отличаться по фармакологической активности от полноразмерной конструкции. Уникальную протеолитически расщепляемую последовательность идентифицируют по биологически активному полипептиду (в данном случае тандемные scFv, которые включают активную часть привлекающего Т-клетки активатора) XPAT. Благодаря своему присутствию во всех вариантах длины XPAT (включая полноразмерный XPAT, его минорные усечения и основные усечения) уникальную протеолитически расщепляемую последовательность можно использовать в качестве эталона для количественной оценки количеств различных продуктов усечения по отношению к общему количеству биологически активного белка.

[62] На фигуре 3 показана возможная конструкция полипептида XTEN со штрих-кодом, путем вставки последовательности, генерирующей штрих-код, в универсальный (или обычный) полипептид XTEN. Иллюстративный универсальный (или обычный) полипептид XTEN (вверху) содержит неперекрывающиеся 12-мерные мотивы в последовательности «BCDABDCDABDCBDCDABDCB», где мотивы последовательности «A», «B», «C» и «D» встречаются 3, 6, 5 и 7 раз, соответственно. Расщепление протеазой Glu-C иллюстративного универсального полипептида XTEN (верхняя панель) не дает уникальных пептидов, за исключением обоих концов («NT» и «CT»). Вставка генерирующей штрих-код последовательности «X» (например, уникальной 12-мерной) в полипептид XTEN приводит к уникальной протеолитически расщепляемой последовательности (или последовательности штрих-кода), которая в полипептиде XTEN больше нигде не встречается. Генерирующая штрих-код последовательность «X» может быть расположена там, где полученный штрих-код отмечает фармакологически релевантную длину полипептида XTEN. Например, полипептид XTEN без штрих-кода может функционально отличаться от соответствующего полипептида XTEN со штрих-кодом. Специалисту в данной области будет понятно, что генерирующая штрих-код

последовательность («X») может быть самой последовательностью штрих-кода. В качестве альтернативы, генерирующая штрих-код последовательность («X») может отличаться от получаемой последовательности штрих-кода. Например, последовательность штрих-кода может перекрываться и, таким образом, содержать часть предшествующего или последующего 12-мерного мотива.

[63] На фигурах 4A-4B показана количественная оценка уровня усечения для N-концевого полипептида XTEN. На фигуре 4A показано, что полипептид XTEN со штрих-кодом (нижняя панель) может быть сконструирован путем замены мотива последовательности в универсальном полипептиде XTEN (верхняя панель) (*например*, третьего мотива последовательности от N-конца, «D») на генерирующий штрих-код мотив «X»; и в этом примере генерирующий штрих-код мотив («X») сам по себе является уникальной протеолитически расщепляемой последовательностью штрих-кода. Как показано на нижней панели на фигуре 4A, штрих-код расположен там, где все формы полипептида XTEN с серьезным усечением не имеют штрих-кода, а все формы полипептида XTEN с ограниченным усечением содержат штрих-код. На фигуре 4B показано относительное количество различных продуктов расщепления в двух разных смесях XPAT. В одной из смесей штрих-код присутствует в 99% конструкций, содержащих биологически активный белок. В другой смеси 13% конструкций не имеют штрих-кода. На фигурах 4A-4B показано применение полипептида XTEN со штрих-кодом для различения двух смесей полипептидов, имеющих фактически одинаковые средние молекулярные массы, но заметно различающиеся по фармакологическим активностям.

[64] На фигуре 5A показана аналитическая эксклюзионная хроматография (SEC) белка XPAT и выявление полноразмерного белка и его усеченных производных. Фракция синтетический белок+усеченные фрагменты включает фрагменты размером с интактный синтетический белок.

[65] На фигуре 5B показано количество пептидов штрих-кода в препаратах XPAT, обнаруженное с помощью масс-спектрометрии. Каждое измерение представляет собой область XIC N-штрих-кода SGPGSTPAE (SEQ ID NO: 8029) и C-штрих-кода GSAPGTE (SEQ ID NO: 8023), нормализованную к пику 400 нМ соответствующего синтетического пептида, меченного тяжелым изотопом.

[66] Файл патента или заявки содержит по меньшей мере один рисунок, выполненный в цвете. Копии публикации этого патента или заявки на патент с цветными рисунками будут предоставлены Ведомством по запросу и после уплаты необходимой пошлины.

ТЕРМИНОЛОГИЯ

[67] Используемые в данном документе следующие термины имеют приписываемые им значения, если не указано иное.

[68] Используемая в данном описании и формуле изобретения форма единственного числа включает ссылки во множественном числе, если контекст явно не требует иного. Например, термин «клетка» включает множество клеток, в том числе их

смеси.

[69] Термины «полипептид», «пептид» и «белок» используют в данном документе взаимозаменяемо для обозначения полимеров из аминокислот любой длины. Полимер может быть линейным или разветвленным, он может содержать модифицированные аминокислоты, и он может прерываться элементами, отличными от аминокислот. Термины также охватывают полимер из аминокислот, который был модифицирован, например, путем образования дисульфидной связи, гликозилирования, липидирования, ацетилирования, фосфорилирования или любых других манипуляций, таких как конъюгация с метящим компонентом.

[70] Используемый в данном документе термин «аминокислота» относится к природным и/или неприродным либо синтетическим аминокислотам, в том числе без ограничения к глицину и обоим из оптических D- или L-изомеров, а также аналогам аминокислот и пептидомиметикам. Для обозначения аминокислот используют стандартные одно- или трехбуквенные коды.

[71] «Клетка-хозяин» включает в себя отдельную клетку или культуру клеток, которая может быть или была реципиентом для векторов человека или животного. Клетки-хозяева включают в себя потомство одной клетки-хозяина. Потомство не обязательно полностью идентично (по морфологическим характеристикам или полному геномному набору ДНК) исходной родительской клетке из-за естественной или генно-инженерной вариации.

[72] «Химерный» белок содержит по меньшей мере один полипептид, содержащий области, находящиеся в последовательности в положении, отличном от того, которое встречается в природе. Области обычно могут существовать в отдельных белках и собраны вместе в слитом полипептиде; или они могут обычно существовать в одном и том же белке, но располагаться в новом порядке расположения в слитом полипептиде. Указанные белки могут быть описаны как «конъюгированные», «связанные», «слитые» или «гибридные» белки; эти термины используют в данном документе взаимозаменяемо, и они относятся к соединению вместе двух или более полипептидных последовательностей любыми способами, включая химическую конъюгацию или рекомбинантные способы. Химерный белок может быть создан, например, путем химического синтеза или путем создания и трансляции полинуклеотида, в котором пептидные области кодируются в необходимом взаимном расположении.

[73] Термины «полинуклеотиды», «нуклеиновые кислоты», «нуклеотиды» и «олигонуклеотиды» используют взаимозаменяемо, и они относятся к полимерной форме нуклеотидов любой длины, либо дезоксирибонуклеотидов, либо рибонуклеотидов, либо их аналогов. Полинуклеотиды могут иметь любую трехмерную структуру и могут выполнять любую функцию, известную или подлежащую открытию или разработке. Полинуклеотид может содержать модифицированные нуклеотиды, такие как метилированные нуклеотиды и аналоги нуклеотидов. Модификации структуры нуклеотидов в случае их наличия могут быть внесены до или после сборки полимера.

Последовательность нуклеотидов может прерываться ненуклеотидными компонентами. Полинуклеотид может быть дополнительно модифицирован после полимеризации, как, например, путем конъюгации с метящим компонентом.

[74] Термин «последовательность, комплементарная полинуклеотиду» обозначает молекулу полинуклеотида, имеющую комплементарную последовательность оснований и обратную ориентацию по сравнению с эталонной последовательностью, где она может гибридизироваться с эталонной последовательностью с полной точностью.

[75] Используемые в данном документе полинуклеотиды, характеризующиеся «гомологией» или являющиеся «гомологичными», предпочтительно являются полинуклеотидами, которые гибридизируются в жестких условиях, определенных в данном документе, и характеризуются по меньшей мере 70%, предпочтительно по меньшей мере 80%, более предпочтительно по меньшей мере 90%, более предпочтительно 95%, более предпочтительно 97%, более предпочтительно 98% и еще более предпочтительно 99% идентичностью последовательностей по отношению к этим последовательностям.

[76] Термины «процент идентичности» и «% идентичности», применительно к полинуклеотидным последовательностям, относятся к процентной доле совпадающих остатков между по меньшей мере двумя полинуклеотидными последовательностями, выровненными с использованием стандартизированного алгоритма. С помощью такого алгоритма можно стандартизированным и воспроизводимым образом вставлять гэпы в сравниваемые последовательности для оптимизации выравнивания между двумя последовательностями и, следовательно, достижения более значимого сравнения двух последовательностей. Процент идентичности можно измерить по длине всей определенной полинуклеотидной последовательности, например, как определено конкретным номером SEQ ID, или можно измерить по меньшей длине, например, по длине фрагмента, взятого из большей определенной полинуклеотидной последовательности, например, фрагмента из по меньшей мере 45, по меньшей мере 60, по меньшей мере 90, по меньшей мере 120, по меньшей мере 150, по меньшей мере 210 или по меньшей мере 450 смежных остатков. Такие значения длины являются исключительно иллюстративными, и следует понимать, что любую длину фрагмента, подкрепленная последовательностями, показанными в данном документе, в таблицах, графических материалах или в перечне последовательностей, можно использовать для описания длины, по которой может быть измерено процентное значение идентичности.

[77] «Процент (%) идентичности аминокислотной последовательности» по отношению к полипептидным последовательностям, идентифицированным в данном документе, идентифицируют как процентную долю аминокислотных остатков в запрашиваемой последовательности, которые идентичны аминокислотным остаткам второй эталонной полипептидной последовательности или ее части после выравнивания последовательностей и введения гэпов, если это необходимо, для достижения максимального процента идентичности последовательностей и без учета каких-либо

консервативных замен как части идентичности последовательности. Выравнивание для целей определения процента идентичности аминокислотных последовательностей может быть достигнуто различными способами, которые находятся в пределах компетенции специалиста в данной области, например, с использованием общедоступного компьютерного программного обеспечения, такого как программное обеспечение BLAST, BLAST-2, ALIGN или Megalign (DNASTAR). Специалисты в данной области могут определить подходящие параметры для измерения выравнивания, включая любые алгоритмы, необходимые для достижения максимального выравнивания по всей длине сравниваемых последовательностей. Процент идентичности можно измерить по длине всей определенной полинуклеотидной последовательности, например, как определено конкретным номером SEQ ID, или можно измерить по меньшей длине, например, по длине фрагмента, взятого из большей определенной полипептидной последовательности, например, фрагмента из по меньшей мере 15, по меньшей мере 20, по меньшей мере 30, по меньшей мере 40, по меньшей мере 50, по меньшей мере 70 или по меньшей мере 150 смежных остатков. Такие значения длины являются исключительно иллюстративными, и следует понимать, что любую длину фрагмента, подкрепленная последовательностями, показанными в данном документе, в таблицах, графических материалах или в перечне последовательностей, можно использовать для описания длины, по которой может быть измерено процентное значение идентичности.

[78] Используемый в данном документе термин «повторяемость» аминокислотной последовательности полипептида XTEN относится к 3-мерной повторяемости и может быть измерена с помощью компьютерных программ или алгоритмов или другими способами, известными в уровне техники. 3-мерную повторяемость аминокислотной последовательности полипептида XTEN можно оценить путем определения числа появлений перекрывающихся 3-мерных последовательностей в полипептиде. Например, полипептид из 200 аминокислотных остатков имеет 198 перекрывающихся 3-аминокислотных последовательностей (3-меров), но количество уникальных 3-мерных последовательностей зависит от степени повторяемости в последовательности. Может быть получена оценка (в дальнейшем «оценка подпоследовательности»), которая отражает степень повторяемости 3-меров в общей полипептидной последовательности. В контексте настоящего изобретения «оценка подпоследовательности» означает сумму появлений каждой уникальной 3-мерной группы в последовательности полипептида из 200 последовательных аминокислот, деленную на абсолютное количество уникальных 3-мерных подпоследовательностей в пределах последовательности из 200 аминокислот. Примеры таких оценок подпоследовательностей, полученных из первых 200 аминокислот повторяющихся и неповторяющихся полипептидов, представлены в примере 73 публикации международной патентной заявки № WO 2010/091122 A1, которая включена посредством ссылки во всей своей полноте. В некоторых вариантах осуществления настоящее изобретение предусматривает полипептиды BPXTEN, каждый из которых содержит по меньшей мере один полипептид XTEN, в котором аминокислотная

последовательность полипептида XTEN может иметь оценку подпоследовательности менее 16, или менее 14, или менее 12, или, более предпочтительно, менее 10.

[79] Термин «практически неповторяющаяся аминокислотная последовательность полипептида XTEN», используемый в данном документе, относится к полипептиду XTEN, где в аминокислотной последовательности полипептида XTEN имеется мало или совсем нет случаев наличия четырех смежных аминокислот, которые являются идентичными по аминокислотным типам, и где аминокислотная последовательность полипептида XTEN имеет оценку подпоследовательности (определенную в предыдущем абзаце данного документа) 12 или 10 или менее, или что в порядке от N-конца к С-концу мотивов последовательности отсутствует паттерн, который составляет эту полипептидную последовательность.

[80] Как указано в данном документе, термин «неперекрывающиеся мотивы последовательности» включает мотивы последовательности, которые полностью не перекрываются, а также мотивы последовательности, которые не перекрываются лишь частично, при условии, что указанные частично неперекрывающиеся мотивы последовательности не перекрываются полностью.

[81] «Вектор» представляет собой молекулу нуклеиновой кислоты, предпочтительно самореплицирующуюся в подходящем хозяине, которая переносит встроенную молекулу нуклеиновой кислоты в клетки-хозяева и/или между клетками-хозяевами. Термин включает в себя векторы, которые функционируют в первую очередь для встраивания ДНК или РНК в клетку, реплицирующиеся векторы, которые функционируют в первую очередь для репликации ДНК или РНК, и векторы экспрессии, которые функционируют для транскрипции и/или трансляции ДНК или РНК. Также включены векторы, которые обеспечивают боле чем одну из вышеуказанных функций. «Вектор экспрессии» представляет собой полинуклеотид, который при введении в соответствующую клетку-хозяина может транскрибироваться и транслироваться в полипептид(-ы). «Система экспрессии» обычно означает подходящую клетку-хозяина, содержащую вектор экспрессии, который может функционировать для получения необходимого продукта экспрессии.

[82] Используемый в данном документе термин «$t_{1/2}$» означает конечный период полувыведения, рассчитываемый как $\ln(2)/K_{el}$. K_{el} представляет собой константу скорости конечного выведения, рассчитанную путем линейной регрессии конечного линейного участка кривой зависимости log концентрации от времени. Период полужизни обычно относится ко времени, требующемуся для того, чтобы половина количества введенного вещества, откладывающегося в живом организме, метаболизировалась или была выведена в ходе нормальных биологических процессов. Термины «$t_{1/2}$», «конечный период полувыведения», «период полувыведения» и «период полужизни в кровотоке» используют в данном документе взаимозаменяемо.

[83] Термины «антиген», «антиген-мишень» или «иммуноген» используют в данном документе взаимозаменяемо для обозначения структуры или детерминанты

связывания, с которой фрагмент антитела или терапевтическое средство на основе фрагмента антитела связывается или в отношении которой обладает специфичностью.

[84] Используемый в данном документе термин «полезная нагрузка» относится к последовательности белка или пептида, обладающей биологической или терапевтической активностью; аналогу фармакофора малых молекул. Примеры полезной нагрузки включают без ограничения цитокины, ферменты, гормоны и факторы крови и роста. Полезные нагрузки могут дополнительно содержать генетически слитые или химически конъюгированные фрагменты, такие как химиотерапевтические средства, противовирусные соединения, токсины или контрастные вещества. Эти конъюгированные фрагменты могут быть присоединены к остальной части полипептида через линкер, который может быть расщепляемым или нерасщепляемым.

[85] Используемые в данном документе термины «лечение» или «проведение лечения», «смягчение» и «облегчение состояния» используют в данном документе взаимозаменяемо, и они относятся к подходу для получения полезных или необходимых результатов, в том числе без ограничения терапевтической пользы и/или профилактической пользы. Под «терапевтической пользой» подразумевают устранение или облегчение первопричинного нарушения, подлежащего лечению. Кроме того, терапевтической пользы достигают при устранении или облегчении одного или более физиологических симптомов, ассоциированных с первопричинным болезненным состоянием, где у человека или животного наблюдают улучшение, несмотря на то, что человек или животное все еще может страдать от первопричинного нарушения. С профилактической целью композиции можно вводить человеку или животному с риском развития определенного болезненного состояния или человеку или животному, сообщающему об одном или более физиологических симптомах заболевания, даже если диагностирование этого заболевания могло быть не произведено.

[86] Используемый в данном документе «терапевтический эффект» относится к физиологическому эффекту, в том числе без ограничения излечению, смягчению, облегчению или предупреждению болезненного состояния у людей или других животных или иному улучшению физического или психического благополучия людей или животных, вызванному слитым полипептидом по настоящему изобретению, кроме способности индуцировать продуцирование антитела к антигенному эпитопу, которым обладает биологически активный белок. Определение терапевтически эффективного количества вполне находится в пределах компетенции специалистов в данной области, особенно в свете подробного раскрытия, представленного в данном документе.

[87] Термины «терапевтически эффективное количество» и «терапевтически эффективная доза», используемые в данном документе, относятся к количеству биологически активного белка, либо отдельно, либо в качестве части композиции на основе слитого белка, которое способно оказывать любой выявляемый благоприятный эффект в отношении любых симптома, аспекта, измеряемого параметра или характеристик болезненного состояния или обусловленного заболеванием условия при

введении человеку или животному в одной дозе или повторных дозах. Такой эффект не обязательно должен быть абсолютным, чтобы быть благоприятным. Болезненное состояние может относиться к нарушению или заболеванию.

[88] Термин «терапевтически эффективный режим введения дозы», используемый в данном документе, относится к схеме последовательного введения доз биологически активного белка либо отдельно, либо в качестве части композиции на основе слитого белка, где дозы вводят в терапевтически эффективных количествах для обеспечения устойчивого благоприятного эффекта в отношении любых симптома, аспекта, измеряемого параметра или характеристик болезненного состояния или обусловленного заболеванием условия.

СЛИТЫЙ ПОЛИПЕПТИД

[89] В данном документе раскрыты полипептиды, содержащие один или более удлиненных рекомбинантных полипептидов (XTEN) (которые более подробно описаны ниже), которые могут быть слиты или иным образом конъюгированы с другим полипептидом, в частности биологически активным полипептидом, где указанные варианты осуществления обозначены в данном документе как BPXTEN.

[90] В некоторых вариантах осуществления полипептид содержит первый полипептид XTEN (такой как полипептид, описанный ниже в разделе «УДЛИНЕННЫЙ РЕКОМБИНАНТНЫЙ ПОЛИПЕПТИД (XTEN)» или описанный где-либо еще в данном документе). В некоторых вариантах осуществления полипептид дополнительно содержит второй полипептид XTEN (такой как полипептид, описанный ниже в разделе «УДЛИНЕННЫЙ РЕКОМБИНАНТНЫЙ ПОЛИПЕПТИД (XTEN)» или описанный где-либо еще в данном документе). В некоторых вариантах осуществления полипептид содержит полипептид XTEN на N-конце или вблизи него («N-концевой XTEN»). В некоторых вариантах осуществления полипептид содержит полипептид XTEN на C-конце или вблизи него («C-концевой XTEN»). В некоторых вариантах осуществления полипептид содержит как N-концевой полипептид XTEN, так и C-концевой полипептид XTEN. В некоторых вариантах осуществления первый полипептид XTEN представляет собой N-концевой полипептид XTEN, а второй полипептид XTEN представляет собой C-концевой полипептид XTEN.

[91] Полипептид может дополнительно содержать биологически активный полипептид («BP»), связанный с полипептидом XTEN, образуя тем самым слитый полипептид, содержащий XTEN, называемый в данном документе полипептидом «BPXTEN».

[92] Полипептид XTEN может содержать один или более фрагментов со штрих-кодом (которые более подробно описано ниже), высвобождаемых (сконфигурированных для высвобождения) из полипептида XTEN при расщеплении слитого полипептида (или BPXTEN) протеазой. В некоторых вариантах осуществления каждый фрагмент со штрих-кодом отличается по последовательности и молекулярной массе от всех других пептидных фрагментов (включая все другие фрагменты со штрих-кодом, если они

присутствуют), которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

[93] (Слитый) полипептид может содержать один или более эталонных фрагментов (которые более подробно описано ниже), высвобождаемых (сконфигурированных для высвобождения) из полипептида, например при расщеплении протеазой, которая высвобождает фрагмент(-ы) со штрих-кодом из полипептида. В некоторых вариантах осуществления каждый эталонный фрагмент может быть единственным эталонным фрагментом, отличающимся по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после расщепления полипептида протеазой.

Удлиненный рекомбинантный полипептид (XTEN)

Длина цепи и аминокислотный состав

[94] В некоторых вариантах осуществления полипептид XTEN содержит по меньшей мере 150 аминокислот. В некоторых вариантах осуществления полипептида XTEN составляет от 150 до 3000 аминокислот в длину или от 150 до 1000 аминокислот в длину, или по меньшей мере 200, по меньшей мере 250, по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 450 или по меньшей мере 500 аминокислот в длину. В некоторых вариантах осуществления по меньшей мере 90% аминокислотных остатков полипептида XTEN представляют собой глицин (G), аланин (A), серин (S), треонин (T), глутаминовую кислоту (E) или пролин (P). В некоторых вариантах осуществления по меньшей мере 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотных остатков полипептида XTEN выбраны из G, A, S, T, E или P. В некоторых вариантах осуществления полипептид XTEN содержит по меньшей мере 4 разных типа аминокислот G, A, S, T, E или P. В некоторых вариантах осуществления полипептид XTEN отличается тем, что он содержит по меньшей мере 150 аминокислот; по меньшей мере 90% аминокислотных остатков полипептида XTEN представляют собой G, A, S, T, E или P, и он содержит по меньшей мере 4 разных типа аминокислот, выбранных из G, A, S, T, E и P, и фактически рандомизирован относительно любого другого неперекрывающегося мотива последовательности, входящего в состав полипептида XTEN. В некоторых вариантах осуществления содержащий XTEN слитый полипептид (*например*, слитый полипептид, содержащий конъюгированный с ним биологически активный полипептид) содержит первый полипептид XTEN и второй полипептид XTEN. В некоторых вариантах осуществления сумма общего количества аминокислот в первом XTEN и общего количества аминокислот во втором полипептиде XTEN составляет по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 500, по меньшей мере 600, по меньшей мере 700 или по меньшей мере 800 аминокислот.

Неперекрывающийся мотив последовательности

[95] В некоторых вариантах осуществления полипептиды XTEN, предусмотренные в данном документе, содержат множество неперекрывающихся мотивов последовательности или образованы из них. В некоторых вариантах осуществления по

меньшей мере один из неперекрывающихся мотивов последовательности является повторяющимся (или повторяется по меньшей мере два раза в XTEN), и где по меньшей мере другой из неперекрывающихся мотивов последовательности является неповторяющимся (или обнаруживается только один раз в XTEN) В некоторых вариантах осуществления множество неперекрывающихся мотивов последовательности содержит набор (повторяющихся) неперекрывающихся мотивов последовательности, где каждый из указанных мотивов последовательности повторяется по меньшей мере два раза в XTEN; и неперекрывающийся (неповторяющийся) мотив последовательности, который встречается (или обнаруживается) только один раз в XTEN. В некоторых вариантах осуществления длина каждого неперекрывающегося мотива последовательности составляет от 9 до 14 (или от 10 до 14, или от 11 до 13) аминокислот. В некоторых вариантах осуществления длина каждого неперекрывающегося мотива последовательности составляет 12 аминокислот. В некоторых вариантах осуществления множество неперекрывающихся мотивов последовательности содержит набор неперекрывающихся (повторяющихся) мотивов последовательности, где каждый из указанных мотивов последовательности повторяется по меньшей мере два раза в XTEN и имеет длину от 9 до 14 аминокислот. В некоторых вариантах осуществления набор (повторяющихся) неперекрывающихся мотивов последовательности включает 12-мерные мотивы последовательности, идентифицированные в данном документе под **SEQ ID NO: 182-203** и **1715-1722** в **таблице 1**. В некоторых вариантах осуществления набор (повторяющихся) неперекрывающихся мотивов последовательности включает 12-мерные мотивы последовательности, идентифицированные в данном документе под **SEQ ID NO: 186-189** в **таблице 1**. В некоторых вариантах осуществления набор (повторяющихся) неперекрывающихся мотивов последовательности предусматривает по меньшей мере два, по меньшей мере три или все четыре 12-мерных мотива последовательности под **SEQ ID NO: 186-189** в **таблице 1**.

Таблица 1. Иллюстративные 12-мерные мотивы последовательности для конструкции XTEN

Семейство мотива*	**SEQ ID NO**	**Аминокислотная последовательность**
AD	182	GESPGGSSGSES
AD	183	GSEGSSGPGESS
AD	184	GSSESGSSEGGP
AD	185	GSGGEPSESGSS
AE, AM	186	GSPAGSPTSTEE
AE, AM, AQ	187	GSEPATSGSETP
AE, AM, AQ	188	GTSESATPESGP
AE, AM, AQ	189	GTSTEPSEGSAP
AF, AM	190	GSTSESPSGTAP
AF, AM	191	GTSTPESGSASP
AF, AM	192	GTSPSGESSTAP
AF, AM	193	GSTSSTAESPGP
AG, AM	194	GTPGSGTASSSP
AG, AM	195	GSSTPSGATGSP

Семейство мотива*	SEQ ID NO	Аминокислотная последовательность
AG, AM	196	GSSPSASTGTGP
AG, AM	197	GASPGTSSTGSP
AQ	198	GEPAGSPTSTSE
AQ	199	GTGEPSSTPASE
AQ	200	GSGPSTESAPTE
AQ	201	GSETPSGPSETA
AQ	202	GPSETSTSEPGA
AQ	203	GSPSEPTEGTSA
BC	1715	GSGASEPTSTEP
BC	1716	GSEPATSGTEPS
BC	1717	GTSEPSTSEPGA
BC	1718	GTSTEPSEPGSA
BD	1719	GSTAGSETSTEA
BD	1720	GSETATSGSETA
BD	1721	GTSESATSESGA
BD	1722	GTSTEASEGSAS

*Обозначает отдельные последовательности мотивов, которые при совместном использовании в различных перестановках дают в результате «семейство последовательности».

Фрагмент со штрих-кодом

[96] В некоторых вариантах осуществления предусмотренные в данном документе полипептиды содержат фрагмент со штрих-кодом (например, первый, второй или третий фрагмент со штрих-кодом полипептида XTEN), высвобождаемый из полипептида при расщеплении протеазой. В некоторых вариантах осуществления фрагмент со штрих-кодом представляет собой часть XTEN, которая включает по меньшей мере часть (неповторяющегося, неперекрывающегося) мотива последовательности, который встречается (или обнаруживается) только один раз в XTEN; и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой. Специалисту в данной области будет понятно, что термин «фрагмент со штрих-кодом» (или «штрих-код», или «последовательность штрих-кода») может относиться либо к части XTEN, идентифицированной в данном документе, слитой с возможностью расщепления внутри полипептида, либо к полученному пептидному фрагменту, высвобождаемому из полипептида.

[97] В некоторых вариантах осуществления фрагмент со штрих-кодом не включает в себя N-концевую аминокислоту или С-концевую аминокислоту полипептида XTEN. Как более подробно описано ниже или описано где-либо в данном документе, в некоторых вариантах осуществления фрагмент со штрих-кодом может высвобождаться (сконфигурирован для высвобождения) при расщеплении слитого полипептида с помощью Glu-C. В некоторых вариантах осуществления фрагмент со штрих-кодом не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой кислоте в полипептиде XTEN. В некоторых вариантах осуществления

фрагмент со штрих-кодом содержит глутаминовую кислоту на С-конце. Специалисту в данной области будет понятно, что С-конец фрагмента со штрих-кодом может относиться к «последнему» (или самому С-концевому) аминокислотному остатку во фрагменте со штрих-кодом при слиянии с возможностью расщепления с полипептидом XTEN, даже если другие аминокислотные остатки «без штрих-кода» расположены с С-конца фрагмента со штрих-кодом в том же полипептиде XTEN. В некоторых вариантах осуществления фрагмент со штрих-кодом содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты. В некоторых вариантах осуществления остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает непосредственно к другому остатку глутаминовой кислоты. В некоторых вариантах осуществления фрагмент со штрих-кодом не включает остаток глутаминовой кислоты в положении, отличном от С-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 до 150 или от 10 до 125 аминокислот либо от N-конца полипептида, либо от С-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен в пределах или на расстоянии 300, 280, 260, 250, 240, 220, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 48, 40, 36, 30, 24, 20, 12 или 10 аминокислот от N-конца полипептида или в месте в диапазоне между любыми из вышеперечисленных. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от N-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от N-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен в пределах или на расстоянии 300, 280, 260, 250, 240, 220, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 48, 40, 36, 30, 24, 20, 12 или 10 аминокислот от С-конца полипептида или в месте в диапазоне между любыми из вышеперечисленных. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от С-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от С-конца полипептида. В некоторых вариантах осуществления фрагмент со штрих-кодом не включает N-концевую аминокислоту или С-концевую аминокислоту полипептида; не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой кислоте в XTEN; содержит глутаминовую кислоту на С-конце; содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты; и (v) расположен на расстоянии от 10 до 150 или от 10 до 125 аминокислот либо от N-конца полипептида, либо от С-конца полипептида. В некоторых вариантах осуществления остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает

непосредственно к другому остатку глутаминовой кислоты. В некоторых вариантах осуществления фрагмент со штрих-кодом не включает остаток глутаминовой кислоты в положении, отличном от С-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин. В некоторых вариантах осуществления для полипептида XTEN со штрих-кодом, слитого с биологически активным полипептидом, по меньшей мере один фрагмент со штрих-кодом (или по меньшей мере два фрагмента со штрих-кодом или три фрагмента со штрих-кодом), содержащийся в XTEN со штрих-кодом, расположен на расстоянии по меньшей мере 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300 аминокислот от биологически активного полипептида. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет по меньшей мере 4, по меньшей мере 5, по меньшей мере 6, по меньшей мере 7 или по меньшей мере 8 аминокислот. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет по меньшей мере 4 аминокислоты. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 аминокислот или в диапазоне между любыми из вышеперечисленных значений. В некоторых вариантах осуществления длина фрагмента со штрих-кодом составляет от 4 до 20, от 5 до 15, от 6 до 12 или от 7 до 10 аминокислот. В некоторых вариантах осуществления фрагмент со штрих-кодом выбран из **SEQ ID NO: 8020-8030 (BAR001-BAR011)** в **таблице 2**.

Таблица 2. Иллюстративные фрагменты со штрих-кодом, высвобождаемые при расщеплении с помощью Glu-C

Аминокислотная последовательность		**SEQ ID NO:**
SPATSGSTPE	BAR001	8020
GSAPATSE	BAR002	8021
GSAPGTATE	BAR003	8022
GSAPGTE	BAR004	8023
PATSGPTE	BAR005	8024
SASPE	BAR006	8025
PATSGSTE	BAR007	8026
GSAPGTSAE	BAR008	8027
SATSGSE	BAR009	8028
SGPGSTPAE	BAR010	8029
SGSE	BAR011	8030

[98] В некоторых вариантах осуществления полипептид XTEN со штрих-кодом содержит только один фрагмент со штрих-кодом. В некоторых вариантах осуществления полипептид XTEN со штрих-кодом содержит набор фрагментов со штрих-кодом, включающий первый фрагмент со штрих-кодом, такой как описанный выше или где-либо еще в данном документе. В этих вариантах осуществления каждый элемент набора последовательностей штрих-кода можно отличить от всех других последовательностей штрих-кода на основе аминокислотной последовательности или молекулярной массы (где эти способы различения разных последовательностей штрих-кода будут связаны). В

некоторых вариантах осуществления набор фрагментов со штрих-кодом содержит второй фрагмент со штрих-кодом (или добавочный фрагмент со штрих-кодом), такой как описанный выше или где-либо еще в данном документе. В некоторых вариантах осуществления набор фрагментов со штрих-кодом содержит третий фрагмент со штрих-кодом, такой как описанный выше или где-либо еще в данном документе. Набор фрагментов со штрих-кодом, слитых с N-концевым полипептидом XTEN, может называться N-концевым набором штрих-кодов («N-концевой набор»). Набор фрагментов со штрих-кодом, слитых с C-концевым полипептидом XTEN, может называться C-концевым набором штрих-кодов («C-концевой набор»). В некоторых вариантах осуществления N-концевой набор содержит первый фрагмент со штрих-кодом и второй фрагмент со штрих-кодом. В некоторых вариантах осуществления N-концевой набор дополнительно содержит третий фрагмент со штрих-кодом. В некоторых вариантах осуществления C-концевой набор содержит первый фрагмент со штрих-кодом и второй фрагмент со штрих-кодом. В некоторых вариантах осуществления C-концевой набор дополнительно содержит третий фрагмент со штрих-кодом. В некоторых вариантах осуществления второй фрагмент со штрих-кодом расположен с N-конца от первого фрагмента со штрих-кодом из того же набора. В некоторых вариантах осуществления второй фрагмент со штрих-кодом расположен с C-конца от первого фрагмента со штрих-кодом из того же набора. В некоторых вариантах осуществления третий фрагмент со штрих-кодом расположен с N-конца как от первого, так и от второго фрагмента со штрих-кодом. В некоторых вариантах осуществления третий фрагмент со штрих-кодом расположен с C-конца как от первого, так и от второго фрагмента со штрих-кодом. В некоторых вариантах осуществления третий фрагмент со штрих-кодом расположен между первым и вторым фрагментами со штрих-кодом. В некоторых вариантах осуществления полипептид содержит набор фрагментов со штрих-кодом, который включает первый фрагмент со штрих-кодом, добавочный (второй) фрагмент со штрих-кодом и по меньшей мере один дополнительный фрагмент со штрих-кодом, где каждый фрагмент со штрих-кодом из набора фрагментов со штрих-кодом представляет собой часть второго полипептида XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

Иллюстративный XTEN со штрих-кодом

[99] Аминокислотные последовательности 13 иллюстративных XTEN со штрих-кодом, содержащих один штрих-код (*например*, SEQ ID NO: 8002-8003, 8005-8009 и 8013), или два штрих-кода (например, SEQ ID NO: 8001, 8004, 8010 и 8012), или три штрих-кода (например, SEQ ID NO: 8011), показаны в **таблице 3а**. Среди этих 13 иллюстративных полипептидов XTEN со штрих-кодом шесть (SEQ ID NO: 8001-8003, 8008-8009 и 8011) могут быть слиты с биологически активным белком на C-конце биологически активного белка, а семь (SEQ ID NO: 8004-8007, 8010 и 8012-8013) могут быть слиты на N-конце биологически активного белка. В некоторых вариантах осуществления полипептид XTEN

характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100% идентичностью последовательности по отношению к последовательности, выбранной из **SEQ ID NO: 8001-8019** в **таблице 3а**.

Таблица 3а. Иллюстративные XTEN со штрих-кодом

SEQ ID NO:	**Тип XTEN**	**№ штрих-кода(-ов)**	**Аминокислотная последовательность**	**Общее количество амино кислот**
8001	C-концевой XTEN	2	PGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGftab TSESATPESGPGSE**PATSGPTE**SGSEPATSG SETPGSPAGSPTSTEEGTSTEPSE**GSAPGTE**	864

			STPSEGSAPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGEPEA	
8002	C-концевой XTEN	1	PGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSE**PATSGPTE**SGSEPATSGSE TPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGEPEA	864
8003	C-концевой XTEN	1	PGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS	864

			PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSE**GSAPGTE**STP SEGSAPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGEPEA	
8004	N-концевой XTEN	2	ASSPAGSPTSTESGTSESATPESGPGTETEPS EGSAPGTSESATPESGPGSEPATSGSETPGT SESATPE**SGPGSTPAE**SGSETPGTSESATPE SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGE**SPATSGSTPE**GTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP	288
8005	N-концевой	1	ASSPAGSPTSTESGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGT	288

	XTEN		SESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGE**SPATSGSTPE**GTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP	
8006	N-концевой XTEN	1	ASSPAGSPTSTESGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGT SESATPE**SGPGSTPAE**SGSETPGTSESATPE SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGEEPATSGSTPEGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP	288
8007	N-концевой XTEN	1	ASSPAGSPTSTESGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGT SESATPE**SGPGSTPAE**SGSETPGTSESATPE SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP	288
8008	C-концевой XTEN	1	PGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP	864

			ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSE**GSAPGTE**STP SEGSAPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPG	
8009	C-концевой XTEN	1	PGSPAGSPTSTEEGTSESATPESGPGSEPATS GSETPGTSESATPESGPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGTSESATPESGPGTSES ATPESGPGSEPATSGSETPGSEPATSGSETP GSPAGSPTSTEEGTSTEPSE**GSAPGTE**STPS EGSAPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAPG	576
8010	N-концевой XTEN	2	SAGSPGSPAGSPTSTEEGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPE**SGPGSTPA** **E**SGSETPGSEPATSGSETPGSPAGSPTSTEE	1152

			GTSESATPESGPGTSTEPSEGSAPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGTSTEPSEGSA PGTSESATPESGPGSE**PATSGSTE**TPGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS ESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSPAGSPTSTEEGS PAGSPTSTEEGSPAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEP ATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGTSTEPSEGSAPGTSTEPSEGSAPGTSES ATPESGPGTESAS	
8011	С-концевой	3	SAGSPGSPAGSPTSTEEGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS	1152

	XTEN		APGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEG SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAP GTSESATPESGPGSEPATSGSETPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGSPAGSPTSTEEGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSESATPESGPGSPAGSPTSTEEGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEP ATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGSE**PATSGSTE**TPGS PAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SE**GSAPGTATE**SPEGSAPGTSESATPESGPG TSTEPSE**GSAPGTSAE**SATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTS ESATPESGPGTESAS	

8012	N-концевой XTEN	2	GSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSE**GSAPATSE**SATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSE**SA** **SPE**SGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAP	864
8013	N-концевой XTEN	1	GSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGSE**SATSGSE** TPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSES	864

			ATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSET PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAP	
8014	N-концевой XTEN (с His-меткой)	1	SPAGSPTSTESGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSES ATPE**SGPGSTPAESGSE**TPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGTSESATPESGPGTSES ATPESGPGSEPATSGSETPGSEPATSGSETP GSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGGSAP	292
8015	C-концевой XTEN	1	PGSPAGSPTSTEEGTSESATPESGPGSEPATS GSETPGTSESATPESGPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSP	582

			TSTEEGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGTSESATPESGPGTSES ATPESGPGSEPATSGSETPGSEPATSGSETP GSPAGSPTSTEEGTSTEPSE**GSAPGTE**STPS EGSAPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAPGEPEA	
8016	C-концевой XTEN	1	TPESGPGTSESATPESGPGSPAGSPTSTEEGT SESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGTSESATPESGPGSEPATSGS ETPGSE**SATSGSE**TPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESA	576
8017	C-	1	GTSTEPSEGSAPGTSESATPESGPGSEPATS	576

	концевой XTEN		GSETPGSEPATSGSETPGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSA PGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGSPAGSPTSTEEGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSA PGTSE**SASPE**SGPGSPAGSPTSTEEGSPAGS PTSTEEGSPAGSPTSTEEGTSESATPESGPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSEPATSGSETPGTSESA TPESGP	
8018	C-концевой XTEN	1	GSETPGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSESATPESGPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGTSTEPSEG SAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGTS ESATPESGPGSE**PATSGSTE**TGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEP	576

			SEGSAPGSPAGSPTSTEEGTSESATPESGPG SEPATS	
8019	C-концевой XTEN	1	EGSAPGTSTEPSEGSAPGTSESATPESGPGT STEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGSPAGSPTSTEEGT SESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGTSE**SASPE**SGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESAT	576

[100] В некоторых вариантах осуществления полипептид XTEN со штрих-кодом может быть получен путем внесения одной или более мутаций в универсальный полипептид XTEN, такой как любой из перечисленных в **таблице 3b**, в соответствии с одним или более из следующих критериев: минимизировать изменение последовательности в полипептиде XTEN, минимизировать изменение аминокислотного состава в полипептиде XTEN, фактически сохранить суммарный заряд полипептида XTEN, фактически сохранить (или улучшить) низкую иммуногенность полипептида XTEN и фактически сохранить (или улучшить) фармакокинетические свойства полипептида XTEN. В некоторых вариантах осуществления аминокислотная последовательность полипептида XTEN характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100% идентичностью последовательности с любой из **SEQ ID NO: 676-734**, перечисленными в **таблице 3b**. В некоторых вариантах осуществления последовательность XTEN, характеризующаяся по меньшей мере 90% (например, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98% или по меньшей мере 99%), но менее чем 100% идентичностью последовательности с любой из **SEQ ID NO:**

676-734, перечисленными в **таблице 3b**, получают при помощи одной или более мутаций (например, менее 10, менее 8, менее 6, менее 5, менее 4, менее 3, менее 2 мутаций) соответствующей последовательности из **таблицы 3b**. В некоторых вариантах осуществления одна или более мутаций предусматривают делецию остатка глутаминовой кислоты, вставку остатка глутаминовой кислоты, замену остатка глутаминовой кислоты или замену на остаток глутаминовой кислоты или любую их комбинацию. В некоторых вариантах осуществления, где аминокислотная последовательность полипептида XTEN отличается, но характеризуется по меньшей мере 90% (например, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98% или по меньшей мере 99%) идентичностью последовательности с любой из **SEQ ID NO: 676-734**, перечисленными в **таблице 3b**, по меньшей мере 80%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 97% или приблизительно 100% различий между аминокислотной последовательностью полипептида XTEN и соответствующей последовательностью в **таблице 3b** предусматривают делецию остатка глутаминовой кислоты, вставку остатка глутаминовой кислоты, замену остатка глутаминовой кислоты или замену на остаток глутаминовой кислоты или любую их комбинацию. В некоторых таких вариантах осуществления по меньшей мере 80%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 97% или приблизительно 100% различий между аминокислотной последовательностью полипептида XTEN и соответствующей последовательностью в **таблице 3b** предусматривают замену остатка глутаминовой кислоты или замену на остаток глутаминовой кислоты или и то, и другое. Термин «замена первой аминокислоты», используемый в данном документе, относится к замещению первого аминокислотного остатка вторым аминокислотным остатком, в результате чего второй аминокислотный остаток занимает место в заменяемом положении в полученной последовательности. Например, «замена глутаминовой кислоты» относится к замене остатка глутаминовой кислоты (E) на остаток, отличный от глутаминовой кислоты (например, серин (S)). Термин «замена на первую аминокислоту», используемый в данном документе, относится к замещению второго аминокислотного остатка первым аминокислотным остатком, в результате чего первой аминокислотный остаток занимает место в заменяемом положении в полученной последовательности. Например, «замена на глутаминовую кислоту» относится к замене остатка, отличного от глутаминовой кислоты (например, серина (S)), на остаток глутаминовой кислоты.

Таблица 3b. Иллюстративные универсальные XTEN, предназначенные для построения XTEN со штрих-кодом

SEQ ID NO	**Название XTEN**	**Аминокислотная последовательность**
676	AE144	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP
677	AE144_1A	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEE

		GTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSET PGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APG
678	AE144_2A	TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSET PGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPG
679	AE144_2B	TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSET PGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPG
680	AE144_3A	SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APG
681	AE144_3B	SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APG
682	AE144_4A	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APG
683	AE144_4B	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APG
684	AE144_5A	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTST EEG
685	AE144_6B	TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGP GSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSA PGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APG
686	AE288_1	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATP ESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAP
687	AE288_2	GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAP
688	AE576	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG

		SAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAP
689	AE624	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSST GSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPE SGPGTSTEPSEGSAP
690	AE864	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAP
691	AE865	GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT

		STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAP
692	AE866	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPG
693	AE1152	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPG

		SPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAP
694	AE144A	STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTE EGS
695	AE144B	SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSET PGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APG
696	AE180A	TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS
697	AE216A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTE PSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEP ATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGSEPATSGSETPGTSESAT
698	AE252A	ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTS ESATPESGPGTSTEPSE
699	AE288A	TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPG SPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESA
700	AE324A	PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPG SPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS
701	AE360A	PESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGS PTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSE PATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPG SEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGTSESAT
702	AE396A	PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESA TPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS

		PAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETP GSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS
703	AE432A	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSET PGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGSEPATS
704	AE468A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEE GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESG PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT
705	AE504A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSA PGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEG SAPGSEPATSGSETPGTSESATPESGPGTSTEPS
706	AE540A	TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEP
707	AE576A	TPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEE

		GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATP ESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATS GSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPAT SGSETPGTSESA
708	AE612A	GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESA T
709	AE648A	PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEP SEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEE GSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESAT PESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT
710	AE684A	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGS ETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGSEPATS
711	AE720A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS

		TEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESAT PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGS PAGSPTSTEEGTSTE
712	AE756A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESAT PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPG TSES
713	AE792A	EGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS PTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTE PSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEP ATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPS
714	AE828A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS

		TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTST EPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSE PATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGSEPATSGSETPGTSESAT
715	AE869	GSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPE SGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGR
716	AE144_R1	SAGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTES ASR
717	AE288_R1	SAGSPTGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGS EPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEG TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGP GSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSA PGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APSASR
718	AE432_R1	SAGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA

		PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST EEGTESASR
719	AE576_R1	SAGSPTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGS ETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPSASR
720	AE864_R1	SAGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESAT PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPG TSESATPESGPGTESASR
721	AE712	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPT STEAHHH
722	AE864_R2	GSPGAGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEP

		ATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPG SPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGTESASR
723	AE288_3	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPG
724	AE284	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATP ESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP SE
725	AE292	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP
726	AE864_2	AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPG SEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA PGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPE SGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEG

		TSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEE GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESG PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGAAEPEA
727	AE867	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGAAEPEA
728	AE867_2	SPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPG
729	AE868	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT

		STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGAAEPEA
730	AE144_7A	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAP
731	AE292	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGGSAP
732	AE293	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGAAEPEA
733	AE300	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGAAEPEA
734	AE584	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP

		GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGAAEPEA

[101] В некоторых вариантах осуществления для конструирования последовательности полипептида XTEN со штрих-кодом выполняют аминокислотные мутации в полипептидах XTEN промежуточной длины по сравнению с приведенными в **таблице 3b**, а также в полипептидах XTEN с большей длиной, чем у приведенных в **таблице 3b**, как, например, такие, при которых один или более 12-мерных мотивов из **таблицы 1** добавляют к N- или C-концу универсального XTEN из **таблицы 3b**.

[102] Дополнительные примеры аминокислотных последовательностей универсального полипептида XTEN, которые могут быть применены в соответствии с настоящим изобретением, раскрыты в публикациях заявок на патент США №№ 2010/0239554 A1, 2010/0323956 A1, 2011/0046060 A1, 2011/0046061 A1, 2011/0077199 A1 или 2011/0172146 A1, или в публикациях международных патентных заявок №№ WO 2010091122 A1, WO 2010144502 A2, WO 2010144508 A1, WO 2011028228 A1, WO 2011028229 A1, WO 2011028344 A2, WO 2014/011819 A2 или WO 2015/023891, раскрытие каждой из которых явно включено в данный документ посредством ссылки.

[103] В некоторых вариантах осуществления полипептид XTEN со штрих-кодом, слитый с полипептидной цепью, примыкающей к N-концу полипептидной цепи («N-концевой XTEN»), может быть присоединен к His-метке, содержащей множество остатков поли(His), включая от шести до восьми остатков His на N-конце, для облегчения очистки слитого полипептида. В некоторых вариантах осуществления полипептид XTEN со штрих-кодом, слитый с полипептидной цепью на C-конце полипептидной цепи («C-концевой полипептид XTEN»), может содержать или быть присоединенным к последовательности EPEA на C-конце для облегчения очистки слитого полипептида. В некоторых вариантах осуществления слитый полипептид содержит как полипептид XTEN со штрих-кодом на N-конце, так и полипептид XTEN со штрих-кодом на C-конце, где XTEN со штрих-кодом на N-конце присоединен к His-метке, содержащей множество остатков поли(His), включая от шести до восьми остатков His на N-конце; и где полипептид XTEN со штрих-кодом на C-конце присоединен к последовательности EPEA на C-конце, с облегчением тем самым очистки слитого полипептида, например, до по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или по меньшей мере 99% чистоты с помощью хроматографических способов, известных в уровне техники, в том числе без ограничения хроматографии IMAC, аффинной матрицы C-tagXL и других подобных способов, в том числе без ограничения описанных ниже в разделе «ПРИМЕРЫ».

РАСЩЕПЛЕНИЕ ПРОТЕАЗОЙ

[104] Фрагмент со штрих-кодом, который описан выше или где-либо еще в данном документе, может быть слит с возможностью расщепления внутри полипептида XTEN и высвобождаться (сконфигурирован для высвобождения) из полипептида XTEN после расщепления полипептида протеазой. В некоторых вариантах осуществления протеаза

представляет собой протеазу Glu-C. В некоторых вариантах осуществления протеаза расщепляет с С-концевой стороны от остатков глутаминовой кислоты, за которыми не следует пролин. Специалисту в данной области будет понятно, что полипептид XTEN со штрих-кодом (полипептид XTEN, который содержит внутри себя фрагмент(-ы) со штрих-кодом) предназначен для достижения высокой эффективности, точности и правильности расщепления протеазой. Например, специалисту в данной области будет понятно, что примыкающие остатки Glu-Glu (EE) в последовательности XTEN могут приводить к различным моделям расщепления при расщеплении с помощью Glu-C. Соответственно, когда протеаза Glu-C применяют для высвобождения штрих-кода, тогда полипептид XTEN со штрих-кодом или фрагмент(-ы) со штрих-кодом могут не содержать какой-либо последовательности Glu-Glu (EE). Специалисту в данной области также понятно, что последовательность дипептида Glu-Pro (EP), если она присутствует в слитом полипептиде, может быть неспособна к расщеплению протеазой Glu-C в процессе высвобождения штрих-кода.

Структурная конфигурация BPXTEN

[105] В некоторых вариантах осуществления слитый белок BPXTEN содержит один полипептид BP и один полипептид XTEN. Такие белки BPXTEN могут иметь по меньшей мере следующие перестановки конфигураций, каждая из которых указана в ориентации от N-конца к С-концу: BP-XTEN; XTEN-BP; BP-S-XTEN и XTEN-S-BP, где «S» представляет собой спейсерную последовательность, указанную ниже.

[106] В некоторых вариантах осуществления белок BPXTEN содержит С-концевой полипептид XTEN и, необязательно, спейсерную последовательность (S) между полипептидом XTEN и полипептидом BP. Такой белок BPXTEN может быть представлен **формулой I** (изображен от N-конца к С-концу):

(BP)-$(S)_x$-(XTEN) (**I**),

где BP представляет собой биологически активный белок, который описан в данном документе ниже; S представляет собой спейсерную последовательность, содержащую от 1 до приблизительно 50 аминокислотных остатков, которая может необязательно включать обеспечивающий высвобождение BP сегмент (как более подробно описано ниже); x равняется или 0, или 1; и XTEN может представлять собой любой описанный в данном документе полипептид XTEN.

[107] В некоторых вариантах осуществления белок BPXTEN содержит N-концевой полипептид XTEN и, необязательно, спейсерную последовательность (S) между полипептидом XTEN и белком BP. Такие белки BPXTEN могут быть представлены **формулой II** (изображены от N-конца к С-концу):

(XTEN)-$(S)_x$-(BP) (**II**),

где BP представляет собой биологически активный белок, который описан в данном документе ниже; S представляет собой спейсерную последовательность, содержащую от 1 до приблизительно 50 аминокислотных остатков, которая может необязательно включать обеспечивающий высвобождение BP сегмент (как более

подробно описано ниже); x равняется или 0, или 1; и XTEN может представлять собой любой полипептид XTEN, который описан в данном документе.

[108] В некоторых вариантах осуществления белок BPXTEN содержит как N-концевой полипептид XTEN, так и C-концевой полипептид XTEN. Такие белки BPXTEN (*например*, XPAT на фигурах 1-2) могут быть представлены **формулой III**:

$(XTEN)-(S)_y-(BP)-(S)_z-(XTEN)$ (**III**),

где BP представляет собой биологически активный белок, который описан в данном документе ниже; S представляет собой спейсерную последовательность, содержащую от 1 до приблизительно 50 аминокислотных остатков, которая может необязательно включать обеспечивающий высвобождение BP сегмент (как более подробно описано ниже); y равняется или 0, или 1; z равняется или 0, или 1; и XTEN может представлять собой любой полипептид XTEN, который описан в данном документе.

БИОЛОГИЧЕСКИ АКТИВНЫЙ ПОЛИПЕПТИД

[109] Биологически активный белок (BP), который может быть слит с одним или более полипептидами XTEN (которые описаны в данном документе), в частности с полипептидами, раскрытыми в данном документе ниже, содержащий последовательности, идентифицированные в данном документе в **таблицах 4a-4h** и **таблицах 6a-6f**, вместе с соответствующими им последовательностями нуклеиновых кислот и аминокислотными последовательностями, хорошо известны в уровне техники. Описания и последовательности этих BP доступны в общедоступных базах данных, таких как базы данных Chemical Abstracts Services (*например*, CAS Registry), GenBank, The Universal Protein Resource (UniProt) и базах данных, предоставляемых по подписке, таких как GenSeq (например, Derwent). Полинуклеотидные последовательности, кодирующие BP, могут быть полинуклеотидными последовательностями дикого типа, кодирующими нативный BP (*например*, или полноразмерный, или зрелый), или в некоторых случаях последовательность может быть вариантом полинуклеотидной последовательности дикого типа (*например*, полинуклеотид, который кодирует биологически активный белок дикого типа), где нуклеотидная последовательность полинуклеотида оптимизирована, например, для экспрессии у конкретных видов; или полинуклеотидом, кодирующим вариант белка дикого типа, такой как сайт-направленный мутант или аллельный вариант. Специалист в данной области вполне может использовать последовательность cDNA дикого типа, или консенсусную последовательность cDNA, или кодон-оптимизированный вариант BP для создания конструкций BPXTEN, предусмотренных в настоящем изобретении, с применением способов, известных в уровне техники, и/или в сочетании с руководством и способами, предусмотренными в данном документе.

[110] BP для включения в белки BPXTEN, раскрытые в данном документе (*например*, слитый полипептид, содержащий по меньшей мере один полипептид BP и по меньшей мере один XTEN), может предусматривать любой белок, предусматривающий биологические, терапевтические, профилактические или диагностические интерес или

функцию, или который применим для опосредования биологической активности или предупреждения или облегчения заболевания, нарушения или состояния при введении человеку или животному. Особенно предпочтительными являются BP, для которых требуется увеличение фармакокинетического параметра, повышенная растворимость, повышенная стабильность, маскирование активности или какое-либо другое улучшенное фармацевтическое свойство, или те BP, для которых увеличение конечного периода полувыведения улучшит эффективность, безопасность или приведет к снижению частоты введения дозы и/или улучшению соблюдения пациентом режима лечения. Таким образом, композиции на основе слитых белков BPXTEN могут быть получены с различными целями, включая улучшение терапевтической эффективности биоактивного соединения, например, за счет увеличения воздействия in vivo или продолжительности времени, в течение которого BPXTEN остается в пределах терапевтического окна при введении человеку или животному по сравнению с BP, не связанным с полипептидом XTEN.

[111] BP может быть нативным полноразмерным белком или может быть фрагментом или вариантом последовательности биологически активного белка, который сохраняет по меньшей мере часть биологической активности нативного белка.

[112] В одном варианте осуществления BP, включенный в композиции для человека или животного, может представлять собой рекомбинантный полипептид с последовательностью, соответствующей белку, встречающемуся в природе. В другом варианте осуществления BP могут представлять собой варианты последовательности, фрагменты, гомологи и миметики природной последовательности, которые сохраняют по меньшей мере часть биологической активности нативного BP. В неограничивающих примерах BP может представлять собой последовательность, которая проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с последовательностью белка, выбранной из **таблиц 4a-4h**. В дополнительных неограничивающих примерах BP может представлять собой биспецифическую последовательность, содержащую первый связывающий домен и второй связывающий домен, где первый связывающий домен, обладающий аффинностью специфического связывания с опухолеспецифическим маркером или антигеном клетки-мишени, проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с парными последовательностями VL и VH антитела к CD3, идентифицированными в **таблице 6f**; и где второй связывающий домен, обладающий аффинностью специфического связывания с эффекторной клеткой, проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с парными последовательностями VL и VH антитела к клетке-

мишени, идентифицированными в **таблице 6а**. В одном варианте осуществления слитый белок BPXTEN может содержать один белок BP, связанный с полипептидом XTEN. В другом варианте осуществления белок BPXTEN может содержать первый BP и вторую молекулу того же BP, что дает в результате слитый белок, содержащий два BP, связанных с одним или более полипептидами XTEN (например, две молекулы глюкагона или две молекулы hGH).

[113] Как правило, BP проявляет специфичность связывания с данной мишенью (или заданным числом мишеней) или другую необходимую биологическую характеристику при применении in vivo или при применении в анализе in vitro. Например, BP может быть агонистом, рецептором, лигандом, антагонистом, ферментом, антителом (например, моно- или биспецифическим) или гормоном. Особый интерес представляют BP, используемые или известные как пригодные для лечения заболевания или нарушения, при которых нативный BP имеет относительно короткий конечный период полувыведения, и для которого улучшение фармакокинетического параметра (который может быть необязательно высвобожден из слитого белка путем расщепления спейсерной последовательности) обеспечит менее частое введение дозы или усиленный фармакологический эффект. Также представляют интерес BP, которые имеют узкое терапевтическое окно между минимальной эффективной дозой или концентрацией в крови (Cmin) и максимально переносимой дозой или концентрацией в крови (Cmax). В таких случаях связывание BP со слитым белком, содержащим выбранную(-ые) полипептидную(-ые) последовательность(-и) XTEN, может привести к улучшению этих свойств, делая их более полезными в качестве терапевтических или профилактических средств по сравнению с BP, не связанными с одним или более полипептидами XTEN.

Пептиды, регулирующие уровень глюкозы

[114] Эндокринные и связанные с ожирением заболевания или нарушения достигли масштабов эпидемии в большинстве развитых стран и представляют собой существенную и растущую нагрузку на здравоохранение в большинстве развитых стран, которые включают большое разнообразие состояний, поражающих органы, ткани и систему кровообращения организма. Особую озабоченность вызывают эндокринные заболевания и нарушения, связанные с ожирением, главным из которых является диабет, одна из основных причин смерти в Соединенных Штатах Америки.

[115] Большинство метаболических процессов в гомеостазе глюкозы и при инсулиновом ответе регулируются множеством пептидов и гормонов, и многие такие пептиды и гормоны, а также их аналоги нашли применение при лечении метаболических заболеваний и нарушений. Многие из этих пептидов имеют тенденцию быть высоко гомологичными друг другу, даже если они обладают противоположными биологическими функциями. Примером пептидов, повышающих уровень глюкозы, является пептидный гормон глюкагон, в то время как пептиды, снижающие уровень глюкозы, включают эксендин-4, глюкагоноподобный пептид 1 и амилин. Однако применение терапевтических пептидов и/или гормонов, даже в сочетании с применением низкомолекулярных

лекарственных средств, имело ограниченный успех в лечении таких заболеваний и нарушений. В частности, оптимизация дозы важна для лекарственных средств и биологических препаратов, применяемых для лечения метаболических заболеваний, особенно лекарственных средств и биологических препаратов с узким терапевтическим окном. Гормоны в целом и пептиды, участвующие в гомеостазе глюкозы, часто имеют узкое терапевтическое окно. Узкое терапевтическое окно в сочетании с тем фактом, что такие гормоны и пептиды обычно имеют короткий период полужизни, что требует частого введения дозы для достижения клинического эффекта, приводит к трудностям в лечении таких пациентов. Хотя химические модификации терапевтического белка, такие как пегилирование, могут изменить скорость его клиренса in vivo и последующее время полужизни в сыворотке, это требует дополнительных производственных стадий и приводит к получению гетерогенного конечного продукта. Кроме того, сообщалось о неприемлемых побочных эффектах при длительном применении. В качестве альтернативы, генетическая модификация путем слияния Fc-домена с терапевтическим белком или пептидом увеличивает размер терапевтического белка, снижая скорость клиренса через почки, и способствует рециркуляции из лизосом с помощью рецептора FcRn. К сожалению, Fc-домен не сворачивается эффективно во время рекомбинантной экспрессии и имеет тенденцию к образованию нерастворимых осадков, известных как тельца включения. Эти тельца включения должны быть солюбилизированы, а функциональный белок должен быть ренатурирован; трудоемкий, неэффективный и дорогостоящий процесс.

[116] Таким образом, одним из аспектов настоящего изобретения является включение пептидов, участвующих в гомеостазе глюкозы, резистентности к инсулину и ожирении (совместно именуемых «пептиды, регулирующие уровень глюкозы»), в слитые белки BPXTEN для создания композиций, пригодных для лечения нарушений, заболеваний и сопутствующих состояний, связанных с нарушением уровня глюкозы, инсулина и ожирением. Подходящие пептиды, регулирующие уровень глюкозы, которые могут быть связаны с полипептидами XTEN, раскрытыми в данном документе, для создания белков BPXTEN, включают все биологически активные полипептиды, inter alia, которые повышают глюкозозависимую секрецию инсулина бета-клетками поджелудочной железы или усиливают действие инсулина. Пептиды, регулирующие уровень глюкозы, также могут включать биологически активные полипептиды, которые стимулируют транскрипцию гена проинсулина в бета-клетках поджелудочной железы. Кроме того, пептиды, регулирующие уровень глюкозы, могут также включать биологически активные полипептиды, которые замедляют время опорожнения желудка и снижают уровень потребления пищи. Пептиды, регулирующие уровень глюкозы, могут также включать биологически активные полипептиды, которые ингибируют высвобождение глюкагона из альфа-клеток островков Лангерганса. В **таблице 4а** представлен неограничивающий перечень последовательностей пептидов, регулирующих уровень глюкозы, которые могут быть включены в состав слитых белков BPXTEN по настоящему изобретению. Пептиды,

регулирующие уровень глюкозы, в композициях на основе BPXTEN по настоящему изобретению, раскрытых в данном документе, могут представлять собой пептиды, которые проявляют по меньшей мере приблизительно 80% идентичность последовательности (например, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности) с аминокислотной последовательностью, выбранной из **таблицы 4а**.

Таблица 4а. Пептиды, регулирующие уровень глюкозы

Название белка (Синоним)	**SEQ ID NO**	**Аминокислотная последовательность**
Адреномедуллин (ADM)	1	YRQSMNNFQGLRSFGCRFGTCTVQKLAHQIYQFTD KDKDNVAPRSKISPQGY
Амилин, крыса	2	KCNTATCATQRLANFLVRSSNNLGPVLPPTNVGSN TY
Амилин, человек	3	KCNTATCATQRLANFLVHSSNNFGAILSSTNVGSN TY
Кальцитонин (hCT)	4	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP
Кальцитонин, лосось	5	CSNLSTCVLGKLSQELHKLQTYPRTNTGSGTP
Кальцитонин-ген родственный пептид (h-CGRP α)	6	ACDTATCVTHRLAGLLSRSGGVVKNMVPTNVGSK AF
Кальцитонин-ген родственный пептид (h-CGRP β)	7	ACNTATCVTHRLAGLLSRSGGMVKSNFVPTNVGS KAF
Холецистокинин (CCK)	8	MNSGVCLCVLMAVLAAGALTQPVPPADPAGSGLQ RAEEAPRRQLRVSQRTDGESRAHLGALLARYIQQA RKAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDF GRRSAEEYEYPS
CCK-33	9	KAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDF
CCK-8	10	DYMGWMDF
Эксендин-3	11	HSDGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGA PPPS
Эксендин-4	12	HGEGTFTSDLSKQMEEEAVR LFIEWLKNGGPSSGAPPPS
FGF-19	13	MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPH VHYGWGDPIRLRHLYTSGPHGLSSCFLRIRADGVV DCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCM GADGKMQGLLQYSEEDCAFEEEIRPDGYNVYRSE KHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMV PEEPEDLRGHLESDMFSSPLETDSMDPFGLVTGLEA VRSPSFEK
FGF-21	14	MDSDETGFEHSGLWVSVLAGLLLGACQAHPIPDSS PLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVG GAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRP DGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHG LPLHLPGNKSPHRDPAPRGPARFLPLPGLPPALPEPP GILAPQPPDVGSSDPLSMVGPSQGRSPSYAS
Гастрин	15	QLGPQGPPHLVADPSKKQGPWLEEEEEAYGWMDF
Гастрин-17	16	DPSKKQGPWLEEEEEAYGWMDF
Желудочный	17	YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKN

ингибирующий полипептид (GIP)		DWKHNITQ
Грелин	18	GSSFLSPEHQRVQQRKESKKPPAKLQPR
Глюкагон	19	HSQGTFTSDYSKYLDSRRAQDFVQWLMNT
Глюкагоноподобный пептид-1 (hGLP-1) (GLP-1; 1-37)	20	HDEFERHAEGTFTSDVSSTLEGQAALEFIAWLVKG RG
GLP-1 (7-36), человек	21	HAEGTFTSDVSSYLEGQAALEFIAWLVKGR
GLP-1 (7-37), человек	22	HAEGTFTSDVSSTLEGQAALEFIAWLVKGRG
GLP-1, лягушка	23	HAEGTYTNDVTEYLEEKAAKEFIEWLIKGKPKKIR YS
Глюкагоноподобный пептид 2 (GLP-2), человек	24	HADGSFSDEMNTILDNLAARDFINWLIETKITD
GLP-2, лягушка	25	HAEGTFTNDMTNYLEEKAAKEFVGWLIKGRP-OH
IGF-1	26	GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSS RRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSA
IGF-2	27	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE
Пептид INGAP (белок, связанный с неогенезом островков)	28	EESQKKLPSSRITCPQGSVAYGSYCYSLILIPQTWSN AELSCQMHFSGHLAFLLSTGEITFVSSLVKNSLTAY QYIWIGLHDPSHGTLPNGSGWKWSSSNVLTFYNW ERNPSIAADRGYCAVLSQKSGFQKWRDFNCENELP YICKFKV
Интермедин (AFP-6)	29	TQAQLLRVGCVLGTCQVQNLSHRLWQLMGPAGR QDSAPVDPSSPHSY
Лептин, человек	30	VPIQKVQDDTKTLIKTIVTRINDISHTQSVSSKQKVT GLDFIPGLHPILTLSKMDQTLAVYQQILTSMPSRNVI QISNDLENLRDLLHVLAFSKSCHLPWASGLETLDSL GGVLEASGYSTEVVALSRLQGSLQDMLWQLDLSP GC
Нейромедин (U-8) свиной	31	YFLFRPRN
Нейромедин (U-9)	32	GYFLFRPRN
Нейромедин (U25) человеческий)	33	FRVDEEFQSPFASQSRGYFLFRPRN
Нейромедин (U25) свиной	34	FKVDEEFQGPIVSQNRRYFLFRPRN
Нейромедин S, человек	35	ILQRGSGTAAVDFTKKDHTATWGRPFFLFRPRN
Нейромедин U, крыса	36	YKVNEYQGPVAPSGGFFLFRPRN
Оксинтомодулин (OXM)	37	HSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRN NIA
Пептид YY (PYY)	38	YPIKPEAPGEDASPEELNRYYASLRHYLNLVTRQR Y
Прамлинтид	39	KCNTATCATNRLANFLVHSSNNFGPILPPTNVGSNT Y-NH2
Урокортин (Ucn-1)	40	DNPSLSIDLTFHLLRTLLELARTQSQRERAEQNRIIF

		DSV
Урокортин (Ucn-2)	41	IVLSLDVPIGLLQILLEQARARAAREQATTNARILAR VGHC
Урокортин (Ucn-3)	42	FTLSLDVPTNIMNLLFNIAKAKNLRAQAAANAHLM AQI

[117] «Адреномедуллин» или «ADM» означает человеческий пептидный гормон адреномедуллин и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого ADM. ADM образуется из препрогормона, состоящего из 185 аминокислот, посредством последовательного ферментативного расщепления и амидирования, в результате чего образуется биологически активный пептид из 52 аминокислот с измеренным периодом полужизни в плазме крови, составляющим 22 минуты. Содержащие ADM слитые белки по настоящему изобретению могут найти конкретное применение при диабете для стимулирующего воздействия на секрецию инсулина островковыми клетками с целью регуляции уровня глюкозы у человека или животных с устойчивой гипотензией. Сообщалось о полной геномной инфраструктуре AM человека (Ishimitsu et al., 1994, Biochem. Biophys. Res. Commun 203:631-639), и были клонированы аналоги пептидов ADM, как описано в патенте США № 6320022.

[118] «Амилин» означает человеческий пептидный гормон, называемый амилином, прамлинтидом, и его видовые вариации, которые описаны в патенте США № 5234906, обладающий по меньшей мере частью биологической активности зрелого амилина. Амилин представляет собой полипептидный гормон, состоящий из 37 аминокислот, который секретируется совместно с инсулином бета-клетками поджелудочной железы в ответ на потребление питательных веществ (Koda et al., 1992, Lancet 339:1179-1180), и, как сообщается, он модулирует несколько ключевых путей метаболизма углеводов, в том числе включение глюкозы в гликоген. Амилинсодержащие слитые белки по настоящему изобретению могут дополнять действие инсулина, который регулирует скорость исчезновения глюкозы из кровотока и ее поглощение периферическими тканями. Аналоги амилина были клонированы, как описано в патентах США №№ 5686411 и 7271238.

[119] Можно создать миметики амилина, сохраняющие биологическую активность. Например, прамлинтид имеет последовательность KCNTATCATNRLANFLVHSSNNFGPILPPTNVGSNTY (SEQ ID NO: 43), где аминокислоты из последовательности амилина крысы заменяют аминокислоты в последовательности амилина человека. В одном варианте осуществления настоящее изобретение предусматривает слитые белки, содержащие миметики амилина с последовательностью $KCNTATCATX_1RLANFLVHSSNNFGX_2ILX_2X_2TNVGSNTY$ (SEQ ID NO: 44), где X_1 независимо представляет собой N или Q, и X_2 независимо представляет собой S, P или G. В одном варианте осуществления миметик амилина, включенный в BPXTEN, может иметь последовательность KCNTATCATNRLANFLVHSSNNFGGILGGTNVGSNTY (SEQ ID NO: 45). В другом варианте осуществления, где миметик амилина используется на С-конце BPXTEN,

миметик может иметь последовательность KCNTATCATNRLANFLVHSSNNFGGILGGTNVGSNTY(NH_2) (SEQ ID NO: 46).

[120] «Кальцитонин» (CT) означает белок кальцитонин человека и его виды и варианты последовательности, в том числе кальцитонин лосося («sCT»), обладающий по меньшей мере частью биологической активности зрелого CT. CT представляет собой пептид из 32 аминокислот, отщепленный от более крупного прогормона щитовидной железы, который, по-видимому, функционирует в нервной и сосудистой системах, но, как сообщается, также является мощным гормональным медиатором рефлекса насыщения. (Рассматривается в Becker, JCEM, 89(4): 1512-1525 (2004) и Sexton, Current Medicinal Chemistry 6: 1067-1093 (1999)). Содержащие кальцитонин слитые белки по настоящему изобретению могут найти конкретное применение для лечения остеопороза и в качестве терапии костной болезни Педжета. Синтетические пептиды кальцитонина были созданы, как описано в патентах США №№ 5175146 и 5364840.

[121] «Кальцитонин-ген родственный пептид» или «CGRP» означает пептид CGRP человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого CGRP, который является представителем семейства пептидов кальцитонина и у людей существует в двух формах, α-CGRP (пептид из 37 аминокислот) и β-CGRP. CGRP характеризуется 43-46% идентичностью последовательности с человеческим амилином. Содержащие CGRP слитые белки по настоящему изобретению могут найти конкретное применение для снижения заболеваемости, связанной с диабетом, облегчения гипергликемии и дефицита инсулина, ингибирования инфильтрации лимфоцитов в островки и защиты бета-клеток от аутоиммунного разрушения. Способы получения синтетического и рекомбинантного CGRP описаны в патенте США № 5374618.

[122] «Холецистокинин» или «CCK» означает пептид CCK человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого CCK. CCK-58 представляет собой зрелую последовательность, в то время как аминокислотная последовательность CCK-33, впервые идентифицированная у людей, является основной формой пептида в кровотоке. Семейство CCK также включает 8-аминокислотный in vivo C-концевой фрагмент («CCK-8»), пентагастрин или CCK-5, представляющий собой C-концевой пептид CCK(29-33), и CCK-4, представляющий собой C-концевой тетрапептид CCK(30-33). CCK представляет собой пептидный гормон желудочно-кишечного тракта, отвечающий за стимуляцию переваривания жиров и белков. Содержащие CCK-33 и CCK-8 слитые белки по настоящему изобретению могут найти конкретное применение для уменьшения повышения глюкозы в кровотоке после приема пищи и потенцирования повышения инсулина в кровотоке. Аналоги CCK-8 были получены, как описано в патенте США № 5631230.

[123] «Эксендин-3» означает пептид, регулирующий уровень глюкозы, выделенный из Heloderma horridum, и варианты его последовательности, обладающие по меньшей мере частью биологической активности зрелого эксендина-3. Амид эксендина-3

является специфической формой антагониста рецептора эксендина, который опосредует увеличение уровня панкреатического цАМФ и высвобождение инсулина и амилазы. Содержащие эксендин-3 слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета и нарушений, связанных с резистентностью к инсулину. Последовательность и способы ее анализа описаны в патенте США № 5424286.

[124] «Эксендин-4» означает пептид, регулирующий уровень глюкозы, обнаруженный в слюне ядозуба Heloderma suspectum, а также его виды и варианты последовательности, и включает нативную последовательность из 39 аминокислот HGEGTFTSDLSKQMEEEAVRLFIEYLKNGGPSSGAPPPS (SEQ ID NO: 47) и гомологичные последовательности и пептидные миметики и их варианты; природные последовательности, такие как от приматов, и не являющиеся природными, обладающие по меньшей мере частью биологической активности зрелого эксендина-4. Эксендин-4 представляет собой инкретиновый полипептидный гормон, который снижает уровень глюкозы в крови, способствует секреции инсулина, замедляет опорожнение желудка и усиливает чувство насыщения, обеспечивая заметное улучшение постпрандиальной гипергликемии. В **таблице 4b** показаны последовательности самых разных видов, тогда как в **таблице 4c** приведен перечень синтетических аналогов GLP-1; все они рассматриваются для применения в белках BPXTEN, описанных в данном документе.

[125] Фактор роста фибробластов 21, или «FGF-21», означает человеческий белок, кодируемый геном FGF-21, или его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого FGF-21. FGF-21 стимулирует поглощение глюкозы адипоцитами, но не другими типами клеток; эффект является аддитивным к активности инсулина. Содержащие FGF-21 слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета, в том числе вызывая повышенный расход энергии, утилизацию жира и экскрецию липидов. FGF-21 был клонирован, как описано в патенте США № 6716626.

[126] «Фактор роста фибробластов 19» или «FGF-19» означает человеческий белок, кодируемый геном FGF-19, или его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого FGF-19. FGF-19 представляет собой белок, принадлежащий к семейству факторов роста фибробластов (FGF). FGF-19 увеличивает экспрессию рецептора лептина в печени, скорость метаболизма, стимулирует поглощение глюкозы адипоцитами и приводит к потере массы тела на модели мышей с ожирением (Fu et al., 2004, Endocrinology 145: 2504-2603). Содержащие FGF-19 слитые белки по настоящему изобретению могут найти конкретное применение для повышения скорости метаболизма и купирования диабета, обусловленного диетой и дефицитом лептина. FGF-19 был клонирован и экспрессирован, как описано в заявке на патент США № 20020042367.

[127] «Гастрин» означает пептид гастрина человека, его усеченные версии, а также его виды и варианты последовательности, обладающие по меньшей мере частью

биологической активности зрелого гастрина. Гастрин встречается в основном в трех формах: гастрин-34 («большой гастрин»); гастрин-17 («маленький гастрин») и гастрин-14 («минигастрин») и имеет гомологию последовательности с CCK. Содержащие гастрин слитые белки по настоящему изобретению могут найти конкретное применение при лечении ожирения и диабета для регуляции уровня глюкозы. Гастрин был синтезирован, как описано в патенте США № 5843446.

[128] «Грелин» означает человеческий гормон, вызывающий чувство насыщения, или его виды и варианты последовательности, включая нативную, процессированную последовательность из 27 или 28 аминокислот и гомологичные последовательности. Уровни грелина увеличиваются перед едой и снижаются после еды, что может привести к повышенному потреблению пищи и увеличению жировой массы за счет действия, оказываемого на уровне гипоталамуса. Содержащие грелин слитые белки по настоящему изобретению могут найти конкретное применение в качестве агонистов; *например*, для избирательной стимуляции моторики GI-тракта при нарушении моторики желудочно-кишечного тракта, для ускорения опорожнения желудка или для стимуляции высвобождения гормона роста. Аналоги грелина с заменами в последовательности или усеченными вариантами, которые описаны в патенте США № 7385026, могут найти конкретное применение в качестве партнеров по слиянию с полипептидами XTEN для применения в качестве антагонистов для улучшения гомеостаза глюкозы, лечения резистентности к инсулину и лечения ожирения. Сообщалось о выделении и характеристике грелина (Kojima et al., 1999, Nature. 402:656-660), а синтетические аналоги были получены путем пептидного синтеза, как описано в патенте США № 6967237.

[129] «Глюкагон» означает регулирующий уровень глюкозы пептид глюкагона человека, или его виды и варианты последовательности, включая нативную последовательность из 29 аминокислот и гомологичные последовательности; природные, такие как от приматов, и не являющиеся природными варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого глюкагона. Используемый в данном документе термин «глюкагон» также включает пептидные миметики глюкагона. Содержащие глюкагон слитые белки по настоящему изобретению могут найти конкретное применение для повышения уровня глюкозы в крови у индивидуумов с существующими запасами гликогена в печени и для поддержания гомеостаза глюкозы при диабете. Глюкагон был клонирован, как описано в патенте США № 4826763.

[130] «GLP-1» означает глюкагоноподобный пептид-1 человека и варианты его последовательности, обладающие по меньшей мере частью биологической активности зрелого GLP-1. Термин «GLP-1» включает человеческий GLP-1(1-37), GLP-1(7-37) и GLP-1(7-36)амид. GLP-1 стимулирует секрецию инсулина, но только в периоды гипергликемии. Безопасность GLP-1 по сравнению с инсулином повышается благодаря этому свойству и наблюдению, что количество секретируемого инсулина пропорционально величине гипергликемии. Биологический период полужизни GLP-1(7-

37)OH составляет всего от 3 до 5 минут (патент США № 5118666). Содержащие GLP-1 слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета и нарушений, связанных с резистентностью к инсулину, для регуляции уровня глюкозы. GLP-1 был клонирован и получены производные, которые описаны в патенте США № 5118666. Неограничивающие примеры последовательностей GLP-1 из самых разных видов показаны в **таблице 4b**, тогда как в **таблице 4c** показаны последовательности ряда синтетических аналогов GLP-1; все из них предназначены для применения в композициях на основе BPXTEN, описанных в данном документе.

Таблица 4b. Примерные встречающиеся в природе гомологи GLP-1 в качестве кандидатов на BP

Название гена	**SEQ ID NO**	**Аминокислотная последовательность**
GLP-1 {лягушка}	48	HAEGTYTNDVTEYLEEKAAKEFIEWLIKGKPKKIRYS
GLP-1a {Xenopus laevis}	49	HAEGTFTSDVTQQLDEKAAKEFIDWLINGGPSKEIIS
GLP-1b {Xenopus laevis}	50	HAEGTYTNDVTEYLEEKAAKEFIIEWLIKGKPK
GLP-1c {Xenopus laevis}	51	HAEGTFTNDMTNYLEEKAAKEFVGWLIKGRPK
Желудочный ингибирующий полипептид {Mus musculus}	52	HAEGTFISDYSIAMDKIRQQDFVNWLL
Глюкозозависимый инсулинотропный полипептид {Equus caballus}	53	HAEGTFISDYSIAMDKIRQQDFVNWLL
Глюкагоноподобный пептид {Petromyzon marinus}	54	HADGTFTNDMTSYLDAKAARDFVSWLARSDKS
Глюкагоноподобный пептид {Anguilla rostrata}	55	HAEGTYTSDVSSYLQDQAAKEFVSWLKTGR
Глюкагоноподобный пептид {Anguilla anguilla}	56	HAEGTYTSDVSSYLQDQAAKEFVSWLKTGR
Глюкагоноподобный пептид {Hydrolagus colliei}	57	HADGIYTSDVASLTDYLKSKRFVESLSNYNKRQNDRRM
Глюкагоноподобный пептид {Amia calva}	58	YADAPYISDVYSYLQDQVAKKWLKSGQDRRE
GLUC_ICTPU/38-65	59	HADGTYTSDVSSYLQEQAAKDFITWLKS
GLUCL_ANGRO/1-28	60	HAEGTYTSDVSSYLQDQAAKEFVSWLKT
GLUC_BOVIN/98-125	61	HAEGTFTSDVSSYLEGQAAKEFIAWLVK
GLUC1_LOPAM/91-118	62	HADGTFTSDVSSYLKDQAIKDFVDRLKA
GLUCL_HYDCO/1-28	63	HADGIYTSDVASLTDYLKSKRFVESLSN

GLUC_CAVPO/53-80	64	HSQGTFTSDYSKYLDSRRAQQFLKWLLN
GLUC_CHIBR/1-28	65	HSQGTFTSDYSKHLDSRYAQEFVQWLMN
GLUC1_LOPAM/53-80	66	HSEGTFSNDYSKYLEDRKAQEFVRWLMN
GLUC_HYDCO/1-28	67	HTDGIFSSDYSKYLDNRRTKDFVQWLLS
GLUC_CALMI/1-28	68	HSEGTFSSDYSKYLDSRRAKDFVQWLMS
GIP_BOVIN/1-28	69	YAEGTFISDYSIAMDKIRQQDFVNWLLA
VIP_MELGA/89-116	70	HADGIFTTVYSHLLAKLAVKRYLHSLIR
PACA_CHICK/131-158	71	HIDGIFTDSYSRYRKQMAVKKYLAAVLG
VIP_CAVPO/45-72	72	HSDALFTDTYTRLRKQMAMKKYLNSVLN
VIP_DIDMA/1-28	73	HSDAVFTDSYTRLLKQMAMRKYLDSILN
EXE1_HELSU/1-28	74	HSDATFTAEYSKLLAKLALQKYLESILG
SLIB_CAPHI/1-28	75	YADAIFTNSYRKVLGQLSARKLLQDIMN
SLIB_RAT/31-58	76	HADAIFTSSYRRILGQLYARKLLHEIMN
SLIB_MOUSE/31-58	77	HVDAIFTTNYRKLLSQLYARKVIQDIMN
PACA_HUMAN/83-110	78	VAHGILNEAYRKVLDQLSAGKHLQSLVA
PACA_SHEEP/83-110	79	VAHGILDKAYRKVLDQLSARRYLQTLMA
PACA_ONCNE/82-109	80	HADGMFNKAYRKALGQLSARKYLHSLMA
GLUC_BOVIN/146-173	81	HADGSFSDEMNTVLDSLATRDFINWLLQ
SECR_CANFA/1-27	82	HSDGTFTSELSRLRESARLQRLLQGLV
SECR_CHICK/1-27	83	HSDGLFTSEYSKMRGNAQVQKFIQNLM
EXE3_HELHO/48-75	84	HSDGTFTSDLSKQMEEEAVRLFIEWLKN

Таблица 4c. Примерные синтетические аналоги GLP-1

SEQ ID NO	**Аминокислотная последовательность**
85	HAEGTFTSDVSSYLEGQAAREFIAWLVKGRG
86	HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG
87	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGKG
88	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKG
89	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKGR
90	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
91	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
92	HAEGTFTSDVSSYLEGQAAREFIAWLVKGKG
93	HAEGTFTSDVSSYLEGQAAKEFIAWLVRGKG
94	HAEGTFTSDVSSYLEGQAAREFIAWLVKGRGRK
95	HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRGRRK
96	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKGRK
97	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKGRRK
98	HGEGTFTSDVSSYLEGQAAREFIAWLVKGRG
99	HGEGTFTSDVSSYLEGQAAKEFIAWLVRGRG
100	HGEGTFTSDVSSYLEGQAAKEFIAWLVKGKG
101	HGEGTFTSDVSSYLEGQAAREFIAWLVRGKG
102	HGEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
103	HGEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
104	HGEGTFTSDVSSYLEGQAAREFIAWLVKGKG
105	HGEGTFTSDVSSYLEGQAAKEFIAWLVRGKG

106	HGEGTFTSDVSSYLEGQAAREFIAWLVKGRGRK
107	HGEGTFTSDVSSYLEGQAAKEFIAWLVRGRGRRK
108	HGEGTFTSDVSSYLEGQAAREFIAWLVRGKGRK
109	HGEGTFTSDVSSYLEGQAAREFIAWLVRGKGRRK
110	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
111	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
112	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
113	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
114	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
115	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
116	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
117	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
118	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
119	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
120	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
121	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
122	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
123	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
124	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
125	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
126	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
127	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
128	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
129	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
130	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
131	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
132	DEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
133	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
134	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
135	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
136	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
137	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
138	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
139	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
140	EFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
141	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
142	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
143	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
144	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
145	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
146	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
147	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
148	FERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
149	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
150	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
151	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
152	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
153	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
154	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
155	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK

156	ERHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
157	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
158	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRK
159	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRRK
160	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREK
161	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFK
162	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPK
163	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEK
164	RHAEGTFTSDVSSYLEGQAAREFIAWLVRGRGRREFPEEK
165	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVKGRGK
166	HDEFERHAEGTFTSDVSSYLEGQAAKEFIAWLVRGRGK
167	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGKGK
168	HAEGTFTSDVSSYLEGQAAREFIAWLVKGRGK
169	HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRGK
170	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKGK
171	HAEGTFTSDVSSYLEGQAAREFIAWLVRGRGK
172	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVKGRGRK
173	HDEFERHAEGTFTSDVSSYLEGQAAKEFIAWLVRGRGRK
174	HDEFERHAEGTFTSDVSSYLEGQAAREFIAWLVRGKGRK
175	HAEGTFTSDVSSYLEGQAAREFIAWLVKGRGRK
176	HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRGRK
177	HAEGTFTSDVSSYLEGQAAREFIAWLVRGKGRK
178	HGEGTFTSDVSSYLEGQAAREFIAWLVKGRGK
179	HGEGTFTSDVSSYLEGQAAREFIAWLVRGKGK

[131] Нативные последовательности GLP могут быть описаны несколькими мотивами последовательности, которые представлены ниже. Буквы в скобках обозначают допустимые аминокислоты в каждом положении последовательности: {HVY} {AGISTV} {DEHQ} {AG} {ILMPSTV} {FLY} {DINST} {ADEKNST} {ADENSTV} {LMVY} {ANRSTY} {EHIKNQRST} {AHILMQVY} {LMRT} {ADEGKQS} {ADEGKNQSY} {AEIKLMQR} {AKQRSVY} {{AILMQSTV} {GKQR} {DEKLQR} {FHLVWY} {ILV} {ADEGHIKNQRST} {ADEGNRSTW} {GILVW} {AIKLMQSV} {ADGIKNQRST} {GKRSY} (SEQ ID NO: 9399). Кроме того, синтетические аналоги GLP-1 могут быть полезны в качестве партнеров по слиянию с полипептидами XTEN для создания белка BPXTEN с биологической активностью, полезной при лечении нарушений, связанных с глюкозой.

[132] «GLP-2» означает глюкагоноподобный пептид-2 человека и варианты его последовательности, обладающие по меньшей мере частью биологической активности зрелого GLP-2. Более конкретно, GLP-2 представляет собой пептид из 33 аминокислот, совместно секретируемый вместе с GLP-1 эндокринными клетками тонкого и толстого кишечника.

[133] «Инсулиноподобный фактор роста 1» или «IGF-1» означает белок IGF-1 человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого IGF-1. IGF-1 состоит из 70 аминокислот и продуцируется главным образом печенью в качестве эндокринного гормона, а также в тканях-мишенях паракринным/аутокринным образом. Содержащие IGF-1 слитые белки по

настоящему изобретению могут найти особое применение при лечении диабета и нарушений, связанных с резистентностью к инсулину, для регуляции уровня глюкозы. IGF-1 был клонирован и экспрессирован в E. coli и дрожжах, как описано в патенте США № 5324639.

[134] «Инсулиноподобный фактор роста 2» или «IGF-2» означает белок IGF-2 человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого IGF-2. IGF-2 был клонирован, как описано в Bell et al., 1985, Proc Natl Acad Sci U S A. 82:6450-4.

[135] «Ассоциированный с неогенезом островков белок» (INGAP) или «фактор роста бета-клеток поджелудочной железы» означает человеческий пептид INGAP и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого INGAP. Содержащие INGAP слитые белки по настоящему изобретению могут найти конкретное применение при лечении или предупреждении диабета и нарушений, связанных с резистентностью к инсулину. INGAP был клонирован и экспрессирован, как описано в R. Rafaeloff et al., 1997, J Clin Invest. 99(9): 2100-2109.

[136] «Интермедин» или «AFP-6» означает пептид интермедина человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого интермедина. Лечение интермедином приводит к снижению артериального давления у людей или животных как в норме, так и при гипертонии, а также к подавлению активности опорожнения желудка, и принимает участие в гомеостазе глюкозы. Содержащие интермедин слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета, нарушений, связанных с резистентностью к инсулину, и ожирения. Пептиды интермединов и варианты были клонированы, как описано в патенте США № 6965013.

[137] «Лептин» означает встречающийся в природе лептин от любого вида, а также биологически активные D-изоформы или их фрагменты и варианты последовательности. Содержащие лептин слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета для регуляции уровня глюкозы, нарушений, связанных с резистентностью к инсулину, и ожирения. Лептин был клонирован, как описано в патенте США № 7112659, а аналоги и фрагменты лептина в патенте США № 5521283, патенте США № 5532336, PCT/US96/22308 и PCT/US96/01471.

[138] «Нейромедин» означает семейство пептидов нейромедина, в том числе пептиды нейромедина U и S, а также их варианты последовательности. В семейство нейромединов U входят различные усеченные или сплайсированные варианты, например FLFHYSKTQKLGKSNVVEELQSPFASQSRGYFLFRPRN (SEQ ID NO: 180). Примером семейства нейромединов S является человеческий нейромедин S с последовательностью ILQRGSGTAAVDFTKKDHTATWGRPFFLFRPRN (SEQ ID NO: 181), в частности его амидная форма. Слитые белки с нейромедином по настоящему изобретению могут найти конкретное применение при лечении ожирения, диабета, снижении потребления пищи и

при других сопутствующих состояниях и нарушениях, которые описаны в данном документе.

[139] «Оксинтомодулин» или «OXM» означает оксинтомодулин человека и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого OXM. OXM представляет собой пептид из 37 аминокислот, продуцируемый в толстой кишке, который содержит последовательность глюкагона из 29 аминокислот, за которой следует карбоксиконцевое удлинение из 8 аминокислот. Содержащие OXM слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета для регуляции уровня глюкозы, нарушений, связанных с резистентностью к инсулину, ожирения, и могут применяться в качестве средств лечения для потери массы тела.

[140] «PYY» означает полипептид человеческого пептида YY и его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого PYY. Содержащие PPY слитые белки по настоящему изобретению могут найти конкретное применение при лечении диабета для регуляции уровня глюкозы, нарушений, связанных с резистентностью к инсулину, и ожирения. Аналоги PYY были получены, как описано в патентах США №№ 5604203, 5574010 и 7166575.

[141] «Урокортин» означает человеческий пептидный гормон урокортин и варианты его последовательности, обладающие по меньшей мере частью биологической активности зрелого урокортина. Существует три урокортина человека: Ucn-1, Ucn-2 и Ucn-3. Дополнительные урокортины и аналоги описаны в патенте США № 6214797. Белки BPXTEN, содержащие урокортин по настоящему изобретению, также могут найти конкретное применение при лечении или предупреждении состояний, ассоциированных со стимуляцией высвобождения ACTH, гипертензии из-за сосудорасширяющих эффектов, воспаления, опосредованного другими причинами, отличными от повышения уровня ACTH, гипертермии, нарушения аппетита, застойной сердечной недостаточности, стресса, тревоги и псориаза. Содержащие урокортин слитые белки также можно комбинировать с модулем натрийуретического пептида, семейством амилинов и семейством эксендинов или модулем семейства GLP1 для обеспечения усиления пользы в отношении сердечно-сосудистой системы, например лечения CHF, в результате обеспечения благотворного сосудорасширяющего действия.

Метаболические заболевания и белки сердечно-сосудистой системы

[142] Метаболические и сердечно-сосудистые заболевания представляют собой существенную нагрузку на здравоохранение в большинстве развитых стран, при этом сердечно-сосудистые заболевания остаются основной причиной смерти и инвалидности в Соединенных Штатах Америки и большинстве европейских стран. Метаболические заболевания и нарушения включают большое разнообразие состояний, поражающих органы, ткани и систему кровообращения организма.

[143] Дислипидемия часто встречается у диабетиков и людей или животных с сердечно-сосудистым заболеванием; обычно характеризуется такими параметрами, как

повышенный уровень триглицеридов в плазме крови, низкий уровень холестерина HDL (липопротеинов высокой плотности), от нормального до повышенного уровня холестерина LDL (липопротеинов низкой плотности) и повышенные уровни частиц LDL малой плотности в крови. Дислипидемия и гипертония вносят основной вклад в повышение частоты коронарных явлений, заболеваний почек и летальных исходов у человека или среди животных с метаболическими заболеваниями, такими как диабет и сердечно-сосудистое заболевание.

[144] Сердечно-сосудистое заболевание может проявляться в виде множества нарушений, симптомов и изменений клинических параметров, затрагивающих сердце, сосуды и системы органов по всему организму, в том числе аневризм, стенокардии, атеросклероза, нарушения мозгового кровообращения (инсульта), цереброваскулярного заболевания, застойной сердечной недостаточности, ишемической болезни сердца, инфаркта миокарда, снижения сердечного выброса и заболевания периферических сосудов, гипертонии, гипотонии, маркеров крови (например, С-реактивный белок, BNP и ферменты, такие как CPK, LDH, SGPT, SGOT) среди прочего.

[145] Большинство метаболических процессов и многие параметры сердечно-сосудистой системы регулируются несколькими пептидами и гормонами («метаболическими белками»), и многие такие пептиды и гормоны, а также их аналоги, нашли применение при лечении таких заболеваний и нарушений. Однако применение терапевтических пептидов и/или гормонов, даже в сочетании с применением низкомолекулярных лекарственных средств, имело ограниченный успех в лечении таких заболеваний и нарушений. В частности, оптимизация дозы важна для лекарственных средств и биологических препаратов, применяемых для лечения метаболических заболеваний, особенно лекарственных средств и биологических препаратов с узким терапевтическим окном. Гормоны в целом и пептиды, участвующие в гомеостазе глюкозы, часто имеют узкое терапевтическое окно. Узкое терапевтическое окно в сочетании с тем фактом, что такие гормоны и пептиды обычно имеют короткий период полужизни, что требует частого введения дозы для достижения клинического эффекта, приводит к трудностям в лечении таких пациентов. Таким образом, сохраняется потребность в терапевтических средствах с повышенной эффективностью и безопасностью при лечении метаболических заболеваний.

[146] Таким образом, одним аспектом настоящего изобретения является включение биологически активных метаболических белков, и их участие или применение при лечении метаболических и сердечно-сосудистых заболеваний и нарушений, в слитые белки BPXTEN для создания композиций, пригодных для лечения таких нарушений, заболевания и сопутствующих состояний. Метаболические белки могут включать любой белок, предусматривающий биологические, терапевтические или профилактические интерес или функцию, который полезен для предупреждения, лечения, преобразования или облегчения метаболического или сердечно-сосудистого заболевания, нарушения или состояния. В **таблице 4d** представлен неограничивающий перечень таких

последовательностей метаболических BP, которые включаются в слитые белки BPXTEN по настоящему изобретению. Метаболические белки композиций на основе BPXTEN по настоящему изобретению могут представлять собой белок, который проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с последовательностью белка, выбранной из **таблицы 4d**.

Таблица 4d. Биологически активные белки для метаболических нарушений и кардиологии

Название белка (Синоним)	**Последовательность**	**SEQ ID NO:**
Антитело к CD3	*См.* патенты США №№ 5885573 и 6491916	
IL-1ra, человек, полной длины	MEICRGLRSHLITLLLFLFHSETICRPSGRKSSKMQAFRIWDV NQKTFYLRNNQLVAGYLQGPNVNLEEKIDVVPIEPHALFLGI HGGKMCLSCVKSGDETRLQLEAVNITDLSENRKQDKRFAFI RSDSGPTTSFESAACPGWFLCTAMEADQPVSLTNMPDEGVM VTKFYFQEDE	1723
IL-1ra, собака	METCRCPLSYLISFLLFLPHSETACRLGKRPCRMQAFRIWDV NQKTFYLRNNQLVAGYLQGSNTKLEEKLDVVPVEPHAVFL GIHGGKLCLACVKSGDETRLQLEAVNITDLSKNKDQDKRFT FILSDSGPTTSFESAACPGWFLCTALEADRPVSLTNRPEEAM MVTKFYFQKE	1724
IL-1ra, кролик	MRPSRSTRRHLISLLLFLFHSETACRPSGKRPCRMQAFRIWD VNQKTFYLRNNQLVAGYLQGPNAKLEERIDVVPLEPQLLFL GIQRGKLCLSCVKSGDKMKLHLEAVNITDLGKNKEQDKRFT FIRSNSGPTTTFESASCPGWFLCTALEADQPVSLTNTPDDSIV VTKFYFQED	1725
IL-1ra, крыса	MEICRGPYSHLISLLLILLFRSESAGHIPAGKRPCKMQAFRIW DTNQKTFYLRNNQLIAGYLQGPNTKLEEKIDMVPIDFRNVFL GIHGGKLCLSCVKSGDDTKLQLEEVNITDLNKNKEEDKRFTF IRSETGPTTSFESLACPGWFLCTTLEADHPVSLTNTPKEPCTV TKFYFQED	1726
IL-1ra, мышь	MEICWGPYSHLISLLLILLFHSEAACRPSGKRPCKMQAFRIW DTNQKTFYLRNNQLIAGYLQGPNIKLEEKIDMVPIDLHSVFL GIHGGKLCLSCAKSGDDIKLQLEEVNITDLSKNKEEDKRFTFI RSEKGPTTSFESAACPGWFLCTTLEADRPVSLTNTPEEPLIVT KFYFQEDQ	1727
Анакинра	MRPSGRKSSKMQAFRIWDVNQKTFYLRNNQLVAGYLQGPN VNLEEKIDVVPIEPHALFLGIHGGKMCLSCVKSGDETRLQLE AVNITDLSENRKQDKRFAFIRSDSGPTTSFESAACPGWFLCT AMEADQPVSLTNMPDEGVMVTKFYFQEDE	1728
α-натрийуретический пептид (ANP)	SLRRSSCFGGRMDRIGAQSGLGCNSFRY	1729
β-	SPKMVQGSGGFGRKMDRISSSSGLGCKVLRRH	1730

натрийуретический пептид, человек (BNP человека)		
Натрийуретический пептид головного мозга, крыса; (BNP крысы)	NSKMAHSSSCFGQKIDRIGAVSRLGCDGLRLF	1731
Натрийуретический пептид С-типа (CNP, свиной)	GLSKGCFGLKLDRIGSMSGLGC	1732
Фактор 2 роста фибробластов (FGF-2)	PALPEDGGSGAFPPGHFKDPKRLYCKNGGFFLRIHPDGRVD GVREKSDPHIKLQLQAEERGVVSIKGVCANRYLAMKEDGRL LASKCVTDECFFFERLESNNYNTYRS RKYTSWYVAL KRTGQYKLGS KTGPGQKAIL FLPMSAKS	1733
Рецептор TNF (TNFR)	LPAQVAFTPYAPEPGSTCRLREYYDQTAQMCCSKCSPGQHA KVFCTKTSDTVCDSCEDSTYTQLWNWVPECLSCGSRCSSDQ VETQACTREQNRICTCRPGWYCALSKQEGCRLCAPLRKCRP GFGVARPGTETSDVVCKPCAPGTFSNTTSSTDICRPHQICNV VAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQ PTPEPSTAPSTSFLLPMGPSPPAEGSTGD	1734

[147] «Антитело к CD3» означает моноклональное антитело к поверхностному белку Т-клеток CD3, его виды и варианты последовательности, а также их фрагменты, в том числе OKT3 (также называемый муромонабом) и гуманизированное моноклональное антитело к CD3 (hOKT31(Ala-Ala)) (Herold et al., 2002, New England Journal of Medicine 346:1692-1698). Содержащие антитело к CD3 слитые белки по настоящему изобретению могут найти конкретное применение для замедления впервые выявленного диабета 1 типа, в том числе применение антитела к CD3 в качестве терапевтического эффектора, а также в качестве нацеливающего фрагмента для второго терапевтического ВР в композиции на основе ВРХТЕN. Последовательности для вариабельной области и создание антител к CD3 описаны в патентах США №№ 5885573 и 6491916.

[148] «IL-1ra» означает человеческий белок-антагонист рецептора IL-1 и его виды и варианты последовательности, в том числе вариант последовательности анакинры (Kineret®), обладающий по меньшей мере частью биологической активности зрелого IL-1ra. Анакинра представляет собой негликозилированный рекомбинантный человеческий IL-1ra и отличается от эндогенного человеческого IL-1ra добавлением N-концевого метионина. Коммерческая версия анакинры продается как Kineret®. Она связывается с рецептором IL-1 с той же авидностью, что и нативный IL-1ra и IL-1b, но не приводит к активации рецептора (сигнальной трансдукции), эффект, связанный с наличием только одного мотива связывания рецептора на IL-1ra по сравнению с двумя такими мотивами на

IL-1α и IL-1β. Анакинра содержит 153 аминокислоты и имеет размер 17,3 кДа, а период полужизни составляет приблизительно 4-6 часов.

[149] Сообщалось об увеличении степени продуцирования IL-1 у пациентов с различными микробными инфекционными заболеваниями и рядом других заболеваний. Содержащие IL-1ra слитые белки по настоящему изобретению могут найти конкретное применение при лечении любого из вышеперечисленных заболеваний и нарушений. IL-1ra был клонирован, как описано в патенте США №№ 5075222 и 6858409.

[150] «Натрийуретические пептиды» означают предсердный натрийуретический пептид (ANP), натрийуретический пептид головного мозга (BNP или натрийуретический пептид B-типа) и натрийуретический пептид C-типа (CNP); как человеческие, так и отличные от человеческих виды и варианты их последовательностей, обладающие по меньшей мере частью биологической активности зрелых аналоговых натрийуретических пептидов. Последовательности пригодных форм натрийуретических пептидов раскрыты в публикации патента США 20010027181. Примеры ANP включают человеческий ANP (Kangawa et al., 1984, BBRC 118:131) или ANP различных видов, в том числе ANP свиней и крыс (Kangawa et al., 1984, BBRC 121:585). Анализ последовательности показал, что препро-BNP состоит из 134 остатков и расщепляется до про-BNP из 108 аминокислот. Отщепление последовательности из 32 аминокислот с C-конца про-BNP дает в результате человеческий BNP (77-108), который является физиологически активной формой в кровотоке. BNP человека из 32 аминокислот включает образование дисульфидной связи (Sudoh et al., 1989, BBRC 159:1420) и патенты США №№ 5114923, 5674710, 5674710 и 5948761. BPXTEN, предусматривающие одну или более натрийуретических функций, могут быть пригодны при лечении гипертонии, индукции диуреза, индукции натрийуреза, расширении или расслаблении сосудистого русла, связывании рецепторов натрийуретического пептида (таких как NPR-A), подавлении секреции альдостерона из надпочечников, лечении сердечно-сосудистых заболеваний и нарушений, уменьшении, прекращении или реверсии ремоделирования сердца после сердечного события или в результате застойной сердечной недостаточности, лечении почечных заболеваний и нарушений; лечении или предупреждении ишемического инсульта и лечении астмы.

[151] «Связывающий гепарин фактор роста 2» или «FGF-2» означает белок FGF-2 человека, а также его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности зрелого аналога. FGF-2 был клонирован, как описано в Burgess, W. H. and Maciag, T., Ann. Rev. Biochem., 58:575-606 (1989); Coulier, F., et al., 1994, Prog. Growth Factor Res. 5:1; и публикации согласно PCT № WO 87/01728.

[152] «Рецептор TNF» означает человеческий рецептор TNF, а также его виды и варианты последовательности, обладающие по меньшей мере частью биологической рецепторной активности зрелого TNFR. Определена рентгеновская кристаллическая структура комплекса, образованного внеклеточным доменом человеческого рецептора p55 TNF и TNFβ (Banner et al., 1993 Cell 73:431, включенная в данный документ посредством ссылки).

Факторы коагуляции

[153] При гемофилии свертываемость крови нарушается из-за недостатка в плазме крови определенных факторов свертывания крови. Человеческий фактор IX (FIX) представляет собой зимоген сериновой протеазы, которая является важным компонентом внутреннего пути каскада свертывания крови. Белки фактора VIIa (FVIIa) нашли применение в лечении эпизодов кровотечения у пациентов с гемофилией A или B с ингибиторами FVIII или FIX и у пациентов с приобретенной гемофилией, а также для предупреждения кровотечения при хирургических вмешательствах или инвазивных процедурах у пациентов с гемофилией A или B с ингибиторами FVIII или FIX. Таким образом, сохраняется потребность в композициях фактора IX и фактора VIIa с увеличенным периодом полужизни и сохранением активности при введении в рамках профилактической и/или терапевтической схемы лечения гемофилии B, а также в составах, которые уменьшают побочные эффекты и которые можно вводить как внутривенным, так и подкожным путями.

[154] Факторы коагуляции для включения в BPXTEN по настоящему изобретению могут предусматривать белки, предусматривающие биологические, терапевтические или профилактические интерес или функцию, которые полезны для предупреждения, лечения, преобразования или облегчения нарушений, заболеваний или недостаточности коагуляции крови. Подходящие белки коагуляции включают биологически активные полипептиды, которые участвуют в каскаде коагуляции в качестве субстратов, ферментов или кофакторов.

[155] В **таблице 4e** представлен неограничивающий перечень последовательностей факторов коагуляции, которые включаются в слитые белки BPXTEN по настоящему изобретению. Факторами коагуляции для включения в BPXTEN по настоящему изобретению может быть белок, который проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с последовательностью белка, выбранной из **таблицы 4e**.

Таблица 4e. Полипептидные последовательности фактора коагуляции

Название BPXTEN	**SEQ ID NO:**	**Аминокислотная последовательность**
Предшественник FIX	1735	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNR PKRYNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTE FWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKN CELDVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKS CEPAVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILD NITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCG GSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRN VIRIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKE YTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATC LRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSF LTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT

FIX, Homo sapiens	1736	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTN IFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLRS TKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGI ISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 4 из патента США № 20080214462	1737	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNR PKRYNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTE FWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKN CELDVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKS CEPAVPFPCGRVSVSQTSKLTRAEAVFPDVDYVNSTEAETILD NITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCG GSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRN VIRIIPHHNFNAAINTYNHDIALLELDEPLVNSYVTPICIADKEY TNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCL RSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFL TGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 6 из патента США № 20080214462	1738	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNR PKRYNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTE FWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKN CELDVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKS CEPAVPFPCGRVSVSQTSKLTRAEAVFPDVDYVNSTEAETILD NITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCG GSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRN VIRIIPHHNFNAAINTYNHDIALLELDEPLVNSYVTPICIADKEY TNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCL RSTKFTIFNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLT GIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 8 из патента США № 20080214462	1739	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNR PKRYNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTE FWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKN CELDVTCNIKNGRCEQFCKNSADNKVVCSCTEG YRLAENQKSCEPAVPFPCGRVSVSQTSKLTRAEAVFPDVDYV NSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLN GKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEET EHTEQKRNVIRIIPHHNFNAAINTYNHDIALLELDEPLVLNSYV TPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRV PLVDATCLRSTKFTIFNNMFCAGFHEGGRDSCQGDSGGPHVT EVEGTSFLTGIVSWGEGCAMKGKYGIYTKVSRYVNWIKEKTK LT
Последовательность 2 из патента США № 7125841	1740	MQRVNMIMAESPSLITICLLGYLLSAECTVFLDHENANKILNR PKRYNSGKLEEFVQGNLERECMEEKCSFEEPREVFENTEKITE FWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKN CELDVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKS CEPAVPFPCGRVSVSQTSKLTRAEAVFPDVDYVNPTEAETILD NITQGTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCG GSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRN VIRAIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIAD

		KEYTNIFLKFGSGYVSGWARVFHKGRSALVLQYLRVPLVDRA TCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGT SFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 1 из патента США № 20080167219	1741	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DATCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTN IFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLRS TKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGI ISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 2 из патента США № 20080167219	1742	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDAPLVLNSYVTPICIADKEYT NIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLR STKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLT GIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 3 из патента США № 20080167219	1743	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DATCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDAPLVLNSYVTPICIADKEYT NIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLR STKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLT GIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 4 из патента США № 20080167219	1744	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DATCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTN IFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLAS TKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGI ISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 5 из патента США № 20080167219	1745	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDAPLVLNSYVTPICIADKEYT NIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLA STKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLT

		GIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Последовательность 6 из патента США № 20080167219	1746	YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFW KQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCEL DATCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEP AVPFPCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNIT QSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFCGGSI VNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDAPLVLNSYVTPICIADKEYT NIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLVDRATCLA STKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLT GIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
Фактор VII/ VIIa	1747	ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFWISY SDGDQCASSPCQNGGSCKDQLQSYICFCLPAFEGRNCETHKD DQLICVNENGGCEQYCSDHTGTKRSCRCHEGYSLLADGVSCT PTVEYPCGKIPILEKRNASKPQGRIVGGKVCPKGECPWQVLLL VNGAQLCGGTLINTIWVVSAAHCFDKIKNWRNLIAVLGEHDL SEHDGDEQSRRVAQVIIPSTYVPGTTNHDIALLRLHQPVVLTD HVVPLCLPERTFSERTLAFVRFSLVSGWGQLLDRGATALELM VLNVPRLMTQDCLQQSRKVGDSPNITEYMFCAGYSDGSKDSC KGDSGGPHATHYRGTWYLTGIVSWGQGCATVGHFGVYTRVS QYIEWLQKLMRSEPRPGVLLRAPFP

[156] «Фактор IX» («FIX») предусматривает белок человеческого фактора IX и его виды и варианты последовательности, обладающие по меньшей мере частью биологической рецепторной активности зрелого фактора IX. В некоторых вариантах осуществления пептид FIX представляет собой структурный аналог или пептидный миметик любого из пептидов FIX, описанных в данном документе, включая последовательности из **таблицы 4е**. В некоторых вариантах осуществления пептид FIX представляет собой структурный аналог или пептидный миметик любого из пептидов FIX, описанных в данном документе, включая последовательности из **таблицы 4е**. В одном конкретном примере настоящего изобретения FIX представляет собой FIX человека. В другом варианте осуществления FIX представляет собой полипептидную последовательность из **таблицы 4е**. Зрелый фактор IX представляет собой одноцепочечный белок из 415 аминокислотных остатков, который содержит приблизительно 17% углеводов по массе (Schmidt 2003, Trends Cardiovasc Med, 13: 39).

[157] В некоторых случаях фактор коагуляции представляет собой фактор IX, вариант последовательности фактора IX или фрагмент фактора IX, такие как иллюстративные последовательности в **таблице 4е**, а также любой белок или полипептид, по сути гомологичный им, чьи биологические свойства приводят к активности фактора IX.

[158] «Фактор VII» (FVII) означает белок человека, а также его виды и варианты последовательности, обладающие по меньшей мере частью биологической активности активированного фактора VII. Фактор VII и рекомбинантный FVIIa человека были введены для применения при неконтролируемом кровотечении у пациентов с гемофилией (с дефицитом фактора VIII или IX), у которых выработались ингибиторы замещаемого фактора коагуляциии. Рекомбинантный человеческий фактор VIIa полезен при лечении неконтролируемого кровотечения у пациентов с гемофилией (с дефицитом фактора VIII

или IX), включая тех, у кого выработались ингибиторы замещаемого фактора коагуляциии. В некоторых вариантах осуществления пептид FVII представляет собой активированную форму (FVIIa), структурный аналог или пептидный миметик любого из описанных в данном документе пептидов FVII, включая последовательности из **таблицы 4e**. Факторы VII и VIIa были клонированы, как описано в патенте США № 6806063 и заявке на патент США № 20080261886.

Белки гормона роста

[159] «Гормон роста» или «GH» означает белок гормона роста человека и его виды и варианты последовательности и включает без ограничения одноцепочечную последовательность человеческого GH из 191 аминокислоты. Настоящее изобретение предполагает включение в BPXTEN любых гомологичных GH последовательностей, фрагментов последовательностей, которые являются природными, например от приматов, млекопитающих (включая домашних животных), и не являющихся природными вариантов последовательностей, которые сохраняют по меньшей мере часть биологической активности или биологической функции GH, и/или которые применимы для предупреждения, лечения, преобразования или облегчения связанного с GH заболевания, дефицита, нарушения или состояния. Последовательности GH, не относящиеся к млекопитающим, хорошо описаны в литературе. Например, выравнивание последовательностей GH рыб можно найти в Genetics and Molecular Biology **2003** 26 p.295-300. Кроме того, нативные последовательности, гомологичные человеческому GH, можно найти при помощи стандартных методик для поиска гомологии, таких как NCBI BLAST.

[160] В одном варианте осуществления GH, включенный в композиции для человека или животного, может представлять собой рекомбинантный полипептид с последовательностью, соответствующей белку, встречающемуся в природе. В другом варианте осуществления GH может представлять собой вариант последовательности, фрагмент, гомолог или миметик природной последовательности, который сохраняет по меньшей мере часть биологической активности нативного GH. В **таблице 4f** представлен не ограничивающий перечень последовательностей GH от широкого круга видов млекопитающих, которые включаются в слитые белки BPXTEN по настоящему изобретению. Любая из этих последовательностей GH или гомологичных производных, сконструированных путем перетасовки индивидуальных мутаций между видами или семействами, может быть применима для слитых белков по настоящему изобретению. GH, который может быть включен в слитый белок BPXTEN, может предусматривать белок, который проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с белком, выбранным из **таблицы 4f.**

Таблица 4f. Аминокислотные последовательности гормона роста различных видов животных

Виды GH	**Аминокислотная последовательность**	**SEQ**

		ID NO:
Человек	FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSF LQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPV QFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSP RTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVET FLRIVQCRSVEGSCGF	1750
Свинья	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1751
Альпака	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERTYIPEGQRYSI QNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLGP VQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILRQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1752
Верблюд	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERTYIPEGQRYSI QNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLGP VQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILRQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1753
Лошадь	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDMELLRFSLLLIQSWLG PVQLLSRVFTNSLVFGTSDRVYEKLRDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1754
Слон	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RPGQVLKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1755
Рыжая лисица	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLVLIQSWLG PLQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1756
Собака	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1757
Кошка	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RGGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1758
Американская норка	FPAMPLSSLFANAVLRAQHLHQLAADTYKDFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDMELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGPILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAETY LRV MKCRRFVESSCAF	1759

Финвал	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1760
Дельфин	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNTQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLGP VQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1761
Гиппопотам	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNTQAAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLGP VQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGSP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1762
Кролик	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDMELLRFSLLLIQSWLG PVQFLSRAFTNTLVFGTSDRVYEKLKDLEEGIQALMRELEDGS PRVGQLLKQTYDKFDTNLRGDDALLKNYGLLSCFKKDLHKAE TYLRV MKCRRFVESSCVF	1763
Крыса	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKEEAQQRTDMELLRFSLLLIQSWLG PVQFLSRIFTNSLMFGTSDRVYEKLKDLEEGIQALMQELEDGSP RIGQILKQTYDKFDANMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFAESSCAF	1764
Мышь	FPAMPLSSLFSNAVLRAQHLHQLAADTYKEFERAYIPEGQRYSI QNAQAAFCFSETIPAPTGKEEAQQRTDMELLRFSLLLIQSWLGP VQFLSRIFTNSLMFGTSDRVYEKLKDLEEGIQALMQELEDGSPR VGQILKQTYDKFDANMRSDDALLKNYGLLSCFKKDLHKAETY LRV MKCRRFVESSCAF	1765
Хомяк	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQTAFCFSETIPAPTGKEEAQQRSDMELLRFSLLLIQSWLGP VQFLSRIFTNSLMFGTSDRVYEKLKDLEEGIQALMQELEDGSPR VGQILKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAETY LRV MKCRRFVESSCAF	1766
Слепыш	FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKEEAQQRSDMELLRFSLLLIQSWLG PVQFLSRVFTNSLVFGTSDRVFEKLKDLEEGIQALMRELEDGSL RAGQLLKQTYDKFDTNMRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1767
Морская свинка	FPAMPLSSLFGNAVLRAQHLHQLAADTYKEFERTYIPEGQRYSI HNTQTAFCFSETIPAPTDKEEAQQRSDVELLHFSLLLIQSWLGP VQFLSRVFTNSLVFGTSDRVYEKLKDLEEGIQALMRELEDGTP RAGQILKQTYDKFDTNLRSNDALLKNYGLLSCFRKDLHRTETY LRV MKCRRFVESSCAF	1768
Бык	AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRY SIQNTQVAFCFSETIPAPTGKNEAQQKSDLELLRISLLLIQSWLG PLQFLSRVFTNSLVFGTSDRVYEKLKDLEEGILALMRELEDGTP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTET YLRV MKCRRFGEASCAF	1769
Овца/ коза	AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRY SIQNTQVAFCFSETIPAPTGKNEAQQKSDLELLRISLLLIQSWLG	1770

	PLQFLSRVFTNSLVFGTSDRVYEKLKDLEEGILALMRELEDVTP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTET YLRV MKCRRFGEASCAF	
Благородный олень	FPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSI QNTQVAFCFSETIPAPTGKNEAQQKSDLELLRISLLLIQSWLGPL QFLSRVFTNSLVFGTSDRVYEKLKDLEEGILALMRELEDGTPR AGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETY LRV MKCRRFGEASCAF	1771
Жираф	AFPAMSLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRY SIQNTQVAFCFSETIPAPTGKNEAQQKSDLELLRISLLLIQSWLG PLQFLSRVFSNSLVFGTSDRVYEKLKDLEEGILALMRELEDGTP RAGQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTET YLRV MKCRRFGEASCAF	1772
Оленек-1	FPAMSLSGLFANAVLRVQHLHQLAADTFKEFERTYIPEGQRYSI QNTQVAFCFSETIPAPTGKNEAQQKSDLELLRISLLLIQSWLGPL QFLSRVFTNSLVFGTSDRVYEKLKDLEEGILALMRELEDGPPRA GQILKQTYDKFDTNMRSDDALLKNYGLLSCFRKDLHKTETYL RV MKCRRFGEASCAF	1773
Толстый лори	FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYS IQNAQAAFCFSETIPAPTGKDEAQQRSDMELLRFSLLLIQSWLG PVQLLSRVFTNSLVLGTSDRVYEKLKDLEEGIQALMRELEDGS PRVGQILKQTYDKFDTNLRSDDALLKNYGLLSCFKKDLHKAET YLRV MKCRRFVESSCAF	1774
Игрунка	FPTIPLSRLLDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSF LQNPQTSLCFSESIPTPASKKETQQKSNLELLRMSLLLIQSWFEP VQFLRSVFANSLLYGVSDSDVYEYLKDLEEGIQTLMGRLEDGS PRTGEIFMQTYRKFDVNSQNNDALLKNYGLLYCFRKDMDKVE TFLRI VQCR-SVEGSCGF	1775
Щекохвостый поссум	FPAMPLSSLFANAVLRAQHLHQLVADTYKEFERTYIPEAQRHSI QSTQTAFCFSETIPAPTGKDEAQQRSDVELLRFSLLLIQSWLSPV QFLSRVFTNSLVFGTSDRVYEKLRDLEEGIQALMQELEDGSSR GGLVLKTTYDKFDTNLRSDEALLKNYGLLSCFKKDLHKAETY LRV MKCRRFVESSCAF	1776
Обезьяна (резус)	FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSF LQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPV QFLRSVFANSLVYGTSYSDVYDLLKDLEEGIQTLMGRLEDGSS RTGQIFKQTYSKFDTNSHNNDALLKNYGLLYCFRKDMDKIETF LRI VQCR-SVEGSCGF	1777

Цитокины

[161] BP может представлять собой цитокин или один или более цитокинов. Цитокины относятся к белкам (например, хемокины, интерфероны, лимфокины, интерлейкины и факторы некроза опухоли), высвобождаемым клетками, которые могут влиять на поведение клеток. Цитокины могут продуцироваться широким спектром клеток, в том числе иммунными клетками, такими как макрофаги, B-лимфоциты, T-лимфоциты и тучные клетки, а также эндотелиальными клетками, фибробластами и различными стромальными клетками. Заданный цитокин может продуцироваться более чем одним типом клеток. Цитокины могут участвовать в выработке системных или местных иммуномодулирующих эффектов.

[162] Определенные цитокины могут действовать как провоспалительные

цитокины. Провоспалительные цитокины относятся к цитокинам, участвующим в индукции или усилении воспалительной реакции. Провоспалительные цитокины могут работать с различными клетками иммунной системы, такими как нейтрофилы и лейкоциты, для создания иммунного ответа. Определенные цитокины могут действовать как противовоспалительные цитокины. Противовоспалительные цитокины относятся к цитокинам, участвующим в уменьшении воспалительной реакции. В некоторых случаях противовоспалительные цитокины могут регулировать ответ провоспалительных цитокинов. Некоторые цитокины могут действовать в качестве как про-, так и противовоспалительных цитокинов.

[163] Цитокины, входящие в состав композиций по настоящему изобретению, могут быть полезны при лечении в различных терапевтических категориях или категориях заболеваний, включая без ограничения рак, ревматоидный артрит, рассеянный склероз, тяжелую миастению, системную красную волчанку, болезнь Альцгеймера, шизофрению, вирусные инфекции (например, хронический гепатит С, СПИД), аллергическую астму, нейродегенеративные процессы сетчатки, метаболические нарушения, инсулинорезистентность и диабетическую кардиомиопатию. Цитокины могут быть особенно полезны при лечении воспалительных состояний и аутоиммунных состояний.

[164] Примеры цитокинов, которые регулируются системами и композициями по настоящему изобретению, включают без ограничения лимфокины, монокины и традиционные полипептидные гормоны, за исключением гормона роста человека. В число цитокинов входят паратиреоидный гормон; тироксин; инсулин; проинсулин; релаксин; прорелаксин; гликопротеиновые гормоны, такие как фолликулостимулирующий гормон (FSH), тиреотропный гормон (TSH) и лютеинизирующий гормон (LH); фактор роста печени; фактор роста фибробластов; пролактин; плацентарный лактоген; фактор некроза опухоли-альфа; мюллеров ингибирующий фактор; мышиный гонадотропин-ассоциированный пептид; ингибин; активин; фактор роста эндотелия сосудов; интегрин; тромбопоэтин (TPO); факторы роста нервов, такие как NGF-альфа; фактор роста тромбоцитов; трансформирующие факторы роста (TGF), такие как TGF-альфа, TGF-бета, TGF-бета1, TGF-бета2 и TGF-бета3; инсулиноподобный фактор роста-I и -II; эритропоэтин (EPO); Flt-3L; фактор стволовых клеток (SCF); остеоиндуктивные факторы; интерфероны (IFN), такие как IFN-α, IFN-β, IFN-γ; колониестимулирующие факторы (CSF), такие как макрофагальный CSF (M-CSF); гранулоцитарно-макрофагальный CSF (GM-CSF); гранулоцитарный CSF (G-CSF); фактор стимуляции макрофагов (MSP); интерлейкины (IL) такие как IL-1, IL-1a, IL-1b, IL-1RA, IL-18, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-12b, IL-13, IL-14, IL-15, IL-16, IL-17, IL-20; фактор некроза опухоли, такой как CD154, LT-бета, TNF-альфа, TNF-бета, 4-1BBL, APRIL, CD70, CD153, CD178, GITRL, LIGHT, OX40L, TALL-1, TRAIL, TWEAK, TRANCE; и другие полипептидные факторы, включая LIF, онкостатин M (OSM) и kit-лиганд (KL). Цитокиновые рецепторы относятся к рецепторным белкам, которые связывают цитокины. Цитокиновые рецепторы могут быть как мембраносвязанными, так и растворимыми.

[165] Полинуклеотид-мишень может кодировать цитокин. Неограничивающие примеры цитокинов включают 4-1BBL, активин βA, активин βB, активин βC, активин βE, артемин (ARTN), BAFF/BLyS/TNFSF138, BMP10, BMP15, BMP2, BMP3, BMP4, BMP5, BMP6, BMP7, BMP8a, BMP8b, костный морфогенический белок 1 (BMP1), CCL1/TCA3, CCL11, CCL12/MCP-5, CCL13/MCP-4, CCL14, CCL15, CCL16, CCL17/TARC, CCL18, CCL19, CCL2/MCP-1, CCL20, CCL21, CCL22/MDC, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CCL3, CCL3L3, CCL4, CCL4L1/LAG-1, CCL5, CCL6, CCL7, CCL8, CCL9, CD153/CD30L/TNFSF8, CD40L/CD154/TNFSF5, CD40LG, CD70, CD70/CD27L/TNFSF7, CLCF1, c-MPL/CD110/TPOR, CNTF, CX3CL1, CXCL1, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, CXCL17, CXCL2/MIP-2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7/Ppbp, CXCL9, EDA-A1, FAM19A1, FAM19A2, FAM19A3, FAM19A4, FAM19A5, Fas-лиганд/FASLG/CD95L/CD178, GDF10, GDF11, GDF15, GDF2, GDF3, GDF4, GDF5, GDF6, GDF7, GDF8, GDF9, нейротрофический фактор глиальной клеточной линии (GDNF), фактор роста и дифференцировки 1 (GDF1), IFNA1, IFNA10, IFNA13, IFNA14, IFNA2, IFNA4, IFNA5/IFNaG, IFNA7, IFNA8, IFNB1, IFNE, IFNG, IFNZ, IFNω/IFNW1, IL11, IL18, IL18BP, IL1A, IL1B, IL1F10, IL1F3/IL1RA, IL1F5, IL1F6, IL1F7, IL1F8, IL1F9, IL1RL2, IL31, IL33, IL6, IL8/CXCL8, ингибин-A, ингибин-B, лептин, LIF, LTA/TNFB/TNFSF1, LTB/TNFC, нейротурин (NRTN), OSM, OX-40L/TNFSF4/CD252, персефин (PSPN), RANKL/OPGL/TNFSF11(CD254), TL1A/TNFSF15, TNFA, TNF-альфа/TNFA, TNFSF10/TRAIL/APO-2L(CD253), TNFSF12, TNFSF13, TNFSF14/LIGHT/CD258, XCL1 и XCL2. В некоторых вариантах осуществления ген-мишень кодирует ингибитор иммунной контрольной точки. Неограничивающие примеры таких ингибиторов иммунной контрольной точки включают PD-1, CTLA-4, LAG3, TIM-3, A2AR, B7-H3, B7-H4, BTLA, IDO, KIR и VISTA. В некоторых вариантах осуществления ген-мишень кодирует альфа-, бета-, гамма- и/или дельта-цепь Т-клеточного рецептора (TCR).

[166] В некоторых случаях цитокин может представлять собой хемокин. Хемокин может быть выбран из группы, включающей без ограничения ARMCX2, BCA-1/CXCL13, CCL11, CCL12/MCP-5, CCL13/MCP-4, CCL15/MIP-5/MIP-1 дельта, CCL16/HCC-4/NCC4, CCL17/TARC, CCL18/PARC/MIP-4, CCL19/MIP-3b, CCL2/MCP-1, CCL20/MIP-3 альфа/MIP3A, CCL21/6Ckine, CCL22/MDC, CCL23/MIP 3, CCL24/эотаксин-2/MPIF-2, CCL25/TECK, CCL26/эотаксин-3, CCL27/CTACK, CCL28, CCL3/Mip1a, CCL4/MIP1B, CCL4L1/LAG-1, CCL5/RANTES, CCL6/C10, CCL8/MCP-2, CCL9, CML5, CXCL1, CXCL10/Crg-2, CXCL12/SDF-1 бета, CXCL14/BRAK, CXCL15/лунгкин, CXCL16/SR-PSOX, CXCL17, CXCL2/MIP-2, CXCL3/GRO гамма, CXCL4/PF4, CXCL5, CXCL6/GCP-2, CXCL9/MIG, FAM19A1, FAM19A2, FAM19A3, FAM19A4/TAFA4, FAM19A5, фракталкин/CX3CL1, I-309/CCL1/TCA-3, IL-8/CXCL8, MCP-3/CCL7, NAP-2/PPBP/CXCL7, XCL2 и Armo IL10.

[167] В **таблице 4g** представлен не ограничивающий перечень таких последовательностей BP, которые включаются в слитые белки BPXTEN по настоящему

изобретению. Метаболические белки композиций на основе BPXTEN по настоящему изобретению могут представлять собой белок, который проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с последовательностью белка, выбранной из **таблицы 4g**.

Таблица 4g. Цитокины для конъюгирования

Название белка (Синоним)	**Аминокислотная последовательность**	**SEQ ID NO:**
Антитело к CD3	*См.* патенты США №№ 5885573 и 6491916	
IL-1ra, человек, полной длины	MEICRGLRSHLITLLLFLFHSETICRPSGRKSSKMQAFRIWDV NQKTFYLRNNQLVAGYLQGPNVNLEEKIDVVPIEPHALFLGI HGGKMCLSCVKSGDETRLQLEAVNITDLSENRKQDKRFAFI RSDSGPTTSFESAACPGWFLCTAMEADQPVSLTNMPDEGVM VTKFYFQEDE	270
IL-1ra, собака	METCRCPLSYLISFLLFLPHSETACRLGKRPCRMQAFRIWDV NQKTFYLRNNQLVAGYLQGSNTKLEEKLDVVPVEPHAVFL GIHGGKLCLACVKSGDETRLQLEAVNITDLSKNKDQDKRFT FILSDSGPTTSFESAACPGWFLCTALEADRPVSLTNRPEEAM MVTKFYFQKE	271
IL-1ra, кролик	MRPSRSTRRHLISLLLFLFHSETACRPSGKRPCRMQAFRIWD VNQKTFYLRNNQLVAGYLQGPNAKLEERIDVVPLEPQLLFL GIQRGKLCLSCVKSGDKMKLHLEAVNITDLGKNKEQDKRFT FIRSNSGPTTTFESASCPGWFLCTALEADQPVSLTNTPDDSIV VTKFYFQED	272
IL-1ra, крыса	MEICRGPYSHLISLLLILLFRSESAGHIPAGKRPCKMQAFRIW DTNQKTFYLRNNQLIAGYLQGPNTKLEEKIDMVPIDFRNVFL GIHGGKLCLSCVKSGDDTKLQLEEVNITDLNKNKEEDKRFTF IRSETGPTTSFESLACPGWFLCTTLEADHPVSLTNTPKEPCTV TKFYFQED	273
IL-1ra, мышь	MEICWGPYSHLISLLLILLFHSEAACRPSGKRPCKMQAFRIW DTNQKTFYLRNNQLIAGYLQGPNIKLEEKIDMVPIDLHSVFL GIHGGKLCLSCAKSGDDIKLQLEEVNITDLSKNKEEDKRFTFI RSEKGPTTSFESAACPGWFLCTTLEADRPVSLTNTPEEPLIVT KFYFQEDQ	274
Анакинра	MRPSGRKSSKMQAFRIWDVNQKTFYLRNNQLVAGYLQGPN VNLEEKIDVVPIEPHALFLGIHGGKMCLSCVKSGDETRLQLE AVNITDLSENRKQDKRFAFIRSDSGPTTSFESAACPGWFLCT AMEADQPVSLTNMPDEGVMVTKFYFQEDE	275
IL-10	MHSSALLCCLVLLTGVRASPGQGTQSENSCTHFPGNLPNML RDLRDAFSRVKTFFQMKDQLDNLLLKESLLEDFKGYLGCQA LSEMIQFYLEEVMPQAENQDPDIKAHVNSLGENLKTLRLRL RRCHRFLPCENKSKAVEQVKNAFNKLQEKGIYKAMSEFDIFI NYIEAYMTMKIRN	276

[168] «IL-1ra» означает человеческий белок-антагонист рецептора IL-1 и его виды и варианты последовательности, в том числе вариант последовательности анакинры (Kineret®), обладающий по меньшей мере частью биологической активности зрелого IL-

1ra. Человеческий IL-1ra представляет собой зрелый гликопротеин, состоящий из 152 аминокислотных остатков. Содержащие IL-1ra слитые белки по настоящему изобретению могут найти конкретное применение при лечении любого из вышеперечисленных заболеваний и нарушений. IL-1ra был клонирован, как описано в патенте США №№ 5075222 и 6858409.

[169] В некоторых случаях BP может представлять собой IL-10. IL-10 может быть эффективным противовоспалительным цитокином, подавляющим продуцирование провоспалительных цитокинов и хемокинов. IL-10 может быть полезен для лечения аутоиммунных заболеваний и воспалительных заболеваний, таких как ревматоидный артрит, рассеянный склероз, тяжелая миастения, системная красная волчанка, болезнь Альцгеймера, шизофрения, аллергическая астма, нейродегенеративные процессы сетчатки и диабет.

[170] В некоторых случаях IL-10 можно модифицировать для повышения стабильности и уменьшения протеолитической деградации. Модификация может представлять собой замещение одной или нескольких амидных связей. В некоторых случаях одна или более амидных связей в основной цепи IL-10 могут быть замещены для достижения вышеупомянутых эффектов. Одна или более амидных связей (–CONH–) в IL-10 могут быть заменены связью, которая является изостерической по отношению к амидной связи, такой как $–CH_2NH–$, $–CH_2S–$, $–CH_2CH_2–$, –CH=CH– (цис и транс), $–COCH_2–$, $–CH(OH)CH_2–$ или $–CH_2SO–$. Кроме того, амидные связи в IL-10 также могут быть заменены восстановленной изостерической псевдопептидной связью. См. Couder et al. (1993) Int. J. Peptide Protein Res. 41:181-184, которая включена в данный документ посредством ссылки во всей своей полноте.

[171] Могут быть замещены одна или более кислых аминокислот, в том числе аспарагиновая кислота, глутаминовая кислота, гомоглутаминовая кислота, тирозин, алкил-, арил-, арилалкил- и гетероарилсульфонамиды 2,4-диаминопропионовой кислоты, орнитин или лизин и тетразолзамещенные алкиламинокислоты; и амидные остатки боковой цепи, такие как аспарагин, глутамин и алкильные или ароматические замещенные производные аспарагина или глутамина; а также гидроксилсодержащие аминокислоты, включая серин, треонин, гомосерин, 2,3-диаминопропионовую кислоту и алкильные или ароматические замещенные производные серина или треонина.

[172] Одна или более гидрофобных аминокислот в IL-10, таких как аланин, лейцин, изолейцин, валин, норлейцин, (S)-2-аминомасляная кислота, (S)-циклогексилаланин или другие простые альфа-аминокислоты, могут быть замещены аминокислотами, в том числе без ограничения с алифатической боковой цепью из C1-C10 атомов углерода, включая алкильные, алкенильные или алкинильные замещения с разветвленной, циклической и прямой цепью.

[173] В некоторых случаях одна или более гидрофобных аминокислот в IL-10, как, например, могут быть замещены заменой ароматически-замещенных гидрофобных аминокислот, включая фенилаланин, триптофан, тирозин, сульфотирозин,

бифенилаланин, 1-нафтилаланин, 2-нафтилаланин, 2-бензотиенилаланин, 3-бензотиенилаланин, гистидин, включая амино-, алкиламино-, диалкиламино-, аза-, галогенированные (фтор, хлор, бром или йод) или алкокси (C_1-C_4)-замещенные формы вышеперечисленных ароматических аминокислот, иллюстративными примерами которых являются: 2-, 3- или 4-аминофенилаланин, 2-, 3- или 4-хлорфенилаланин, 2-, 3- или 4-метилфенилаланин, 2-, 3- или 4-метоксифенилаланин, 5-амино-, 5-хлор-, 5-метил- или 5-метокситриптофан, 2'-, 3'- или 4'-амино-, 2'-, 3'- или 4'-хлор-, 2-, 3- или 4-бифенилаланин, 2'-, 3'- или 4'-метил-, 2-, 3- или 4-бифенилаланин и 2- или 3-пиридилаланин.

[174] Одна или более гидрофобных аминокислот в IL-10, таких как фенилаланин, триптофан, тирозин, сульфотирозин, бифенилаланин, 1-нафтилаланин, 2-нафтилаланин, 2-бензотиенилаланин, 3-бензотиенилаланин, гистидин, в том числе амино-, алкиламино-, диалкиламино-, аза-, галогенированные (фтор, хлор, бром или йод) или алкокси-, могут быть замещены ароматическими аминокислотами, включая: 2-, 3- или 4-аминофенилаланин, 2-, 3- или 4-хлорфенилаланин, 2-, 3- или 4-метилфенилаланин, 2-, 3- или 4-метоксифенилаланин, 5-амино-, 5-хлор-, 5-метил- или 5-метокситриптофан, 2'-, 3'- или 4'-амино-, 2'-, 3'- или 4'-хлор-, 2-, 3- или 4-бифенилаланин, 2'-, 3'- или 4'-метил-, 2-, 3- или 4-бифенилаланин и 2- или 3-пиридилаланин.

[175] Могут быть замещены аминокислоты, содержащие основные боковые цепи, в том числе аргинин, лизин, гистидин, орнитин, 2,3-диаминопропионовая кислота, гомоаргинин, в том числе алкильные, алкенильные или арилзамещенные производные предыдущих аминокислот. Примерами являются N-эпсилон-изопропил-лизин, 3-(4-тетрагидропиридил)-глицин, 3-(4-тетрагидропиридил)-аланин, N, N-гамма, гамма'-диэтилгомоаргинин, альфа-метил-аргинин, альфа-метил-2,3-диаминопропионовая кислота, альфа-метилгистидин и альфа-метилорнитин, где алкильная группа занимает про-(R) положение альфа-углерода. Модифицированный IL-10 может включать амиды, образованные из любой комбинации алкильных, ароматических, гетероароматических производных, орнитина или 2,3-диаминопропионовой кислоты, карбоновых кислот или любых из многих хорошо известных активированных производных, таких как хлорангидриды, активные сложные эфиры, активные азолиды и родственные производные, лизин и орнитин.

[176] В некоторых случаях IL-10 может содержать одну или более встречающихся в природе L-аминокислот, синтетических L-аминокислот и/или D-энантиомеров аминокислоты. Полипептид IL-10 может содержать одну или более из следующих аминокислот: ω-аминодекановая кислота, ω-аминотетрадекановая кислота, циклогексилаланин, α,γ-диаминомасляная кислота, α,β-диаминопропионовая кислота, δ-аминовалериановая кислота, трет-бутилаланин, трет-бутилглицин, N-метилизолейцин, фенилглицин, циклогексилаланин, норлейцин, нафтилаланин, орнитин, цитруллин, 4-хлорфенилаланин, 2-фторфенилаланин, пиридилаланин, 3-бензотиенилаланин, гидроксипролин, β-аланин, о-аминобензойная кислота, м-аминобензойная кислота, п-аминобензойная кислота, м-аминометилбензойная кислота, 2,3-диаминопропионовая

кислота, α-аминоизомасляная кислота, N-метилглицин (саркозин), 3-фторфенилаланин, 4-фторфенилаланин, пеницилламин, 1,2,3,4-тетрагидроизохинолин-3-карбоновая кислота, β-2-тиенилаланин, метионинсульфоксид, гомоаргинин, N-ацетиллизин, 2,4-диаминомасляная кислота, ро-аминофенилаланин, N-метилвалин, гомоцистеин, гомосерин, ε-аминогексановая кислота, ω-аминогексановая кислота, ω-аминогептановая кислота, ω-аминооктановая кислота и 2,3-диаминомасляная кислота.

[177] IL-10 может содержать остаток цистеина или цистеин, который может действовать в качестве линкера в отношении другого пептида через дисульфидную связь или обеспечивать циклизацию полипептида IL-10. Способы введения цистеина или аналога цистеина известны в уровне техники; см., например, патент США № 8067532. Полипептид IL-10 может быть подвергнут циклизации. Другие способы циклизации включают введение оксимного линкера или лантионинового линкера; см., например, патент США № 8044175. Можно использовать и/или вводить любую комбинацию аминокислот (или фрагментов, не являющихся аминокислотными), которая может образовывать циклизирующую связь. Циклизирующая связь может быть получена с помощью любой комбинации аминокислот (или с аминокислотой и $–(CH_2)_nCO–$ или $–(CH_2)_nC_6H_4–CO–$) с функциональными группами, которые допускают введение мостика. Некоторыми примерами являются дисульфиды, дисульфидные миметики, такие как $–(CH_2)_n$-карбо-мостик, тиоацеталь, тиоэфирные мостики (цистатионин или лантионин) и мостики, содержащие сложные и простые эфиры.

[178] IL-10 может быть замещен N-алкильной, арильной или поперечной сшивкой основной цепи для создания лактамов и других циклических структур, C-концевых гидроксиметильных производных, о-модифицированных производных, N-концевых модифицированных производных, включая замещенные амиды, такие как алкиламиды и гидразиды. В некоторых случаях полипептид IL-10 является ретроинверсным аналогом.

[179] IL-10 может быть нативным белком, пептидным фрагментом IL-10 или модифицированным пептидом, обладающим по меньшей мере частью биологической активности нативного IL-10. IL-10 можно модифицировать для улучшения внутриклеточного поглощения. Одной из таких модификаций может быть присоединение домена белковой трансдукции. Домен белковой трансдукции может быть присоединен к C-концу IL-10. В качестве альтернативы, домен белковой трансдукции может быть присоединен к N-концу IL-10. Домен белковой трансдукции может быть присоединен к IL-10 посредством ковалентной связи. Домен белковой трансдукции может быть выбран из любой из последовательностей, перечисленных в **таблице 4h**.

Таблица 4h. Иллюстративные домены белковой трансдукции

SEQ ID NO	**Аминокислотная последовательность**
277	YGRKKRRQRRR;
278	RRQRRTSKLMKR
279	GWTLNSAGYLLGKINLKALAALAKKIL
280	KALAWEAKLAKALAKALAKHLAKALAKALKCEA
281	RQIKIWFQNRRMKWKK

SEQ ID NO	Аминокислотная последовательность
282	YGRKKRRQRRR
283	RKKRRQRRR
284	YGRKKRRQRRR
285	RKKRRQRR
286	YARAAARQARA
287	THRLPRRRRRR
288	GGRRARRRRRR

[180] BP композиций для человека или животного не ограничены нативными полноразмерными полипептидами, но также включают рекомбинантные версии, а также их биологически и/или фармакологически активные варианты или фрагменты. Например, квалифицированному специалисту будет понятно, что в BP могут быть сделаны различные аминокислотные замены для создания вариантов без отклонения от сути настоящего изобретения в отношении биологической активности или фармакологических свойств BP. Примеры консервативных замен аминокислот в полипептидных последовательностях показаны в **таблице 5**. Однако, в вариантах осуществления BPXTEN, в которых идентичность последовательности BP составляет менее 100% по сравнению с конкретной последовательностью, раскрытой в данном документе, настоящее изобретение предусматривает замену на любую из 19 других природных L-аминокислот для данного аминокислотного остатка данного BP, который может находиться в любом положении внутри последовательности BP, включая смежные аминокислотные остатки. Если какая-либо конкретная замена приводит к нежелательному изменению биологической активности, тогда можно использовать альтернативную аминокислоту, а конструкцию оценивать при помощи способов, описанных в данном документе, или с использованием любой из методик и рекомендаций для консервативных и неконсервативных мутаций, изложенных, например, в патенте США № 5364934, содержание которого включено посредством ссылки во всей своей полноте, или при помощи способов, общеизвестных для специалистов в данной области. Кроме того, варианты также могут включать, например, полипептиды, в которых один или более аминокислотных остатков добавлены или удалены на N- или C-конце полноразмерной нативной аминокислотной последовательности BP, которая сохраняет по меньшей мере часть биологической активности нативного пептида.

Таблица 5. Иллюстративные консервативные аминокислотные замены

Исходный остаток	Иллюстративные замены
Ala (A)	val; leu; ile
Arg (R)	lys; gin; asn
Asn (N)	gin; his; Iys; arg
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn
Glu (E)	asp
Gly (G)	pro
His (H)	asn: gin: Iys: arg
xIle (I)	leu; val; met; ala; phe: норлейцин

Leu (L)	норлейцин: ile: val; met; ala: phe
Lys (K)	arg: gin: asn
Met (M)	leu; phe; ile
Phe (F)	leu: val: ile; ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr(Y)	trp: phe: thr: ser
Val (V)	ile; leu; met; phe; ala; норлейцин

[181] В некоторых вариантах осуществления BP, включенный в полипептид BPXTEN, может иметь последовательность, которая проявляет по меньшей мере приблизительно 80% идентичность последовательности с последовательностью из **таблиц 4a-4h**, альтернативно, по меньшей мере приблизительно 81%, или приблизительно 82%, или приблизительно 83%, или приблизительно 84%, или приблизительно 85%, или приблизительно 86%, или приблизительно 87%, или приблизительно 88%, или приблизительно 89%, или приблизительно 90%, или приблизительно 91%, или приблизительно 92%, или приблизительно 93%, или приблизительно 94%, или приблизительно 95%, или приблизительно 96%, или приблизительно 97%, или приблизительно 98%, или приблизительно 99% или 100% идентичность последовательности с последовательностью из **таблиц 4a-4h**. В некоторых вариантах осуществления BP, включенный в BPXTEN, может представлять собой биспецифическую последовательность, содержащую первый связывающий домен и второй связывающий домен, где первый связывающий домен, обладающий аффинностью специфического связывания с опухолеспецифическим маркером или антигеном клетки-мишени, проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с парными последовательностями VL и VH антитела к CD3, выбранными из **таблицы 6f**; и где второй связывающий домен, обладающий аффинностью специфического связывания с эффекторной клеткой, проявляет по меньшей мере приблизительно 80% идентичность последовательности или, альтернативно, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичность последовательности с парными последовательностями VL и VH антитела к клетке-мишени, выбранными из **таблицы 6a**. Активность BP согласно вышеизложенным вариантам осуществления можно оценить с помощью анализов или измеренных или определенных параметров, которые описаны в данном документе, и те последовательности, которые сохраняют по меньшей мере приблизительно 40%, или приблизительно 50%, или приблизительно 55%, или приблизительно 60%, или приблизительно 70%, или приблизительно 80%, или приблизительно 90%, или приблизительно 95% или большую активность по сравнению с соответствующей нативной последовательностью BP, можно считать подходящими для включения в

BPXTEN для человека или животного. Обнаружено, что BP, сохраняющий подходящий уровень активности, может быть связан с одним или более полипептидами XTEN, описанными выше или где-либо еще в данном документе. В одном варианте осуществления BP, который, как обнаружено, сохраняет подходящий уровень активности, может быть связан с одним или более полипептидами XTEN, обладающими по меньшей мере приблизительно 80% идентичностью последовательности (например, по меньшей мере приблизительно 81%, по меньшей мере приблизительно 82%, по меньшей мере приблизительно 83%, по меньшей мере приблизительно 84%, по меньшей мере приблизительно 85%, по меньшей мере приблизительно 86%, по меньшей мере приблизительно 87%, по меньшей мере приблизительно 88%, по меньшей мере приблизительно 89%, по меньшей мере приблизительно 90%, по меньшей мере приблизительно 91%, по меньшей мере приблизительно 92%, по меньшей мере приблизительно 93%, по меньшей мере приблизительно 94%, по меньшей мере приблизительно 95%, по меньшей мере приблизительно 96%, по меньшей мере приблизительно 97%, по меньшей мере приблизительно 98%, по меньшей мере приблизительно 99% или 100% идентичностью последовательности) с последовательностью из **таблиц 3a-3b**, давая в результате химерный слитый белок.

Привлекающие Т-клетки активаторы

[182] Дополнительные формулы структурной конфигурации BPXTEN относятся к икстенилированным активируемым протеазой привлекающим Т-клетки активаторам («XPAT»), где BP представляет собой биспецифическое антитело (например, биспецифический привлекающий Т-клетки активатор). В некоторых вариантах осуществления композиция на основе XPAT содержит первую часть, содержащую первый связывающий домен и второй связывающий домен, вторую часть, содержащую сегмент для обеспечения высвобождения, и третью часть, содержащую объемообразующий фрагмент XTEN. В некоторых вариантах осуществления композиция на основе XPAT имеет конфигурацию **формулы Ia** (обозначена от N-конца к С-концу):

(первая часть)-(вторая часть)-(третья часть) (**Ia**),

где первая часть является биспецифической, содержащей два scFv, где первый связывающий домен обладает аффинностью специфического связывания с опухолеспецифическим маркером или антигеном клетки-мишени, и второй связывающий домен обладает аффинностью специфического связывания с эффекторной клеткой; вторая часть содержит сегмент для обеспечения высвобождения (RS), способный расщепляться протеазой млекопитающих (как более подробно описано в данном документе ниже, протеаза может быть опухоле- или антиген-специфичной, тем самым активируясь); и третья часть представляет собой объемообразующий фрагмент. В приведенном выше варианте осуществления связывающие домены первой части могут располагаться в порядке (VL-VH)1-(VL-VH)2, где «1» и «2» представляют собой первый и второй связывающие домены соответственно, или (VL-VH)1-(VH-VL)2, или (VH-VL)1-(VL-VH)2, или (VH-VL)1-(VH-VL)2, где спаренные связывающие домены связаны полипептидным

линкером (как более подробно описано в данном документе ниже). В одном варианте осуществления альтернативы для первой части VL и VH идентифицированы в **таблицах 6a-6f**; альтернативы для RS идентифицированы в последовательностях, представленных в **таблицах 8a-8b** (как более подробно описано в данном документе ниже); и альтернативы для объемообразующего фрагмента идентифицированы в данном документе как: XTEN; альбумин-связывающий домен; альбумин; IgG-связывающий домен; полипептиды, состоящие из пролина, серина и аланина; жирная кислота; Fc-домен; полиэтиленгликоль (PEG), PLGA и гидроксиэтилкрахмал. При необходимости, объемообразующий фрагмент представляет собой XTEN, обладающий по меньшей мере приблизительно 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью, идентифицированной с помощью последовательностей, представленных в **таблицах 3a-3b**. В приведенных выше вариантах осуществления композиция представляет собой рекомбинантный слитый белок. В другом варианте осуществления части связаны посредством химической конъюгации.

[183] В другом варианте осуществления композиция на основе XPAT имеет конфигурацию **формулы IIa** (обозначена от N-конца к C-концу):

(третья часть)-(вторая часть)-(первая часть) (**IIa**),

где первая часть является биспецифической, содержащей два scFv, где первый связывающий домен обладает аффинностью специфического связывания с опухолеспецифическим маркером или антигеном клетки-мишени, и второй связывающий домен обладает аффинностью специфического связывания с эффекторной клеткой; вторая часть содержит сегмент для обеспечения высвобождения (RS), способный расщепляться протеазой млекопитающих; и третья часть представляет собой объемообразующий фрагмент. В приведенном выше варианте осуществления связывающие домены первой части могут располагаться в порядке (VL-VH)1-(VL-VH)2, где «1» и «2» представляют собой первый и второй связывающие домены соответственно, или (VL-VH)1-(VH-VL)2, или (VH-VL)1-(VL-VH)2, или (VH-VL)1-(VH-VL)2, где спаренные связывающие домены связаны полипептидным линкером, как описано в данном документе ниже. В одном варианте осуществления альтернативы для первой части VL и VH идентифицированы в **таблицах 6a-6f**; альтернативы для RS идентифицированы в последовательностях, представленных в **таблицах 8a-8b**; и альтернативы для объемообразующего фрагмента идентифицированы в данном документе как: XTEN; альбумин-связывающий домен; альбумин; IgG-связывающий домен; полипептиды, состоящие из пролина, серина и аланина; жирная кислота и Fc-домен. При необходимости, объемообразующий фрагмент представляет собой XTEN, обладающий по меньшей мере приблизительно 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью, выбранной из группы последовательностей, представленных в **таблицах 3a-3b**. В приведенных выше вариантах осуществления композиция представляет собой рекомбинантный слитый белок. В другом варианте осуществления части связаны посредством химической конъюгации.

[184] В другом варианте осуществления композиция на основе XPAT имеет конфигурацию **формулы IIIa** (обозначена от N-конца к C-концу):

(пятая часть)-(четвертая часть)-(первая часть)-(вторая часть)-(третья часть) (**IIIa**),

где первая часть является биспецифической, содержащей два scFv, где первый связывающий домен обладает аффинностью специфического связывания с опухолеспецифическим маркером или антигеном клетки-мишени и второй связывающий домен обладает аффинностью специфического связывания с эффекторной клеткой; вторая часть содержит сегмент для обеспечения высвобождения (RS), способный расщепляться протеазой млекопитающих; третья часть представляет собой объемообразующий фрагмент; четвертая часть содержит сегмент для обеспечения высвобождения (RS), способный расщепляться протеазой млекопитающих, который может быть идентичным или отличаться от второй части; и пятая часть представляет собой объемообразующий фрагмент, который может быть идентичным или может отличаться от третьей части. В приведенном выше варианте осуществления связывающие домены первой части могут располагаться в порядке (VL-VH)1-(VL-VH)2, где «1» и «2» представляют собой первый и второй связывающие домены соответственно, или (VL-VH)1-(VH-VL)2, или (VH-VL)1-(VL-VH)2, или (VH-VL)1-(VH-VL)2, где спаренные связывающие домены связаны полипептидным линкером, как описано в данном документе ниже. В приведенных выше вариантах осуществления альтернативы для RS идентифицированы в последовательностях, представленных в **таблицах 8a-8b**. В приведенных выше вариантах осуществления альтернативы для объемообразующего фрагмента идентифицированы в данном документе как: XTEN; альбумин-связывающий домен; альбумин; IgG-связывающий домен; полипептиды, состоящие из пролина, серина и аланина; жирная кислота и Fc-домен. При необходимости, объемообразующий фрагмент представляет собой XTEN, обладающий по меньшей мере приблизительно 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью, выбранной из группы последовательностей, представленных в **таблицах 3a-3b**. В приведенных выше вариантах осуществления композиция представляет собой рекомбинантный слитый белок. В другом варианте осуществления части связаны посредством химической конъюгации.

[185] Композиции для человека или животного, исходя из их конструкции и конкретных компонентов, предпочтительно обеспечивают биспецифические терапевтические средства, которые обладают большей селективностью, большим периодом полужизни и приводят к меньшей токсичности и меньшему количеству побочных эффектов после их расщепления протеазами, обнаруженными в ассоциированных с мишенью тканях или в тканях, ставших нездоровыми в результате заболевания, где композиции для человека или животного имеют улучшенный терапевтический индекс по сравнению с композициями биспецифических антител, известными в уровне техники. Такие композиции пригодны при лечении определенных заболеваний, включая без ограничения рак, как указано в данном документе. Не

ограничиваясь какой-либо механистической теорией, специалисту в данной области будет понятно, что композиции по настоящему изобретению обеспечивают такое снижение неспецифических взаимодействий за счет комбинации механизмов, которые включают стерические затруднения из-за локализации связывающих доменов на объемных молекулах XTEN, при этом гибкие, неструктурированные характеристики полипептидов XTEN, будучи привязанными к композиции, способны колебаться и перемещаться вокруг связывающих доменов, обеспечивая блокировку между композицией и тканями или клетками, а также обеспечивая снижение способности интактной композиции проникать в клетку или ткань из-за большой молекулярной массы (обусловленной фактической молекулярной массой полипептида(-ов) XTEN и из-за большого гидродинамического радиуса неструктурированных полипептидов XTEN) по сравнению с размером отдельных связывающих доменов. Однако, композиции сконструированы таким образом, что в непосредственной близости от ткани или клетки-мишени, несущей или секретирующей протеазу, способную расщеплять RS, или при интернализации в клетку или ткань-мишень, когда связывающий домен связывает лиганд, биспецифические связывающие домены высвобождаются из основной массы XTEN под действием протеазы (протеаз), устраняя барьер стерических препятствий, и более свободно проявляют свой фармакологический эффект. Композиции для человека или животного находят применение при лечении множества состояний, когда необходима селективная доставка терапевтической композиции биспецифического антитела в клетку, ткань или орган. В одном варианте осуществления ткань-мишень представляет собой рак, который может представлять собой лейкоз, лимфому или опухоль органа или системы.

Связывающие домены

[186] Настоящее изобретение предусматривает применение одноцепочечных связывающих доменов, таких как, без ограничения, Fv, Fab, Fab', Fab'-SH, $F(ab')_2$, линейные антитела, однодоменное антитело, однодоменное верблюжье антитело, одноцепочечные молекулы антител (scFv) и диатела, способные связывать лиганды или рецепторы, ассоциированные с эффекторными клетками и антигенами пораженных тканей или клеток, которые представляют собой формы рака, опухоли или другие злокачественные ткани. В некоторых вариантах осуществления биспецифическое антитело содержит первый связывающий домен со специфичностью связывания с маркером клетки-мишени и второй связывающий домен со специфичностью связывания с антигеном эффекторной клетки. В некоторых вариантах осуществления первый и второй связывающие домены могут быть каркасами, не являющимися антителами, такими как антикалины, аднектины, финомеры, аффилины, аффитела, центирины, дарпины. В других вариантах осуществления связывающий домен для опухолевой клетки-мишени представляет собой вариабельный домен Т-клеточного рецептора, который был сконструирован для связывания МНС, нагруженного пептидным фрагментом белка, который избыточно экспрессируется опухолевыми клетками. В некоторых вариантах осуществления композиции на основе XPAT разработаны с учетом локализации протеазы

ткани-мишени, а также присутствия той же протеазы в здоровых тканях, не предназначенных для воздействия, а также присутствия лиганда-мишени в здоровой ткани, но большего присутствия лиганда в нездоровой ткани-мишени, чтобы обеспечить широкое терапевтическое окно. «Терапевтическое окно» относится к наибольшей разнице между минимальной эффективной дозой и максимальной переносимой дозой для данной терапевтической композиции. Чтобы помочь в достижении широкого терапевтического окна, связывающие домены первой части композиций экранируют за счет близости объемообразующего фрагмента (например, полипептида XTEN), где аффинность связывания интактной композиции для одного или обоих из лигандов уменьшается по сравнению с композицией, которая была расщеплена протеазой млекопитающих, тем самым освобождая первую часть от экранирующих эффектов объемообразующего фрагмента.

[187] Что касается одноцепочечных связывающих доменов, которые хорошо известны в уровне техники, Fv представляет собой минимальный фрагмент антитела, который содержит полный сайт распознавания и связывания антигена, состоящий из димера одной тяжелой (VH) и одной легкой цепи вариабельного домена (VL) в нековалентной ассоциации. Внутри каждой цепи VH и VL находятся три определяющие комплементарность области (CDR), которые взаимодействуют, чтобы сформировать сайт связывания антигена на поверхности димера VH-VL; шесть CDR связывающего домена придают специфичность связывания антигена антителу или одноцепочечному связывающему домену. В некоторых случаях создают scFv, каждый из которых имеет 3, 4 или 5 CHR в каждом связывающем домене. Каркасные последовательности, фланкирующие CDR, имеют третичную структуру, которая по сути консервативна в нативных иммуноглобулинах у разных видов, а каркасные остатки (FR) служат для удержания CDR в их соответствующей ориентации. Константные домены не требуются для функции связывания, но могут способствовать стабилизации взаимодействия VH-VL. В некоторых вариантах осуществления домен сайта связывания полипептида может представлять собой пару доменов VH-VL, VH-VH или VL-VL одного и того же или разных иммуноглобулинов, однако обычно предпочтительно создавать одноцепочечные связывающие домены с использованием соответствующих цепей VH и VL из исходного антитела. Порядок доменов VH и VL в полипептидной цепи не является ограничивающим для настоящего изобретения; указанный порядок доменов можно изменить, как правило, без какой-либо потери функции, но следует понимать, что домены VH и VL расположены таким образом, что сайт связывания антигена может правильно сворачиваться. Таким образом, одноцепочечные связывающие домены вариантов осуществления биспецифических scFv композиций для человека или животного могут располагаться в порядке $(VL\text{-}VH)^1\text{-}(VL\text{-}VH)^2$, где «1» и «2» представляют собой первый и второй связывающие домены соответственно, или $(VL\text{-}VH)^1\text{-}(VH\text{-}VL)^2$, или $(VH\text{-}VL)^1\text{-}(VL\text{-}VH)^2$, или $(VH\text{-}VL)^1\text{-}(VH\text{-}VL)^2$, где спаренные связывающие домены связаны полипептидным линкером, который описан в данном документе ниже.

[188] Таким образом, расположение связывающих доменов в иллюстративном биспецифическом одноцепочечном антителе, раскрытом в данном документе, может быть таким, при котором первый связывающий домен расположен с С-конца по отношению ко второму связывающему домену. Расположение V-цепей может быть следующим: VH(поверхностный антиген клетки-мишени)-VL(поверхностный антиген клетки-мишени)-VL(антиген эффекторной клетки)-VH(антиген эффекторной клетки), VH(поверхностный антиген клетки-мишени)-VL(поверхностный антиген клетки-мишени)-VH(антиген эффекторной клетки)-VL(антиген эффекторной клетки), VL(поверхностный антиген клетки-мишени)-VH(поверхностный антиген клетки-мишени)-VL(антиген эффекторной клетки)-VH(антиген эффекторной клетки) или VL(поверхностный антиген клетки-мишени)-VH(поверхностный антиген клетки-мишени)-VH(антиген эффекторной клетки)-VL(антиген эффекторной клетки). Для расположения, при котором второй связывающий домен расположен с N-конца по отношению к первому связывающему домену, возможен следующий порядок: VH (антиген эффекторной клетки)-VL(антиген эффекторной клетки)-VL(поверхностный антиген клетки-мишени)-VH(поверхностный антиген клетки-мишени), VH(антиген эффекторной клетки)-VL(антиген эффекторной клетки)-VH(поверхностный антиген клетки-мишени)-VL(поверхностный антиген клетки-мишени), VL(антиген эффекторной клетки)-VH(антиген эффекторной клетки)-VL(поверхностный антиген клетки-мишени)-VH(поверхностный антиген клетки-мишени) или VL(антиген эффекторной клетки)-VH(антиген эффекторной клетки)-VH(поверхностный антиген клетки-мишени)-VL(поверхностный антиген клетки-мишени). Используемые в данном документе термины «с N-конца по отношению к» или «с С-конца по отношению к» и их грамматические варианты обозначают относительное расположение в первичной аминокислотной последовательности, а не размещение на абсолютном N- или С-конце биспецифического одноцепочечного антитела. Следовательно, в качестве неограничивающего примера, первый связывающий домен, который «расположен с С-конца по отношению к второму связывающему домену», означает, что первое связывание расположено с карбоксильной стороны второго связывающего домена внутри биспецифического одноцепочечного антитела, и не исключает возможности того, что дополнительная последовательность, например His-метка, или другое соединение, такое как радиоизотоп, расположен на С-конце биспецифического одноцепочечного антитела.

[189] В одном варианте осуществления композиции на основе сборки химерных полипептидов содержат первую часть, содержащую первый связывающий домен и второй связывающий домен, где каждый из указанных связывающих доменов представляет собой scFv, и где каждый scFv содержит один VL и один VH. В другом варианте осуществления композиции на основе сборки химерных полипептидов содержат первую часть, содержащую первый связывающий домен и второй связывающий домен, где указанные связывающие домены имеют конфигурацию диатела и где каждый домен содержит один домен VL и один VH. В приведенных выше вариантах осуществления первый домен

обладает специфичностью связывания с опухолеспецифическим маркером или антигеном клетки-мишени, а второй связывающий домен обладает специфичностью связывания с антигеном эффекторной клетки. В одном варианте осуществления вышеизложенного антиген эффекторной клетки экспрессируется на эффекторной клетке или внутри нее. В одном варианте осуществления антиген эффекторной клетки экспрессируется на Т-клетке, такой как CD4+, CD8+ или естественная клетка-киллер (NK). В другом варианте осуществления антиген эффекторной клетки экспрессируется на В-клетке, тучной клетке, дендритной клетке или миелоидной клетке. В одном варианте осуществления антиген эффекторной клетки представляет собой CD3, антиген кластера дифференцировки 3 цитотоксической Т-клетки. В некоторых вариантах осуществления вышеизложенного первый связывающий домен проявляет специфичность связывания с опухолеспецифическим маркером, ассоциированным с опухолевой клеткой. В одном варианте осуществления связывающий домен обладает аффинностью связывания с опухолеспецифическим маркером, где опухолевая клетка может включать, без ограничения, клетки опухоли из стромальных клеток, опухоли из фибробластов, опухоли из миофибробластов, опухоли из глиальных клеток, опухоли из эпителиальных клеток, опухоли из жировых клеток, опухоли из иммунных клеток, опухоли из сосудистых клеток и опухоли из гладкомышечных клеток. В одном варианте осуществления опухолеспецифический маркер или антиген клетки-мишени может представлять собой альфа-4-интегрин, Ang2, B7-H3, B7-H6, CEACAM5, cMET, CTLA4, FOLR1, EpCAM, CCR5, CD19, HER2, HER2 neu, HER3, HER4, HER1 (EGFR), PD-L1, PSMA, CEA, TROP-2, MUC1 (муцин), MUC-2, MUC3, MUC4, MUC5AC, MUC5B, MUC7, MUC16 βhCG, Льюис-Y, CD20, CD33, CD38, CD30, CD56 (NCAM), CD133, ганглиозид GD3; 9-O-ацетил-GD3, GM2, Globo H, фукозил GM1, GD2, карбоангидразу IX, CD44v6, нектин-4, «Sonic Hedgehog» (Shh), Wue-1, антиген 1 плазматической клетки, хондроитинсульфат протеогликан меланомы (MCSP), CCR8, 6-трансмембранный эпителиальтный антиген предстательной железы (STEAP), мезотелин, антиген A33, антиген стволовых клеток предстательной железы (PSCA), Ly-6, десмоглеин 4, фетальный ацетилхолиновый рецептор (fnAChR), CD25, раковый антиген 19-9 (CA19-9), раковый антиген 125 (CA-125), рецептор мюллерова ингибирующего фактора типа II (MISIIR), сиалированный антиген Tn (s TN), антиген активации фибробластов (FAP), эндосиалин (CD248), вариант III рецептора эпидермального фактора роста (EGFRvIII), опухолеассоциированный антиген L6 (TAL6), SAS, CD63, TAG72, антиген Томсена-Фриденрайха (TF-антиген), рецептор инсулиноподобного фактора роста I (IGF-IR), антиген Cora, CD7, CD22, CD70, CD79a, CD79b, G250, MT-MMP, антиген F19, CA19-9, CA-125, альфа-фетопротеин (AFP), VEGFR1, VEGFR2, DLK1, SP17, ROR1 и EphA2. В одном варианте осуществления первым связывающим доменом, который проявляет аффинность связывания с CD70, является его природный лиганд, CD27, а не фрагмент антитела. В другом варианте осуществления первым связывающим доменом, который проявляет аффинность связывания с B7-H6, является его природный лиганд Nkp30, а не фрагмент антитела.

[190] Варианты осуществления scFv композиций на основе XPAT по настоящему изобретению содержат первый связывающий домен и второй связывающий домен, где домены VL и VH получены из моноклональных антител со специфичностью связывания с опухолеспецифическим маркером или антигеном клетки-мишени и антигенами эффекторной клетки соответственно. В других случаях каждый из первого и второго связывающих доменов содержит шесть CDR, полученных из моноклональных антител со специфичностью связывания с маркером клетки-мишени, таким как опухолеспецифический маркер, и антигенами эффекторной клетки соответственно. В других вариантах осуществления первый и второй связывающие домены первой части композиций для человека или животного могут иметь 3, 4 или 5 CHR в каждом связывающем домене. В других вариантах осуществления варианты осуществления настоящего изобретения предусматривают первый связывающий домен и второй связывающий домен, где каждый содержит область CDR-H1, область CDR-H2, область CDR-H3, область CDR-L1, область CDR-L2 и область CDR-H3, где каждая из указанных областей получена из моноклонального антитела, способного связывать опухолеспецифический маркер или антиген клетки-мишени и антигены эффекторной клетки, соответственно. В одном варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, в которой второй связывающий домен содержит области VH и VL, полученные из моноклонального антитела, способного связывать CD3 человека. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен scFv содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к CD3, представленного в **таблице 6а**. В другом аспекте варианты осуществления второго домена по настоящему изобретению содержат область CDR-H1, область CDR-H2, область CDR-H3, область CDR-L1, область CDR-L2 и область CDR-H3, где каждая из указанных областей получена из моноклонального антитела, представленного в **таблице 6а**. В приведенных выше вариантах осуществления домены VH и/или VL могут быть сконфигурированы в виде scFv, диатела, однодоменного антитела или однодоменного верблюжьего антитела.

[191] В других вариантах осуществления вторые домены композиций для человека или животного получены из антитела к CD3, представленного в **таблице 6а**. В одном варианте осуществления вышеизложенного второй домен композиции для человека или животного содержит парные последовательности областей VL и VH антитела к CD3, представленного в **таблице 6а**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%,

или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к CD3 huUCHT1 из **таблицы 6a**. В приведенных выше вариантах осуществления домены VH и/или VL могут быть сконфигурированы в виде scFv, части диатела, однодоменного антитела или однодоменного верблюжьего антитела.

[192] В других вариантах осуществления scFv первого домена композиции получают из антитела к опухолевым клеткам, представленного в **таблице 6f**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где первый связывающий домен содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к опухолевым клеткам, представленного в **таблице 6f**. В одном варианте осуществления вышеизложенного первый домен указанных композиций содержит парные последовательности областей VL и VH раскрытого в данном документе антитела к опухолевым клеткам. В приведенных выше вариантах осуществления домены VH и/или VL могут быть сконфигурированы в виде scFv, части диатела, однодоменного антитела или однодоменного верблюжьего антитела.

[193] В другом варианте осуществления композиции на основе сборки химерных полипептидов содержат первую часть, содержащую первый связывающий домен и второй связывающий домен, где указанные связывающие домены имеют конфигурацию диатела и каждый из указанных связывающих доменов содержит один домен VL и один домен VH. В одном варианте осуществления варианты осуществления диатела по настоящему изобретению предусматривают первый связывающий домен и второй связывающий домен, где домены VL и VH получены из моноклональных антител со специфичностью связывания с опухолеспецифическим маркером или антигеном клетки-мишени и антигеном эффекторной клетки, соответственно. В другом варианте осуществления варианты осуществления диатела по настоящему изобретению предусматривают первый связывающий домен и второй связывающий домен, где каждый содержит область CDR-H1, область CDR-H2, область CDR-H3, область CDR-L1, область CDR-L2 и область CDR-H3, где каждая из указанных областей получена из моноклонального антитела, способного связывать опухолеспецифический маркер или антиген клетки-мишени и антиген эффекторной клетки, соответственно. Предусматривается, что варианты осуществления диатела по настоящему изобретению включают первый связывающий домен и второй связывающий домен, где домены VL и VH получены из моноклональных антител со специфичностью связывания с опухолеспецифическим маркером или антигеном клетки-мишени и антигеном эффекторной клетки, соответственно. В другом аспекте варианты осуществления диатела по настоящему изобретению предусматривают первый связывающий домен и второй связывающий домен, где каждый содержит область CDR-H1, область CDR-H2, область CDR-H3, область CDR-L1, область CDR-L2 и область CDR-

H3, где каждая из указанных областей получена из моноклонального антитела, способного связывать опухолеспецифический маркер или антиген клетки-мишени и антиген эффекторной клетки, соответственно. В одном варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен диатела содержит парные области VH и VL, полученные из моноклонального антитела, способного связывать CD3 человека. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен диатела содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к CD3, представленного в **таблице 6a**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен диатела содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной последовательностям VL и VH антитела huUCHT1, представленного в **таблице 6a**. В других вариантах осуществления второй домен диатела композиции получен из антитела к CD3, описанного в данном документе. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где первый связывающий домен диатела содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной последовательностям VL и VH антитела к опухолевым клеткам, представленного в **таблице 6f**. В других вариантах осуществления первый домен диатела композиции получен из антитела к опухолевым клеткам, описанного в данном документе.

[194] Терапевтические моноклональные антитела, из которых могут быть получены VL, и VH, и домены CDR для композиций для человека или животного, известны в уровне техники. Последовательности вышеуказанных антител можно получить из общедоступных баз данных, патентов или литературных ссылок. Кроме того, неограничивающие примеры моноклональных антител и последовательностей VH и VL из антител к CD3 приведены в **таблице 6a** и неограничивающие примеры моноклональных антител и последовательностей VH и VL к маркерам рака, опухоли или клеток-мишеней приведены в **таблице 6f**.

Связывающие домены к CD3

[195] В некоторых вариантах осуществления настоящее изобретение предусматривает композиции на основе сборки химерных полипептидов, содержащие связывающий домен первой части с аффинностью связывания с Т-клетками. В одном варианте осуществления связывающий домен второй части содержит VL и VH,

полученные из моноклонального антитела к антигену CD3. В другом варианте осуществления связывающий домен содержит VL и VH, полученные из моноклонального антитела к CD3-эпсилон и CD3-дельта. Моноклональные антитела к CD3 neu известны в уровне техники. Иллюстративные неограничивающие примеры последовательностей VL и VH моноклональных антител к CD3 представлены в **таблице 6а**. В одном варианте осуществления настоящее изобретение предусматривает сборку химерного полипептида, предусматривающую связывающий домен с аффинностью связывания с CD3, содержащий последовательности VL и VH антитела к CD3, представленные в **таблице 6а**. В другом варианте осуществления настоящее изобретение предусматривает сборку химерного полипептида, предусматривающую связывающий домен первой части с аффинностью связывания с CD3-эпсилон, содержащий последовательности VL и VH антитела к CD3-эпсилон, представленные в **таблице 6а**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где второй связывающий домен scFv первой части содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к CD3 huUCHT1 из **таблицы 6а**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, содержащую связывающий домен с аффинностью связывания с CD3, включающий область CDR-L1, область CDR-L2, область CDR-L3, область CDR-H1, область CDR-H2 и область CDR-H3, каждая из которых получена из соответствующих последовательностей VL и VH антитела к CD3, представленных в **таблице 6а**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, содержащую связывающий домен с аффинностью связывания с CD3, включающий область CDR-L1, область CDR-L2, область CDR-L3, область CDR-H1, область CDR-H2 и область CDR-H3, где последовательности CDR представляют собой RASQDIRNYLN (SEQ ID NO: 8034), YTSRLES (SEQ ID NO: 8035),QQGNTLPWT (SEQ ID NO: 8036), GYSFTGYTMN (SEQ ID NO: 8037), LINPYKGVST (SEQ ID NO: 8038) и SGYYGDSDWYFDV (SEQ ID NO: 8039)

[196] Комплекс CD3 представляет собой группу молекул клеточной поверхности, которая связывается с Т-клеточным антигенным рецептором (TCR) и участвует в экспрессии TCR на клеточной поверхности и в каскаде сигнальной трансдукции, который возникает, когда комплекс пептид:лиганд MHC связывается с TCR. Как правило, когда антиген связывается с Т-клеточным рецептором, CD3 посылает сигналы через клеточную мембрану в цитоплазму внутри Т-клетки. Это вызывает активацию Т-клеток, которые быстро делятся с образованием новых Т-клеток, сенсибилизированных для атаки на конкретный антиген, который воздействовал на TCR. Комплекс CD3 состоит из молекулы CD3-эпсилон вместе с четырьмя другими связанными с мембраной полипептидами (CD3-гамма, -дельта, -зета и -бета). У человека CD3-эпсилон кодируется геном CD3E на хромосоме 11. Внутриклеточные домены каждой из цепей CD3 содержат

иммунорецепторные тирозиновые активирующие мотивы (ITAM), которые служат точкой зарождения для внутриклеточного механизма сигнальной трансдукции при взаимодействии с Т-клеточным рецептором.

[197] Ряд терапевтических стратегий модулирует Т-клеточный иммунитет путем целенаправленного воздействия на передачу сигналов TCR, в частности, моноклональные антитела (mAb) к человеческому CD3, которые широко применяются клинически в иммуносупрессивных схемах. CD3-специфическое мышиное mAb OKT3 было первым mAb, лицензированным для применения у людей (Sgro, C. Side-effects of a monoclonal antibody, muromonab CD3/orthoclone OKT3: bibliographic review. Toxicology 105:23-29, 1995), и широко применяется клинически в качестве иммуносупрессивного средства при трансплантации (Chatenoud, Clin. Transplant 7:422-430, (1993); Chatenoud, Nat. Rev. Immunol. 3:123-132 (2003); Kumar, Transplant. Proc. 30:1351-1352 (1998)), диабете 1 типа и псориазе. Важно отметить, что mAb к CD3 могут индуцировать частичную передачу сигналов Т-клетками и клональную анергию (Smith, JA, Nonmitogenic Anti-CD3 Monoclonal Antibodies Deliver a Partial T Cell Receptor Signal and Induce Clonal Anergy J. Exp. Med. 185:1413-1422 (1997)). OKT3 было описано в литературе как митоген Т-клеток, а также как мощный убийца Т-клеток (Wong, JT. The mechanism of anti-CD3 monoclonal antibodies. Mediation of cytolysis by inter-T cell bridging. Transplantation 50:683-689 (1990)). В частности, исследования Wong продемонстрировали, что при связывании с помощью CD3 Т-клеток и клеток-мишеней можно добиться уничтожения мишени, и что ни FcR-опосредованная ADCC, ни фиксация комплемента не требуются бивалентным mAb к CD3 для лизиса клеток-мишеней.

[198] OKT3 проявляет как митогенную активность, так и активность по уничтожению Т-клеток в зависимости от времени; после ранней активации Т-клеток, приводящей к высвобождению цитокинов, позже при дальнейшем введении OKT3 блокирует все известные функции Т-клеток. Именно из-за этого более позднего блокирования функции Т-клеток OKT3 нашло такое широкое применение в качестве иммуносупрессанта в схемах лечения для уменьшения или даже устранения отторжения ткани аллотрансплантата. Другие антитела, специфичные к молекуле CD3, раскрыты в Tunnacliffe, Int. Immunol. 1 (1989), 546-50, в WO2005/118635 и WO2007/033230 описаны моноклональные антитела к CD3-эпсилон человека, в патенте США № 5821337 описаны последовательности VL и VH мышиного моноклонального антитела к CD3 UCHT1 (muxCD3, Shalaby et al., J. Exp. Med. 175, 217-225 (1992)) и гуманизированный вариант этого антитела (hu UCHT1), а в заявке на патент США № 20120034228 раскрыты связывающие домены, способные связываться с эпитопом цепи CD3-эпсилон человека и примата, не являющегося шимпанзе.

Таблица 6а. Моноклональные антитела и последовательности к CD3

Название клона	**Название антитела**	**Мишень**	**Последовательность VH**	**SEQ ID NO: (полный**	**SEQ ID NO: (CDR)**	**Последовательность VL**	**SEQ ID NO: (полный**	**SEQ ID NO: (CDR)**

				длины)			**длины)**	
huOKT3		CD3	QVQLVQSGGGVVQPGRSLRLSCKAS**GYTFTRYTMH**WVRQAPGKGLEWIG**YINPSRGYTNYNQKVKD**RFTISRDNSKNTAFLQMDSLRPEDTGVYFCAR**YYDDHYCLDY**WGQGTPVTVSS	8040	8900, 8901, 8902	DIQMTQSPSSLSASVGDRVTITC**SASSSVSYMN**WYQQTPGKAPKRWIY**DTSKLAS**GVPSRFSGSGSGTDYTFTISSLQPEDIATYYC**QQWSSNPFT**FGQGTKLQITR	8049	8908, 8909, 8910
huUCHT1		CD3	EVQLVESGGGLVQPGGSLRLSCAAS**GYSFTGYTMN**WVRQAPGKGLEWVA**LINPYKGVST**YNQKFKDRFTISVDKSKNTAYLQMNSLRAEDTAVYYCAR**SGYYGDSDWYFDV**WGQGTLVTVSS	8041	8037, 8038, 8039	DIQMTQSPSSLSASVGDRVTITC**RASQDIRNYLN**WYQQKPGKAPKLLIY**YTSRLES**GVPSRFSGSGSGTDYTLTISSLQPEDFATYYC**QQGNTLPWT**FGQGTKVEIK	8050	8034, 8035, 8036
hu12F6		CD3	QVQLVQSGGGVVQPGRSLRLSCKAS**GYTFTSYTMH**WVRQAPGKGLEWIG**YINPSSGYTKYNQKFKD**RFTISADKSKSTAFLQMDSLRPEDTGVYFCAR**WQDYDVYFDY**WGQGTPVTVSS	8042	8903, 8904, 8905	DIQMTQSPSSLSASVGDRVTMTC**RASSSVSYMH**WYQQTPGKAPKPWIY**ATSNLAS**GVPSRFSGSGSGTDYTLTISSLQPEDIATYYC**QQWSSNPPT**FGQGTKLQITR	8051	8911, 8912, 8913
mOKT3		CD3	QVQLQQSGAELARPGASVKMSCKA	8043	8900, 8906, 8902	QIVLTQSPAIMSASPGEKVTMT	8052	8908, 8909, 8910

			S**GYTFTRY TMH**WVKQ RPGQGLEW IG**YINPSRG YTNYNQK FKD**KATLT TDKSSSTA YMQLSSLT SEDSAVYY CAR**YYDD HYCLDY**W GQGTTLTV SS			C**SASSSVS YMN**WYQ QKSGTSP KRWIY**DT SKLAS**GV PAHFRGS GSGTSYS LTISGME AEDAATY YC**QQWS SNPFT**FGS GTKLEINR		
MT10 3	blinat umo mab	CD3	DIKLQQSG AELARPGA SVKMSCKT S**GYTFTRY TMH**WVKQ RPGQGLEW IG**YINPSRG YTNYNQK FKD**KATLT TDKSSSTA YMQLSSLT SEDSAVYY CAR**YYDD HYCLDY**W GQGTTLTV SS	8044	8900, 8906, 8902	DIQLTQSP AIMSASP GEKVTMT C**RASSSV SYMN**WY QQKSGTS PKRWIY**D TSKVAS**G VPYRFSG SGSGTSYS LTISSMEA EDAATYY C**QQWSS NPLT**FGA GTKLELK	8053	8914, 8915, 8916
MT11 0	solito mab	CD3	DVQLVQSG AEVKKPGA SVKVSCKA S**GYTFTRY TMH**WVRQ APGQGLE WIG**YINPS RGYTNYA DSVKG**RFT ITTDKSTST AYMELSSL RSEDTATY YCAR**YYD DHYCLDY** WGQGTTV TVSS	8045	8900, 8907, 8902	DIVLTQSP ATLSLSPG ERATLSC **RASQSVS YMN**WYQ QKPGKAP KRWIY**DT SKVAS**GV PARFSGS GSGTDYS LTINSLEA EDAATYY C**QQWSS NPLT**FGG GTKVEIK	8054	8917, 8915, 8916
CD3.7		CD3	EVQLVESG GGLVQPGG SLKLSCAA SGFTFNKY AMNWVRQ APGKGLE WVARIRSK	8046		QTVVTQE PSLTVSPG GTVTLTC GSSTGAV TSGYYPN WVQQKP GQAPRGL	8055	

			YNNYATY YADSVKDR FTISRDDSK NTAYLQM NNLKTEDT AVYYCVR HGNFGNSY ISYWAYW GQGTLVTV SS			IGGTKFLA PGTPARFS GSLLGGK AALTLSG VQPEDEA EYYCALW YSNRWVF GGGTKLT VL		
CD3.8		CD3	EVQLVESG GGLVQPGG SLRLSCAA SGFTFNTY AMNWVRQ APGKGLE WVGRIRSK YNNYATY YADSVKGR FTISRDDSK NTLYLQM NSLRAEDT AVYYCVR HGNFGNSY VSWFAYW GQGTLVTV SS	8047		QAVVTQE PSLTVSPG GTVTLTC GSSTGAV TTSNYAN WVQQKP GQAPRGL IGGTNKR APGVPAR FSGSLLG GKAALTL SGAQPED EAEYYCA LWYSNL WVFGGGT KLTVL	8056	
CD3.9		CD3	EVQLLESG GGLVQPGG SLKLSCAA SGFTFNTY AMNWVRQ APGKGLE WVARIRSK YNNYATY YADSVKDR FTISRDDSK NTAYLQM NNLKTEDT AVYYCVR HGNFGNSY VSWFAYW GQGTLVTV SS	773		ELVVTQE PSLTVSPG GTVTLTC RSSTGAV TTSNYAN WVQQKP GQAPRGL IGGTNKR APGTPAR FSGSLLG GKAALTL SGVQPED EAEYYCA LWYSNL WVFGGGT KLTVL	772	
CD3.10		CD3	EVKLLESG GGLVQPKG SLKLSCAA SGFTFNTY AMNWVRQ APGKGLE WVARIRSK YNNYATY	8048		QAVVTQE SALTTSPG ETVTLTC RSSTGAV TTSNYAN WVQEKPD HLFTGLIG GTNKRAP	8057	

			YADSVKDR FTISRDDSQ SILYLQMN NLKTEDTA MYYCVRH GNFGNSYV SWFAYWG QGTLVTVS S			GVPARFS GSLIGDK AALTITG AQTEDEA IYFCALW YSNLWVF GGGTKLT VL		

* подчеркнутые последовательности, если они есть, представляют собой CDR внутри VL и VH

АНТИГЕНСВЯЗЫВАЮЩИЕ ФРАГМЕНТЫ ДЛЯ CD3 НА КЛЕТКАХ

[199] В другом аспекте настоящее изобретение относится к антигенсвязывающим фрагментам (AF2), характеризующимся аффинностью специфичного связывания с антигеном эффекторных клеток, которые могут быть включены в состав любого из описанных в данном документе вариантов осуществления композиции для человека или животного. В некоторых случаях антиген эффекторных клеток экспрессируется на поверхности эффекторной клетки, выбранной из плазматической клетки, Т-клетки, В-клетки, цитокин-индуцированной клетки-киллера (CIK-клетки), тучной клетки, дендритной клетки, регуляторной Т-клетки (RegT-клетки), хелперной Т-клетки, миелоидной клетки и NK-клетки.

[200] Различные AF2, которые связывают антигены эффекторных клеток, особенно применимы для спаривания с антигенсвязывающим фрагментом с аффинностью связывания с антигенами EGFR, ассоциированными с пораженной клеткой или тканью, в форматах композиции для осуществления уничтожения клеток, направленного на пораженную клетку или ткань. Специфичность связывания может быть определена с помощью областей, определяющих комплементарность, или CDR, таких как CDR легкой цепи или CDR тяжелой цепи. Во многих случаях специфичность связывания определяется с помощью CDR легкой цепи и CDR тяжелой цепи. Заданная комбинация CDR тяжелой цепи и CDR легкой цепи обеспечивает получение заданного связывающего кармана, который придает более высокую аффинность и/или специфичность в отношении антигена эффекторных клеток по сравнению с другими эталонными антигенами. Получаемые биспецифические композиции, имеющие первый антигенсвязывающий фрагмент (AF1) для EGFR, связанный с помощью короткого гибкого пептидного линкера со вторым антигенсвязывающим фрагментом (AF2) со специфичностью связывания с антигеном эффекторных клеток, являются биспецифическими, при этом каждый антигенсвязывающий фрагмент характеризуется аффинностью специфичного связывания с его соответствующими лигандами. Специалисту в данной области будет понятно, что в таких композициях AF1, направленный против EGFR в пораженной ткани, используется в комбинации с AF2, направленным против маркера эффекторных клеток, чтобы привести эффекторную клетку в непосредственную близость к клетке пораженной ткани для осуществления цитолиза клетки пораженной ткани. Кроме того, AF1 и AF2 включают в

состав специально разработанных полипептидов, содержащих расщепляемые сегменты для обеспечения высвобождения и XTEN, чтобы придать характеристики пролекарства композициям, которые активируются при высвобождении слитых AF1 и AF2 после расщепления сегментов для обеспечения высвобождения, когда они находятся вблизи пораженной ткани, имеющей протеазы, способные расщеплять сегменты для обеспечения высвобождения в одном или более местоположениях в последовательности сегмента для обеспечения высвобождения.

[201] В одном варианте осуществления AF2 в композициях для человека или животного характеризуется аффинностью связывания с антигеном эффекторных клеток, экспрессируемым на поверхности Т-клетки. В другом варианте осуществления AF2 в композициях для человека или животного характеризуется аффинностью связывания с CD3. В другом варианте осуществления AF2 в композициях для человека или животного характеризуется аффинностью связывания с представителем комплекса CD3, который включает в себя в отдельной форме или в форме независимой комбинации все известные субъединицы CD3 комплекса CD3; например, CD3-эпсилон, CD3-дельта, CD3-гамма, CD3-дзета, CD3-альфа и CD3-бета. В другом варианте осуществления AF2 характеризуется аффинностью связывания с CD3-эпсилон, CD3-дельта, CD3-гамма, CD3-дзета, CD3-альфа или CD3-бета.

[202] Антигенсвязывающие фрагменты, рассматриваемые в настоящем изобретении, по происхождению могут быть получены из встречающегося в природе антитела или его фрагмента, не встречающегося в природе антитела или его фрагмента, гуманизированного антитела или его фрагмента, синтетического антитела или его фрагмента, гибридного антитела или его фрагмента или сконструированного антитела или его фрагмента. Способы получения антитела к заданному маркеру-мишени хорошо известны из уровня техники. Например, моноклональные антитела могут быть получены с использованием гибридомного способа, впервые описанного Kohler et al., Nature, 256:495 (1975), или могут быть получены с помощью способов рекомбинантных ДНК (патент США № 4816567). Структура антител и их фрагментов, вариабельные области тяжелых и легких цепей антитела (VH и VL), одноцепочечные вариабельные области (scFv), области, определяющие комплементарность (CDR), и доменные антитела (dAb) хорошо изучены. Способы получения полипептида, имеющего необходимый антигенсвязывающий фрагмент с аффинностью связывания с заданным антигеном, известны из уровня техники.

[203] Специалисту в данной области будет понятно, что использование термина «антигенсвязывающие фрагменты» для раскрытых в данном документе вариантов осуществления композиции предполагает включение в него частей или фрагментов антител, которые сохраняют способность связывать антигены, являющиеся лигандами соответствующего интактного антитела. В таких вариантах осуществления антигенсвязывающий фрагмент может представлять собой без ограничения CDR и промежуточные каркасные области, вариабельные или гипервариабельные области легкой и/или тяжелой цепей антитела (VL, VH), вариабельные фрагменты (Fv), Fab'-фрагменты,

F(ab')2-фрагменты, Fab-фрагменты, одноцепочечные антитела (scAb), антитела VHH верблюдовых, одноцепочечный вариабельный фрагмент (scFv), линейные антитела, однодоменное антитело, области, определяющие комплементарность (CDR), доменные антитела (dAb), однодоменные иммуноглобулины из тяжелых цепей типа BHH или BNAR, однодоменные иммуноглобулины из легких цепей или другие полипептиды, известные из уровня техники, содержащие фрагмент антитела, способный связывать антиген. Антигенсвязывающие фрагменты, имеющие CDR-H и CDR-L, могут быть представлены в конфигурации в ориентации (CDR-H)-(CDR-L) или (CDR-H)-(CDR-L) от N-конца к C-концу. VL и VH двух антигенсвязывающих фрагментов также могут быть представлены в конфигурации одноцепочечного диатела; т. е. VL и VH из AF1 и AF2 представлены в конфигурации с линкерами подходящей длины, обеспечивающей их расположение в виде диатела.

[204] Различные CD3-связывающие AF2 по настоящему изобретению были специально модифицированы для повышения их стабильности в описанных в данном документе вариантах осуществления полипептидов. Белковая агрегация антител продолжает оставаться серьезной проблемой для возможности их разработки и остается основным направлением деятельности при получении антител. Агрегация антитела может быть вызвана частичным разворачиванием его доменов, что приводит к ассоциации мономер-мономер с последующим зародышеобразованием и ростом агрегатов. Хотя на склонность антител и белков на основе антител к агрегации могут влиять внешние экспериментальные условия, она сильно зависит от внутренних свойств антител, определяемых их последовательностями и структурами. Хотя хорошо известно, что белки в своих свернутых состояниях стабильны лишь в незначительной степени, зачастую менее понятно, что большинство белков по своей природе склонны к агрегации в их развернутых или частично развернутых состояниях, и получаемые агрегаты могут быть чрезвычайно стабильными и долгоживущими. Также было показано, что снижение склонности к агрегации сопровождается увеличением титра экспрессии, что демонстрирует, что снижение агрегации белка является благоприятным на протяжении всего процесса разработки и может приводить к более эффективному пути к клиническим исследованиям. В случае с терапевтическими белками агрегаты являются значительным фактором риска пагубных иммунных реакций у пациентов и могут образовываться с помощью различных механизмов. Контроль агрегации может позволить улучшить стабильность, технологичность, коэффициенты потери, безопасность, составление, титры, иммуногенность и растворимость белков. Внутренние свойства белков, такие как размер, гидрофобность, электростатические характеристики и распределение заряда, играют важные роли в растворимости белков. Было показано, что низкая растворимость терапевтических белков вследствие поверхностной гидрофобности затрудняет разработку состава и может приводить к неудовлетворительному биораспределению, нежелательному фармакокинетическому поведению и иммуногенности in vivo. Уменьшение общей поверхностной гидрофобности моноклональных антител-кандидатов также может

обеспечивать преимущества и экономию средств, связанные с очисткой и режимами введения доз. Отдельные аминокислоты могут быть идентифицированы с помощью структурного анализа как содействующие потенциальной агрегации в антителе и могут быть локализованы в CDR, а также в каркасных областях. В частности, можно предположить, что остатки имеют высокий риск приведения к проблемам гидрофобности в данном антителе. В одном варианте осуществления в настоящем изобретении предусмотрен AF2, обладающий способностью специфично связывать CD3, при этом AF2 имеет по меньшей мере одну аминокислотную замену гидрофобной аминокислоты в каркасной области по сравнению с исходным антителом или фрагментом антитела, где гидрофобная аминокислота выбрана из изолейцина, лейцина или метионина. В другом варианте осуществления AF2 для CD3 имеет по меньшей мере две аминокислотные замены гидрофобных аминокислот в одной или более каркасных областях, где гидрофобные аминокислоты выбраны из изолейцина, лейцина или метионина.

[205] Изменения суммарного электрического заряда полипептида, в частности, в отношении антител или фрагментов антител, содержащих конкретные варианты осуществления настоящего изобретения, изложенные в данном документе, учитывались при разработке последовательностей AF2 вариантов осуществления, описанных в данном документе, где отдельные аминокислотные замены были сделаны относительно исходного антитела, используемого в качестве отправной точки. Релевантной исходя из этих соображений относительно конструкции является изоэлектрическая точка (pI) полипептида, которая представляет собой pH, при котором антитело или фрагмент антитела не имеет суммарного электрического заряда. Антитело или фрагмент антитела обычно имеют суммарный положительный заряд, который имеет тенденцию коррелировать с повышенным клиренсом в крови и удержанием в тканях, обычно с более коротким периодом полужизни, тогда как суммарный отрицательный заряд приводит к уменьшению поглощения тканями и более длительному периоду полужизни. Можно производить манипуляции с этим зарядом посредством мутаций остатков каркасных областей. Изоэлектрическая точка полипептида может быть определена математически (например, путем вычислений) или экспериментально при помощи анализа in vitro. В некоторых вариантах осуществления изоэлектрические точки AF1 и AF2 разработаны так, что они находятся в конкретном диапазоне друг относительно друга, что тем самым способствует стабильности.

[206] В одном варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, содержащий CDR-L и CDR-H, где AF2 (a) специфично связывается с кластером дифференцировки 3 (CD3) в Т-клеточном рецепторе и (b) содержит CDR-H1, CDR-H2 и CDR-H3, имеющие соответственно аминокислотные последовательности под SEQ ID NO: 742, 743 и 744. В другом варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, содержащий CDR-L и CDR-

H, где AF2 (a) специфично связывается с кластером дифференцировки 3 (CD3) в Т-клеточном рецепторе; (b) содержит CDR-H1, CDR-H2 и CDR-H3, имеющие соответственно аминокислотные последовательности под SEQ ID NO: 742, 743 и 744; и (c) содержит CDR-L, где CDR-L содержит CDR-L1, имеющую аминокислотную последовательность под SEQ ID NO: 735 или 736, CDR-L2, имеющую аминокислотную последовательность под SEQ ID NO: 738 или 739, и CDR-L3, имеющую аминокислотную последовательность под SEQ ID NO: 740. В другом варианте осуществления вышеупомянутые варианты осуществления AF2 из данного абзаца дополнительно содержат каркасные области легкой цепи (FR-L) и каркасные области тяжелой цепи (FR-H), где AF2 содержит FR-L1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 746, FR-L2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательности под SEQ ID NO: 747, FR-L3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под любым из SEQ ID NO: 748-751, FR-L4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 754, FR-H1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 755 или SEQ ID NO: 756, FR-H2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 759, FR-H3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 760, и FR-H4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 764. В другом варианте осуществления AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов содержит каркасные области легкой цепи (FR-L) и каркасные области тяжелой цепи (FR-H), где AF2 содержит FR-L1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 746, FR-L2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность

последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 747, FR-L3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 748, FR-L4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 754, FR-H1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 755, FR-H2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 759, FR-H3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 760, и FR-H4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 764. В другом варианте осуществления AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов содержит каркасные области легкой цепи (FR-L) и каркасные области тяжелой цепи (FR-H), где AF2 содержит FR-L1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 746, FR-L2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 747, FR-L3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 749, FR-L4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 754, FR-H1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 755, FR-H2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной

последовательностью под SEQ ID NO: 759, FR-H3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 760, и FR-H4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 764. В другом варианте осуществления AF2 из описанных в данном документе вариантов осуществления полипептидов для человека или животного содержит каркасные области легкой цепи (FR-L) и каркасные области тяжелой цепи (FR-H), где AF2 содержит FR-L1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 746, FR-L2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 747, FR-L3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 750, FR-L4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 754, FR-H1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 755, FR-H2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 759, FR-H3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 760, и FR-H4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 764. В другом варианте осуществления AF2 из описанных в данном документе вариантов осуществления полипептидов для человека или животного содержит каркасные области легкой цепи (FR-L) и каркасные области тяжелой цепи (FR-H), где AF2 содержит FR-L1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 746, FR-L2, демонстрирующую по меньшей мере

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 747, FR-L3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 751, FR-L4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 754, FR-H1, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 756, FR-H2, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 759, FR-H3, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 760, и FR-H4, демонстрирующую по меньшей мере 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичность последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 764.

[207] В другом варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, где AF2 содержит аминокислотную последовательность вариабельной области тяжелой цепи (VH), характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичностью последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 766 или SEQ ID NO: 769. В другом варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, где AF2 содержит аминокислотную последовательность вариабельной области легкой цепи (VL), характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичностью последовательности или являющуюся идентичной с аминокислотной последовательностью под любым из SEQ ID NO: 765, 767, 768, 770 или 771. В другом варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, где AF2 содержит аминокислотную последовательность вариабельной области тяжелой цепи (VH), характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичностью последовательности или являющуюся идентичной с аминокислотной последовательностью под SEQ ID NO: 766 или SEQ ID NO: 769, и аминокислотную

последовательность вариабельной области легкой цепи (VL), характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентичностью последовательности или являющуюся идентичной с аминокислотной последовательностью под любым из SEQ ID NO: 765, 767, 768, 770 или 771.

[208] В другом варианте осуществления в настоящем изобретении предусмотрен AF2 для применения в любом из описанных в данном документе вариантов осуществления полипептидов, где AF2 содержит аминокислотную последовательность, характеризующуюся по меньшей мере 95%, 96%, 97%, 98%, 99% идентичностью последовательности или являющуюся идентичной по отношению к аминокислотной последовательности под любым из SEQ ID NO: 776-780.

[209] В другом аспекте в настоящем изобретении предусмотрены антигенсвязывающие фрагменты AF2, которые связываются с белковым комплексом CD3 и которые характеризуются повышенной стабильностью по сравнению с CD3-связывающими антителами или антигенсвязывающими фрагментами, известными из уровня техники. Кроме того, антигенсвязывающие фрагменты для CD3 по настоящему изобретению разработаны для придания более высокой степени стабильности композициям на основе химерных биспецифических антигенсвязывающих фрагментов, в которые они интегрированы, что приводит к улучшенной экспрессии и извлечению слитого белка, увеличению срока хранения и повышению стабильности при введении человеку или животному. В одном подходе AF2 для CD3 по настоящему изобретению разработаны таким образом, что они характеризуются более высокой степенью термической стабильности по сравнению с определенными CD3-связывающими антителами и антигенсвязывающими фрагментами, известными из уровня техники. В результате этого AF2 для CD3, используемый в качестве компонентов композиций на основе химерных биспецифических антигенсвязывающих фрагментов, в которые они интегрированы, демонстрирует благоприятные фармацевтические свойства, включая высокую термостабильность и низкую склонность к агрегации, что приводит к улучшенным экспрессии и извлечению в ходе изготовления и хранения, а также способствует длительному периоду полужизни в сыворотке крови. Биофизические свойства, такие как термостабильность, часто ограничены вариабельными доменами антител, которые сильно различаются по своим внутренним свойствам. Высокая термическая стабильность часто ассоциирована с высокими уровнями экспрессии и другими необходимыми свойствами, включая меньшую подверженность агрегации (Buchanan A, et al. Engineering a therapeutic IgG molecule to address cysteinylation, aggregation and enhance thermal stability and expression. MAbs 2013; 5:255). Термическую стабильность определяют путем измерения «температуры плавления» (T_m), которая определяется как температура, при которой половина молекул денатурируется. Температура плавления каждого гетеродимера указывает на его термическую стабильность. Из уровня техники известны анализы in vitro для определения T_m, в том числе способы, описанные в примерах, приведенных ниже. Точку плавления гетеродимера

можно измерить с помощью таких методик, как дифференциальная сканирующая калориметрия (Chen et al (2003) Pharm Res 20:1952-60; Ghirlando et al (1999) Immunol Lett 68:47-52). В качестве альтернативы, термическую стабильность гетеродимера можно измерить с помощью кругового дихроизма (Murray et al. (2002) J. Chromatogr Sci 40:343-9), или как описано в примерах, приведенных ниже.

[210] Кривые тепловой денатурации связывающих CD3 фрагментов и биспецифических антител к CD3, содержащих указанный связывающий фрагмент для CD3 и эталонный связывающий фрагмент по настоящему изобретению, демонстрируют, что конструкции по настоящему изобретению являются более устойчивыми к тепловой денатурации, чем антигенсвязывающий фрагмент, состоящий из последовательности, показанной под SEQ ID NO: 781, или контрольное биспецифическое антитело, где указанный контрольный биспецифический антигенсвязывающий фрагмент содержит SEQ ID NO: 781 и эталонный антигенсвязывающий фрагмент, который связывается с вариантом осуществления EGFR, описанным в данном документе. В одном варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2 для CD3 из описанных в данном документе вариантов осуществления, где T_m AF2 на по меньшей мере 2°C выше, или на по меньшей мере 3°C выше, или на по меньшей мере 4°C выше, или на по меньшей мере 5°C выше, или на по меньшей мере 6°C выше, или на по меньшей мере 7°C выше, или на по меньшей мере 8°C выше, или на по меньшей мере 9°C выше, или на по меньшей мере 10°C выше, чем T_m антигенсвязывающего фрагмента, состоящего из последовательности под SEQ ID NO: 781, как определено по увеличению температуры плавления в анализе in vitro.

[211] В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который специфично связывает CD3 человека или макака-крабоеда с константой диссоциации (K_d), составляющей от приблизительно 10 нМ до приблизительно 400 нМ, или от приблизительно 50 нМ до приблизительно 350 нМ, или от приблизительно 100 нМ до 300 нМ, как определено в анализе связывания антигена in vitro, предусматривающем антиген CD3 человека или макака-крабоеда. В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который специфично связывает CD3 человека или макака-крабоеда с константой диссоциации (K_d) ниже чем приблизительно 10 нМ, или приблизительно 50 нМ, или приблизительно 100 нМ, или приблизительно 150 нМ, или приблизительно 200 нМ, или приблизительно 250 нМ, или приблизительно 300 нМ, или приблизительно 350 нМ, или ниже чем приблизительно 400 нМ, как определено в анализе связывания антигена in vitro. Для ясности, антигенсвязывающий фрагмент с K_d, составляющей 400, связывает свой лиганд слабее, чем фрагмент с K_d, составляющей 10 нМ. В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления

композиции для человека и животных содержат AF2, который специфично связывает CD3 человека или макака-крабоеда с аффинностью связывания, которая в по меньшей мере 2 раза, 3 раза, 4 раза, 5 раз, 6 раз, 7 раз, 8 раз, 9 раз или по меньшей мере 10 раз слабее, чем у антигенсвязывающего фрагмента, состоящего из аминокислотной последовательности под SEQ ID NO: 781, как определено с помощью соответствующих констант диссоциации (K_d) в анализах связывания антигена in vitro. В другом варианте осуществления в настоящем изобретении предусмотрены биспецифические полипептиды, содержащие AF2, который демонстрирует аффинность связывания с CD3, которая в по меньшей мере 2 раза, 3 раза, 4 раза, 5 раз, 6 раз, 7 раз, 8 раз, 9 раз, 10 раз, 20 раз, 50 раз, 100 раз или по меньшей мере 1000 раз слабее по сравнению с таковой у описанных в данном документе вариантов осуществления AF1 для EGFR, которые включены в состав полипептидов человека или животного, как определено с помощью соответствующих констант диссоциации (K_d) в анализе связывания антигена in vitro. Аффинность связывания композиций для человека или животного с лигандами-мишенями можно анализировать с использованием анализов связывания или конкурентного связывания, таких как анализы Biacore с иммобилизованными на чипе рецепторами или связывающими белками или анализы ELISA, описанные в патенте США № 5534617, анализы, описанные в примерах в данном документе, радиорецепторные анализы или другие анализы, известные из уровня техники. Затем может быть определена константа аффинности связывания с использованием стандартных способов, таких как анализ по Скэтчарду, описанный в van Zoelen, et al., Trends Pharmacol Sciences (1998) 19(12):487, или другие способы, известные из уровня техники.

[212] В связанном аспекте в настоящем изобретении предусмотрен AF2, который связывается с CD3 и включен в состав композиций на основе химерных биспецифических полипептидов, которые разработаны таким образом, что они имеют изоэлектрическую точку (pI), что придает повышенную стабильность композициям по настоящему изобретению по сравнению с соответствующими композициями, содержащими CD3-связывающие антитела или антигенсвязывающие фрагменты, известные из уровня техники. В одном варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который связывается с CD3, где AF2 демонстрирует pI, составляющую от 6,0 до 6,6 включительно. В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который связывается с CD3, где AF2 демонстрирует pI, которая на по меньшей мере 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9 или 1,0 единицы pH ниже, чем pI эталонного антигенсвязывающего фрагмента, состоящего из последовательности, показанной под SEQ ID NO: 781. В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который связывается с CD3, слитый с AF1, который связывается с антигеном EGFR, где AF2 демонстрирует pI, которая находится в

пределах по меньшей мере 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4 или 1,5 единицы pH от pI AF1, который связывает антиген EGFR или его эпитоп. В другом варианте осуществления полипептиды из любого из описанных в данном документе вариантов осуществления композиции для человека или животного содержат AF2, который связывается с CD3, слитый с AF1, который связывается с антигеном EGFR, где AF2 демонстрирует pI, которая находится в пределах от по меньшей мере приблизительно 0,1 до приблизительно 1,5, или от по меньшей мере приблизительно 0,3 до приблизительно 1,2, или от по меньшей мере приблизительно 0,5 до приблизительно 1,0, или от по меньшей мере приблизительно 0,7 до приблизительно 0,9 единицы pH от pI AF1. В частности, предполагается, что при такой конструкции, в которой pI двух антигенсвязывающих фрагментов находится в таких диапазонах, получаемые слитые антигенсвязывающие фрагменты будут придавать более высокую степень стабильности композициям на основе химерных биспецифических антигенсвязывающих фрагментов, в которые они интегрированы, что приводит к улучшенной экспрессии и повышенному извлечению слитого белка в растворимой неагрегированной форме, увеличенному сроку хранения составленных композиций на основе химерных биспецифических полипептидов и повышенной стабильности при введении композиции человеку или животному. Другими словами, наличие AF2 и AF1 в пределах относительно узкого диапазона pI может обеспечивать возможность выбора буфера или другого раствора, в котором как AF2, так и AF1 являются стабильными, что тем самым способствует общей стабильности композиции.

[213] В определенных вариантах осуществления VL и VH антигенсвязывающих фрагментов слиты с помощью относительно длинных линкеров, содержащих 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 или 35 гидрофильных аминокислот, которые при соединении друг с другом обладают характеристикой гибкости. В одном варианте осуществления VL и VH из любого из описанных в данном документе вариантов осуществления scFv связаны с помощью относительно длинных линкеров из гидрофильных аминокислот, которые представляют собой GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 8058)

TGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGT (SEQ ID NO: 8059),

GATPPETGAETESPGETTGGSAESEPPGEG (SEQ ID NO: 8060) или

GSAAPTAGTTPSASPAPPTGGSSAAGSPST (SEQ ID NO: 8061). В другом варианте осуществления AF1 и AF2 связаны друг с другом с помощью короткого линкера из гидрофильных аминокислот, имеющего 3, 4, 5, 6 или 7 аминокислот. В одном варианте осуществления последовательности короткого линкера представляют собой SGGGGS (SEQ ID NO: 8062), GGGGS (SEQ ID NO: 8063), GGSGGS (SEQ ID NO: 8064), GGS или GSP. В другом варианте осуществления в настоящем изобретении предусмотрены композиции, содержащие одноцепочечное диатело, в котором после сворачивания первый домен (VL или VH) спарен с последним доменом (VH или VL) с образованием одного scFv, а два домена в середине спарены с образованием другого scFv, в котором первый и второй домены, а также третий и последний домены слиты друг с другом с помощью

одного из указанных выше коротких линкеров, а второй и третий вариабельные домены слиты с помощью одного из указанных выше относительно длинных линкеров. Как будет понятно специалисту в данной области, выбор короткого линкера и относительно длинного линкера должен предотвратить неправильное спаривание прилежащих вариабельных доменов, тем самым облегчая образование конфигурации одноцепочечного диатела, содержащей VL и VH первого антигенсвязывающего фрагмента и второго антигенсвязывающего фрагмента.

Таблица 6b. Иллюстративные последовательности CDR для CD3

Конструкция	**ОБЛАСТЬ CDR**	**Аминокислотная последовательность**	**SEQ ID NO:**
3.23, 3.30, 3.31, 3.32	CDR-L1	RSS**N**GAVT**S**SNYAN	735
3.24	CDR-L1	RSS**N**G**E**VTTSNYAN	736
3.33, 3.9	CDR-L1	RSSTGAVTTSNYAN	737
3.23, 3.30, 3.31, 3.32, 3.9, 3.33	CDR-L2	GTNKRAP	738
3.24	CDR-L2	GT**I**KRAP	739
3.23, 3.24, 3.30, 3.31, 3.32	CDR-L3	ALWY**P**NLWVF	740
3.33, 3.9	CDR-L3	ALWYSNLWVF	741
3.23, 3.24, 3.30, 3.31, 3.32, 3.9	CDR-H1	GFTFNTYAMN	742
3.23, 3.24, 3.30, 3.31, 3.32, 3.9	CDR-H2	RIRSKYNNYATYYADSVKD	743
3.23. 3.24, 3.30, 3.31, 3.32	CDR-H3	H**E**NFGNSYVSWFA**H**	744
3.9	CDR-H3	HGNFGNSYVSWFAY	745

Таблица 6c. Иллюстративные последовательности FR для CD3

Конструкция	ОБЛАСТЬ FR	Аминокислотная последовательность	SEQ ID NO:
3.23, 3.24, 3.30, 3.31, 3.32, 3.9, 3.33	FR-L1	ELVVTQEPSLTVSPGGTVTLTC	746
3.23, 3.24, 3.30, 3.31, 3.32, 3.9, 3.33	FR-L2	WVQQKPGQAPRGLIG	747
3.23, 3.24	FR-L3	GTPARFSGSLLGGKAALTLSGVQPEDEA**V**YYC	748
3.30	FR-L3	GTPARFSGS**S**LGGKAALTLSGVQPEDEA**V**YYC	749
3.31	FR-L3	GTPARFSGSLLGG**S**AALTLSGVQPEDEA**V**YYC	750
3.32	FR-L3	GTPARFSGS**S**LGG**S**AALTLSGVQPEDEA**V**YYC	751
3.9	FR-L3	GTPARFSGSLLGGKAALTLSGVQPEDEAEYYC	752
3.33	FR-L3	GTPARFSGS**S**LGG**S**AALTLSGVQPEDEAEYYC	753
3.23, 3.24, 3.30,	FR-L4	GGGTKLTVL	754

Конструкция	ОБЛАСТЬ FR	Аминокислотная последовательность	SEQ ID NO:
3.31, 3.32, 3.9			
3.23, 3.24	FR-H1	EVQLLESGGG**I**VQPGGSLKLS CAAS	755
3.30, 3.31, 3.32	FR-H1	EVQL**Q**ESGGG**I**VQPGGSLKLS CAAS	756
3.33	FR-H1	EVQL**Q**ESGGGLVQPGGSLKL SCAAS	757
3.9	FR-H1	EVQLLESGGGLVQPGGSLKLS CAAS	758
3.23, 3.24, 3.30, 3.31, 3.32, 3.9, 3.33	FR-H2	WVRQAPGKGLEWVA	759
3.23, 3.24, 3.30, 3.31, 3.32	FR-H3	RFTISRDDSKNT**V**YLQMNNL KTEDTAVYYCVR	760
3.9	FR-H3	RFTISRDDSKNTAYLQMNNL KTEDTAVYYCVR	762
3.23, 3.24, 3.30, 3.31, 3.32, 3.9, 3.33	FR-H4	WGQGTLVTVSS	764

Таблица 6d. Иллюстративные последовательности VL и VH

Конструкция	ОБЛАСТЬ	Аминокислотная последовательность	SEQ ID NO:
3.23	VL	ELVVTQEPSLTVSPGGTVTLTC RSS**N**GAVT**S**SNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GSLLGGKAALTLSGVQPEDEA **V**YYCALWY**P**NLWVFGGGTKL TVL	765
3.23, 3.24	VH	EVQLLESGGG**I**VQPGGSLKLSC AASGFTFNTYAMNWVRQAPG KGLEWVARIRSKYNNYATYYA DSVKDRFTISRDDSKNT**V**YLQ MNNLKTEDTAVYYCVRH**E**NF GNSYVSWFA**H**WGQGTLVTVS S	766
3.24	VL	ELVVTQEPSLTVSPGGTVTLTC RSS**N**G**E**VTTSNYANWVQQKPG QAPRGLIGGT**I**KRAPGTPARFS GSLLGGKAALTLSGVQPEDEA **V**YYCALWY**P**NLWVFGGGTKL TVL	767
3.30	VL	ELVVTQEPSLTVSPGGTVTLTC RSS**N**GAVT**S**SNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GS**S**LGGKAALTLSGVQPEDEA **V**YYCALWY**P**NLWVFGGGTKL TVL	768
3.30, 3.31, 3.32	VH	EVQL**Q**ESGGG**I**VQPGGSLKLSC AASGFTFNTYAMNWVRQAPG	769

		KGLEWVARIRSKYNNYATYYA DSVKDRFTISRDDSKNT**V**YLQ MNNLKTEDTAVYYCVRH**E**NF GNSYVSWFA**H**WGQGTLVTVS S	
3.31	VL	ELVVTQEPSLTVSPGGTVTLTC RSS**N**GAVT**S**SNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GSLLGG**S**AALTLSGVQPEDEA**V** YYCALWY**P**NLWVFGGGTKLT VL	770
3.32	VL	ELVVTQEPSLTVSPGGTVTLTC RSS**N**GAVT**S**SNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GS**S**LGG**S**AALTLSGVQPEDEA**V** YYCALWY**P**NLWVFGGGTKLT VL	771
3.9	VL	ELVVTQEPSLTVSPGGTVTLTC RSSTGAVTTSNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GSLLGGKAALTLSGVQPEDEA EYYCALWYSNLWVFGGGTKL TVL	7**72**
3.9	VH	EVQLLESGGGLVQPGGSLKLSC AASGFTFNTYAMNWVRQAPG KGLEWVARIRSKYNNYATYYA DSVKDRFTISRDDSKNTAYLQ MNNLKTEDTAVYYCVRHGNF GNSYVSWFAYWGQGTLVTVSS	773
3.33	VL	ELVVTQEPSLTVSPGGTVTLTC RSSTGAVTTSNYANWVQQKPG QAPRGLIGGTNKRAPGTPARFS GSSLGGSAALTLSGVQPEDEAE YYCALWYSNLWVFGGGTKLT VL	774
3.33	VH	EVQLQESGGGLVQPGGSLKLS CAASGFTFNTYAMNWVRQAP GKGLEWVARIRSKYNNYATYY ADSVKDRFTISRDDSKNTAYLQ MNNLKTEDTAVYYCVRHGNF GNSYVSWFAYWGQGTLVTVSS	775

Таблица 6e. Иллюстративные последовательности scFv

Конструкция	Аминокислотная последовательность	SEQ ID NO:
3.23	ELVVTQEPSLTVSPGGTVTLTCRSS**N**GAVT**S**SNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGSLLGGKAALTLSGVQPE DEA**V**YYCALWY**P**NLWVFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGEVQLLESGGG**I**VQPGGSLKLSCAASG FTFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADS VKDRFTISRDDSKNT**V**YLQMNNLKTEDTAVYYCVRH**E**NFG NSYVSWFA**H**WGQGTLVTVSS	776
3.24	ELVVTQEPSLTVSPGGTVTLTCRSS**N**G**E**VTTSNYANWVQQK	777

	PGQAPRGLIGGTIKRAPGTPARFSGSLLGGKAALTLSGVQPE DEA**V**YYCALWY**P**NLWVFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGEVQLLESGGG**I**VQPGGSLKLSCAASG FTFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADS VKDRFTISRDDSKNT**V**YLQMNNLKTEDTAVYYCVRH**E**NFG NSYVSWFA**H**WGQGTLVTVSS	
3.30	ELVVTQEPSLTVSPGGTVTLTCRSS**N**GAVT**S**SNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGS**S**LGGKAALTLSGVQPE DEA**V**YYCALWY**P**NLWVFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGEVQL**Q**ESGGG**I**VQPGGSLKLSCAASG FTFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADS VKDRFTISRDDSKNT**V**YLQMNNLKTEDTAVYYCVRH**E**NFG NSYVSWFA**H**WGQGTLVTVSS	778
3.31	ELVVTQEPSLTVSPGGTVTLTCRSS**N**GAVT**S**SNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGSLLGG**S**AALTLSGVQPE DEA**V**YYCALWY**P**NLWVFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGEVQL**Q**ESGGG**I**VQPGGSLKLSCAASG FTFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADS VKDRFTISRDDSKNT**V**YLQMNNLKTEDTAVYYCVRH**E**NFG NSYVSWFA**H**WGQGTLVTVSS	779
3.32	ELVVTQEPSLTVSPGGTVTLTCRSS**N**GAVT**S**SNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGS**S**LGG**S**AALTLSGVQPE DEA**V**YYCALWY**P**NLWVFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGEVQL**Q**ESGGG**I**VQPGGSLKLSCAASG FTFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADS VKDRFTISRDDSKNT**V**YLQMNNLKTEDTAVYYCVRH**E**NFG NSYVSWFA**H**WGQGTLVTVSS	780
3.9	ELVVTQEPSLTVSPGGTVTLTCRSSTGAVTTSNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGSLLGGKAALTLSGVQPE DEAEYYCALWYSNLWVFGGGTKLTVLGATPPETGAETESPG ETTGGSAESEPPGEGEVQLLESGGGLVQPGGSLKLSCAASGF TFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADSV KDRFTISRDDSKNTAYLQMNNLKTEDTAVYYCVRHGNFGN SYVSWFAYWGQGTLVTVSS	781
3.33	ELVVTQEPSLTVSPGGTVTLTCRSSTGAVTTSNYANWVQQK PGQAPRGLIGGTNKRAPGTPARFSGSSLGGSAALTLSGVQPE DEAEYYCALWYSNLWVFGGGTKLTVLGATPPETGAETESPG ETTGGSAESEPPGEGEVQLQESGGGLVQPGGSLKLSCAASGF TFNTYAMNWVRQAPGKGLEWVARIRSKYNNYATYYADSV KDRFTISRDDSKNTAYLQMNNLKTEDTAVYYCVRHGNFGN SYVSWFAYWGQGTLVTVSS	782
4.11	QSVLTQPPSASGTPGQRVTISCSGSSSNIGSNYVYWYQQLPG TAPKLLIYRNNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEA DYYCAAWDDSLSGLWVFGGGTKLTVLGATPPETGAETESPG ETTGGSAESEPPGEGQVQLQQWGGGLVKPGGSLRLSCAASG FTFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKGR FTISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWGQ GTLVTVSS	783
4.12	QAGLTQPPSASGTPGQRVTLSCSGSYSNIGTYYVYWYQQLP GTAPKLLIYSNDQRLSGVPDRFSGSKSGTSASLAISGLQSEDE AAYYCAAWDDSLNGWAFGGGTKLTVLGATPPETGAETESP GETTGGSAESEPPGEGQVQLQQWGGGLVKPGGSLRLSCAAS	784

	GFTFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKG RFTISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWG QGTLVTVSS	
4.13	QPGLTQPPSASGTPGQRVTLSCSGRSSNIGSYYVYWYQHLPG MAPKLLIYRNSRRPSGVPDRFSGSKSGTSASLVISGLQSDDEA DYYCAAWDDSLKSWVFGGGTKLTVLGATPPETGAETESPGE TTGGSAESEPPGEGQVQLQQWGGGLVKPGGSLRLSCAASGF TFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKGRF TISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWGQG TLVTVSS	785
4.14	QSVLTQPPSASGTPGQRVTISCSGSSSNIGTNYVYWYQQFPG TAPKLLIYSNNQRPSGVPDRFSGSKSGTSGSLAISGLQSEDEA DYSCAAWDDSLNGWVFGGGTKLTVLGATPPETGAETESPGE TTGGSAESEPPGEGQVQLVQWGGGLVKPGGSLRLSCAASGF TFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKGRF TISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWGQG TLVTVSS	786
4.15	QPGLTQPPSASGTPGQRVTISCSGSSSNIGSNYVYWYQQLPG TAPKLLIYRNNQRPSGVPDRLSGSKSGTSASLAISGLRSEDEA DYYCAAWDDSLSGWVFGGGTKLTVLGATPPETGAETESPGE TTGGSAESEPPGEGQVQLVQWGGGLVKPGGSLRLSCAASGF TFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKGRF TISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWGQG TLVTVSS	787
4.16	QAVLTQPPSASGTPGQRVTISCSGSSSNIGSYYVYWYQQVPG AAPKLLMRLNNQRPSGVPDRFSGAKSGTSASLVISGLRSEDE ADYYCAAWDDSLSGQWVFGGGTKLTVLGATPPETGAETES PGETTGGSAESEPPGEGQVQLQQWGGGLVKPGGSLRLSCAA SGFTFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVK GRFTISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNW GQGTLVTVSS	788
4.17	QAGLTQPPSASGTPGQRVTISCSGSSSNIGSNYVYWYQQLPG TAPKLLIYRNNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEA DYYCATWDASLSGWVFGGGTKLTVLGATPPETGAETESPGE TTGGSAESEPPGEGEVQLVQWGGGLVKPGGSLRLSCAASGF TFSSYSMNWVRQAPGKGLEWVSRINSDGSSTNYADSVKGRF TISRDNAKNTLYLQMNSLRAEDTAVYYCARELRWGNWGQG TLVTVSS	789

Связывающие домены к EpCAM

[214] В некоторых вариантах осуществления настоящее изобретение предусматривает композиции на основе сборки химерных полипептидов, содержащие связывающий домен с аффинностью связывания с опухолеспецифическим маркером EpCAM. В одном варианте осуществления связывающий домен содержит VL и VH, полученные из моноклонального антитела к EpCAM. Моноклональные антитела к EpCAM известны в уровне техники. Иллюстративные неограничивающие примеры моноклональных антител к EpCAM и их последовательностей VL и VH представлены в **таблице 6f**. В одном варианте осуществления настоящее изобретение предусматривает

композицию на основе сборки химерного полипептида, содержащую связывающий домен с аффинностью связывания с опухолеспецифическим маркером EpCAM, содержащий последовательности VL и VH к EpCAM, представленные в **таблице 6f**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, где первый связывающий домен первой части содержит области VH и VL, где каждая из областей VH и VL проявляет по меньшей мере приблизительно 90%, или 91%, или 92%, или 93%, или 94%, или 95%, или 96%, или 97%, или 98%, или 99% идентичность или является идентичной парным последовательностям VL и VH антитела к EpCAM 4D5MUCB, представленного в **таблице 6f**. В другом варианте осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, содержащую связывающий домен с аффинностью связывания с опухолеспецифическим маркером, включающий область CDR-L1, область CDR-L2, область CDR-L3, область CDR-H1, область CDR-H2 и область CDR-H3, каждая из которых получена из соответствующих последовательностей VL и VH, представленных в **таблице 6f**.

Таблица 6f. Моноклональные антитела и последовательности к клеткам-мишеням

Торговое наименование	**Название антитела**	**Мишень**	**Последовательность VH**	**SEQ ID NO: (полноразмерная)**	**SEQ ID NO: (CDR)**	**Последовательность VL**	**SEQ ID NO: (полноразмерная)**	**SEQ ID NO: (CDR)**
Tysabri™	natalizumab	Alpha 4 Integrin	QVQLVQSGAEVKKPGASVKVSCKASGFNIK**DTYIH**WVRQAPGQRLEWMG**RIDPANGYTKYDPKFQG**RVTITADTSASTAYMELSSLRSEDTAVYYCAR**EGYYGNYGVYAMDY**WGQGTLVTVSS	8065	8918, 8919, 8920	DIQMTQSPSSLSASVGDRVTITC**KTSQDINKYMA**WYQQTPGKAPRLLIH**YTSALQP**GIPSRFSGSGSGRDYTFTISSLQPEDIATYYC**LQYDNLWT**FGQGTKVEIK	8164	9178, 9179, 9180
REGN910	nesvacumab	Ang2	EVQLVESGGGLVQPGGSLRLSCAAS**GFTFSSYDIH**WVRQATGKGLEWVSAI**GPAGDTYYPGS**	8066	8921, 8922, 8923	EIVLTQSPGTLSLSPGERATLSCRA**SQSVSSTYLA**WYQQKPGQAPRLLIY**GASSRAT**GIPDRFSGS	8165	9181, 9182, 9183

			VKGRFTISR ENAKNSLY LQMNSLRA GDTAVYYC AR**GLITFG GLIAPFD**Y WGQGTLV TVSS			GSGTDFTL TISRLEPED FAVYYC**Q HYDNSQ**TF GQGTKVEI K		
hMFE2 3		CEA	QVKLEQSG AEVVKPGA SVKLSCKA S**GFNIKDS** YMHWLRQ GPGQRLEW IGWI**DPEN GD**TEYAPK FQGKATFT TDTSANTA YLGLSSLRP EDTAVYYC NEG**TPTGP YYFD**YWG QGTLVTVS S	8067	8924, 8925, 8926	ENVLTQSP SSMSASVG DRVNIACS A**SSSVS**YM HWFQQKP GKSPKLWI YSTSN**LAS** GVPSRFSGS GSGTDYSL TISSMQPED AATYYCQQ **RSSYPL**TF GGGTKLEI K	8166	9184, 9185, 9186
M5A (humani zed T84.66)		CEA	EVQLVESG GGLVQPGG SLRLSCAA SGFNIK**DT YMH**WVRQ APGKGLEW VA**RIDPAN GNSKYADS VKG**RFTIS ADTSKNTA YLQMNSLR AEDTAVYY CAP**FGYYV SDYAMAY** WGQGTLV TVSS	8068	8927, 8928, 8929	DIQLTQSPS SLSASVGD RVTITC**RA GESVDIFG VGFLH**WY QQKPGKAP KLLIY**RAS NLES**GVPS RFSGSGSRT DFTLTISSL QPEDFATY YC**QQTNE DPYT**FGQG TKVEIK	8167	9187, 9188, 9189
M5B (humani zed T84.66)		CEA	EVQLVESG GGLVQPGG SLRLSCAA SGFNIK**DT YM**HWVRQ APGKGLEW VA**RIDPAN GNSKYVP KFQG**RATI SADTSKNT AYLQMNSL RAEDTAVY	8069	8930, 8931, 8929	DIQLTQSPS SLSASVGD RVTITC**RA GESVDIFG VGFLH**WY QQKPGKAP KLLIY**RAS NLES**GVPS RFSGSGSRT DFTLTISSL QPEDFATY YC**QQTNE**	8167	9187, 9188, 9189

			YCAP**FGYY VSDYAMA Y**WGQGTL VTVSS			**DPYT**FGQG TKVEIK		
CEA-Cide	Labetuzumab (MN-14)	CEACAM5	EVQLVESG GGVVQPGR SLRLSCSAS GFDFT**TYW MS**WVRQA PGKGLEWI G**EIHPDSS TINYAPSL KD**RFTISRD NAKNTLFL QMDSLRPE DTGVYFCA S**LYFGFPW FAY**WGQG TPVTVSS	8070	8932, 8933, 8934	DIQLTQSPS SLSASVGD RVTITC**KA SQDVGTSV A**WYQQKP GKAPKLLI Y**WTSTRH T**GVPSRFS GSGSGTDF TFTISSLQP EDIATYYC **QQYSLYRS** FGQGTKVE IK	8168	9190, 9191, 9192
CEA-Scan	arcitumomab	CEACAM5	EVKLVESG GGLVQPGG SLRLSCATS **GFTFTDYY MN**WVRQP PGKALEWL G**FIGNKAN GYTTEYSA S**VKGRFTIS RDKSQSILY LQMNTLRA EDSA**TYYC TRDR**GLRF YFDYWGQ GTTLTVSS	8071	8935, 8936, 8937	QTVLSQSP AILSASPGE KVTMTC**R ASSSVTYI HW**YQQKP GSSPKSWI YA**TSNLAS G**VPARFSG SGSGTSYSL TISRVEAED AATYYC**Q HWSSKPPT** FGGGTKLEI KR	8169	9193, 9194, 9195
MT110		CEACAM5	EVQLVESG GGLVQPGR SLRLSCAA SGFTVS**SY WMH**WVR QAPGKGLE WVG**FIRNK ANGGTTE YAASVKG** RFTISRDDS KNTLYLQM NSLRAEDT AVYYCAR **DRGLRFYF DY**WGQGT TVTVSS	8072	8938, 8939, 8940	QAVLTQPA SLSASPGAS ASLTC**TLR RGINVGAY SIY**WYQQK PGSPPQYLL R**YKSDSDK QQGS**GVSS RFSASKDA SANAGILLI SGLQSEDE ADYYC**MI WHSGASA V**FGGGTKL TVL	8170	9196, 9197, 9198
MT103	blinatumomab	CD19	QVQLQQSG AELVRPGS	8073	8941, 8942,	DIQLTQSPA SLAVSLGQ	8171	9199, 9200,

			SVKISCKAS GYAFS**SYW MN**WVKQR PGQGLEWI G**QIWPGD GDTNYNG KFKG**KATL TADESSST AYMQLSSL ASEDSAVY FCAR**RETT TVGRYYY AMDY**WGQ GTTVTVSS		8943	RATISC**KAS QSVDYDG DSY**LNWY QQIPGQPPK LLI**YDASN LVS**GIPPRF SGSGSGTD FTLNIHPVE KVDAATY HC**QQSTED PWT**FGGGT KLEIK		9201
Arzerra	ofatumumab	CD20	EVQLVESG GGLVQPGR SLRLSCAA SGFTFN**DY AMH**WVRQ APGKGLEW VS**TISWNS GSIGYADS VKG**RFTIS RDNAKKSL YLQMNSLR AEDTALYY CAK**DIQYG NYYYGMD V**WGQGTT VTVSS	8074	8944, 8945, 8946	EIVLTQSPA TLSLSPGER ATLSC**RAS QSVSSYLA** WYQQKPG QAPRLLIY **DASNRAT**G IPARFSGSG SGTDFTLTI SSLEPEDFA VYYC**QQR SNWPIT**FG QGTRLEIK	8172	9202, 9203, 9204
Bexxar™	tositumomab	CD20	QAYLQQSG AELVRPGA SVKMSCKA SGYTFT**SY NMH**WVKQ TPRQGLEW IG**AIYPGN GDTSYNQ KFKG**KATL TVDKSSST AYMQLSSL TSEDSAVY FCAR**VVYY SNSYWYFD V**WGTGTT VTVSG	8075	8947, 8948, 8949	QIVLSQSPA ILSASPGEK VTMTC**RAS SSVSYMH** WYQQKPG SSPKPWIY **APSNLAS**G VPARFSGS GSGTSYSL TISRVEAED AATYYC**Q QWSFNPPT** FGAGTKLE LK	8173	9205, 9206, 9207
GAZYVA	Obinutuzumab	CD20	QVQLVQSG AEVKKPGS SVKVSCKA SGYAFS**YS WIN**WVRQ APGQGLEW	8076	8950, 8951, 8952	DIVMTQTP LSLPVTPGE PASISC**RSS KSLLHSNG ITYLY**WYL QKPGQSPQ	8174	9208, 9209, 9210

			MG**RIFPGD GDTDYNG KFKG**RVTI TADKSTST AYMELSSL RSEDTAVY YCAR**NVFD GYWLVY**W GQGTLVTV SS			LLIY**QMSN LVS**GVPDR FSGSGSGT DFTLKISRV EAEDVGVY YC**AQNLEL PYT**FGGGT KVEIK		
	Ocrelizumab/ 2H7 v16	CD20	EVQLVESG GGLVQPGG SLRLSCAA S**GYTFTSY N**MHWVRQ APGKGLEW VGA**IYPGN GDT**SYNQK FKGRFTISV DKSKNTLY LQMNSLRA EDTAVYYC AR**VVYYSN SYWYFDV** WGQGTLV TVSS	8077	8953, 8954, 8949	DIQMTQSP SSLSASVG DRVTITC**R ASSSVSYM H**WYQQKP GKAPKPLI Y**APSNLAS** GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ WSFNPPT**F GQGTKVEI K	8175	8911, 9206, 9207
Rituxan™	rituximab	CD20	QVQLQQPG AELVKPGA SVKMSCKA S**GYTFTSY N**MHWVKQ TPGRGLEW IGA**IYPGN GDT**SYNQK FKGKATLT ADKSSSTA YMQLSSLT SEDSAVYY C**ARSTYYG GDWYFNV** WGAGTTV TVSA	8078	8953, 8954, 8955	QIVLSQSPA ILSASPGEK VTMTCRAS **SSVSY**IHW FQQKPGSS PKPWIY**AT S**NLASGVP VRFSGSGS GTSYSLTIS RVEAEDAA TYYC**QQW TSNPPT**FG GGTKLEIK	8176	9211, 9212, 9213
Zevalin™	ibritumomab tieuxetan	CD20	QAYLQQSG AELVRPGA SVKMSCKA S**GYTFTSY N**MHWVKQ TPRQGLEW IG**AIYPGN GDTSYNQ KF**KGKATL TVDKSSST	8079	8953, 8956, 8949	QIVLSQSPA ILSASPGEK VTMTC**RAS SSVSYMH** WYQQKPG SSPKPWIY **APSNLAS**G VPARFSGS GSGTSYSL TISRVEAED	8173	8911, 9206, 9207

			AYMQLSSL TSEDSAVY FCAR**VVYY SNSYWYFD V**WGTGTT VTVSA			AATYYC**Q QWSFNPPT** FGAGTKLE LK		
Mylotarg	Gemtuzumab (hP67.6)	CD33	QLVQSGAE VKKPGSSV KVSCKAS**G YTITDSNIH** WVRQAPG QSLEWIG**YI YPYNGGT DYNQKFK N**RATLTVD NPTNTAYM ELSSLRSED TDFYYCVN **GNPWLAY** WGQGTLV TVSS	8080	8957, 8958, 8959	DIQLTQSPS TLSASVGD RVTITC**RAS ESLDNYGI RFLT**WFQQ KPGKAPKL LMY**AASN QGS**GVPSR FSGSGSGTE FTLTISSLQ PDDFATYY C**QQTKEV PWS**FGQGT KVEVK	8177	9214, 9215, 9216
Daratumumab		CD38	EVQLLESG GGLVQPGG SLRLSCAV S**GFTFNSF A**MSWVRQ APGKGLEW VSA**ISGSG GGT**YYAD SVKGRFTIS RDNSKNTL YLQMNSLR AEDTAVYF C**AKDKIL WFGEPVF DY**WGQGT LVTVSS	8081	8960, 8961, 8962	EIVLTQSPA TLSLSPGER ATLSCRAS **QSVSSY**LA WYQQKPG QAPRLLIY **DAS**NRATG IPARFSGSG SGTDFTLTI SSLEPEDFA VYYC**QQR SNWPPT**FG QGTKVEIK	8178	9217, 9218, 9219
	1F6	CD70	QIQLVQSG PEVKKPGE TVKISCKAS **GYTFTNY GMN**WVKQ APGKGLK WMG**WINT YTGEPTYA DAFKG**RFA FSLETSAST AYLQINNL KNEDTATY FCAR**DYGD YGMDY**WG QGTSVTVS	8082	8963, 8964, 8965	DIVLTQSPA SLAVSLGQ RATISC**RAS KSVSTSGY SFMH**WYQ QKPGQPPK LLIY**LASNL ES**GVPARF SGSGSGTD FTLNIHPVE EEDAATY YC**QHSREV PWT**FGGGT KLEIK	8179	9220, 9221, 9222

			S					
	2F2	CD70	QVQLQQSG TELMTPGA SVTMSCKT S**GYTFSTY WIE**WVKQ RPGHGLEW IG**EILGPSG YTDYNEKF KA**KATFTA DTSSNTAY MQLSSLAS EDSAVYYC AR**WDRLY AMDY**WGG GTSVTVSS	8083	8966, 8967, 8968	DIVLTQSPA SLTVSLGQ KTTISC**RAS KSVSTSGY SFMH**WYQ LKPGQSPK LLIY**LASDL PS**GVPARF SGSGSGTD FTLKIHPVE EEDAATY YC**QHSREI PYT**FGGGT KLEIT	8180	9220, 9223, 9224
	2H5	CD70	QVQLVESG GGVVQPGR SLRLSCAA SGFTFS**SYI MH**WVRQA PGKGLEWV A**VISYDGR NKYYADS VKG**RFTIS RDNSKNTL YLQMNSLR AED TAVYYCAR **DTDGYDF DY**WGQGT LVTVSS	8084	8969, 8970, 8971	EIVLTQSPA TLSLSPGER ATLSC**RAS QSVSSYLA** WYQQKPG QAPRLLI**Y DASNRA**TG IPARFSGSG SGTDFTLTI SSLEPEDFA VYYC**QQ RTNWPLT**F GGGTKVEI K	8181	9202, 9225, 9226
	10B4	CD70	QIQLVESG GGVVQPGR SLRLSCAA SGFTFG**YY AMH**WVRQ APGKGLEW VA**VISYDG SIKYYADS VKG**RFTIS RDNSKNTL YLQMNSLR AED TAVYYCAR **EGPYSNYL DY**WGQGT LVTVSS	8085	8972, 8973, 8974	AIQLTQSPS SLSASVGD RVTITC**RAS QGISSALA** WYQQKPG KAPKFLIY **DASSLES**G VPSRFSGSG SGTDFTLTI SSLQPEDFA TYYC**QQ FNSYPFT**F GPGTKVDI K	8182	9227, 9228, 9229
	8B5	CD70	QVQLVESG GGVVQPGR SLRLSCATS GFTFS**DYG**	8086	8975, 8976, 8977	DIQMTQSP SSLSASVG DRVTITC**R ASQGISSW**	8183	9230, 9231, 9232

			MHWVRQAPGKGLEWVA**VIWYDGSNKYYADSVKG**RFTISRDNSKKTLSLQMNSLRAED TAVYYCAR**DSIMVRGDY**WGQGTLVTVSS			**LA**WYQQKPEKAPKSLIY**AASSLQS**GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQYNSYPLT**FGGGTKVEIK		
	18E7	CD70	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**DHGMH**WVRQAPGKGLEWVA**VIWYDGSNKYYADSVKG**RFTISRDNSKNTLYLQMNSLRAED TAVYYCAR**DSIMVRGDY**WGQGTLVTVSS	8087	8978, 8976, 8977	DIQMTQSPSSLSASVGDRVTITC**RASQGISSWLA**WYQQKPEKAPKSLIY**AASSLQS**GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQYNSYPLT**FGGGTKVEIK	8183	9230, 9231, 9232
	69A7	CD70	QVQLQESGPGLVKPSETLSLTCTVSGGSVS**SDYYYWS**WIRQPPGKGLEWLG**YIYYSGSTNYNPSLKS**RVTISVDTSKNQFSLKLRSVTTA DTAVYYCAR**GDGDYGGNCF**DYWGQGTLVTVSS	8088	8979, 8980, 8981	EIVLTQSPATLSLSPGERATLSC**RASQSVSSYLA**WYQQKPGQAPRLLIF**DASNRAT**GIPARFSGSGSGTDFTLTISSLEPEDFAVYYC**QQRSNWPLT**FGGGTKVEIK	8184	9202, 9203, 9233
CE-355621		cMET	QVQLVQSGAEVKKPGASVKVSCKASGYTFT**SYGFS**WVRQAPGQGLEWMG**WISASN**	8089	8982, 8983, 8984	DIQMTQSPSSVSASVGDRVTITC**RASQGINTWLA**WYQQKPGKAPKLLIY**AASSLKS**	8185	9234, 9235, 9236

			GNTYYAQ KLQGRVT MTTDTSTS TAYMELRS LRSDDTAV YYCAR**VY ADYADY**W GQGTLVTV SS			GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ ANSFPLT**F GGGTKVEI K		
LY2875 358	emibetu zumab	cMET	QVQLVQSG AEVKKPGA SVKVSCKA S**GYTFTDY Y**MHWVRQ APGQGLEW MGR**VNPN RRGTT**YN QKFEGRVT MTTDTSTS TAYMELRS LRSDDTAV YYC**ARAN WLDY**WGQ GTTVTVSS	8090	8985, 8986, 8987	DIQMTQSP SSLSASVG DRVTITCSV S**SSVSSIY**L HWYQQKP GKAPKLLI Y**STS**NLAS GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QV YSGYPLT**F GGGTKVEI K	8186	9237, 9238, 9239
MetMA b	onartuz umab	cMET	EVQLVESG GGLVQPGG SLRLSCAA SGYTFT**SY WLH**WVRQ APGKGLEW VG**MIDPSN SDTRFNPN FKD**RFTISA DTSKNTAY LQMNSLRA EDTAVYYC **ATYRSYVT PLDY**WGQ GTLVTVSS	8091	8988, 8989, 8990	DIQMTQSP SSLSASVG DRVTITC**K SSQSLLYT SSQKNYLA** WYQQKPG KAPKLLIY **WASTRES**G VPSRFSGSG SGTDFTLTI SSLQPEDFA TYYC**QQY YAYPWT**F GQGTKVEI K	8187	9240, 9241, 9242
	tremeli mumab (CP-675206, or 11.2.1)	CTLA 4	QVQLVESG GGVVQPGR SLRLSCAA S**GFTFSSY GMH**WVRQ APGKGLEW VA**VIWYD GSNKYYA DSV**KGRFTI SRDNSKNT LYLQMNSL RAEDTAVY YCAR**DPRG**	8092	8991, 8992, 8993	DIQMTQSP SSLSASVG DRVTITC**R ASQSINSY LD**WYQQK PGKAPKLLI Y**AASSLQS** GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ YYSTPFT**F GPGTKVEI	8188	9243, 9231, 9244

			ATLYYYY YGMDVWG QGTTVTVS S			K		
Yervoy	Ipilimumab 10D1	CTLA 4	QVQLVESG GGVVQPGR SLRLSCAA SGFTFS**SYT MH**WVRQA PGKGLEWV T**FISYDGN NKYYADS VKG**RFTIS RDNSKNTL YLQMNSLR AEDTAIYY CAR**TGWL GPFDY**WG QGTLVTVS S	8093	8994, 8995, 8996	EIVLTQSPG TLSLSPGER ATLSC**RAS QSVGSSYL A**WYQQKP GQAPRLLI Y**GAFSRAT** GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q QYGSSPW T**FGQGTKV EIK	8189	9245, 9246, 9247
AGS16F	H16-7.8	ENPP3	QVQLQESG PGLVKPSQ TLSLTCTVS GGSIS**SGG YY**WSWIRQ HPGKGLEW IG**IIYYSGS TYYNPSLK S**RVTISVDT SKNQFSLK LNSVTAAD TAVFYCAR **VAIVTTIP GGMDV**WG QGTTVTVS S	8094	8997, 8998, 8999	EIVLTQSPD FQSVTPKE KVTITC**RA SQSIGISLH** WYQQKPD QSPKLLIK**Y ASQSFS**GV PSRFSGSGS GTDFTLTIN SLEAEDAA TYYC**HQSR SFPWT**FGQ GTKVEIK	8190	9248, 9249, 9250
MT110	solitomab	EpCAM	EVQLLEQS GAELVRPG TSVKISCKA SGYAFT**NY WLG**WVKQ RPGHGLEW IG**DIFPGSG NIHYNEKF KG**KATLTA DKSSSTAY MQLSSLTF EDSAVYFC AR**LRNWD EPMDY**WG QGTTVTVS S	8095	9000, 9001, 9002	ELVMTQSP SSLTVTAG EKVTMSC**K SSQSLLNS GNQKNYL T**WYQQKP GQPPKLLIY **WASTRES**G VPDRFTGS GSGTDFTL TISSVQAED LAVYYC**Q NDYSYPLT** FGAGTKLEI K	8191	9251, 9241, 9252

MT201	Adecatumumab	EpCAM	EVQLLESGGGVVQPGRSLRLSCAASGFTFS**SYGMH**WVRQAPGKGLEWVA**VISYDGSNKYYADSVKG**RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK**DMGWGSGWRPYYYYGMDV**WGQGTTVTVSS	8096	9003, 9004, 9005	ELQMTQSPSSLSASVGDRVTITC**RTSQSISSYLN**WYQQKPGQPPKLLIY**WASTRES**GVPDRFSGSGSGTDFTLTISSLQPEDSATYYC**QQSYDIPYT**FGQGTKLEIK	8192	9253, 9241, 9254
Panorex	Edrecolomab Mab CO17-1A	EpCAM	QVQLQQSGAELVRPGTSVKVSCKAS**GYAFTNYL**IEWVKQRPGQGLEWIGV**INPGSGGT**NYNEKFKGKATLTADKSSSTAYMQLSSLTSDDSAVYFC**ARDGPWFAY**WGQGTLVTVSA	8097	9006, 9007, 9008	NIVMTQSPKSMSMSVGERVTLTCKAS**ENVVTY**VSWYQQKPEQSPKLLIY**GAS**NRYTGVPDRFTGSGSATDFTLTISSVQAEDLADYHC**GQGYSYPYT**FGGGTKLEIK	8193	9255, 9256, 9257
	tucotuzumab	EpCAM	QIQLVQSGPELKKPGETVKISCKAS**GYTFTNYGMN**WVRQAPGKGLKWMG**WINTYTGEPTYAD**DFKGRFVFSLETSASTAFLQLNNLRSEDTATYFCVRFI**SKGDY**WGQGTSVTVSS	8098	8963, 9009, 9010	QILLTQSPAIMSASPGEKVTMTC**SASSSVS**YMLWYQQKPGSSPKPWIF**DTSNLAS**GFPARFSGSGSGTSYSLIISSMEAEDAATYYC**HQRSGYPYT**FGGGTKLEIK	8194	9258, 9259, 9260
UBS-54		EpCAM	VQLQQSDAELVKPGASVKISCKAS**GYTFTDHAIH**WVKQ	8099	9011, 9012, 9013	DIVMTQSPDSLAVSLGERATINC**KSSQSVLYSSNNKNYLA**	8195	9261, 9241, 9262

			NPEQGLEW IG**YFSPGN DDFKYNE**R FKGKATLT ADKSSSTA YVQLNSLT SEDSAVYF CTR**SLNMA Y**WGQGTS VTVSS			WYQQKPG QPPKLLIY **WASTRES**G VPDRFSGS GSGTDFTL TISSLQAED VAVYYC**Q QYYSYPLT** FGGGTKVK ES		
3622W94	323/A3	EpCAM	EVQLVQSG PEVKKPGA SVKVSCKA S**GYTFTNY GMN**WVRQ APGQGLEW MG**WINTY TGEPTYGE** DFKGRFAF SLDTSAST AYMELSSL RSEDTAVY FCARFG**NY VDY**WGQG SLVTVSS	8100	8963, 9014, 9015	DIVMTQSP LSLPVTPGE PASISC**RSSI NKKGSNGI TY**LYWYL QKPGQSPQ LLIYQMSN LASGVPDR F**SGSGS**GT DFTLKISRV EAEDVGVY YC**AQNLEI PRT**FGQGT KVEIK	8196	9263, 9264, 9265
4D5MOCBv2		EpCAM	EVQLVQSG PGLVQPGG SVRISCAAS GYTFT**NYG MN**WVKQA PGKGLEW MG**WINTY TGESTYAD SFKG**RFTF SLDTSASA AYLQINSL RAEDTAVY YCAR**FAIK GD**YWGQG TLLTVSS	8101	9016, 9017, 9018	DIQMTQSP SSLSASVG DRVTITC**RS TKSLLHSN GITYLY**WY QQKPGKAP KLLIY**QMS NLAS**GVPS RFSSSGSGT DFTLTISSL QPEDFATY YC**AQNLEI PRT**FGQGT KVEIK	8197	9266, 9267, 9265
4D5MOCB		EpCAM	EVQLVQSG PGLVQPGG SVRISCAAS GYTFT**NYG MN**WVKQA PGKGLEW MG**WINTY TGESTYAD SFKG**RFTF SLDTSASA AYLQINSL RAEDTAVY	8101	9016, 9017, 9019	DIQMTQSP SSLSASVG DRVTITC**RS TKSLLHSN GITYLY**WY QQKPGKAP KLLIY**QMS NLAS**GVPS RFSSSGSGT DFTLTISSL QPEDFATY YC**AQNLEI**	8198	9266, 9267, 9265

			YCAR**FAIK GDY**WGQG TLLTVSS			**PRT**FGQGT KVELK		
MEDI-547	1C1	EphA2	EVQLLESG GGLVQPGG SLRLSCAA SGFTFS**HY MMA**WVR QAPGKGLE WVS**RIGPS GGPTHYA DSVKG**RFT ISRDNSKNT LYLQMNSL RAEDTAVY YCAGYDSG **YDYVAVA GPAEYFQ H**WGQGTL VTVSS	8102	9020, 9021, 9022	DIQMTQSP SSLSASVG DRVTITC**R ASQSISTW LA**WYQQK PGKAPKLLI Y**KASNLHT** GVPSRFSGS GSGTEFSLT ISGLQPDDF ATYYC**QQ YNSYSRT**F GQGTKVEI K	8199	9268, 9269, 9270
MORAb-003	farletuzumab	FOLR1	EVQLVESG GGVVQPGR SLRLSCSAS **GFTFSGYG** LSWVRQAP GKGLEWV AM**ISSGGS YT**YYADSV KGRFAISR DNAKNTLF LQMDSLRP EDTGVYFC **ARHGDDP AWF**AYWG QGTPVTVS S	8103	9023, 9024, 9025	DIQLTQSPS SLSASVGD RVTITCSVS **SSISSNN**LH WYQQKPG KAPKPWIY **GTS**NLASG VPSRFSGSG SGTDYTFTI SSLQPEDIA TYYC**QQW SSYPYMYT** FGQGTKVE IK	8200	9271, 9272, 9273
M9346A	huMOV19 (vLCv1.00)	FOLR1	QVQLVQSG AEVVKPGA SVKISCKAS GYTFT**GYF MN**WVKQS PGQSLEWI G**RIHPYDG DTFYNQKF QG**KATLTV DKSSNTAH MELLSLTS EDFAVYYC TR**YDGSRA MDY**WGQG TTVTVSS	8104	9026, 9027, 9028	DIVLTQSPL SLAVSLGQ PAIISC**KAS QSVSFAGT SLMH**WYH QKPGQQPR LLIY**RASN LEA**GVPDR FSGSGSKT DFTLNISPV EAEDAATY YC**QQSREY PYT**FGGGT KLEIK	8201	9274, 9275, 9276
M9346	huMOV	FOLR	QVQLVQSG	8105	9026,	DIVLTQSPL	8202	9274,

A	19 (vLCv1.60)	1	AEVVKPGASVKISCKASGYTFT**GYFMN**WVKQSPGQSLEWIG**RIHPYDGDTFYNQKFQG**KATLTVDKSSNTAHMELLSLTSEDFAVYYCTR**YDGSRAMD**YWGQGTTVTVSS		9027, 9029	SLAVSLGQPAIISC**KASQSVSFAGTSLMH**WYHQKPGQQPRLLIY**RASNLEA**GVPDRFSGSGSKTDFTLTISPVEAEDAATYYC**QQSREYPYT**FGGGTKLEIK		9275, 9276
26B3.F2		FOLR1	GPELVKPGASVKISCKASDYSFT**GYFMN**WVMQSHGKSLEWIG**RIFPYNGDTFYNQKFKG**RATLTVDKSSSTAHMELRSLASEDSAVYFCAR**GTHYFDY**WGQGTTLTVSS	8106	9026, 9030, 9031	PASLSASVGETVTITC**RTSENIFSYLA**WYQQKQGISPQLLVY**NAKTLAE**GVPSRFSGSGSGTQFSLKINSLQPEDFGSYYC**QHHYAFPWT**FGGGSKLEIK	8203	9277, 9278, 9279
RG7686	GC33	GPC3	QVQLVQSGAEVKKPGASVKVSCKASGYTFT**DYEMH**WVRQAPGQGLEWMG**ALDPKTGDTAYSQKFKG**RVTLTADKSTSTAYMELSSLTSEDTAVYYCTR**FYSYTY**WGQGTLVTVSS	8107	9032, 9033, 9034	DVVMTQSPLSLPVTPGEPASISC**RSSQSLVHSNGNTYLH**WYLQKPGQSPQLLIY**KVSNRFS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**SQNTHVPPT**FGQGTKLEIK	8204	9280, 9281, 9282
	4A6	GPC3	EVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIA**WVRQMPGKGLEWMG**IIFPGDSD**	8108	9035, 9036, 9037	EIVLTQSPGTLSLSPGERATLSC**RAVQSVSSSYLA**WYQQKPGQAPRLLIY**GASSRAT**	8205	9283, 9182, 9284

			TRYSPSFQ GQVTISAD RSIRTAYLQ WSSLKASD TALYYCAR **TREGYFDY** WGQGTLV TVSS			GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q QYGSSPT**F GGGTKVEI K		
	11E7	GPC3	EVQLVQSG AEVKKPGE SLKISCKGS GYSFT**NYW IA**WVRQMP GKGLEWM G**IIYPGDSD TRYSPSFQ G**QVTISAD KSIRTAYL QWSSLKAS D TAMYYCA R**TREGYFD Y**WGQGTL VTVSS	8109	9038, 9039, 9037	EIVLTQSPG TLSLSPGER ATLSC**RAS QSVSSSYL A**WYQQKP GQAPRLLI Y**GASSRAT** GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q QYGSSPT**F GGGTKVEI K	8206	9285, 9182, 9284
	16D10	GPC3	EVQLVQSG ADVTKPGE SLKISCKVS GYRFT**NY WIG**WMRQ MSGKGLE WMG**IIYPG DSDTRYSP SFQG**HVTI SADKSINT AYLRWSSL KASD TAIYYCAR **TREGFFDY** WGQGTPVT VSS	8110	9040, 9039, 9041	EILLTQSPG TLSLSPGER ATLSC**RAS QSVSSSYL A**WYQQKP GQAPRLLI Y**GASSRAT** GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q QYGSSPT**F GQGTKVEI K	8207	9285, 9182, 9284
AMG-595		HER1(EGFR)	QVQLVESG GGVVQSGR SLRLSCAA S**GFTFRNY GMH**WVRQ APGKGLEW **VAVIWYD GSDKYYA DSVRG**RFT ISRDNSKNT LYLQMNSL RAEDTAVY	8111	9042, 9043, 9044	DTVMTQTP LSSHVTLG QPASISC**RS SQSLVHSD GNTYLS**W LQQRPGQP PRLLIY**RIS RRFS**GVPD RFSGSGAG TDFTLEISR VEAEDVGV YYC**MQST**	8208	9286, 9287, 9288

			YCARDGY **DILTGNPR DFDY**WGQ GTLVTVSS			**HVPRT**FGQ GTKVEIK		
Erubitux™	cetutximab	HER1(EGFR)	QVQLKQSG PGLVQPSQ SLSITCTVS **GFSLTNYG** VHWVRQSP GKGLEWL GV**IWSGG NT**DYNTPF TSRLSINKD NSKSQVFF KMNSLQSN DTAIYYC**A RALTYYD YEFAY**WG QGTLVTVS A	8112	9045, 9046, 9047	DILLTQSPV ILSVSPGER VSFSCRAS **QSIGTN**IH WYQQRTN GSPRLLIK**Y AS**ESISGIPS RFSGSGSG TDFTLSINS VESEDIAD YYC**QQNN NWPTT**FG AGTKLELK	8209	9289, 9290, 9291
GA201	Imgatuzumab	HER1(EGFR)	QVQLVQSG AEVKKPGS SVKVSCKA SGFTFT**DY KIH**WVRQ APGQGLEW MG**YFNPNS GYSTYAQ KFQG**RVTI TADKSTST AYMELSSL RSEDTAVY YCAR**LSPG GYYVMDA**WGQGTTV TVSS	8113	9048, 9049, 9050	DIQMTQSP SSLSASVG DRVTITC**R ASQGINNY LN**WYQQK PGKAPKRLI Y**NTNNLQ T**GVPSRFS GSGSGTEF TLTISSLQP EDFATYYC **LQHNSFPT** FGQGTKLEI K	8210	9292, 9293, 9294
Humax	zalutumumab	HER1(EGFR)	QVQLVESG GGVVQPGR SLRLSCAA SGFTFS**TY GMH**WVRQ APGKGLEW VA**VIWDD GSYKYYG DSVKG**RFT ISRDNSKNT LYLQMNSL RAEDTAVY YCAR**DGIT MVRGVMK DYFDY**WG QGTLVTVS	8114	9051, 9052, 9053	AIQLTQSPS SLSASVGD RVTITC**RAS QDISSALV** WYQQKPG KAPKLLIY **DASSLES**G VPSRFSGSE SGTDFTLTI SSLQPEDFA TYYC**QQFN SYPLT**FGG GTKVEIK	8211	9295, 9228, 9296

			S					
IMC-11F8	necitumumab	HER1(EGFR)	QVQLQESGPGLVKPSQTLSLTCTVSGGSIS**SGDYYWS**WIRQPPGKGLEWIG**YIYYSGSTDYNPSLKS**RVTMSVDTSKNQFSLKVNSVTAADTAVYYCAR**VSIFGVGTFDY**WGQGTLVTVSS	8115	9054, 9055, 9056	EIVMTQSPATLSLSPGERATLSC**RASQSVSSYLA**WYQQKPGQAPRLLIY**DASNRAT**GIPARFSGSGSGTDFTLTISSLEPEDFAVYYC**HQYGSTPLT**FGGGTKAEIK	8212	9202, 9203, 9297
MM-151	P1X	HER1(EGFR)	QVQLVQSGAEVKKPGSSVKVSCKASGGTFS**SYAIS**WVRQAPGQGLEWMG**SIIPIFGTVNYAQKFQG**RVTITADESTSTAYMELSSLRSEDTAVYYCAR**DPSVNLYWYFDL**WGRGTLVTVSS	8116	9057, 9058, 9059	DIQMTQSPSTLSASVGDRVTITC**RASQSISSWWA**WYQQKPGKAPKLLIY**DASSLES**GVPSRFSGSGSGTEFTLTISSLQPDDFATYYC**QQYHAHPTT**FGGGTKVEIK	8213	9298, 9228, 9299
MM-151	P2X	HER1(EGFR)	QVQLVQSGAEVKKPGSSVKVSCKASGGTFG**SYAIS**WVRQAPGQGLEWMG**SIIPIFGAANPAQKSQG**RVTITADESTSTAYMELSSLRSEDTAVYYCAK**MGRGKVAFDI**WGQGTMVTVSS	8117	9057, 9060, 9061	DIVMTQSPDSLAVSLGERATINC**KSSQSVLYSPNNKNYLA**WYQQKPGQPPKLLIY**WASTRES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQYYGSPIT**FGGGTKVEIK	8214	9300, 9241, 9301
MM-151	P3X	HER1(EGFR)	QVQLVQSGAEVKKPGASVKVSCKASGYAFT**SY**	8118	9062, 9063, 9064	EIVMTQSPATLSVSPGERATLSC**RASQSVSSN**	8215	9302, 9303, 9304

			GINWVRQ APGQGLEW MG**WISAY NGNTYYA QKLRG**RV TMTTDTST STAYMELR SLRSDDTA VYYCAR**DL GGYGSGS VPFDP**WG QGTLVTVS S			**LA**WYQQK PGQAPRLLI Y**GASTRAT** GIPARFSGS GSGTEFTLT ISSLQSEDF AVYYC**QD YRTWPRR V**FGGGTKV EIK		
TheraCIM	nimotuzumab	HER1(EGFR)	QVQLQQSG AEVKKPGS SVKVSCKA SGYTFT**NY YIY**WVRQA PGQGLEWI G**GINPTSG GSNFNEKF KT**RVTITA DESSTTAY MELSSLRS EDTAFYFC TR**QGLWF DSDGRGFD F**WGQGTT VTVSS	8119	9065, 9066, 9067	DIQMTQSP SSLSASVG DRVTITC**RS SQNIVHSN GNTYLD**W YQQTPGKA PKLLIY**KVS NRFS**GVPS RFSGSGSG TDFTFTISS LQPEDIATY YC**FQYSHV PWT**FGQGT KLQIT	8216	9305, 9281, 9306
Vectibix™	panitumimab	HER1(EGFR)	QVQLQESG PGLVKPSE TLSLTCTVS **GGSVSSGD YY**WTWIR QSPGKGLE WIGH**IYYS GNT**NYNPS LKSRLTISI DTSKTQFS LKLSSVTA ADTAIYYC **VRDRVTG AFDI**WGQG TMVTVSS	8120	9068, 9069, 9070	DIQMTQSP SSLSASVG DRVTITCQ AS**QDISNY** LNWYQQK PGKAPKLLI Y**DAS**NLET GVPSRFSGS GSGTDFTF TISSLQPEDI ATYFC**QHF DHLPLA**FG GGTKVEIK	8217	9307, 9218, 9308
07D06		HER1(EGFR)	QIQLVQSG PELKKPGE TVKISCKAS **GYTFTEYP** IHWVKQAP GKGFKWM GM**IYTDIG KP**TYAEEF	8121	9071, 9072, 9073	DVVMTQTP LSLPVSLG DQASISCRS S**QSLVHSN GNTY**LHW YLQKPGQS PKLLIY**KVS NR**FSGVPD	8218	9309, 9310, 9311

			KGRFAFSL ETSASTAY LQINNLKN EDTATYFC **VRDRYDSL FDY**WGQG TTLTVSS			RFSGSGSG TDFTLKISR VEAEDLGV YFC**SQSTH VPWT**FGG GTKLEIK		
12D03		HER1(EGFR)	EMQLVESG GGFVKPGG SLKLSCAA SGFAFS**HY DMS**WVRQ TPKQRLEW VA**YIASGG DITYYADT VKG**RFTIS RDNAQNTL YLQMSSLK SEDTAMFY CSR**SSYGN NGDALDF** WGQGTSVT VSS	8122	9074, 9075, 9076	DVVMTQTP LSLPVSLG DQASISC**RS SQSLVHSN GNTYLH**W YLQKPGQS PKLLIY**KVS NRFS**GVPD RFSGSGSG TDFTLKISR VEAEDLGV YFC**SQSTH VLT**FGSGT KLEIK	8219	9280, 9281, 9312
	C1	HER2	QVQLVESG GGLVQPGG SLRLSCAA SGFTFS**SYA MG**WVRQA PGKGLEWV S**SISGSSRY IYYADSVK G**RFTISRDN SKNTLYLQ MNSLRAED TAVYYCAK **MDASGSYF NF**WGQGT LVTVSS	8123	9077, 9078, 9079	QSPSFLSAF VGDRITITC **RASPGIRN YLA**WYQQ KPGKAPKL LIY**AASTL QS**GVPSRF SGSGSGTD FTLTISSLQ PEDFATYY C**QQYNSYP LS**FGGGTK VEIK	8220	9313, 9314, 9315
Erbicin		HER2	QVQLLQSA AEVKKPGE SLKISCKGS GYSFT**SYW IG**WVRQM PGKGLEW MG**IIYPGD SDTRYSPS FQG**QVTIS ADKSISTA YLQWSSLK ASDTAVYY CAR**WRDS PL**WGQGTL	8124	9080, 9039, 9081	QAVVTQEP SFSVSPGGT VTLTC**GLS SGSVSTSY YPS**WYQQT PGQAPRTLI Y**STNTRSS** GVPDRFSG SILGNKAA LTITGAQA DDESDYYC **VLYMGSG QYV**FGGGT KLTVL	8221	9316, 9317, 9318

			VTVSS					
Herceptin	trastuzumab	HER2	EVQLVESGGGLVQPGGSLRLSCAAS**GFNIKDTY**IHWVRQAPGKGLEWVAR**IYPTNGYT**RYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYC**SRWGGDGFYAMDY**WGQGTLVTVSS	8125	9082, 9083, 9084	DIQMTQSPSSLSASVGDRVTITCRAS**QDVNTA**VAWYQQKPGKAPKLLIY**SAS**FLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYC**QQHYTTPPT**FGQGTKVEIK	8222	9319, 9320, 9321
MAGH22	margetuximab	HER2	QVQLQQSGPELVKPGASLKLSCTAS**GFNIKDTY**IHWVKQRPEQGLEWIGR**IYPTNGYT**RYDPKFQDKATITADTSSNTAYLQVSRLTSEDTAVYYC**SRWGGDGFYAMDY**WGQGASVTVSS	8126	9082, 9083, 9084	DIVMTQSHKFMSTSVGDRVSITCKAS**QDVNTA**VAWYQQKPGHSPKLLIY**SAS**FRYTGVPDRFTGSRSGTDFTFTISSVQAEDLAVYYC**QQHYTTPPT**FGGGTKVEIK	8223	9319, 9320, 9321
MM-302	F5	HER2	QVQLVESGGGLVQPGGSLRLSCAASGFTFR**SY****AMS**WVRQAPGKGLEWVS**AISGRG****DNTYYADS****VKG**RFTISRDNSKNTLYLQMNSLRAEDTAVYYC**AKMTSN**	8127	9085, 9086, 9087	QSVLTQPPSVSGAPGQRVTISC**TGSSSNIGAGYG****VH**WYQQLPGTAPKLLIY**GNTNRPS**GVPDRFSGFKSGTSASLAITGLQAEDEADYYC**QFYDSSLS****GWV**FGGGTKLTVL	8224	9322, 9323, 9324

			AFAFDYW GQGTLVTV SS					
Perjeta	pertuzumab	HER2	EVQLVESG GGLVQPGG SLRLSCAA S**GFTFTDY T**MDWVRQ APGKGLEW VAD**VNPNS GGS**IYNQR FKGRFTLS VDRSKNTL YLQMNSLR AEDTAVYY C**ARNLGPS FYFDY**WG QGTLVTVS S	8128	9088, 9089, 9090	DIQMTQSP SSLSASVG DRVTITCK ASQ**DVSIG** VAWYQQK PGKAPKLLI Y**SAS**YRYT GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ YYIYPYT**F GQGTKVEI K	8225	9325, 9320, 9326
MM-121/ SAR256 212		HER3	EVQLLESG GGLVQPGG SLRLSCAA SGFTFS**HY VMA**WVRQ APGKGLEW VS**SISSSGG WTLYADS VKG**RFTIS RDNSKNTL YLQMNSLR AEDTAVYY CTR**GLKM ATIFDY**WG QGTLVTVS S	8129	9091, 9092, 9093	QSALTQPA SVSGSPGQ SITISC**TGT SSDVGSYN VVS**WYQQ HPGKAPKL IIY**EVSQRP S**GVSNRFS GSKSGNTA SLTISGLQT EDEADYYC **CSYAGSSIF VI**FGGGTK VTVL	8226	9327, 9328, 9329
MEHD7 945A	Duligotumab	HER1(EGFR) /HER3	EVQLVESG GGLVQPGG SLRLSCAA SGFTLS**GD WIH**WVRQ APGKGLEW VG**EISAAG GYTDYADS VKG**RFTIS ADTSKNTA YLQMNSLR AEDTAVYY CAR**ESRVS FEAAMDY** WGQGTLV TVSS	8130	9094, 9095, 9096	DIQMTQSP SSLSASVG DRVTITC**R ASQNIATD VA**WYQQK PGKAPKLLI Y**SASFLYS** GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ SEPEPYT**F GQGTKVEI K	8227	9330, 9331, 9332

MM-111		HER2/3	QVQLQESGGGLVKPGGSLRLSCAASGFTFS**SYWMS**WVRQAPGKGLEWVA**NINRDGSASYYVDS****VKG**RFTISRDDAKNSLYLQMNSLRAEDTAVYYCAR**DRGVGYFDL**WGRGTLVTVSS	8131	9097, 9098, 9099	QSALTQPASVSGSPGQSITISC**TGTSSDVGGYNFVS**WYQQHPGKAPKLMIY**DVSDRPS**GVSDRFSGSKSGNTASLIISGLQADDEADYYC**SSYGSSSTHVI**FGGGTKVTVL	8228	9333, 9334, 9335
MM-111		HER2/3	QVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIA**WVRQMPGKGLEYMG**LIYPGDSDTKYSPSF****QG**QVTISVDKSVSTAYLQWSSLKPSDSAVYFCAR**HDVGYCTDRTCAKWPEWLGV**WGQGTLVTVSS	8132	9035, 9100, 9101	QSVLTQPPSVSAAPGQKVTISC**SGSSSNIGNNYVS**WYQQLPGTAPKLLIY**DHTNRPA**GVPDRFSGSKSGTSASLAISGFRSEDEADYYC**ASWDYTLSGWV**FGGGTKLTVL	8229	9336, 9337, 9338
	Hu3S193	Lewis-Y	EVQLVESGGGVVQPGRSLRLSCSTSGFTFS**DYYMY**WVRQAPGKGLEWVA**YMSNVGAITDYPDT****VKG**RFTISRDNSKNTLFLQMDSLRPEDTGVYFCAR**GTRD**	8133	9102, 9103, 9104	DIQMTQSPSSLSASVGDRVTITC**RSSQRIVHSNGNTYLE**WYQQTPGKAPKLLIY**KVSNRFS**GVPSRFSGSGSGTDFTFTISSLQPEDIATYYC**FQGSHVPFT**FGQGT	8230	9339, 9281, 9340

			GSWFAYW GQGTPVTV SS			KLQIT		
BAY 94-9343	anetumab ravtansine	Mesothelin	QVELVQSG AEVKKPGE SLKISCKGS **GYSFTSYW** IGWVRQAP GKGLEWM GI**IDPGDSR T**RYSPSFQ GQVTISAD KSISTAYLQ WSSLKASD TAMYYC**A RGQLYGG TYMDG**WG QGTLVTVS S	8134	9105, 9106, 9107	DIALTQPAS VSGSPGQSI TISCTGT**SS DIGGYNS**V SWYQQHP GKAPKLMI Y**GVN**NRPS GVSNRFSG SKSGNTAS LTISGLQAE DEADYYC**S SYDIESATP V**FGGGTKL TVL	8231	9341, 9342, 9343
	SS1	Mesothelin	QVQLQQSG PELEKPGA SVKISCKAS GYSFTGYT MNWVKQS HGKSLEWI GLITPYNG ASSYNQKF RGKATLTV DKSSSTAY MDLLSLTS EDSAVYFC ARGGYDG RGFDYWG QGTTVTVS S	8135		DIELTQSPA IMSASPGE KVTMTCSA SSSVSYMH WYQQKSG TSPKRWIY DTSKLASG VPGRFSGS GSGNSYSL TISSVEAED DATYYCQQ WSGYPLTF GAGTKLEI K	8232	
		Mesothelin	QVYLVESG GGVVQPGR SLRLSCAA SGITFS**IYG MH**WVRQA PGKGLEWV A**VIWYDGS HEYYADSV KG**RFTISR DNSKNTLY LLMNSLRA ED TAVYYCAR **DGDYYDS GSPLDY**W GQGTLVTV	8136	9108, 9109, 9110	EIVLTQSPA TLSLSPGER ATLSC**RAS QSVSSYLA** WYQQKPG QAPRLLIY **DASNRAT**G IPARFSGSG SGTDFTLTI SSLEPEDFA VYYC**QQ RSNWPLT**F GGGTKVEI K	8233	9202, 9203, 9233

			SS					
		Mesothelin	QVHLVESGGGVVQPGRSLRLSCVASGITF**RIYGMH**WVRQAPGKGLEWVA**VLWYDGSHEYYADSVKG**RFTISRDNSKNTLYLQMNSLRAEDTAIYYCAR**DGDYYDSGSPLDY**WGQGTLVTVSS	8137	9111, 9112, 9110	EIVLTQSPATLSLSPGERATLSC**RASQSVSSYLA**WYQQKPGQAPRLLIY**DASNRAT**GIPARFSGSGSGTDFTLTISSLEPEDFAVYYC**QQRSNWPLT**FGGGTKVEIK	8233	9202, 9203, 9233
		Mesothelin	EVHLVESGGGLVQPGGSLRLSCAASGFTFS**RYWMS**WVRQAQGKGLEWVA**SIKQAGSEKTYVDSVKG**RFTISRDNAKNSLSLQMNSLRAEDTAVYYCAR**EGAYYYDSASYYPYYYYYSMDV**WGQGTTVTVSS	8138	9113, 9114, 9115	EIVLTQSPGTLSLSPGERATLSC**RASQSVSSSYLA**WYQQKPGQAPRLLIY**GASSRAT**GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYC**QQYGSSQYT**FGQGTKLEIK	8234	9285, 9182, 9344
MORAb-009	amatuximab	Mesothelin	QVQLQQSGPELEKPGASVKISCKASGYSFT**GYTMN**WVKQSHGKSLEWIG**LITPYNGASSYNQKFRG**KATLTVDKSSSTAYMDLLSLTSEDSAVYFCAR**GGYDGRGFDY**WGSGTPVTVSS	8139	9116, 9117, 9118	DIELTQSPAIMSASPGEKVTMTC**SASSSVSYMH**WYQQKSGTSPKRWIY**DTSKLAS**GVPGRFSGSGSGNSYSLTISSVEAEDDATYYC**QQWSKHPLT**FGSGTKVEIK	8235	9345, 8909, 9346

hPAM4		MUC-1	EVQLQESGPELVKPGASVKMSCKASGYTFP**SYVLH**WVKQKPGQGLEWIG**YINPYNDGTQYNEKFKG**KATLTSDKSSSTAYMELSRLTSEDSAVYYCAR**GFGGSYGFAY**WGQGTLITVSA	8140	9119, 9120, 9121	DIVMTQSPAIMSASPGEKVTMTC**SASSSVSSSYLY**WYQQKPGSSPKLWIY**STSNLAS**GVPARFSGSGSGTSYSLTISSMEAEDAASYFC**HQWNRYPYT**FGGGTKLEIK	8236	9347, 9348, 9349
hPAM4-Cide	clivatuzumab	MUC1	QVQLQQSGAEVKKFGASVKVSCEASGYTFP**SYVLH**WVKQAPGQGLEWIG**YINPYNDGTQTNKKFKG**KATLTRDTSINTAYMELSRLRSDDTAVYYCAR**GFGGSYGFAY**NGQGTLVTVSS	8141	9119, 9122, 9121	DIQLTQSPSSLSASVGDRVTMTC**SASSSVSSSYLY**WYQQKPGKAPKLWIY**STSNLAS**GVPARFSGSGSGTDFTLTISSLQPEDSASYFC**HQWNRYPYT**FGGGTRLEIK	8237	9347, 9348, 9349
SAR566658	huDS6v1.01	MUC1	QAQLQVSGAEVVKPGASVKMSCKASGYTFT**SYNMH**WVKQTPGQGLEWIG**YIYPGNGATNYNQKFQG**KATLTADTSSSTAYMQISSLTSEDSAVYFCAR**GDSVPFAY**WGQGTLVTVSA	8142	8947, 9123, 9124	EIVLTQSPATMSASPGERVTITC**SAHSSVSFMH**WFQQKPGTSPKLWIY**STSSLAS**GVPARFGGSGSGTSYSLTISSMEAEDAATYYC**QQRSSFPLT**FGAGTKLELK	8238	9350, 9351, 9352
Theragyn	Pemtumomab muHMFG1	MUC1	QVQLQQSGAELMKPGASVKISCKATGYTFS**AYWIE**WVKQRP	8143	9125, 9126, 9127	DIVMSQSPSSLAVSVGEKVTMSC**KSSQSLLYSSNQKIYLA**	8239	9353, 9354, 9355

			GHGLEWIG **EILPGSNN SRYNEKFK G**KATFTAD TSSNTAYM QLSSLTSED SAVYYCSR **SYDFAWF AY**WGQGT PVTVSA			WYQQKPG QSPKLLIY W**ASTRES**G VPDRFTGG GSGTDFTL TISSVKAED LAVYYC**Q QYYRYPR T**FGGGTKL EIK		
Therex	Sontuzumab huHMFG1 AS1402 R1150	MUC1	QVQLVQSG AEVKKPGA SVKVSCKA SGYTFS**AY WIE**WVRQ APGKGLEW VG**EILPGS NNSRYNEK FKG**RVTVT RDTSTNTA YMELSSLR SEDTAVYY CAR**SYDFA WFAY**WGQ GTLVTVSS	8144	9125, 9126, 9127	DIQMTQSP SSLSASVG DRVTITC**K SSQSLLYSS NQKIYLA** WYQQKPG KAPKLLIY W**ASTRES**G VPSRFSGSG SGTDFTFTI SSLQPEDIA TYYC**QQY YRYPRT**FG QGTKVEIK	8240	9353, 9354, 9355
MDX-1105 or BMS-936559		PD-L1	QVQLVQSG AEVKKPGS SVKVSCKT SGDTFS**TY AIS**WVRQA PGQGLEW MG**GIIPIFG KAHYAQK FQG**RVTIT ADESTSTA YMELSSLR SEDTAVYF CAR**KFHFV SGSPFGMD V**WGQGTT VTVSS	8145	9128, 9129, 9130	EIVLTQSPA TLSLSPGER ATLSC**RAS QSVSSYLA** WYQQKPG QAPRLLIY **DASNRAT**G IPARFSGSG SGTDFTLTI SSLEPEDFA VYYC**QQR SNWPT**FGQ GTKVEIK	8241	9202, 9203, 9356
MEDI-4736	durvalumab	PD-L1	EVQLVESG GGLVQPGG SLRLSCAA S**GFTFSRY W**MSWVRQ APGKGLEW VAN**IKQDG SEK**YYVDS VKGRFTISR DNAKNSLY LQMNSLRA	8146	9131, 9132, 9133	EIVLTQSPG TLSLSPGER ATLSCRAS **QRVSSSY**L AWYQQKP GQAPRLLI Y**DAS**SRAT GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q**	8242	9357, 9218, 9358

			EDTAVYYC **AREGGWF GELAFDY** WGQGTLV TVSS			**QYGSLPW T**FGQGTKV EIK		
MPDL3 280A	atezoliz umab	PD-L1	EVQLVESG GGLVQPGG SLRLSCAA S**GFTFSDS W**IHWVRQ APGKGLEW VAW**ISPYG GST**YYADS VKGRFTISA DTSKNTAY LQMNSLRA EDTAVYYC **ARRHWPG GFDY**WGQ GTLVTVSS	8147	9134, 9135, 9136	DIQMTQSP SSLSASVG DRVTITCR AS**QDVSTA** VAWYQQK PGKAPKLLI Y**SAS**FLYS GVPSRFSGS GSGTDFTL TISSLQPED FATYYC**QQ YLYHPAT**F GQGTKVEI K	8243	9359, 9320, 9360
MSB00 10718C	aveluma b	PD-L1	EVQLLESG GGLVQPGG SLRLSCAA S**GFTFSSYI** MMWVRQA PGKGLEWV SS**IYPSGGI T**FYADTVK GRFTISRDN SKNTLYLQ MNSLRAED TAVYYC**AR IKLGTVTT VDY**WGQG TLVTVSS	8148	9137, 9138, 9139	QSALTQPA SVSGSPGQ SITISCTGT**S SDVGGYN Y**VSWYQQ HPGKAPKL MIY**DVS**NR PSGVSNRFS GSKSGNTA SLTISGLQA EDEADYYC **SSYTSSSTR V**FGTGTKV TVL	8244	9361, 9362, 9363
MLN59 1		PSMA	EVQLVQSG PEVKKPGA TVKISCKTS **GYTFTEYT IH**WVKQAP GKGLEWIG **NINPNNGG TTYNQKFE D**KATLTVD KSTDTAYM ELSSLRSED TAVYYCAA **GWNFDY**W GQGTLLTV SS	8149	9140, 9141, 9142	DIQMTQSP SSLSTSVGD RVTLTC**KA SQDVGTA VD**WYQQK PGPSPKLLI Y**WASTRH T**GIPSRFSG SGSGTDFT LTISSLQPE DFADYYC**Q QYNSYPLT** FGPGTKVD IK	8245	9364, 9365, 9232
MT112	pasotuxi zumab	PSMA	QVQLVESG GGLVKPGE	8150	9143, 9144,	DIQMTQSP SSLSASVG	8246	9366, 9320,

			SLRLSCAA SGFTFSDY YMYWVRQ APGKGLEW VAIISDGG YYTYYSDII KGRFTISRD NAKNSLYL QMNSLKAE DTAVYYCA RGFPLLRH GAMDYWG QGTLVTVS S		9145	DRVTITCK ASQNVDTN VAWYQQK PGQAPKSLI YSASYRYS DVPSRFSGS ASGTDFTL TISSVQSED FATYYCQQ YDSYPYTF GGGTKLEI K		9367
		ROR1	QEQLVESG GRLVTPGG SLTLSCKAS GFDFSAYY MSWVRQA PGKGLEWI ATIYPSSG KTYYATW VNGRFTISS DNAQNTV DLQMNSLT AAD RATYFCAR DSYADDG ALFNIWGP GTLVTISS	8151	9146, 9147, 9148	ELVLTQSPS VSAALGSP AKITCTLSS AHKTDTID WYQQLQG EAPRYLMQ VQSDGSYT KRPGVPDR FSGSSSGA DRYLIIPSV QADDEAD Y YCGADYIG GYVFGGGT QLTVTG	8247	9368, 9369, 9370
		ROR1	EVKLVESG GGLVKPGG SLKLSCAA SGFTFSSYA MSWVRQIP EKRLEWVA SISRGGTT YYPDSVKG RFTISRDNV RNILYLQM SSLRSEDT AMYYCGR YDYDGYY AMDYWGQ GTSVTVSS	8152	9085, 9149, 9150	DIKMTQSP SSMYASLG ERVTITCK ASPDINSY LSWFQQKP GKSPKTLIY RANRLVD GVPSRFSG GGSGQDYS LTINSLEYE DMGIYYCL Q YDEFPYTF GGGTKLEM K	8248	9371, 9372, 9373
		ROR1	QSLEESGG RLVTPGTP LTLTCTVS GIDLNSHW MSWVRQA PGKGLEWI GIIAASGST	8153	9151, 9152, 9153	ELVMTQTP SSVSAAVG GTVTINCQ ASQSIGSY LAWYQQK PGQPPKLLI YYASNLAS	8249	9374, 9375, 9376

			YYANWAK **G**RFTISKTS TTVDLRIAS PTTEDTAT Y FCAR**DYGD** **YRLVTFNI** WGPGTLVT VSS			GVPSRFSGS GSGTEYTL TISGVQRE DAATYYC**L** **G** **SLSNSDNV** FGGGTELEI L		
		ROR1	QSVKESEG DLVTPAGN LTLTCTAS GSDIN**DYPI** **S**WVRQAPG KGLEWIG**F** **INSGGSTW** **YASWVKG** RFTISRTST TVDLKMTS LTTDDTAT Y FCAR**GYST** **YYCDFNI**W GPGTLVTIS S	8154	9154, 9155, 9156	ELVMTQTP SSTSGAVG GTVTINC**Q** **ASQSIDSN** **LA**WFQQKP GQPPTLLIY **RASNLAS**G VPSRFSGSR SGTEYTLTI SGVQREDA ATYYC**LG** **GVGNVSY** **RTS**FGGGT EVVVK	8250	9377, 9378, 9379
CC49 (Human ized)		TAG- 72	QVQLVQSG AEVVKPGA SVKISCKAS GYTFT**DHA** **IH**WVKQNP GQRLEWIG **YFSPGNDD** **FKYNERFK** **G**KATLTAD TSASTAYV ELSSLRSED TAVYFCTR **SLNMAY**W GQGTLVTV SS	8155	9157, 9158, 9013	DIVMSQSP DSLAVSLG ERVTLNCK SS**QSLLYS** **GNQKNYL** **A**WYQQKP GQSPKLLIY **WASARES** GVPDRFSG SGSGTDFT LTISSVQAE DVAVYYC **QQYYSYPL** **T**FGAGTKL ELK	8251	9380, 9381, 9262
	Murine A1	TPBG/ 5T4	QIQLVQSG PELKKPGE TVKISCKAS **GYTFTNFG** **MN**WVKQG PGEGLKW MG**WINTN** **TGEPRYAE** **EFKG**RXAF SLETTASTA YLQINNLK NEDTATYF	8156	9159, 9160, 9161	SIVMTQTP KFLLVSAG DRVTITC**K** **ASQSVSND** **VA**WYQQK PGQSPKLLI N**FATNRYT** GVPNRFTG SGYGTDFT FTISTVQAE DLALYFC**Q** **Q**	8252	9382, 9383, 9384

			CAR**DWDG AYFFDY**W GQGTTLTV SS			**DYSSPWT**F GGGTKLEI K		
	Murine A2	TPBG/ 5T4	QVQLQQSR PELVKPGA SVKMSCKA S**GYTFTDY VIS**WVKQR TGQGLEWI G**EIYPGSN SIYYNEKF KG**RATLTA DKSSSTAY MQLSSLTS EDSAVYFC AM**GGNYG FDY**WGQG TTLTVSS	8157	9162, 9163, 9164	SVIMSRGQI VLTQSPAI MSASLGER VTLTC**TAS SSVNSNYL H**WYQQKP GSSPKLWI Y**STSNLAS** GVPARFSG SGSGTSYSL TISSMEAED AATYYC**H QYHRSPLT** FGAGTKLE LK	8253	9385, 9348, 9386
	Murine A3	TPBG/ 5T4	EVQLVESG GGLVQPKG SLKLSCAA S**GFTFNTY AMN**WVRQ APGKGLEW VA**RIRSKS NNYATYY ADSVKD**RF TISRDDSQS MLYLQMN NLKTEDTA MYXCVR**Q WDYDVRA MNY**WGQG TSVTVSS	8158	742, 9165, 9166	DIVMTQSHI FMSTSVGD RVSITC**KAS QDVDTAV A**WYQQKP GQSPKLLIY **WASTRLT** GVPDRFTG SGSGTDFT LTISNVQSE DLADYFC**Q Q YSSYPYT**F GGGTKLEI K	8254	9387, 9388, 9389
IMMU-132	hRS-7	TROP-2	QVQLQQSG SELKKPGA SVKVSCKA SGYTFT**NY GMN**WVKQ APGQGLK WMG**WINT YTGEPTYT DDFKG**RFA FSLDTSVST AYLQISSLK ADDTAVYF CAR**GGFGS SYWYFDV** WGQGSLVT VSS	8159	9016, 9167, 9168	DIQLTQSPS SLSASVGD RVSITC**KAS QDVSIAVA** WYQQKPG KAPKLLIY**S ASYRYT**GV PDRFSGSGS GTDFTLTIS SLQPEDFA VYYC**QQH YITPLT**FG AGTKVEIK	8255	9390, 9391, 9392
IMC-	icrucum	VEGF	QAQVVESG	8160	9169,	EIVLTQSPG	8256	9285,

18F1	ab	R1	GGVVQSGR SLRLSCAA S**GFAFSSY GMHW**VRQ APGKGLEW VA**VIWYD GSNKYYA DSVRG**RFT ISRDNSENT LYLQMNSL RAEDTAVY YCAR**DHY GSGVHHY FYYGLDV** WGQGTTV TVSS		9170, 9171	TLSLSPGER ATLSC**RAS QSVSSSYL A**WYQQKP GQAPRLLI Y**GASSRAT** GIPDRFSGS GSGTDFTL TISRLEPED FAVYYC**Q QYGSSPLT** FGGGTKVE IK		9182, 9393
Cyramza	ramucirumab	VEGFR2	EVQLVQSG GGLVKPGG SLRLSCAA S**GFTFSSYS MN**WVRQA PGKGLEWV S**SISSSSSYI YYADSVK G**RFTISRDN AKNSLYLQ MNSLRAED TAVYYCAR **VTDAFDI**W GQGTMVT VSSA	8161	9172, 9173, 9174	DIQMTQSP SSVSASIGD RVTITC**RAS QGIDN**WL GWYQQKP GKAPKLLI Y**DASNLDT** GVPSRFSGS GSGTYFTL TISSLQAED FAVYFC**QQ AKAFPPT**F GGGTKVDI K	8257	9394, 9395, 9396
g165DFM-PEG	alacizumabpegol	VEGFR2	EVQLVESG GGLVQPGG SLRLSCAA S**GFTFSSY G**MSWVRQ APGKGLEW VAT**ITSGG SYT**YYVDS VKGRFTISR DNAKNTLY LQMNSLRA EDTAVYYC **VRIGEDAL DY**WGQGT LVTVSS	8162	9175, 9176, 9177	DIQMTQSP SSLSASVG DRVTITCR AS**QDIAGS** LNWLQQKP GKAIKRLIY **ATS**SLDSG VPKRFSGS RSGSDYTL TISSLQPED FATYYC**LQ YGSFPPT**F GQGTKVEI K	8258	9397, 9212, 9398
Imclone 6.64		VEGFR2	KVQLQQSG TELVKPGA SVKVSCKA SGYIFTEYII HWVKQRS GQGLEWIG	8163		DIVLTQSPA SLAVSLGQ RATISCRAS ESVDSYGN SFMHWYQ QKPGQPPK	8259	

			WLYPESNII KYNEKFKD KATLTADK SSSTVYME LSRLTSEDS AVYFCTRH DGTNFDY WGQGTTLT VSSA			LLIYRASNL ESGIPARFS GSGSRTDF TLTINPVEA DDVATYYC QQSNEDPL TFGAGTKL ELK		

* подчеркнутые и выделенные жирным шрифтом последовательности, если они есть, представляют собой CDR внутри VL и VH

[215] Адгезивная молекула эпителиальных клеток (EpCAM, также известная как антиген 17-1A) представляет собой интегрированный в мембрану гликопротеин с молекулярной массой 40 кДа, состоящий из 314 аминокислот, экспрессируемый в определенных типах эпителия и во многих видах карциномы человека (*см.,* Balzar, The biology of the 17-1A antigen (Ep-CAM), J. Mol. Med. 1999, 77:699-712). Из-за своего эпителиального происхождения опухолевые клетки большинства карцином экспрессируют EpCAM на своей поверхности (больше, чем нормальные, здоровые клетки), включая большинство клеток первичной, метастатической и диссеминированной немелкоклеточной карциномы легкого (Passlick, B., et al. The 17-1A antigen is expressed on primary, metastatic and disseminated non-small cell lung carcinoma cells. Int. J. Cancer 87(4):548-552, 2000), аденокарциномы желудка и гастроэзофагеального перехода (Martin, I.G., Expression of the 17-1A antigen in gastric and gastro-oesophageal junction adenocarcinomas: a potential immunotherapeutic target? J Clin Pathol 1999;52:701-704) и рак молочной железы и колоректальный рак (Packeisen J, et al. Detection of surface antigen 17-1A in breast and colorectal cancer. Hybridoma. 1999 18(1):37-40), при раке молочной железы избыточная экспрессия EpCAM на опухолевых клетках является предиктором выживания (Gastl, Lancet. 2000, 356, 1981-1982). Из-за своего эпителиального происхождения опухолевые клетки большинства карцином экспрессируют EpCAM на своей поверхности.

[216] В одном из вариантов осуществления в данном документе предусмотрены композиции на основе сборки биспецифического химерного полипептида, первая часть которых имеет связывающий домен, специфичный для EpCAM, и связывающий домен, специфичный для CD3. Техническая проблема, которую необходимо было решить, заключалась в обеспечении средств и способов получения улучшенных композиций, проявляющих свойства хорошо переносимых и более удобных лекарственных препаратов (менее частое введение дозы) для эффективного лечения и/или облегчения опухолевых заболеваний. Решение указанной технической проблемы достигается с помощью вариантов осуществления, раскрытых в данном документе и охарактеризованных в формуле изобретения.

[217] Соответственно, в некоторых вариантах осуществления настоящее изобретение относится к композициям на основе сборки химерных полипептидов, где указанная композиция предусматривает первую часть, содержащую композицию на

основе биспецифического одноцепочечного антитела, содержащую по меньшей мере два связывающих домена, при этом один из указанных доменов связывается с антигеном эффекторной клетки, таким как антиген CD3, а второй домен связывается с антигеном EpCAM, где указанные связывающие домены содержат VL и VH, специфичные для EpCAM, и VL и VH, специфичные для антигена CD3 человека. Предпочтительно в варианте осуществления указанный связывающий домен, специфичный для EpCAM, характеризуется значением K_d от более чем 10^{-7} до 10^{-10} M, как определено в анализе связывания in vitro. В одном варианте осуществления вышеизложенного связывающие домены имеют формат scFv. В другом варианте осуществления вышеизложенного связывающие домены имеют формат одноцепочечного диатела.

[218] В некоторых вариантах осуществления настоящее изобретение предусматривает композиции на основе сборки химерного полипептида, содержащие связывающий домен первой части, обладающий аффинностью связывания с опухолеспецифическим маркером, и второй связывающий домен, связывающийся с антигеном эффекторной клетки, таким как антиген CD3. Опухолеспецифические маркеры, предусмотренные этими вариантами осуществления настоящего изобретения, включают без ограничения CCR5, CD19, HER-2, HER-3, HER-4, EGFR, PSMA, CEA, MUC1, MUC2, MUC3, MUC4, MUC5AC, MUC5B, MUC7, βhCG, Льюис-Y, CD-20, CD33, CD30, ганглиозид GD3, 9-O-ацетил-GD3, Globo H, фукозил GM1, GD-2, карбоангидразу IX, CD44v6, «Sonic Hedgehog», Wue-1, антиген плазматической клетки 1, хондроитинсульфат протеогликан меланомы, CCR8, 6-трансмембранный эпителиальтный антиген предстательной железы (STEAP), мезотелин, антиген A33, антиген стволовых клеток предстательной железы (PSCA), LY-6, SAS, десмоглеин 4, фетальный ацетилхолиновый рецептор, CD-25, раковый антиген 19-9 (CA 19-9), раковый антиген 125 (CA-125), рецептор мюллерова ингибирующего фактора типа II (MISIIR), сиалированный антиген Tn, антиген активации фибробластов (FAP), эндосиалин (CD248), вариант III рецептора эпидермального фактора роста (EGFRvIII), опухолеассоциированный антиген L6 (TAL6), CD-63, TAG-72, антиген Томсена-Фриденрайха (TF-антиген), рецептор инсулиноподобного фактора роста I (IGF-IR), антиген Cora, CD7, CD22, CD79a, CD79b, G250, F19, EphA2 и MT-MM. В определенных вариантах осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, содержащую связывающий домен первой части с аффинностью связывания с опухолеспецифическим маркером, содержащий последовательности VL и VH к маркеру. Иллюстративные неограничивающие примеры последовательностей VL и VH, специфичных для некоторых из этих опухолевых маркеров, представлены в **таблице 6f**. В других вариантах осуществления настоящее изобретение предусматривает композицию на основе сборки химерного полипептида, содержащую связывающий домен первой части с аффинностью связывания с опухолеспецифическим маркером, включающий область CDR-L1, область CDR-L2, область CDR-L3, область CDR-H1, область CDR-H2 и область CDR-H3, каждая из которых получена из соответствующих последовательностей VL и

VH. Предпочтительно в вариантах осуществления указанное связывание характеризуется значением K_d от более чем 10^{-7} до 10^{-10} M, как определено в анализе связывания in vitro.

[219] В частности, предполагается, что композиция на основе сборки химерного полипептида может содержать любой из вышеуказанных связывающих доменов или их варианты последовательности, при условии, что варианты проявляют специфичность связывания в отношении описанного антигена. В одном варианте осуществления вариант последовательности может быть создан путем замены аминокислоты в последовательности VL или VH другой аминокислотой. В вариантах на основе делеции удаляют один или более аминокислотных остатков в последовательности VL или VH, как описано в данном документе. Таким образом, варианты на основе делеции включают все фрагменты полипептидной последовательности связывающего домена. В вариантах на основе замены один или более аминокислотных остатков полипептида VL или VH (или CDR) удаляют и заменяют альтернативными остатками. В одном аспекте замены являются консервативными по своей природе, и консервативные замены этого типа хорошо известны в уровне техники. Кроме того, конкретно предполагается, что композиции, содержащие первый и второй связывающие домены, раскрытые в данном документе, можно использовать в любом из способов, раскрытых в данном документе.

Неструктурированная конформация

[220] Как правило, полипептидный компонент XTEN слитых белков, раскрытых в данном документе, сконструирован так, чтобы вести себя как денатурированные пептидные последовательности в физиологических условиях, несмотря на увеличенную длину полимера. «Денатурированный» описывает состояние пептида в растворе, характеризующееся большой конформационной свободой пептидной основной цепи. Большинство пептидов и белков принимают денатурированную конформацию в присутствии высоких концентраций денатурантов или при повышенной температуре. Пептиды в денатурированной конформации имеют, например, характерные спектры кругового дихроизма (CD) и характеризуются отсутствием дальнодействующих взаимодействий, как определено с помощью NMR. «Денатурированную конформацию» и «неструктурированную конформацию» используют в данном документе как синонимы. В некоторых случаях настоящее изобретение предусматривает полипептиды XTEN, которые в физиологических условиях могут напоминать денатурированные последовательности, в значительной степени лишенные вторичной структуры. В других случаях полипептиды XTEN могут практически не иметь вторичной структуры в физиологических условиях. Используемое в данном контексте выражение «в значительной степени лишенный» означает, что менее 50% аминокислотных остатков XTEN каждого полипептида XTEN вносят вклад во вторичную структуру, как измерено или определено способами, описанными в данном документе. Используемое в данном контексте выражение «практически лишенный» означает, что по меньшей мере приблизительно 60%, или приблизительно 70%, или приблизительно 80%, или приблизительно 90%, или приблизительно 95% или по меньшей мере приблизительно 99% аминокислотных

остатков XTEN из последовательности XTEN не вносят вклад во вторичную структуру, как измерено или определено способами, описанными в данном документе.

[221] В уровне техники было разработано множество способов для определения наличия или отсутствия вторичных и третичных структур в данном полипептиде. В частности, вторичная структура XTEN может быть измерена спектрофотометрически, например, с помощью спектроскопии кругового дихроизма в «дальней УФ» области спектра (190-250 нм). Элементы вторичной структуры, такие как альфа-спираль и бета-лист, обуславливают характерную форму и величину спектров CD. Вторичная структура полипептидной последовательности также может быть предсказана с помощью определенных компьютерных программ или алгоритмов, таких как хорошо известный алгоритм Чоу-Фасмана (Chou, P. Y., et al. (1974) Biochemistry, 13: 222-45) и алгоритм Гарнье-Осгуторпа-Робсона («GOR») (Garnier J, Gibrat JF, Robson B. (1996), GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol 266:540-553), как описано в публикации заявки на патент США № 20030228309A1. Для данной последовательности алгоритмы могут предсказать, существует ли какая-либо вторичная структура или ее вообще нет, что выражается в виде общего количества и/или процентного содержания остатков последовательности, образующих, например, альфа-спирали или бета-листы, или процентного содержания остатков последовательности, которые по прогнозам приведут к образованию статистического клубка (в котором отсутствует вторичная структура).

[222] В некоторых случаях полипептиды XTEN, используемые в композициях на основе слитых белков по настоящему изобретению, могут иметь процентное содержание альфа-спирали в диапазоне от 0% до менее чем приблизительно 5%, как определено при помощи алгоритма Чоу-Фасмана. В других случаях полипептиды XTEN, содержащие композиции на основе слитых белков, могут иметь процентное содержание бета-листов в диапазоне от 0% до менее чем приблизительно 5%, как определено при помощи алгоритма Чоу-Фасмана. В некоторых случаях последовательности XTEN композиций на основе слитых белков могут иметь процентное содержание альфа-спиралей в диапазоне от 0% до менее чем приблизительно 5% и процентное содержание бета-листов в диапазоне от 0% до менее чем приблизительно 5%, как определено при помощи алгоритма Чоу-Фасмана. В предпочтительных вариантах осуществления полипептиды XTEN, содержащие композиции на основе слитых белков, могут иметь процентное содержание альфа-спиралей менее чем приблизительно 2% и процентное содержание бета-листьев менее чем приблизительно 2%. В других случаях последовательности XTEN композиций на основе слитых белков могут иметь высокую степень процентного содержания статистического клубка, как определено при помощи алгоритма GOR. В некоторых вариантах осуществления полипептид XTEN может содержать по меньшей мере приблизительно 80%, более предпочтительно по меньшей мере приблизительно 90%, более предпочтительно по меньшей мере приблизительно 91%, более предпочтительно по меньшей мере приблизительно 92%, более предпочтительно по меньшей мере

приблизительно 93%, более предпочтительно по меньшей мере приблизительно 94%, более предпочтительно по меньшей мере приблизительно 95%, более предпочтительно по меньшей мере приблизительно 96%, более предпочтительно по меньшей мере приблизительно 97%, более предпочтительно по меньшей мере приблизительно 98% и, наиболее предпочтительно, по меньшей мере приблизительно 99% статистического клубка, как определено при помощи алгоритма GOR.

Суммарный заряд

[223] В других случаях полипептиды XTEN могут иметь неструктурированные характеристики, придаваемые путем включения аминокислотных остатков с суммарным зарядом и/или уменьшения доли гидрофобных аминокислот в полипептиде XTEN. Общий суммарный заряд и плотность суммарного заряда можно контролировать путем изменения содержания заряженных аминокислот в полипептидах XTEN. В некоторых случаях плотность суммарного заряда XTEN композиций может быть выше +0,1 или ниже -0,1 зарядов/остаток. В других случаях суммарный заряд полипептида XTEN может составлять приблизительно 0%, приблизительно 1%, приблизительно 2%, приблизительно 3%, приблизительно 4%, приблизительно 5%, приблизительно 6%, приблизительно 7%, приблизительно 8%, приблизительно 9%, приблизительно 10% приблизительно 11%, приблизительно 12%, приблизительно 13%, приблизительно 14%, приблизительно 15%, приблизительно 16%, приблизительно 17%, приблизительно 18%, приблизительно 19%, или приблизительно 20% или более.

[224] Поскольку большинство тканей и поверхностей человека или животного имеют суммарный отрицательный заряд, полипептиды XTEN могут быть сконструированы таким образом, чтобы иметь суммарный отрицательный заряд для сведения к минимуму неспецифических взаимодействий между композициями, содержащими полипептид XTEN, и различными поверхностями, такими как кровеносные сосуды, здоровые ткани или различные рецепторы. Чтобы не быть связанными какой-либо конкретной теорией, полипептид XTEN может принимать открытые конформации из-за электростатического отталкивания между отдельными аминокислотами полипептида XTEN, которые по отдельности несут высокий суммарный отрицательный заряд и которые распределены по последовательности полипептида XTEN. Такое распределение суммарного отрицательного заряда в длинах удлиненных последовательностей полипептида XTEN может привести к неструктурированной конформации, что, в свою очередь, может привести к эффективному увеличению гидродинамического радиуса. Соответственно, в одном варианте осуществления настоящее изобретение предусматривает полипептиды XTEN, в которых полипептиды XTEN содержат приблизительно 8, 10, 15, 20, 25 или даже приблизительно 30% глутаминовой кислоты. Полипептиды XTEN композиций по настоящему изобретению обычно не содержат положительно заряженных аминокислот или имеют их низкое содержание. В некоторых случаях полипептиды XTEN могут содержать менее чем приблизительно 10% аминокислотных остатков с положительным зарядом, или менее чем приблизительно 7%,

или менее чем приблизительно 5%, или менее чем приблизительно 2% аминокислотных остатков с положительным зарядом. Однако в настоящем изобретении рассматриваются конструкции, в которых ограниченное количество аминокислот с положительным зарядом, таких как лизин, может быть включено в полипептиды XTEN для обеспечения конъюгации между эпсилон-амином лизина и реакционноспособной группой пептида, линкерным мостиком, или реакционноспособной группой лекарственного средства или малой молекулы, которая должна быть конъюгирована с основной цепью полипептида XTEN. Согласно вышеизложенному можно сконструировать слитые белки, которые содержат один или более полипептидов XTEN, биологически активный белок плюс химиотерапевтическое средство, применимое для лечения метаболических заболеваний или нарушений, где максимальное количество молекул средства, включенных в компонент полипептида XTEN, определяется количеством остатков лизина или других аминокислот с реакционноспособными боковыми цепями (*например*, цистеина), включенных в XTEN.

[225] В некоторых случаях полипептид XTEN может содержать заряженные остатки, разделенные другими остатками, такими как серин или глицин, что может привести к лучшему поведению при экспрессии или очистке. На основе суммарного заряда полипептиды XTEN композиций для человека или животного могут иметь изоэлектрическую точку (pI) 1,0, 1,5, 2,0, 2,5, 3,0, 3,5, 4,0, 4,5, 5,0, 5,5, 6,0 или даже 6,5. В предпочтительных вариантах осуществления полипептид XTEN имеет изоэлектрическую точку от 1,5 до 4,5. В этих вариантах осуществления XTEN, включенный в композиции на основе слитых белков BPXTEN по настоящему изобретению, будет нести суммарный отрицательный заряд в физиологических условиях, что может способствовать неструктурированной конформации и уменьшенному связыванию компонента полипептида XTEN с белками и тканями млекопитающих.

[226] Поскольку гидрофобные аминокислоты могут придавать структуру полипептиду, настоящее изобретение предусматривает, что содержание гидрофобных аминокислот в полипептидах XTEN обычно составляет менее 5%, или менее 2%, или менее 1% содержания гидрофобных аминокислот. В одном варианте осуществления содержание аминокислот метионина и триптофана в компоненте, представляющем собой XTEN, слитого белка BPXTEN обычно составляет менее 5% или менее 2% и наиболее предпочтительно менее 1%. В другом варианте осуществления полипептид XTEN имеет последовательность, которая содержит менее 10% аминокислотных остатков с положительным зарядом, или менее приблизительно 7%, или менее приблизительно 5%, или менее приблизительно 2% аминокислотных остатков с положительным зарядом, сумма остатков метионина и триптофана составляет менее 2%, а сумма остатков аспарагина и глутамина составляет менее 10% от всего полипептида XTEN.

УВЕЛИЧЕННЫЙ ГИДРОДИНАМИЧЕСКИЙ РАДИУС

[227] В некоторых вариантах осуществления полипептид XTEN может иметь большой гидродинамический радиус, что придает соответствующую увеличенную

кажущуюся молекулярную массу слитому белку BPXTEN, который включает в себя полипептид XTEN. Связывание полипептида XTEN с последовательностями BP может привести к получению композиций на основе BPXTEN, которые могут иметь увеличенные гидродинамические радиусы, увеличенную кажущуюся молекулярную массу и повышенный коэффициент кажущейся молекулярной массы по сравнению с BP, не связанным с полипептидом XTEN. Например, в терапевтических применениях, в которых желателен пролонгированный период полужизни, композиции, в которых полипептид XTEN с большим гидродинамическим радиусом включен в слитый белок, содержащий один или более BP, могут эффективно увеличивать гидродинамический радиус композиции свыше размера гломерулярной поры, составляющего приблизительно 3-5 нм (соответствует кажущейся молекулярной массе приблизительно 70 кДа) (Caliceti. 2003. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. Adv. Drug Deliv. Rev. 55:1261-1277), приводя к снижению почечного клиренса белков в кровотоке. Чтобы не быть связанными какой-либо конкретной теорией, полипептид XTEN может принимать открытые конформации из-за электростатического отталкивания между отдельными зарядами пептида или присущей ему гибкости, сообщаемой определенными аминокислотами в последовательности, которые не обладают потенциалом для придания вторичной структуры. Открытая, удлиненная и неструктурированная конформация полипептида XTEN может иметь больший пропорциональный гидродинамический радиус по сравнению с полипептидами с сопоставимой длиной последовательности и/или молекулярной массой, которые имеют вторичную и/или третичную структуру, такими как типичные глобулярные белки. Способы определения гидродинамического радиуса хорошо известны из уровня техники, например, с применением эксклюзионной хроматографии (SEC), как описано в патентах США №№ 6406632 и 7294513. Добавление увеличивающихся длин полипептидов XTEN приводит к пропорциональному увеличению параметров гидродинамического радиуса, кажущейся молекулярной массы и коэффициента кажущейся молекулярной массы, что позволяет адаптировать BPXTEN к необходимым характеристическим пороговым значениям кажущейся молекулярной массы или гидродинамического радиуса. Соответственно, в определенных вариантах осуществления слитый белок BPXTEN может быть сконфигурирован с полипептидом XTEN, где слитый белок может иметь гидродинамический радиус по меньшей мере приблизительно 5 нм, или по меньшей мере приблизительно 8 нм, или по меньшей мере приблизительно 10 нм, или 12 нм, или по меньшей мере приблизительно 15 нм. В приведенных выше вариантах осуществления большой гидродинамический радиус, обеспечиваемый полипептидами XTEN в слитом белке BPXTEN, может приводить к снижению почечного клиренса полученного слитого белка, что приводит к соответствующему увеличению конечного периода полувыведения, увеличению среднего времени пребывания и/или снижению скорости почечного клиренса.

[228] В другом варианте осуществления полипептид XTEN выбранной длины и последовательности может быть селективно включен в BPXTEN для создания слитого

белка, который в физиологических условиях имеет кажущуюся молекулярную массу по меньшей мере приблизительно 150 кДа, или по меньшей мере приблизительно 300 кДа, или по меньшей мере приблизительно 400 кДа, или по меньшей мере приблизительно 500 кДа, или по меньшей мере приблизительно 600 кДа, или по меньшей мере приблизительно 700 кДа, или по меньшей мере приблизительно 800 кДа, или по меньшей мере приблизительно 900 кДа, или по меньшей мере приблизительно 1000 кДа, или по меньшей мере приблизительно 1200 кДа, или по меньшей мере приблизительно 1500 кДа, или по меньшей мере приблизительно 1800 кДа, или по меньшей мере приблизительно 2000 кДа, или по меньшей мере приблизительно 2300 кДа или более. В другом варианте осуществления полипептид XTEN выбранной длины и последовательности может быть селективно связан с BP с получением слитого белка BPXTEN, который в физиологических условиях имеет коэффициент кажущейся молекулярной массы по меньшей мере три, в качестве альтернативы, по меньшей мере четыре, в качестве альтернативы, по меньшей мере пять, в качестве альтернативы, по меньшей мере шесть, в качестве альтернативы, в по меньшей мере восемь, в качестве альтернативы, по меньшей мере 10, в качестве альтернативы, по меньшей мере 15, или коэффициент кажущейся молекулярной массы по меньшей мере 20 или более. В другом варианте осуществления слитый белок BPXTEN в физиологических условиях имеет коэффициент кажущейся молекулярной массы, который составляет от приблизительно 4 до приблизительно 20, или составляет от приблизительно 6 до приблизительно 15, или составляет от приблизительно 8 до приблизительно 12, или составляет от приблизительно 9 до приблизительно 10 относительно фактической молекулярной массы слитого белка. В некоторых вариантах осуществления (слитый) полипептид демонстрирует коэффициент кажущейся молекулярной массы в физиологических условиях, превышающий приблизительно 6.

Увеличенный конечный период полувыведения

[229] В некоторых вариантах осуществления (слитый) полипептид имеет конечный период полувыведения, который в по меньшей мере два раза дольше, или в по меньшей мере три раза дольше, или в по меньшей мере четыре раза дольше, или в по меньшей мере пять раз дольше в сравнении с биологически активным полипептидом, не связанным с полипептидом XTEN. В некоторых вариантах осуществления (слитый) полипептид имеет конечный период полувыведения, который в по меньшей мере два раза дольше в сравнении с биологически активным полипептидом, не связанным с полипептидом XTEN.

[230] Введение терапевтически эффективной дозы любого из вариантов осуществления слитых белков BPXTEN, описанных в данном документе, человеку или животному, нуждающемуся в этом, может привести к увеличению времени в по меньшей мере два раза, или в по меньшей мере три раза, или в по меньшей мере четыре раза, или в по меньшей мере пять раз или более, проводимого в пределах терапевтического окна для слитого белка по сравнению с соответствующим BP, не связанным с полипептидом XTEN и вводимым в сравнимой дозе человеку или животному.

Низкая иммуногенность

[231] В другом аспекте настоящее изобретение предусматривает композиции, в которых полипептиды XTEN характеризуются низкой степенью иммуногенности или являются практически неиммуногенными. Несколько факторов могут способствовать низкой иммуногенности полипептидов XTEN, *например* практически неповторяющаяся последовательность, ее неструктурированная конформация, высокая степень растворимости, низкая степень самоагрегации или ее отсутствие, низкая степень наличия протеолитических сайтов внутри последовательности или их отсутствие, а также низкая степень наличия эпитопов в полипептидах XTEN или их отсутствие.

[232] Специалисту в данной области будет понятно, что, как правило, полипептид, содержащий высоко повторяющиеся короткие аминокислотные последовательности (например, где последовательность длиной 200 аминокислот содержит в среднем 20 или более повторов из ограниченного набора 3- или 4-меров) и/или имеющий смежные повторяющиеся аминокислотные остатки (например, где 5- или 6-мерные последовательности имеют идентичные аминокислотные остатки), имеет тенденцию к агрегации или образованию структур более высокого порядка или к образованию контактов, приводящих к формированию кристаллических или псевдокристаллических структур.

[233] В некоторых вариантах осуществления полипептиды XTEN являются практически неповторяющимися, где аминокислотная последовательность XTEN не содержит трех смежных аминокислот, которые относятся к идентичным типам аминокислот, за исключением случаев, когда аминокислота представляет собой серин, и в этом случае не более трех смежных аминокислот могут быть сериновыми остатками; и где аминокислотная последовательность XTEN не содержит 3-аминокислотных последовательностей (3-меров), которые встречаются более 16, более 14, более 12 или более 10 раз в последовательности полипептида XTEN длиной 200 аминокислот. Специалисту в данной области будет понятно, что такие фактически неповторяющиеся последовательности имеют меньшую склонность к агрегации и, таким образом, позволяют конструировать длинные последовательности XTEN с относительно низкой частотой заряженных аминокислот, которые бы, вероятно, агрегировались, если бы последовательности или аминокислотные остатки в остальном были более повторяющимися.

[234] Конформационные эпитопы образуются за счет областей поверхности белка, которые состоят из множества прерывистых аминокислотных последовательностей белкового антигена. Точная укладка белка приводит эти последовательности в четко определенные, стабильные пространственные конфигурации или эпитопы, которые могут быть распознаны как «чужие» гуморальной иммунной системой хозяина, что приводит к выработке антител к белку или запуску клеточно-опосредованного иммунного ответа. В последнем случае на иммунный ответ на белок у индивидуума сильно влияет распознавание Т-клеточного эпитопа, которое зависит от специфичности связывания пептидов аллотипа HLA-DR у этого индивидуума. Взаимодействие пептидного комплекса

MHC класса II с родственным Т-клеточным рецептором на поверхности Т-клетки вместе с перекрестным связыванием некоторых других корецепторов, таких как молекула CD4, может индуцировать активированное состояние Т-клетки. Активация приводит к высвобождению цитокинов, которые дополнительно активируют другие лимфоциты, такие как В-клетки, для выработки антител или активируют Т-киллерные клетки в качестве полноценного клеточного иммунного ответа.

[235] Способность пептида связывать данную молекулу MHC класса II для презентации на поверхности APC (антигенпрезентирующей клетки) зависит от ряда факторов; особенно от его первичной последовательности. В одном варианте осуществления более низкая степень иммуногенности может достигаться путем конструирования полипептидов XTEN, которые устойчивы к процессингу антигена в антигенпрезентирующих клетках, и/или выбора последовательностей, которые плохо связываются с рецепторами MHC. Настоящее изобретение предусматривает слитые белки BPXTEN с практически неповторяющимися полипептидами XTEN, предназначенными для уменьшения связывания с рецепторами MHC II, а также для предотвращения образования эпитопов для связывания Т-клеточного рецептора или антитела, что приводит к низкой степени иммуногенности. Избегание иммуногенности частично является прямым результатом конформационной гибкости полипептидов XTEN; *m. e.*, отсутствия вторичной структуры из-за выбора и порядка аминокислотных остатков. Например, особый интерес представляют последовательности, имеющие низкую тенденцию к принятию компактно свернутых конформаций в водном растворе или в физиологических условиях, которые могут приводить к конформационным эпитопам. Введение слитых белков, содержащих полипептид(-ы) XTEN, с использованием обычных терапевтических методик и доз, как правило, не приводит к образованию нейтрализующих антител к полипептиду XTEN, а также может снижать иммуногенность партнера BP по слиянию в композициях на основе BPXTEN.

[236] В одном варианте осуществления полипептиды XTEN, применяемые в слитых белках человека или животного, могут практически не содержать эпитопов, распознаваемых Т-клетками человека. Устранение таких эпитопов с целью создания менее иммуногенных белков было раскрыто ранее; см., например, WO 98/52976, WO 02/079232 и WO 00/3317, которые включены в данный документ посредством ссылки. Описаны анализы эпитопов Т-клеток человека (Stickler, M., et al. (2003) J Immunol Methods, 281: 95-108). Особый интерес представляют пептидные последовательности, которые могут быть олигомеризованы без образования Т-клеточных эпитопов или последовательностей, не являющихся человеческими. Это может быть достигнуто путем тестирования прямых повторов этих последовательностей в отношении присутствия Т-клеточных эпитопов и в отношении наличия 6-15-мерных и, в частности, 9-мерных последовательностей, которые не являются человеческими, а затем путем изменения конструкции полипептида XTEN для устранения или нарушения последовательности эпитопа. В некоторых случаях полипептиды XTEN являются практически неиммуногенными из-за ограничения числа

эпитопов полипептида XTEN, которые, как предполагается, связываются с рецепторами MHC. Со снижением количества эпитопов, способных связываться с рецепторами MHC, происходит сопутствующее снижение потенциала активации Т-клеток, а также функции Т-клеточных хелперов, снижение активации или повышенной регуляции В-клеток и снижение выработки антител. Низкую степень наличия предсказанных Т-клеточных эпитопов можно определить с помощью алгоритмов прогнозирования эпитопов, таких как, например, TEPITOPE (Sturniolo, T., et al. (1999) Nat Biotechnol, 17: 555-61), как показано в **примере 74** публикации международной патентной заявки № WO 2010/144502 A2, которая включена посредством ссылки во всей своей полноте. Оценка TEPITOPE для данного пептидного каркаса в белке представляет собой логарифм K_d (константа диссоциации, аффинность, скорость диссоциации) связывания этого пептидного каркаса с несколькими наиболее распространенными аллелями MHC человека, как описано в Sturniolo, T. et al. (1999) Nature Biotechnology 17:555). Оценка колеблется в диапазоне по меньшей мере 20 логарифмических единиц, от приблизительно 10 до приблизительно -10 (соответствует ограничениям связывания с K_d от $10e^{10}$ до K_d $10e^{-10}$), и может быть снижена за счет исключения гидрофобных аминокислот, которые могут служить якорными остатками во время размещения пептида на MHC, таких как M, I, L, V, F. В некоторых вариантах осуществления полипептид XTEN, включенный в BPXTEN, не имеет предсказанного Т-клеточного эпитопа при оценке TEPITOPE приблизительно -5 или выше, или -6 или выше, или -7 или выше, или -8 или выше, или при оценке TEPITOPE -9 или выше. Используемая в данном документе оценка «-9 или выше» будет включать в себя оценки TEPITOPE от 10 до -9 включительно, но не будет охватывать оценку -10, поскольку -10 меньше, чем -9.

[237] В другом варианте осуществления полипептиды XTEN по настоящему изобретению, в том числе те, которые включены в слитые белки BPXTEN для человека или животного, можно сделать практически неиммуногенными за счет рестрикции известных протеолитических сайтов из последовательности полипептида XTEN, уменьшая процессинг полипептидов XTEN до небольших пептидов, которые могут связываться с рецепторами MHC II. В другом варианте осуществления полипептид XTEN можно сделать практически неиммуногенным за счет применения последовательности, практически лишенной вторичной структуры, придающей устойчивость ко многим протеазам благодаря высокой энтропии структуры. Соответственно, сниженная оценка TEPITOPE и устранение известных протеолитических сайтов из полипептидов XTEN может привести к тому, что композиции на основе XTEN-полипептида, включая полипептиды XTEN композиций на основе слитых белков BPXTEN, практически не смогут связываться с рецепторами млекопитающих, включая рецепторы иммунной системы. В одном варианте осуществления полипептид XTEN слитого белка BPXTEN может характеризоваться связыванием с рецептором млекопитающего с K_d, составляющей >100 нМ, или K_d, составляющей более 500 нМ, или K_d, составляющей более 1 мкМ в отношении полипептидного рецептора на поверхности клетки млекопитающего или в

кровотоке.

[238] Кроме того, практически неповторяющаяся последовательность и соответствующее отсутствие эпитопов в таких вариантах осуществления полипептидов XTEN могут ограничивать способность В-клеток связываться с полипептидами XTEN или активироваться ими. В то время как полипептид XTEN может вступать в контакт со многими различными В-клетками в пределах своей удлиненной последовательности, каждая отдельная В-клетка может вступать только в один или небольшое число контактов с отдельным полипептидом XTEN. В результате полипептиды XTEN обычно могут иметь гораздо меньшую склонность стимулировать пролиферацию В-клеток и, следовательно, иммунный ответ. В одном варианте осуществления BPXTEN может иметь пониженную иммуногенность по сравнению с соответствующим неслитым BP. В одном варианте осуществления введение до трех парентеральных доз BPXTEN млекопитающему может привести к обнаружению IgG к BPXTEN при разведении сыворотки 1:100, но не при разведении 1:1000. В другом варианте осуществления введение до трех парентеральных доз BPXTEN млекопитающему может привести к обнаружению IgG к BP при разведении сыворотки 1:100, но не при разведении 1:1000. В другом варианте осуществления введение до трех парентеральных доз BPXTEN млекопитающему может привести к обнаружению IgG к XTEN при разведении сыворотки 1:100, но не при разведении 1:1000. В предшествующих вариантах осуществления млекопитающее может представлять собой мышь, крысу, кролика или макака-крабоеда.

[239] Дополнительным признаком определенных вариантов осуществления полипептидов XTEN с практически неповторяющимися последовательностями по сравнению с полипептидами с менее неповторяющимися последовательностями (например, с последовательностями, имеющими три идентичных смежных аминокислоты) может быть то, что неповторяющиеся полипептиды XTEN образуют более слабые контакты с антителами (*например*, моновалентные взаимодействия), что приводит к меньшей вероятности иммунного клиренса, где композиции на основе BPXTEN могут оставаться в кровотоке в течение увеличенного периода времени.

[240] В некоторых вариантах осуществления (слитый) полипептид является менее иммуногенным по сравнению с биологически активным полипептидом, не связанным с полипептидом XTEN, где иммуногенность устанавливают путем измерения продукции антител IgG, которые избирательно связываются с биологически активным полипептидом, после введения сравнимых доз человеку или животному.

Спейсеры и сегмент для обеспечения высвобождения BP

[241] В некоторых вариантах осуществления по меньшей мере часть биологической активности соответствующего BP сохраняется интактным BPXTEN. В некоторых вариантах осуществления компонент BP либо становится биологически активным, либо имеет повышенную биологическую активность после его высвобождения из полипептида(-ов) XTEN путем расщепления необязательной последовательности расщепления, включенной в спейсерные последовательности в BPXTEN, как более

подробно описано в данном документе ниже.

[242] Любая группа спейсерных последовательностей является необязательной в слитых белках, охватываемых настоящим изобретением. Спейсер может быть предусмотрен для усиления экспрессии слитого белка из клетки-хозяина или для уменьшения стерических затруднений там, где компонент BP может принимать необходимую третичную структуру и/или соответствующим образом взаимодействовать со своей молекулой-мишенью. О спейсерах и способах идентификации требуемых спейсеров см., например, George, et al. (2003) Protein Engineering 15:871-879, специально включенную в данный документ посредством ссылки. В одном варианте осуществления спейсер содержит одну или более пептидных последовательностей длиной от 1 до 50 аминокислотных остатков, или приблизительно от 1 до 25 остатков, или приблизительно от 1 до 10 остатков. Спейсерные последовательности, исключая сайты расщепления, могут предусматривать любую из 20 природных L-аминокислот и предпочтительно содержат гидрофильные аминокислоты, которые стерически не затруднены и которые могут включать без ограничения глицин (G), аланин (A), серин (S), треонин (T), глутаминовую кислоту (E) и пролин (P). В некоторых вариантах осуществления спейсер может представлять собой полиглицины или полиаланины или преимущественно смесь комбинаций остатков глицина и аланина. Спейсерный полипептид, за исключением последовательности расщепления, в значительной степени практически лишен вторичной структуры. В одном варианте осуществления одна или обе спейсерные последовательности в композиции на основе слитого белка BPXTEN могут дополнительно содержать последовательность расщепления, которая может быть идентичной или может быть различной, где на последовательность расщепления может воздействовать протеаза для высвобождения BP из слитого белка.

[243] В некоторых случаях включение последовательности расщепления в BPXTEN предназначено для обеспечения высвобождения BP, который становится активным или более активным после его высвобождения из полипептида XTEN. Последовательности расщепления располагаются достаточно близко к последовательностям BP, обычно в пределах 18, или 12, или 6, или 2 аминокислот от конца последовательности BP, где любые оставшиеся остатки, присоединенные к BP после расщепления, не мешают заметно активности (*например*, такой как связывание с рецептором) BP, но при этом обеспечивают достаточный доступ протеазе, чтобы иметь возможность осуществлять расщепление последовательности расщепления. В некоторых вариантах осуществления сайт расщепления представляет собой последовательность, которая может быть расщеплена протеазой, эндогенной для человека или животного-млекопитающего, где BPXTEN может расщепляться после введения человеку или животному. В таких случаях BPXTEN может служить пролекарством или депо для BP в кровотоке. Примеры сайтов расщепления, предусмотренных настоящим изобретением, включают без ограничения полипептидную последовательность, расщепляемую эндогенной протеазой млекопитающих, выбранной из FXIa, FXIIa, калликреина, FVIIa,

FIXa, FXa, FIIa (тромбина), эластазы-2, гранзима B, MMP-12, MMP-13, MMP-17 или MMP-20, или протеазами организмов, не относящихся к млекопитающим, такими как TEV, энтерокиназа, протеаза PreScission™ (протеаза 3C риновируса) и сортаза A. Последовательности, которые, как известно, расщепляются вышеуказанными протеазами, известны в уровне техники. Иллюстративные последовательности расщепления и сайты расщепления в последовательностях представлены в **таблице 7a**, также как и варианты последовательностей. Например, тромбин (активированный фактор свертывания II) действует на последовательность LTPRSLLV (SEQ ID NO: 222) {Rawlings N.D., et al. (2008) Nucleic Acids Res., **36**: D320}, которая будет расщепляться после аргинина в положении 4 в данной последовательности. Активный FIIa вырабатывается при расщеплении FII под действием FXa в присутствии фосфолипидов и кальция и расположен ниже фактора IX в каскаде коагуляции. После активации его естественная роль в коагуляции заключается в расщеплении фибриногена, который затем, в свою очередь, начинает формирование сгустка. Активность FIIa строго контролируется и возникает только тогда, когда коагуляция необходима для надлежащего гемостаза. Однако, поскольку коагуляция является у млекопитающих непрерывным процессом, при включении последовательности LTPRSLLV (SEQ ID NO: 223) в BPXTEN между BP и полипептидом XTEN полипептиды XTEN будут удаляться из прилегающих BP одновременно с активацией либо внешнего, либо внутреннего каскада коагуляции, когда коагуляция физиологически необходима, высвобождая тем самым BP с течением времени. Аналогичным образом, включение в BPXTEN других последовательностей, на которые действуют эндогенные протеазы, могло бы обеспечить замедленное высвобождение BP, что в определенных случаях может обеспечить более высокую степень активности BP из «пролекарственной» формы BPXTEN.

[244] В некоторых случаях только две или три аминокислоты, фланкирующие обе стороны сайта расщепления (всего от четырех до шести аминокислот), будут включены в последовательность расщепления. В других случаях известная последовательность расщепления может иметь одну или более делеций или вставок, или одну, или две, или три аминокислотные замены для любой одной, двух или трех аминокислот в известной последовательности, где делеции, вставки или замены приводят к уменьшению или повышению чувствительности, но не отсутствию чувствительности к протеазе, давая возможность адаптировать скорость высвобождения BP из XTEN. Иллюстративные замены показаны в **таблице 7a**.

Таблица 7a. Последовательности расщепления протеазой для высвобождения BP

Протеаза, действующая на последовательность	**SEQ ID NO**	**Иллюстративная последовательность расщепления**	**Минимальный сайт расщепления***
FXIa	224	KLTR↓VVGG	KD/FL/T/R↓VA/VE/GT/GV
FXIIa	225	TMTR↓IVGG	NA
Калликреин	226	SPFR↓STGG	-/-/FL/RY↓SR/RT/-/-
FVIIa	227	LQVR↓IVGG	NA

Протеаза, действующая на последовательность	SEQ ID NO	Иллюстративная последовательность расщепления	Минимальный сайт расщепления*
FIXa	228	PLGR↓IVGG	-/-/G/R↓-/-/-/-
FXa	229	IEGR↓TVGG	IA/E/GFP/R↓STI/VFS/-/G
FIIa (тромбин)	230	LTPR↓SLLV	-/-/PLA/R↓SAG/-/-/-
Эластаза-2	231	LGPV↓SGVP	-/-/-/VIAT↓-/-/-/-
Гранзим-B	232	VAGD↓SLEE	V/-/-/D↓-/-/-/-
MMP-12	233	GPAG↓LGGA	G/PA/-/G↓L/-/G/- (SEQ ID NO: 241)
MMP-13	234	GPAG↓LRGA	G/P/-/G↓L/-/GA/- (SEQ ID NO: 242)
MMP-17	235	APLG↓LRLR	-/PS/-/-↓LQ/-/LT/-
MMP-20	236	PALP↓LVAQ	NA
TEV	237	ENLYFQ↓G	ENLYFQ↓G/S (SEQ ID NO: 243)
Энеокиназа	238	DDDK↓IVGG	DDDK↓IVGG (SEQ ID NO: 244)
Протеаза 3C (PreScission™)	239	LEVLFQ↓GP	LEVLFQ↓GP (SEQ ID NO: 245)
Сортаза A	240	LPKT↓GSES	L/P/KEAD/T↓G/-/EKS/S (SEQ ID NO: 246)

↓ указывает сайт расщепления;

NA: не применимо;

* перечисление нескольких аминокислот до, между или после косой черты указывает альтернативные аминокислоты, на которые может быть проведена замена в данном положении;

«-» указывает, что любая аминокислота может быть заменой для соответствующей аминокислоты, указанной в средней колонке

[245] В некоторых вариантах осуществления слитый белок BPXTEN может содержать спейсерные последовательности, которые могут дополнительно содержать одну или более последовательностей расщепления, сконфигурированных для высвобождения BP из слитого белка под действием протеазы. В некоторых вариантах осуществления одна или более последовательностей расщепления могут представлять собой последовательность, характеризующуюся по меньшей мере приблизительно 80% (например, по меньшей мере приблизительно 85%, по меньшей мере приблизительно 90%, по меньшей мере приблизительно 95%, по меньшей мере приблизительно 96%, по меньшей мере приблизительно 97%, по меньшей мере приблизительно 98%, по меньшей мере приблизительно 99% или 100%) идентичностью последовательности с последовательностью, представленной в **таблице 7a**.

[246] В некоторых вариантах осуществления настоящее изобретение предусматривает пептиды сегмента для обеспечения высвобождения BP (или сегмент для обеспечения высвобождения (RS)), которые являются субстратами для одной или более протеаз млекопитающих, ассоциированных с пораженными тканями или клетками или

продуцируемых ими, обнаруживаемых вблизи пораженных тканей. Такие протеазы могут включать в себя без ограничения классы протеаз, такие как металлопротеиназы, цистеиновые протеазы, аспартатпротеазы и сериновые протеазы, в том числе без ограничения протеазы, представленные в **таблице 7b**. RS применимы, среди прочего, для включения в рекомбинантные полипептиды человека или животного, для придания формата пролекарства, который может быть активирован путем расщепления RS протеазами млекопитающих. Как описано в данном документе, RS включены в состав композиций на основе рекомбинантных полипептидов человека или животного, связывая включенные в их состав связывающие фрагменты с XTEN (конфигурации которых более подробно описаны ниже), где при расщеплении RS под действием одной или более протеаз, для которых RS являются субстратами, связывающие фрагменты и XTEN высвобождаются из композиции, и связывающие фрагменты, больше не экранированные с помощью XTEN, восстанавливают свой полный потенциал связывания со своими лигандами. В этих композициях на основе рекомбинантных полипептидов, содержащих первый и второй фрагмент антитела, композиции также упоминаются в данном документе как активируемые композиции на основе антител (AAC).

Таблица 7b. Протеазы тканей-мишеней

Класс протеаз	**Протеаза**
Металлопротеиназы	Меприн
	Неприлизин (CD10)
	PSMA
	BMP-1
	Дизинтегрин и металлопротеиназы (ADAM)
	ADAM8
	ADAM9
	ADAM10
	ADAM12
	ADAM15
	ADAM17 (TACE)
	ADAM19
	ADAM28 (MDC-L)
	ADAM с мотивами тромбоспондина (ADAMTS)
	ADAMTS1
	ADAMTS4
	ADAMTS5
	Матричные металлопротеиназы (MMP)
	MMP-1 (коллагеназа 1)
	MMP-2 (желатиназа A)
	MMP-3 (m1)
	MMP-7 (матрилизин 1)
	MMP-8 (коллагеназа 2)
	MMP-9 (желатиназа B)
	MMP-10 (стромелизин 2)
	MMP-11 (стромелизин 3)
	MMP-12 (эластаза макрофагов)
	MMP-13 (коллагеназа 3)

	MMP-14 (MT1-MMP)
	MMP-15 (MT2-MMP)
	MMP-19
	MMP-23 (CA-MMP)
	MMP-24 (MT5-MMP)
	MMP-26 (матрилизин 2)
	MMP-27 (CMMP)
Цистеиновые протеазы	Легумаин
	Цистеиновые катепсины
	Катепсин B
	Катепсин C
	Катепсин K
	Катепсин L
	Катепсин S
	Катепсин X
Аспартатпротеазы	Катепсин D
	Катепсин E
	Секретаза
Сериновые протеазы	Урокиназа (uPA)
	Тканевой активатор плазминогена (tPA)
	Плазмин
	Тромбин
	Простатоспецифический антиген (PSA, KLK3)
	Эластаза нейтрофилов человека (HNE)
	Эластаза
	Триптаза
	Трансмембранные сериновые протеазы типа II (TTSP)
	DESC1
	Гепсин (HPN)
	Матриптаза
	Матриптаза-2
	TMPRSS2
	TMPRSS3
	TMPRSS4 (CAP2)
	Белок активации фибробластов (FAP)
	Пептидаза, родственная калликреину (семейство KLK)
	KLK4
	KLK5
	KLK6
	KLK7
	KLK8
	KLK10
	KLK11
	KLK13
	KLK14

[247] В одном варианте осуществления настоящее изобретение предусматривает активируемые рекомбинантные полипептиды, содержащие последовательность первого сегмента для обеспечения высвобождения (RS1), характеризующуюся по меньшей мере 88%, или по меньшей мере 94% или 100% идентичностью последовательности при

оптимальном выравнивании по отношению к последовательности, выбранной из последовательностей, представленных в **таблице 8a**, где RS1 представляет собой субстрат для одной или более протеаз млекопитающих. В других вариантах осуществления настоящее изобретение предусматривает активируемые рекомбинантные полипептиды, содержащие RS1 и последовательность второго сегмента для обеспечения высвобождения (RS2), каждая из которых характеризуется по меньшей мере 88%, или по меньшей мере 94% или 100% идентичностью последовательности при оптимальном выравнивании по отношению к последовательности, выбранной из последовательностей, представленных в **таблице 8a**, где каждый из RS1 и RS2 является субстратом для одной или более протеаз млекопитающих. В другом варианте осуществления настоящее изобретение предусматривает активируемые рекомбинантные полипептиды, содержащие первую последовательность RS (RS1), характеризующуюся по меньшей мере 90%, по меньшей мере 93%, по меньшей мере 97% или 100% идентичностью при оптимальном выравнивании по отношению к последовательности, выбранной из последовательностей, представленных в **таблице 8b**, где RS представляет собой субстрат для одной или более протеаз млекопитающих. В других вариантах осуществления настоящее изобретение предусматривает активируемые рекомбинантные полипептиды, содержащие RS1 и последовательность второго сегмента для обеспечения высвобождения (RS2), каждая из которых характеризуется по меньшей мере 88%, или по меньшей мере 94%, или 100% идентичностью последовательности при оптимальном выравнивании по отношению к последовательности, выбранной из последовательностей, представленных в **таблице 8b**, где каждый из RS1 и RS2 является субстратом для одной или более протеаз млекопитающих. В вариантах осуществления активируемых рекомбинантных полипептидов, содержащих RS1 и RS2, два сегмента для обеспечения высвобождения могут быть идентичными или их последовательности могут быть разными.

[248] В настоящем изобретении рассматриваются сегменты для обеспечения высвобождения, которые являются субстратами для одного, двух или трех различных классов протеаз, выбранных из металлопротеиназ, цистеиновых протеаз, аспартатпротеаз и сериновых протеаз, включая протеазы, представленные в **таблице 7b**. В конкретном признаке RS служат субстратами для протеаз, находящихся в тесной ассоциации или совместно локализованных с пораженными тканями или клетками, такими как без ограничения опухоли, раковые клетки и воспаленные ткани, и при расщеплении RS связывающие фрагменты, которые в иных случаях экранированы с помощью XTEN в композициях на основе рекомбинантных полипептидов человека или животного (и, таким образом, характеризуются более низкой аффинностью связывания со своими соответствующими лигандами), высвобождаются из композиции и восстанавливают свой полный потенциал связывания с лигандами клеток-мишеней и/или эффекторных клеток. В другом варианте осуществления RS композиций на основе рекомбинантных полипептидов человека или животного содержит аминокислотную последовательность, которая является субстратом для клеточной протеазы, локализованной в клетке-мишени, включая без

ограничения протеазы, представленные в **таблице 7b**. В другом конкретном признаке композиций на основе рекомбинантных полипептидов человека или животного RS, которые являются субстратами для двух или трех классов протеаз, разрабатывали с последовательностями, которые способны расщепляться в различных местоположениях последовательности RS различными протеазами. Таким образом, RS, которые являются субстратами для двух, трех или более классов протеаз, имеют два, три или множество различных сайтов расщепления в последовательности RS, но расщепление одной протеазой, тем не менее, приводит к высвобождению связывающих фрагментов и XTEN из композиции на основе рекомбинантных полипептидов, содержащей RS.

[249] В одном варианте осуществления RS по настоящему изобретению для включения в композиции на основе рекомбинантных полипептидов человека или животного представляет собой субстрат для одной или более протеаз, включая меприн, неприлизин (CD10), PSMA, BMP-1, дизинтегрин и металлопротеиназы (ADAM), ADAM8, ADAM9, ADAM10, ADAM12, ADAM15, ADAM17 (TACE), ADAM19, ADAM28 (MDC-L), ADAM с мотивами тромбоспондина (ADAMTS), ADAMTS1, ADAMTS4, ADAMTS5, MMP-1 (коллагеназа 1), матриксную металлопротеиназу-1 (MMP-1), матриксную металлопротеиназу-2 (MMP-2, желатиназа A), матриксную металлопротеиназу-3 (MMP-3, стромелизин 1), матриксную металлопротеиназу-7 (MMP-7, матрилизин 1), матриксную металлопротеиназу-8 (MMP-8, коллагеназа 2), матриксную металлопротеиназу-9 (MMP-9, желатиназа B), матриксную металлопротеиназу-10 (MMP-10, стромелизин 2), матриксную металлопротеиназу-11 (MMP-11, стромелизин 3), матриксную металлопротеиназу-12 (MMP-12, эластаза макрофагов), матриксную металлопротеиназу-13 (MMP-13, коллагеназа 3), матриксную металлопротеиназу-14 (MMP-14, MT1-MMP), матриксную металлопротеиназу-15 (MMP-15, MT2-MMP), матриксную металлопротеиназу-19 (MMP-19), матриксную металлопротеиназу-23 (MMP-23, CA-MMP), матриксную металлопротеиназу-24 (MMP-24, MT5-MMP), матриксную металлопротеиназу-26 (MMP-26, матрилизин 2), матриксную металлопротеиназу-27 (MMP-27, CMMP), легумаин, катепсин B, катепсин C, катепсин K, катепсин L, катепсин S, катепсин X, катепсин D, катепсин E, секретазу, урокиназу (uPA), тканевой активатор плазминогена (tPA), плазмин, тромбин, простатоспецифический антиген (PSA, KLK3), эластазу нейтрофилов человека (HNE), эластазу, триптазу, трансмембранные сериновые протеазы II типа (TTSP), DESC1, гепсин (HPN), матриптазу, матриптазу-2, TMPRSS2, TMPRSS3, TMPRSS4 (CAP2), белок активации фибробластов (FAP), пептидазу, родственную калликреину (семейство KLK), KLK4, KLK5, KLK6, KLK7, KLK8, KLK10, KLK11, KLK13 и KLK14. В одном варианте осуществления RS представляет собой субстрат для ADAM17. В одном варианте осуществления RS представляет собой субстрат для BMP-1. В одном варианте осуществления RS представляет собой субстрат для катепсина. В одном варианте осуществления RS представляет собой субстрат для HtrA1. В одном варианте осуществления RS представляет собой субстрат для легумаина. В одном варианте осуществления RS представляет собой субстрат для MMP-1. В одном варианте

осуществления RS представляет собой субстрат для MMP-2. В одном варианте осуществления RS представляет собой субстрат для MMP-7. В одном варианте осуществления RS представляет собой субстрат для MMP-9. В одном варианте осуществления RS представляет собой субстрат для MMP-11. В одном варианте осуществления RS представляет собой субстрат для MMP-14. В одном варианте осуществления RS представляет собой субстрат для uPA. В одном варианте осуществления RS представляет собой субстрат для матриптазы. В одном варианте осуществления RS представляет собой субстрат для MT-SP1. В одном варианте осуществления RS представляет собой субстрат для эластазы нейтрофилов. В одном варианте осуществления RS представляет собой субстрат для тромбина. В одном варианте осуществления RS представляет собой субстрат для TMPRSS3. В одном варианте осуществления RS представляет собой субстрат для TMPRSS4. В одном варианте осуществления RS композиций на основе рекомбинантных полипептидов человека или животного является субстратом для по меньшей мере двух протеаз, которые представляют собой легумаин, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA и матриптазу. В другом варианте осуществления RS композиций на основе рекомбинантных полипептидов человека или животного представляет собой субстрат для легумаина, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA и матриптазы.

Таблица 8а. Последовательности сегмента для обеспечения высвобождения BP.

Название	**ID конструкции**	**Аминокислотная последовательность**	**SEQ ID NO:**
RSR-1517	AC1611	EAGRSANHEPLGLVAT	8261
BSRS-A1	AC1605	ASGRSTNAGPSGLAGP	8262
BSRS-A2	AC1606	ASGRSTNAGPQGLAGQ	8263
BSRS-A3	AC1607	ASGRSTNAGPPGLTGP	8264
VP-1	AC1608	ASSRGTNAGPAGLTGP	8265
RSR-1752	AC1609	ASSRTTNTGPSTLTGP	8266
RSR-1512	AC1610	AAGRSDNGTPLELVAP	8267
RSR-1517	AC1611	EAGRSANHEPLGLVAT	8261
VP-2	AC1612	ASGRGTNAGPAGLTGP	8268
RSR-1018	AC1613	LFGRNDNHEPLELGGG	8269
RSR-1053	AC1614	TAGRSDNLEPLGLVFG	8270
RSR-1059	AC1615	LDGRSDNFHPPELVAG	8271
RSR-1065	AC1616	LEGRSDNEEPENLVAG	8272
RSR-1167	AC1617	LKGRSDNNAPLALVAG	8273
RSR-1201	AC1618	VYSRGTNAGPHGLTGR	8274
RSR-1218	AC1619	ANSRGTNKGFAGLIGP	8275
RSR-1226	AC1620	ASSRLTNEAPAGLTIP	8276
RSR-1254	AC1621	DQSRGTNAGPEGLTDP	8277
RSR-1256	AC1622	ESSRGTNIGQGGLTGP	8278
RSR-1261	AC1623	SSSRGTNQDPAGLTIP	8279
RSR-1293	AC1624	ASSRGQNHSPMGLTGP	8280
RSR-1309	AC1625	AYSRGPNAGPAGLEGR	8281
RSR-1326	AC1626	ASERGNNAGPANLTGF	8282
RSR-1345	AC1627	ASHRGTNPKPAILTGP	8283
RSR-1354	AC1628	MSSRRTNANPAQLTGP	8284

RSR-1426	AC1629	GAGRTDNHEPLELGAA	8285
RSR-1478	AC1630	LAGRSENTAPLELTAG	8286
RSR-1479	AC1631	LEGRPDNHEPLALVAS	8287
RSR-1496	AC1632	LSGRSDNEEPLALPAG	8288
RSR-1508	AC1633	EAGRTDNHEPLELSAP	8289
RSR-1513	AC1634	EGGRSDNHGPLELVSG	8290
RSR-1516	AC1635	LSGRSDNEAPLELEAG	8291
RSR-1524	AC1636	LGGRADNHEPPELGAG	8292
RSR-1622	AC1637	PPSRGTNAEPAGLTGE	8293
RSR-1629	AC1638	ASTRGENAGPAGLEAP	8294
RSR-1664	AC1639	ESSRGTNGAPEGLTGP	8295
RSR-1667	AC1640	ASSRATNESPAGLTGE	8296
RSR-1709	AC1641	ASSRGENPPPGGLTGP	8297
RSR-1712	AC1642	AASRGTNTGPAELTGS	8298
RSR-1727	AC1643	AGSRTTNAGPGGLEGP	8299
RSR-1754	AC1644	APSRGENAGPATLTGA	8300
RSR-1819	AC1645	ESGRAANTGPPTLTAP	8301
RSR-1832	AC1646	NPGRAANEGPPGLPGS	8302
RSR-1855	AC1647	ESSRAANLTPPELTGP	8303
RSR-1911	AC1648	ASGRAANETPPGLTGA	8304
RSR-1929	AC1649	NSGRGENLGAPGLTGT	8305
RSR-1951	AC1650	TTGRAANLTPAGLTGP	8306
RSR-2295	AC1761	EAGRSANHTPAGLTGP	8307
RSR-2298	AC1762	ESGRAANTTPAGLTGP	8308
RSR-2038	AC1679	TTGRATEAANLTPAGLTGP	8309
RSR-2072	AC1680	TTGRAEEAANLTPAGLTGP	8310
RSR-2089	AC1681	TTGRAGEAANLTPAGLTGP	8311
RSR-2302	AC1682	TTGRATEAANATPAGLTGP	8312
RSR-3047	AC1697	TTGRAGEAEGATSAGATGP	8313
RSR-3052	AC1698	TTGEAGEAANATSAGATGP	8314
RSR-3043	AC1699	TTGEAGEAAGLTPAGLTGP	8315
RSR-3041	AC1700	TTGAAGEAANATPAGLTGP	8316
RSR-3044	AC1701	TTGRAGEAAGLTPAGLTGP	8317
RSR-3057	AC1702	TTGRAGEAANATSAGATGP	8318
RSR-3058	AC1703	TTGEAGEAAGATSAGATGP	8319
RSR-2485	AC1763	ESGRAANTEPPELGAG	8320
RSR-2486	AC1764	ESGRAANTAPEGLTGP	8321
RSR-2488	AC1688	EPGRAANHEPSGLTEG	8322
RSR-2599	AC1706	ESGRAANHTGAPPGGLTGP	8323
RSR-2706	AC1716	TTGRTGEGANATPGGLTGP	8324
RSR-2707	AC1717	RTGRSGEAANETPEGLEGP	8325
RSR-2708	AC1718	RTGRTGESANETPAGLGGP	8326
RSR-2709	AC1719	STGRTGEPANETPAGLSGP	8327
RSR-2710	AC1720	TTGRAGEPANATPTGLSGP	8328
RSR-2711	AC1721	RTGRPGEGANATPTGLPGP	8329
RSR-2712	AC1722	RTGRGGEAANATPSGLGGP	8330
RSR-2713	AC1723	STGRSGESANATPGGLGGP	8331
RSR-2714	AC1724	RTGRTGEEANATPAGLPGP	8332
RSR-2715	AC1725	ATGRPGEPANTTPEGLEGP	8333
RSR-2716	AC1726	STGRSGEPANATPGGLTGP	8334

RSR-2717	AC1727	PTGRGGEGANTTPTGLPGP	8335
RSR-2718	AC1728	PTGRSGEGANATPSGLTGP	8336
RSR-2719	AC1729	TTGRASEGANSTPAPLTEP	8337
RSR-2720	AC1730	TYGRAAEAANTTPAGLTAP	8338
RSR-2721	AC1731	TTGRATEGANATPAELTEP	8339
RSR-2722	AC1732	TVGRASEEANTTPASLTGP	8340
RSR-2723	AC1733	TTGRAPEAANATPAPLTGP	8341
RSR-2724	AC1734	TWGRATEPANATPAPLTSP	8342
RSR-2725	AC1735	TVGRASESANATPAELTSP	8343
RSR-2726	AC1736	TVGRAPEGANSTPAGLTGP	8344
RSR-2727	AC1737	TWGRATEAPNLEPATLTTP	8345
RSR-2728	AC1738	TTGRATEAPNLTPAPLTEP	8346
RSR-2729	AC1739	TQGRATEAPNLSPAALTSP	8347
RSR-2730	AC1740	TQGRAAEAPNLTPATLTAP	8348
RSR-2731	AC1741	TSGRAPEATNLAPAPLTGP	8349
RSR-2732	AC1742	TQGRAAEAANLTPAGLTEP	8350
RSR-2733	AC1743	TTGRAGSAPNLPPTGLTTP	8351
RSR-2734	AC1744	TTGRAGGAENLPPEGLTAP	8352
RSR-2735	AC1745	TTSRAGTATNLTPEGLTAP	8353
RSR-2736	AC1746	TTGRAGTATNLPPSGLTTP	8354
RSR-2737	AC1747	TTARAGEAENLSPSGLTAP	8355
RSR-2738	AC1748	TTGRAGGAGNLAPGGLTEP	8356
RSR-2739	AC1749	TTGRAGTATNLPPEGLTGP	8357
RSR-2740	AC1750	TTGRAGGAANLAPTGLTEP	8358
RSR-2741	AC1751	TTGRAGTAENLAPSGLTTP	8359
RSR-2742	AC1752	TTGRAGSATNLGPGGLTGP	8360
RSR-2743	AC1753	TTARAGGAENLTPAGLTEP	8361
RSR-2744	AC1754	TTARAGSAENLSPSGLTGP	8362
RSR-2745	AC1755	TTARAGGAGNLAPEGLTTP	8363
RSR-2746	AC1756	TTSRAGAAENLTPTGLTGP	8364
RSR-2747	AC1757	TYGRTTTPGNEPPASLEAE	8365
RSR-2748	AC1758	TYSRGESGPNEPPPGLTGP	8366
RSR-2749	AC1759	AWGRTGASENETPAPLGGE	8367
RSR-2750	AC1760	RWGRAETTPNTPPEGLETE	8368
RSR-2751	AC1765	ESGRAANHTGAEPPELGAG	8369
RSR-2754	AC1801	TTGRAGEAANLTPAGLTES	8370
RSR-2755	AC1802	TTGRAGEAANLTPAALTES	8371
RSR-2756	AC1803	TTGRAGEAANLTPAPLTES	8372
RSR-2757	AC1804	TTGRAGEAANLTPEPLTES	8373
RSR-2758	AC1805	TTGRAGEAANLTPAGLTGA	8374
RSR-2759	AC1806	TTGRAGEAANLTPEGLTGA	8375
RSR-2760	AC1807	TTGRAGEAANLTPEPLTGA	8376
RSR-2761	AC1808	TTGRAGEAANLTPAGLTEA	8377
RSR-2762	AC1809	TTGRAGEAANLTPEGLTEA	8378
RSR-2763	AC1810	TTGRAGEAANLTPAPLTEA	8379
RSR-2764	AC1811	TTGRAGEAANLTPEPLTEA	8380
RSR-2765	AC1812	TTGRAGEAANLTPEPLTGP	8381
RSR-2766	AC1813	TTGRAGEAANLTPAGLTGG	8382
RSR-2767	AC1814	TTGRAGEAANLTPEGLTGG	8383
RSR-2768	AC1815	TTGRAGEAANLTPEALTGG	8384

RSR-2769	AC1816	TTGRAGEAANLTPEPLTGG	8385
RSR-2770	AC1817	TTGRAGEAANLTPAGLTEG	8386
RSR-2771	AC1818	TTGRAGEAANLTPEGLTEG	8387
RSR-2772	AC1819	TTGRAGEAANLTPAPLTEG	8388
RSR-2773	AC1820	TTGRAGEAANLTPEPLTEG	8389

Таблица 8b. Последовательности сегмента для обеспечения высвобождения

Название	**Аминокислотная последовательность**	**SEQ ID NO:**	**Название**	**Аминокислотная последовательность**	**SEQ ID NO:**
RSN-0001	GSAPGSAGGYAELR MGGAIATSGSETPG T	8390	RSC-0001	GTAEAASASGGSAG GYAELRMGGAIPGS P	8635
RSN-0002	GSAPGTGGGYAPLR MGGGAATSGSETPG T	8391	RSC-0002	GTAEAASASGGTGG GYAPLRMGGGAPGS P	8636
RSN-0003	GSAPGAEGGYAALR MGGEIATSGSETPGT	8392	RSC-0003	GTAEAASASGGAEG GYAALRMGGEIPGS P	8637
RSN-0004	GSAPGGPGGYALLR MGGPAATSGSETPG T	8393	RSC-0004	GTAEAASASGGGPG GYALLRMGGPAPGS P	8638
RSN-0005	GSAPGEAGGYAFLR MGGSIATSGSETPGT	8394	RSC-0005	GTAEAASASGGEAG GYAFLRMGGSIPGSP	8639
RSN-0006	GSAPGPGGGYASLR MGGTAATSGSETPG T	8395	RSC-0006	GTAEAASASGGPGG GYASLRMGGTAPGS P	8640
RSN-0007	GSAPGSEGGYATLR MGGAIATSGSETPG T	8396	RSC-0007	GTAEAASASGGSEG GYATLRMGGAIPGS P	8641
RSN-0008	GSAPGTPGGYANLR MGGGAATSGSETPG T	8397	RSC-0008	GTAEAASASGGTPG GYANLRMGGGAPG SP	8642
RSN-0009	GSAPGASGGYAHLR MGGEIATSGSETPGT	8398	RSC-0009	GTAEAASASGGASG GYAHLRMGGEIPGS P	8643
RSN-0010	GSAPGGTGGYGELR MGGPAATSGSETPG T	8399	RSC-0010	GTAEAASASGGGTG GYGELRMGGPAPGS P	8644
RSN-0011	GSAPGEAGGYPELR MGGSIATSGSETPGT	8400	RSC-0011	GTAEAASASGGEAG GYPELRMGGSIPGSP	8645
RSN-0012	GSAPGPGGGYVELR MGGTAATSGSETPG T	8401	RSC-0012	GTAEAASASGGPGG GYVELRMGGTAPGS P	8646
RSN-0013	GSAPGSEGGYLELR MGGAIATSGSETPG T	8402	RSC-0013	GTAEAASASGGSEG GYLELRMGGAIPGSP	8647
RSN-0014	GSAPGTPGGYSELR MGGGAATSGSETPG T	8403	RSC-0014	GTAEAASASGGTPG GYSELRMGGGAPGS P	8648
RSN-0015	GSAPGASGGYTELR MGGEIATSGSETPGT	8404	RSC-0015	GTAEAASASGGASG GYTELRMGGEIPGSP	8649

RSN-0016	GSAPGGTGGYQELRMGGPAATSGSETPGT	8405	RSC-0016	GTAEAASASGGGTGGYQELRMGGPAPGSP	8650
RSN-0017	GSAPGEAGGYEELRMGGSIATSGSETPGT	8406	RSC-0017	GTAEAASASGGEAGGYEELRMGGSIPGSP	8651
RSN-0018	GSAPGPGIGPAELRMGGTAATSGSETPGT	8407	RSC-0018	GTAEAASASGGPGIGPAELRMGGTAPGSP	8652
RSN-0019	GSAPGSEIGAAELRMGGAIATSGSETPGT	8408	RSC-0019	GTAEAASASGGSEIGAAELRMGGAIPGSP	8653
RSN-0020	GSAPGTPIGSAELRMGGGAATSGSETPGT	8409	RSC-0020	GTAEAASASGGTPIGSAELRMGGGAPGSP	8654
RSN-0021	GSAPGASIGTAELRMGGEIATSGSETPGT	8410	RSC-0021	GTAEAASASGGASIGTAELRMGGEIPGSP	8655
RSN-0022	GSAPGGTIGNAELRMGGPAATSGSETPGT	8411	RSC-0022	GTAEAASASGGGTIGNAELRMGGPAPGSP	8656
RSN-0023	GSAPGEAIGQAELRMGGSIATSGSETPGT	8412	RSC-0023	GTAEAASASGGEAIGQAELRMGGSIPGSP	8657
RSN-0024	GSAPGPGGPYAELRMGGTAATSGSETPGT	8413	RSC-0024	GTAEAASASGGPGGPYAELRMGGTAPGSP	8658
RSN-0025	GSAPGSEGAYAELRMGGAIATSGSETPGT	8414	RSC-0025	GTAEAASASGGSEGAYAELRMGGAIPGSP	8659
RSN-0026	GSAPGTPGVYAELRMGGGAATSGSETPGT	8415	RSC-0026	GTAEAASASGGTPGVYAELRMGGGAPGSP	8660
RSN-0027	GSAPGASGLYAELRMGGEIATSGSETPGT	8416	RSC-0027	GTAEAASASGGASGLYAELRMGGEIPGSP	8661
RSN-0028	GSAPGGTGIYAELRMGGPAATSGSETPGT	8417	RSC-0028	GTAEAASASGGGTGIYAELRMGGPAPGSP	8662
RSN-0029	GSAPGEAGFYAELRMGGSIATSGSETPGT	8418	RSC-0029	GTAEAASASGGEAGFYAELRMGGSIPGSP	8663
RSN-0030	GSAPGPGGYYAELRMGGTAATSGSETPGT	8419	RSC-0030	GTAEAASASGGPGGYYAELRMGGTAPGSP	8664
RSN-0031	GSAPGSEGSYAELRMGGAIATSGSETPGT	8420	RSC-0031	GTAEAASASGGSEGSYAELRMGGAIPGSP	8665
RSN-0032	GSAPGTPGNYAELRMGGGAATSGSETPGT	8421	RSC-0032	GTAEAASASGGTPGNYAELRMGGGAPGSP	8666
RSN-0033	GSAPGASGEYAELRMGGEIATSGSETPGT	8422	RSC-0033	GTAEAASASGGASGEYAELRMGGEIPGSP	8667
RSN-0034	GSAPGGTGHYAELRMGGPAATSGSETPGT	8423	RSC-0034	GTAEAASASGGGTGHYAELRMGGPAPGSP	8668
RSN-0035	GSAPGEAGGYAEAR	8424	RSC-0035	GTAEAASASGGEAG	8669

	MGGSIATSGSETPGT			GYAEARMGGSIPGSP	
RSN-0036	GSAPGPGGGYAEVRMGGTAATSGSETPGT	8425	RSC-0036	GTAEAASASGGPGGGYAEVRMGGTAPGSP	8670
RSN-0037	GSAPGSEGGYAEIRMGGAIATSGSETPGT	8426	RSC-0037	GTAEAASASGGSEGGYAEIRMGGAIPGSP	8671
RSN-0038	GSAPGTPGGYAEFRMGGGAATSGSETPGT	8427	RSC-0038	GTAEAASASGGTPGGYAEFRMGGGAPGSP	8672
RSN-0039	GSAPGASGGYAEYRMGGEIATSGSETPGT	8428	RSC-0039	GTAEAASASGGASGGYAEYRMGGEIPGSP	8673
RSN-0040	GSAPGGTGGYAESRMGGPAATSGSETPGT	8429	RSC-0040	GTAEAASASGGGTGGYAESRMGGPAPGSP	8674
RSN-0041	GSAPGEAGGYAETRMGGSIATSGSETPGT	8430	RSC-0041	GTAEAASASGGEAGGYAETRMGGSIPGSP	8675
RSN-0042	GSAPGPGGGYAELAMGGTRATSGSETPGT	8431	RSC-0042	GTAEAASASGGPGGGYAELAMGGTRPGSP	8676
RSN-0043	GSAPGSEGGYAELVMGGARATSGSETPGT	8432	RSC-0043	GTAEAASASGGSEGGYAELVMGGARPGSP	8677
RSN-0044	GSAPGTPGGYAELLMGGGRATSGSETPGT	8433	RSC-0044	GTAEAASASGGTPGGYAELLMGGGRPGSP	8678
RSN-0045	GSAPGASGGYAELIMGGERATSGSETPGT	8434	RSC-0045	GTAEAASASGGASGGYAELIMGGERPGSP	8679
RSN-0046	GSAPGGTGGYAELWMGGPRATSGSETPGT	8435	RSC-0046	GTAEAASASGGGTGGYAELWMGGPRPGSP	8680
RSN-0047	GSAPGEAGGYAELSMGGSRATSGSETPGT	8436	RSC-0047	GTAEAASASGGEAGGYAELSMGGSRPGSP	8681
RSN-0048	GSAPGPGGGYAELTMGGTRATSGSETPGT	8437	RSC-0048	GTAEAASASGGPGGGYAELTMGGTRPGSP	8682
RSN-0049	GSAPGSEGGYAELQMGGARATSGSETPGT	8438	RSC-0049	GTAEAASASGGSEGGYAELQMGGARPGSP	8683
RSN-0050	GSAPGTPGGYAELNMGGGRATSGSETPGT	8439	RSC-0050	GTAEAASASGGTPGGYAELNMGGGRPGSP	8684
RSN-0051	GSAPGASGGYAELEMGGERATSGSETPGT	8440	RSC-0051	GTAEAASASGGASGGYAELEMGGERPGSP	8685
RSN-0052	GSAPGGTGGYAELRPGGPIATSGSETPGT	8441	RSC-0052	GTAEAASASGGGTGGYAELRPGGPIPGSP	8686

RSN-0053	GSAPGEAGGYAELR AGGSAATSGSETPG T	8442	RSC-0053	GTAEAASASGGEAG GYAELRAGGSAPGS P	8687
RSN-0054	GSAPGPGGGYAELR LGGTIATSGSETPGT	8443	RSC-0054	GTAEAASASGGPGG GYAELRLGGTIPGSP	8688
RSN-0055	GSAPGSEGGYAELRI GGAAATSGSETPGT	8444	RSC-0055	GTAEAASASGGSEG GYAELRIGGAAPGSP	8689
RSN-0056	GSAPGTPGGYAELR SGGGIATSGSETPGT	8445	RSC-0056	GTAEAASASGGTPG GYAELRSGGGIPGSP	8690
RSN-0057	GSAPGASGGYAELR NGGEAATSGSETPG T	8446	RSC-0057	GTAEAASASGGASG GYAELRNGGEAPGS P	8691
RSN-0058	GSAPGGTGGYAELR QGGPIATSGSETPGT	8447	RSC-0058	GTAEAASASGGGTG GYAELRQGGPIPGSP	8692
RSN-0059	GSAPGEAGGYAELR DGGSAATSGSETPG T	8448	RSC-0059	GTAEAASASGGEAG GYAELRDGGSAPGS P	8693
RSN-0060	GSAPGPGGGYAELR EGGTIATSGSETPGT	8449	RSC-0060	GTAEAASASGGPGG GYAELREGGTIPGSP	8694
RSN-0061	GSAPGSEGGYAELR HGGAAATSGSETPG T	8450	RSC-0061	GTAEAASASGGSEG GYAELRHGGAAPGS P	8695
RSN-0062	GSAPGTPGGYAELR MPGGIATSGSETPGT	8451	RSC-0062	GTAEAASASGGTPG GYAELRMPGGIPGSP	8696
RSN-0063	GSAPGASGGYAELR MAGEAATSGSETPG T	8452	RSC-0063	GTAEAASASGGASG GYAELRMAGEAPGS P	8697
RSN-0064	GSAPGGTGGYAELR MVGPIATSGSETPGT	8453	RSC-0064	GTAEAASASGGGTG GYAELRMVGPIPGSP	8698
RSN-0065	GSAPGEAGGYAELR MLGSAATSGSETPG T	8454	RSC-0065	GTAEAASASGGEAG GYAELRMLGSAPGS P	8699
RSN-0066	GSAPGPGGGYAELR MIGTIATSGSETPGT	8455	RSC-0066	GTAEAASASGGPGG GYAELRMIGTIPGSP	8700
RSN-0067	GSAPGSEGGYAELR MYGAIATSGSETPG T	8456	RSC-0067	GTAEAASASGGSEG GYAELRMYGAIPGS P	8701
RSN-0068	GSAPGTPGGYAELR MSGGAATSGSETPG T	8457	RSC-0068	GTAEAASASGGTPG GYAELRMSGGAPGS P	8702
RSN-0069	GSAPGASGGYAELR MNGEIATSGSETPGT	8458	RSC-0069	GTAEAASASGGASG GYAELRMNGEIPGSP	8703
RSN-0070	GSAPGGTGGYAELR MQGPAATSGSETPG T	8459	RSC-0070	GTAEAASASGGGTG GYAELRMQGPAPGS P	8704
RSN-0071	GSAPGANHTPAGLT GPGARATSGSETPG T	8460	RSC-0071	GTAEAASASGGANH TPAGLTGPGARPGSP	8705
RSN-0072	GSAPGANTAPEGLT GPSTRATSGSETPGT	8461	RSC-0072	GTAEAASASGGANT APEGLTGPSTRPGSP	8706
RSN-0073	GSAPGTGAPPGGLT	8462	RSC-0073	GTAEAASASGGTGA	8707

	GPGTRATSGSETPGT			PPGGLTGPGTRPGSP	
RSN-0074	GSAPGANHEPSGLTEGSPRATSGSETPGT	8463	RSC-0074	GTAEAASASGGANHEPSGLTEGSPRPGSP	8708
RSN-0075	GSAPGANTEPPELGAGTERATSGSETPGT	8464	RSC-0075	GTAEAASASGGANTEPPELGAGTERPGSP	8709
RSN-0076	GSAPGASGPPPGLTGPPGRATSGSETPGT	8465	RSC-0076	GTAEAASASGGASGPPPGLTGPPGRPGSP	8710
RSN-0077	GSAPGASGTPAPLGGEPGRATSGSETPGT	8466	RSC-0077	GTAEAASASGGASGTPAPLGGEPGRPGSP	8711
RSN-0078	GSAPGPAGPPEGLETEAGRATSGSETPGT	8467	RSC-0078	GTAEAASASGGPAGPPEGLETEAGRPGSP	8712
RSN-0079	GSAPGPTSGQGGLTGPESRATSGSETPGT	8468	RSC-0079	GTAEAASASGGPTSGQGGLTGPESRPGSP	8713
RSN-0080	GSAPGSAGGAANLVRGGAIATSGSETPGT	8469	RSC-0080	GTAEAASASGGSAGGAANLVRGGAIPGSP	8714
RSN-0081	GSAPGTGGGAAPLVRGGGAATSGSETPGT	8470	RSC-0081	GTAEAASASGGTGGGAAPLVRGGGAPGSP	8715
RSN-0082	GSAPGAEGGAAALVRGGEIATSGSETPGT	8471	RSC-0082	GTAEAASASGGAEGGAAALVRGGEIPGSP	8716
RSN-0083	GSAPGGPGGAALLVRGGPAATSGSETPGT	8472	RSC-0083	GTAEAASASGGGPGGAALLVRGGPAPGSP	8717
RSN-0084	GSAPGEAGGAAFLVRGGSIATSGSETPGT	8473	RSC-0084	GTAEAASASGGEAGGAAFLVRGGSIPGSP	8718
RSN-0085	GSAPGPGGGAASLVRGGTAATSGSETPGT	8474	RSC-0085	GTAEAASASGGPGGGAASLVRGGTAPGSP	8719
RSN-0086	GSAPGSEGGAATLVRGGAIATSGSETPGT	8475	RSC-0086	GTAEAASASGGSEGGAATLVRGGAIPGSP	8720
RSN-0087	GSAPGTPGGAAGLVRGGGAATSGSETPGT	8476	RSC-0087	GTAEAASASGGTPGGAAGLVRGGGAPGSP	8721
RSN-0088	GSAPGASGGAADLVRGGEIATSGSETPGT	8477	RSC-0088	GTAEAASASGGASGGAADLVRGGEIPGSP	8722
RSN-0089	GSAPGGTGGAGNLVRGGPAATSGSETPGT	8478	RSC-0089	GTAEAASASGGGTGGAGNLVRGGPAPGSP	8723
RSN-0090	GSAPGEAGGAPNLVRGGSIATSGSETPGT	8479	RSC-0090	GTAEAASASGGEAGGAPNLVRGGSIPGSP	8724
RSN-0091	GSAPGPGGGAVNLVRGGTAATSGSETPGT	8480	RSC-0091	GTAEAASASGGPGGGAVNLVRGGTAPGSP	8725
RSN-0092	GSAPGSEGGALNLVRGGAIATSGSETPGT	8481	RSC-0092	GTAEAASASGGSEGGALNLVRGGAIPGSP	8726
RSN-0093	GSAPGTPGGASNLVRGGGAATSGSETPGT	8482	RSC-0093	GTAEAASASGGTPGGASNLVRGGGAPGSP	8727

RSN-0094	GSAPGASGGATNLVRGGEIATSGSETPGT	8483	RSC-0094	GTAEAASASGGASGGATNLVRGGEIPGSP	8728
RSN-0095	GSAPGGTGGAQNLVRGGPAATSGSETPGT	8484	RSC-0095	GTAEAASASGGGTGGAQNLVRGGPAPGSP	8729
RSN-0096	GSAPGEAGGAENLVRGGSIATSGSETPGT	8485	RSC-0096	GTAEAASASGGEAGGAENLVRGGSIPGSP	8730
RSN-1517	GSAPEAGRSANHEPLGLVATATSGSETPGT	8486	RSC-1517	GTAEAASASGEAGRSANHEPLGLVATPGSP	8731
BSRS-A1	GSAPASGRSTNAGPSGLAGPATSGSETPGT	8487	BSRS-A1	GTAEAASASGASGRSTNAGPSGLAGPPGSP	8732
BSRS-A2	GSAPASGRSTNAGPQGLAGQATSGSETPGT	8488	BSRS-A2	GTAEAASASGASGRSTNAGPQGLAGQPGSP	8733
BSRS-A3	GSAPASGRSTNAGPPGLTGPATSGSETPGT	8489	BSRS-A3	GTAEAASASGASGRSTNAGPPGLTGPPGSP	8734
VP-1	GSAPASSRGTNAGPAGLTGPATSGSETPGT	8490	VP-1	GTAEAASASGASSRGTNAGPAGLTGPPGSP	8735
RSN-1752	GSAPASSRTTNTGPSTLTGPATSGSETPGT	8491	RSC-1752	GTAEAASASGASSRTTNTGPSTLTGPPGSP	8736
RSN-1512	GSAPAAGRSDNGTPLELVAPATSGSETPGT	8492	RSC-1512	GTAEAASASGAAGRSDNGTPLELVAPPGSP	8737
RSN-1517	GSAPEAGRSANHEPLGLVATATSGSETPGT	8486	RSC-1517	GTAEAASASGEAGRSANHEPLGLVATPGSP	8731
VP-2	GSAPASGRGTNAGPAGLTGPATSGSETPGT	8493	VP-2	GTAEAASASGASGRGTNAGPAGLTGPPGSP	8738
RSN-1018	GSAPLFGRNDNHEPLELGGGATSGSETPGT	8494	RSC-1018	GTAEAASASGLFGRNDNHEPLELGGGPGSP	8739
RSN-1053	GSAPTAGRSDNLEPLGLVFGATSGSETPGT	8495	RSC-1053	GTAEAASASGTAGRSDNLEPLGLVFGPGSP	8740
RSN-1059	GSAPLDGRSDNFHPPELVAGATSGSETPGT	8496	RSC-1059	GTAEAASASGLDGRSDNFHPPELVAGPGSP	8741
RSN-1065	GSAPLEGRSDNEEPENLVAGATSGSETPGT	8497	RSC-1065	GTAEAASASGLEGRSDNEEPENLVAGPGSP	8742
RSN-1167	GSAPLKGRSDNNAPLALVAGATSGSETPGT	8498	RSC-1167	GTAEAASASGLKGRSDNNAPLALVAGPGSP	8743
RSN-1201	GSAPVYSRGTNAGPHGLTGRATSGSETP	8499	RSC-1201	GTAEAASASGVYSRGTNAGPHGLTGRPG	8744

	GT			SP	
RSN-1218	GSAPANSRGTNKGFAGLIGPATSGSETPGT	8500	RSC-1218	GTAEAASASGANSRGTNKGFAGLIGPPGSP	8745
RSN-1226	GSAPASSRLTNEAPAGLTIPATSGSETPGT	8501	RSC-1226	GTAEAASASGASSRLTNEAPAGLTIPPGSP	8746
RSN-1254	GSAPDQSRGTNAGPEGLTDPATSGSETPGT	8502	RSC-1254	GTAEAASASGDQSRGTNAGPEGLTDPPGSP	8747
RSN-1256	GSAPESSRGTNIGQGGLTGPATSGSETPGT	8503	RSC-1256	GTAEAASASGESSRGTNIGQGGLTGPPGSP	8748
RSN-1261	GSAPSSSRGTNQDPAGLTIPATSGSETPGT	8504	RSC-1261	GTAEAASASGSSSRGTNQDPAGLTIPPGSP	8749
RSN-1293	GSAPASSRGQNHSPMGLTGPATSGSETPGT	8505	RSC-1293	GTAEAASASGASSRGQNHSPMGLTGPPGSP	8750
RSN-1309	GSAPAYSRGPNAGPAGLEGRATSGSETPGT	8506	RSC-1309	GTAEAASASGAYSRGPNAGPAGLEGRPGSP	8751
RSN-1326	GSAPASERGNNAGPANLTGFATSGSETPGT	8507	RSC-1326	GTAEAASASGASERGNNAGPANLTGFPGSP	8752
RSN-1345	GSAPASHRGTNPKPAILTGPATSGSETPGT	8508	RSC-1345	GTAEAASASGASHRGTNPKPAILTGPPGSP	8753
RSN-1354	GSAPMSSRRTNANPAQLTGPATSGSETPGT	8509	RSC-1354	GTAEAASASGMSSRRTNANPAQLTGPPGSP	8754
RSN-1426	GSAPGAGRTDNHEPLELGAAATSGSETPGT	8510	RSC-1426	GTAEAASASGGAGRTDNHEPLELGAAPGSP	8755
RSN-1478	GSAPLAGRSENTAPLELTAGATSGSETPGT	8511	RSC-1478	GTAEAASASGLAGRSENTAPLELTAGPGSP	8756
RSN-1479	GSAPLEGRPDNHEPLALVASATSGSETPGT	8512	RSC-1479	GTAEAASASGLEGRPDNHEPLALVASPGSP	8757
RSN-1496	GSAPLSGRSDNEEPLALPAGATSGSETPGT	8513	RSC-1496	GTAEAASASGLSGRSDNEEPLALPAGPGSP	8758
RSN-1508	GSAPEAGRTDNHEPLELSAPATSGSETPGT	8514	RSC-1508	GTAEAASASGEAGRTDNHEPLELSAPPGSP	8759
RSN-1513	GSAPEGGRSDNHGPLELVSGATSGSETPGT	8515	RSC-1513	GTAEAASASGEGGRSDNHGPLELVSGPGSP	8760
RSN-1516	GSAPLSGRSDNEAPLELEAGATSGSETPG	8516	RSC-1516	GTAEAASASGLSGRSDNEAPLELEAGPGS	8761

	T			P	
RSN-1524	GSAPLGGRADNHEPPELGAGATSGSETPGT	8517	RSC-1524	GTAEAASASGLGGRADNHEPPELGAGPGSP	8762
RSN-1622	GSAPPPSRGTNAEPAGLTGEATSGSETPGT	8518	RSC-1622	GTAEAASASGPPSRGTNAEPAGLTGEPGSP	8763
RSN-1629	GSAPASTRGENAGPAGLEAPATSGSETPGT	8519	RSC-1629	GTAEAASASGASTRGENAGPAGLEAPPGSP	8764
RSN-1664	GSAPESSRGTNGAPEGLTGPATSGSETPGT	8520	RSC-1664	GTAEAASASGESSRGTNGAPEGLTGPPGSP	8765
RSN-1667	GSAPASSRATNESPAGLTGEATSGSETPGT	8521	RSC-1667	GTAEAASASGASSRATNESPAGLTGEPGSP	8766
RSN-1709	GSAPASSRGENPPPGGLTGPATSGSETPGT	8522	RSC-1709	GTAEAASASGASSRGENPPPGGLTGPPGSP	8767
RSN-1712	GSAPAASRGTNTGPAELTGSATSGSETPGT	8523	RSC-1712	GTAEAASASGAASRGTNTGPAELTGSPGSP	8768
RSN-1727	GSAPAGSRTTNAGPGGLEGPATSGSETPGT	8524	RSC-1727	GTAEAASASGAGSRTTNAGPGGLEGPPGSP	8769
RSN-1754	GSAPAPSRGENAGPATLTGAATSGSETPGT	8525	RSC-1754	GTAEAASASGAPSRGENAGPATLTGAPGSP	8770
RSN-1819	GSAPESGRAANTGPPTLTAPATSGSETPGT	8526	RSC-1819	GTAEAASASGESGRAANTGPPTLTAPPGSP	8771
RSN-1832	GSAPNPGRAANEGPPGLPGSATSGSETPGT	8527	RSC-1832	GTAEAASASGNPGRAANEGPPGLPGSPGSP	8772
RSN-1855	GSAPESSRAANLTPPELTGPATSGSETPGT	8528	RSC-1855	GTAEAASASGESSRAANLTPPELTGPPGSP	8773
RSN-1911	GSAPASGRAANETPPGLTGAATSGSETPGT	8529	RSC-1911	GTAEAASASGASGRAANETPPGLTGAPGSP	8774
RSN-1929	GSAPNSGRGENLGAPGLTGTATSGSETPGT	8530	RSC-1929	GTAEAASASGNSGRGENLGAPGLTGTPGSP	8775
RSN-1951	GSAPTTGRAANLTPAGLTGPATSGSETPGT	8531	RSC-1951	GTAEAASASGTTGRAANLTPAGLTGPPGSP	8776
RSN-2295	GSAPEAGRSANHTPAGLTGPATSGSETPGT	8532	RSC-2295	GTAEAASASGEAGRSANHTPAGLTGPPGSP	8777
RSN-2298	GSAPESGRAANTTPAGLTGPATSGSETPGT	8533	RSC-2298	GTAEAASASGESGRAANTTPAGLTGPPGSP	8778

RSN-2038	GSAPTTGRATEAANLTPAGLTGPATSGSETPGT	8534	RSC-2038	GTAEAASASGTTGRATEAANLTPAGLTGPPGSP	8779
RSN-2072	GSAPTTGRAEEAANLTPAGLTGPATSGSETPGT	8535	RSC-2072	GTAEAASASGTTGRAEEAANLTPAGLTGPPGSP	8780
RSN-2089	GSAPTTGRAGEAANLTPAGLTGPATSGSETPGT	8536	RSC-2089	GTAEAASASGTTGRAGEAANLTPAGLTGPPGSP	8781
RSN-2302	GSAPTTGRATEAANATPAGLTGPATSGSETPGT	8537	RSC-2302	GTAEAASASGTTGRATEAANATPAGLTGPPGSP	8782
RSN-3047	GSAPTTGRAGEAEGATSAGATGPATSGSETPGT	8538	RSC-3047	GTAEAASASGTTGRAGEAEGATSAGATGPPGSP	8783
RSN-3052	GSAPTTGEAGEAANATSAGATGPATSGSETPGT	8539	RSC-3052	GTAEAASASGTTGEAGEAANATSAGATGPPGSP	8784
RSN-3043	GSAPTTGEAGEAAGLTPAGLTGPATSGSETPGT	8540	RSC-3043	GTAEAASASGTTGEAGEAAGLTPAGLTGPPGSP	8785
RSN-3041	GSAPTTGAAGEAANATPAGLTGPATSGSETPGT	8541	RSC-3041	GTAEAASASGTTGAAGEAANATPAGLTGPPGSP	8786
RSN-3044	GSAPTTGRAGEAAGLTPAGLTGPATSGSETPGT	8542	RSC-3044	GTAEAASASGTTGRAGEAAGLTPAGLTGPPGSP	8787
RSN-3057	GSAPTTGRAGEAANATSAGATGPATSGSETPGT	8543	RSC-3057	GTAEAASASGTTGRAGEAANATSAGATGPPGSP	8788
RSN-3058	GSAPTTGEAGEAAGATSAGATGPATSGSETPGT	8544	RSC-3058	GTAEAASASGTTGEAGEAAGATSAGATGPPGSP	8789
RSN-2485	GSAPESGRAANTEPPELGAGATSGSETPGT	8545	RSC-2485	GTAEAASASGESGRAANTEPPELGAGPGSP	8790
RSN-2486	GSAPESGRAANTAPEGLTGPATSGSETPGT	8546	RSC-2486	GTAEAASASGESGRAANTAPEGLTGPPGSP	8791
RSN-2488	GSAPEPGRAANHEPSGLTEGATSGSETPGT	8547	RSC-2488	GTAEAASASGEPGRAANHEPSGLTEGPGSP	8792
RSN-2599	GSAPESGRAANHTGAPPGGLTGPATSGSETPGT	8548	RSC-2599	GTAEAASASGESGRAANHTGAPPGGLTGPPGSP	8793
RSN-2706	GSAPTTGRTGEGANATPGGLTGPATSGSETPGT	8549	RSC-2706	GTAEAASASGTTGRTGEGANATPGGLTGPPGSP	8794
RSN-2707	GSAPRTGRSGEAANETPEGLEGPATSGSETPGT	8550	RSC-2707	GTAEAASASGRTGRSGEAANETPEGLEGPPGSP	8795

RSN-2708	GSAPRTGRTGESANETPAGLGGPATSGSETPGT	8551	RSC-2708	GTAEAASASGRTGRTGESANETPAGLGGPPGSP	8796
RSN-2709	GSAPSTGRTGEPANETPAGLSGPATSGSETPGT	8552	RSC-2709	GTAEAASASGSTGRTGEPANETPAGLSGPPGSP	8797
RSN-2710	GSAPTTGRAGEPANATPTGLSGPATSGSETPGT	8553	RSC-2710	GTAEAASASGTTGRAGEPANATPTGLSGPPGSP	8798
RSN-2711	GSAPRTGRPGEGANATPTGLPGPATSGSETPGT	8554	RSC-2711	GTAEAASASGRTGRPGEGANATPTGLPGPPGSP	8799
RSN-2712	GSAPRTGRGGEAANATPSGLGGPATSGSETPGT	8555	RSC-2712	GTAEAASASGRTGRGGEAANATPSGLGGPPGSP	8800
RSN-2713	GSAPSTGRSGESANATPGGLGGPATSGSETPGT	8556	RSC-2713	GTAEAASASGSTGRSGESANATPGGLGGPPGSP	8801
RSN-2714	GSAPRTGRTGEEANATPAGLPGPATSGSETPGT	8557	RSC-2714	GTAEAASASGRTGRTGEEANATPAGLPGPPGSP	8802
RSN-2715	GSAPATGRPGEPANTTPEGLEGPATSGSETPGT	8558	RSC-2715	GTAEAASASGATGRPGEPANTTPEGLEGPPGSP	8803
RSN-2716	GSAPSTGRSGEPANATPGGLTGPATSGSETPGT	8559	RSC-2716	GTAEAASASGSTGRSGEPANATPGGLTGPPGSP	8804
RSN-2717	GSAPPTGRGGEGANTTPTGLPGPATSGSETPGT	8560	RSC-2717	GTAEAASASGPTGRGGEGANTTPTGLPGPPGSP	8805
RSN-2718	GSAPPTGRSGEGANATPSGLTGPATSGSETPGT	8561	RSC-2718	GTAEAASASGPTGRSGEGANATPSGLTGPPGSP	8806
RSN-2719	GSAPTTGRASEGANSTPAPLTEPATSGSETPGT	8562	RSC-2719	GTAEAASASGTTGRASEGANSTPAPLTEPPGSP	8807
RSN-2720	GSAPTYGRAAEAANTTPAGLTAPATSGSETPGT	8563	RSC-2720	GTAEAASASGTYGRAAEAANTTPAGLTAPPGSP	8808
RSN-2721	GSAPTTGRATEGANATPAELTEPATSGSETPGT	8564	RSC-2721	GTAEAASASGTTGRATEGANATPAELTEPPGSP	8809
RSN-2722	GSAPTVGRASEEANTTPASLTGPATSGSETPGT	8565	RSC-2722	GTAEAASASGTVGRASEEANTTPASLTGPPGSP	8810
RSN-2723	GSAPTTGRAPEAANATPAPLTGPATSGSETPGT	8566	RSC-2723	GTAEAASASGTTGRAPEAANATPAPLTGPPGSP	8811
RSN-2724	GSAPTWGRATEPANATPAPLTSPATSGSETPGT	8567	RSC-2724	GTAEAASASGTWGRATEPANATPAPLTSPPGSP	8812

RSN-2725	GSAPTVGRASESANATPAELTSPATSGSETPGT	8568	RSC-2725	GTAEAASASGTVGRASESANATPAELTSPPGSP	8813
RSN-2726	GSAPTVGRAPEGANSTPAGLTGPATSGSETPGT	8569	RSC-2726	GTAEAASASGTVGRAPEGANSTPAGLTGPPGSP	8814
RSN-2727	GSAPTWGRATEAPNLEPATLTTPATSGSETPGT	8570	RSC-2727	GTAEAASASGTWGRATEAPNLEPATLTTPPGSP	8815
RSN-2728	GSAPTTGRATEAPNLTPAPLTEPATSGSETPGT	8571	RSC-2728	GTAEAASASGTTGRATEAPNLTPAPLTEPPGSP	8816
RSN-2729	GSAPTQGRATEAPNLSPAALTSPATSGSETPGT	8572	RSC-2729	GTAEAASASGTQGRATEAPNLSPAALTSPPGSP	8817
RSN-2730	GSAPTQGRAAEAPNLTPATLTAPATSGSETPGT	8573	RSC-2730	GTAEAASASGTQGRAAEAPNLTPATLTAPPGSP	8818
RSN-2731	GSAPTSGRAPEATNLAPAPLTGPATSGSETPGT	8574	RSC-2731	GTAEAASASGTSGRAPEATNLAPAPLTGPPGSP	8819
RSN-2732	GSAPTQGRAAEAANLTPAGLTEPATSGSETPGT	8575	RSC-2732	GTAEAASASGTQGRAAEAANLTPAGLTEPPGSP	8820
RSN-2733	GSAPTTGRAGSAPNLPPTGLTTPATSGSETPGT	8576	RSC-2733	GTAEAASASGTTGRAGSAPNLPPTGLTTPPGSP	8821
RSN-2734	GSAPTTGRAGGAENLPPEGLTAPATSGSETPGT	8577	RSC-2734	GTAEAASASGTTGRAGGAENLPPEGLTAPPGSP	8822
RSN-2735	GSAPTTSRAGTATNLTPEGLTAPATSGSETPGT	8578	RSC-2735	GTAEAASASGTTSRAGTATNLTPEGLTAPPGSP	8823
RSN-2736	GSAPTTGRAGTATNLPPSGLTTPATSGSETPGT	8579	RSC-2736	GTAEAASASGTTGRAGTATNLPPSGLTTPPGSP	8824
RSN-2737	GSAPTTARAGEAENLSPSGLTAPATSGSETPGT	8580	RSC-2737	GTAEAASASGTTARAGEAENLSPSGLTAPPGSP	8825
RSN-2738	GSAPTTGRAGGAGNLAPGGLTEPATSGSETPGT	8581	RSC-2738	GTAEAASASGTTGRAGGAGNLAPGGLTEPPGSP	8826
RSN-2739	GSAPTTGRAGTATNLPPEGLTGPATSGSETPGT	8582	RSC-2739	GTAEAASASGTTGRAGTATNLPPEGLTGPPGSP	8827
RSN-2740	GSAPTTGRAGGAANLAPTGLTEPATSGSETPGT	8583	RSC-2740	GTAEAASASGTTGRAGGAANLAPTGLTEPPGSP	8828
RSN-2741	GSAPTTGRAGTAENLAPSGLTTPATSGSETPGT	8584	RSC-2741	GTAEAASASGTTGRAGTAENLAPSGLTTPPGSP	8829

RSN-2742	GSAPTTGRAGSATNLGPGGLTGPATSGSETPGT	8585	RSC-2742	GTAEAASASGTTGRAGSATNLGPGGLTGPPGSP	8830
RSN-2743	GSAPTTARAGGAENLTPAGLTEPATSGSETPGT	8586	RSC-2743	GTAEAASASGTTARAGGAENLTPAGLTEPPGSP	8831
RSN-2744	GSAPTTARAGSAENLSPSGLTGPATSGSETPGT	8587	RSC-2744	GTAEAASASGTTARAGSAENLSPSGLTGPPGSP	8832
RSN-2745	GSAPTTARAGGAGNLAPEGLTTPATSGSETPGT	8588	RSC-2745	GTAEAASASGTTARAGGAGNLAPEGLTTPPGSP	8833
RSN-2746	GSAPTTSRAGAAENLTPTGLTGPATSGSETPGT	8589	RSC-2746	GTAEAASASGTTSRAGAAENLTPTGLTGPPGSP	8834
RSN-2747	GSAPTYGRTTTPGNEPPASLEAEATSGSETPGT	8590	RSC-2747	GTAEAASASGTYGRTTTPGNEPPASLEAEPGSP	8835
RSN-2748	GSAPTYSRGESGPNEPPPGLTGPATSGSETPGT	8591	RSC-2748	GTAEAASASGTYSRGESGPNEPPPGLTGPPGSP	8836
RSN-2749	GSAPAWGRTGASENETPAPLGGEATSGSETPGT	8592	RSC-2749	GTAEAASASGAWGRTGASENETPAPLGGEPGSP	8837
RSN-2750	GSAPRWGRAETTPNTPPEGLETEATSGSETPGT	8593	RSC-2750	GTAEAASASGRWGRAETTPNTPPEGLETEPGSP	8838
RSN-2751	GSAPESGRAANHTGAEPPELGAGATSGSETPGT	8594	RSC-2751	GTAEAASASGESGRAANHTGAEPPELGAGPGSP	8839
RSN-2754	GSAPTTGRAGEAANLTPAGLTESATSGSETPGT	8595	RSC-2754	GTAEAASASGTTGRAGEAANLTPAGLTESPGSP	8840
RSN-2755	GSAPTTGRAGEAANLTPAALTESATSGSETPGT	8596	RSC-2755	GTAEAASASGTTGRAGEAANLTPAALTESPGSP	8841
RSN-2756	GSAPTTGRAGEAANLTPAPLTESATSGSETPGT	8597	RSC-2756	GTAEAASASGTTGRAGEAANLTPAPLTESPGSP	8842
RSN-2757	GSAPTTGRAGEAANLTPEPLTESATSGSETPGT	8598	RSC-2757	GTAEAASASGTTGRAGEAANLTPEPLTESPGSP	8843
RSN-2758	GSAPTTGRAGEAANLTPAGLTGAATSGSETPGT	8599	RSC-2758	GTAEAASASGTTGRAGEAANLTPAGLTGAPGSP	8844
RSN-2759	GSAPTTGRAGEAANLTPEGLTGAATSGSETPGT	8600	RSC-2759	GTAEAASASGTTGRAGEAANLTPEGLTGAPGSP	8845
RSN-2760	GSAPTTGRAGEAANLTPEPLTGAATSGSETPGT	8601	RSC-2760	GTAEAASASGTTGRAGEAANLTPEPLTGAPGSP	8846

RSN-2761	GSAPTTGRAGEAANLTPAGLTEAATSGSETPGT	8602	RSC-2761	GTAEAASASGTTGRAGEAANLTPAGLTEAPGSP	8847
RSN-2762	GSAPTTGRAGEAANLTPEGLTEAATSGSETPGT	8603	RSC-2762	GTAEAASASGTTGRAGEAANLTPEGLTEAPGSP	8848
RSN-2763	GSAPTTGRAGEAANLTPAPLTEAATSGSETPGT	8604	RSC-2763	GTAEAASASGTTGRAGEAANLTPAPLTEAPGSP	8849
RSN-2764	GSAPTTGRAGEAANLTPEPLTEAATSGSETPGT	8605	RSC-2764	GTAEAASASGTTGRAGEAANLTPEPLTEAPGSP	8850
RSN-2765	GSAPTTGRAGEAANLTPEPLTGPATSGSETPGT	8606	RSC-2765	GTAEAASASGTTGRAGEAANLTPEPLTGPPGSP	8851
RSN-2766	GSAPTTGRAGEAANLTPAGLTGGATSGSETPGT	8607	RSC-2766	GTAEAASASGTTGRAGEAANLTPAGLTGGPGSP	8852
RSN-2767	GSAPTTGRAGEAANLTPEGLTGGATSGSETPGT	8608	RSC-2767	GTAEAASASGTTGRAGEAANLTPEGLTGGPGSP	8853
RSN-2768	GSAPTTGRAGEAANLTPEALTGGATSGSETPGT	8609	RSC-2768	GTAEAASASGTTGRAGEAANLTPEALTGGPGSP	8854
RSN-2769	GSAPTTGRAGEAANLTPEPLTGGATSGSETPGT	8610	RSC-2769	GTAEAASASGTTGRAGEAANLTPEPLTGGPGSP	8855
RSN-2770	GSAPTTGRAGEAANLTPAGLTEGATSGSETPGT	8611	RSC-2770	GTAEAASASGTTGRAGEAANLTPAGLTEGPGSP	8856
RSN-2771	GSAPTTGRAGEAANLTPEGLTEGATSGSETPGT	8612	RSC-2771	GTAEAASASGTTGRAGEAANLTPEGLTEGPGSP	8857
RSN-2772	GSAPTTGRAGEAANLTPAPLTEGATSGSETPGT	8613	RSC-2772	GTAEAASASGTTGRAGEAANLTPAPLTEGPGSP	8858
RSN-2773	GSAPTTGRAGEAANLTPEPLTEGATSGSETPGT	8614	RSC-2773	GTAEAASASGTTGRAGEAANLTPEPLTEGPGSP	8859
RSN-3047	GSAPTTGRAGEAEGATSAGATGPATSGSETPGT	8538	RSC-3047	GTAEAASASGTTGRAGEAEGATSAGATGPPGSP	8783
RSN-2783	GSAPEAGRSAEATSAGATGPATSGSETPGT	8615	RSC-2783	GTAEAASASGEAGRSAEATSAGATGPPGSP	8860
RSN-3107	GSAPSASGTYSRGESGPGSPATSGSETPGT	8616	RSC-3107	GTAEAASASGSASGTYSRGESGPGSPPGSP	8861
RSN-3103	GSAPSASGEAGRTDTHPGSPATSGSETPGT	8617	RSC-3103	GTAEAASASGSASGEAGRTDTHPGSPPGSP	8862

RSN-3102	GSAPSASGEPGRAAEHPGSPATSGSETPGT	8618	RSC-3102	GTAEAASASGSASGEPGRAAEHPGSPPGSP	8863
RSN-3119	GSAPSPAGESSRGTTIAGSPATSGSETPGT	8619	RSC-3119	GTAEAASASGSPAGESSRGTTIAGSPPGSP	8864
RSN-3043	GSAPTTGEAGEAAGLTPAGLTGPATSGSETPGT	8540	RSC-3043	GTAEAASASGTTGEAGEAAGLTPAGLTGPPGSP	8785
RSN-2789	GSAPEAGESAGATPAGLTGPATSGSETPGT	8620	RSC-2789	GTAEAASASGEAGESAGATPAGLTGPPGSP	8865
RSN-3109	GSAPSASGAPLELEAGPGSPATSGSETPGT	8621	RSC-3109	GTAEAASASGSASGAPLELEAGPGSPPGSP	8866
RSN-3110	GSAPSASGEPPELGAGPGSPATSGSETPGT	8622	RSC-3110	GTAEAASASGSASGEPPELGAGPGSPPGSP	8867
RSN-3111	GSAPSASGEPSGLTEGPGSPATSGSETPGT	8623	RSC-3111	GTAEAASASGSASGEPSGLTEGPGSPPGSP	8868
RSN-3112	GSAPSASGTPAPLTEPPGSPATSGSETPGT	8624	RSC-3112	GTAEAASASGSASGTPAPLTEPPGSPPGSP	8869
RSN-3113	GSAPSASGTPAELTEPPGSPATSGSETPGT	8625	RSC-3113	GTAEAASASGSASGTPAELTEPPGSPPGSP	8870
RSN-3114	GSAPSASGPPPGLTGPPGSPATSGSETPGT	8626	RSC-3114	GTAEAASASGSASGPPPGLTGPPGSPPGSP	8871
RSN-3115	GSAPSASGTPAPLGGEPGSPATSGSETPGT	8627	RSC-3115	GTAEAASASGSASGTPAPLGGEPGSPPGSP	8872
RSN-3125	GSAPSPAGAPEGLTGPAGSPATSGSETPGT	8628	RSC-3125	GTAEAASASGSPAGAPEGLTGPAGSPPGSP	8873
RSN-3126	GSAPSPAGPPEGLETEAGSPATSGSETPGT	8629	RSC-3126	GTAEAASASGSPAGPPEGLETEAGSPPGSP	8874
RSN-3127	GSAPSPTSGQGGLTGPGSEPATSGSETPGT	8630	RSC-3127	GTAEAASASGSPTSGQGGLTGPGSEPPGSP	8875
RSN-3131	GSAPSESAPPEGLETESTEPATSGSETPGT	8631	RSC-3131	GTAEAASASGSESAPPEGLETESTEPPGSP	8876
RSN-3132	GSAPSEGSEPLELGAASETPATSGSETPGT	8632	RSC-3132	GTAEAASASGSEGSEPLELGAASETPPGSP	8877
RSN-3133	GSAPSEGSGPAGLEAPSETPATSGSETPGT	8633	RSC-3133	GTAEAASASGSEGSGPAGLEAPSETPPGSP	8878
RSN-3138	GSAPSEPTPPASLEAEPGSPATSGSETPGT	8634	RSC-3138	GTAEAASASGSEPTPPASLEAEPGSPPGSP	8879

[250] В другом аспекте RS для включения в рекомбинантные полипептиды человека или животных могут быть разработаны с избирательной чувствительностью, чтобы они характеризовались разными скоростями расщепления и разными значениями

эффективности расщепления различными протеазами, для которых они являются субстратами. Поскольку данная протеаза может быть обнаружена в различных концентрациях в пораженных тканях, в том числе без ограничения в опухоли, раковых клетках крови или воспаленной ткани или очаге воспаления, по сравнению со здоровыми тканями или кровотоком, в настоящем изобретении предусмотрены RS, которые имели индивидуальные аминокислотные последовательности, сконструированные так, чтобы они характеризовались более высокой или более низкой эффективностью расщепления для данной протеазы, чтобы обеспечить преимущественное превращение рекомбинантного полипептида из формы пролекарства в активную форму (т. е. путем отделения и высвобождения связывающих фрагментов и XTEN из рекомбинантного полипептида после расщепления RS), когда он находится вблизи клетки- или ткани-мишени и совместно локализованных с ней протеаз, по сравнению со скоростью расщепления RS в здоровой ткани или кровотоке, где высвобождающиеся связывающие фрагменты антитела обладают большей способностью связываться с лигандами в пораженных тканях по сравнению с формой пролекарства, которая остается в кровотоке. Посредством таких избирательных конструкций терапевтический индекс получаемых композиций может быть улучшен, что приведет к уменьшению побочных эффектов по сравнению с традиционными терапевтическими препаратами, которые не предусматривают такую сайт-специфическую активацию.

[251] Как используется в данном документе, эффективность расщепления определяется как значение $\log_2$ соотношения процентной доли расщепленного тестируемого субстрата, содержащего RS, и процентной доли расщепленного контрольного субстрата AC1611, когда каждый из них подвергается действию фермента протеазы в биохимических анализах (более подробно см. в примерах), в которых проводят реакцию, где начальная концентрация субстрата составляет 6 мкМ, реакционные смеси инкубируют при 37°C в течение 2 часов перед остановкой путем добавления EDTA, при этом количество продуктов расщепления и нерасщепленного субстрата анализируют посредством SDS-PAGE в невосстанавливающих условиях, чтобы установить соотношение процентных долей расщепления. Эффективность расщепления рассчитывают следующим образом: $Log_2\left(\frac{\%\text{ расщепленного субстрата, представляющего интерес}}{\%\text{ расщепленного }AC1611\ i\text{в том же эксперементе}}\right)$. Таким образом, эффективность расщепления -1 означает, что количество расщепленного тестируемого субстрата составляло 50% по сравнению с таким количеством контрольного субстрата, тогда как эффективность расщепления +1 означает, что количество расщепленного тестируемого субстрата составляло 200% по сравнению с таким количеством контрольного субстрата. Более высокая скорость расщепления тестируемой протеазой по сравнению с контролем приведет к более высокой эффективности расщепления, а более низкая скорость расщепления тестируемой протеазой по сравнению

с контролем приведет к более низкой эффективности расщепления. Как подробно описано в примерах, было установлено, что контрольная последовательность RS AC1611 (RSR-1517), имеющая аминокислотную последовательность EAGRSANHEPLGLVAT (SEQ ID NO: 8261), характеризуется надлежащей базовой эффективностью расщепления протеазами легумаином, MMP-2, MMP-7, MMP-9, MMP-14, uPA и матриптазой при тестировании в биохимических анализах in vitro для определения скоростей расщепления отдельными протеазами. Посредством избирательной замены аминокислот в отдельных местоположениях в пептидах RS создавали библиотеки RS и оценивали против панели из 7 протеаз (что более подробно описано в примерах), в результате чего получали профили, которые использовали для установления методических рекомендаций для соответствующих аминокислотных замен с целью достижения RS с необходимыми значениями эффективности расщепления. При получении RS с необходимыми значениями эффективности расщепления предпочтительными являются замены с использованием гидрофильных аминокислот A, E, G, P, S и T, однако могут быть заменены другие L-аминокислоты в заданных положениях в целях корректировки эффективности расщепления при условии, что RS сохраняет по меньшей мере некоторую восприимчивость к расщеплению протеазой. Консервативные замены аминокислот в пептиде для сохранения или воздействия на активность находятся в пределах знаний и возможностей специалиста в данной области. В одном варианте осуществления настоящее изобретение предусматривает RS, при этом RS расщепляется протеазой, выбранной из легумаина, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA или матриптазы, с в по меньшей мере 0,2 $\log_2$, или 0,4 $\log_2$, или 0,8 $\log_2$, или 1,0 $\log_2$ более высокой эффективностью расщепления в биохимическом конкурентном анализе in vitro по сравнению с расщеплением той же протеазой контрольной последовательности RSR-1517, имеющей последовательность EAGRSANHEPLGLVAT (SEQ ID NO: 8261). В другом варианте осуществления настоящее изобретение предусматривает RS, при этом RS расщепляется протеазой, выбранной из легумаина, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA или матриптазы, с в по меньшей мере 0,2 $\log_2$, или 0,4 $\log_2$, или 0,8 $\log_2$, или 1,0 $\log_2$ более низкой эффективностью расщепления в биохимическом конкурентном анализе in vitro по сравнению с расщеплением той же протеазой контрольной последовательности RSR-1517, имеющей последовательность EAGRSANHEPLGLVAT (SEQ ID NO: 8261). В одном варианте осуществления настоящее изобретение предусматривает RS, при этом скорость расщепления RS протеазой, выбранной из легумаина, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA или матриптазы, в по меньшей мере 2 раза, или в по меньшей мере 4 раза, или в по меньшей

мере 8 раз, или в по меньшей мере 16 раз быстрее по сравнению с контрольной последовательностью RSR-1517, имеющей последовательность EAGRSANHEPLGLVAT (SEQ ID NO: 8261). В другом варианте осуществления настоящее изобретение предусматривает RS, при этом скорость расщепления RS протеазой, выбранной из легумаина, MMP-1, MMP-2, MMP-7, MMP-9, MMP-11, MMP-14, uPA или матриптазы, в по меньшей мере 2 раза, или в по меньшей мере 4 раза, или в по меньшей мере 8 раз, или в по меньшей мере 16 раз медленнее по сравнению с контрольной последовательностью RSR-1517, имеющей последовательность EAGRSANHEPLGLVAT (SEQ ID NO: 8261).

[252] В другом аспекте настоящее изобретение предусматривает AAC, содержащую несколько RS, где каждая последовательность RS выбрана из группы последовательностей, представленных в **таблице 8a**, и RS связаны друг с другом при помощи 1-6 аминокислот, выбранных из глицина, серина, аланина и треонина. В одном варианте осуществления AAC содержит первый RS и второй RS, отличающийся от первого RS, где каждая последовательность RS выбрана из группы последовательностей, представленных в **таблице 8a**, и RS связаны друг с другом при помощи 1-6 аминокислот, выбранных из глицина, серина, аланина и треонина. В другом варианте осуществления AAC содержит первый RS, второй RS, отличный от первого RS, и третий RS, отличный от первого и второго RS, где каждая последовательность выбрана из группы последовательностей, представленных в **таблице 8a**, и первый, и второй, и третий RS связаны друг с другом при помощи 1-6 аминокислот, выбранных из глицина, серина, аланина и треонина. В частности, предполагается, что несколько RS из AAC могут быть объединены с образованием последовательности, которая может быть расщеплена несколькими протеазами с различной скоростью или эффективностью расщепления. В другом варианте осуществления настоящее изобретение предусматривает AAC, содержащую RS1 и RS2, выбранные из группы последовательностей, представленных в **таблицах 8a-8b**, и XTEN 1 и XTEN 2, такие как те, которые описаны в данном документе выше или где-либо еще, где RS1 слит между XTEN1 и связывающими фрагментами, а RS2 слит между XTEN2 и связывающими фрагментами. Предполагается, что такие композиции будут легче расщепляться пораженными тканями-мишенями, которые экспрессируют множество протеаз, по сравнению со здоровыми тканями или в нормальной системе кровообращения, в результате чего полученные фрагменты, несущие связывающие фрагменты, будут легче проникать в ткань-мишень; например опухоль, и иметь повышенную способность связывать и соединять клетку-мишень и эффекторную клетку (или только клетку-мишень в случае AAC, сконструированной с одним связывающим фрагментом).

[253] RS по настоящему изобретению пригодны для включения в рекомбинантные полипептиды как терапевтические средства для лечения форм рака, аутоиммунных заболеваний, воспалительных заболеваний и других состояний, при которых необходима локальная активация рекомбинантного полипептида. Композиции для человека или

животного направлены на удовлетворение нереализованных потребностей и превосходят их по одному или более аспектам, включая увеличенный конечный период полувыведения, направленную доставку и улучшенное терапевтическое соотношение с пониженной токсичностью для здоровых тканей по сравнению с традиционными терапевтическими средствами на основе антител или терапевтическими средствами на основе биспецифических антител, которые активны при введении.

[254] В некоторых вариантах осуществления (слитый) полипептид содержит первый сегмент для обеспечения высвобождения (RS1), расположенный между (первым) XTEN и биологически активным полипептидом. В некоторых вариантах осуществления полипептид дополнительно содержит второй сегмент для обеспечения высвобождения (RS2), расположенный между биологически активным полипептидом и вторым XTEN. В некоторых вариантах осуществления RS1 и RS2 являются идентичными по последовательности. В некоторых вариантах осуществления RS1 и RS2 являются неидентичными по последовательности. В некоторых вариантах осуществления RS1 содержит аминокислотную последовательность, характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности по отношению к последовательности, представленной в **таблицах 8a-8b**. В некоторых вариантах осуществления RS2 содержит аминокислотную последовательность, характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности по отношению к последовательности, представленной в **таблицах 8a-8b**. В некоторых вариантах осуществления каждый из RS1 и RS2 представляет собой субстрат для расщепления несколькими протеазами в одном, двух или трех сайтах расщепления в каждой последовательности сегмента для обеспечения высвобождения.

[255] Эталонный фрагмент

[256] В некоторых вариантах осуществления (слитый) полипептид дополнительно содержит один или более эталонных фрагментов, высвобождаемых из полипептида при расщеплении протеазой. В некоторых вариантах осуществления каждый из одного или более эталонных фрагментов содержит часть биологически активного полипептида. В некоторых вариантах осуществления один или более эталонных фрагментов представляют собой один эталонный фрагмент, который отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после расщепления полипептида протеазой.

СМЕСЬ ПОЛИПЕПТИДОВ

[257] Раскрытое в данном документе предусматривает смесь, содержащую множество полипептидов различной длины, где смесь предусматривает первый набор полипептидов и второй набор полипептидов. В некоторых вариантах осуществления каждый полипептид из первого набора полипептидов содержит фрагмент со штрих-кодом, который (a) высвобождается из полипептида при расщеплении протеазой и (b) имеет последовательность и молекулярную массу, которые отличаются от последовательности и

молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов. В некоторых вариантах осуществления во втором наборе полипептидов отсутствует фрагмент со штрих-кодом первого набора полипептидов. В некоторых вариантах осуществления как первый набор полипептидов, так и второй набор полипептидов каждый содержит эталонный фрагмент, который (a) является общим для первого набора полипептидов и второго набора полипептидов и (b) может высвобождаться при расщеплении протеазой. В некоторых вариантах осуществления отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,70. В некоторых вариантах осуществления отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8, 0,9, 0,95 или 0,98. В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде из первого набора полипептидов и второго набора полипептидов. В некоторых вариантах осуществления протеаза представляет собой протеазу, которая расщепляет с С-концевой стороны от остатков глутаминовой кислоты. В некоторых вариантах осуществления протеаза представляет собой протеазу Glu-C. В некоторых вариантах осуществления протеаза не является трипсином. В некоторых вариантах осуществления полипептиды с различными значениями длины предусматривают полипептиды, содержащие по меньшей мере один удлиненный рекомбинантный полипептид (XTEN), такой как любой, описанный в данном документе выше или описанный где-либо еще в этом документе. В некоторых вариантах осуществления первый набор полипептидов включает полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида. В некоторых вариантах осуществления полноразмерный полипептид представляет собой (слитый) полипептид, такой как любой из описанных в данном документе выше или описанных где-либо еще в данном документе. В некоторых вариантах осуществления фрагмент со штрих-кодом лишен (не содержит) как N-концевой аминокислоты, так и С-концевой аминокислоты полноразмерного полипептида. В некоторых вариантах осуществления в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-, так и С-концевого усечения полноразмерного полипептида. В некоторых вариантах осуществления первый набор полипептидов и второй набор полипептидов могут отличаться по одному или более фармакологическим свойствам. Неограничивающие иллюстративные свойства включают.

СПОСОБ ОПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК ПОЛИПЕПТИДОВ

[258] Раскрытое в данном документе предусматривает способ оценки в смеси, содержащей полипептиды различной длины, относительного количества первого набора полипептидов в смеси ко второму набору полипептидов в смеси, где (1) каждый полипептид из первого набора полипептидов имеет общий фрагмент со штрих-кодом, который встречается один и только один раз в полипептиде, и (2) каждый полипептид из второго набора полипептидов не имеет фрагмента со штрих-кодом, который является

общим для полипептидов из первого набора, где каждый из отдельных полипептидов как из первого набора полипептидов, так и из второго набора полипептидов, содержит эталонный фрагмент. Способ может предусматривать приведение смеси в контакт с протеазой для получения множества протеолитических фрагментов, полученных в результате расщепления первого набора полипептидов и второго набора полипептидов, где множество протеолитических фрагментов включает множество эталонных фрагментов и множество фрагментов со штрих-кодом. Способ может дополнительно предусматривать определение отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов с оценкой тем самым относительного количества первого набора полипептидов ко второму набору полипептидов. В некоторых вариантах осуществления фрагмент со штрих-кодом встречается не более одного раза в каждом полипептиде первого набора полипептидов. В некоторых вариантах осуществления эталонный фрагмент встречается не более одного раза в каждом полипептиде из первого набора полипептидов и второго набора полипептидов. В некоторых вариантах осуществления множество протеолитических фрагментов включает множество эталонных фрагментов и множество фрагментов со штрих-кодом. В некоторых вариантах осуществления протеаза расщепляет первый и второй наборы полипептидов (или полипептиды различной длины) с С-концевой стороны от остатков глутаминовой кислоты, за которыми не следует остаток пролина. В некоторых вариантах осуществления протеаза представляет собой протеазу Glu-C. В некоторых вариантах осуществления протеаза не является трипсином. В некоторых вариантах осуществления стадия определения отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов включает количественную оценку фрагментов со штрих-кодом и эталонных фрагментов в смеси после ее контакта с протеазой. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют на основе их соответствующих масс. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью масс-спектрометрии. В некоторых вариантах осуществления фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью жидкостной хроматографии-масс-спектрометрии (LC-MS). В некоторых вариантах осуществления стадия определения отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает изобарическое мечение. В некоторых вариантах осуществления стадия определения отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает введение в смесь одного или обоих из меченного изотопом эталонного фрагмента и меченного изотопом фрагмента со штрих-кодом. В некоторых вариантах осуществления полипептиды с различными значениями длины предусматривают полипептиды, которые содержат по меньшей мере один удлиненный рекомбинантный полипептид (XTEN), который описан в данном документе выше или описан где-либо еще в данном документе. В некоторых вариантах осуществления XTEN отличается тем, что (i) содержит по меньшей мере 150 аминокислот; (ii) по меньшей мере 90% аминокислотных остатков XTEN выбраны из глицина (G), аланина (A), серина (S),

треонина (T), глутаминовой кислоты (E) и пролина (P); и (iii) он содержит по меньшей мере 4 разных типа аминокислот, выбранных из G, A, S, T, E и P. В некоторых вариантах осуществления фрагмент со штрих-кодом, если он присутствует, является частью XTEN. В некоторых вариантах осуществления смесь полипептидов с различными значениями длины предусматривает полипептид, как любой из описанных в данном документе выше или описанных где-либо еще в данном документе. В некоторых вариантах осуществления полипептиды различной длины предусматривают полноразмерный полипептид и его усеченные фрагменты. В некоторых вариантах осуществления полипептиды различной длины состоят фактически из полноразмерного полипептида и его усеченных фрагментов. В некоторых вариантах осуществления в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, C-концевого усечения или как N-, так и C-концевого усечения полноразмерного полипептида. В некоторых вариантах осуществления полноразмерный полипептид представляет собой полипептид, который описан в данном документе выше или описан где-либо еще в данном документе. В некоторых вариантах осуществления отношение количества фрагментов со штрих-кодом к эталонным фрагментам составляет более 0,5, 0,6, 0,7, 0,8, 0,9, 0,95, 0,98 или 0,99.

Количественное определение пептидов на основе изобарического мечения

[259] В некоторых вариантах осуществления изобарическое мечение можно использовать для определения отношения фрагментов со штрих-кодом к эталонным фрагментам. Специалисту будет понятно, что изобарическое мечение представляет собой стратегию на основе масс-спектрометрии, используемую в количественной протеомике, где пептиды или белки (или их части) метят различными химическими группами, которые являются изобарическими (одинаковыми по массе), но различаются по распределению тяжелых изотопов по своей структуре. Эти метки, обычно называемые тандемными масс-метками, сконструированы таким образом, что массовая метка расщепляется в определенной линкерной области при высокоэнергетической диссоциации, индуцированной столкновением (CID) во время тандемной масс-спектрометрии, давая тем самым репортерные ионы разной массы. Специалисту с обычными навыками будет понятно, что одними из наиболее распространенных изобарических меток являются метки, реакционноспособные в отношении аминов.

[260] С помощью расширенных возможностей обнаружения и количественной оценки продуктов усечения (например, с помощью изобарического мечения) можно получать знания, которые способны помочь в разработке производственных процессов с включением стадий очистки, чтобы свести к минимуму присутствие нежелательных вариантов в очищенном лекарственном веществе/продукте.

ПОЛУЧЕНИЕ РЕКОМБИНАТНЫХ ЕДИНИЦ

[261] Настоящее изобретение в данном документе включает нуклеиновую кислоту. Нуклеиновая кислота может предусматривать полинуклеотид (или полинуклеотидную последовательность), кодирующий (слитый) полипептид, такой как любой из описанных в данном документе выше или описанных где-либо еще в данном документе; или

нуклеиновая кислота может предусматривать обратно-комплементарную последовательность такого полинуклеотида (или полинуклеотидной последовательности).

[262] Настоящее изобретение в данном документе включает вектор экспрессии, который содержит полинуклеотидную последовательность, такую как любая из описанных в предыдущем абзаце, и регуляторную последовательность, функционально связанную с полинуклеотидной последовательностью.

[263] Настоящее изобретение в данном документе включает клетку-хозяина, содержащую вектор экспрессии, такой как описанный в предыдущем абзаце. В некоторых вариантах осуществления клетка-хозяин является прокариотической. В некоторых вариантах осуществления клетка-хозяин представляет собой E. coli. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку млекопитающего.

[264] В другом аспекте в настоящем изобретении предусмотрены способы изготовления композиций для человека или животного. В одном варианте осуществления способ включает культивирование клетки-хозяина, содержащей конструкцию нуклеиновой кислоты, которая кодирует полипептид или содержащую XTEN композицию по любому из вариантов осуществления, описанных в данном документе, в условиях, которые способствуют экспрессии полипептида или слитого полипептида BPXTEN, с последующим извлечением полипептида или слитого полипептида BPXTEN с использованием стандартных способов очистки (например, колоночной хроматографии, HPLC и т. п.), где композицию извлекают, при этом по меньшей мере 70%, или по меньшей мере 80%, или по меньшей мере 90%, или по меньшей мере 95%, или по меньшей мере 97%, или по меньшей мере 99% связывающих фрагментов экспрессируемых полипептида или слитого полипептида BPXTEN являются правильно свернутыми. В другом варианте осуществления способа получения извлекают экспрессируемые полипептид или слитый полипептид BPXTEN, в которых по меньшей мере 90%, или по меньшей мере 95%, или по меньшей мере 97%, или по меньшей мере 99% полипептида или слитого полипептида BPXTEN извлекают в мономерной растворимой форме.

[265] В другом аспекте настоящее изобретение относится к способам получения полипептида и слитого полипептида BPXTEN при высоких уровнях экспрессии функционального белка в процессе ферментации с использованием E. coli или клетки-хозяина, являющейся клеткой млекопитающего, а также получения векторов экспрессии, кодирующих конструкции, применимые в способах получения цитотоксически активных композиций на основе конструкций полипептидов при высоких уровнях экспрессии. В одном варианте осуществления способ включает стадии 1) получения полинуклеотида, кодирующего полипептиды из любого из вариантов осуществления, раскрытых в данном документе, 2) клонирования полинуклеотида в вектор экспрессии, который может представлять собой плазмиду или другой вектор под контролем подходящих последовательностей транскрипции и трансляции для экспрессии белка в биологической системе на высоком уровне, 3) трансформации подходящей клетки-хозяина вектором

экспрессии и 4) культивирования клетки-хозяина в традиционных питательных средах в условиях, подходящих для экспрессии композиции на основе полипептида. Там, где это необходимо, клетка-хозяин представляет собой E. coli. Благодаря данному способу экспрессия полипептида приводит к титрам ферментации, составляющим по меньшей мере 0,05 г/л, или по меньшей мере 0,1 г/л, или по меньшей мере 0,2 г/л, или по меньшей мере 0,3 г/л, или по меньшей мере 0,5 г/л, или по меньшей мере 0,6 г/л, или по меньшей мере 0,7 г/л, или по меньшей мере 0,8 г/л, или по меньшей мере 0,9 г/л, или по меньшей мере 1 г/л, или по меньшей мере 2 г/л, или по меньшей мере 3 г/л, или по меньшей мере 4 г/л, или по меньшей мере 5 г/л продукта экспрессии клетки-хозяина, и при этом по меньшей мере 70%, или по меньшей мере 80%, или по меньшей мере 90%, или по меньшей мере 95%, или по меньшей мере 97%, или по меньшей мере 99% экспрессируемого белка являются правильно свернутыми. Используемый в данном документе термин «правильно свернутый» означает, что компоненты композиции, представляющие собой антигенсвязывающие фрагменты, обладают способностью специфично связывать свой лиганд-мишень. В другом варианте осуществления в настоящем изобретении предусмотрен способ получения полипептида или слитого полипептида BPXTEN, при этом способ включает культивирование в ферментационной реакционной смеси клетки-хозяина, которая содержит вектор, кодирующий полипептид, включающий в себя полипептид или слитый полипептид BPXTEN, в условиях, эффективных для экспрессии полипептидного продукта при концентрации, составляющей более чем приблизительно 10 миллиграммов/грамм сухой массы клетки-хозяина (мг/г), или по меньшей мере приблизительно 250 мг/г, или приблизительно 300 мг/г, или приблизительно 350 мг/г, или приблизительно 400 мг/г, или приблизительно 450 мг/г, или приблизительно 500 мг/г указанного полипептида, при достижении ферментационной реакционной смесью оптической плотности, составляющей по меньшей мере 130 при длине волны 600 нм, и где антигенсвязывающие фрагменты экспрессируемого белка являются правильно свернутыми. В другом варианте осуществления в настоящем изобретении предусмотрен способ получения полипептида или слитого полипептида BPXTEN, при этом способ включает культивирование в ферментационной реакционной смеси клетки-хозяина, которая содержит вектор, кодирующий композицию, в условиях, эффективных для экспрессии полипептидного продукта при концентрации, составляющей более чем приблизительно 10 миллиграммов/грамм сухой массы клетки-хозяина (мг/г), или по меньшей мере приблизительно 250 мг/г, или приблизительно 300 мг/г, или приблизительно 350 мг/г, или приблизительно 400 мг/г, или приблизительно 450 мг/г, или приблизительно 500 мг/г указанного полипептида, при достижении ферментационной реакционной смесью оптической плотности, составляющей по меньшей мере 130 при длине волны 600 нм, и где экспрессируемый полипептидный продукт является растворимым.

ФАРМАЦЕВТИЧЕСКАЯ КОМПОЗИЦИЯ

[266] Раскрытое в данном документе предусматривает фармацевтическую

композицию, содержащую полипептид BPXTEN, такой как любой из описанных в данном документе выше или описанных где-либо еще в данном документе, и одно или более фармацевтически приемлемых вспомогательных веществ. В некоторых вариантах осуществления фармацевтическая композиция составлена для внутрикожного, подкожного, внутривенного, внутриартериального, внутрибрюшного, внутрибрюшинного, интравитреального, интратекального или внутримышечного введения. В некоторых вариантах осуществления фармацевтическая композиция представлена в жидкой форме. В некоторых вариантах осуществления фармацевтическая композиция находится в устройстве, которое имплантируют в глаз или другую часть тела. В некоторых вариантах осуществления фармацевтическая композиция находится в предварительно заполненном шприце для однократной инъекции. В некоторых вариантах осуществления фармацевтическая композиция составлена в виде лиофилизированного порошка, подлежащего восстановлению перед введением.

[267] В некоторых вариантах осуществления дозу вводят внутрикожно, подкожно, внутривенно, интравитреально (или вводят в глаз иным образом), внутриартериально, интраабдоминально, внутрибрюшинно, интратекально или внутримышечно. В некоторых вариантах осуществления фармацевтическую композицию вводят с помощью устройства, имплантированного в глаз или другую часть тела. В некоторых вариантах осуществления человек или животное представляют собой мышь, крысу, обезьяну или человека.

[268] Фармацевтические композиции можно вводить для терапии любым подходящим путем. Кроме того, фармацевтические композиции могут также содержать другие фармацевтически активные соединения или множество соединений по настоящему изобретению.

[269] В некоторых вариантах осуществления фармацевтическую композицию можно вводить в терапевтически эффективной дозе. В некоторых случаях вышеизложенного терапевтически эффективная доза приводит к выигрышу во времени, проведенном в пределах терапевтического окна для слитого белка, по сравнению с соответствующим BP из слитого белка, не связанным со слитым белком и вводимым в сравнимой дозе человеку или животному.

[270] В другом варианте осуществления настоящее изобретение предусматривает способ лечения заболевания, нарушения или состояния, включающий введение фармацевтической композиции, описанной выше, человеку или животному с использованием нескольких последовательных доз фармацевтической композиции, вводимых с использованием терапевтически эффективного режима введения дозы.

[271] Полипептиды BPXTEN по настоящему изобретению могут быть составлены в соответствии с известными способами для получения фармацевтически полезных композиций, при этом полипептид комбинируют в смеси с фармацевтически приемлемой средой-носителем, таким как водные растворы или буферы, фармацевтически приемлемые суспензии и эмульсии. Терапевтические составы готовят для хранения путем смешивания активного ингредиента, имеющего необходимую степень чистоты, с необязательными

физиологически приемлемыми носителями, вспомогательными веществами или стабилизаторами, как описано в Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980).

ФАРМАЦЕВТИЧЕСКИЕ НАБОРЫ

[272] В другом аспекте настоящее изобретение предусматривает набор для облегчения применения полипептидов BPXTEN. В одном варианте осуществления набор содержит, по меньшей мере, в первом контейнере: (a) некоторое количество композиции на основе слитого белка BPXTEN, достаточное для лечения заболевания, состояния или нарушения при введении человеку или животному, нуждающемуся в этом; и (b) некоторое количество фармацевтически приемлемого носителя; вместе в составе, готовом для инъекции или для восстановления стерильной водой, буфером или декстрозой; вместе с этикеткой, идентифицирующей лекарственное средство на основе BPXTEN и условия хранения и обработки, и с листом утвержденных показаний для лекарственного средства, инструкциями по восстановлению и/или введению лекарственного средства на основе BPXTEN для применения с целью предупреждения и/или лечения утвержденного показания, соответствующей дозой и информацией по безопасности, а также информацией, идентифицирующей партию и срок годности лекарственного средства. В другом варианте осуществления вышеизложенного набор может содержать второй контейнер, который может содержать подходящий разбавитель для композиции на основе BPXTEN, который обеспечивает пользователя соответствующей концентрацией BPXTEN для доставки человеку или животному.

СПОСОБ ЛЕЧЕНИЯ

[273] Раскрытое в данном документе предусматривает применение полипептида, такого как любой из описанных в данном документе выше или описанных где-либо еще в данном документе, при получении лекарственного препарата для лечения заболевания у человека или животного. В некоторых вариантах осуществления конкретное заболевание, подлежащее лечению, зависит от выбора биологически активных белков. В некоторых вариантах осуществления заболевание представляет собой рак.

[274] Раскрытое в данном документе предусматривает способ лечения заболевания у человека или животного, при этом способ предусматривает введение человеку или животному, нуждающемуся в этом, одной или более терапевтически эффективных доз фармацевтической композиции, такой как любая из описанных в данном документе выше или описанных где-либо еще в данном документе. В некоторых вариантах осуществления заболевание представляет собой рак. В некоторых вариантах осуществления фармацевтическую композицию вводят человеку или животному в виде одной или более терапевтически эффективных доз, вводимых в соответствии с режимом дозирования. В некоторых вариантах осуществления человек или животное представляют собой мышь, крысу, обезьяну или человека.

[275] Далее приведены примеры композиций и оцениваний композиций по настоящему изобретению. Следует понимать, что различные другие варианты

осуществления могут быть реализованы на практике с учетом общего описания, приведенного выше.

ПРИМЕРЫ

Пример 1. Разработка XTEN со штрих-кодом при помощи минимальных мутаций универсального XTEN

[276] Данный пример иллюстрирует типичный подход к разработке полипептида XTEN со штрих-кодом путем внесения минимальной мутации(-й) аминокислотной последовательности универсального полипептида XTEN (такого как из **таблицы 3b** в данном документе выше). Соответствующие критерии для выполнения минимальной(-ых) мутации(-й) включают один или более из следующего: (a) свести к минимуму изменение последовательности соответствующего полипептида XTEN; (b) свести к минимуму изменение аминокислотного состава в соответствующем полипептиде XTEN; (c) фактически поддерживать суммарный заряд в соответствующем полипептиде XTEN; (d) фактически поддерживать низкую иммуногенность соответствующего полипептида XTEN; (e) фактически поддерживать фармакокинетические свойства, обеспечиваемые полипептидом XTEN.

[277] Например, XTEN со штрих-кодом конструировали путем выполнения одной или более мутаций, включающих делецию остатка глутаминовой кислоты, вставку остатка глутаминовой кислоты, замену остатка глутаминовой кислоты или замену на остаток глутаминовой кислоты, или любую их комбинацию для универсальных XTEN в **таблице 9**.

Таблица 9. Четыре универсальных XTEN, используемые для проектирования полипептидов XTEN со штрих-кодом

SEQ ID NO:	**Название XTEN**	**Аминокислотная последовательность**
676	AE144	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP
686	AE288_1	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP
688	AE576	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE

		GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTE EGTSESATPESGPGTSTEPSEGSAP
690	AE864	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTE EGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTS ESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPG SEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP

Пример 2. Анализ последовательности полипептидов XTEN со штрих-кодом и их отбор для слияния с биологически активным полипептидом («BP»)

[278] Этот пример иллюстрирует конструкцию и выбор полипептида XTEN со штрих-кодом (и сборку более чем одного XTEN со штрих-кодом в набор) для слияния с биологически активным полипептидом. В зависимости от расположения фрагмента(-ов) со штрих-кодом в XTEN и способа слияния полипептида(-ов) XTEN с биологически активным белком с образованием конструкции, содержащей полипептид XTEN (например, икстенилированный активируемый протеазой привлекающий Т-клетки активатор (XPAT)), фрагмент(-ы) со штрих-кодом может(-гут) указывать на усечение(-я) полипептида XTEN.

[279] Анализ in silico расщепления с помощью GluC проводили на двух иллюстративных полипептидах XTEN (XTEN864 и XTEN288_1) для количественного определения высвобождаемых пептидных фрагментов после полного расщепления с помощью GluC полипептида XTEN. В анализе in silico принимается во внимание то, что в отношении полипептида XTEN, содержащего последовательные остатки глутаминовой кислоты (например, «EE»), GluC может расщеплять после любого из остатков глутаминовой кислоты. Как показано в результатах, обобщенных ниже в **таблице 10**, 10-мерная пептидная последовательность «TPGTSTEPSE (SEQ ID NO: 8880)» и 14-мерная пептидная последовательность «GSAPGSEPATSGSE (SEQ ID NO: 8881)» встречаются один и только один раз в более длинном XTEN864, в то время как все остальные пептидные последовательности встречаются в XTEN864 два или более раз. И 14-мерная

пептидная последовательность «GSAPGSEPATSGSE (SEQ ID NO: 8881)» также встречается один и только один раз в более коротком XTEN288_1.

[280] Уникальность штрих-кода-кандидата оценивается по отношению ко всем другим пептидным фрагментам, высвобождаемым из конструкции, содержащей полипептид XTEN. Соответственно, последовательность штрих-кода в одном полипептиде XTEN не может встречаться где-либо еще в конструкции, содержащей полипептид XTEN, включая любой другой содержащийся в ней полипептид XTEN, любой содержащийся в ней биологически активный белок или любую связь между его соседними компонентами. Например, в **таблице 11** показана таблица «уникальности» пептидов для набора из двух полипептидов XTEN. Благодаря своему присутствию как в XTEN864, так и в XTEN288, 14-мерная пептидная последовательность «GSAPGSEPATSGSE (SEQ ID NO: 8881)» не уникальна для набора полипептидов XTEN, включающих как XTEN864, так и XTEN288, и, таким образом, не может использоваться в качестве штрих-кода для обнаружения усечений в полипептидных продуктах, которые содержат оба из двух полипептидов XTEN.

[281] Выбор штрих-кода (или набора штрих-кодов) может дополнительно включать в себя идентификацию и определение надлежащего местоположения(-й) или положения(-й) кандидвтного(-ых) штрих-кода(-ов) в полипептиде XTEN. Местоположение или положение кандидатного штрих-кода может быть ассоциировано с фармакологически значимой информацией о полипептиде XTEN (и конструкции, содержащей полипептид XTEN, в целом), такой как усечение полипептида XTEN за пределами критической длины и/или делеция(-и) в полипептиде XTEN. 10-мерный пептид «TPGTSTEPSE (SEQ ID NO: 8880)» может служить подходящим фрагментом со штрих-кодом, если XTEN864 будет помещен на N-конце продукта, содержащего полипептид XTEN, и если усечение 238 аминокислот с N-конца продукта существенно не влияет на фармакологические свойства продукта.

Таблица 10. Примерные последовательности XTEN для анализа расщепления с помощью GluC

Иллюстративный XTEN	Аминокислотная последовательность	SEQ ID NO:
XTEN864	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE**TPGTSTEPSE**GSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG	8882

	PGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGS PTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEE GTSTEPSEGSAPGTSTEPSE**GSAPGSEPATSGSE**TPGTSESATP ESGPGTSTEPSEGSAP	
AE288_1	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSE**GSAPGSEPATSGSE**TPGT SESATPESGPGTSTEPSEGSAP	8883

Таблица 11. Анализ «уникальности» пептидов

Пептидные фрагменты	SEQ ID NO:	XTEN AE288_1	XTEN864	Оба
SATPE	8884	9	23	32
SGPGSEPATSGSE	8885	4	9	13
GSAPGTSTEPSE	8886	1	10	11
TPGTSE	8887	4	7	11
SGPGTSTEPSE	8888	2	8	10
GSAPGTSE	8889	1	8	9
GTSTEPSE	8890	2	5	7
GSAPGSPAGSPTSTEE	8891	1	4	5
GSPAGSPTSTEE	186	1	4	5
SGPGTSE	8892	2	3	5
SGPGSPAGSPTSTEE	8893	1	3	4
TPGSEPATSGSE	8894	1	2	3
TPGSPAGSPTSTEE	8895	1	2	3
GSAPGSEPATSGSE	8881	1	1	2
TPGTSTEPSE	8880		1	1

Все подчеркнутые последовательности продуцируют уникальные пептиды GluC.

Зона, отличная от XTEN, подчеркнута и выделена курсивом

Пептиды штрих-кода выделены жирным шрифтом

[282] Иллюстративные пептидные последовательности штрих-кода продемонстрированы ниже в **таблице 12**. Эти последовательности штрих-кодов должны быть фланкированы в соответствии со структурной формулой (I):

AAA—Glu—пептид штрих-кода—BBB,

где «AAA» представляет собой Gly, Ala, Ser, Thr или Pro, а «BBB» представляет собой Gly, Ala, Ser или Thr, сконфигурированные для облегчения эффективного высвобождения пептида штрих-кода путем расщепления с помощью GluC. Примечательно, что вставка каждого пептида штрих-кода в XTEN может привести к дополнительным уникальным последовательностям, непосредственно предшествующим

или следующим за вставленными пептидами штрих-кода.

Таблица 12. Список подходящих пептидов штрих-кода

Пептид(-ы) каднидатного штрих-кода	SEQ ID NO:	Появление в XTEN864	Появление в XTEN288_1	Появление в XTEN1152
SPATSGSTPE	8020	0	0	
GSAPATSE	8021	0	0	0
GSAPGTATE	8022	0	0	0
GSAPGTE	8023	0	0	0
PATSGPTE	8024	0	0	0
SASPE	8025	0	0	0
PATSGSTE	8026	0	0	0
GSAPGTSAE	8027			
SATSGSE	8028	0	0	0
SGPGSTPAE	8029	0	0	
SGSE	8030	0		

Пример 3. Конструирование и отбор XTEN в конструкциях на основе икстенилированных полипептидов с полной последовательностью

[283] Этот пример иллюстрирует конструирование полной последовательности полипептидной конструкции, содержащей два полипептида XTEN, один на N-конце, а другой на С-конце.

[284] В приведенной ниже **таблице 13** показаны полипептиды XTEN, используемые в иллюстративном BPXTEN со штрих-кодом (содержащем полипептиды XTEN со штрих-кодом как на N-, так и на С-конце) и эталонном BPXTEN (содержащем универсальные XTEN как на N-, так и на С-конце). В иллюстративном BPXTEN со штрих-кодом полипептид XTEN со штрих-кодом (**SEQ ID NO: 8014**) слит с N-концом BP, а другой полипептид XTEN со штрих-кодом (**SEQ ID NO: 8015**) слит с С-концом BP. В эталонном BPXTEN полипептид XTEN «Ref-N» (**SEQ ID NO: 8896**) слит с N-концом BP, а полипептид XTEN «Ref-C» (**SEQ ID NO: 8897**) слит с С-концом BP. Полипептид XTEN «Ref-N» (**SEQ ID NO: 8896**) сопоставим по длине с полипептидом XTEN со штрих-кодом **SEQ ID NO: 8014**; а полипептид XTEN «Ref-C» (**SEQ ID NO:** 8897) сопоставим по длине с полипептидом XTEN со штрих-кодом **SEQ ID NO: 8015**. Каждый BPXTEN со штрих-кодом и эталонный BPXTEN содержит эталонную последовательность в компоненте, представляющем собой BP. Эталонная последовательность уникальна и отличается по молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из соответствующего BPXTEN при полном расщеплении протеазой GluC (например, согласно **примеру 5**). Уникальность эталонной последовательности оценивают по отношению ко всем другим пептидным фрагментам, высвобождаемым из конструкции на основе BPXTEN.

Таблица 13. Примерные наборы N- и С-концевых XTEN, используемые в полноразмерных конструкциях на основе BPXTEN.

SEQ ID NO:	**Тип XTEN**	**Аминокислотная последовательность**	**Общее количество аминокислот**
8014 (из **таблицы 3а**)	N-концевой XTEN	SPAGSPTSTESGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES**GPGSTPAESGSE**TPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPA	292
8015 (из **таблицы 3а**)	C-концевой XTEN	PGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSE**GSAPGTE**STPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGEPEA	582
8896 Ref-N	N-концевой XTEN	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPA	292

		GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPG SEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGGSAP	
8897 Ref-C	C-концевой XTEN	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGSEPAT SGSETPGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETP GTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSESATPESGPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPT STEEGTSESATPESGPGTSTEPSEGSAPGAAEPE A	584

Пример 4. Рекомбинантное конструирование и получение икстенилированных слитых полипептидов со штрих-кодом

[285] Примеры 4a-4b иллюстрируют рекомбинантное конструирование, получение и очистку полноразмерных полипептидов, содержащих полипептид(-ы) XTEN со штрих-кодом, с использованием раскрытых в данном документе способов.

Пример 4a. Икстенилированный слитый полипептид, содержащий XTEN со штрих-кодом на С-конце

[286] ЭКСПРЕССИЯ. Конструкции, кодирующие икстенилированный слитый полипептид, содержащий одноцепочечный вариабельный фрагмент (scFv) к EpCAM и последовательность XTEN со штрих-кодом длиной 864 аминокислоты (**SEQ ID NO: 8008**) на С-конце, экспрессируют в запатентованном штамме AmE098 E. coli и распределяют в периплазму с использованием N-концевой секреторной лидерной последовательности (MKKNIAFLLASMFVFSIATNAYA-) (SEQ ID NO: 8898), которая расщепляется во время транслокации. Ферментационные культуры выращивают на комплексной среде, не содержащей компонентов животного происхождения, при 37°C; и температуру сдвигают

до 26°C до начала истощения запасов фосфатов. В ходе сбора материала цельный ферментационный бульон центрифугируют для осаждения клеток. При сборе материала регистрируют общий объем и сырую массу клеток (WCW; соотношение осадка и надосадочной жидкости), и осажденные клетки собирают и замораживают при -80°C.

[287] ИЗВЛЕЧЕНИЕ. Осадок замороженных клеток ресуспендируют в буфере для лизиса (17,7 мМ лимонной кислоты, 22,3 мМ Na_2 HPO_4 , 75 мМ NaCl, 2 мМ EDTA, pH 4,0), ориентируясь на 30% сырой массы клеток. Ресуспензии дают уравновеситься при pH 4, затем гомогенизируют в два прохода при 800±50 бар, при этом температуру на выходе контролируют и поддерживают на уровне 15±5°C. Подтверждают, что pH гомогената находится в пределах указанного диапазона (pH 4,0±0,2).

[288] ОСВЕТЛЕНИЕ. Чтобы снизить примеси эндотоксина и клеток-хозяев, гомогенат подвергают низкотемпературной (10±5°C) кислотной (pH 4,0±0,2) флокуляции в течение ночи (15-20 часов). Чтобы удалить нерастворимую фракцию, флокулированный гомогенат центрифугируют в течение 40 минут при 16900 об/мин при 2-8°C, а надосадочную жидкость сохраняют. Надосадочную жидкость разбавляют водой Milli-Q (MQ) примерно в 3 раза, затем доводят до 7±1 мСм/см с помощью 5 М NaCl. Для удаления нуклеиновой кислоты, липидов и эндотоксина и для использования в качестве вспомогательного фильтрующего средства надосадочную жидкость доводят до 0,1% (м/м) диатомовой земли. Чтобы поддерживать фильтрующую добавку во взвешенном состоянии, надосадочную жидкость перемешивают лопастной мешалкой и оставляют для уравновешивания в течение 30 минут. Комплект фильтров, состоящий из глубинного фильтра, за которым следует фильтр 0,22 мкм, собирают, а затем промывают с помощью MQ. Надосадочную жидкость прокачивают через комплект фильтров, модулируя поток для поддержания перепада давления на уровне 25±5 фунтов на квадратный дюйм. Чтобы довести составную буферную систему (на основе соотношения лимонной кислоты и Na_2 HPO_4) до необходимого диапазона для хроматографии с захватом, фильтрат доводят до 500 мМ Na_2 HPO_4 , где конечное соотношение Na_2 HPO_4 к лимонной кислоте составляет 9,33:1, а pH буферного фильтрата поддерживают в пределах указанного диапазона (pH 7,0±0,2).

[289] ОЧИСТКА

[290] Захват с AEX. Для отделения димеров, агрегатов и больших усечений от мономерного продукта, а также для удаления эндотоксина и нуклеиновых кислот применяют анионообменную (AEX) хроматографию для захвата электроотрицательного C-концевого домена XTEN. В данном случае применяют неподвижную фазу AEX1 (GE Q Sepharose FF), подвижную фазу A для AEX1 (12,2 мМ Na_2 HPO_4 , 7,8 мМ NaH_2 PO_4 , 40 мМ NaCl) и подвижную фазу B для AEX1 (12,2 мМ Na_2 HPO_4 , 7,8 мМ NaH_2 PO_4 , 500 мМ NaCl). Колонку уравновешивают подвижной фазой A для AEX1. На основании концентрации общего белка, измеренной с помощью анализа с бицинхониновой кислотой (BCA), фильтрат загружают в колонку с целевой концентрацией смолы 28±4 г/л, вытесняют подвижной фазой A для AEX1, затем промывают с шагом до 30% B.

Связанный материал элюируют с градиентом от 30% B до 60% B при 20 CV. Фракции собирают аликвотами по 1 CV в то время как A220 выше (локального) исходного уровня на ≥ 100 произвольные мЕд оптической плотности. Фракции элюирования анализируют и объединяют на основе SDS-PAGE и SE-HPLC.

[291] Промежуточная очистка IMAC. Для обеспечения целостности C-конца применяют аффинную хроматографию с иммобилизованными металлами (IMAC) для захвата C-концевой полигистидиновой метки (His(6) (SEQ ID NO: 8031)). В данном случае применяют неподвижную фазу для IMAC (GE IMAC Sepharose FF), подвижную фазу A для IMAC (18,3 мМ Na_2 HPO_4 , 1,7 мМ NaH_2 PO_4 , 500 мМ NaCl, 1 мМ имидазола) и подвижную фазу B для IMAC (18,3 мМ Na_2 HPO_4 , 1,7 мМ NaH_2 PO_4 , 500 мМ NaCl, 500 мМ имидазола). Колонку заполняют раствором цинка и уравновешивают подвижной фазой A для IMAC. Пул AEX1 доводят до pH 7,8±0,1, 50±5 мСм/см (с 5 M NaCl) и 1 мМ имидазола, загружают в колонку IMAC с целевой концентрацией смолы 2 г/л и вытесняют подвижной фазой A для IMAC до тех пор, пока поглощение при 280 нм (A280) не возвращается к (локальному) исходному уровню. Связанный материал элюируют с шагом до 25% подвижной фазы B для IMAC. Сбор элюата IMAC начинают, когда A280 выше (локального) исходного уровня на ≥ 10 произвольных мЕд оптической плотности, направляют в контейнер, предварительно обогащенный EDTA, достаточным для доведения 2 CV до 2 мМ EDTA, и прекращают после того, как собрано 2 CV. Элюат анализируют с помощью SDS-PAGE.

[292] Промежуточная очистка с белком L. Для обеспечения целостности N-конца белок L применяют для захвата каппа-доменов, находящихся вблизи N-конца молекулы BPXTEN (в частности, scFv к EpCAM). В данном случае применяют неподвижную фазу для белка L (GE Capto L), подвижную фазу A для белка L (16,0 мМ лимонной кислоты, 20,0 мМ Na_2 HPO_4 , pH 4,0±0,1), подвижную фазу B для белка L (29,0 мМ лимонной кислоты, 7,0 мМ Na_2 HPO_4 , pH 2,60±0,02) и подвижную фазу C для белка L (3,5 мМ лимонной кислоты, 32,5 мМ Na_2 HPO_4 , 250 мМ NaCl, pH 7,0±0,1). Колонку уравновешивают подвижной фазой C для белка L. Значение pH элюата IMAC доводят до pH 7,0±0,1 и 30±3 мСм/см (с 5 M NaCl и MQ) и загружают в колонку с белком L с целевой концентрацией смолы 2 г/л, затем вытесняют подвижной фазой C для белка L до тех пор, пока поглощение при 280 нм (A280) не возвращается к (локальному) исходному уровню. Колонку промывают подвижной фазой A для белка L, а подвижные фазы A и B для белка L используют для элюирования при низком pH. Связанный материал элюируют примерно при pH 3,0 и собирают в контейнер, предварительно обогащенный одной частью 0,5 M Na_2 HPO_4 на каждые 10 частей собранного объема. Фракции анализируют с помощью SDS-PAGE.

[293] Очистка с HIC. Для разделения N-концевых вариантов (4 остатка на абсолютном N-конце не являются существенными для связывания белка L) и общих конформационных вариантов применяют хроматографию гидрофобного взаимодействия (HIC). В данном случае применяют неподвижную фазу для HIC (GE Capto Phenyl ImpRes),

подвижную фазу A для HIC (20 мМ гистидина, 0,02% (мас./об.) полисорбата 80, pH 6,5±0,1) и подвижную фазу B для HIC (1 М аммония сульфата, 20 мМ гистидина, 0,02% (мас./об.) полисорбата 80, pH 6,5±0,1). Колонку уравновешивают подвижной фазой B для HIC. Скорректированный элюат после белка L загружают в колонку HIC с целевой концентрацией смолы 2 г/л и вытесняют подвижной фазой B для HIC до тех пор, пока поглощение при 280 нм (A280) не вернется к (локальному) исходному уровню. Колонку промывают 50% B. Связанный материал элюируют градиентом от 50% B до 0% B при 75 CV. Фракции собирают аликвотами по 1 CV в то время как A280 выше (локального) исходного уровня на ≥ 3 произвольные мЕд оптической плотности. Фракции элюирования анализируют и объединяют на основе SE-HPLC и HI-HPLC.

[294] СОСТАВЛЕНИЕ. Для замены продукта в буфере для состава и доведения продукта до целевой концентрации (0,5 г/л) снова применяют обмен анионов для захвата C-концевого XTEN. В данном случае применяют неподвижную фазу для AEX2 (GE Q Sepharose FF), подвижную фазу A для AEX2 (20 мМ гистидина, 40 мМ NaCl, 0,02% (мас./об.) полисорбата 80, pH 6,5±0,2), подвижную фазу B для AEX2 (20 мМ гистидина, 1 М NaCl, 0,02% (мас./об.) полисорбата 80, pH 6,5±0,2) и подвижную фазу C для AEX2 (12,2 мМ Na_2 HPO_4 , 7,8 мМ NaH_2 PO_4 , 40 мМ NaCl, 0,02% (мас./об.) полисорбата 80, pH 7,0±0,2). Колонку уравновешивают подвижной фазой C для AEX2. Пул HIC доводят до pH 7,0±0,1 и 7±1 мСм/см (с MQ) и загружают в колонку AEX2 с целевой концентрацией смолы 2 г/л, затем вытесняют подвижной фазой C для AEX2 до тех пор, пока A280 не вернется к (локальному) исходному уровню. Колонку промывают подвижной фазой A для AEX2 (20 мМ гистидина, 40 мМ NaCl, 0,02% (мас./об.) полисорбата 80, pH 6,5±0,2). Подвижные фазы A и B для AEX2 применяют для создания стадии с {NaCl} и эффективного элюирования. Связанный материал элюируют с шагом до 38% подвижной фазы B для AEX2. Сбор элюата AEX2 начинают, когда A280 выше (локального) исходного уровня на ≥ 5 произвольных мЕд оптической плотности, и прекращают после того, как собраны 2 объема колонки. Элюат AEX2 фильтруют через фильтр 0,22 мкм в BSC, делят на аликвоты, маркируют и хранят при температуре -80°C в виде нерасфасованной лекарственной субстанции (BDS). Нерасфасованную лекарственную субстанцию (BDS) подтверждают при помощи различных аналитических способов на соответствие всем критериям выпуска партии. Общее качество анализируют с помощью SDS-PAGE, отношение мономера к димеру и агрегату анализируют с помощью SE-HPLC, а качество N-конца и гомогенность продукта анализируют с помощью HI-HPLC.

Пример 4b. Икстенилированный слитый полипептид, содержащий XTEN со штрих-кодом на C-конце и другой XTEN со штрих-кодом на N-конце

[295] ЭКСПРЕССИЯ. Конструкцию, кодирующую икстенилированный слитый полипептид, содержащий одноцепочечный вариабельный фрагмент (scFv) к EGFR, XTEN со штрих-кодом длиной 864 аминокислоты (**SEQ ID NO: 8008**) на C-конце и XTEN со штрих-кодом длиной 288 аминокислот (**SEQ ID NO: 8007**) на N-конце, экспрессируют в запатентованном штамме AmE098 E. coli и распределяют в периплазму с использованием

N-концевой секреторной лидерной последовательности (MKKNIAFLLASMFVFSIATNAYA-) (SEQ ID NO: 8898), которая расщепляется во время транслокации. Ферментационные культуры выращивают на комплексной среде, не содержащей компонентов животного происхождения, при 37°C; и температуру сдвигают до 26°C до начала истощения запасов фосфатов. В ходе сбора материала цельный ферментационный бульон центрифугируют для осаждения клеток. При сборе материала регистрируют общий объем и сырую массу клеток (WCW; соотношение осадка и надосадочной жидкости), и осажденные клетки собирают и замораживают при -80°C.

[296] ИЗВЛЕЧЕНИЕ. Осадок замороженных клеток ресуспендируют в буфере для лизиса (100 мМ лимонной кислоты), ориентируясь на 30% сырой массы клеток. Ресуспензии дают уравновеситься при pH 4,4, затем гомогенизируют при 17000±200 бар, при этом температуру на выходе контролируют и поддерживают на уровне 15±5°C. Подтверждают, что pH гомогената находится в пределах указанного диапазона (pH 4,4±0,1).

[297] ОСВЕТЛЕНИЕ. Чтобы снизить примеси эндотоксина и клеток-хозяев, гомогенат подвергают низкотемпературной (10±5°C) кислотной (pH 4,4±0,1) флокуляции в течение ночи (15-20 часов). Чтобы удалить нерастворимую фракцию, флокулированный гомогенат центрифугируют в течение 40 минут при 8000 об/мин при 2-8°C и надосадочную жидкость сохраняют. Для удаления нуклеиновой кислоты, липидов и эндотоксина и для использования в качестве вспомогательного фильтрующего средства надосадочную жидкость доводят до 0,1% (м/м) диатомовой земли. Чтобы поддерживать фильтрующую добавку во взвешенном состоянии, надосадочную жидкость перемешивают лопастной мешалкой и оставляют для уравновешивания в течение 30 минут. Комплект фильтров, состоящий из глубинного фильтра, за которым следует фильтр 0,22 мкм, собирают, а затем промывают с помощью MQ. Надосадочную жидкость прокачивают через комплект фильтров, модулируя поток для поддержания перепада давления на уровне 25±5 фунтов на квадратный дюйм.

[298] ОЧИСТКА

[299] Захват с белком L. Для удаления белков клетки-хозяина, эндотоксина и нуклеиновой кислоты применяют белок L для захвата каппа-домена, присутствующего в scFv к aEGFR молекулы BPXTEN. В данном случае применяют неподвижную фазу для белка L (Tosoh TP AF-rProtein L-650F), подвижную фазу A для белка L (11,5 мМ лимонной кислоты, 24,5 мМ Na_2 HPO_4 , 125 мМ NaCl, 0,005% полисорбата 80, pH 5,0) и подвижную фазу B для белка L (11 мМ фосфорной кислоты, 0,005% полисорбата 80, pH 2,0). Колонку уравновешивают подвижной фазой A для белка L. Фильтрат доводят до pH 5,5±0,2 и загружают в колонку с белком L с целевой концентрацией смолы 2-4 г/л, затем вытесняют подвижной фазой A для белка L до тех пор, пока поглощение при 280 нм (A280) не возвращается к (локальному) исходному уровню. Связанный материал элюируют подвижной фазой B и собирают в виде фракции 2 CV с предварительным добавлением 0,4 CV 0,5 M Na_2 HPO_4 и анализируют с помощью SDS-PAGE.

[300] Промежуточная очистка IMAC. Для обеспечения целостности N-конца применяют аффинную хроматографию с иммобилизованными металлами (IMAC) для захвата N-концевой полигистидиновой метки (His(6); (SEQ ID NO: 8031)) молекулы слитого полипептида. В данном случае применяют неподвижную фазу для IMAC (GE IMAC Sepharose FF), подвижную фазу A для IMAC (12,2 мМ Na_2HPO_4 , 7,8 мМ NaH_2PO_4 , 500 мМ NaCl, 0,005% полисорбата 80, pH 7,0) и подвижную фазу B для IMAC (50 мМ гистидина, 200 мМ NaCl, 0,005% полисорбата 80, pH 6,5). Колонку уравновешивают подвижной фазой A для IMAC. Элюат после белка L доводят до pH 7,8±0,1 и 50±5 мСм/см (с 5 M NaCl). Скорректированный пул после белка L загружают в колонку IMAC с целевой концентрацией смолы 2 г/л и вытесняют подвижной фазой A для IMAC до тех пор, пока поглощение при 280 нм (A280) не вернется к (локальному) исходному уровню. Связанный материал элюируют подвижной фазой B для IMAC. Элюат после IMAC собирают в виде фракции 2 CV с предварительным добавлением 0,02 CV 200 мМ EDTA и анализируют с помощью SDS-PAGE.

[301] Промежуточная очистка с C-меткой. Для обеспечения целостности C-конца применяют аффинную хроматографию с C-меткой для захвата C-концевой метки -EPEA (SEQ ID NO: 8033). В данном случае применяют неподвижную фазу для C-метки (Thermo C-tagXL), подвижную фазу A для C-метки (50 мМ гистидина, 200 мМ NaCl, 0,005% полисорбата 80, pH 6,5) и подвижную фазу B для C-метки (20 мМ Tris, 0,6 M $MgCl_2$, 0,005% полисорбата 80, pH 7,0). Колонку уравновешивают подвижной фазой A для C-метки. Элюат после IMAC загружают в колонку с C-меткой с целевой концентрацией смолы 2 г/л и вытесняют подвижной фазой A для C-метки до тех пор, пока поглощение при 280 нм (A280) не вернется к (локальному) исходному уровню. Связанный материал элюируют подвижной фазой B для C-метки. Элюат после C-метки собирают в виде фракции 2 CV и анализируют с помощью SDS-PAGE.

[302] Очистка с AEX. Для отделения димеров и агрегатов от мономерного продукта применяют анионообменную хроматографию (AEX) для захвата электроотрицательных N- и C-концевых доменов XTEN. В данном случае применяют неподвижную фазу для AEX1 (BIA QA-80), подвижную фазу A для AEX1 (50 мМ гистидина, 200 мМ NaCl, 0,005% полисорбата 80, pH 6,5) и подвижную фазу B для AEX1 (50 мМ гистидина, 500 мМ NaCl, 0,005% полисорбата 80, pH 6,5). Колонку уравновешивают подвижной фазой A для AEX. Элюат C-метки разбавляют с помощью MQ до 10 мСм/см, загружают с целевой концентрацией смолы 2 г/л и вытесняют подвижной фазой A для AEX до тех пор, пока поглощение при 280 нм не вернется к (локальному) исходному уровню. Связанный материал элюируют градиентом от 0% B до 100% B при 60 CV. Фракции собирают аликвотами по 1 CV в то время как A280 выше (локального) исходного уровня на $\geq$ 2 произвольные мЕд оптической плотности. Элюированные фракции анализируют с помощью SDS-PAGE и SE-HPLC, а фракции, содержащие $\geq$ 98% мономера, объединяют (пул AEX) для дальнейшей обработки.

[303] СОСТАВЛЕНИЕ. Для замены продукта в буфере для состава и доведения

продукта до целевой концентрации (0,5 г/л) применяют ультрафильтрацию/диафильтрацию (UF/DF). Используя мембрану 10 кДа с площадью 0,1 м² и целевое давление TMP 15 фунтов на квадратный дюйм, пул AEX концентрируют до 0,5 г/л, затем разбавляют в 10 раз буфером для состава (50 мМ гистидина, 200 мМ NaCl, 0,005% полисорбата 80, pH 6,5). Пул AEX концентрируют в 10 раз и разбавляют в 10 раз еще два раза. Восстановленный составленный продукт фильтруют через фильтр 0,22 мкм в BSC, делят на аликвоты, маркируют и хранят при температуре -80°C в виде нерасфасованной лекарственной субстанции (BDS). BDS подтверждают при помощи различных аналитических способов на соответствие всем критериям выпуска партии. Общее качество анализируют с помощью SDS-PAGE, отношение мономера к димеру и агрегату анализируют с помощью SE-HPLC, а качество N-конца и гомогенность продукта анализируют с помощью HI-HPLC. Идентичность подтверждают при помощи ESI-MS.

Пример 5. Высвобождение пептидов штрих-кода при расщеплении протеазами

[304] Данный пример иллюстрирует высвобождение фрагмента(-ов) со штрих-кодом и эталонного(-ых) фрагмента(-ов) из смеси полипептидов, которая содержит различные длины или усеченные формы конструкции, содержащей XTEN, при помощи способов, раскрытых в данном документе.

[305] Образец конструкции, содержащей XTEN, восстанавливают и алкилируют посредством последовательной инкубации в DTT, а затем в йодацетамиде. Затем в образцах заменяют буфер и проводят обессоливание с помощью центрифужного картриджа для разделения по размеру. К образцам добавляют протеазу Glu-C при соотношении фермента к субстрату 1:5 и образцы инкубируют при 37°C для расщепления. Затем образцы доводят до температуры 4°C, чтобы остановить протеолитическую реакцию, и помещают во флаконы автоматического пробоотборника для анализа.

Пример 6. Обнаружение и количественное определение пептида(-ов) штрих-кода и эталонного(-ых) пептида(-ов)

[306] Данный пример иллюстрирует способы масс-спектрометрии, используемые для получения количественных измерений отдельных пептидов штрих-кода. Способ мониторинга одновременно протекающих реакций при помощи LC (PRM) запрограммирован в масс-спектрометре высокого разрешения с определением точной массы (HRAM). В отличие от традиционных способов масс-спектрометрии со сбором в зависимости от данных (DDA), способы PRM фокусируются на определенном наборе из 15-30 пептидов за один прогон, секвенируя каждый с помощью MS-MS один раз за рабочий цикл. Таким образом, этот способ генерирует хроматограммы экстрагированных ионов (XIC) для нефрагментированных ионов-предшественников интактного пептида, а также для каждого иона-фрагмента пептида для подтверждения его последовательности. XIC ионов-фрагментов часто являются более чувствительными и избирательно количественными, чем фрагменты ионов-предшественников. Используемый способ LC-PRM включает легкие и тяжелые версии семи пептидов штрих-кода. Площади хроматографических пиков всех ионов-фрагментов этих 14 пептидов измеряют после

получения, и для количественного измерения применяют преобладающий ион-фрагмент. Затем рассчитывают отношения площадей пиков пептидов штрих-кода XTEN к пептидам штрих-кода PAT в отношении относительного содержания XTEN:PAT в различных точках молекул XTEN.

Пример 7. Маркировка стабильными изотопами для количественного определения пептидов с помощью масс-спектрометрии (MS)

[307] Данный пример иллюстрирует схему мечения стабильными изотопами, позволяющую проводить абсолютное (а не относительное) количественное определение пептидов штрих-кода из полипептидов, содержащих XTEN. Используют стандартную количественную схему аминокислот с тяжелой меткой, где синтетические аналоги пептидов штрих-кода, в которых С-концевая глутаминовая кислота заменена аналогом с тяжелой меткой $(^{13}C)_5H_7(^{15}N)O_3$, закупаются у специализированного поставщика. Калибровочную кривую строят, когда известное количество полипептида, содержащего штрих-код XTEN, серийно разбавляют в матрице, где синтетический пептид с тяжелой меткой поддерживается при постоянной концентрации. Точное количественное определение можно выполнить путем калибровки отношений «тяжелых» и «легких» площадей хроматографических пиков по кривой относительно исследовательских образцов, содержащих такой же пиковый уровень пептида с тяжелой меткой.

Пример 8. Количественная оценка усечения содержащего XTEN полипептида

[308] Данный пример иллюстрирует количественную оценку вариантов длины или вариантов усечения в смеси полипептидов, содержащих XTEN.

[309] Например, пептид со штрих-кодом «SGPGSTPAESGSE» (SEQ ID NO: 8899) расположен в 76 аминокислотах от примерной последовательности BPXTEN со штрих-кодом, описанной и полученной в **примере 3**, для обозначения сильного усечения XTEN на N-конце BPXTEN. Также рассмотрим потенциальный фрагмент со штрих-кодом «SPAGSPTSTESGTSE» (SEQ ID NO: 8260), расположенный на N-конце. Отношение измерения содержания каждого пептида штрих-кода к уникальной последовательности эталонного пептида из последовательности биологически активного белка (например, фрагмента scFv) в соответствии с процедурой из примера 6 указывает на общее количество полноразмерных полипептидов и вариантов, имеющих усечения, которые могут повлиять на фармакологическую эффективность смеси образцов. Измерение содержания по меньшей мере одного эталонного фрагмента применяют для определения общего количества всех вариантов полипептида в смеси образцов. Соответственно, разница в содержании между эталонным фрагментом и фрагментами со штрих-кодом дает информацию о количестве усеченных вариантов полипептида. Данные LC-MS анализируют для определения соотношения количества фрагмента со штрих-кодом и эталонного фрагмента, что указывает на относительное количество фармакологически эффективных вариантов в полипептидной смеси.

[310] Набор из двух (или трех) штрих-кодов применяют для указания различных уровней усечения полипептида. Данные LC-MS используют для определения

соотношения количества каждого фрагмента со штрих-кодом и количества эталонного фрагмента, тем самым количественно определяя распределение вариантов усечения в полипептидной смеси.

Пример 9. Количественная оценка усечения содержащего XTEN полипептида

[311] Этапом очистки AC2329 является анионообменная хроматография QIR. Вторая половина основного пика элюирования из этой колонки содержит полноразмерный белок, тогда как первая половина пика содержит смесь полноразмерного белка, а также множества усеченных форм. Для оценки пептида штрих-кода брали две фракции. Одна фракция включала только вторую половину пика, до этого называемую «полноразмерной» фракцией. Вторая фракция включала первую половину пика, до этого называемую фракцией «синтетический белок и усеченные фрагменты», как представлено в данном документе:

[312] Аналитическая эксклюзионная хроматография (SEC) с наложением полноразмерной (синяя) фракции и фракции полноразмерная+усеченные фрагменты (черная). Очевидно, что фракция синтетический белок+усеченные фрагменты включает фрагменты такого же размера, как и интактный синтетический белок, а также много более мелких компонентов, на что указывает большое пологое плечо с правой стороны пика.

[313] Полноразмерный синтетический белок и полноразмерный синтетический белок и усеченные фрагменты смешивали в соотношениях, описанных в матрице разведения.

Матрица разбавления

Разбавление	1	2	3	4	5	6	7	8	9
Соединение 1 (нМ)	400	350	300	250	200	150	100	50	0
Соединение 1 усеч. (нМ)	0	50	100	150	200	250	300	350	400
Всего нМ	400	400	400	400	400	400	400	400	400
Всего объем	60	60	60	60	60	60	60	60	60

[314] Затем каждый из образцов по 400 нМ расщепляли с помощью 1 мг/мл GluC (Roche 10791156001) в реакционном буфере, содержащем 50 мМ Tris-HCl, pH 7,5+0,1% Rapigest (Waters 186001861). Расщепление с помощью GluC проводили в течение ночи при температуре 37°C во встряхивающем инкубаторе. После расщепления добавляли муравьиную кислоту до конечной концентрации 10% и инкубировали при температуре 37°C в течение 45 мин. Осадок с Rapigest удаляли центрифугированием при 16000 × g в течение 10 мин и добавляли стандартную смесь «тяжелых» пептидов до конечной концентрации 400 нМ. Анализ каждого гидролизата с помощью LC-MS проводили, как описано. Каждый образец измеряли в двух повторностях.

Анализ с помощью LC-MS

[315] Растворы анализировали способом 30-минутного градиента на колонке Waters HSS-T3 (176003994) с использованием системы UHPLC Vanquish (Thermo), соединенной с масс-спектрометром Q-Exactive Plus (Thermo). Способ MS заключался в способе на десяти лучших DDA, в котором десять лучших пептидов секвенировали с помощью анализа MSMS после каждого сканирования MS. Полученные файлы данных

обрабатывали с использованием программного обеспечения Skyline Software (MacCoss Lab, UW), в котором можно было рассчитать и измерить нормализованные по «тяжелому» пептиду концентрации каждого пептида штрих-кода.

[316] Каждое измерение, как показано на фигуре 5B, представляет собой область XIC N-штрих-кода SGPGSTPAE (SEQ ID NO: 8029) и C-штрих-кода GSAPGTE (SEQ ID NO: 8023), нормализованную к пику 400 нМ соответствующего синтетического пептида, меченного тяжелым изотопом. Разведение 1 имеет наименьшее количество усеченных фрагментов, а разведение 9 имеет наибольшее их количество. Эти данные показывают, что пептид N-штрих-кода измерен в относительно одинаковом количестве в образцах, содержащих усеченные и неусеченные фрагменты. Однако, во фракции, содержащей усечение, наблюдается уменьшение содержания пептида C-штрих-кода. Это свидетельствует о том, что терминация трансляции является самым сильным фактором, способствующим образованию усеченных видов, поскольку трансляция у прокариот инициируется на N-конце.

[317] После подробного описания настоящего изобретения со ссылками на его конкретные аспекты и/или варианты осуществления становится очевидным, что возможны модификации и варианты без отклонения от объема настоящего изобретения, определенного в прилагаемой формуле изобретения. Более конкретно, хотя некоторые аспекты настоящего изобретения могут быть идентифицированы в данном документе как особенно выгодные, предполагается, что настоящее изобретение не ограничено этими конкретными аспектами настоящего изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Полипептид, содержащий удлиненный рекомбинантный полипептид (XTEN), содержащий

(a) удлиненный рекомбинантный полипептид (XTEN), содержащий

(i) набор неперекрывающихся мотивов последовательности, где каждый неперекрывающийся мотив последовательности указанного набора повторяется по меньшей мере два раза в полипептиде XTEN; и

(ii) дополнительный неперекрывающийся мотив последовательности, который встречается только один раз в полипептиде XTEN; и

(b) первый фрагмент со штрих-кодом, высвобождаемый из полипептида при частичном или полном расщеплении протеазой, где первый фрагмент со штрих-кодом представляет собой часть полипептида XTEN, содержащую мотив последовательности, который встречается только один раз в полипептиде XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой;

где фрагмент со штрих-кодом не включает в себя N-концевую аминокислоту или C-концевую аминокислоту полипептида.

2. Полипептид по п. 1, где фрагмент со штрих-кодом не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой кислоте в полипептиде XTEN.

3. Полипептид по любому из пп. 1-2, где фрагмент со штрих-кодом содержит глутаминовую кислоту на C-конце.

4. Полипептид по любому из пп. 1-3, где фрагмент со штрих-кодом содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты.

5. Полипептид по п. 4, где остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает непосредственно к другому остатку глутаминовой кислоты.

6. Полипептид по любому из пп. 1-4, где фрагмент со штрих-кодом не включает остаток глутаминовой кислоты в положении, отличном от C-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин.

7. Полипептид по пп. 1-6, где фрагмент со штрих-кодом расположен на расстоянии от 10 аминокислот до 150 аминокислот от N-конца или C-конца полипептида.

8. Полипептид по любому из пп. 1-7, где мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 182-203 и 1715-1722.

9. Полипептид по п. 8, где мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 186-189.

10. Полипептид по п. 9, где набор неперекрывающихся мотивов последовательности включает по меньшей мере два, по меньшей мере три или все четыре мотива последовательности, идентифицированных под SEQ ID NO: 186-189.

11. Полипептид, содержащий удлиненный рекомбинантный полипептид (XTEN), содержащий первый фрагмент со штрих-кодом, высвобождаемый из полипептида при расщеплении протеазой, который является частью полипептида XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой; и где фрагмент со штрих-кодом:

(i) не включает N-концевую аминокислоту или C-концевую аминокислоту

полипептида;

(ii) не включает глутаминовую кислоту, которая непосредственно примыкает к другой глутаминовой

кислоте в полипептидах XTEN;

(iii) содержит глутаминовую кислоту на C-конце;

(iv) содержит N-концевую аминокислоту, которой непосредственно предшествует остаток глутаминовой кислоты;

и

(v) расположен на расстоянии от 10 аминокислот до 125 аминокислот от N-конца или

C-конца полипептида.

12. Полипептид по п. 11, где остаток глутаминовой кислоты, который предшествует N-концевой аминокислоте, не примыкает непосредственно к другому остатку глутаминовой кислоты.

13. Полипептид по п. 11 или п. 12, где фрагмент со штрих-кодом не включает остаток глутаминовой кислоты в положении, отличном от C-конца фрагмента со штрих-кодом, за исключением случаев, когда сразу за глутаминовой кислотой следует пролин.

14. Полипептид по п. 11, где полипептид XTEN содержит множество неперекрывающихся мотивов последовательности, где каждый из указанных мотивов последовательности имеет длину от 9 до 14 аминокислот.

15. Полипептид по п. 12, где мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 182-203 и 1715-1722.

16. Полипептид по п. 15, где мотивы последовательности из набора неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 186-189.

17. Полипептид по п. 16, где набор неперекрывающихся мотивов последовательности включает по меньшей мере два, по меньшей мере три или все четыре мотива последовательности под SEQ ID NO: 186-189.

18. Полипептид по любому из пп. 1-17, где по меньшей мере 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99% или 100% аминокислотных остатков полипептида XTEN представляют собой глицин (G), аланин (A), серин (S), треонин (T), глутаминовую кислоту (E) или пролин (P).

19. Полипептид по любому из пп. 1-17, где длина полипептида XTEN составляет от 150 до 3000 аминокислот.

20. Полипептид по п. 19, где длина полипептида XTEN составляет от 150 до 1000 аминокислот.

21. Полипептид по любому из пп. 1-20, где фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от N-конца полипептида.

22. Полипептид по любому из пп. 1-21, где фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от N-конца белка.

23. Полипептид по любому из пп. 1-22, где фрагмент со штрих-кодом расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от C-конца полипептида.

24. Полипептид по п. 23, где фрагмент со штрих-кодом расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от C-конца белка.

25. Полипептид по любому из пп. 1-24, где длина фрагмента со штрих-кодом составляет по меньшей мере 4 аминокислоты.

26. Полипептид по любому из пп. 1-24, где длина фрагмента со штрих-кодом составляет от 4 до 20, от 5 до 15, от 6 до 12 или от 7 до 10 аминокислот.

27. Полипептид по любому из пп. 1-26, где фрагмент со штрих-кодом идентифицирован под SEQ ID NO: 8020-8030 (BAR001-BAR011).

28. Полипептид по любому из пп. 1-27, где полипептид дополнительно содержит второй фрагмент со штрих-кодом, где второй фрагмент со штрих-кодом представляет собой часть полипептида XTEN, которая содержит мотив последовательности, который встречается только один раз в полипептиде XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

29. Полипептид по п. 28, где полипептид дополнительно содержит третий фрагмент со штрих-кодом, где третий фрагмент со штрих-кодом представляет собой часть полипептида XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

30. Полипептид по любому из пп. 1-29, где полипептид XTEN характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100% идентичностью последовательности с последовательностью, идентифицированной под SEQ ID NO: 8001-8019.

31. Полипептид по любому из пп. 1-31, где длина полипептида XTEN составляет по

меньшей мере 200, по меньшей мере 250, по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 450 или по меньшей мере 500 аминокислот.

32. Биологически активный полипептид, связанный с полипептидом по пп. 1-31.

33. Биологически активный полипептид по п. 32, где полипептид XTEN имеет проксимальный конец и дистальный конец по отношению к биологически активному полипептиду, где проксимальный конец расположен относительно дистального конца ближе к биологически активному полипептиду, и где фрагмент со штрих-кодом расположен в области полипептида XTEN, длина которого при измерении от дистального конца составляет от 5% до 50%, от 7% до 40% или от 10% до 30% длины полипептида XTEN.

34. Биологически активный полипептид по любому из п. 32 или п. 33, дополнительно содержащий один или более эталонных фрагментов, высвобождаемых из полипептида при расщеплении протеазой, где каждый из одного или более эталонных фрагментов содержит часть биологически активного полипептида.

35. Биологически активный полипептид по п. 34, где один или более эталонных фрагментов представляют собой один эталонный фрагмент, который отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после расщепления полипептида протеазой.

36. Биологически активный полипептид по любому из пп. 32-35, дополнительно содержащий первый сегмент для обеспечения высвобождения (RS1), расположенный между полипептидом XTEN и биологически активным полипептидом.

37. Биологически активный полипептид по п. 36, где RS1 содержит аминокислотную последовательность, характеризующуюся по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью, идентифицированной в данном документе в таблицах 8a-8b.

38. Биологически активный полипептид по любому из пп. 32-37, где биологически активный полипептид идентифицирован в данном документе под любой одной или комбинацией последовательностей в таблицах 4a-4h и 6a-6f.

39. Биологически активный полипептид по любому из пп. 32-38, где полипептид имеет конечный период полувыведения, который в по меньшей мере два раза дольше по сравнению с биологически активным полипептидом, не связанным с каким-либо полипептидом XTEN.

40. Биологически активный полипептид по любому из пп. 32-39, где полипептид является менее иммуногенным по сравнению с биологически активным полипептидом, не связанным с каким-либо полипептидом XTEN, где иммуногенность устанавливают путем измерения уровня продуцирования антител IgG, которые избирательно связываются с биологически активным полипептидом, после введения сравнимых доз человеку или животному.

41. Биологически активный полипептид по любому из пп. 32-40, где полипептид

демонстрирует коэффициент кажущейся молекулярной массы в физиологических условиях, превышающий приблизительно 6.

42. Биологически активный полипептид по любому из пп. 32-41, дополнительно содержащий второй полипептид XTEN, где первый полипептид XTEN расположен с N-конца от биологически активного полипептида, а второй полипептид XTEN расположен с С-конца от биологически активного полипептида.

43. Биологически активный полипептид по п. 42, дополнительно содержащий второй сегмент для обеспечения высвобождения (RS2), расположенный между биологически активным полипептидом и вторым XTEN.

44. Биологически активный полипептид по п. 43, где RS1 и RS2 являются идентичными по последовательности.

45. Биологически активный полипептид по п. 43, где RS1 и RS2 являются неидентичными по последовательности.

46. Биологически активный полипептид по любому из пп. 43-45, где каждый из RS1 и RS2 представляет собой субстрат для расщепления несколькими протеазами в одном, двух или трех сайтах расщепления в каждой последовательности сегмента для обеспечения высвобождения.

47. Биологически активный полипептид по любому из пп. 42-46, где полипептид содержит добавочный фрагмент со штрих-кодом, который представляет собой часть второго XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

48. Биологически активный полипептид по п. 47, где добавочный фрагмент со штрих-кодом не включает С-концевую аминокислоту полипептида.

49. Биологически активный полипептид по п. 47 или п. 48, где добавочный фрагмент со штрих-кодом содержит остаток глутаминовой кислоты на С-конце.

50. Биологически активный полипептид по любому из пп. 47-49, где добавочный фрагмент со штрих-кодом второго XTEN расположен в пределах 200, в пределах 150, в пределах 100 или в пределах 50 аминокислот от С-конца полипептида.

51. Биологически активный полипептид по любому из пп. 47-50, где добавочный фрагмент со штрих-кодом второго XTEN расположен на расстоянии от 10 до 200, от 30 до 200, от 40 до 150 или от 50 до 100 аминокислот от С-конца полипептида.

52. Биологически активный полипептид по любому из пп. 47-51, где длина добавочного фрагмента со штрих-кодом составляет от 4 до 20, от 5 до 15, от 6 до 12 или от 7 до 10 аминокислот.

53. Биологически активный полипептид по любому из пп. 47-52, где добавочный фрагмент со штрих-кодом идентифицирован под SEQ ID NO: 8020-8030 (BAR001-BAR011).

54. Биологически активный полипептид по любому из пп. 47-53, дополнительно содержащий набор фрагментов со штрих-кодом, который включает добавочный фрагмент

со штрих-кодом и по меньшей мере один дополнительный фрагмент со штрих-кодом, где каждый фрагмент со штрих-кодом из набора фрагментов со штрих-кодом представляет собой часть второго полипептида XTEN и отличается по последовательности и молекулярной массе от всех других пептидных фрагментов, которые могут высвобождаться из полипептида после полного расщепления полипептида протеазой.

55. Биологически активный полипептид по любому из пп. 42-55, где второй XTEN идентифицирован под SEQ ID NO: 8001-8019.

56. Биологически активный полипептид по любому из пп. 47-55, где добавочный фрагмент со штрих-кодом не включает остаток глутаминовой кислоты, который непосредственно примыкает к другому остатку глутаминовой кислоты в полипептиде.

57. Биологически активный полипептид по любому из пп. 42-56, где по меньшей мере 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотных остатков второго полипептида XTEN представляют собой глицин (G), аланин (A), серин (S), треонин (T), глутаминовую кислоту (E) или пролин (P).

58. Биологически активный полипептид по любому из пп. 42-57, где сумма общего количества аминокислот в первом полипептиде XTEN и общего количества аминокислот во втором полипептиде XTEN составляет по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 500, по меньшей мере 600, по меньшей мере 700 или по меньшей мере 800 аминокислот.

59. Биологически активный полипептид по любому из пп. 42-58, где второй полипептид XTEN содержит множество неперекрывающихся мотивов последовательности, где каждый мотив последовательности находится во втором полипептиде XTEN и имеет длину от 9 до 14 аминокислот.

60. Биологически активный полипептид по п. 59, где для второго полипептида XTEN мотивы последовательности из множества неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 182-203 и 1715-1722.

61. Биологически активный полипептид по п. 60, где мотивы последовательности из множества неперекрывающихся мотивов последовательности идентифицированы под SEQ ID NO: 186-189.

62. Биологически активный полипептид по п. 59 или п. 60, где для второго полипептида XTEN множество неперекрывающихся мотивов последовательности предусматривает по меньшей мере два, по меньшей мере три или все четыре из следующих мотивов: SEQ ID NO: 186-189.

63. Биологически активный полипептид по любому из пп. 42-62, где длина второго полипептида XTEN составляет от 150 до 3000 аминокислот.

64. Биологически активный полипептид по п. 63, где длина второго полипептида XTEN составляет от 150 до 1000 аминокислот.

65. Биологически активный полипептид по любому из пп. 42-64, где второй полипептид XTEN характеризуется по меньшей мере 90%, по меньшей мере 92%, по меньшей мере 95%, по меньшей мере 98%, по меньшей мере 99% или 100%

идентичностью последовательности с последовательностью, идентифицированной под SEQ ID NO: 8001-8019.

66. Биологически активный полипептид по любому из пп. 42-64, где длина второго полипептида XTEN составляет по меньшей мере 200, по меньшей мере 250, по меньшей мере 300, по меньшей мере 350, по меньшей мере 400, по меньшей мере 450 или по меньшей мере 500 аминокислот.

67. Смесь, содержащая множество полипептидов различной длины, где смесь предусматривает

первый набор полипептидов, где каждый полипептид из первого набора полипептидов содержит фрагмент со штрих-кодом, который высвобождается из полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов; и

второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов;

где каждый из как первого набора полипептидов, так и второго набора полипептидов содержит эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и может высвобождаться при расщеплении протеазой;

где отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,70.

68. Смесь по п. 67, где отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8, 0,9 или 0,95.

69. Смесь по п. 67 или п. 68, где эталонный фрагмент встречается дважды в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

70. Смесь по п. 67, где первый набор полипептидов содержит полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида.

71. Смесь по п. 67, где полноразмерный полипептид представляет собой полипептид по любому из пп. 1-68.

72. Смесь по любому из пп. 70-71, где фрагмент со штрих-кодом не содержит N-концевую аминокислоту и С-концевую аминокислоту полноразмерного полипептида.

73. Смесь по любому из пп. 67-72, где в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-концевого, так и С-концевого усечения полноразмерного полипептида.

74. Нуклеиновая кислота, содержащая полинуклеотид, кодирующий полипептид по любому из пп. 1-75, или обратно-комплементарную последовательность этого полинуклеотида.

75. Вектор экспрессии, содержащий полинуклеотидную последовательность по п.

74 и регуляторную последовательность, функционально связанную с полинуклеотидной последовательностью.

76. Клетка-хозяин, содержащая вектор экспрессии по п. 75.

77. Клетка-хозяин по п. 76, где клетка-хозяин является прокариотической.

78. Клетка-хозяин по п. 77, где клетка-хозяин представляет собой E. coli.

79. Клетка-хозяин по п. 78, где клетка-хозяин представляет собой клетку млекопитающего.

80. Фармацевтическая композиция, содержащая полипептид по любому из пп. 1-73 и одно или более из фармацевтически приемлемых вспомогательных веществ.

81. Применение полипептида по любому из пп. 1-68 или смесей на его основе по пп. 67-73 при получении лекарственного препарата для лечения заболевания у человека или животного.

82. Применение по п. 81, где заболевание представляет собой рак.

83. Способ лечения заболевания у человека или животного, при этом способ предусматривает введение человеку или животному, нуждающемуся в этом, одной или более терапевтически эффективных доз фармацевтической композиции по п. 80.

84. Способ по п. 83, где заболевание представляет собой формы рака.

85. Способ по п. 83, где человек или животное представляет собой человека.

86. Способ оценки в смеси, содержащей полипептиды различной длины, относительного количества первого набора полипептидов в смеси ко второму набору полипептидов в смеси, где каждый полипептид первого набора полипептидов имеет общий фрагмент со штрих-кодом, который встречается один раз в полипептиде, и каждый полипептид второго набора полипептидов не имеет фрагмента со штрих-кодом, который является общим для полипептидов из первого набора, где каждый из отдельных полипептидов как из первого набора полипептидов, так и из второго набора полипептидов, содержит эталонный фрагмент, при этом способ предусматривает

приведение смеси в контакт с протеазой для получения множества протеолитических фрагментов, полученных в результате расщепления первого набора полипептидов и второго набора полипептидов, где множество протеолитических фрагментов включает

множество эталонных фрагментов и

множество фрагментов со штрих-кодом; и

определение отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов с оценкой тем самым относительных количеств первого набора полипептидов ко второму набору полипептидов.

87. Способ по п. 86, где эталонный фрагмент встречается не более одного раза в каждом полипептиде из первого набора полипептидов и второго набора полипептидов.

88. Способ по п. 86 или п. 87, где протеаза расщепляет полипептиды различной длины с С-концевой стороны от остатков глутаминовой кислоты, за которыми не следует остаток пролина.

89. Способ по любому из пп. 86-88, где протеаза представляет собой протеазу Glu-C.

90. Способ по любому из пп. 86-89, где протеаза не является трипсином.

91. Способ по любому из пп. 86-90, где стадия определения отношения количества фрагментов со штрих-кодом к количеству эталонных фрагментов предусматривает количественную оценку фрагментов со штрих-кодом и эталонных фрагментов в смеси после ее контакта с протеазой.

92. Способ по п. 91, где фрагменты со штрих-кодом и эталонные фрагменты идентифицируют на основе их соответствующих масс.

93. Способ по п. 91 или п. 92, где фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью масс-спектрометрии.

94. Способ по любому из пп. 91-93, где фрагменты со штрих-кодом и эталонные фрагменты идентифицируют с помощью жидкостной хроматографии-масс-спектрометрии (LC-MS).

95. Способ по любому из пп. 86-94, где стадия определения отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает изобарическое мечение.

96. Способ по любому из пп. 86-95, где стадия определения отношения фрагментов со штрих-кодом к эталонным фрагментам предусматривает введение в смесь одного или обоих из меченного изотопом эталонного фрагмента и меченного изотопом фрагмента со штрих-кодом.

97. Способ по п. 96, где фрагмент со штрих-кодом, если он присутствует, представляет собой часть полипептида XTEN.

98. Способ по любому из пп. 86-97, где смесь полипептидов с различными значениями длины предусматривает полипептид по любому из пп. 1-68.

99. Способ по любому из пп. 86-98, где полипептиды различной длины предусматривают полноразмерный полипептид и его усеченные фрагменты.

100. Способ по п. 99, где полипептиды различной длины представляют собой полноразмерный полипептид и его усеченные фрагменты.

101. Способ по любому из пп. 86-100, где в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, C-концевого усечения или как N-, так и C-концевого усечения полноразмерного полипептида.

102. Способ по п. 101, где полноразмерный полипептид представляет собой полипептид по любому из пп. 1-68.

103. Способ по любому из пп. 86-102, где отношение количества фрагментов со штрих-кодом к эталонным фрагментам составляет более 0,5, 0,6, 0,7, 0,8, 0,9 или 0,95.

104. Смесь, содержащая множество полипептидов различной длины, где смесь предусматривает

первый набор полипептидов, где каждый полипептид из первого набора полипептидов содержит фрагмент со штрих-кодом, который высвобождается из

полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов; и

второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов;

где каждый из как первого набора полипептидов, так и второго набора полипептидов содержит

эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и может высвобождаться при расщеплении протеазой;

где количество эталонных фрагментов, определенное количественно в смеси полипептидов после расщепления протеазой, равно сумме количества первого и второго наборов полипептидов в смеси, и количество фрагментов со штрих-кодом, определенное количественно в смеси полипептидов после расщепления протеазой, равно количеству первого набора полипептидов в смеси.

105. Смесь по п. 104, где если первый набор полипептидов содержит один эталонный фрагмент, то отношение первого набора полипептидов к полипептидам в смеси, содержащей эталонный фрагмент, составляет более 0,7.

106. Смесь по п. 105, где отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8, 0,9 или 0,95.

107. Смесь по любому из пп. 104-106, где эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

108. Смесь по любому из пп. 104-106, где эталонный фрагмент встречается дважды в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

109. Смесь по любому из пп. 104-106, где первый набор полипептидов содержит полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида.

110. Смесь по п. 104 или п. 105, где полноразмерный полипептид представляет собой полипептид по любому из пп. 1-66.

111. Смесь по п. 108 или п. 109, где фрагмент со штрих-кодом не содержит N-концевую аминокислоту и С-концевую аминокислоту полноразмерного полипептида.

112. Смесь по любому из пп. 104-111, где в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-концевого, так и С-концевого усечения полноразмерного полипептида.

113. Смесь по любому из пп. 104-106, где эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

114. Смесь по п. 104 или п. 105, где количество эталонных фрагментов в первом наборе полипептидов может отличаться от количества эталонных фрагментов во втором

наборе полипептидов, но их количество в каждом полипептиде каждого набора должно быть одинаковым.

115. Смесь по п. 108, где каждый из эталонных фрагментов в полипептидах смеси имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов.

116. Смесь, содержащая множество полипептидов различной длины, где смесь предусматривает

первый набор полипептидов, где каждый полипептид из первого набора полипептидов содержит

фрагмент со штрих-кодом, который может высвобождаться из полипептида при расщеплении протеазой и имеет последовательность и молекулярную массу, отличающиеся от последовательности и молекулярной массы всех других фрагментов, которые могут высвобождаться из первого набора полипептидов; и

второй набор полипептидов, в котором отсутствует фрагмент со штрих-кодом первого набора полипептидов;

где каждый из как первого набора полипептидов, так и второго набора полипептидов содержит эталонный фрагмент, который является общим для первого набора полипептидов и второго набора полипептидов и может высвобождаться при расщеплении протеазой;

где отношение первого набора полипептидов к полипептидам в смеси имеет формулу:

[полипептиды, содержащие штрих-код]/[(полипептиды, содержащие эталонные пептиды) x N],

где N представляет собой количество появлений эталонного пептида, который высвобождается из каждого полипептида в смеси.

117. Смесь по п. 116, где если первый набор полипептидов содержит один эталонный фрагмент, то отношение первого набора полипептидов к полипептидам в смеси, содержащей эталонный фрагмент, составляет более 0,7.

118. Смесь по п. 117, где отношение первого набора полипептидов к полипептидам, содержащим эталонный фрагмент, составляет более 0,8, 0,9 или 0,95.

119. Смесь по любому из пп. 116-118, где эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

120. Смесь по любому из пп. 116-118, где эталонный фрагмент встречается дважды в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

121. Смесь по п. 115 или п. 116, где первый набор полипептидов содержит полноразмерный полипептид, где фрагмент со штрих-кодом представляет собой часть полноразмерного полипептида.

122. Смесь по любому из пп. 116-118, где полноразмерный полипептид представляет собой полипептид по любому из пп. 1-66.

123. Смесь по п. 120 или п. 121, где фрагмент со штрих-кодом не содержит N-концевую аминокислоту и С-концевую аминокислоту полноразмерного полипептида.

124. Смесь по любому из пп. 115-123, где в смеси полипептидов с различными значениями длины они отличаются друг от друга из-за N-концевого усечения, С-концевого усечения или как N-концевого, так и С-концевого усечения полноразмерного полипептида.

125. Смесь по любому из пп. 115-117, где эталонный фрагмент встречается не более одного раза в каждом полипептиде первого набора полипептидов и второго набора полипептидов.

126. Смесь по п. 115 или п. 116, где количество эталонных фрагментов в первом наборе полипептидов может отличаться от количества эталонных фрагментов во втором наборе полипептидов, но их количество в каждом полипептиде каждого набора должно быть одинаковым.

127. Смесь по п. 120, где каждый из эталонных фрагментов в полипептидах смеси имеет последовательность и молекулярную массу, которые отличаются от последовательности и молекулярной массы всех других фрагментов.

128. Способ определения целостности последовательности полипептидов, содержащих первый набор полипептидов в смеси полипептидов по пп. 104-127, при этом способ предусматривает стадии расщепления смеси полипептидов протеазой, которая высвобождает фрагмент со штрих-кодом и эталонный фрагмент из первого набора полипептидов и высвобождает эталонный фрагмент из второго набора полипептидов, и определения отношения фрагментов со штрих-кодом из первого набора полипептидов к эталонным фрагментам из первого и второго набора полипептидов, где целостность последовательности полипептидов первого набора полипептидов выявляют путем сравнения соотношения фрагментов с ожидаемым соотношением фрагментов на основе количества фрагментов со штрих-кодом и эталонных фрагментов в полипептидах, включающих первый и второй наборы полипептидов.

129. Способ по п. 128, где фрагменты со штрих-кодом и эталонные фрагменты выявляют с помощью LC/MS.

130. Способ по п. 129, где в смесь полипептидов добавляют известное количество стандартного материала, где стандартный материал содержит меченную изотопами версию указанной смеси множества полипептидов различной длины.

131. Способ по п. 130, где меченную изотопами версию указанной смеси полипептидов различной длины добавляют к смеси перед расщеплением указанной протеазой.

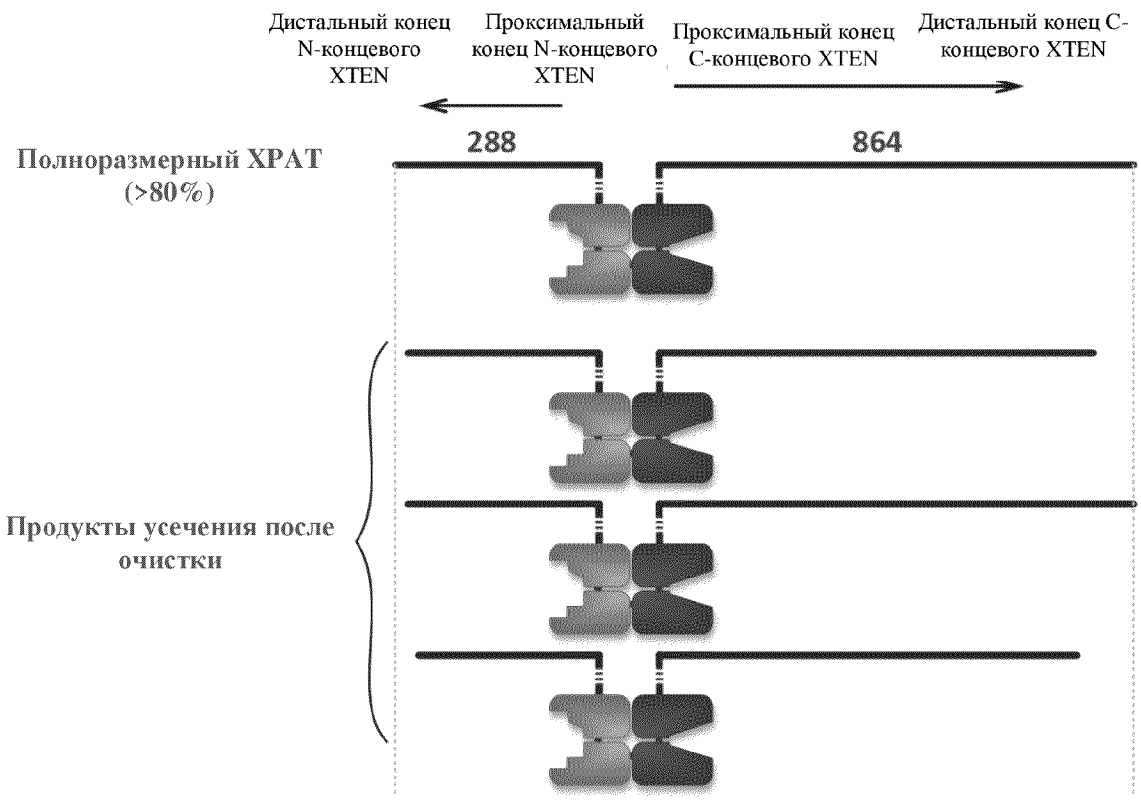
132. Способ по п. 130, где меченную изотопами версию указанной смеси полипептидов с различными значениями длины расщепляют указанной протеазой перед добавлением к смеси по п. 129 после того, как эта смесь была расщеплена указанной протеазой.

133. Способ по любому из пп. 130-132, дополнительно предусматривающий

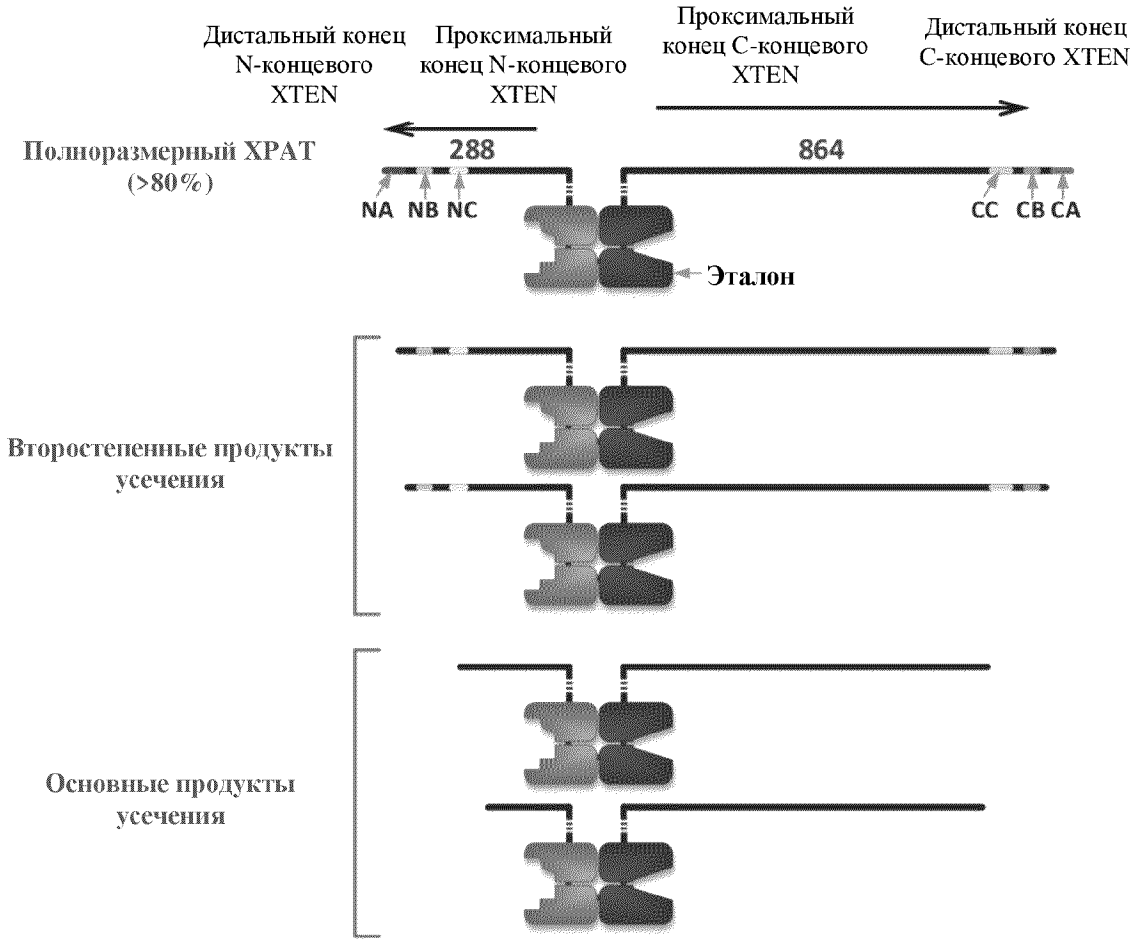
количественную оценку обнаруженных изотопно различимых количеств фрагментов со штрих-кодом, эталонных фрагментов или того и другого.

По доверенности

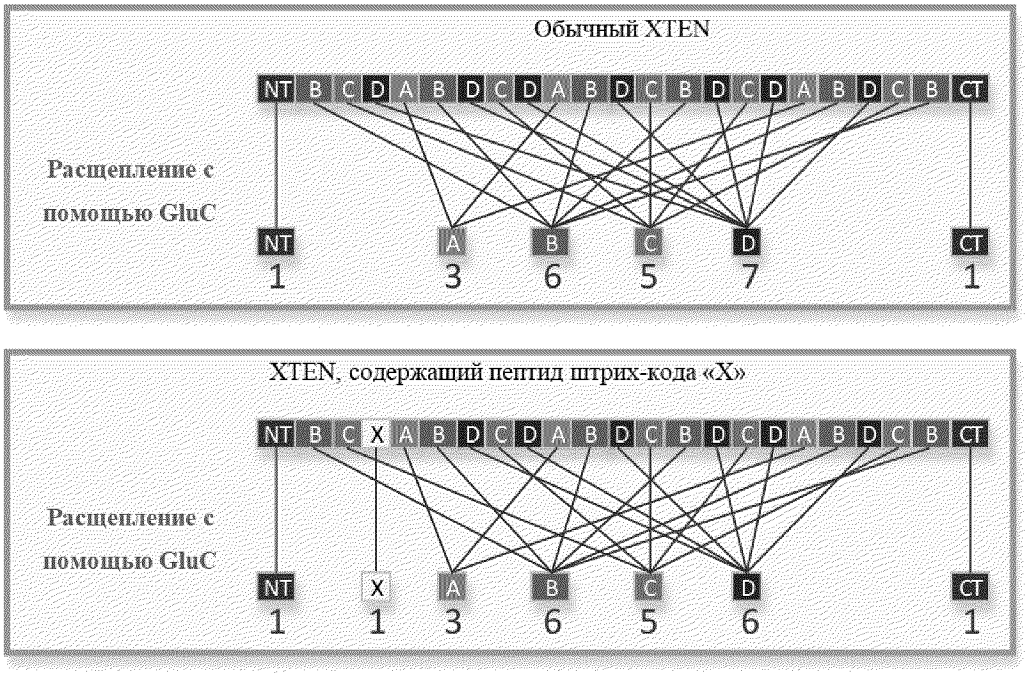
Фигура 1


Дистальный конец N-концевого XTEN
Проксимальный конец N-концевого XTEN
Проксимальный конец C-концевого XTEN
Дистальный конец C-концевого XTEN
288
864
Полноразмерный XPAT (>80%)
Продукты усечения после очистки

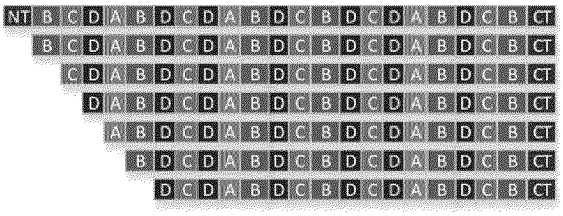
Фигура 2


Дистальный конец N-концевого XTEN
Проксимальный конец N-концевого XTEN
Проксимальный конец C-концевого XTEN
Дистальный конец C-концевого XTEN
Полноразмерный XPAT (>80%)
288
864
NA NB NC
CC CB CA
Эталон
Второстепенные продукты усечения
Основные продукты усечения

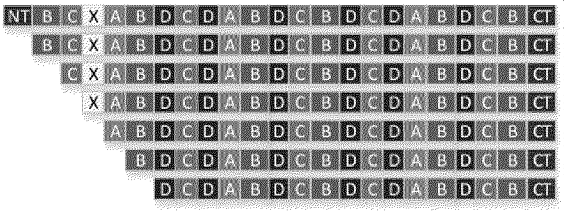
Фигура 3


Обычный XTEN
NT B C D A B D C D A B D C B D C D A B D C B CT
Расщепление с помощью GluC
NT A B C D CT
1 3 6 5 7 1
XTEN, содержащий пептид штрих-кода «X»
NT B C X A B D C D A B D C B D C D A B D C B CT
Расщепление с помощью GluC
NT X A B C D CT
1 1 3 6 5 6 1

Фигура 4А

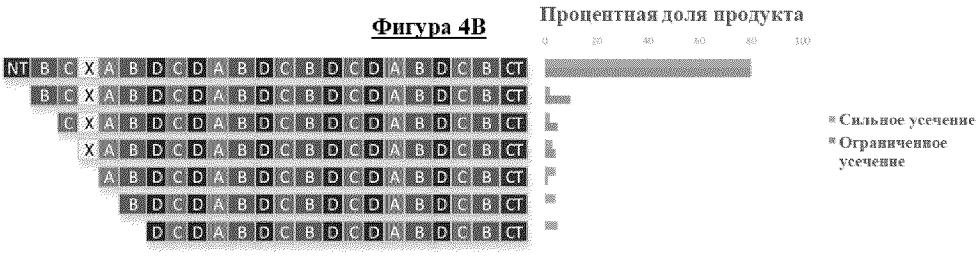


NT	A	B	C	D	CT	X
1	3	6	5	7	1	
	3	6	5	7	1	
	3	5	5	7	1	
	3	5	4	7	1	
	3	5	4	6	1	
	2	4	4	6	1	
	2	3	4	6	1	



NT	A	B	C	D	CT	X	
1	3	6	5	6	1	1	Наличие пептида штрих-кода = ограниченное усечение
	3	6	5	6	1	1	
	3	5	5	6	1	1	
	3	5	4	6	1	1	
	3	5	4	6	1		Отсутствие пептида штрих-кода = сильное усечение
	2	4	4	6	1		
	2	3	4	6	1		

Фигура 4B



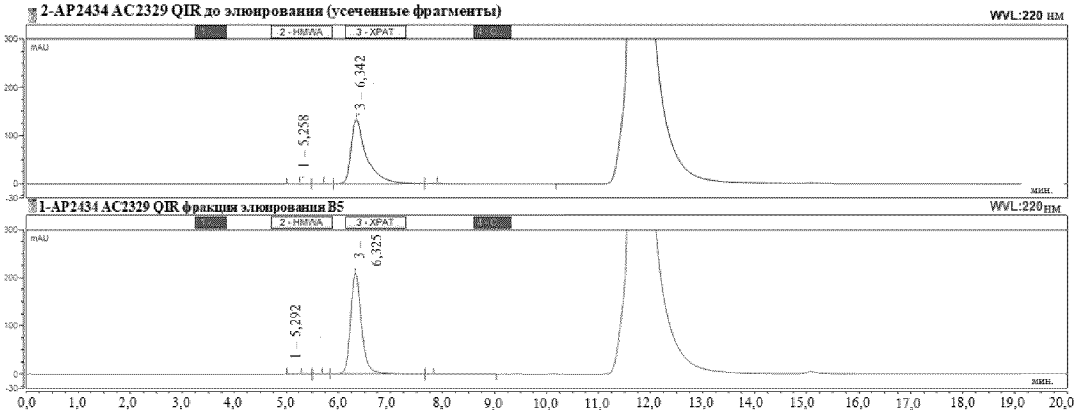
Относительное содержание продуктов расщепления

	NT	A	B	C	D	CT	X
Ограниченное	0,80	1,00	0,98	0,99	1,00	1,00	0,99
Сильное усечение	0,80	0,97	0,95	0,97	1,00	1,00	0,87

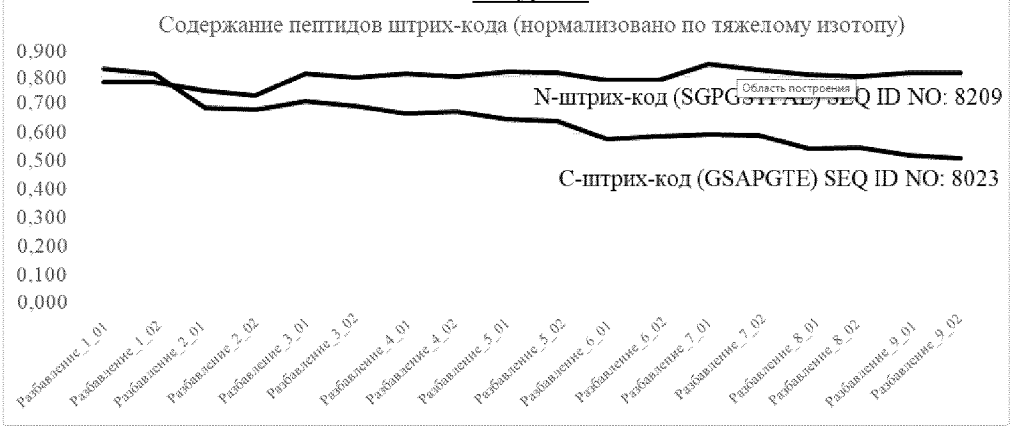
99% XTEN имеет ограниченное усечение

13% XTEN имеет сильное усечение

Фигура 5A


2-AP2434 AC2329 QIR до элюирования (усеченные фрагменты)
WVL:220 нм
2 - HMWA
3 - XPAT
mAU
1 — 5,258
3 — 6,342
мин.
1-AP2434 AC2329 QIR фракция элюирования B5
WVL:220 нм
2 - HMWA
3 - XPAT
mAU
1 — 5,292
3 — 6,325
мин.
0,0 1,0 2,0 3,0 4,0 5,0 6,0 7,0 8,0 9,0 10,0 11,0 12,0 13,0 14,0 15,0 16,0 17,0 18,0 19,0 20,0

Фигура 5B


Содержание пептидов штрих-кода (нормализовано по тяжелому изотопу)
0,900
0,800
0,700
0,600
0,500
0,400
0,300
0,200
0,100
0,000
Область построения
N-штрих-код (SGPG[illegible]) SEQ ID NO: 8209
C-штрих-код (GSAPGTE) SEQ ID NO: 8023
Разбавление_1_01
Разбавление_1_02
Разбавление_2_01
Разбавление_2_02
Разбавление_3_01
Разбавление_3_02
Разбавление_4_01
Разбавление_4_02
Разбавление_5_01
Разбавление_5_02
Разбавление_6_01
Разбавление_6_02
Разбавление_7_01
Разбавление_7_02
Разбавление_8_01
Разбавление_8_02
Разбавление_9_01
Разбавление_9_02