

МИНИГЕН АНТИТРОМБИНА III ЧЕЛОВЕКА С ОПТИМИЗИРОВАННЫМ ПАТТЕРНОМ СПЛАЙСИНГА

© 2019 г. М. В. Шепелев^а, * Е. К. Саакян^а, С. В. Калиниченко^а, И. В. Коробко^а

^аИнститут биологии гена Российской академии наук, Москва, 119334 Россия

*e-mail: mshepelev@mail.ru

Поступила в редакцию 16.11.2018 г.

После доработки 16.11.2018 г.

Принята к печати 11.12.2018 г.

Антитромбин III (АТ3) – белок из суперсемейства ингибиторов сериновых протеаз (серпинов), является основным физиологическим антикоагулянтом. На основе гена *SERPINC1* сконструирован миниген, кодирующий значимый для медицины и биотехнологии белок АТ3 человека. Для этого использовали минимизацию размера протяженных интронов с сохранением последовательностей, фланкирующих сайты сплайсинга. Анализ паттерна сплайсинга минигена АТ3 выявил один корректный транскрипт и два альтернативных, образующихся вследствие либо пропуска второго и третьего экзона минигена, либо встраивания аберрантного экзона путем использования криптических сайтов сплайсинга в первом интроне минигена. Паттерн сплайсинга минигена АТ3 успешно оптимизирован с помощью сайт-направленного мутагенеза криптических сайтов сплайсинга, что позволило предотвратить формирование альтернативных транскриптов. Присутствие в первом интроне минигена криптических сайтов сплайсинга, подтвержденное с использованием программы Human Splicing Finder v.3.1, указывает на принципиальную возможность выявления сайтов альтернативного сплайсинга в минимизированных или гибридных интронах минигенов и их удаления с помощью мутагенеза еще до экспериментальной проверки паттерна сплайсинга. Описанный подход к дизайну минигенов вкпе с биоинформатическим анализом нуклеотидных последовательностей интронов может использоваться при конструировании минигенов для получения трансгенных животных с продукцией экономически значимых рекомбинантных белков в молоке.

Ключевые слова: миниген, антитромбин III, паттерн сплайсинга, аберрантный экзон, альтернативный транскрипт

DOI: 10.1134/S0026898419030170

ВВЕДЕНИЕ

Белок антитромбин III (АТ3), входящий в суперсемейство ингибиторов сериновых протеаз (серпинов), является основным физиологическим антикоагулянтом, ингибирующим тромбин и факторы свертывания Ха и IXa [1]. Как наследственная, так и приобретенная недостаточность АТ3 у человека проявляется прежде всего венозными тромбозами [1]. Для профилактики тромбозов при хирургических вмешательствах или родах пациентам с недостаточностью АТ3 показаны препараты АТ3, который получают из крови человека или из молока трансгенных коз [2, 3].

Производство рекомбинантных белков в молоке трансгенных животных представляет актуальное направление биотехнологии [4]. Рекомбинантный АТ3 под коммерческим названием Atryn® был первым одобренным FDA (Управление по контролю качества пищевых продуктов и лекар-

ственных препаратов США) лекарственным средством, получаемым из молока трансгенных животных [5]. Белоккодирующая последовательность, используемая для получения трансгенных животных, продуцирующих рекомбинантные белки, может входить в состав полноразмерной копии гена, кДНК или минигена, содержащего гетерологичные, гибридные или минимизированные нативные интроны целевого гена [4]. В частности, трансгенные козы, молоко которых содержит рекомбинантный АТ3 человека, получены с использованием генетической конструкции, несущей кДНК АТ3 человека [6]. Тем не менее, считается, что присутствие интронов положительно влияет на уровень экспрессии трансгена и продукцию рекомбинантного белка [7, 8]. Однако полноразмерные копии генов человека зачастую имеют очень большой размер (до десятков тысяч пар нуклеотидов), что затрудняет проведение генно-инженер-

Сокращения: АТ3 – антитромбин III; НТО – нетранслируемая область.

ных манипуляций и обуславливает необходимость минимизации размера генетической конструкции, в том числе с помощью минигена. В частности, описанный ранее миниген, кодирующий АТЗ человека, содержал лишь часть первого интрона гена [9], а его экспрессию изучали только на клетках в культуре.

Цель данной работы состояла в конструировании минигена, кодирующего белок АТЗ, с использованием стратегии, предполагающей удаление наиболее протяженных интронов гена с сохранением коротких последовательностей (около 150 п.н.), фланкирующих сайты сплайсинга.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Штаммы *Escherichia coli*, выделение ДНК. Для генно-инженерных манипуляций и получения плазмидной ДНК в препаративных количествах использовали штамм *E. coli* XL-1 Blue MRF' ("Stratagene", США). Плазмидную ДНК выделяли с помощью наборов GenElute Plasmid Miniprep Kit ("Sigma-Aldrich", США) или GeneJET Plasmid Miniprep Kit ("Thermo Fisher Scientific", США) в соответствии с протоколами производителей. Геномную ДНК человека, используемую в качестве матрицы в ПЦР, выделяли из крови здорового донора с помощью набора реагентов QIAamp DNA Blood Mini Kit ("Qiagen", Германия). Геномную ДНК мышей гибридной линии F₁(СВА × С57BL/6), служащую матрицей в ПЦР, выделяли из фрагмента хвоста с помощью набора GenElute Mammalian Genomic DNA Miniprep kit ("Sigma-Aldrich"). Нуклеотидные последовательности всех амплифицированных в ПЦР фрагментов ДНК и фрагментов минигена после проведения мутагенеза в соответствующих плаزمидях верифицировали путем секвенирования с помощью набора реактивов ABI PRISM® BigDye™ Terminator v. 3.1 с последующим анализом продуктов реакции на автоматическом секвенаторе Applied Biosystems 3730 DNA Analyzer в ЦКП "Геном" Института молекулярной биологии им. В.А. Энгельгардта РАН, Россия.

Конструирование минигена, кодирующего анти-тромбин III человека. Фрагменты генома человека, используемые для конструирования минигена, кодирующего АТЗ, амплифицировали с помощью ПЦР. При необходимости введения сайтов рестрикции эти фрагменты реамплифицировали с использованием соответствующих праймеров. Все ПЦР-продукты клонировали в вектор pCR-Blunt ("Invitrogen", США), при необходимости выбирая клоны с определенной ориентацией вставки. Описание амплифицированных фрагментов генома человека, последовательности олигонуклеотидных праймеров и названия плазмид приведены в табл. 1. В нуклеотидной последовательности плазмиды рhAT3ex35-3 при секвенировании обнаружен сайт рестрикции PstI, отсутствующий в ре-

ференсной геномной последовательности гена *SERPINC1* человека и представленный, вероятно, полиморфизмом. Референсную нуклеотидную последовательность восстанавливали с помощью сайт-направленного мутагенеза по протоколу QuikChange ("Agilent Technologies", США) с олигонуклеотидными праймерами 5'-ccagaggtgctgcaagagtgctggatga-3' и 5'-tcattccagccactcttcagcacctctgg-3'.

Фрагмент генома мыши размером 2209 п.н. ("левое плечо" геномного локуса гена *Serpinc1*, кодирующего АТЗ), включавший 3'-часть первого интрона транскрипта АТЗ мыши (NM_080844.4) (2004 п.н.), второй экзон (143 п.н.) и 5'-часть второго интрона (62 п.н.) амплифицировали в ПЦР с использованием праймеров 5'-ctggaagctcaggaagataac-3' и 5'-gagagtccttcagaggtcaga-3'. Далее 5'-область левого плеча АТЗ мыши реамплифицировали с праймерами 5'-aatccttagcggccgcataaagg-3' и 5'-gagagtccttcagaggtcaga-3'. ПЦР-продукт клонировали в вектор pCR-Blunt и получили плазмиду рmATLA-Not. 3'-Область левого плеча реамплифицировали с праймерами 5'-сааactcgaggctagcccgtcagag-3' и 5'-gagagtccttcagaggtcaga-3' и клонировали в вектор pCR-Blunt, получив плазмиду рmAT3LA.

Далее с помощью ПЦР сконструировали гибридную молекулу ДНК, включающую "левое плечо" геномного локуса АТЗ мыши, слитое с последовательностью первого экзона гена АТЗ человека (рис. 1з). С этой целью частично перекрывающиеся фрагменты левого плеча АТЗ и первого экзона АТЗ человека амплифицировали с использованием праймеров 5'-сааactcgaggctagcccgtcagag-3' и 5'-ctatcacattggaatacatggccgatgctc-3' на матрице плазмиды рmAT3LA и праймеров 5'-taggagcatcgccatgtattccaatgtga-3' и 5'-acggatccttgactgtaactaccag-3' на матрице плазмиды рhAT3ex1. Полученные ПЦР-продукты смешивали, гибридизовали и проводили три цикла ПЦР без добавления праймеров для элонгации гибридных молекул ДНК, сформированных в результате отжига перекрывающихся последовательностей исходных продуктов ПЦР. Далее добавляли фланкирующие праймеры 5'-сааactcgaggctagcccgtcagag-3' и 5'-acggatccttgactgtaactaccag-3' и амплифицировали полноразмерный гибридный фрагмент ДНК, который клонировали в вектор pCR-Blunt. Таким образом получили плазмиду рAT3LAhyb. Далее фрагмент плазмиды рAT3LAhyb XhoI (обработан фрагментом Кленова)—BamHI клонировали в плазмиду рBlue-scriptSKII(-)("Stratagene", США) по сайтам XbaI (обработан фрагментом Кленова)—BamHI и получили плазмиду рBI-AT3LAhyb. Затем фрагмент NotI—NheI, включающий 5'-область геномного локуса АТЗ мыши, клонировали из плазмиды рATLA-Not по сайтам NotI—NheI в вектор рBI-AT3LAhyb, получив плазмиду рATLA, содержащую левое плечо геномного локуса АТЗ мыши, слитое с последовательностью первого экзона гена *SERPINC1* человека.

Таблица 1. Фрагменты гена *SERPINC1* человека, амплифицированные с помощью ПЦР для конструирования минигена

Фрагмент гена <i>SERPINC1</i>	Праймеры 5' → 3'	ПЦР-продукт, п.н.	Плазида	Сайты клонирования
3'-Часть первого экзона (94 п.н.) + 5'-часть первого интрона (146 п.н.)	F: gccctcagatcacactatctcc R: aatgagctcttccaacagctc	239	phAT3ex1	
3'-Часть первого интрона (303 п.н.) + экзон 2 (367 п.н.) + 5'-часть второго интрона (102 п.н.)	F: cacaacaaggaagctctctcttc R: tagtggcctgcagctgttgg	772	-----	
3'-Часть первого интрона (232 п.н.) + экзон 2 (367 п.н.) + 5'-часть второго интрона (102 п.н.)	F: atc agatctt ctctaatccaacatg R: tagtggcctgcagctgttgg	701	phAT3ex2-BgP	BgIII (введен) – PstI (эндогенный)
3'-Часть второго интрона (237 п.н.) + экзон 3 (216 п.н.) + интрон 3 (907 п.н.) + 5'-часть экзона 4 (120 п.н.)	F: cttaggggcaacctagtctc R: aaccagaccagaacagtgag	1480	-----	
3'-Часть интрона 3 (146 п.н.) + экзон 3 (216 п.н.) + интрон 3 (907 п.н.) + 5'-часть экзона 4 (120 п.н.)	F: gagcattt ctgctg ctccagg R: aaccagaccagaacagtgag	1389	phAT3ex35-5	PstI (введен) – SacI (эндогенный)
Экзон 4 (138 п.н.) + интрон 4 (811 п.н.) + экзон 5 (391 п.н.) + 5'-часть интрона 5 (113 п.н.)	F: gaaaatgcagagacaatccagag R: cagtatccaggagctctgac	1453	-----	
3'-Часть экзона 4 (62 п.н.) + интрон 4 (811 п.н.) + экзон 5 (391 п.н.) + 5'-часть интрона 5 (113 п.н.)	F: ttcctccggaagccatcaatg R: caggatcc aggagctctgac	1377	phAT3ex35-3	SacI (эндогенный) – BamHI (введен)
3'-Часть интрона 2 (146 п.н.) + экзон 3 (216 п.н.) + интрон 3 (907 п.н.) + интрон 4 (811 п.н.) + экзон 5 (391 п.н.) + 5'-часть интрона 5 (113 п.н.)	Фрагмент SacI-SacI клонировали из плазмиды phAT3ex35-3 в плазмиду phAT3ex35-5 по сайту SacI	2722	phAT3ex35-PB	PstI (введен) – BamHI (введен)
3'-Часть интрона 5 (151 п.н.) + экзон 6 (65 п.н.) + 5'-часть интрона 6 (164 п.н.)	F: catgtgagagtattaggtg R: aggcagaacaatcacaccatg	380	-----	
3'-Часть интрона 5 (13 п.н.) + экзон 6 (65 п.н.) + 5'-часть интрона 6 (164 п.н.)	F: aggg agatct aggtattctct R: aggcagaacaatcacaccatg	362	phATex6-BgSp	BgIII (введен) – SphI (эндогенный)
3'-Часть интрона 6 (145 п.н.) + 5'-часть экзона 7 (216 п.н.), включая стоп-кодон и 39 п.н. из 3'-НТО	F: tctggagggaatccaacctg R: aaacsaanaataggagaaggtgctc	361	-----	
3'-Часть интрона 6 (145 п.н.) + 5'-часть экзона 7 (209 п.н.), включая стоп-кодон и 32 п.н. из 3'-НТО	F: tctggagg gcatg caacctg R: cagctg acasaatctagaaataggagaag-gtgc	354	phATex7-SpXS	SphI (введен) – XbaI/SaII (введены)

Примечание. Сайты узнавания эндонуклеаз рестрикции, введенные в праймеры, выделены полужирным курсивом.

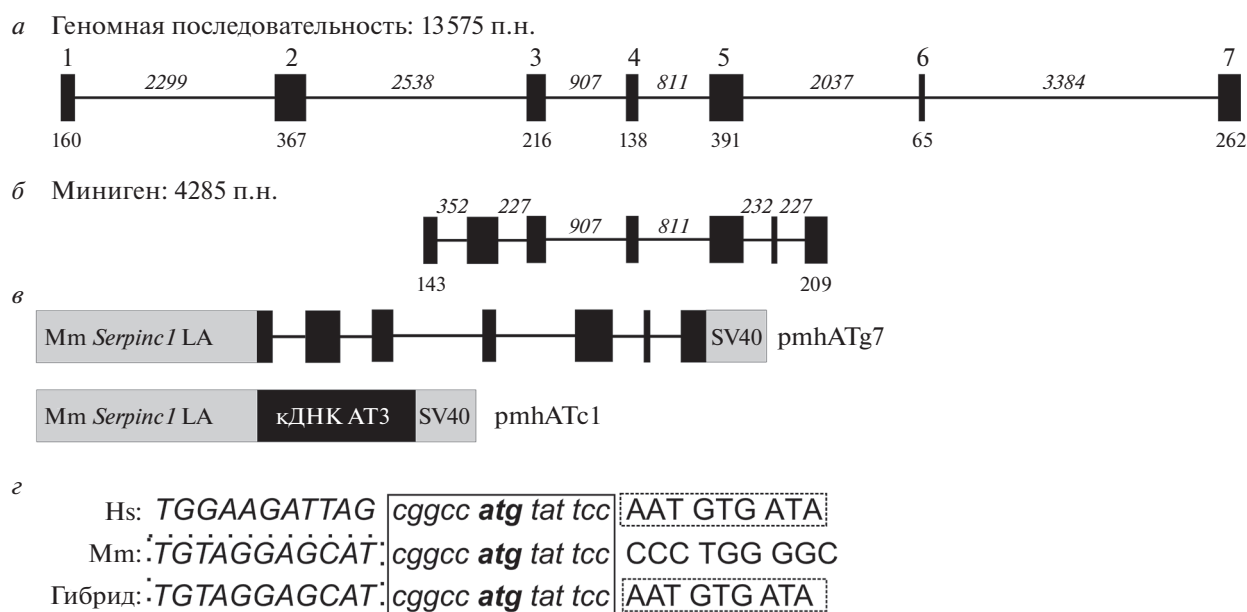


Рис. 1. Схемы экзон-интронной структуры гена *SERPINC1* человека, минигена АТ3, плазмидных конструкций и области формирования гибридной последовательности ДНК. *а* – Схема экзон-интронной структуры гена *SERPINC1* человека. Геномная последовательность экзонов 1–7 имеет длину 13 575 п.н. Экзоны показаны черными прямоугольниками, над и под которыми приведены их номера и длина (п.н.) соответственно. Интроны показаны тонкими черными линиями, длины интронов (п.н.) приведены сверху. *б* – Схема экзон-интронной структуры минигена АТ3. Экзоны и интроны обозначены, как в панели *а*. Указаны длины лишь тех экзонов (1 и 7), размер которых был уменьшен при конструировании минигена. Длины всех интронов минигена указаны сверху. *в* – Схема плазмидных конструкций, кодирующих миниген АТ3 (pmhATg7) и кДНК АТ3 (pmhATc1). Экзоны и кДНК показаны черными прямоугольниками, интроны – тонкими черными линиями, левое плечо геномного локуса АТ3 мыши (LA, left arm) и последовательность полиаденилирования вируса SV40 показаны серыми прямоугольниками. *г* – Нуклеотидные последовательности участка формирования гибридной молекулы между левым плечом геномного локуса АТ3 мыши и первым экзоном или кДНК гена АТ3 человека. Идентичные нуклеотиды выделены курсивом и сплошной рамкой. Кодон инициации трансляции показан полужирным шрифтом. Нуклеотидные последовательности из левого плеча геномного локуса АТ3 мыши, присутствующие в гибридной молекуле, выделены рамкой из точек. Последовательности первого экзона/кДНК АТ3 человека, присутствующие в гибридной молекуле, выделены пунктирной рамкой. Mm – мышь, Hs – человек.

Сборку минигена АТ3 осуществляли путем последовательного клонирования фрагментов, содержащих экзоны и интроны гена *SERPINC1* человека (табл. 1). В частности, фрагмент EcoRI–PstI, содержащий экзон 2, клонировали из плазмиды phATex2-BgP по сайтам EcoRI–PstI в вектор pBluescriptSKII(–) и получили плазмиду phATg1. Затем фрагмент PstI–BamHI плазмиды phATex35-PB, содержащий экзоны 3–5, клонировали в плазмиду phATg1 по сайтам PstI–BamHI и получили плазмиду phATg2. Из плазмиды phATg2 удаляли сайт рестрикции SalI, разрезая ее рестриктазой SalI, обрабатывая фрагментом Кленова с последующим лигированием плазмиды саму на себя. Фрагмент BamHI–NotI, содержащий экзон 7, клонировали из плазмиды phATex7-SpXS в вектор pBluescriptSKII(–) по сайтам BamHI–NotI с получением плазмиды phATg3. После этого фрагмент XhoI–SphI, содержащий экзон 6, клонировали из плазмиды phATex6-BgSp в плазмиду phATg3 по сайтам XhoI–SphI, получая таким образом плазмиду phATg4. Затем фрагмент BglII–NotI, содержащий экзоны 6 и 7, клонировали из плазмиды

phATg4 в плазмиду phATg2 по сайтам BamHI–NotI. В результате получили плазмиду phATg5, которую затем по сайтам XbaI–SalI клонировали последовательность полиаденилирования вируса SV40 из вектора pGL3-Control (“Promega”, США) и получили плазмиду phATg6.

Для получения конечной плазмиды, кодирующей миниген АТ3, фрагмент BglII–SalI, содержащий экзоны 2–7 минигена АТ3 и сигнал полиаденилирования SV40, клонировали из плазмиды phATg6 по сайтам BamHI–SalI в плазмиду pATLA, содержащую гибридную последовательность левого плеча геномного локуса гена *Serpinc1* мыши и первого экзона гена *SERPINC1* человека, и получили плазмиду pmhATg7 (схема плазмидной конструкции показана на рис. 1*в*).

Клонирование кДНК АТ3 человека. кДНК АТ3 человека амплифицировали в виде двух фрагментов на матрице первых цепей кДНК из печени человека. 5'-Область кДНК амплифицировали с использованием олигонуклеотидных праймеров 5'-gcctcagatcacactatctcc-3' и 5'-aacagcaccagaa-cagtggag-3' и клонировали в вектор pCR-Blunt, по-

экзона (AccLoop), или с их комбинацией (DonMut AccLoop).

Далее фрагменты минигена AT3 *NheI*–*Cfr9I* из плазмиды pVK-ATg7NC, содержащей соответствующие мутации (AccMut, DonLoop и AccMut-DonLoop), клонировали в вектор pmhATg7 по сайтам *NheI*–*Cfr9I*. Затем в полученные плазмиды клонировали фрагменты минигена AT3 *Cfr9I*-*EcoRV* из плазмид pVI-ATg7CRV, содержащих мутации DonMut, AccLoop и DonMutAccLoop. Таким образом получили плазмиды, кодирующие миниген с мутациями критических сайтов сплайсинга в первом интроне (Spl), содержащие последовательности для формирования шпилек (Loop) или их комбинацию (Spl Loop).

Культуры клеток и трансфекция. Клетки линии HEK293 культивировали в среде DMEM (“HyClone”, США) с добавлением 10% фетальной бычьей сыворотки (“HyClone”), 100 ед./мл пенициллина и 100 мкг/мл стрептомицина (“Gibco”, США). Для трансфекции клеток плазмидной ДНК использовали реагент EcoTransfect (“OZ Bioscience”, Франция).

Выделение РНК, обратная транскрипция и ПЦР. Клетки HEK293 рассеивали по 5×10^5 клеток в лунку шестилучной планшета, на следующий день трансфицировали соответствующими плазидами (2 мкг на лунку) либо оставляли не трансфицированными в качестве отрицательного контроля. Суммарную РНК выделяли через 48 ч после трансфекции с использованием набора RNeasy Mini Kit (“Qiagen”) и обрабатывали ДНКазой из набора RNase-Free DNase Set (“Qiagen”) для минимизации содержания примесей геномной и плазмидной ДНК в препарате РНК. Реакцию обратной транскрипции проводили в объеме 10 мкл с использованием 0.5 мкг РНК и мастермикса SuperScript™ IV VIL0™ Master Mix with ezDNase™ Enzyme (“Invitrogen”), включая обработку РНК ферментом ezDNase™ в соответствии с протоколом производителя. Продукты реакции обратной транскрипции (первые цепи кДНК) использовали в качестве ПЦР-матрицы.

Полноразмерный транскрипт минигена амплифицировали с использованием ДНК-полимеразы Platinum™ Taq DNA Polymerase High Fidelity (“Invitrogen”, #11304011) и олигонуклеотидных праймеров 5'-atctctgtaggagcctcgcc-3' и 5'-cgactctagaataggaagagg-3', специфичных для транскрипта минигена, в реакции объемом 25 мкл, содержащей 1 × буфер High Fidelity PCR buffer (60 mM Трис-SO₄ (pH 8.9) и 18 mM (NH₄)₂SO₄), 2 mM MgSO₄, 400 нМ dNTP, 400 нМ прямого и обратного праймеров, 1 ед. акт. ДНК-полимеразы и 1 мкл реакционной смеси обратной транскрипции в качестве матрицы. ПЦР проводили в следующих условиях: 94°C, 2 мин; 35 циклов: 94°C, 30 с, 50°C, 30 с и 68°C, 2 мин. Продукты амплифика-

ции разделяли в 0.9%-ном ТАЕ-агарозном геле и визуализировали, окрашивая бромистым этидием. Результаты электрофореза документировали и анализировали с помощью системы гель-документации ChemiDoc XRS (“Bio-Rad”, США) и программного обеспечения QuantityOne 4.6.2 (“Bio-Rad”). Для анализа нуклеотидных последовательностей транскриптов минигена фрагменты ДНК очищали из агарозного геля с помощью набора QiaexII Gel Extraction kit (“Qiagen”), клонировали в вектор pJET1 (“Thermo Fisher Scientific”) и секвенировали.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Конструирование минигена, кодирующего антитромбин III человека

Для валидации стратегии минимизации нативных интронов, используемых в генетической конструкции для трансгенеза, сконструирован миниген, кодирующий белок AT3 человека. Ген *SERPINC1* человека (Gene ID: 432), кодирующий AT3, расположен на хромосоме 1 и содержит семь экзонов и шесть интронов. Длина геномной последовательности, включающей первый и последний экзоны гена, составляет 13575 п.н. В то же время, длина кодирующей последовательности AT3 (изоформа 1, NM_000488.3) составляет 1392 п.н. (464 аминокислотных остатка). При этом наиболее протяженные первый, второй, пятый и шестой интроны гена *SERPINC1* имеют длину 2299, 2538, 2037 и 3384 п.н. соответственно. Считается, что основные элементы интронов, обеспечивающие правильный сплайсинг транскрипта, расположены в пределах 100 п.н. от донорного и акцепторного сайтов [9, 10]. Поэтому мы использовали стратегию конструирования минигена, которая предполагает удаление большей части протяженных интронов гена *SERPINC1* с сохранением небольших участков этих интронов (около 100–150 п.н.), прилежащих к сайтам сплайсинга. Третий и четвертый интроны гена *SERPINC1* имеют относительно небольшую протяженность (907 и 811 п.н. соответственно), поэтому для создания минигена использовали интактные нуклеотидные последовательности этих интронов (рис. 1а, б).

Для проверки функциональности минигена AT3 на трансгенных животных в конструкцию, кодирующую миниген, ввели последовательность из геномного локуса AT3 мыши, включающую протяженный первый интрон гена *Serpinc1*. Эта последовательность содержит регуляторные элементы для активации транскрипции минигена в составе плазмид при транзитной трансфекции в клетки млекопитающих. Она может быть использована как левое плечо гомологии для встраивания минигена AT3 человека в геномный локус AT3 мыши с

помощью CRISPR/Cas9-опосредованной гомологичной рекомбинации. Для этого создана конструкция, содержащая геномную последовательность локуса АТЗ мыши, слитую с первым экзоном гена АТЗ человека (рис. 1а). В результате первый экзон минигена АТЗ длиной 143 п.н. включает 97 п.н. 5'-области второго экзона гена *Serpinc1* мыши, 14 п.н. вблизи кодона инициации трансляции, идентичных у человека и мыши, и 32 п.н. из 3'-области первого экзона гена *SERPINC1* человека (рис. 1б, в). Таким образом сконструирован миниген для экспрессии АТЗ человека под контролем регуляторных последовательностей гена *Serpinc1* мыши. Этот миниген содержит минимизированные первый, второй, пятый и шестой интроны гена *SERPINC1* и последовательность полиаденилирования вируса SV40 (рис. 1б, в). При этом длина белоккодирующей части генетической конструкции (включая экзоны 1–7) уменьшилась в 3 раза (13575–4285 п.н.). Сходным образом создана контрольная плазмида, несущая кДНК АТЗ человека под контролем регуляторных элементов гена *Serpinc1* мыши. Схемы конструкций, кодирующих миниген АТЗ человека (плазмида pmhATg7) и кДНК АТЗ человека (pmhATc1), показаны на рис. 1в.

Анализ паттерна сплайсинга транскрипта минигена антитромбина III

Для анализа функциональности минигена и правильности сплайсинга первичного транскрипта, кодируемого минигеном, плазмидные конструкции трансфицировали в клетки линии НЕК293, выделяли суммарную РНК, проводили ОТ-ПЦР и амплифицировали полноразмерный транскрипт АТЗ с помощью ПЦР с праймерами, специфичными к транскрипту АТЗ, кодируемому минигеном. Обнаружено, что с плазмиды pmhATc1 экспрессируется транскрипт ожидаемого размера 1457 п.н., тогда как с плазмиды pmhATg7 экспрессируются три транскрипта (рис. 2б, дорожка pmhATg7), один из которых по размеру соответствует корректному транскрипту АТЗ (рис. 2б, положение полосы в геле указано *). Однако обнаружен также длинный транскрипт (рис. 2б, положение полосы в геле указано **) и укороченный транскрипт длиной порядка 800 п.н. (рис. 2б, положение полосы в геле указано ***). Индивидуальные полосы были выделены из геля, клонированы и секвенированы. Установлено, что эта полоса (*) действительно представляет собой правильно сплайсированный транскрипт минигена АТЗ, полоса (**) представляет собой транскрипт минигена, содержащий aberrantный экзон длиной 98 н., расположенный между первым и вторым экзонами минигена, а полоса (***) соответствует альтернативно сплайсированному транскрипту, у которого пропущены второй и третий экзоны. При

этом в обнаруженных альтернативных транскриптах нарушена рамка считывания белка АТЗ.

В базе данных GenBank представлены две изоформы мРНК АТЗ человека. Изоформа 1 (NM_000488.3) кодирует полноразмерный белок (464 аминокислотных остатка), тогда как изоформа 2 (NM_001365052 или XM_005245198) предположительно может кодировать укороченный белок, из которого удалены 48 N-концевых аминокислот, а трансляция начинается с остатка метионина в положении 49 (NP_001351981.1). Анализ нуклеотидной последовательности транскрипта NM_001365052 выявил aberrantный экзон длиной 169 п.н., локализованный между первым и вторым экзонами гена *SERPINC1*, который нарушает рамку считывания АТЗ за счет внесения нескольких стоп-кодонов. Таким образом, альтернативным иницирующим кодоном может быть лишь метионин в позиции 49. Однако несмотря на присутствие транскрипта NM_001365052 и белка NP_001351981.1 в базах данных, нам не удалось найти опубликованных экспериментальных доказательств обнаружения этих молекул.

Сравнение транскрипта NM_001365052 и выявленного нами альтернативного транскрипта с aberrantным экзонами длиной 98 п.н. показало, что донорный сайт сплайсинга aberrantного экзона транскрипта NM_001365052 совпадает с донорным сайтом сплайсинга aberrantного экзона длиной 98 п.н. из первого интрона минигена. Таким образом, мы обнаружили, что первый интрон гена *SERPINC1* человека действительно содержит криптографический сайт сплайсинга, который в определенных условиях может функционировать как донорный, приводя к инсерции aberrantного экзона в транскрипт. Мы не смогли проверить функциональность акцепторного сайта сплайсинга aberrantного экзона из транскрипта NM_001365052, так как область интрона, включающая данный сайт, была удалена при конструировании минигена.

Оптимизация паттерна сплайсинга транскрипта минигена антитромбина III

Нами проведен анализ нуклеотидной последовательности первого интрона минигена длиной 356 п.н. и фланкирующих экзонных последовательностей с помощью программы Human Splicing Finder v.3.1 (<http://www.umd.be/HSF3>). Установлено, что среди всех обнаруженных потенциальных сайтов сплайсинга наибольшее консенсусное значение (число от 0 до 100, значимыми считаются результаты с показателем 80 и более) имеют донорный сайт первого экзона минигена и акцепторный сайт второго экзона, которые являются сайтами сплайсинга корректного транскрипта гена *SERPINC1* (табл. 2). При этом криптографические сайты сплайсинга в интроне минигена также имели высокие консенсусные значения (табл. 2). Можно предполо-

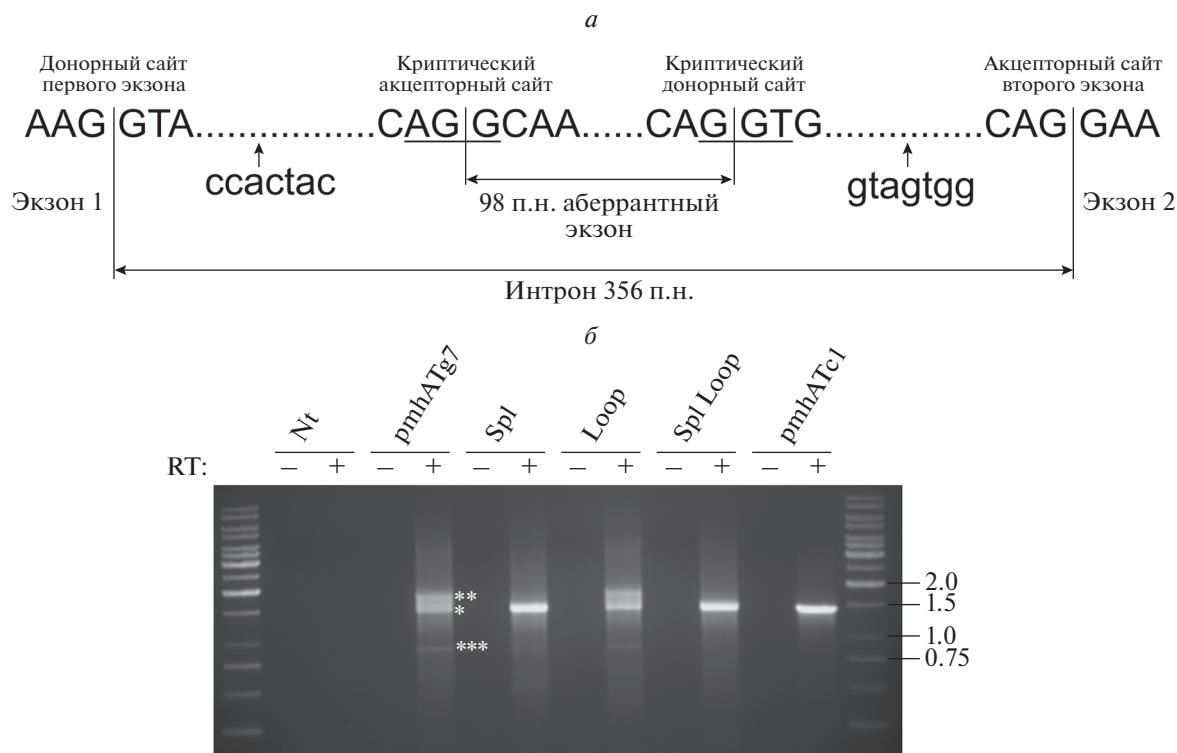


Рис. 2. Анализ и оптимизация паттерна сплайсинга минигена AT3 человека в клетках линии HEK293. *a* – Схема сайтов сплайсинга первого интрона минигена. Показаны нуклеотидные последовательности донорного сайта сплайсинга первого экзона минигена, акцепторного сайта сплайсинга второго экзона минигена, критических акцепторного и донорного сайтов сплайсинга в первом интроне минигена. Вертикальные полосы показывают границы экзонов и интронов. Длина первого интрона минигена – 356 п.н., длина aberrantного экзона – 98 п.н. Подчеркнуты нуклеотидные последовательности AGG и GGT, соответствующие критическим сайтам сплайсинга, которые при оптимизации паттерна сплайсинга заменили на TCT и CTA соответственно. Места встраивания комплементарных последовательностей для формирования стабильной шпильчатой структуры (ccactac, через 21 п.н. от границы первого экзона и интрона; gtagtgg, за 18 п.н. до точки ветвления вблизи акцепторного сайта второго экзона) показаны стрелками. *b* – Электрофоретический анализ в 0.9%-ном TAE-агарозном геле амплифицированных в ПЦР транскриптов минигена AT3, детектируемых в клетках HEK293. Клетки HEK293 трансфицировали плазмидами для экспрессии минигена AT3 (pmhATg7), кДНК AT3 (pmhATc1) и минигена с мутациями критических сайтов сплайсинга (Spl) с введенными в первый интрон последовательностями для формирования шпилек (Loop) или с их комбинацией (Spl Loop). Nt – не-трансфицированные клетки. Положение фрагментов ДНК, соответствующее различным транскриптам, показано звездочками. * Корректно сплайсированный транскрипт AT3 (ожидаемый размер продукта ПЦР 1457 п.н.). ** Транскрипт с инсерцией 98 п.н. в первом интроне минигена. *** Укороченный транскрипт с пропущенными экзонами 2 и 3. В качестве матрицы в ПЦР использовали образцы, полученные в реакции обратной транскрипции с добавлением (RT+) или без добавления (RT-) обратной транскриптазы. Справа указаны длины фрагментов ДНК-маркера O'GeneRuler 1kb DNA Ladder ("Thermo Fisher Scientific") (т.п.н.).

жить, что удаление большей части первого интрона гена *SERPINC1* позволяет эффективно протекать альтернативному сплайсингу между донорным сайтом первого экзона минигена и критическим акцепторным сайтом в первом интроне, и, соответственно, между критическим донорным сайтом в первом интроне и акцепторным сайтом во втором экзоне минигена.

Для оптимизации паттерна сплайсинга минигена AT3 использовали два подхода: (а) сайт-направленный мутагенез критических сайтов сплайсинга в первом интроне минигена; (б) введение в первый интрон вблизи корректных донорного и акцепторного сайтов сплайсинга комплементарных нуклеотидных последовательностей, потенци-

ально способных формировать вторичные структуры (шпильки) в молекуле пре-мРНК, тем самым стерически сближая сайты сплайсинга и способствуя правильному протеканию сплайсинга [9]. С помощью сайт-направленного мутагенеза созданы плазмидные конструкции с мутациями критических сайтов сплайсинга (Spl), с последовательностями для формирования вторичных структур (Loop) и с их комбинацией (Spl Loop). Анализ с помощью программы Human Splicing Finder v.3.1, проведенный после внесения мутаций и введения последовательностей для формирования шпилек, показал, что оба критических сайта сплайсинга успешно удалены из интрона (данные не показаны). Анализ паттерна сплайсинга в

Таблица 2. Анализ нуклеотидной последовательности первого интрона минигена с помощью программы Human Splicing Finder v.3.1

Позиция нуклеотида	Тип сайта сплайсинга	Потенциальный сайт сплайсинга	Консенсусное значение (0–100)
25	Донорный	AAGgtaaga	96.71
34	Донорный	GGGgtgagc	90.05
82	Акцепторный	aaggactcacagGA	81.08
129	Акцепторный	tctccctgtagTT	80.09
135	Акцепторный	tggtagttacagTC	81.69
172	Акцепторный*	gaccacgtccagGC	84.1
242	Акцепторный	cttggttcaaagGA	80.43
270	Акцепторный	tgccacaccagGT	85.95
279	Донорный*	CAGgtgggc	88.33
318	Донорный	GTGgtgggc	81.82
368	Акцепторный	gtctcctgcagGA	94.08

Примечание. Показаны результаты с консенсусным значением выше 80. Экзонные последовательности выделены прописным шрифтом. Донорный сайт первого экзона минигена и акцепторный сайт второго экзона показаны полужирным. Криптические сайты сплайсинга, используемые в альтернативном транскрипте со вставкой aberrантного экзона длиной 98 п.н., показаны полужирным курсивом и отмечены звездочкой.

случае плазмидных конструкций, транзитно трансфицированных в клетки HEK293, позволил установить, что мутагенез криптических сайтов сплайсинга в первом интроне минигена эффективно предотвращает формирование альтернативных транскриптов, включая формирование укороченного транскрипта с пропущенными вторым и третьим экзонами (рис. 2б, дорожки Spl или Spl Loop). Введение комплементарных последовательностей для формирования шпилек не влияло существенно на образование альтернативных транскриптов (рис. 2б, дорожка Loop).

Таким образом, путем удаления наиболее протяженных интронов гена *SERPINC1* человека с сохранением последовательностей длиной около 150 п.н. вблизи сайтов сплайсинга нами был сконструирован функциональный миниген АТЗ человека. Анализ характера сплайсинга минигена АТЗ в клетках млекопитающих выявил один корректный транскрипт и два альтернативных со вставкой aberrантного экзона между первым и вторым экзонами минигена и с пропущенными вторым и третьим экзонами. Оптимизация паттерна сплайсинга минигена с помощью сайт-направленного мутагенеза криптических сайтов в первом интроне позволила эффективно предотвратить формирование альтернативных транскриптов, тогда как подход, основанный на стерическом сближении корректных сайтов сплайсинга путем введения последовательностей для формирования вторичных структур в пре-мРНК, оказался неэффективным. Кроме того, показана функциональность одного из криптических сайтов сплайсинга, находящихся в первом интроне гена *SERPINC1* человека. Анализ нуклеотидной

последовательности минимизированного первого интрона минигена с помощью программы Human Splicing Finder v.3.1 выявил присутствие в нем криптических сайтов сплайсинга. Таким образом, возможные сайты альтернативного сплайсинга минигена в минимизированных или гибридных интронах можно выявить еще до экспериментальной проверки паттерна сплайсинга и эффективно удалить их с помощью мутагенеза.

В целом, нами показана возможность дизайна минигенов путем удаления наиболее протяженных интронов. Описанный подход вкупе с биоинформатическим анализом нуклеотидных последовательностей интронов минигена может применяться при конструировании минигенов, используемых для получения трансгенных животных с продукцией экономически значимых рекомбинантных белков в молоке. Кроме того, минигены с интронами гена *SERPINC1* человека могут использоваться для изучения влияния интронных мутаций на характер сплайсинга транскрипта АТЗ, которые встречаются у пациентов с недостаточностью АТЗ.

В работе использована инфраструктура Центра коллективного пользования Института биологии гена Российской академии наук “Биология живой клетки и биомедицинские нанотранспортеры лекарств”.

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 16-14-00150).

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Luxembourg B., Delev D., Geisen C., Spannagl M., Krause M., Miesbach W., Heller C., Bergmann F., Schmeink U., Grossmann R., Lindhoff-Last E., Seifried E., Oldenburg J., Pavlova A. (2011) Molecular basis of antithrombin deficiency. *Thromb. Haemost.* **105**, 635–646.
2. Konkle B.A., Bauer K.A., Weinstein R., Greist A., Holmes H.E., Bonfiglio J. (2003) Use of recombinant human antithrombin in patients with congenital antithrombin deficiency undergoing surgical procedures. *Transfusion.* **43**, 390–394.
3. Paidas M.J., Triche E.W., James A.H., DeSancho M., Robinson C., Lazarchick J., Ornaghi S., Frieling J. (2016) Recombinant human antithrombin in pregnant patients with hereditary antithrombin deficiency: integrated analysis of clinical data. *Am. J. Perinatol.* **33**, 343–349.
4. Shepelev M.V., Kalinichenko S.V., Deykin A.V., Korobko I.V. (2018) Production of recombinant proteins in the milk of transgenic animals: current state and prospects. *Acta Naturae.* **10**, 40–47.
5. Lavine G. (2009) FDA approves first biological product derived from transgenic animal. *Am. J. Health Syst. Pharm.* **66**(6), 518. doi 10.2146/news090023
6. Edmunds T., Van Patten S.M., Pollock J., Hanson E., Bernasconi R., Higgins E., Manavalan P., Ziomek C., Meade H., McPherson J.M., Cole E.S. (1998) Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin. *Blood.* **91**, 4561–4571.
7. Palmiter R.D., Sandgren E.P., Avarbock M.R., Allen D.D., Brinster R.L. (1991) Heterologous introns can enhance expression of transgenes in mice. *Proc. Natl. Acad. Sci. USA.* **88**, 478–482.
8. Choi T., Huang M., Gorman C., Jaenisch R. (1991) A generic intron increases gene expression in transgenic mice. *Mol. Cell. Biol.* **11**, 3070–3074.
9. Raker V.A., Mironov A.A., Gelfand M.S., Pervouchine D.D. (2009) Modulation of alternative splicing by long-range RNA structures in *Drosophila*. *Nucl. Acids Res.* **37**, 4533–4544.
10. Gelfand M.S. (1989) Statistical analysis of mammalian pre-mRNA splicing sites. *Nucl. Acids Res.* **17**, 6369–6382.

HUMAN ANTITHROMBIN III MINIGENE WITH OPTIMIZED SPLICING PATTERN

M. V. Shepelev¹*, E. K. Saakian¹, S. V. Kalinichenko¹, and I. V. Korobko¹

¹*Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia*

*e-mail: mshepelev@mail.ru

Antithrombin III (AT3) belongs to the superfamily of serine protease inhibitors and is a major anticoagulant under physiological conditions. As a model we used *SERPINC1* gene, encoding human AT3 protein that is valuable for medicine and biotechnology, and constructed minigene by minimizing the size of lengthy introns and preserving splicing site-flanking sequences. Analysis of splicing pattern of minigene allowed identifying correct AT3 transcript and two alternatively spliced transcripts that were formed due to skipping of second and third exons of minigene or due to insertion of an aberrant exon via usage of cryptic splicing sites in the first intron of minigene. Splicing pattern of AT3 minigene transcript was successfully optimized via site-directed mutagenesis of cryptic splicing sites thus completely preventing alternative transcripts generation. Analysis with Human Splicing Finder v.3.1 software confirmed the presence of cryptic splicing sites in the first intron of minigene thus demonstrating the principal opportunity to identify putative alternative splicing sites in minimized or hybrid introns in minigenes and allowing for deletion of such sites via mutagenesis before experimentally testing the splicing pattern of the minigene. The described approach to design of minigenes together with the bioinformatical analysis of the nucleotide sequences of minigene introns could be used in the field of minigene construction for generation of transgenic animals producing economically valuable proteins in the milk.

Keywords: minigene, antithrombin III, splicing pattern, aberrant exon, alternative transcript