

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

На правах рукописи

Маслакова Айтсана Алексеевна
Структурно-функциональный анализ транскриптов
гена *SERPINA1*: поиск альтернативных продуктов
трансляции - изоформ и С-концевых пептидов альфа1-
антитрипсина человека

1.5.3 «Молекулярная биология»

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата биологических наук

Научный руководитель:
кандидат биологических наук
Рубцов Михаил Александрович

Москва-2023

Оглавление

СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ	4
1. ВВЕДЕНИЕ	6
АКТУАЛЬНОСТЬ ТЕМЫ ИССЛЕДОВАНИЯ И СТЕПЕНЬ ЕЕ РАЗРАБОТАННОСТИ	6
ЦЕЛЬ И ЗАДАЧИ ДИССЕРТАЦИОННОЙ РАБОТЫ	8
НАУЧНАЯ НОВИЗНА	9
НАУЧНАЯ И ПРАКТИЧЕСКАЯ ЗНАЧИМОСТЬ.....	10
ЛИЧНЫЙ ВКЛАД АВТОРА.....	10
МЕТОДОЛОГИЯ И МЕТОДЫ ИССЛЕДОВАНИЯ	11
ПОЛОЖЕНИЯ, ВЫНОСИМЫЕ НА ЗАЩИТУ.....	11
СТЕПЕНЬ ДОСТОВЕРНОСТИ РЕЗУЛЬТАТОВ И АПРОБАЦИЯ РАБОТЫ.....	11
ПУБЛИКАЦИИ.....	11
УЧАСТИЕ В КОНФЕРЕНЦИЯХ.....	12
2. ОБЗОР ЛИТЕРАТУРЫ	14
2.1. АЛЬФА1-АНТИТРИПСИН. КРАТКИЙ ОБЗОР СУПЕРСЕМЕЙСТВА СЕРПИНОВ.....	14
2.2. ОСОБЕННОСТИ СИНТЕЗА И “ЖИЗНЕННЫЙ ЦИКЛ” АЛЬФА1-АНТИТРИПСИНА.....	17
2.2.1. Структура гена <i>SERPINA1</i> . Структура транскриптов гена <i>SERPINA1</i> . Регуляция экспрессии и тканеспецифическая экспрессия гена <i>SERPINA1</i>	17
2.2.2. Структура и фолдинг полноразмерного ААТ. Гликоформы ААТ. Конформеры ААТ ..	21
2.2.3. Молекулярные механизмы антипротеолитической активности ААТ. Механизм утилизации комплекса ААТ-протеаза	24
2.3. МОЛЕКУЛЯРНЫЕ МЕХАНИЗМЫ АНТИТРИПСИНОВОЙ НЕДОСТАТОЧНОСТИ ААТ НА ПРИМЕРЕ Z-ВАРИАНТА ААТ.....	28
2.4. МНОЖЕСТВЕННОСТЬ “ПОБОЧНЫХ” ФУНКЦИЙ ААТ.....	31
2.4.1. ААТ ингибирует протеазы из других семейств.....	31
2.4.2. Биологические функции ААТ, не зависящие от анти-протеолитической активности	33
2.5. ААТ - ОПУХОЛЕВЫЙ БИОМАРКЕР	41
2.6. С-КОНЦЕВОЙ ДОМЕН ААТ, КОДИРУЕМЫЙ ПЯТЫМ ЭКЗОНОМ ГЕНА <i>SERPINA1</i> , - ИСТОЧНИК БИОЛОГИЧЕСКИ АКТИВНЫХ ПЕПТИДОВ.....	45
3. МАТЕРИАЛЫ И МЕТОДЫ	56
3.1. КУЛЬТИВИРОВАНИЕ КЛЕТОЧНЫХ ЛИНИЙ И ПЕРВИЧНЫХ КУЛЬТУР	56
3.2. ПЦР В РЕАЛЬНОМ ВРЕМЕНИ (ПЦР-РВ).....	56
3.3. ОПРЕДЕЛЕНИЕ СОДЕРЖАНИЯ ААТ В СЕКРЕТОМАХ КЛЕТОЧНЫХ ЛИНИЙ МЕТОДОМ ИММУНОБЛОТТИНГ	58
3.4. НОЗЕРН ГИБРИДИЗАЦИЯ	59
3.4.1. Конструирование ДНК-матриц для синтеза кРНК.....	59
3.4.2. Синтез смысловой кРНК	60
3.4.3. Синтез радиоактивно меченого антисмыслового кРНК зонда	60
3.4.4. Препараты суммарной РНК клеточных линий и тканей.....	60
3.4.5. Электрофоретическое разделение РНК	61
3.4.6. Перенос РНК на мембрану.....	61
3.4.7. Гибридизация.....	61
3.5. ДЕТЕКЦИЯ ЭКСПРЕССИИ ГЕНА <i>SERPINA1</i> С ПРОТИВОПОЛОЖНОЙ ЦЕПИ ДНК В ЛОКУСЕ ЭКЗОНА 5.....	61
3.6. 3'-RACE (RAPID AMPLIFICATION OF cDNA ENDS)	62
3.7. 5'-RACE	62
3.8. ИЗВЛЕЧЕНИЕ И ОБРАБОТКА ПИКОВ FANTOM5 CAGE И ДАННЫХ ENCODE CHIP-SEQ.....	63
3.9. АНАЛИЗ ПОТЕНЦИАЛА К ТРАНЛЯЦИИ АЛЬТЕРНАТИВНЫХ ОРС ИНДИВИДУАЛЬНЫХ ДЛИННЫХ И КОРОТКИХ ТРАНСКРИПТОВ ГЕНА <i>SERPINA1</i>	63
3.9.1. Конструирование плазмид.....	63
3.9.2. Трансфекция.....	65
3.9.3. Подготовка секретомов.....	65
3.9.4. Проточная цитофлуориметрия и сортировка клеток.....	65
3.9.5. Подготовка лизатов клеток.....	65
3.9.6. Иммуноблоттинг.....	65
3.9.7. Дегликозилирование белков ферментом PNFaseF.....	66

3.9.8.	Сайт-направленный мутагенез.....	67
3.10.	ИЗУЧЕНИЕ ВНУТРИКЛЕТОЧНОЙ ЛОКАЛИЗАЦИИ ХИМЕРНЫХ ПРОДУКТОВ ТРАНСЛЯЦИИ	67
3.10.1.	Конфокальная микроскопия на живых клетках	67
3.10.2.	Конфокальная микроскопия на препаратах фиксированных клеток.....	68
3.11.	ДЕТЕКЦИЯ И ИЗУЧЕНИЕ ВНУТРИКЛЕТОЧНОЙ ЛОКАЛИЗАЦИИ ПРЕДПОЛАГАЕМЫХ ЭНДОГЕННЫХ ПРОДУКТОВ ТРАНСЛЯЦИИ.....	68
3.11.1.	С длинных транскриптов <i>SERPINA1</i>	68
3.11.2.	С коротких транскриптов <i>SERPINA1</i>	69
3.12.	ЭВОЛЮЦИОННЫЙ АНАЛИЗ.....	71
3.12.1.	Сборка и выравнивание последовательностей белков	71
3.12.2.	Тестирование на давление отбора	71
3.12.3.	Оценка максимального правдоподобия и тест отношения правдоподобия	71
3.13.	НОКАУТ ЭКЗОНА 5 ГЕНА <i>SERPINA1</i> С ПОМОЩЬЮ ТЕХНОЛОГИИ CRISPR/Cas9	72
4.	РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ.....	74
4.1.	АНАЛИЗ ПОТЕНЦИАЛЬНЫХ АЛЬТЕРНАТИВНЫХ САЙТОВ ИНИЦИАЦИИ ТРАНСЛЯЦИИ В РАМКЕ С ГЛАВНОЙ ОПС <i>SERPINA1</i>	76
4.2.	ВЫБОР МОДЕЛЬНОЙ КЛЕТОЧНОЙ ЛИНИИ.....	78
4.2.1.	Определение относительного уровня экспрессии гена <i>SERPINA1</i> в клеточных линиях, первичной культуре и тканях человека методом ПЦР в реальном времени.....	78
4.2.2.	Детекция длинных и коротких транскриптов <i>SERPINA1</i> с помощью нозерн-гибридизации	80
4.2.3.	Экспрессия антисмысловой цепи ДНК в локусе пятого экзона гена <i>SERPINA1</i> с образованием предполагаемого антисмыслового транскрипта	82
4.2.4.	Антисмысловой транскрипт - возможный положительный регулятор синтеза ААТ. Содержание ААТ в секретах клеточных линий.....	85
4.3.	ИДЕНТИФИКАЦИЯ СТРУКТУРЫ 3'- И 5'-КОНЦОВ ДЛИННЫХ И КОРОТКИХ ТРАНСКРИПТОВ.....	87
4.3.1.	Идентификация структуры 3'-концов транскриптов методом 3'-RACE.....	87
4.3.2.	Идентификация структуры 5'-концов транскриптов методом 5'-RACE.....	91
4.4.	ДОСЬЕ <i>SERPINA1</i> , ИЗВЛЕЧЕННОЕ ИЗ ДАННЫХ CAGE И CHIP-SEQ, ВЫЯВЛЯЕТ ВОЗМОЖНЫЙ МЕХАНИЗМ ОБРАЗОВАНИЯ КОРОТКИХ ТРАНСКРИПТОВ.....	96
4.5.	ИЗУЧЕНИЕ ТРАНСЛЯЦИИ ИНДИВИДУАЛЬНЫХ ИЗОФОРМ ДЛИННОЙ мРНК <i>SERPINA1</i> В ЖИВЫХ КЛЕТКАХ <i>IN VITRO</i> : ПОИСК ПРОДУКТОВ ТРАНСЛЯЦИИ И ИДЕНТИФИКАЦИЯ АЛЬТЕРНАТИВНЫХ ОПС	100
4.5.1.	С длинных транскриптов <i>SERPINA1</i> синтезируется предполагаемая несекреторная изоформа ААТ.....	100
4.5.2.	Сайт-направленный мутагенез плазмидных конструкций ранжирует внутриклеточную протеоформу в несекреторную изоформу ААТ	104
4.5.3.	Альтернативный сайт инициации трансляции <i>SERPINA1</i> и размер альтернативной ОПС консервативны среди млекопитающих	106
4.5.4.	Изучение внутриклеточной локализации химерных продуктов трансляции с длинных транскриптов и предполагаемой эндогенной внутриклеточной изоформы ААТ. 107	
4.6.	ИЗУЧЕНИЕ ТРАНСЛЯЦИИ ИНДИВИДУАЛЬНЫХ КОРОТКИХ ТРАНСКРИПТОВ <i>SERPINA1</i> В ЖИВЫХ КЛЕТКАХ <i>IN VITRO</i> : ПОИСК ПРОДУКТОВ ТРАНСЛЯЦИИ	113
4.6.1.	Короткие транскрипты <i>SERPINA1</i> транслируются в живых клетках.....	113
4.6.2.	Сочетание 3'-проксимальных альтернативных ОПС <i>SERPINA1</i> представлено у приматов. икОПС <i>SERPINA1</i> - объект текущей эволюции млекопитающих	118
4.6.3.	Изучение внутриклеточной локализации химерных продуктов трансляции с коротких транскриптов и предполагаемых эндогенных С-концевых пептидов ААТ	120
	ЗАКЛЮЧЕНИЕ	126
	ВЫВОДЫ.....	127
	ПРИЛОЖЕНИЕ.....	128
	СПИСОК ЛИТЕРАТУРЫ	141
	БЛАГОДАРНОСТИ.....	167

Список используемых сокращений

русскоязычные сокращения

5'-ОРС - кОРС, расположенная в 5'-НТО
5'/3'-НТО - 5'/3'-нетранслируемая область
а.о. - аминокислотный остаток
ААТ - альфа1-антитрипсин
ААТр - очищенный альфа1-антитрипсин из плазмы крови человека
альтОРС - альтернативная ОРС
г/л - грамм/литр
Е5-эпитоп - эпитоп, кодируемый экзоном 5
ед/мл - единиц/миллилитр
икОРС - изоформная короткая ОРС
ИФА - иммуноферментный анализ
кДНК - комплементарная ДНК
Ки/ммоль - кюри/миллимоль
кОРС - короткая ОРС
кРНК - комплементарная РНК
мин - минута
мкг/мл - микрограмм/миллилитр
мкКи/мкл - микрокюри/микролитр
мкл - микролитр
мкм - микрометр
мкМ – микромоль/литр
мРНК - матричная РНК
н.о. - нуклеотидный остаток
нм - нанометр
ОРС - открытая рамка считывания
п.н. — пара (пары) нуклеотидов
ПЦР — полимеразная цепная реакция
ПЦР-РВ - ПЦР в реальном времени
с - секунда
см - сантиметр

Сп - название С-концевого пептида
ААТ соответствует n-ному количеству остатков аминокислот с С-конца белка
т.п.н. - тысяч пар нуклеотидов
ЭПР - эндоплазматический ретикулум

англоязычные сокращения

5'/3'-RACE - 5'/3'-Rapid Amplification of cDNA ends
APA - alternative polyadenylation site
BSA - bovine serum albumin
CAGE - Cap Analysis of Gene Expression
Cas9 — CRISPR-associated protein 9
cpm - counts per minute (количество импульсов в минуту)
CRISPR — clustered regularly interspaced short palindromic repeats
DAGI - donkey anti-goat IgG
DAMI - donkey anti-mouse IgG
DAPI - 4',6-diamidino-2-phenylindole
DARI - donkey anti-rabbit IgG
DLT - DU145 long transcript
DTT - dithiothreitol
E. coli — *Escherichia coli*
EDTA - ethylenediaminetetraacetic acid
eGFP — enhanced green fluorescent protein
EGTA - ethylene glycol tetraacetic acid
EX - экзон
Hepes- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLT - HepG2 long transcript
HRP - horseradish peroxidase
IgG - иммуноглобулин класса G
KLH - keyhole limpet hemocyanin
NGt(Rb)IgG - normal goat(rabbit) IgG
NHS - N-hydroxysuccinimide

PAS - polyadenylation signal
PBS — phosphate-buffered saline
pI - isoelectric point (изоэлектрическая точка)
PMSF - phenylmethylsulfonyl fluoride
RCL - reactive center loop
SDS - sodium dodecyl sulfate

ST - short transcript
TB - transport buffer
TBS - tris-buffered saline
TMB - 3,3',5,5'-Tetramethylbenzidine
TPM - tags per million
Tris - tris(hydroxymethyl)aminomethan
uORF - upstream open reading frame

1. ВВЕДЕНИЕ

Актуальность темы исследования и степень ее разработанности

Ген *SERPINA1* кодирует альфа1-антитрипсин (ААТ) - главный антипротеолитический агент, поддерживающий в организме баланс протеаза–ингибитор. ААТ является архетипическим представителем надсемейства SERPIN (SERine Protease INhibitor). Основная физиологическая функция ААТ – защита эластических волокон в альвеолах легких от чрезмерного протеолиза эластазой нейтрофилов. К другим функциям ААТ можно отнести противовоспалительную и иммунорегуляторную [1, 2], также показана его противоапоптотическая [3, 4], противовирусная [5, 6, 7] и даже шапероноподобная [8, 9] активности.

Концентрация ААТ в крови максимальна в сравнении с другими известными секреторными серпинами и в норме составляет около 20-30 мкМ (~1,5 г/л) [10], что обеспечивает 90% антипротеолитической емкости сыворотки крови человека [11]. “Фабриками” организма, производящими ААТ на уровне его потребностей (70-80%), являются гепатоциты, а ген *SERPINA1* - один из наиболее интенсивно экспрессируемых генов в печени [12]. ААТ секретируется также моноцитами, макрофагами, нейтрофилами, альвеолярными эпителиальными клетками, клетками эпителия роговицы и кишечника [13, 14, 15, 16, 17].

ААТ является одним из наиболее изученных серпинов, поскольку aberrantная экспрессия гена *SERPINA1* ассоциирована с развитием ряда патологических состояний. Особое внимание исследователей обусловлено тем, что в 1963 году Лаурел и Эриксон обнаружили связь отсутствия ААТ в сыворотке с развитием острой обструктивной болезни легких [18]. С тех пор и, можно считать, по настоящее время изучение серпина в основном сосредоточено на его роли в заболевании человека - альфа1-антитрипсиновой недостаточности. Дефицит ААТ - наиболее распространенная форма *серпинопатии* - заболевания, характеризующегося аномальным внутриклеточным накоплением серпинов вследствие полимеризации молекул. Дефицит ААТ обусловлен мутациями в гене *SERPINA1* и является наследуемым фактором риска развития заболеваний легких и печени (таких, как ранняя хроническая обструктивная болезнь легких, цирроз печени у детей и взрослых) и, реже, рецидивирующего панникулита, системного васкулита, гломерулонефрита и прочих воспалительных и аутоиммунных заболеваний [19]. По данным PubMed количество работ, в которых упоминается ААТ, составляет 5238, из них 1981 - за последние 10 лет. Проблеме антитрипсиновой недостаточности посвящены 2434 работы начиная с 1964 г., из них около 1164 работ опубликовано за последние 10 лет.

Дефицит ААТ может также являться фактором риска развития сахарного диабета 1 и 2 типов [20, 21] и ряда злокачественных новообразований: опухолей печени [22, 23, 24],

лёгких [25], мочевого пузыря [26], желчного пузыря [27], толстой кишки [28, 29] и злокачественной лимфомы [30].

Развитие онкологических заболеваний часто сопровождается активацией (синтез *de novo*) или увеличением экспрессии гена *SERPINA1* и, как следствие, повышением концентрации белка в крови и тканевых жидкостях. Этот феномен выявлен для опухолей печени, поджелудочной железы, легких, простаты, желудка, кишечника, гортани, молочной железы, шейки матки, мочевого пузыря [31, 32, 33, 34, 35], а экспрессия гена *SERPINA1* выявлена *in vitro* в ряде клеточных линий опухолевого происхождения [36, 37, 38]. В PubMed индексируется 835 работ (с начала 1970-х годов), в которых ААТ упоминается в контексте канцерогенеза, из которых всего 261 работ приходится на последние 10 лет, в то время как пик исследований по теме начинается с середины 1990-х годов и продолжается в текущее время.

На основании опубликованных работ изменение экспрессии гена *SERPINA1* можно расценивать как следствие (*или же одной из причин?*) трансформации клетки. ААТ может быть вовлечен в универсальные механизмы канцерогенеза, однако вопрос о роли ААТ в этом процессе не имеет окончательного ответа, что неудивительно, поскольку ААТ является многофункциональным белком, и, возможно, некоторые его функции, а также его изо(протео)формы и их функции остаются неизвестными. Помимо секреторного ААТ, опухолевые клетки, вероятно, производят практически не охарактеризованные внутриклеточные протеоформы. Так, клеточная линия опухоли молочной железы MCF7 производит *предполагаемый* цитоплазматический вариант ААТ, препятствующий аутофагической гибели клеток [39]. ААТ *детектирован* в ядрах клеток на ранней стадии развития опухоли яичника (“опухоли с низким злокачественным потенциалом”) [40]. К сожалению, структура предполагаемых внутриклеточных протеоформ ААТ и механизм их происхождения неизвестны.

Ген *SERPINA1* транскрибируется тканеспецифически с двух альтернативных промоторов [41, 42] в избыточное количество сплайс-изоформ мРНК (более 11 вариантов), отличающихся структурой 5'-НТО, ввиду чего *SERPINA1* входит в 0,5% генов человека, обладающих наибольшей транскрипционной сложностью [43]. Изменение структуры 5'-НТО, наряду с индивидуальной доступностью стартовых кодонов основной и вышерасположенных коротких открытых рамок считывания ввиду различной пространственной структуры РНК, играет роль в посттранскрипционной регуляции экспрессии гена *SERPINA1*, влияя на эффективность трансляции [44]. Считается, что со всех альтернативно сплайсирующихся мРНК транслируется один и тот же белок, состоящий из 394 а.о. (после разрезания сигнального пептида в 24 а.о.) и подвергающийся N-гликозилированию по мере его прохождения секреторного пути в клетке.

Ограниченный протеолиз полноразмерного белка обнажает «спящие» функциональные мотивы в виде С-концевых пептидов ААТ и расширяет многогранные

свойства антитрипсина. ААТ является предшественником ряда пептидов, последовательность которых кодируется экзоном 5 гена *SERPINA1*: опухоль-ассоциированных CRISPP пептидов [45, 46, 47], SPAAT (short piece of alpha 1-antitrypsin) [48, 49, 50], VIRIP (virus-inhibitory peptide) [51], C36 and C41/42 (название соответствует полному количеству остатков аминокислот с С-конца белка) [52, 53, 54], CAAP48 (C-terminal alpha1-antitrypsin peptide 4,789 kDa) [55] и nNIF (neonatal NET-Inhibitory Factor; где NET - Neutrophil Extracellular Trap) [56]. Пептиды ААТ, наряду с синтетическими гомологами C105Y [57] и C26 [58], обладают широким спектром биологической активности, включая иммуномодулирующую, противовирусную и митогенную. Протеолитический путь является единственно известным механизмом образования пептидов.

Открытие новых свойств ААТ и его фрагментов указывает на то, что ему отведена более широкая роль, не ограничивающаяся единственной функцией, заложенной в названии, а изучение особенностей экспрессии гена *SERPINA1*, повышающих разнообразие белковых продуктов, является важной научной задачей. Выявление альтернативных изо(протео)форм ААТ и изучение молекулярных механизмов их происхождения представляет интерес не только для фундаментальной науки, но и для решения прикладных задач, например таких, как разработка современных диагностических и терапевтических средств для лечения заболеваний, возможно связанных не только с дефицитом ААТ, но и имеющих более сложные механизмы развития и фатальные последствия.

Цель и задачи диссертационной работы

Цель работы: идентификация спектра изоформ мРНК гена *SERPINA1*, продуцируемых опухолевыми клеточными линиями человека, и их продуктов трансляции.

Для достижения цели были поставлены следующие **задачи**:

1. Выбор модельной клеточной линии. Скрининг культивируемых клеточных линий человека: определение относительной экспрессии длинных и предполагаемых коротких транскриптов гена *SERPINA1*, содержащих и не содержащих главную ОРС, соответственно, и содержания ААТ в секретах.

2. Идентификация спектра изоформ длинных мРНК *SERPINA1*: детекция в препаратах суммарных РНК и определение структуры их 3'- и 5'-концов.

3. Проверка гипотезы о существовании внутриклеточных изоформ ААТ: трансляция индивидуальных длинных мРНК *SERPINA1* в живых клетках *in vitro*, поиск продуктов трансляции и идентификация альтернативных ОРС.

4. Изучение внутриклеточной локализации химерных продуктов трансляции и предполагаемых эндогенных внутриклеточных изоформ ААТ, транслируемых с длинных транскриптов.

5. Проверка гипотезы об альтернативном механизме образования С-концевых пептидов ААТ - синтезе с коротких транскриптов, содержащих экзон 5 гена *SERPINA1*:

детекция коротких транскриптов в препаратах суммарных РНК и определение структуры их 3'- и 5'-концов.

6. Изучение трансляции индивидуальных коротких транскриптов в живых клетках *in vitro*, поиск продуктов трансляции. Изучение внутриклеточной локализации химерных продуктов трансляции и предполагаемых эндогенных протеоформ ААТ.

Научная новизна

Работа посвящена актуальной теме - исследованию экспрессии гена *SERPINA1*. Считается, что единственным белковым продуктом трансляции транскриптов гена является секреторный гликозилированный альфа1-антитрипсин. Прицельных исследований экспрессии гена на уровне трансляции, а именно возможности синтеза альтернативных изо(протео)форм ААТ, не зависящих от протеолиза белка-предшественника, и условий (структуры транскриптов), при котором этот синтез реализуется, ранее не проводилось. Исходя из структуры гена, а именно наличия ряда АТG-кодонов, распределенных по всей кодирующей области, в рамке с главным и в подходящем для инициации трансляции контексте Козак, мы предположили, что на уровне трансляции возможны два сценария. Поскольку главный стартовый кодон расположен в субоптимальном контексте Козак, с длинных транскриптов, с одной стороны, возможно сканирование с утечкой и инициация с одного из нижележащих альтернативных стартов, а с другой - нельзя исключить инициации с вышележащих неканонических стартов, имеющих в структуре некоторых 5'-НТО. Во втором предполагаемом сценарии реализуются альтернативные старт-кодоны из экзона 5 с образованием биоактивных С-концевых пептидов ААТ. Последнее становится возможным только в случае исключения главной открытой рамки считывания (ОРС), то есть генерирования коротких транскриптов, содержащих экзон 5.

В данной работе мы провели скрининг опухолевых клеточных линий человека методами ПЦР в реальном времени, иммуноблоттинга, а с помощью нозерн-гибридизации детектировали длинные транскрипты *SERPINA1*, содержащие главную ОРС, и впервые - короткие транскрипты, содержащие экзон 5 и не содержащие главную ОРС. На кДНК выбранных клеточных линий DU145 (опухоль эпителия простаты) и HepG2 (опухоль печени) определили первичную структуру 5'-концов длинных и коротких транскриптов, а также показали, что транскрипты в основном полиаденилированы по проксимальному альтернативному сайту полиаденилирования, приводящему к укорочению 3'-НТО.

С помощью подхода *in vitro* мы впервые продемонстрировали, что с длинных транскриптов независимо от структуры 5'-НТО помимо секреторного ААТ транслируется также внутриклеточный и не гликозилированный ААТ с альтернативного внутреннего старт-кодона AUG-2. Химерные продукты трансляции локализуются в разных клеточных компартментах: внутриклеточный ААТ-eGFP, в отличие от секреторного, накапливающегося в ЭПР, мигрирует между ядром и цитоплазмой. Мы детектировали предполагаемый эндогенный внутриклеточный ААТ в ядерных спеклах.

Данные FANTOM5 CAGE (Cap Analysis of Gene Expression), извлеченные для *SERPINA1*, позволили нам предположить, что короткие транскрипты генерируются необычным путем в результате посттранскрипционного расщепления сплайсированной мРНК, транскрибируемой в основном с гепато-специфического промотора. В работе впервые получены доказательства в пользу гипотезы альтернативного происхождения С-концевых пептидов ААТ - трансляции с коротких транскриптов. С помощью подхода *in vitro* мы продемонстрировали, что с коротких гепато-специфических транскриптов транслируется четыре С-концевые протеоформы ААТ, две из которых - с изоформных (по отношению к главной) коротких ОРС в экзоне 5. Химеры более удлиненных С-концевых фрагментов, транслированные со стартов в экзоне 3, не секретируются и, по-видимому, локализируются в цитоплазме. Химеры пептидных протеоформ ААТ не только секретируются, но также проникают и накапливаются в ядрах клеток, в то время как предполагаемые соответствующие эндогенные С-концевые пептиды ААТ - в ядрышках.

Экспрессия гена *SERPINA1* сложнее, чем считалось ранее. Наши данные свидетельствуют о повышении биологического разнообразия продуктов гена *SERPINA1* за счет сканирования с утечкой вдоль длинных транскриптов и инициации трансляции с внутреннего альтернативного старт-кодона либо укорочения транскриптов для реализации 3'-проксимальных альтернативных стартовых кодонов.

Научная и практическая значимость

Полученные в работе данные представляют большой интерес, прежде всего, для фундаментальной науки, и расширяют представления об экспрессии гена *SERPINA1*. Обнаружение ряда альтернативных протеоформ ААТ и механизмов их образования, а в перспективе - их физиологических функций в клетке и в организме в целом, открывает новые возможности для более глубокого понимания развития заболеваний: например, изучения взаимосвязи экспрессии *SERPINA1* и канцерогенеза, а также мутаций в гене *SERPINA1* (в том числе ассоциированных с дефицитом ААТ) и риска развития патологических состояний организма, - что может дать представление о прогрессировании заболевания, его тяжести и индивидуальной изменчивости. Нельзя исключать и возможности использования полученных данных в качестве отправного пункта для последующих исследований, имеющих прикладное значение, вероятнее всего, связанное с развитием диагностических и терапевтических приемов и методов.

Личный вклад автора

Вклад соискателя заключался в выборе предмета и направлений исследования, анализе литературных данных, постановке задач, планировании (соучастие в планировании нозерн-гибридизации, 3'- и 5'-RACE) и проведении экспериментов (за исключением конфокальной микроскопии, клеточного сортирования, конъюгации пептида Е5 с белком-носителем и иммунизации животных при получении антител, дизайна и сборки некоторых плазмидных конструкций и анализа *in silico*), обработке полученных экспериментальных

данных и их интерпретации. Соискатель писал текст всех тезисов и статей, в которых указан первым автором, и занимался их подачей на конференции и в журналы.

Методология и методы исследования

В работе использовались современные методы клеточной (выращивание клеточных культур, клонирование клеток для выведения гомозиготных нокаутов по экзону 5, получение клеточных лизатов и секретомов), молекулярной биологии (ПЦР-РВ, нозерн-гибридизация, 3'- и 5'-RACE, методы молекулярного клонирования, геномное редактирование CRISPR/Cas9), иммунобиохимические методы (иммуноблоттинг, ИФА, получение антител с заданной специфичностью, иммунофлуоресценция), методы цито- и иммунофлуоресцентной микроскопии, прижизненного наблюдения с использованием современного оборудования, методы математического и статистического анализа.

Положения, выносимые на защиту

1. Ген *SERPINA1* экспрессируется в опухолевых клеточных линиях человека в длинные и короткие транскрипты, содержащие и не содержащие главную ОРС, соответственно.

2. С длинных транскриптов *SERPINA1* помимо секреторной синтезируется внутриклеточная изоформа ААТ за счет инициации трансляции с альтернативного стартового кодона.

3. Внутриклеточная изоформа ААТ локализуется в клеточных компартментах, не характерных для секреторной.

4. Короткие транскрипты, содержащие экзон 5 *SERPINA1* - источник усеченных протеоформ ААТ, в том числе С-концевых пептидов ААТ.

5. Продукты трансляции с коротких транскриптов поступают в клеточные компартменты, не характерные для секреторной изоформы.

6. Биологическое разнообразие продуктов гена *SERPINA1* повышается за счет реализации альтернативных сайтов инициации трансляции в составе длинных (5'-проксимальный сайт) либо коротких (3'-проксимальные сайты) транскриптов.

Степень достоверности результатов и апробация работы

Результаты работы были получены с использованием современных методик и на современном оборудовании. Результаты статистически достоверны и воспроизводимы. Основные положения и выводы исследования изложены в 5 статьях в рецензируемых научных изданиях, индексируемых в базах данных WoS, Scopus, RSCI и рекомендованных для защиты в диссертационном совете МГУ. Основные результаты работы были представлены на 9 конференциях.

Публикации

1. Maslakova A.A.; Golyshev S.A.; Potashnikova D.M.; Moisenovich A.M.; Orlovsky I.V.; Smirnova O.V.; Rubtsov M.A. *SERPINA1* long transcripts produce non-secretory alpha1-

antitrypsin isoform: *In vitro* translation in living cells // Int J Biol Macromol. 2023. Vol. 241. P. 124433. IF Web of Science = 8,2

2. **Maslakova A.A.**; Didych D.A.; Golyshev S.A.; Katrukha I.A.; Viushkov V.S.; Zamalutdinov A.V.; Potashnikova D.M.; Rubtsov M.A.; Smirnova O.V.; Orlovsky I.V. Towards unveiling the nature of short *SERPINA1* transcripts: Avoiding the main ORF control to translate alpha1-antitrypsin C-terminal peptides // Int J Biol Macromol. 2022. Vol. 203. P. 703-717. IF Web of Science = 8,2

3. **Maslakova A.A.**; Telkov M.V.; Orlovsky I.V.; Sokolova O.S. Comparative analysis of *SERPINA1* gene expression in tumor cell lines // Moscow University biological sciences bulletin. 2015. Vol. 70, № 3. P. 127–131. IF Scopus = 0,9. Статья доступна на русском языке в журнале Вестник Московского университета. Серия 16: Биология. 2015. Т. 70, № 3. С. 26–31.

4. **Докрунова А.А. (Маслакова А.А.)**; Соколова О.С. Содержание α 1-антитрипсина в сыворотке при заболеваниях предстательной железы // Урология. 2012. Т. 5. С. 77-80. IF Scopus = 0,293; IF РИНЦ Web of Science = 0,94

5. **Докрунова А.А. (Маслакова А.А.)**; Уласова Н.Ю.; Крамина Т.Е.; Альбицкий И.А.; Хасигов П.З.; Соколова О.С. Полиморфизм альфа1-антитрипсина при раке и гиперплазии предстательной железы // Молекулярная медицина. 2010, Т. 6. С. 43-49. IF РИНЦ Web of Science = 0,5

Участие в конференциях

1. **Maslakova A.A.**; Golyshev S.A.; Potashnikova D.M.; Moisenovich A.M.; Orlovsky I.V.; Rubtsov M.A.; Smirnova O.V. Non-secretory alpha1-antitrypsin isoform produced from *SERPINA1* long transcripts // 2022 ASCB Annual Meeting. Mol Biol Cell Suppl. 2022. Vol. 34, № 2. P.584.- P1966.

2. **Maslakova A.A.**; Funtova Y.S.; Rubtsov M.A.; Viushkov V.S.; Orlovsky I.V.; Smirnova O.V. Hormones as effective regulators of short *SERPINA1* transcripts encoding alpha1-antitrypsin bioactive peptides *in vitro* // Molecular Biology of the Cell. 2020. Vol. 31, № 26. P. 2863. - P1380.

3. **Maslakova A.A.**; Golyshev S.A.; Rubtsov M.A.; Smirnova O.V.; Orlovsky, I.V. Alpha1-antitrypsin C-terminal peptide instantly stimulates hepatocyte metabolism // ASCB/EMBO Meeting. Mol Biol Cell Suppl. 2018. Vol. 29, № 26. P. 3063. - P1715.

4. **Maslakova A.A.**; Rubtsov M.A.; Efimova V.S.; Orlovsky, I.V.; Sokolova O.S.; Smirnova O.V. Alpha1-antitrypsin-derived C-terminal peptide is a potent oxidative stress inhibitor // ASCB/EMBO Meeting. Mol Biol Cell Suppl. 2017. Vol. 28, № 26. P. 3727. - P1259.

5. **Maslakova A.A.**; Efimova V.S.; Maslakov A.S.; Spangenberg V.E.; Rubtsov M.A.; Orlovsky I.V. VIRIP – an anti-HIV host peptide output hypothesis // Abstracts of the 30th international conference on antiviral research (ICAR). 2017. P.104.

6. **Maslakova A.A.**; Sokolova O.S.; Orlovsky I.V. Alpha1-antitrypsin-derived C-terminal peptide activity at subnanomolar to micromolar concentrations in human cell lines: a mitogen and a metabolism activator // *Molecular Biology of the Cell (Suppl. Data)*. 2016. Vol. 27, № 25. P. 3947. – P1169.

7. **Maslakova A.A.**; Spangenberg V.E.; Sokolova O.S.; Orlovsky I.V. Search and analysis of short transcripts from *SERPINA1* gene locus // XI international scientific-practical conference “High Technologies, Basic and Applied Researches in Physiology and Medicine”. 2016. P. 186-192.

8. **Dokrunova A.A.**; Sokolova O.S. Alpha1-antitrypsin in prostate cancer // Тезисы докладов “The 1st Multidisciplinary Symposium “Molecular oncology: from laboratory bench to medicine”, раздел "Novel approaches in cancer diagnostics", изд-во Института молекулярной биологии и генетики НАН Украины (Киев, Украина). 2012. С.67.

9. **Докрунова А.А.**; Соколова О.С. Альфа1-антитрипсин как потенциальный маркер рака предстательной железы // Тезисы докладов XIX Международной конференции студентов, аспирантов и молодых учёных «Ломоносов» секции “Фундаментальная медицина” подсекции “Клиническая медицина”. Электронный сборник тезисов. 2012.

2. ОБЗОР ЛИТЕРАТУРЫ

Регуляция протеолитических ферментов эндогенными ингибиторами является необходимым условием поддержания гомеостаза в организме. ААТ является важным компонентом сыворотки крови и тканевых жидкостей, подавляющим чрезмерную активность протеолитических ферментов, выделяемых активированными иммунными клетками, преимущественно нейтрофилами. Альфа1-антитрипсин - третий по распространенности сывороточный белок. Это свидетельствует о его жизненно важной роли - поддержании баланса в организме, и как показывают многочисленные работы, не только между протеолизом и защитой от него. Разнообразие функций и ролей этого белка в организме обеспечивается как тонко настроенной регуляцией экспрессии кодирующего его гена *SERPINA1*, так и биохимическими и физико-химическими свойствами самой молекулы ААТ, позволяющей ей взаимодействовать со структурно не связанными друг с другом молекулами и подвергаться посттрансляционным модификациям, модулирующим свойства белка. В обзоре литературы мы постарались охватить данные об особенностях экспрессии гена *SERPINA1*, свойствах и функциях ААТ, зависящих и не зависящих от его антипротеолитической активности, его роли в норме и при патологии, уделяя особое внимание причинам и механизмам антитрипсиновой недостаточности, а также роли ААТ и его С-концевых фрагментов в канцерогенезе.

2.1 Альфа1-антитрипсин. Краткий обзор суперсемейства серпинов

Антипротеолитические свойства крови были замечены Ферми и Перносси еще в 1894 г. [59]. Поскольку было выдвинуто предположение, что ингибирующий компонент крови является *антителом* [60], его назвали “*антитрипсин*” [61], что впоследствии признали ошибочным [62]. Возможность идентификации молекулы, придающей крови эти свойства, представилась лишь полвека спустя, в связи с развитием лабораторных методов. В 1955 г. Шульце выделил и охарактеризовал белок, который позже назвал “альфа1-антитрипсин”, ввиду его присутствия в альфа1-глобулиновой фракции и способности ингибировать трипсин поджелудочной железы [63, 64]. Исходное название является несколько историческим: позже у белка появилось альтернативное название - “альфа1-ингибитор протеиназы” - ввиду его основной установленной физиологической функции - ингибировать эластазу нейтрофилов [65]. Острый интерес к ААТ и в последствии к серпинам в целом был вызван открытием в 1963 году, когда Лаурел и Эриксон с помощью электрофоретического разделения сыворотки крови пациентов обнаружили связь отсутствия фракции ААТ с развитием острой обструктивной болезни легких [18].

Альфа1-антитрипсин - архетипический представитель самого большого надсемейства ингибиторов сериновых протеаз SERPIN (акроним от *serine proteinase inhibitor*) [66, 67]. Существование целого надсемейства белков признали с обнаружением высокой степени гомологии между овальбумином, ААТ и антитромбином (30–50%) [68].

Аббревиатура SERPIN была первоначально предложена Каррэллом и Трэвисом в 1985 г. [66], поскольку первые идентифицированные серпины из плазмы млекопитающих действуют на химотрипсин-подобные сериновые протеазы.

Серпиноподобные гены обнаружены у животных, растений, бактерий, архей и поксвирусов [69, 70, 71]. Серпины имеются у всех многоклеточных эукариот: например, у *Homo sapiens*, *Drosophila melanogaster*, *Arabidopsis thaliana* и *Caenorhabditis elegans* 37, 13, 29 и около 9 серпиноподобных генов, соответственно [69, 72, 73], - а также обнаружены у одноклеточных эукариот [74]. 219 серпинов эукариот классифицированы в 16 кладов (- групп, возникших из одного предшественника) (clades A-P) по гомологии [69, 72]. Единого предкового гена-гомолога серпинов эукариот среди прокариот не обнаружено [69, 75]. Серпины в основном обнаруживаются у эукариот, у прокариот представлены спорадически и обладают существенной степенью гомологии с серпинами позвоночных клада В (внутриклеточных серпинов, - эволюционно более древних, чем секреторные), ввиду чего предположили, что прокариоты приобрели серпины посредством горизонтального переноса генов между царствами от эукариот [70, 76]. Однако, такой механизм не объясняет тот факт, что гены серпинов обнаруживаются у свободноживущих, в т.ч. экстремофильных, прокариот [75]. В недавнем наиболее обширном филогенетическом исследовании серпинов проанализировали более 18000 уникальных белковых последовательностей (лишь малая часть из которых имеет экспериментальное подтверждение), извлеченных из общедоступных белковых баз данных, и пришли к выводу, что серпины про- и эукариот эволюционировали конвергентно [75]. В этом же исследовании предложили расширить классификацию серпинов, добавив клады Q-U, а некоторые серпины в клады не классифицируются вовсе и представляют “бесхозные” (orphan) последовательности (серпины грибов, некоторых прокариот и нехордовых многоклеточных животных) [75].

В отличие от животных, у растений отсутствуют химотрипсин-подобные протеазы (за исключением таковой, найденной в пыльце амброзии), и их серпины выполняют защитную функцию от экзогенных протеаз и патогенов [77]. У позвоночных и беспозвоночных животных серпины являются критически важными регуляторами основных протеолитических каскадов организма, включая свертывание крови и гемолимфы, активацию системы комплемента, фибринолиз, апоптоз, ангиогенез и воспалительные процессы [78].

Серпины относительно крупные молекулы размером 350-500 аминокислотных остатка (а.о.), в сравнении, например, с панкреатическим секреторным ингибитором трипсина человека (PSTI/SPINK1, serine protease inhibitor Kazal type I), состоящим из 56 а.о. [79]. В последовательности серпинов из разных филогенетических групп 51 а.о. высококонсервативны (встречаются в >70% последовательностей), значительная доля из которых является частью гидрофобного кора белка, что указывает на сохранение их третичной структуры [69]. Действительно, несмотря на функциональное расхождение

серпинов в процессе эволюции, они сворачиваются в консервативную метастабильную структуру [69]. Благодаря такой особой структуре ингибирующие серпины, в отличие от небольших «жестких» ингибиторов протеаз (таких как ингибиторы семейства Казала или Кунитца) известны своим необычным механизмом действия - т. н. субстратно-суицидальным механизмом ингибирования. В момент узнавания протеазой-мишенью молекула серпина претерпевает существенные и сложные конформационные изменения, в результате которых необратимо ингибирует целевую протеазу, разрушая ее каталитический центр [80]. Ввиду своей уникальной природы, серпины представляют интерес в контексте изучения общей структуры и укладки белков [69].

Ингибирующая специфичность серпинов определяется последовательностью петли реакционного центра (RCL - reactive center loop), в особенности одним остатком аминокислоты (P1) в активном сайте молекулы, экспонированном в составе петли, который узнается в качестве субстрата соответствующей протеазой [66, 81, 82]. Так, остаток метионина (M358) в активном сайте ААТ узнается эластазой, аргинина (R393) в антитромбине – тромбином (таблица 2.1). Установлено, что детерминанты специфичности более сложны, чем сама последовательность RCL, и представляют сложные взаимодействия между динамикой внутри и вокруг RCL и локальной электростатикой [83], а также вспомогательные экзосайты, которые стабилизируют промежуточный нековалентный комплекс серпин-протеаза (Михаэлис-подобный комплекс) перед расщеплением RCL [84].

Clade; Serpin	P4	P3	P2	P1-P1'	P2'	P3'	P4'	Primary target
A1; α 1-PI	A	I	P	M-S	I	P	P	Neutrophil elastase
A3; α 1-ACT	I	T	L	L-S	A	L	V	Cathepsin G
A4; kallistatin	I	K	F	F-S	A	Q	T	Tissue kallikrein
A5; PCI	F	T	F	R-S	A	R	L	Activated protein C
A10; PZ	I	T	A	Y-S	M	P	P	IXa & XIa
B1; MNEI	A	T	F	C-M	L	M	P	Neutrophil elastase
B2; PAI-2	M	T	G	R-T	G	H	G	plasminogen activator
B3; SCCA1	G	F	G	S-S	P	A	S	Cathepsin K, L & S
B6; PI6	M	M	M	R-C	A	R	F	plasmin
C1; antithrombin III	I	A	G	R-S	L	N	P	IIa & IXa
D1; heparin cofactor II	F	M	P	L-S	T	Q	V	IIa
E1; PAI-1	V	S	A	R-M	A	P	E	Plasminogen activator
E2; protease nexin 1	L	I	A	R-S	S	P	P	Plasminogen activator
F2; α 2-antiplasmin	A	M	S	R-M	S	L	S	plasmin
G1; C1-inhibitor	S	V	A	R-T	L	L	V	C1
I1; neuroserpin	A	I	S	R-M	A	V	L	Plasmin & plasminogen activator

↑
Scissile bond

Таблица 2.1. Остаток аминокислоты в активном сайте различных серпинов, определяющий их специфичность к протеазам-мишеням. Остатки, фланкирующие активный сайт P1-P1', также важны для мульти специфичности к протеазам. Таблица полностью взята из [81].

Среди известных 37 серпинов человека, 30 выполняют основную функцию в регуляции протеолитической активности [73]. Большинство ингибирующих серпинов человека инактивируют сериновые химотрипсин-подобные сериновые протеазы. Однако некоторые проявляют и межклассовую специфичность. К примеру, SERPINB9 ингибирует цистеиновую протеазу каспазу 1 [85], SERPINB3 - папаино-подобные цистеиновые протеазы катепсины L, K и S [86]. Тем не менее, ферментативная химия аналогична, и

механизм ингибирования серпинами одинаков для обоих классов протеаз [87]. Кроме того, не все представители суперсемейства обладают классическими ингибиторными свойствами. Так, ангиотензиноген (*SERPINA8*) регулирует кровяное давление [65], 47 кДа белок теплового шока HSP47 (*SERPINH1*) является проколлаген/коллаген-специфическим шапероном [88], кортикостероид-связывающий (*SERPINA6*) и тироксин-связывающий (*SERPINA7*) глобулины человека участвуют в транспорте гормонов [89, 90]. Интересно, что механизм высвобождения гормона также зависит от динамики конформационных изменений серпина [91].

Чрезвычайное влияние серпинов на множественные молекулярные регуляторные пути проявляется в генетических аномалиях, при которых регуляция нормальных функций серпинами утрачивается. Серпинопатии возникают в результате точечных мутаций, приводящих к образованию упорядоченных полимеров, которые удерживаются в синтезирующей клетке. Это вызывает «усиление токсической функции» из-за накопления белка и «потерю функции» в результате дефицита растворимого мономерного ингибитора [92]. Утрата серпиновой активности может иметь серьезные патологические последствия. Например, у пациентов с дефицитом α 1-антитрипсина (ААТ/*SERPINA1*) развивается цирроз печени из-за токсического действия полимерных агрегатов [93] и эмфизема легких из-за неконтролируемой активности эластазы нейтрофилов [94]. Дефицит ингибитора С1-эстеразы (*C1INH/SERPING1*) приводит к приступам ангионевротического отека из-за чрезмерного образования брадикинина контактной системой плазмы [95], в то время как у пациентов с низким уровнем антитромбина (АТIII/*SERPINC1*) повышен риск ишемического инсульта, тромбоза глубоких вен или легочная эмболия из-за повышенной активности системы свертывания [96]. Мутации в гене *SERPINI1*, кодирующем нейросерпин, ассоциированы с эпилепсией и деменцией [97]. Таким образом, серпины представляют большой клинический интерес.

2.2 Особенности синтеза и “жизненный цикл” альфа1-антитрипсина

2.2.1 Структура гена *SERPINA1*. Структура транскриптов гена *SERPINA1*. Регуляция экспрессии и тканеспецифическая экспрессия гена *SERPINA1*

Название «SERPIN» было выделено как символ всех генов серпинов для людей и других видов. Полная последовательность гена, кодирующего ААТ человека, идентифицирована в 1984 г. [98]. Ген был назван *SERPINA1*, что отражает принадлежность ААТ к кладу А, включающему секреторные серпины. Ген *SERPINA1* находится в кластере генов серпинов 14q32.1 в длинном плече хромосомы 14 (рисунок 2.1, А). Кластер общей протяженностью ~370 т.п.н. включает 11 генов и условно поделен на проксимальный, центральный и дистальный субкластеры, включающие 4, 3 и 4 гена, соответственно, что составляет $\approx 1/3$ от общего количества генов серпинов человека [42]. Ген *SERPINA1*

находится в проксимальном субкластере и в геномном браузере UCSC Genome Browser (<http://genome.ucsc.edu>) расположен в участке 14q32.13 с координатами chr14:94,376,747-94,390,692 (протяженность гена 13946 п.н.) - (рисунок 2.1, А), на минус-цепи ДНК 14й хромосомы, содержит 6 интронов, 3 нетранслируемых альтернативных экзона (IA, IB и IC, образующих 5'-НТО транскриптов), 4 кодирующих экзона (экзоны II-V) и 3'-нетранслируемую область (3'-НТО) в составе экзона V общей протяженностью 1702 п.н. (рисунок 2.1, Б).

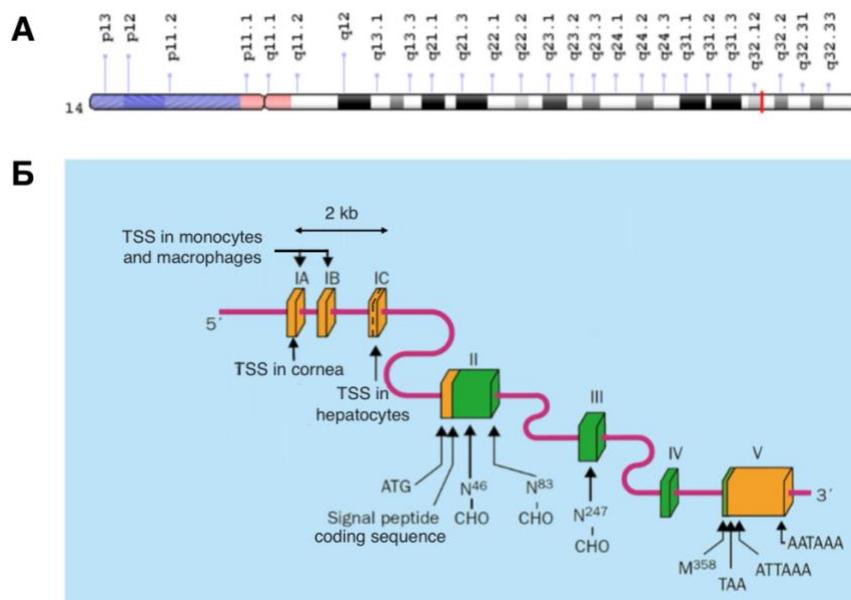


Рисунок 2.1. А) Локализация гена *SERPINA1* на 14 хромосоме. Изображение получено в онлайн-программе Genome Decoration Page/NCBI. Б) Схема строения гена *SERPINA1*: кубики — экзоны; зеленым цветом обозначены кодирующие экзоны; оранжевым - нетранслируемые экзоны (соответствующие 5'-НТО), участок, кодирующий сигнальный пептид в экзоне II и участок экзона V, образующий 3'-НТО. Экзоны и интроны отображены

пропорционально относительным длинам. TSS - сайты старта транскрипции [41, 99, 100]. Иллюстрация взята из [101] с некоторыми изменениями.

Особенностью гена *SERPINA1* является сложность механизмов, регулирующих его экспрессию. Экспрессия гена контролируется двумя промоторами, влияющими на выбор старта транскрипции. Гепато-специфический (проксимальный) промотор находится выше старта транскрипции в экзоне IC и содержит практически канонический ТАТА-бокс в положении -25/-20, мутация которого полностью ингибирует транскрипцию гена [102, 103]. В пределах 137-нуклеотидной последовательности выше старта транскрипции в экзоне IC располагается гепато-специфический проксимальный цис-регуляторный элемент (-137/-37), включающий два энхансера в положениях -86/-58 (для связывания с транскрипционным фактором HNF-1 α - hepatocyte nuclear factor 1 α) и -125/-100 (HNF-4), действующие синергически [103, 104, 105, 106, 107, 108]. Указанные транскрипционные факторы являются положительными регуляторами и первичными детерминантами тканеспецифической экспрессии гена *SERPINA1* в клетках печени [109]. Проксимальный промотор используется в тканях печени, кишечника и легких [42]. В энтероцитах задействуются те же транс-действующие факторы, что и в клетках печени

(HNF-1 α и HNF-4) [110], а в альвеолярных эпителиальных клетках HNF-1 β играет ключевую роль [15].

мРНК *SERPINA1* - одна из наиболее распространенных среди всех мРНК печени [12], чему способствует не только особая гепато-специфическая структура хроматина всего кластера 14q32.1, содержащего большое количество ДНКзаI-гиперчувствительных сайтов, но и сила гепато-специфического промотора [111]. Участок ДНК протяженностью в 2,3 т.п.н. между -8,1-5,8 т.п.н. выше гена *SERPINA1* регулирует не только активацию самого гена, но и отвечает за ремоделирование хроматина во всем проксимальном субкластере 14q32.1 [111].

Стимуляция экспрессии гена происходит в основном под действием провоспалительных цитокинов, таких как IL-6, IL-1 β , TNF α , онкостатин М (OSM) или бактериальный липополисахарид [112, 113]. OSM-чувствительный элемент (3'-энхансер) расположен между +11536 и +11670 п.н. от старта транскрипции в гепатоцитах, IL-6-чувствительный элемент - между +11167 и +11536 п.н. [114]. Энхансеры в составе репортерной конструкции работают синергически как между собой, так и с проксимальным промотором: ко-стимуляция цитокинами вызывает максимальный 5-кратный отклик репортерного гена, в то время как при стимуляции цитокинами в отдельности эффект превышал базовую экспрессию в 3 раза. Известно, что транскрипционные факторы Oct-1 и STAT3 опосредуют эффект цитокинов и связываются с IL-6 и OSM-чувствительными элементами, соответственно [115, 114]. IL-6 рассматривается как основной активатор экспрессии гена в гепатоцитах [116], в то время как OSM - в альвеолярном легочном эпителии [113]. Таким образом, ген *SERPINA1* чувствителен к провоспалительным сигналам, вызывающим системный ответ острой фазы, а его продукт - альфа1-антитрипсин - белком острой фазы. В процессе воспаления, а также в 3ем триместре беременности [117] концентрация белка в крови может повышаться в несколько раз.

Второй (дистальный) промотор расположен на 2000 п.н. выше печеночного и задействуется в экспрессии гена в моноцитах, макрофагах и клетках роговицы [41, 42] (рисунок 2.1, Б). Оба промотора работают тканеспецифически, в результате чего в гепатоцитах и энтероцитах транскрипты начинаются с середины экзона 1С, в то время как в моноцитах и макрофагах - с экзонов 1А и 1В [41, 99] (рисунок 2.1, Б).

Два альтернативных промотора, отвечающих за выбор разных стартов транскрипции, шесть альтернативных донорных и три альтернативных акцепторных сайтов сплайсинга приводят к тому, что ген *SERPINA1* транскрибируется в избыточное количество изоформ мРНК (более 11 вариантов) и входит в 0,5% генов человека, обладающих наибольшей транскрипционной сложностью (рисунок 2.2) [43, 44]. При этом, порядка 95% мультиэкзонных генов человека альтернативно сплайсируются [118], а среднее количество изоформ мРНК, приходящихся на один ген, находится в пределах 3-7 [119, 120]. Стоит отметить, что в композицию на рисунке 2.2 не включен моноцит-специфический

транскрипт, начинающийся с экзона 1В, - он также не представлен в базе данных Ensembl, но транскрипция с экзона 1В описана в литературе [41, 99]. Кроме того, основной гепато-специфический транскрипт (NM_000295.5) содержит укороченный вариант экзона 1С (рисунок 2.2, В), и его TSS приходится на середину экзона 1С [99]. Таким образом, ожидаемое количество изоформ мРНК превышает 11 вариантов, большая часть из которых генерируется за счет альтернативного сплайсинга.

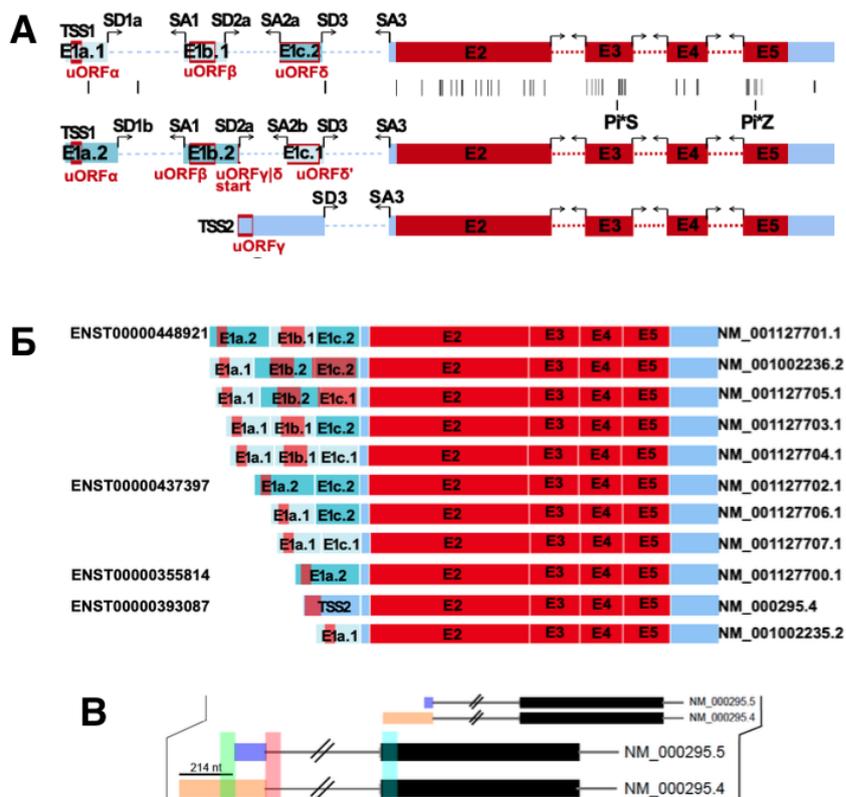


Рисунок 2.2. С гена *SERPINA1* продуцируются как минимум 11 изоформ мРНК, большая часть из которых генерируется за счет альтернативного сплайсинга. А) Схема структуры экзонов с отображением альтернативных доноров (SD) и акцепторов (SA) сплайсинга. Кодирующая область обозначена красным, нетранслируемые области - голубым. Красными рамками обозначены 5'-ОРС (uORF): 5'-ОРС $\delta\delta'$ охватывает сплайс-стыковку двух экзонов и представлен только в транскриптах, содержащих экзон E1b.2. TSS - сайт старта транскрипции. Черными вертикальными линиями обозначены аллели, ассоциированные с дефицитом ААТ (согласно базе данных Human Gene Mutation Database), включая наиболее часто встречающиеся P δ *S и P δ *Z аллели. Б) Схема 11 вариантов мРНК *SERPINA1*: слева даны наименования Ensembl, справа - NCBI. 5'-ОРС обозначены светло-красными прямоугольниками в области 5'-НТО. Рисунки А и Б полностью взяты из [44]. В) Транскрипты, начинающиеся с экзона 1С, представлены двумя вариантами, отличающимися на 214 н.о. за счет старта транскрипции с двух разных TSS (NM_000295.5 и NM_000295.4). Рисунок взят из [121].

Заметной особенностью альтернативного сплайсинга *SERPINA1* является то, что все варианты мРНК различаются только структурой 5'-НТО. Изменение первичной структуры 5'-НТО приводит к перетасовке коротких вышерасположенных рамок считывания (5'-ОРС) или, в случае короткой ведущей последовательности, как в гепато-специфическом транскрипте NM_000295.5, к их отсутствию (рисунок 2.2). Изменения в первичной структуре 5'-НТО, наряду с индивидуальной доступностью стартовых кодонов основной и

вышестоящих рамок считывания ввиду различной пространственной структуры РНК, играет роль в посттранскрипционной регуляции экспрессии гена *SERPINA1*, влияя на эффективность трансляции [44]. Вклад в разнообразие вариантов транскриптов может вносить также выбор альтернативных сайтов полиаденилирования [38, 122]. Считается, что, поскольку все варианты транскриптов содержат главную ОПС, у здорового взрослого человека со всех мРНК транслируется один и тот же секреторный белок [44].

2.2.2 Структура и фолдинг полноразмерного ААТ. Гликоформы ААТ.

Конформеры ААТ

Выбор первого альтернативного нетранслируемого экзона не влияет на структуру белкового продукта. Протяженность кодирующей области гена *SERPINA1* составляет 1257 п.н., включая стоп-кодон. Предшественник ААТ состоит из 418 а.о., первые 24 из которых являются сигнальным пептидом секреции. Фолдинг ААТ происходит в ЭПР, где от молекулы предшественника отщепляется сигнальный пептид и новосинтезированная полипептидная цепь подвергается посттрансляционной модификации - N-гликозилированию. N-гликаны обеспечивают контроль качества фолдинга ААТ через зависимые от гликозилирования шапероны кальретикулин и кальнексин [123]. Гликозилирование необходимо для эффективной секреции зрелой молекулы ААТ из клетки, зрелый белок состоит из 394 а.о. с молекулярной массой порядка 52-56 кДа, 15% из которых приходится на остатки сахаров [124, 125, 126].

Наиболее часто встречающаяся аллель *SERPINA1* дикого типа Pi*(M) (Proteinase inhibitor M allele) обеспечивает оптимальный уровень ААТ в сыворотке порядка 1,5 г/л. Известны девять гликоформ М-варианта ААТ, M0-M8, разделяющихся с помощью двумерного электрофореза и отличающихся в основном структурой олигосахаридов, присоединенных по остаткам аспарагина в положениях 46, 83 (кодируемыми экзоном 2) и 247 (экзоном 3) (нумерация относительно зрелой молекулы ААТ, в которой сигнальный пептид отщеплен), из которых количественно преобладают M4 и M6 (рисунок 2.3). Изоэлектрические точки гликоформ ААТ варьируют в пределах рН 4,2-4,9 со средним интервалом в 0,05 рI-единиц [127]. При электрофоретическом разделении ААТ мигрирует в α 1-глобулиновой фракции, что отражено в названии белка.

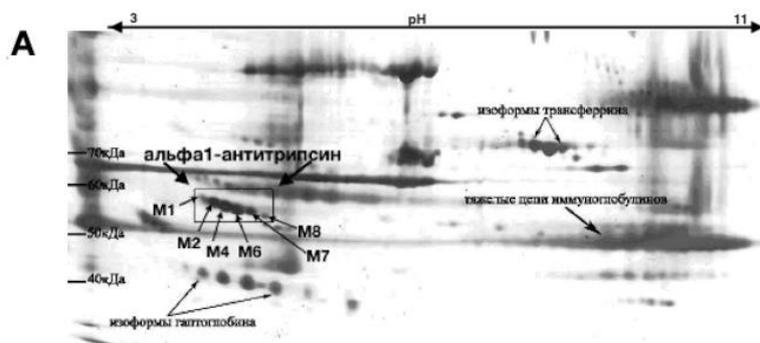
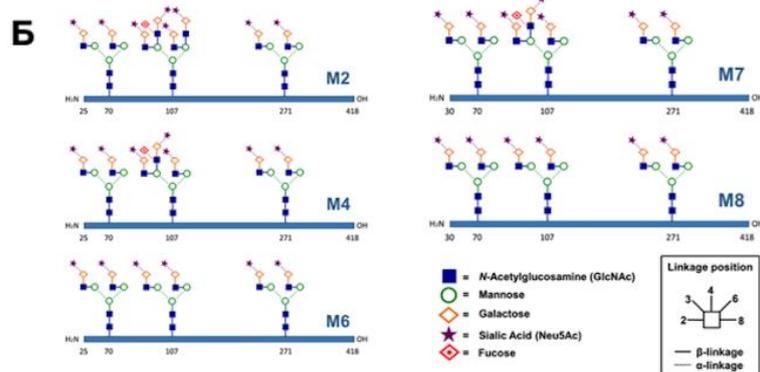


Рисунок 2.3. А) Гликоформы ААТ на карте двумерного электрофореза. Рисунок взят из [128]. Б) Структура остатков сахаров в разных гликоформах ААТ. Остатки аспарагина пронумерованы относительно предшественника ААТ. Рисунок взят из [126].



Во всех гликоформах в положениях 46 и 247 присоединены биветвляющиеся остатки сахаров, а в положении 83 - гликаны различной структуры и ветвления, определяющие ту или иную гликоформу (рисунок 2.3, Б). Как и в случае других гликозилируемых белков, первичный этап гликозилирования молекулы ААТ происходит в ЭПР, а завершающее ветвление гликанов - в аппарате Гольджи. Гликоформы М7 и М8 по структуре сахаров идентичны М4 и М6, соответственно, но их аминокислотная последовательность усечена с N-конца на пять аминокислот [127]. Гликоформы М0, М1, М3 и М5 являются минорными, чаще всего не детектируются с помощью 2D-электрофореза ввиду низкой представленности и сложнее идентифицируются методами протеомного анализа [126]. На рисунке 2.4, А изображена гликоформа М4 с сохранением относительного размера полипептидного остова ААТ и N-гликанов.

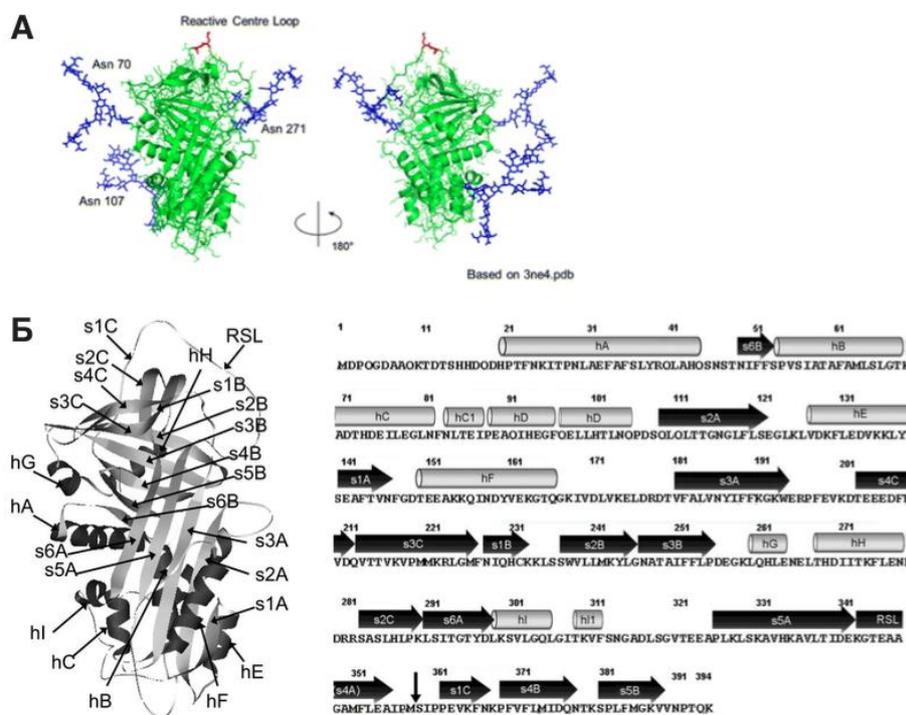


Рисунок 2.4. А) Модель гликозилированной молекулы ААТ на основе структуры гликоформы М4. Зеленым - полипептидный остов, синим - остатки сахаров, красным - активный сайт молекулы (M358-S359). Остатки аспарагина пронумерованы относительно предшественника ААТ. Рисунок взят из [126]. Б) слева - Структура нативной молекулы ААТ (PDB, 1QLP), справа - Аминокислотная последовательность ААТ с обозначением элементов вторичной структуры: h-helix; s-strand; стрелкой обозначен сайт расщепления целевой протеазой. Рисунок взят из [129] с некоторыми изменениями.

Особенностью молекул серпинов, отличающей их от большинства других белков, является их способность сворачиваться в энергетически не оптимальную конформацию и существовать в таком виде длительное время. ААТ имеет упорядоченную структуру, состоящую из трех β -листов (листы А-С) и девяти α -спиралей, окружающих “каркас” из β -листов (рисунок 2.4, Б), что необходимо для надежной фиксации незащищенной петли реакционного центра RCL в оптимальной “рабочей” конформации [66].

Спонтанное переключение в более стабильную, но неактивную латентную конформацию молекулы ААТ (рисунок 2.5, А, Б) затруднено высоким кинетическим барьером, ввиду чего время полужизни белка в сыворотке длительное и составляет порядка 4-5 дней [130], в отличие, например, от другого серпина - ингибитора активатора пламиногена (PAI-1, *SERPINE1*), функциональное время полужизни которого составляет лишь 1-2 часа ввиду спонтанной инактивации [131]. ААТ можно “заставить” принять латентную конформацию путем нагревания при высоких температурах в 0,7 М цитрате в течение 12 часов [132].

Нативная конформация рассматривается как кинетически “пойманное” промежуточное состояние, за счет которого серпины выполняют антипротеолитические функции. Известно, что кинетический барьер в молекуле ААТ обеспечивается оптимальной длиной RCL, при увеличении которой молекула спонтанно инактивируется [133].

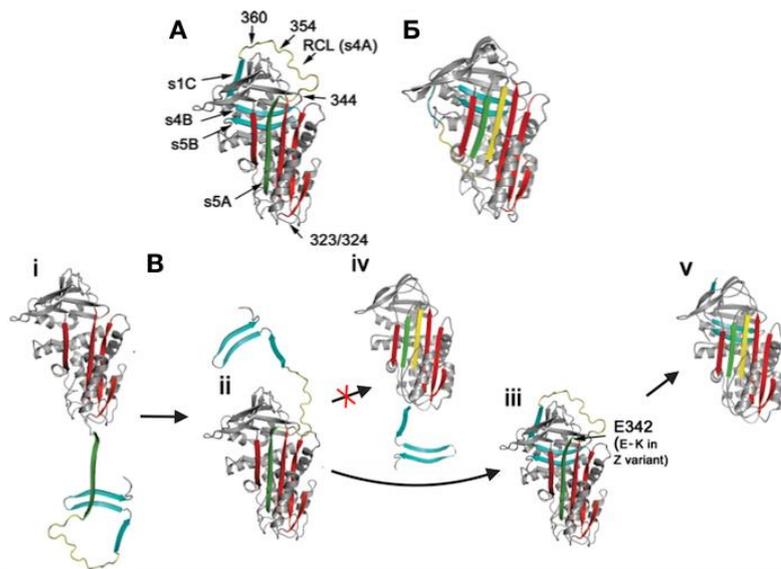


Рисунок 2.5. Две основные конформации ААТ - нативная (А) и латентная (Б). На рисунке А указаны ключевые структурные элементы ААТ. В) Последовательность сворачивания белка, способствующая фиксации молекулы ААТ в метастабильной конформации. Рисунок полностью взят из [134].

Переход из одного состояния в другое происходит за счет встраивания петли реакционного центра RCL в бета-складчатый слой А в виде дополнительной, четвертой бета-цепи (s4A, выделена желтым цветом на рисунке 2.5, А, Б), в то время как бета-тяж s1C выходит из состава бета-листа С и экспонируется. Точные подсчеты разницы свободной энергии Гиббса между двумя конформерами ААТ отсутствуют, но известно, что разница температур денатурации обоих конформеров ААТ составляет порядка 17°C [134].

Удивительно, что для эффективного сворачивания в “рабочее” состояние не требуется помощи шаперонов или продоменов (pro regions), поэтому, денатурированные из нативной либо латентной конформации, серпины способны снова ренатурировать в метастабильный конформер (за исключением таких серпинов, как С1-ингибитор (*SERPING1*), нативная конформация которого стабилизирована дисульфидной связью [135]). На рисунке 2.5, В представлена последовательность сворачивания ААТ в нативную конформацию, проверенная экспериментально [134], ключевым этапом которой является фиксация RCL за счет предварительного встраивания бета тяжа s5A в бета-слой А, и лишь затем - встраиванием бета-тяжей s1C/s4B/s5B в белковый каркас, которое неэффективно без предварительной укладки тяжа s5A. Это связано с тем, что тяж s5A в структуре бета-слоя А оптимизирует гидрофобные взаимодействия между бета-слоями А и В [134].

2.2.3 Молекулярные механизмы антипротеолитической активности ААТ. Механизм утилизации комплекса ААТ-протеаза

ААТ - чрезвычайно эффективный ингибитор сериновых протеаз. Специфической мишенью ААТ является эластаза нейтрофилов, аффинность к которой максимальна среди других протеолитических ферментов и константа скорости ассоциации с которой одна из наиболее высоких в природе ($K_{acc} = 6.5 \times 10^7 \text{ M}^{-1} \text{ сек}^{-1}$) [136]. Помимо ингибирования избытка эластазы нейтрофилов, поджелудочной железы и бактерий, ААТ также способен

инактивировать протеиназу-3 и катепсин G, выделяемые нейтрофилами [137]; химазу тучных клеток [138]; калликреины 7 и 14 [139]; трипсин и химотрипсин поджелудочной железы [140]; гранзим B из T-лимфоцитов [141]; сериновые протеазы коагуляционного каскада (плазмин и тромбин при очень высоких концентрациях ингибитора [142], урокиназу [143] и фактор Ха [144]), а также каталитический домен трансмембранной сериновой протеазы матриптазы [145] и TMPRSS2, который ферментативно праймирует шиповидный белок SARS-CoV-2 для слияния вируса с мембраной клетки [7]. ААТ обеспечивает около 90% антипротеолитической емкости сыворотки крови человека, остальные 10% приходятся на альфа2-макроглобулин [11].

Механизм ингибирования сериновой протеазы молекулой ААТ подробно изучен. RCL ААТ - специализированная экспонированная на поверхности молекулы структура (а.о. 345-360 зрелого ААТ), выполняющая роль псевдо-субстрата для протеазы-мишени. Ключевыми остатками аминокислот, служащими приманкой, являются P1-P1' остатки M358-S359 (в соответствии с номенклатурой, предложенной [146]). Остаток M358 является важным элементом, обеспечивающим специфичность ААТ к протеазам. Так, мутантная форма ААТ-Питтсбург (Pittsburg), в которой остаток метионина заменен на аспарагин (M358R) в активном сайте молекулы (подобно молекуле антитромбина III/*SERPINC1*, содержащей R393 в активном сайте), утрачивает специфичность к эластазе нейтрофилов, но приобретает способность эффективно ингибировать аргинин-специфические ферменты - факторы свертывания крови тромбин, факторы XIa и XIIa и калликреин, приводя к смертельным случаям кровопотери [147, 148]. Замена одной аминокислоты усиливает активность ААТ-Питтсбург по сравнению с ААТ М-вариантом в 4000 раз в отношении тромбина, что приблизительно в 40 раз эффективней ингибирования антитромбином III [149].

Сериновые протеазы катализируют расщепление пептидной связи посредством двухстадийного процесса [150]. Первоначально каталитический серин триады активного центра осуществляет нуклеофильную атаку на пептидную связь субстрата. Это высвобождает новый N-конец субстрата (то есть C-концевую часть субстрата) и образует ковалентную сложноэфирную связь между ферментом и субстратом. Этот ковалентный комплекс между ферментом и субстратом называется промежуточным соединением ацил-фермент. Для стандартных субстратов эфирная связь далее гидролизуется во второй стадии процесса (реакции деацилирования), и новый C-конец субстрата (то есть N-концевая часть субстрата) высвобождается для завершения катализа. Однако в случае ингибирующих серпинов это не так.

Вслед за “стыковкой” протеазы с остатком P1 в составе RCL, фермент расщепляет пептидную связь P1-P1', образуя интермедиат ААТ P1-ацил-протеаза. Однако, расщепление активного сайта ААТ сопровождается мгновенным изменением конформации молекулы с метастабильной (напряженной, Stressed) в гиперстабильную релаксированную

(Relaxed), в которой энергия Гиббса минимальна. Такое переключение конформации называют S-к-R переходом и оно одинаково как для расщепленной, так и для нерасщепленной латентной формы ААТ (сравните рисунки 2.5, Б с рисунком 2.6). Энергия, необходимая для конформационного перехода, берется из метастабильности нативной конформации молекулы.

Мгновенное конформационное реагирование расщепленного ААТ приводит к быстрому перемещению “ковалентно пойманной” протеазы вслед за сформированным бета-тяжем S4A сверху-вниз относительно молекулы серпина (рисунок 2.6). Расстояние, на которое смещается остаток P1, составляет 71 Å [80]. Каталитическая триада фермента разупорядочивается (ингибирование путем искажения - “inhibition by distortion”), - протеаза необратимо инактивируется [80, 151]. Таким образом, механизм ингибирования субстратно-суицидальный, - одна молекула ААТ ингибирует одну молекулу протеазы.

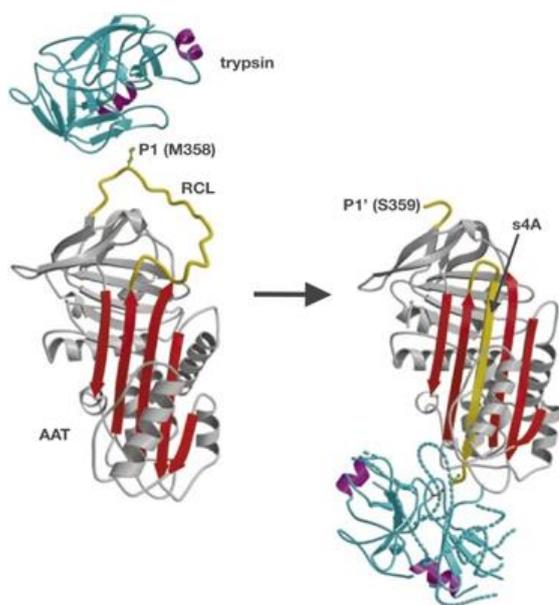


Рисунок 2.6. Образование комплекса ААТ-трипсин. Для сравнения ориентации молекулы трипсина альфа-спирали выделены фиолетовым цветом. Области разупорядоченной структуры трипсина в составе комплекса показаны в виде прерывистых витков, спроецированных из нативной структуры трипсина. Иллюстрация взята из [80].

При сравнении структур ингибиторных серпинов обнаружили, что остатки аминокислот P8-P12 (нумерация отображена в таблице 2.1) в области шарнира (P8-P15 hinge region, структурный компонент RCL) имеют малые боковые цепи и представлены преимущественно остатками аланина [152]. Установили, что в случае ААТ замена остатков в положениях P10 и P12 на остатки с более громоздкой структурой ингибиторные свойства молекулы снижаются: активная эластаза медленно высвобождается из комплекса с расщепленным ААТ либо константа скорости ассоциации комплекса снижается в 12 раз, соответственно [152]. Также показали, что помимо остатков P1-P1' для ингибирования важен остаток P2 (P357), замена которого на остатки с более громоздкой боковой цепью, а в особенности - на отрицательно заряженные остатки, приводило к существенному снижению образования ингибиторного комплекса [153].

Выведение ингибиторного комплекса ААТ-протеаза из кровообращения происходит в печени. В работе [154] показано, что в качестве лиганда для узнавания комплекса SEC-рецепторами-мусорщиками (scavenger Serpin-Enzyme Complex receptor) на поверхности гепатоцитов служит пентамотив *FVFLM* в составе бета-тяжа s4B С-концевого домена (рисунок 2.4, Б), спрятанный в нативном ААТ, но экспонированный и доступный для узнавания после реакции ингибирования. Взаимодействие с SEC-рецепторами стимулирует экспрессию *SERPINA1* в гепатоцитах по принципу положительной обратной связи [155]. Таким образом, С-концевой домен (отныне - *С-концевой пептид* ААТ), начинающийся с остатка P1' (S359), удерживается в составе ингибиторного комплекса, но лишен прочной ковалентной связи с ним. В условиях избытка протеаз-мишеней комплекс подвергается дальнейшему расщеплению и пептид высвобождается. Комплекс серпин-фермент интернализуется рецептор-опосредованным эндоцитозом и подвергается деградации в лизосомах [154]. Установлено, что SEC-рецептором в клетках гепатоцитов и фибробластов является LRP1 рецептор (low density lipoprotein receptor-related protein 1), член семейства рецепторов липопротеинов низкой плотности [156, 157].

Стоит отметить, что N-гликозилирование ААТ продлевает период полувыведения белка из кровотока (к примеру, за счет маскирования потенциальных сайтов протеолитического расщепления и препятствуя его агрегации/полимеризации) и обеспечивает его противовоспалительные свойства (см. раздел 2.4.2.3), но не влияет на ингибирующую способность как таковую [126, 158, 159, 160]. Молекула рекомбинантного ААТ, в которой отсутствуют N-гликаны, сохраняет свои антипротеолитические свойства, поскольку рекомбинантный негликозилированный ААТ, вводимый пациентам посредством аэрозолизации, образует ингибирующий комплекс *in vivo* [161].

Роль ААТ не ограничивается ингибированием *внеклеточных* сериновых протеаз: обнаружили участие *предполагаемой цитоплазматической* формы ААТ (происхождение которой в работе не обсуждается), продуцируемой опухолевыми клетками, в ингибировании процесса аутофагии, в котором ААТ инактивирует трипсиноподобные ферменты [39]. С одной стороны, возможно проникновение ААТ в клетку извне, которое зависит от экспрессии рецепторов-мусорщиков. Так, продемонстрировано, что ААТ проникает извне в некоторые типы клеток - через легочный эндотелий в альвеолярный эпителий посредством трансцитоза [162]; интернализация ААТ в легочный эндотелий опосредуется рецептором-мусорщиком В типа I (SR-BI) [163]; рецептор LRP1 способствует клатрин-зависимому эндоцитозу ААТ в Т-клетки, в которых ААТ проходит путь эндосомализосома-цитозоль [164]; богатые холестерином липидные рафты рассматриваются как возможные места стыковки ААТ [165]. С другой стороны, нельзя исключить существование цитоплазматической, несекретируемой изоформы белка.

Установлены некоторые факторы, не обусловленные мутантными формами белка, но влияющие на ингибиторные свойства ААТ. Ингибиторный механизм зависит от остатков

метионинов, которые легко окисляются [166]. Окисление остатков Met351 *либо* Met358 приводит к полной утрате ингибиторной способности [167]. Гидрофобные желчные кислоты специфически и необратимо связываются с ААТ в высокомолекулярный комплекс и приводят к потере его ингибирующей способности, а также к полимеризации ААТ [168]. Молекула ААТ подвержена протеолитическому воздействию такими металлопротеазами, как коллагеназа и желатиназа нейтрофилов, стромелизин соединительной ткани, что также приводит к потере ингибирующей активности [169, 170] и высвобождению биоактивных С-концевых пептидов ААТ (см. раздел 2.6).

Полинуклеотиды тРНК и синтетическая ДНК связываются с эластазой, защищая ее от ингибирования ААТ в 30 и 3 раза, соответственно [171]. ДНК длиной 30 п.н. защищает катепсин G от ингибирования ААТ более чем в 3000 раз, в отличие от протеиназы 3 [172]. Гепарин связывается с катепсином G и препятствует образованию комплекса протеаза-ААТ, в то время как на ингибирование эластазы нейтрофилов влияние не столь ярко выражено [173]. Позже обнаружили и прямое взаимодействие ААТ (через K154, K155, K174 в составе лизин-богатой альфа-спирали F - рисунок 2.4, Б) с гепарином с образованием нековалентного комплекса, вследствие чего существенно возрастает активность ААТ против эластазы поджелудочной железы и трипсина [174]. Таким образом, ДНК и гепарин способны *избирательно* модулировать специфичность ААТ.

2.3 Молекулярные механизмы антитрипсиновой недостаточности ААТ на примере Z-варианта ААТ

К снижению анти-протеолитической активности ААТ приводят не только “внешние” факторы, но и “внутренние” - мутации в гене. Ген *SERPINA1* высоко полиморфен. Известны более 200 вариантов аллелей гена *SERPINA1*, и новые, крайне редкие, варианты аллелей периодически обнаруживаются методами секвенирования нового поколения [175, 176]. Аллели передаются по аутосомно-кодминантному менделевскому типу наследования. Нормальная аллель обозначается Pⁱ*(M), а наиболее часто встречающиеся “дефицитные” аллели - Z и S. Большинство генотипов представлены комбинацией трех аллелей: MM генотип обнаруживается у 90-95% популяции (100% от нормального уровня ААТ в сыворотке); 5-10% популяции имеют генотипы MS (80% от нормального уровня ААТ) и более редкие - SS (60%), MZ (55%), SZ (40%) и ZZ (15%) [177]. Таким образом, Z-аллель обуславливает наиболее острую нехватку сывороточного ААТ, представляя наибольший интерес в клинической практике [130]. Концентрация ААТ в крови у Z-гомозигот составляет 3-7 мкМ, в то время как минимальный защитный пороговый уровень ААТ - 11 мкМ [178].

Важную роль в укладке третичной структуры молекулы ААТ играют два солевых мостика (E342-K290 и E264-K387) [130]. Мутации, нарушающие эти взаимодействия, приводят к серьезным патологическим нарушениям структуры и, как следствие, секреции ААТ. S-аллель несет однонуклеотидную замену, приводящую к мутации E264V, которая

расположена в области “затвора” (shutter region, рисунок 2.7, А) и нарушает не только солевой мостик, но и образование ключевой водородной связи с остатком Y38, что приводит к повышенной деградации белка в гепатоцитах, умеренной полимеризации S-варианта ААТ и не ассоциировано с риском заболеваний печени и легких даже в гомозиготном состоянии [179]. Pi*(SZ) генотип способствует развитию обструктивной болезни легких у курящих, поскольку курение приводит к окислению M358 активного сайта ААТ и снижает антипротеолитический потенциал пула ААТ [177].

Z-вариант ААТ обладает пониженными антипротеолитическими свойствами и крайне склонен к полимеризации, которая является причиной “конформационных” расстройств, называемых *серпинопатиями* [180]. Около 10-15% ААТ-Z сворачиваются правильно и в результате секретируются, 70% в итоге подвергается ЭПР-ассоциированной протеасомной деградации в гепатоцитах (ERAD - ER associated degradation), 15-20% - полимеризуются, частично деградируются в процессе аутофагии либо образуют устойчивые включения, обнаруживаемые по окраске остатков сахаров (периодической кислотой - красителем Шиффа) и устойчивости к ферменту, расщепляющему гликоген (диастаза-устойчивые) [181, 182]. Включения цитотоксичны, приводят к повреждениям и заболеваниям печени у 10% новорожденных - гомозигот по Z-аллели и развитию фатального младенческого цирроза, а снижение сывороточного ААТ способствует развитию обструктивной болезни легких [178]. 96% людей с заболеваниями, вызванными дефицитом ААТ, являются носителями генотипа Pi*(ZZ), остальные 4% - комбинации “редких” и “нулл” (особо редких, приводящих к отсутствию *секреторной формы* ААТ) аллелей с аллелями S и Z [177].

Z-аллель несет однонуклеотидную замену, приводящую к мутации E342K в важной структурной области молекулы ААТ у основания RCL - “области разрыва” (breach region) (рисунок 2.7, А), в результате чего укладка ААТ в метастабильную конформацию неэффективна, а конформер-интермедиат, содержащий приоткрытый бета-слой А, стабилизируется. Это приводит к тому, что ААТ-Z в мономерном состоянии нестабилен. Было предложено три модели механизма его полимеризации. Первая модель предполагает, что раскрытый бета-лист А акцептирует мотив RCL соседней молекулы, формируя димер петля-лист (рисунок 2.7, Б), наращивание цепи приводит к образованию более длинных олигомеров с упорядоченной структурой [19]. Пептиды-миметики мотива RCL ингибируют полимеризацию Z-варианта, что свидетельствует в пользу предлагаемой модели [178]. Терапевтическое применение подобных пептидов обсуждается в работе [183].

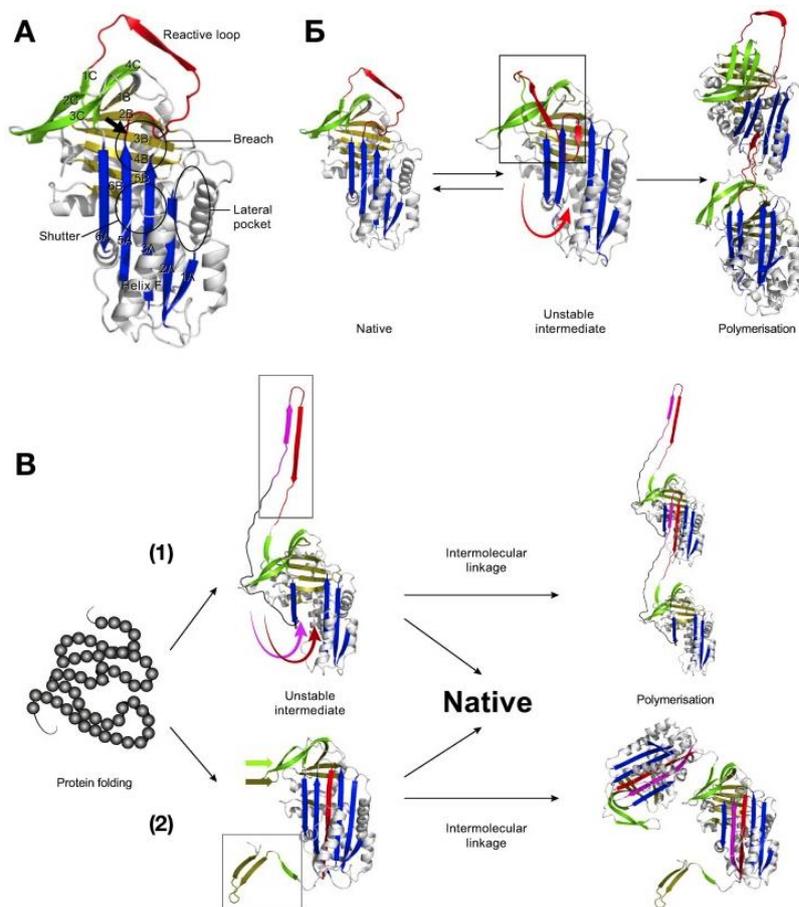


Рисунок 2.7. Модель полимеризации Z-варианта ААТ. А) Кристаллическая структура М-варианта ААТ. Черной стрелкой отмечен остаток Е342; Б) Модель полимеризации Z-варианта ААТ - петля-лист. В) Альтернативные пути полимеризации ААТ: (1) инсерция бета-шпильки s4A-RCL и (2) обмен С-концевым доменом (s1C/s4B/s5B). Рисунок взят из [19].

Образование внутриклеточных включений Z-варианта ААТ усиливается при повышении температуры тела - пирексии. Это наблюдение привело к изучению полимеризации ААТ-Z *in vitro* при нагревании и при денатурации хаотропными агентами (гуанидином или мочевиной) и обнаружению двух дополнительных механизмов полимеризации ААТ-Z - межмолекулярной инсерции бета-шпильки s4A-RCL (при денатурации гуанидином) и межмолекулярным обменом С-концевым доменом s1C/s4B/s5B комбинированным с латентно-подобной аутоинсерцией RCL (при термической денатурации) (рисунок 2.7, В (1) и (2), соответственно). Полимеры *in vivo* близки по структуре к полимерам, полученным нагреванием *in vitro*, поскольку оба типа узнаются моноклональными антителами 2C1, направленными к нео-эпитопу (последовательность которого не указана), экспонирующемуся в результате полимеризации *in vivo* и ассоциированному с заболеванием - дефицитом ААТ [184]. Стоит отметить, что полимеры *in vitro* охотно образуют кольцевые три-, тетра- и пентамеры [185, 186], в отличие от полимеров *ex vivo*, 80% которых не замкнуты [178, 187]. Несмотря на это, именно последняя модель (рисунок 2.7, В (2)) объясняет совокупность характеристик полимеров, выделенных *ex vivo* и полученных нагреванием *in vitro*, которые, действительно, практически идентичны [187].

В отличие от других неправильно свернутых мутантных форм (к примеру, ААТ нулл-ГонгКонг и Саар), ААТ-Z не вызывает стрессорный ответ на неправильно свернутые

белки (UPR - unfolded protein response) в люмене ЭПР, вероятно, ввиду высокой упорядоченности Z-полимеров [188]. Однако клетки, синтезирующие ААТ-Z, гиперчувствительны к ЭПР-стрессу, вызывающему активацию ответа на неправильно свернутые белки, что, как считается, является причиной гепатотоксичности дефицита ААТ [189, 190]. В недавнем исследовании раскрыли механизм гиперчувствительности к ЭПР-стрессу: полимеры ААТ-Z изменяют биофизические свойства белковой среды люмена ЭПР, претерпевая фазовый переход жидкость:твердое тело и становясь молекулярным ситом, замедляющим подвижность белков ЭПР за счет молекулярной фильтрации, зависящей от размера [191]. Фазовый переход инициируется во время ЭПР-стресса через ATF6-опосредованный UPR. Отверждению, удлинению полимеров и как следствие увеличению жесткости белкового матрикса способствует лектиновый шаперон ЭПР - кальретикулин (экспрессия которого контролируется транскрипционным фактором ATF6), который в результате иммобилизуется ситом из полимеров ААТ-Z. Таким образом, иммобилизация важных белковых компонентов ЭПР, таких как шапероны, лежит в основе гиперчувствительности к ЭПР-стрессу при тяжелом дефиците ААТ. Кроме того, дефицит ААТ приводит к гиперактивации системы комплемента [192], синтез Z-мутанта активирует NFκB и последующую секрецию провоспалительных цитокинов, что усугубляет патогенность этой формы [189, 188].

На сегодняшний день единственной утвержденной терапией дефицита ААТ с 1980-х годов является пожизненное еженедельное введение ААТ, очищенного из сыворотки доноров, которая замедляет прогрессирование эмфиземы легких [193]. В настоящее время ведутся клинические исследования потенциальных препаратов, способных снизить гепатоцитарную нагрузку полимерами ААТ-Z, среди которых можно выделить вещества-усилители аутофагии (карбамазепин, флуфеназин), малые молекулы для коррекции фолдинга ААТ-Z и препараты на основе РНК-интерференции [160]. В нашей стране диагностика дефицита ААТ и организация помощи пациентам практически не разрабатывались, однако интерес к этой проблеме в РФ активизировался, чему способствовало включение заболеваний, связанных с дефицитом ААТ, в группу орфанных болезней и появление реальных перспектив лечения людей с таким заболеванием [194].

2.4 Множественность “побочных” функций ААТ

2.4.1 ААТ ингибирует протеазы из других семейств

Специфичность ААТ не ограничивается сериновыми протеазами. На мышинной модели *in vivo* показали, что ААТ обладает кардиопротекторными свойствами и защищает миокард от острого ишемического и реперфузионного повреждения, ингибируя провоспалительную каспазу I (*cysteine-aspartic proteases*) [195], участвующую в цитокиновой сигнализации и относящуюся к семейству цистеин-зависимых аспартат-направленных протеаз. ААТ также ингибирует эффекторные, но не инициаторные, каспазы 3, 6 и 7, защищая клетки от апоптоза [3]. Ингибирование каспазы 3 продемонстрировано на

мышинных альвеолярных эпителиальных клетках, а также в бесклеточной *in vitro* системе с очищенными белками, позволившей выявить, что ингибиторные свойства проявляются молекулой ААТ в нативной/метастабильной конформации [3]. Делеция RCL либо окисление остатков метионинов ААТ значительно снижает ингибиторные свойства по отношению к эффекторным каспазам, в то время как мутантная форма ААТ-Z сохраняет активность [196]. Максимальное ингибирование наблюдали по отношению к каспазе 6 в бесклеточной системе *in vitro*, а также на мышах *in vivo* [196]. В другой работе представлены несколько противоречивые данные о том, что ААТ, как нативный, так и окисленный (утративший анти-эластазную активность), одинаково эффективно ингибирует каспазу 8 (инициаторную) и 3, тем самым повышая выживаемость гепатоцитов в мышинной модели острой печеночной недостаточности [197]. ААТ также ингибирует кальций-зависимую цистеиновую протеазу кальпаин I (μ -кальпаин), индуцируя поляризацию и случайную миграцию нейтрофилов и снижение их адгезии к фибронектину [198], что подразумевает снижение способности нейтрофилов мигрировать по направлению к хемоаттрактантам [199].

ААТ проявляет ингибирующие свойства и по отношению к некоторым металлопротеазам. Нативный ААТ конкурентно ингибирует ADAM-17, также известный как TNF- α конвертирующий фермент (TACE), в бесклеточной системе *in vitro*, в то время как другие формы (окисленная, полимеризованная либо протеолитически инактивированная (С-концевой пептид С36, образующийся после расщепления ААТ)) проявляли существенно сниженные (полимеризованная форма) либо полностью отсутствующие (остальные формы) ингибиторные свойства [200]. Константа скорости ассоциации с ADAM-17 при эквимоллярной концентрации с ингибитором составляет порядка $2.23 \times 10^4 \text{ M}^{-1}\text{сек}^{-1}$. Ингибируя ADAM-17, ААТ сохраняет мембранную форму Fc-рецептора, заякоренного на мембране гликозилфосфатидилинозитолом (Fc γ RIIIb-рецептора), интактной (ААТ напрямую взаимодействует с рецептором и заякоривается на мембране нейтрофилов), и снижает тем самым хемотаксис нейтрофилов, а в условиях дефицита ААТ рецептор отщепляется протеазой от поверхности нейтрофилов, что приводит к увеличению хемотаксиса нейтрофилов в ответ на стимуляцию растворимым иммунным комплексом [200]. Ингибирование ADAM17 наблюдали также и в эндотелиоцитах легких, в результате чего снижалась секреция TNF- α [201]. Стоит отметить, что ААТ может быть *опосредованно* задействован в ингибировании других металлопротеаз, таких как MMP9 (желатиназа В), так как, ингибируя эластазу, препятствует активации соответствующего профермента [202].

2.4.2 Биологические функции ААТ, не зависящие от анти-протеолитической активности

2.4.2.1 ААТ - внеклеточный шаперон

Подобно некоторым другим белкам острой фазы, таким как фибриноген, α 2-макроглобулин, гаптоглобин и α 1-кислый гликопротеин, ААТ также проявляет шапероноподобную активность, ввиду чего его можно назвать *внеклеточным шапероном* [8]. ААТ обнаруживается в местах патологических белковых отложений при различных заболеваниях (болезнь Альцгеймера, боковой амиотрофический склероз, атеросклероз, псевдоэксфолиативный синдром) [8]. Повышенная агрегация β -амилоидного пептида при болезни Альцгеймера ассоциирована с пониженным содержанием ААТ (наряду с цистатином С и транстиретином) в спинно-мозговой жидкости [203]. Содержание ААТ снижается в мозге у людей с болезнью Гентингтона, которая также сопряжена с образованием амилоидоподобных отложений, а на мышинной модели этого заболевания продемонстрировали снижение ААТ на протяжении всей болезни с ранним снижением в тканях мозга, а на терминальной стадии - в печени, сердце и яичках [204]. Наряду с альбумином и иммуноглобулинами, ААТ ингибирует полимеризацию β -амилоидного пептида [205], а вместе с α 1-кислым гликопротеином - агрегацию амилоидогенного предсердного натрийуретического пептида [206], и, наконец, ААТ препятствует агрегации собственных амилоидогенных С-концевых пептидов [207]. Низкий уровень экспрессии *SERPINA1* ассоциирован с развитием семейной амилоидной полинейропатии, связанной с отложением агрегатов мутантных форм транстиретина; ААТ *in vitro* напрямую взаимодействует с транстиретином и препятствует его агрегации [9].

In vitro, ААТ эффективно ингибирует как термически, так и химически индуцированную агрегацию различных тестируемых белков, включая внутриклеточные ферменты алкогольдегидрогеназу, альдолазу, карбоангидразу, каталазу, цитратсинтазу, енолазу, глутатион-S-трансферазу, L-лактатдегидрогеназу [8], - что, несомненно, расширяет его потенциальную значимость в качестве шаперона. Автор предполагает, что именно конформационная пластичность нативной молекулы ААТ позволяет ей выполнять двойную функцию - ингибирование протеаз и шаперонную. Zsila F. умозрительно приходит к выводу, что за шаперонную активность ААТ скорее всего отвечает не подвижный бета-складчатый слой А, который может акцептировать короткие гидрофобные пептиды [208], а скорее протяженная *расщелина*, формирующаяся альфа-спиралями А, G/H и бета-слоем В (рисунок 2.4, Б), по аналогии с известным серпином-шапероном HSP47/*SERPINH1*, предположительно акцептирующим структурные пептидные компоненты проколлагена/коллагена в соответствующую пептидосвязывающую канавку. HSP47 имеет более узкоспецифическую направленность по отношению к субстрату и необходим для поддержания правильной конформации коллагена. В пользу проводимой аналогии свидетельствует тот факт, что мутантные формы HSP47, в которых deletированы ключевые

компоненты, образующие расщелину, утрачивают способность связывать коллаген [209]. Позже получили кристаллическую структуру свободного серпина HSP47 и в его комплексе с субстратом и показали, что шаперон работает в виде димера голова-к-голове и главные структурные компоненты (остатки L381, Y383 - участвуют в гидрофобных взаимодействиях, D385 - образует солевой мостик с субстратом) входят в состав C-концевого домена HSP47 [210]. Интересно, что для шаперонной активности HSP47, который не проявляет антипротеолитических свойств, не требуется значительных конформационных перестроек, и RCL, хоть и находится вблизи ключевых взаимодействий, по-видимому, участия в них не принимает [210].

2.4.2.2 ААТ защищает от окислительного стресса

ААТ содержит экспонированные остатки метионина и один остаток цистеина, которые могут быть подвержены окислению [211, 212]. Следовательно, ААТ может действовать как прямой “поглотитель” активных форм кислорода (АФК) и снижать уровень АФК, высвобождаемых нейтрофилами [167, 213]. Противоокислительные свойства ААТ, по-видимому, проявляются не только через прямое окисление самой молекулы, но имеют более сложные механизмы действия. Так, ААТ проявляет полезные антиоксидантные свойства в контексте преэклампсии в моделях *in vitro* и *in vivo* за счет ингибирования сигнального пути p38MAPK [214]. Это приводит к повышению активности антиоксидантных ферментов, снижению маркеров окислительного стресса и улучшению клеточной функции [214]. Антиоксидантные свойства сохраняются также у C-концевого пептида ААТ (С36) даже при наномолярных концентрациях [215]. Нейтрофилы, выделенные у пациентов с дефицитом ААТ, имеют повышенное содержание связанных с плазматической мембраной p67phox и p47phox компонентов NADPH-оксидазы, а также в них значительно увеличена продукция O_2^- , в то время как ААТ снижает мембранную транслокацию phox-компонентов NADPH-оксидазы и продукцию O_2^- [216]. Нейтрофилы, выделенные у доноров с генотипом Pi*ZZ, продуцируют большее количество перекиси водорода, пероксинитрита и оксида азота и меньшее количество антиоксидантных ферментов каталазы, супероксиддисмутазы и глутатионредуктазы, имеют сниженный потенциал мембраны митохондрий [217]. Таким образом, окислительный стресс является одним из факторов патогенеза при дефиците ААТ.

2.4.2.3 Иммуномодулирующие свойства ААТ

Помимо антипротеолитической роли ААТ, второй широко признанной является его иммуномодулирующая роль. О взаимосвязи ААТ и воспалительных процессов известно еще с тех пор, как обнаружили связь между развитием патологических состояний и фактором риска их возникновения - наследуемым дефицитом ААТ. У людей с дефицитом ААТ наблюдается низкий уровень контроля над такими провоспалительными медиаторами, как IL-1 β , IL-6, TNF- α и IL-8 [218], вследствие чего развиваются различные заболевания, такие как васкулит, бронхоэктатическая болезнь, панникулит [219].

Добавление ААТ к мононуклеарным клеткам периферической крови снижает воспалительный ответ - экспрессию IL-8, IL-6, TNF- α и IL-1 β [220]. ААТ ингибирует синтез и секрецию TNF- α и IL-1 β , индуцированную активацией толл-подобных рецепторов TLR-3, -4, -7/8 соответствующими лигандами, но повышает продукцию противовоспалительного цитокина IL-10 в моноцитах и дендритных клетках, независимо от сохранения ингибиторных свойств, поскольку эффект сохраняется и в случае полимеризованного и окисленного ААТ [221, 222]. Также есть свидетельства о том, что окисленный ААТ активирует популяцию моноцитов, усиливая воспалительный ответ [223], и стимулирует легочные эпителиальные клетки к секреции хемоаттрактанта - моноцитарного хемотаксического белка-1 (MCP-1/CCL2) [211], - таким образом, про-/противовоспалительные свойства ААТ могут зависеть от посттрансляционной модификации молекулы. Рекомбинантный ААТ, заведомо лишенный антипротеолитических свойств путем слияния ААТ с Fc-участком IgG1, так же, как и нативный ААТ, на 80% снижает экспрессию рецепторов TLR2, TLR4 и MHCII на поверхности островковых макрофагов, более того, противовоспалительные свойства рекомбинантного проявлялись при концентрации в 40-100 раз ниже нативного [1], по-видимому, ввиду иной трехмерной конфигурации слитого белка и экспонированию соответствующих активных сайтов молекулы ААТ.

ААТ подавляет ВИЧ-индуцированную активацию транскрипционного фактора NF- κ B, активирующего гены провоспалительных цитокинов, в опухолевой линии моноцитов U1 [5]. Механизм ингибирования ВИЧ-индуцированной активации NF- κ B установлен в работе на CD4⁺ Т-лимфоцитах, демонстрирующей, что ААТ стабилизирует и увеличивает время полужизни его ингибитора I κ B α , маскирующего сигнал ядерной локализации фактора [224]. Это происходит путем прямого связывания с I κ B α и изменения его паттерна убиквитинирования с преобладающей связи K48 на связь K63 предположительно за счет изменения характера последующего взаимодействия ингибитора с убиквитин-лигазным комплексом, а не ингибирования ферментативной активности протеасомы [224]. ААТ при этом проникает извне в CD4⁺ Т-лимфоциты через рецептор LRP1 посредством клатрин-зависимого эндоцитоза и в итоге высвобождается из лизосом в цитозоль [164].

ААТ проявляет также антиген-специфическую иммунорегуляторную активность, отличную от иммуносупрессии, переводя иммунные клетки в толерогенный фенотип, что было продемонстрировано в ряде моделей *in vivo*. Поскольку ААТ снижает уровень провоспалительных цитокинов и повышает уровень противовоспалительных, но в то же время не влияет на уровень провоспалительного цитокина IL-2, обладающего толерогенными свойствами, результирующий цитокиновый профиль способствует приобретению иммунными клетками толерогенного фенотипа [225]. ААТ переводит активность дендритных клеток в полузрелый, IL-10- и CCR7-положительный толерогенный фенотип, который способствует разрешению воспалительных процессов [226]. В мышинной

модели диабета 1 типа монотерапия ААТ после аллотрансплантации островков бета-клеток способствует выработке длительной *донор-специфической* (антиген-специфической) иммунотолерантности - популяции Treg клеток в организме реципиента, которая характеризуется набором иммуносупрессорных маркеров (таких как Foxp3, CTLA-4, TGF-beta, IL-10 и антагонист рецептора IL-1 (IL-1RA)) и низкой либо нулевой экспрессией провоспалительных цитокинов [227]. Для ААТ-опосредованного увеличения Treg-клеток необходима интактность воспалительных сигнальных путей, - ААТ, по-видимому, требует стимулированного микроокружения для ускорения разрешения воспаления, что отличает его от классических противовоспалительных средств [226, 228].

ААТ не способствует восприимчивости к вирусам и возникновению опухолей - скорее наоборот [5, 225, 229], а значит, не снижает активность натуральных киллеров, в то время как модулирует их взаимодействие с дендритными клетками, повышая выживаемость бета-клеток островков Лангерганса [230, 231]. ААТ избирательно поддерживает метаболическую активность иммунных клеток (натуральных киллеров, дендритных и T-регуляторных, но не эффекторных, клеток), способствует защите от реакции “трансплантат-против-хозяина” и усилению реакции “трансплантат-против-лейкемии” в мышинной модели трансплантации костного мозга [225]. Вероятность развития реакции “трансплантат-против-хозяина” обратно коррелирует с уровнем ААТ в крови доноров [225]. ААТ ограничивает пролиферацию В-лимфоцитов и переключение изотипа антител в сторону IgG, сохраняя при этом продукцию IgM антител и повышает количество циркулирующих Treg клеток при аллотрансплантации кожи, снижая отторжение [232]. Терапия ААТ у лиц без дефицита ААТ может препятствовать прогрессированию заболевания при диабете 1 и 2 типа, остром инфаркте миокарда, ревматоидном артрите, воспалительном заболевании кишечника, муковисцидозе, отторжении трансплантата, реакции “трансплантат-против-хозяина” и рассеянном склерозе [233]. Уровень ААТ повышается в третьем триместре беременности [117], что подчеркивает его значимость как толерогенного агента [234]. Низкие концентрации ААТ в плазме крови повышают риск возникновения самопроизвольных абортов [235] и преждевременного разрыва плодных оболочек [236].

В то время как ААТ преобразует макрофаги и дендритные клетки в M2-подобный противовоспалительный и толерогенный фенотипы, соответственно, ингибируя аутоиммунную активность NK-клеток, в специфических условиях, таких как опухоль, ААТ, наоборот, повышает инфильтрацию опухолевой ткани M1-подобными провоспалительными макрофагами наряду с CD8+ T-клетками, ингибируя рост и метастазирование опухоли *in vivo*, - то есть иммуномодулирующая активность ААТ зависит от биохимического контекста [229]. Было выявлено, что модифицированная нитрозилированная форма ААТ (S-NO-ААТ), в которой ААТ может существовать в опухолевом микроокружении, способствует дифференциации макрофагов в

провоспалительный M1-фенотип [229]. Нитрозилирование единственного остатка цистеина C232 в зрелой молекуле ААТ (S-NO-ААТ), происходящее в условиях избытка NO также при инфекциях, не сказывается на эффективности ингибиторных свойств, более того, модифицированный таким образом ААТ обладает провоспалительными свойствами и проявляет в 10-1000 раз более активные бактериостатические свойства, чем NO, S-NO-альбумин или S-NO-глутатион, и является своего рода “NO-отводом” [237, 238].

Вышеперечисленные данные об иммуномодулирующей активности ААТ свидетельствуют в пользу того, что белок способствует распознаванию иммунной системой ситуаций, когда уместно остро реагировать на “подлинные” угрозы, а когда каскады иммунного реагирования, преумноженные петлей положительной обратной связи, чрезмерны и приводят к сопутствующим повреждениям здоровых клеток [234].

Иммуномодулирующие свойства ААТ реализуются через множественные пути. Противовоспалительные свойства ААТ обеспечиваются, по-видимому, сайтами в структуре молекулы вне RCL [239, 240]. ААТ выполняет “побочные” биологические функции во многом благодаря тому, что его структурные компоненты - отличные от задействованных при ингибировании протеазы - взаимодействуют с другими молекулами, как белковой, так и не белковой природы, преимущественно через электростатические и гидрофобные взаимодействия, в том числе медиаторами воспаления и специфическими рецепторами. Так, выяснили, что именно за счет образования нековалентного комплекса между ААТ и фактором системы комплемента анафилотоксином C3a (а не отдельных компонентов вне комплекса), который контаминирует в некоторых коммерческих препаратах ААТ, повышается синтез противовоспалительного цитокина IL-10, маркера иммунной толерантности [241].

Ингибирующее действие ААТ на воспалительный ответ нейтрофилов, помимо указанных выше механизмов (ингибирование кальпаина I и ADAM17) происходит также напрямую и не зависимо от антипротеолитических свойств, через эффективное прямое связывание и инактивацию IL-8, - сильного хемотаксического лиганда, взаимодействие которого со своим рецептором CXCR1 на поверхности нейтрофилов приводит к их активации [200]. Под действием ААТ наблюдали ингибирование полимеризации F-актина у лидирующего края клеток, что дополнительно свидетельствует об ингибировании хемотаксического каскада. Интересно, что взаимодействие с IL-8 происходит специфически, поскольку сывороточный альбумин похожие свойства не проявлял, и *через остатки сахаров* ААТ, поскольку рекомбинантный негликозилированный ААТ утрачивал эти свойства [200].

ААТ также напрямую связывает другой хемотаксический агент - лейкотриен В4 (LTB4), препятствуя его взаимодействию с рецептором и последующей адгезии и дегрануляции нейтрофилов [242]. Окисленный ААТ сохраняет связывающие свойства, пептид С36 их не проявляет, и, в отличие от IL-8, остатки сахаров не принимают участия во

взаимодействии с LTB₄ [242]. Наиболее вероятно, что LTB₄ стыкуется в гидрофобный кармашек ААТ, состоящий из бета-тяжа s2A и альфа-спиралей D и E (рисунок 2.4, Б) [242].

На мышинной модели эмфиземы легких, индуцированной эластазой, показали, что связывание ААТ с липопротеином высокой плотности усиливает защитный эффект ААТ (снижение уровня провоспалительных цитокинов, количества нейтрофилов и макрофагов в бронхоальвеолярной лаважной жидкости) [243]. Также показано, что ААТ ослабляет воспаление, вызванное повреждением, путем связывания с некоторыми молекулярными фрагментами, ассоциированными с повреждениями (DAMPs), включая gp96, HSP70 и HMGB1, *вне клетки* [244, 245, 246]. Открытым остается вопрос, способен ли ААТ связывать эти молекулы двойственной функциональности *внутри клетки* и какие именно сайты ААТ ответственны за связывание с DAMPs? Исследования в этом направлении активно ведутся [226].

ААТ связывается с гемом с высокой константой ($K(D) 2.0 \times 10^{-8} M$), близкой к таковой для сывороточного альбумина ($1.26 \times 10^{-8} M$) [247]. По-видимому, это связывание обеспечивает способность ААТ ингибировать провоспалительные реакции в клетках эндотелия, вызванные свободным гемом. В этом процессе ААТ ингибирует гемопосредованное защелачивание лизосом [248]. Интересно, что при добавлении ААТ совместно с гемом существенно повышается интернализация ААТ в клетки эндотелия [248], что может свидетельствовать об изменении конформационного состояния ААТ в комплексе с гемом и открытии пентамотива для связывания с SEC-рецептором.

Отдельного внимания заслуживает взаимодействие ААТ с жирными кислотами. Было обнаружено, что препараты очищенного ААТ из плазмы доноров отличаются по составу: широко используемый в заместительной терапии препарат Prolastin, в отличие от препарата Zemaïra, содержит ААТ в комплексе с такими ненасыщенными жирными кислотами, как линолевая и олеиновая. Именно в такой формуле ААТ повышает продукцию ангиопоэтин-подобного белка 4 (Angptl4, который является также сильным противовоспалительным фактором), белок 4 связывающий жирные кислоты (FABP4) и транслоказу жирных кислот CD36 [249]. Эффекты комплекса ААТ с жирными кислотами опосредуются через рецепторы, активируемые пролифераторами пероксисом (peroxysome proliferator-activated receptors — PPARs), — группы транскрипционных факторов (PPAR β/δ и PPAR γ) [249], выполняющих противовоспалительную роль в ряде заболеваний с фоновым воспалительным компонентом [250].

Интересно отметить, что ААТ *in vitro* напрямую взаимодействует с полностью транс-ретиноевой кислотой (ПТРК) - метаболитом витамина А, образуя нековалентный комплекс со стехиометрическим соотношением 1:2 (предположительно через остатки W194 и W238, анти-протеолитические свойства ААТ в комплексе сохраняются) и существенно повышая его растворимость в воде (от 0,2 до 87 мкМ) и время жизни в водном растворе (до нескольких месяцев) [251]. Биологическая значимость комплексообразования не известна,

но можно предположить, что, во-первых, ААТ снижает цитотоксичность ПТРК, и во-вторых, поскольку оба компонента комплекса имеют однонаправленное действие и способствуют установлению толерогенного фенотипа иммунных клеток [252, 253, 254], их объединение в комплекс может усиливать действие обоих.

Учитывая способность ААТ связывать ненасыщенные жирные кислоты и ПТРК, вероятно, что ААТ, находясь в комплексе с двумя лигандами одновременно (что требует проверки!), становится “гетеролигандом” (или гетеробифункциональным индуктором) - что менее вероятно, либо является “доставщиком” разных лигандов (в итоге уступая в конкуренции за лиганд с рецептором) и способствует коактивации гетеродимеров рецепторов, активируемых пролиферацией пероксисом (PPAR) и рецепторов ретиноевой кислоты (RAR) с ретиноидными X-рецепторами (RXR). Известно, что коактивация ядерных рецепторов в состоянии ДНК-связанного гетеродимера PPAR/RXR сигергестически усиливает транскрипционный отклик в генах-мишенях [255]. Не исключено, что комплекс ААТ-жирные кислоты является прямым лигандом PPAR рецепторов - вероятность этого обсуждается в работе [249]. С другой стороны, в работе [251] ААТ помимо ПТРК связывает также коричную, но не миристиновую, кислоту, что подчеркивает его селективность во взаимодействиях с лигандами. Несмотря на то, что в работе не рассматривались изомеры ПТРК (9-цис-ретиноевая кислота - агонист RXR рецепторов), не исключено, что ААТ связывается и с ними, и наиболее интересно было бы изучить его взаимодействие с 9-цис-13,14-дигидроретиноевой кислотой - физиологически значимым эндогенным активатором RXR рецепторов млекопитающих [256], являющихся облигаторным гетеропартнером для других ядерных рецепторов. Также стоит отметить, что гем, с которым ААТ взаимодействует со сравнительно высокой константой, является лигандом для ядерных рецепторов REV-ERBa и REV-ERBβ, которые являются связующим звеном между циркадными ритмами и метаболизмом млекопитающих [257].

Механизм иммуномодулирующего действия ААТ может быть также связан с влиянием ААТ на работу белков, ассоциированных с холестерин-богатыми липидными рафтами, поскольку известно, что ААТ аккумулируется на липидных рафтах на поверхности моноцитов [165], а также связывается с липопротеинами низкой и высокой плотности, а окисленный ААТ детектирован в комплексе с липопротеинами низкой плотности *in vivo* [258, 259]. Известно, что ААТ напрямую взаимодействует с LRP1-рецептором [156, 164], который, помимо участия в эндоцитозе своих лигандов, также проводит сигналы внутрь клетки [260]. Запуск сигнальных путей через рецептор LRP1 ингибируется при дезинтеграции липидных рафтов [261].

LRP1 широко представлен в моноцитах, макрофагах и дендритных клетках [262], составляющих, помимо нейтрофилов [263], основные клеточные мишени ААТ. Действительно, агонисты рецептора ингибируют провоспалительный ответ в макрофагах, а антагонисты его активируют [264]. LRP1 ингибирует LPS-индуцированный

провоспалительный ответ (активацию JNK-сигнального каскада), а в присутствии лиганда - аполипопротеина E - этот эффект усиливается [265]. Лактоферрин - другой лиганд - через LRP1-опосредованный эндоцитоз проникает в клетку, связывается с TRAF6, ингибируя активацию NFκB и продукцию воспалительных цитокинов [266]. LRP1 подавляет активность дендритных клеток при аллергической астме [267]. Действительно, было продемонстрировано прямое участие рецептора LRP1 в запуске противовоспалительных (ингибирование активации NF-κB после стимуляции провоспалительными агентами в модели *in vitro*) и противоапоптотических путей (через фосфорилирование протеин-киназы Akt в *in vivo* модели ишемии-реперфузии миокарда) под действием синтетического пептида SP16, гомологичного C-концевому домену ААТ [268].

Однако, для узнавания рецептором необходим доступ пептида FVFLM [154] в C-концевом домене ААТ. С одной стороны, противовоспалительные каскады могут запускаться новообразованными комплексами ААТ-протеаза, в которых пептид экспонирован, при их локальной утилизации клетками, экспрессирующими рецептор. С другой стороны, возникает вопрос, может ли свободная молекула ААТ принимать подходящую конформацию, без образования ингибиторного комплекса? В зрелой молекуле ААТ всего 9 остатков метионина, из которых метионин из пептида находится в положении 374. В работе [269] изучали подверженность остатков метионина окислению пероксидом водорода при разных рН. При физиологическом рН окислению подвергается пять остатков, в то время как M374 *доступен* для окисления только при *кислом* рН. Также, умеренно кислая среда вызывает *локальные* конформационные изменения структуры ААТ около единственного остатка цистеина C232, усиливая его окисляемость [270]. С учетом того, что ААТ сохраняет *глобальную* третичную структуру и антипротеолитические свойства в широком спектре рН [271], вместе, эти данные свидетельствуют в пользу того, что *в кислой среде частично окисленная молекула ААТ* изменяет конформацию и пептид экспонируется. Реализация иммуномодулирующей активности ААТ через пептид объяснило бы и отсутствие необходимости сохранения антипротеолитических свойств. Известно, что при воспалительных процессах (а также в опухолях) среда закисляется, то есть локально, в местах воспаления молекула ААТ может окисляться, принимая “подходящую” конформацию, и преобразовываться в *сигнальную* молекулу, настраивая местный иммунитет.

При воспалении LRP1 рецептор подвергается протеолитическому расщеплению, внеклеточный домен sLRP1 высвобождается и активирует p38MAPK и JNK каскадный путь, усиливая провоспалительные реакции. Интересно, что главным ферментом, осуществляющим расщепление, является “шеддаза” ADAM17 [272]. ААТ, ингибируя металлопротеазу [200], может являться регулятором физиологической активности рецептора.

Помимо действия через LRP1, недавно было обнаружено, что противовоспалительная активность ААТ отчасти опосредуется также через его прямое взаимодействие с глюкокортикоидным рецептором в цитоплазме макрофагов [273]. Вклад в противовоспалительную активность ААТ также вносят и конформационные сайты вне RCL молекулы, предсказанные *in silico* и подтвержденные с использованием синтетических пептидомиметиков, действующих схожим образом с ААТ из плазмы доноров, а в некоторых случаях превосходящих по амплитуде действия свой прототип [240]. Предполагается, что эти сайты участвуют в межмолекулярных взаимодействиях, однако, партнеры по связыванию не известны.

2.5 ААТ - опухолевый биомаркер

Ключевой стадией опухолевой прогрессии можно считать перестройку ядерных процессов трансформированных клеток и формирование механизмов управления стромой, способствующих разрушению внеклеточного матрикса и поддержанию процессов инвазии и метастазирования [274]. Протеолитическое ремоделирование внеклеточного матрикса рассматривается как предпосылка и следствие миграции инвазивных клеток опухоли [275], а окооклеточный протеолиз играет важную роль в мезенхимальном типе движения неопластических клеток, претерпевших эпителиально-мезенхимальный переход [276]. Исходя из проканцерогенного вклада протеолитических процессов и анти-протеолитической роли ААТ в организме, можно ожидать, что ААТ обладает противоканцерогенными свойствами. Однако, имеющиеся данные свидетельствуют об обратном.

Развитие онкологических заболеваний часто сопровождается активацией экспрессии гена *SERPINA1* и повышением количества белка в крови и тканевых жидкостях. Белок начинает синтезироваться *de novo* также в тканях, для которых в норме это не свойственно. Этот феномен выявлен для опухолей печени [277], поджелудочной железы [278], легких [31, 279], простаты [32, 280], желудка [33], гортани [281], толстой и прямой кишки [34, 35], молочной железы [282, 283], шейки матки [284] и мочевого пузыря [285]. Иммуногистохимические исследования показывают, что пациенты с ААТ-положительными аденокарциномами желудка [286] и легких [287, 288, 289] имеют наименее благоприятный прогноз выживания. Высокий *сывороточный* уровень ААТ является диагностическим и прогностическим фактором и коррелирует с менее благоприятным прогнозом при раке печени [277], поджелудочной железы [278] и легких, где также наблюдается положительная корреляция между сывороточным уровнем и иммуногистохимическим окрашиванием тканей [279], в то время как снижение уровня ААТ до нормы может являться маркером ответа пациента на лечение [32, 282, 290]. Все перечисленные свидетельства, с одной стороны, могут отражать провоспалительный ответ организма на патологию, а с другой - нельзя исключить участие ААТ в канцерогенезе в качестве *проканцерогенной функциональной молекулы*.

Прямая корреляция уровня экспрессии *SERPINA1* и опухолевой прогрессии справедлива не для всех типов рака. Так, для опухоли яичников установлена обратная корреляция уровня экспрессии гена *SERPINA1* (определенная методом ПЦР в реальном времени на образцах ткани) и стадии рака, а в опухолях с низким злокачественным потенциалом (LMP) ААТ чаще детектируется в ядре, чем в инвазивных опухолях (TOV) (происхождение ядерной протеоформы ААТ не известно) [40]. Измерение концентрации ААТ в сыворотке в этой работе не проводилось. Уровень экспрессии гена *SERPINA1* в образцах ткани погранично-инвазивных опухолей (но относящихся к группе TOV) ближе к группе LMP, ввиду чего авторы предполагают участие *SERPINA1* в опухолевой прогрессии. Интенсивность окрашивания ААТ в цитоплазме опухолевых клеток коррелирует со степенью злокачественности. Авторы объясняют это тем, что менее агрессивные опухолевые клетки более эффективно секретируют ААТ, в то время как более агрессивные клетки накапливают “неактивный” белок в цитоплазме. Данное наблюдение, и в особенности присутствие ААТ-подобного белка в ядрах трансформированных клеток, наводит на мысль о том, что при опухолевой трансформации образуются неизвестные протеоформы ААТ.

Работ по установлению роли ААТ в опухолевой прогрессии немного, и некоторые имеющиеся в настоящее время исследования могут вводить в заблуждение, поскольку сфокусированы на изучении противоопухолевых свойств экзогенно вводимого, коммерчески доступного очищенного ААТ из плазмы со “здоровым” структурным фенотипом (ниже описаны изменения в профиле гликозилирования ААТ при раке). Так, сывороточный ААТ и его вариант без С-концевого домена ингибируют рост опухоли и прорастание опухоли сосудами в мышинной модели *in vivo* [291]. Сывороточный и искусственно экспрессируемый ААТ ингибирует секрецию митогена TGF- α , пролиферацию клеток опухоли молочной железы MCF7 и образование колоний в мягком агаре [292, 293]. Сывороточный ААТ и синтетический пептид С36, хоть и повышают продукцию VEGF, но снижают пролиферацию (ААТ) и инвазию (С36) клеток немелкоклеточного рака легких *in vitro* [294]. Однако при *длительном* росте в среде с добавлением ААТ из плазмы в течение 3 недель клетки немелкоклеточного рака легких активнее пролиферировали и проявляли устойчивость к стауроспорин-индуцируемому апоптозу: ААТ блокировал ингибирование АКТ/МАРК сигнального пути, активацию прокаспазы-3 и аутофагии [4].

Недавние исследования наиболее близко подошли к раскрытию роли ААТ в процессе канцерогенеза, поскольку в них изучали *эндогенно* синтезируемый опухолевой клеткой ААТ, и установили, что ААТ не побочный продукт, а непосредственный участник злокачественных процессов. В одной работе сравнили две клеточные линии аденокарциномы легких - с низким и высоким метастатическим потенциалом - по профилю секретируемых белков и обнаружили корреляцию между повышенной продукцией ААТ (в

том числе путем подавления или активации экспрессии) и приобретением более агрессивного опухолевого фенотипа, связанного с миграцией и инвазией, но не пролиферативной активностью [295]. Показали, что ААТ необходим для сборки околоточного фибронектина FN1, необходимого в процессе метастазирования для адгезии и запуска экстравазации опухолевой клетки в эндотелий капилляров через рецептор - сериновую экзопептидазу дипептидилпептидаза IV (DPP IV), однако механизм участия ААТ в этом процессе в данной работе не раскрыт [295]. В другой работе подтвердили участие ААТ в метастазировании клеток аденокарциномы легких и продемонстрировали роль ААТ в качестве непрямого положительного регулятора экспрессии фибронектина, и показали, что околоточный фибронектин способствует метастазированию через связывание с интегрином $\alpha 5$ на поверхности эндотелиальных клеток [289]. Авторы предполагают механистическое участие ААТ в экспрессии фибронектина путем ингибирования его протеолитической деградации, однако антипротеолитические свойства ААТ, продуцируемого изученными в этой работе клетками, продемонстрированы, к сожалению, не были. В обеих перечисленных работах по раку легких подтвердили также и прогностическую значимость ААТ при этом заболевании, а в последней выявили прямую корреляцию между ААТ-положительными опухолями и частотой метастазирования в регионарные лимфатические узлы. На клеточных линиях опухоли толстого кишечника показали, что экспрессия *SERPINA1* положительно регулирует уровень фибронектина, в то время как обратной зависимости не обнаружено [296]. Также продемонстрировали роль ААТ, искусственно экспрессируемого клетками нормального эпителия легких, в качестве положительного регулятора уровня транскрипционного (STAT5B) и трансляционного (EEF1A2) факторов, а также белка везикулярного транспорта GOPC, ингибирующего экспрессию белка аутофагии BECN1, тем самым способствуя выживаемости клеток; гиперэкспрессия ААТ также повышала активность кэп-зависимой трансляции в среднем на 50%, интенсивность ангиогенеза и клеточной адгезии через увеличение экспрессии метастатического белка тромбоспондина 1 (THBS1) [297]. И наоборот, нокаунт *SERPINA1* снижал пролиферацию, метастазирование и адгезивные свойства клеточной линии A549 и опухолевых клеток в мышинной модели *K-ras*^{LA1} немелкоклеточного рака легких *in vivo* [297].

ААТ индуцирует эпителиально-мезенхимальный и эндотелиально-мезенхимальный переход, способствует устойчивости клеток опухоли легких к противоопухолевому препарату цисплатину и поддержанию клеточной стволовости [298]. Результаты по миграции клеточных линий и сублиний опухоли яичника предполагают, что ААТ способствует также мезенхимально-амебoidalному переходу, - более подвижному фенотипу опухолевых клеток, не требующему протеолиза внеклеточного матрикса [40] и являющемуся частью программы эпителиально-мезенхимального перехода [299].

При опухолевой прогрессии модулируется структура сывороточного ААТ, что может сказываться на его функциональности: изменяется тип гликозилирования молекулы ААТ, ассоциированный с неблагоприятным прогнозом, в сторону большего ветвления гликанов, корового фукозилирования, терминального сиалирования и фукозилирования с образованием эпитопа “сиалил-Льюис х” (который состоит из остатка сиаловой кислоты, присоединенной к α -1,3-фукозилированному N-ацетилглюкозамину через остаток галактозы α -2,3-гликозидной связью) [128, 300, 301, 302], - потенциально способствующего экстравазации в процессе метастазирования (поскольку ААТ может заякориться на мембране клеток). Вклад в повышение концентрации ААТ, в том числе ААТ с измененным типом гликозилирования, при развитии онкологических заболеваний могут вносить и опухолевые клетки. Клеточные линии, выведенные из опухолей легкого, почки и кишечника, также секретируют ААТ с измененным гликофенотипом [37]. Продукция ААТ выведенными из опухолей клеточными линиями выявлена и в других работах [38, 295, 303, 304, 305]. Кроме того, секретируя провоспалительные цитокины, опухолевые клетки могут дистанционно влиять на продукцию ААТ в печени. ААТ, секретируемый опухолевыми клетками, может сохранять толерогенные свойства, описанные ранее, способствуя уклонению опухолевых клеток от иммунного надзора и опухолевой прогрессии [306], а ААТ с более разветвленными гликанами (не взаимодействующими с лектином конканавалином А, специфичным к биветвящимся сахарам) эффективно ингибирует цитотоксическое действие натуральных киллеров, играющих основную роль в защите организма от трансформированных клеток [307].

Исходя из противоопухолевых свойств “нормального” ААТ, можно предположить, что на ранних стадиях канцерогенеза, когда циркулирующий ААТ еще не претерпел структурные, опухоль-ассоциированные, изменения на системном уровне, белок препятствует опухолевой прогрессии, преимущественно за счет антипротеолитической активности, но впоследствии его свойства меняются в сторону онкогенных.

Сохраняет ли опухоль-ассоциированный ААТ антипротеолитические свойства? Wozniak A. с соавторами изучили соотношение протеолитической активности катепсина D и антипротеолитических свойств ААТ как независимые параметры, так как ААТ не ингибирует этот фермент, в сыворотке крови пациентов с опухолью молочной железы. Авторы выявили, что высокая активность фермента и низкая активность ААТ по сравнению с нормой наблюдается у пациентов до операции, в то время как спустя полгода после удаления опухоли соотношение изменилось на противоположное [308]. Вторая работа этих же авторов подтверждает первые наблюдения, но у пациентов с опухолью шейки матки: восстановление активности ААТ до нормы происходит спустя полгода после брахитерапии [309]. Высокая активность катепсина D и низкая активность ААТ предлагается в качестве маркера на степень малигнизации опухоли молочной железы, рецидива и множественного метастазирования при опухоли шейки матки [308, 309]. Высокая ингибиторная активность

ААТ является маркером хорошего прогноза при опухоли простаты [310]. У пациентов с плоскоклеточным раком пищевода снижаются ингибиторные свойства ААТ, несмотря на общий повышенный уровень [311]. Эндогенный ААТ, секретируемый клетками гепатоцеллюлярной карциномы HepG2, проявляет сниженные противотрипсиновые свойства по сравнению с эквимоллярным количеством рекомбинантного ААТ, продуцируемого этими же клетками, или ингибитора из плазмы [312]. ААТ, секретируемый клеточной линией меланомы ME 1477, не обладает ингибиторными свойствами [313]. Снижение ингибиторной способности ААТ при развитии онкологических заболеваний может быть связано с протеолитическим расщеплением и инактивацией ААТ с образованием С-концевых пептидов ААТ.

2.6 С-концевой домен ААТ, кодируемый пятым экзоном гена *SERPINA1*, - источник биологически активных пептидов

Функционирование белка зависит от его поверхности, обращенной к внешней среде и доступной для распознавания другими молекулами. Конформация белка предопределяется его первичной структурой, значительная часть которой скрыта в ядре белка в его пространственной конфигурации. Изменение тонкой структуры белка ввиду посттрансляционных модификаций (ПТМ) может в конечном итоге привести к переключению его функционального экстерьера. “Экстремальным” инструментом для извлечения спящих функциональных мотивов белка является ограниченный протеолиз - необратимая ПТМ, которая извлекает более диффузные усеченные фрагменты белка окружающему биохимическому контексту. Альфа1-антитрипсин, ингибитор протеаз, может быть отнесен к ярким примерам белков, чья функциональная двойственность полностью проявляется, когда он подвергается протеолизу.

В условиях избытка специфических (таргетных) протеаз, а также при действии различных тканевых металлопротеаз, продуцируемых иммунными клетками в местах воспаления и опухолевыми клетками, происходит расщепление молекулы ААТ в сайте активного центра (M358-S359), а также в близлежащих к нему сайтах на большой и малые фрагменты, - таким образом, активность ингибитора подавляется (рисунок 2.8) [314, 315]. Инактивация и расщепление ААТ не только сдвигает баланс протеаза–антипротеаза в сторону протеолиза, но также способствует образованию различных С-концевых пептидов ААТ, проявляющих биологическую активность и обладающих свойствами, не характерными для крупного аналога. Структура и свойства пептидов собраны в таблице 2.2.

название	структура	где обнаружен	установленная роль или предположения авторов	источник (DOI/PMID)
CRISPP (Cancer-associated, SCM-Recognition, Immunedefense-suppressing, and Serine Protease-Protecting)	C29 (R/V/S/M)IPP(E/D)VKFNKPFV FLMIEQNTK(V/T/R)PLFMGK; C35 (V/R/S)IPPEVKFNKPFVFLMI EQNTKSPLFMGKVVNPT(Q/E)	периферическая кровь больных онкологией различного генеза; не выявляются у здоровых людей	Пептиды вызывают структурные изменения в цитоплазматическом матриксе (SCM - Structuredness of the Cytoplasmic Matrix) в лимфоцитах от доноров с опухолевыми заболеваниями. Пептиды полностью ингибируют цитотоксичность клеток NK и LAK в количестве 35 фмоль на эффекторную клетку в течение 3 часов <i>in vitro</i> и снижают секрецию TNFα. Защищают трипсин, катепсин G, плазмин и эластазу от ингибирования полноразмерным ААТ, предположительно из-за образования защитного комплекса протеаза-CRISPP, предотвращая стыковку ААТ, но сохраняя протеолитическую активность фермента. Фиолетовый - "эпитоп распознавания опухолевого SCM"; курсив - эпитоп-супрессор NK/LAK; синий - эпитоп защиты сериновой протеазы.	Cercek, L. & Cercek B., 1992, 1993 (PMID: 1473120; PMID: 7691406; PMID: 1473120)
C36; C42 эндогенные	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK; C42 LEAIPMSIPPEVKFNKPFVFL MIEQNTKSPLFMGKVVNPTQ K	в фосфолипидной фракции селезенки (C36) и желчи (C36, C42)	Роль не установлена; предполагается влияние пептидов на формы холестерина	Johansson J. et al., 1992 (DOI: 10.1016/0014-5793(92)80234-8)
сигнальный пептид (TIMP-1)- C37	C37 MSIPPEVKFNKPFVFLMIEQN TKSPLFMGKVVNPTQK	оверэкспрессирован в клеточной линии аденокарциномы поджелудочной железы SUIT-2	У голых мышей <i>in vivo</i> сублиния с оверэкспрессией образует в 1,5-3 раза большие опухоли с большей инвазивностью и метастазами в лимфатические узлы. Преимущество роста отсутствует у мышей с делецией NK, то есть действие опосредовано ингибированием NK-клеток.	Kataoka H. et al., 1999 (DOI: 10.1016/s0002-9440(10)65292-3)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK; C5 FVFLM; C20	<i>in vitro</i>	C5 и C36 образуют амилоидные фибриллы <i>in vitro</i> , полноразмерный ААТ из плазмы ингибирует образование фибрилл и деполимеризует ранее сформированные фибриллы; C20 не образует фибрилл <i>in vitro</i> . ААТ в присутствии C5 и C36 полимеризуется и повышается его термостабильность.	Janciauskiene S. et al., 1995 (DOI: 10.1515/bchm3.1995.376.7.415)

	IEQNTKSPLFMGKVVNPTQK			
C36 синтетический в виде фибрилл	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	Фибриллы C36 повышают экспрессию мРНК ЛПНП-рецептора (LDLR) за счет увеличения ядерной формы фактора транскрипции SREBP-1, в результате чего увеличивается связывание ЛПНП с ЛПНП-рецептором и его поглощение клетками HepG2, эффект ингибируется антителами к рецептору и холодным ЛПНП; ингибирует деградацию 125I-ЛПНП и синтез желчных кислот и не влияет на биосинтез холестерина. Таким образом, фибриллы C36 способствуют захвату ЛПНП клетками гепатоцитов без его последующего метаболизма и нарушают гомеостаз холестерина - проявляют цитотоксические свойства. Фибриллы обладают атерогенными свойствами: они повышают связывание ЛПНП моноцитами, экспрессию ЛПНП-рецептора, CD36, активность глутатионредуктазы, синтез ААТ, и в течение 24 часов под влиянием фибрилл моноциты дифференцируются в макрофаги.	Janciauskiene S., Lindgren S., 1999; Janciauskiene S. et al., 1999 (DOI:10.1002/hep.51 0290217; DOI:10.1016/s0021- 9150(99)00194-x)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	Пептид является хемоаттрактантом нейтрофилов и активатором их адгезии, дегрануляции и продукции супероксидного анион-радикала.	Janciauskiene S. et al., 2004 (DOI:10.1016/j.bbrc. 2004.01.055)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	C36 устраняет эффект среды, кондиционированной полиморфноядерными нейтрофилами (PNM), усиливающей инвазивность клеток опухоли легкого. C36 снижает инвазивность клеток опухоли на 51,4%.	Zelvyte I., et al, 2004 (DOI: 10.1186/1475- 2867-4-7)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	Действие пептида в течение суток повышают экспрессию MCP-1 (моноцитарный хемоаттрактантный белок 1), секрецию TNF α и IL-6, интерстициальной коллагеназы (MMP1) и желатиназы В (MMP9), увеличивает потребление кислорода на 30% (полноразмерный ААТ ингибирует на 50%).	Moraga F. et al., 2001 (DOI: 10.1006/abbi.2000.22 11)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	Пептид индуцирует продукцию желатиназы В, MCP-1, IL-6 моноцитами, повышенное потребление кислорода через рецептор CD36 (связывание окисленных ЛПНП), так как эффекты исчезают при преинкубации с анти-CD36 антителами, при этом индукция секреции TNF α опосредуется через LDLR. .	Janciauskiene S. et al., 2001 (DOI:10.1016/s0021- 9150(00)00767-x)

C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	C36, в отличие от нативного ААТ, повышает уровень TGF-beta 1, секрецию TNF α и IL-6 клетками аденокарциномы поджелудочной железы. C36, in contrast to native AAT, increases the level of TGF-beta 1, secretion of TNF α and IL-6 by pancreatic adenocarcinoma cells.	Zelvyte I. et al., 2003 (PMID: 12894502)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	C36, в отличие от нативного ААТ, увеличивает включение 3Н-тимидина в клетки меланомы ME 1477 на 35% (ААТ снижает на 53%), уровень проколлагеназы-1 повышает в 1,5 раза (ААТ снижает в 1,5 раза).	Zelvyte I. et al., 2002 (DOI: 10.1016/s0361-090x(02)00090-9)
C36 синтетический	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	<i>in vitro</i>	5 мкМ C36 активирует пролиферацию, инвазивность и продукцию IL-6 в опухолевых клетках молочной железы MDA-MB468, в отличие от отсутствия или противоположных эффектов нативного ААТ.	Zelvyte I. et al., 2003 (DOI: 10.1097/00008469-200304000-00005)
C36 синтетический и предположитель- но эндогенный	C36 SIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK; C20 IEQNTKSPLFMGKVVNPTQK	ткань легкого предположительн о пептид- положительна (3 из 20 образцов, 2 из которых с ХОБЛ, ИГХ моноклональным и антителами 5G11, иммуноблоттинг не проводился); исследования <i>in vitro</i>	C36 активирует сигнальный путь LPS, но с меньшей амплитудой, чем LPS, и снижает эффект LPS при костимуляции: увеличивает секрецию моноцитами IL-8, TNF α , IL-6 и IL-1 β , а также активацию NFkB; увеличение секреции TNF α опосредовано рецепторами CD14 и TLR4 (так как антитела к рецепторам блокируют действие пептида и LPS (но прямое взаимодействие C36 с ними не показано)), ферментами сигнального каскада p38/MAPK, так как действие блокируется специфическими ингибиторами; C20 не похожих эффектов не проявляет.	Subramaniam D. et al., 2006 (DOI: 10.1016/j.biocel.2005.09.021)
SPAAT (Short Peptide from AAT)	C44 MFLEAIPMSIPPEVKFNKPFV FLMIEQNTKSPLFMGKVVNP TQK	выделен из ткани плаценты	Плацентарный пептид SPAAT ингибирует химотрипсин, нейтрофильную и панкреатическую эластазу, но не ингибирует трипсин. В отличие от нативной ААТ, SPAAT обратимо и конкурентно ингибирует химотрипсин. Связывается с белками внеклеточного матрикса, в основном с коллагеном III типа и ламинином-1, защищая их от протеолиза.	Niemann MA et al., 1992, 1997 (DOI: 10.1016/s0934-8832(11)80066-1; DOI: 10.1016/s0167-4838(97)00034-4; PMID: 9257191)

<p>nNIF (неонатальный фактор, ингибирующий NET) 4-6 кДа пептиды; NET - внеклеточные ловушки нейтрофилов</p>	<p>C29 KFNKPFVFLMIEQNTKSPLF MGKVVNPTQ; +CRISPP и SPAAT пептиды изучались в работе</p>	<p>Пептиды идентифицировал и в пуповинной крови методами tandemной масс-спектрометрии, ИП и вестерн-блоттинга полиАТ против С-конца ААТ (обнаружены пептиды 4-6 кДа). Синтетический С29 показал все эффекты.</p>	<p>Природный ингибитор образования внеклеточных ловушек LPS-индуцированными нейтрофилами; пептид экспрессируется и действует до нескольких дней после рождения ребенка (как доношенного, так и недоношенного). Пептид CRISPP проявляет такой же эффект, а CRISPP-SCR (состав такой же, но последовательность случайная) - нет. SPAAT ингибирует, но в меньшей степени. Все пептиды не действуют на ранее существовавшие ловушки (не проявляют ДНКазную активность). Ингибирует PAD4 (цитруллинирование гистонов). nNIF не ингибирует эластазу нейтрофилов.</p>	<p>Yost CC et al., 2016 (DOI: 10.1172/JCI83873)</p>
<p>СААР48 (C-terminal AAT Peptide 4,789 kDa)</p>	<p>СААР48 C42 LEAIPMSIPPEVKFNKPFVFL MIEQNTKSPLFMGKVVNPTQ K; СААР47 (SNP: E > D замена) LEAIPMSIPPEVKFNKPFVFL MIDQNTKSPLFMGKVVNPTQ K; Hairpin peptide PFVFLMIEQNTKSPLFMGKV VNPTQK; u-hairpin LEAIPMSIPPEVKFNK; VIRIP LEAIPMSIPPEVKFNKPFVF; scrambled KKFISPFVKMPNFVTSGPVIP QNEKFLMEQPKLMVPITLEN A</p>	<p>биомаркер сепсиса; изучен <i>in vitro</i></p>	<p>СААР48 активирует полиморфноядерные нейтрофилы, вызывая окислительный взрыв; СААР47 слабо активирует, VIRIP и пептид с случайной последовательностью (scramble) не активируют. Домен бета-шпильки отвечает за активность СААР48. СААР48 в концентрации 5-100 мкМ вызывает апоптоз нейтрофилов, а СААР47 - в концентрации 40-100 мкМ. СААР47/48 увеличивают хемотаксис нейтрофилов на 22-30%, секрецию воспалительного хемокина MIP-1α (CCL3) при 40-100 мкМ, VIRIP и случайный пептид не влияют на этот параметр.</p>	<p>Blaurock N. et al., 2016 (DOI: 10.1155/2016/6129437)</p>

CAAP48 (C-terminal Peptide 4,789 kDa)	C42 LEAIPMSIPPEVKFNKPFVFL MIEQNTKSPLFMGKVVNPTQ K	биомаркер сепсиса; <i>in vitro</i>	CAAP48 вызывает дисфункцию клеток печени в модели <i>in vitro</i> («печень на чипе»): накопление внутриклеточных ферментов аспаргат- и аланин-трансаминазы, нарушает активность цитохрома P450 3A4, снижает экспрессию белка 2, ассоциированного с множественной лекарственной устойчивостью (MRP2); нарушает целостность эндотелиального слоя, снижая экспрессию и вызывая перераспределение VE-кадгерина, белка плотных контактов зонального окклюдина-1, а также уменьшая полимеризацию F-актина. Индуцирует секрецию про- и противовоспалительных цитокинов.	Blaurock-Möller N. et al., 2019 (DOI: 10.3389/fimmu.2019.00273)
VIRIP (virus-inhibitory peptide)	VIRIP LEAIPMSIPPEVKFNKPFVF	пептид выделен из гемофильтрата, полученного от больных хронической болезнью почек	Эффективно ингибирует проникновение вируса ВИЧ-1 в клетку-хозяин, взаимодействуя с gp41 и предотвращая его интеграцию в клеточную мембрану. Мотивы, необходимые для ингибирования, выделены курсивом. Пептид также ингибирует устойчивые к антиретровирусным препаратам вирусные штаммы.	Münch J. et al., 2007 (DOI: 10.1016/j.cell.2007.02.042)
C10 эндогенный	MGKVVNPTQK	пептид идентифицирован методом масс-спектрометрии в пептидоме мочи беременных женщин, вынашивающих плод с синдромом Дауна	Высокий уровень пептида ААТ в моче у женщины с плодами с синдромом Дауна указывают на то, что женщины с плодами с синдромом Дауна имеют иммунный ответ на аномальные плоды. Высокий уровень пептида ААТ в моче может быть связан с воспалением, связанным с аномальными плодами.	Shan D. et al., 2019 (DOI: 10.1089/omi.2019.0098)

SA1-III синтетический C10	Ac-MGKVVNPTQK	in vitro	Пептид в микромолярных концентрациях вызывает увеличение уровня растворимого коллагена I типа в культуральной среде фибробластов. Значительного изменения жизнеспособности или пролиферации клеток фибробластов под действием пептида не наблюдалось. SA1-III снижает деградацию коллагена I типа, предположительно либо за счет индукции синтеза ингибитора эластазы ААТ [Janciauskiene S., Lindgren S., 1999], либо за счет конкуренции за протеолитический фермент, поскольку он содержит предполагаемые сайты расщепления эластазой (отмечен голубым цветом) и коллагеназой (зеленым цветом) [Niemann MA et al., 1997]. Нет исследований, подтверждающих, что C10 может вести себя как сигнальная молекула в пути синтеза коллагена.	Pascarella S. et al., 2016 (DOI: 10.1002/cmdc.201500472)
C21 эндогенный	MIEQNTKSPLFMGKVVNPTQK	пептид среди других преимущественно С-концевых пептидов ААТ идентифицирован методом масс-спектрометрии в пептидоме мочи беременных с преэклампсией (ПЭ)	Преэклампсия (ПЭ) проявляется повышенным артериальным давлением (систолическое ≥ 140 мм рт.ст. или диастолическое ≥ 90 мм рт.ст.) и протеинурией $\geq 0,3$ г/л в сочетании с нарушением функции почек и печени. Оценка <i>SERPINA1</i> в моче может использоваться в качестве диагностического теста на тяжесть преэклампсии с очень высокой специфичностью. Обнаружение пептидов <i>SERPINA1</i> в моче важно для дифференциации хронической артериальной гипертензии (ХАГ) и сочетанной ПЭ (ПЭ с ХАГ) с настоящей ПЭ. Плацента может быть источником пептидов <i>SERPINA1</i> в моче. Предполагается, что пептиды образовались в результате протеолиза ААТ металлопротеазами.	Kononikhin A.S. et al., 2016; Starodubtseva N. et al., 2020 (DOI: 10.1016/j.jprot.2016.04.024; DOI: 10.3390/ijms21030914)
C21 эндогенный	MIEQNTKSPLFmGKVVNPTQK (m-oxidized methionine); MIEQNTKSPLFMGKVVNPTQK	пептидом мочи терминальной стадии почечной недостаточности	Уровень C21 в моче повышен в 11-14 раз у пациентов с терминальной стадией почечной недостаточности по сравнению со стабильным заболеванием. Изменения концентрации пептида были обнаружены при длительном исследовании больных с аутосомно-доминантным поликистозом почек, у которых развилась терминальная стадия заболевания со снижением скорости клубочковой фильтрации и утратой почечной функции. Этот пептид был одним из 20 наиболее важных пептидов, участвующих в изменении экскреции при этом заболевании. Авторы	Pejchinovski M. et al., 2017 (DOI: 10.1093/ndt/gfw243)

			предполагают, что пептид образовался путем протеолиза.	
C21 эндогенный	MIEQNTKSPLFMGKVVNPTQ K	ткань почечно- клеточного рака	Пептид был идентифицирован как ассоциированный с опухолью эпитоп (лиганд HLA-DR), выделенный из ткани почечно-клеточной карциномы пациента RCC190.	Maria-Dorothea Nastke. Dissertation "T cell epitopes from viral and tumor associated antigens: Induction and Analysis of antigen- specific T cells" (2005)

Таблица 2.2. С-концевые пептиды ААТ и их свойства.

Главной отличительной особенностью С-концевых пептидов от полноразмерного белка является их вклад в воспалительный процесс: пептиды являются хемоаттрактантами и активаторами нейтрофилов (этимися свойствами обладает также и комплекс ААТ-протеаза [316]) [317], активируют секрецию провоспалительных цитокинов и хемокинов моноцитами через рецепторы LDLR и CD36 (связывающий окисленный LDL) [318, 319], запускают LPS-сигнальный каскад в моноцитах [320], - таким образом, можно утверждать, что нативный ААТ и С-концевые фрагменты обладают антагонистическим действием.

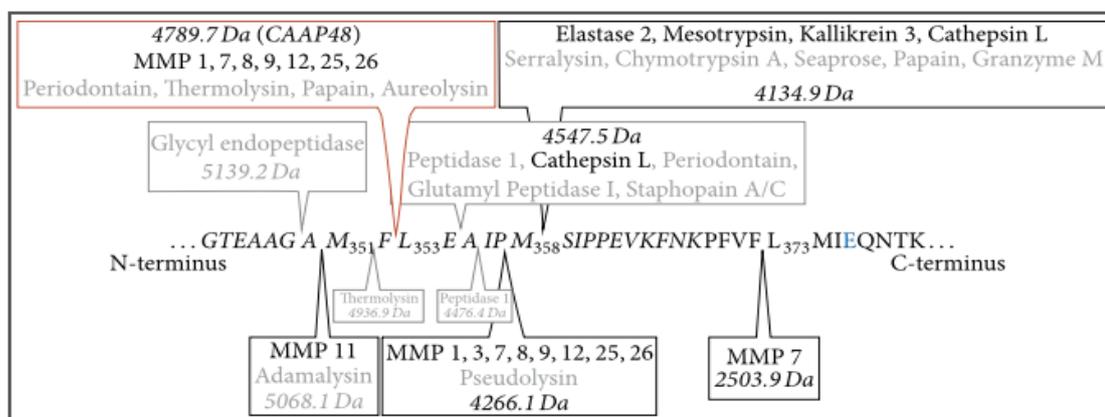


Рисунок 2.8. Сайты протеолитического расщепления молекулы ААТ в С-концевом домене, приводящего к образованию биоактивных пептидов. Курсивом указаны сайты расщепления, установленные экспериментально. Протеазы человека отмечены черным, других видов - серым. Для каждого сайта расщепления указана молекулярная масса образуемого С-концевого пептида. Красная рамка - расщепление, приводящее к образованию маркера сепсиса СААР48. Рисунок полностью взят из [55]. MMP7-матрилизин; MMP3-стромелизин1; MMP11-стромелизин3; MMP8-коллагеназа нейтрофилов; MMP1-интерстициальная коллагеназа.

В норме С-концевые пептиды ААТ обнаружены в фосфолипидной фракции селезенки (С36) и желчи (С36, С42) [321]; в ткани плаценты пептид С44 (SPAАТ) защищает внеклеточный матрикс от протеолиза [50], а в пуповинной крови пептиды массой 4-6 кДа, названные nNIF (neonatal NET-inhibitory factor), эффективно ингибируют образование внеклеточных ДНК-ловушек нейтрофилами, тем самым препятствуя их провоспалительной гиперреакции на границе мать-ребенок и защищая плод от патологических последствий [56]. Пептид VIRIP (С20), выделенный из гемофильтрата, полученного от пациентов с хронической болезнью почек, препятствует проникновению вируса ВИЧ в клетку хозяина, избирательно взаимодействуя с вирусным белком gp41 гидрофобными мотивами (таблица 2.2) и инактивируя тем самым процесс слияния вируса с мембраной клетки [51]. Позже обнаружили, что ВИЧ-ингибирующими свойствами обладает и полноразмерный ААТ (механизм тот же) [6]. Пептид СААР48 (С42) предложен в качестве маркера сепсиса, обладает провоспалительными свойствами и может являться одним из инициаторов печеночной недостаточности, ассоциированной с сепсисом [55, 322]. С-концевой фрагмент ААТ С36 обнаружен в атеросклеротических бляшках [52], а вместе с С42 - у пациентов с тяжелым течением COVID-19 и бактериальным легочным сепсисом [323]. Недавно был разработан мультиплексный подход к обнаружению С-концевых пептидов в биологических

жидкостях для мониторинга течения патологических состояний, связанных с повышением уровня пептидов-производных ААТ [324].

Интересно отметить, что все обнаруженные пептиды (кроме VIRIP и SA1-III/C10), функция которых установлена, содержат пентамотив FVFLM, и можно ожидать, что их физиологическое действие опосредуется активацией рецепторов LRP1, LDL, CD36, предположительно CD14/TLR4 и, не исключено, FPR1 [156, 157, 318, 319, 320, 325]. С учетом того, что пептид С36 является хемотаксическим агентом нейтрофилов и содержит мотивы, напоминающие хемотаксический трипептид, - SIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK - возникает вопрос, являются ли мотивы в контексте структуры С36 мимикрирующими fMLF? С этого ракурса хемоаттрактантные свойства С36 в работах не рассматривались. Возможность участия специфического рецептора FPR1 подтверждается тем, что повышение концентрации внутриклеточных катионов кальция под действием синтетического пентамотива FVFLM блокируется при действии антагониста известного хемотаксического агента бактериального происхождения fMLF (N-формил-метионил-лейцил-фенилаланин), - ВосMLF (трет-бутилоксикарбонил-метионил-лейцил-фенилаланин) [325]. В работе Джослин с соавторами обнаружили, что комплекс ААТ-протеаза, в котором пентамотив экспонирован, также активирует хемотаксис нейтрофилов [316]. Однако авторы утверждают, что данный эффект опосредован SEC-рецептором, а клетки, предынкубированные с синтетическими аналогами пентамотива, сохранили “чувствительность” к лиганду fMLF. Последнее можно объяснить также вытеснением пентамотива из комплекса с рецептором (FPR1) ввиду более высокой константы связывания последнего с fMLF, - было бы интересно провести этот же эксперимент, но с добавлением хемоаттрактантов в обратном порядке.

CRISPP пептиды (cancer-associated, SCM-recognition, immune defense-suppressing, and serine protease-protecting) обнаружены в периферической крови пациентов с онкологическими заболеваниями различного происхождения. Пептиды обладают проонкогенными свойствами: ингибируют цитотоксическое действие NK/LAK (lymphokine-activated killer cell) клеток *in vitro*, защищают сериновые протеазы от инактивации ААТ *in vitro*, повышают пролиферацию (оцененную по уровню встраивания ³H-тимидина) EGF-стимулированных первичных гепатоцитов крысы и РНА-стимулированных лимфоцитов периферической крови человека на 130 и 250%, соответственно [45, 46, 47, 326]. С-концевой пептид С36 обнаружен в желудочном соке при опухоли желудка [53], а С41/С42 - в протоковой жидкости при опухоли молочной железы [54]. С36 *in vitro* проявляет противоположные нативному ААТ, проопухолевые, эффекты на клеточных линиях карциномы молочной железы и меланомы, повышая секрецию провоспалительных цитокинов, проферментов металлопротеаз, активируя пролиферацию и инвазивность [313, 327]. При оверэкспрессии секреторного варианта С-концевого

фрагмента ААТ С37 линией аденокарциномы поджелудочной железы наблюдали увеличение размеров опухолей при подкожном введении суспензий опухолевых клеток голым мышам в 1,5-3 раза и уровня метастазирования в лимфатические узлы по сравнению с контрольной группой [328]. Авторы связывают данный эффект с подавлением активации натуральных киллеров в присутствии С-концевого фрагмента белка *in vivo*, так как преимущества в росте отсутствовали в NK-делетированных мышах, что подтверждает *in vitro* наблюдения Церцекос [328]. Таким образом, С-концевые пептиды ААТ, содержащие ключевые эпитопы (таблица 2.2, CRISPP), обладают проопухолевыми свойствами.

3. МАТЕРИАЛЫ И МЕТОДЫ

3.1 Культивирование клеточных линий и первичных культур

В работе использовались перевиваемые клеточные линии человека и первичная культура HUVES. Клеточные линии были любезно предоставлены коллегами из ИМБ РАН имени В.А. Энгельгардта и ФГБУ «НМИЦ онкологии им. Н.Н. Блохина»; эндотелиальные клетки вены пуповины человека (HUVES) были любезно предоставлены коллегами из ФГБУ ФНКЦ ФХМ им. Ю.М. Лопухина ФМБА России. Прикрепленные клеточные линии: опухоли эпителия предстательной железы DU145 (гормон-нечувствительная линия, выведена из метастазов в мозг), PC-3 (гормон-нечувствительная, из метастазов в кость), LNCaP (гормон-чувствительная, из метастазов в лимфоузел); опухоли эпителия молочной железы MDA-MB-231 (трижды негативная), T-47D (HER2-отрицательная) и MCF7 (все три линии выведены из плеврального выпота), ZR-75-1, BT-20 (трижды негативная); спонтанно иммортализованная линия эпителия молочной железы HBL-100 (выведена из грудного молока); эпителиоподобная линия опухоли печени HepG2; опухоли эпителия поджелудочной железы VxPC-3; фибробластоподобная клеточная линия меланомы MeWo. Суспензионные клеточные линии: линии Т-лимфобластного лейкоза Jurkat и эритромиелоидного лейкоза K562. Прикрепленные клетки культивировали в пластиковых чашках с различной площадью роста (Greiner Bio-One, Австрия), суспензионные клетки растили в матрасах объемом 50 мл (Greiner Bio-One, Австрия) в среде DMEM/F12 (DU145, PC-3, MCF-7, BT-20 и HepG2) либо RPMI 1640 (LNCaP, MDA-MB-231, T-47D, HBL-100, ZR-75-1, VxPC-3, K562, Jurkat и MeWo), с добавлением 10% (по объему) эмбриональной телячьей сыворотки (Gibco, США), 1% (по объему) раствора антибиотик/антимикотик (конечные концентрации: пенициллин 100 ед/мл, стрептомицин 100 мкг/мл, амфотерицин В 0,25 мкг/мл) (Gibco, США). Первичную культуру HUVES растили в специализированной среде (EGM™-2 Endothelial Cell Growth Media, Lonza, США). Все клетки культивировали в CO₂-инкубаторе (Binder, Германия) при 37°C во влажной атмосфере с 5% содержанием CO₂. Прикрепленные культуры обрабатывали раствором 0,05% трипсина-EDTA (Gibco, США) и перевивали в соотношении 1:4-1:6 каждые 3-4 дня.

3.2 ПЦР в реальном времени (ПЦР-РВ)

Прикрепленные клетки сажали на 3,5 см культуральные чашки и по достижении ≈70-80% конфлюентности лизировали реагентом Trizol (Life Technologies Invitrogen, США); суспензионные клетки центрифугировали 5 минут при 400 g перед лизисом. Чашка с прикрепленными клетками RWPE-1 в конфлюентности ≈70%, выращенными в соответствии с рекомендациями Американской коллекции клеточных культур (ATCC-LGC), была любезно предоставлена лабораторией проф. Воробьева И. А. (кафедры клеточной биологии и гистологии биологического факультета МГУ). Суммарную РНК выделяли в соответствии с протоколом производителя Trizol. Использовали также

коммерческие препараты суммарной РНК тканей человека: каталожные номера, лоты и производители указаны в пункте 3.4.4. Концентрация препаратов РНК оценивалась спектрофотометрически (“Nanodrop”, Thermo Fisher Scientific, США) при длине волны 260 нм, а также по соотношению A(260)/A(280) нм и A(260)/A(230) нм, а целостность - электрофоретически в 1% агарозном геле (Рисунок 1П (Б)). Для синтеза кДНК использовали систему обратной транскрипции QuantiTect (Quagen, Германия), предполагающую предварительное удаление геномной ДНК перед синтезом кДНК, реакцию с 1 мкг суммарной РНК проводили в соответствии с протоколом производителя. Реакцию без добавления обратной транскриптазы использовали для контроля отсутствия геномной ДНК в образцах РНК.

Для определения уровня экспрессии стыка EX2/3 использовали следующие праймеры: прямой 5'-TGTGGATTTGGTCAAGG-3', обратный - 5'-ССТТCACGGTGGTCAAC-3'; пептид-кодирующей области EX5: прямой 5'-GCCATACCCATGTCTATC-3', обратный 5'-GAGCGAGAGGCAGTTA-3'. Для изучения уровней экспрессии транскриптов, полиаденилированных по разным сайтам (APA1-3, Alternative PolyAdenylation site) использовали следующие праймеры: (APA1) прямой праймер 5'-GCCATACCCATGTCTATC-3', обратный - 5'- TTGGGTGGGATTCACCAC-3'; (APA2) прямой праймер для APA1, обратный - 5'-СТСААССТТСТТТААТГТС-3'; (APA3) прямой праймер 5'- GACCTGAGAGTCTGAAGA-3', обратный - 5'-СТТГААСАТСАТАССААСТС-3'. В качестве референсного гена использовали *GAPDH*: прямой праймер 5'-CGTCAAGGCTGAGAAC-3'; обратный - 5'-АСТССАСГАСГТАСТС-3'. ПЦР-РВ проводили в термоциклере CFX96 Real-Time PCR Detection System (Bio-Rad, США), с набором реагентов Maxima SYBR Green qPCR Master Mix (2X) (Thermo Fisher Scientific, США) в 20 мкл. Специфичность амплификации оценивали в режиме кривых плавления ампликонов, продукты амплификации анализировали в 2% агарозном геле (Рисунки 1П (В) и 2П (Б) приложения, ожидаемые длины продуктов амплификации подписаны на рисунках). Условия ПЦР-РВ: преденатурация 94 °С 10 мин; 45 циклов: 94 °С -10 с, 57,6 °С - 20 с, 72 °С - 30 с; 94 °С 10 с; режим кривых плавления: постепенное нагревание от 55 до 94 °С с шагом в 0,5 °С/10 с. Все эксперименты выполнены в трех параллельных повторах. Синтез дезоксирибоолигонуклеотидов (олигонуклеотидов) выполнен в компании “Синтол” (Москва, Россия).

Эффективность амплификации рассчитывали по тангенсу угла наклона стандартной кривой ($\text{tg}\alpha$), полученной серией из пяти последовательных четырехкратных разведений смеси кДНК клеток HepG2, по формуле: $E = 10^{-1/\text{tg}\alpha}$ или $E(\%) = (E - 1) * 100 \%$. Для этого с 1 мкл смеси кДНК проводили ПЦР-РВ с каждой парой праймеров в трех повторностях, по их результатам вычисляли среднее значение порогового цикла C_t для каждой пары праймеров и для каждого разведения. Строили график зависимости среднего C_t от

десятичного логарифма концентрации кДНК. Исходная концентрация кДНК условно принималась равной $1 \cdot 10^6$.

Для расчета относительной экспрессии гена использовали формулу Пфаффа с поправкой на эффективность амплификации [329]: $E(GAPDH)^{<Ct>(GAPDH)} / E(target)^{<Ct>(target)}$, где E – эффективность амплификации с соответствующей пары праймеров, а <Ct> - среднее значение порогового цикла по трем повторам. Количественный подсчет экспрессии длинных и коротких транскриптов проводили в отдельности. Для этого вычисляли количество матрицы таргета на одну матрицу GAPDH, показатель для EX2/3 - экспрессия длинных транскриптов, а разница между показателями EX5 и EX2/3 - экспрессия коротких транскриптов. Уровни экспрессии длинных и коротких транскриптов *SERPINA1* в разных образцах нормировали на образец с минимальной экспрессией - РС3 EX2/3. Аналогичным образом рассчитали в отдельности экспрессию транскриптов, полиаденилированных по разным сайтам, данные нормировали на образец с минимальной экспрессией - DU145 АРА3. Обработка данных амплификации в ПЦР-РВ участков гена *SERPINA1* выполнена в программе MS Excel (Microsoft, США). Статистический анализ выполнен с использованием GraphPad Prism 8.0.1 (GraphPad Software Inc., США). Данные, представленные как среднее \pm стандартное отклонение, были проанализированы с помощью однофакторного дисперсионного анализа (ANOVA), за которым следовали множественные сравнительные тесты Тьюки. Различия со значениями $p \leq 0,05$ считались статистически значимыми.

3.3 Определение содержания ААТ в секретах клеточных линий методом иммуноблоттинг

Клеточные линии выращивали на 14,5 см культуральных чашках в полной ростовой среде до достижения $\approx 60\%$ конfluence. Клетки были тщательно и аккуратно промыты 3 раза по 3 минуты в предварительно подогретом DPBS (с содержанием катионов Ca^{2+} и Mg^{2+}) (Gibco, США), и полная среда была заменена на бессывороточную среду с добавлением раствора пенициллина (до 100 ед/мл) и стрептомицина (100 мкг/мл) (Gibco, США). Секреты были насыщены в течение 48 часов, собраны и центрифугированы при 500 g в течение 10 минут; затем надосадочные жидкости собраны и центрифугированы при 16000 g в течение 20 минут. Секреты концентрировали в центрифужном концентраторе Vivaspin 15R Centrifugal Filter Unit с пороговым значением 2 кДа (VS15RH91, Sartorius, Германия). Общую концентрацию белка измеряли с помощью набора для колориметрического определения белка на основе бигинхониновой кислоты (Pierce BCA Protein Assay Kit, #23227, Thermo Scientific, США), и концентраты хранились при температуре $-80^\circ C$ до использования. 25 мкг секреторных белков, разведенных в 1x буфере для нанесения, нагревали при $95^\circ C$ в течение 5 минут и разделили с помощью денатурирующего электрофореза в полиакриламидном геле с 5% концентрирующим и 12,5% разделяющим гелями с использованием системы электрофореза Mini-PROTEAN

(Bio-Rad, США). Белки перенесли на нитроцеллюлозную мембрану с размером пор 0,2 мкм (#1620112, Bio-Rad, Hercules, CA, США) в системе для полусухого переноса Trans-Blot SD semi-dry transfer cell (Bio-Rad, США). Мембраны блокировали в 4% BSA в PBS. Сигналы от ААТ детектировали с помощью первичных антител против полноразмерного белка в разведении 1:5000 (LS-B1454, LifeSpan Biosciences, США) в паре со вторичными антителами в разведении 1:7000 (ab6741, Abcam, Великобритания) в 2% BSA в PBST (PBS с добавлением 0,05% Tween-20). Сигнал проявляли с использованием субстрата Clarity Western ECL (#1705060, Bio-Rad, США). Мембраны сканировали с использованием системы ChemiDoc XRS+ (BioRad, США). Коммерческий ААТ 178251 (ААТр, Calbiochem, Германия) использовался в качестве положительного контроля. Precision Plus Protein WesternC стандарты (#1610376, Bio-Rad, США) использовали в качестве маркеров молекулярных масс. Представление данных и статистический анализ проводились, как описано выше в пункте 3.2. Связь между уровнями мРНК *SERPINA1* и белка оценивалась с использованием коэффициента корреляции Пирсона в GraphPad Prism 8.0.1 (GraphPad Software Inc., США).

3.4 Нозерн гибридизация

3.4.1 Конструирование ДНК-матриц для синтеза кРНК

Фрагмент гена *SERPINA1* длиной 1104 п.н. включающий экзон 5 и смежный интрон, амплифицировали из геномной ДНК HepG2 с использованием прямого праймера 5'-АТСАСТААGGТСТТСАГСА-3' и обратного праймера 5'-TGAGGAGCGAGAGGCAGT-3' (праймеры синтезировали в компании "Синтол" (Москва, Россия)) в реакции ПЦР с использованием Taq ДНК-полимеразы (10342020, Invitrogen, США). Амплифицированный фрагмент очистили экстракцией фенола/хлороформа и осаждением этанолом. Очищенный фрагмент клонировали в плазмидный вектор pTZ57R/T из системы TA-клонирования (InsTAclone™ PCR Cloning Kit, #K1214, Thermo Scientific, США) и трансформировали в штамм *E. coli* JM109 в строгом соответствии с протоколом производителя. Бактериальные клоны скринировали с помощью ПЦР для определения ориентации вставки относительно T7-промотора (*плазмидная конструкция с прямой ориентацией вставки служила для синтеза положительного контроля смысловой кРНК; в обратной ориентации - для антисмыслового кРНК-зонда*) с использованием комбинации вектор- и ген-специфических праймеров. Рекомбинантные плазмиды с прямой и обратной ориентацией фрагмента гена *SERPINA1* выделили из бактериальных клонов с использованием набора Plasmid Mini Kit (12143, Qiagen) в строгом соответствии с протоколом производителя. Для получения более короткой вставки, содержащей экзон 5, очищенную плазмиду с прямой ориентацией вставки обработали эндонуклеазами XhoI (ER0691, Thermo Scientific, США) и SalI (ER0645, Thermo Scientific, США), а плазмиду с обратной ориентацией вставки - эндонуклеазами XhoI и XbaI (ER0681, Thermo Scientific), за которыми последовала обработка фрагментом Кленова. Векторы с оставшимися последовательностями вставки (в прямой и обратной

ориентации - по 260 п.н.) были лигированы сами на себя, трансформированы в штамм *E. coli* JM109, и плазмиды были выделены из индивидуальных клонов, как описано выше.

3.4.2 Синтез смысловой кРНК

1 мкг очищенных рекомбинантных плазмид, линейаризованных эндонуклеазой EcoRI (ER0271, Thermo Scientific, США), с прямой ориентацией фрагмента гена *SERPINA1* использовались в качестве матрицы для синтеза смысловой кРНК с последовательностью *SERPINA1* протяженностью 1104 и 260 н.о. в реакции транскрипции с использованием T7 РНК-полимеразы (EP0111, Thermo Scientific) в строгом соответствии с протоколом производителя. Реакционную смесь обрабатывали фенол-хлороформной экстракцией и осаждением этанолом. Полученные кРНК использовались в качестве положительного контроля в нозерн гибридизации.

3.4.3 Синтез радиоактивно меченого антисмыслового кРНК зонда

1 мкг очищенных рекомбинантных плазмид, линейаризованных эндонуклеазой EcoRI (ER0271, Thermo Scientific, США), с обратной ориентацией фрагмента гена *SERPINA1* использовались в качестве матрицы для синтеза антисмыслового кРНК зонда длиной 260 н.о. с координатами в геномном браузере UCSC Browser: (chr14:94,378,432-94,378,692) с использованием набора для транскрипции T7/SP6 transcription Kit (P1460, Promega) в строгом соответствии с протоколом производителя. [α -³²P]GTP с молярной активностью 4000 Ки/ммоль и объемной активностью 50 мкКи/мкл был получен от ЦКП "ФОСФОР" РАН (Москва, Россия). Включение меченого [α -³²P]GTP определяли по протоколу, предоставленному производителем Promega, и общая активность пробы составила $21,9 \cdot 10^6$ cpm.

3.4.4 Препараты суммарной РНК клеточных линий и тканей

Культивирование клеток и получение препаратов суммарной РНК обозначены в пунктах 3.1 и 3.2. Суммарную РНК обрабатывали ДНКазой I в присутствии 2,5 mM Mg²⁺ (EN0523, Thermo Scientific, США), выделяли фенол-хлороформной экстракцией в смеси с изоамиловым спиртом (в соотношении 50:49:1 по объему), преципитировали этанолом в присутствии 0,5 M ацетата аммония, растворяли в воде без нуклеаз. Концентрацию и целостность препаратов суммарной РНК анализировали как описано в пункте 3.2.

В работе также использовали коммерческие препараты суммарной РНК из тканей человека (Каталожный №, лот № (один или несколько)): фетальной почки (636584, 1005012), фетальной печени (636540, 1101452A), нормальной печени (636531, 1109024A), нормального тонкого кишечника (636539, 1202353A), нормальной почки (636529, 9110240A), опухоли желудка (636629, 1205234A), опухоли молочной железы (636635, 1109330A, 1109288A), опухоли почки (636632, 1102289A, 1202890A, 1202938A), нормальной скелетной мышцы (636534, 1108079A), нормальной простаты (636550, 1106275A), нормальной молочной железы (636576, 1109023A) были получены от Clontech Laboratories, Inc (США); нормальной печени (AM7960, 1203009), нормальной простаты

(AM7988, 1202007), нормальной молочной железы (AM6952, 1203016) были получены от Ambion (США); нормальной простаты (R1234201- P, B509110) был получен от BioChain (США).

3.4.5 Электрофоретическое разделение РНК

10 мкг суммарных РНК клеточных линий или тканей, 4 мкл РНК-маркера (HR RiboRuler, SM1821, Thermo Scientific, США) и контрольные смысловые кРНК были разделены с помощью электрофореза в 1,5% агарозном геле с использованием набора NorthernMax™-Gly Kit (AM1946, Thermo Scientific, США), в котором в качестве денатурирующего агента служит смесь глиоксаля с ДМСО, в строгом соответствии с протоколом производителя. Целостность РНК верифицировали в геле под ультрафиолетовым светом.

3.4.6 Перенос РНК на мембрану

РНК перенесли “сверху-вниз” на мембрану из нейлона BrightStar™-Plus Positively Charged Nylon Membrane (AM10102, Thermo Scientific, США) по протоколу производителя набора NorthernMax™-Gly Kit (AM1946, Thermo Scientific, США). Перенос оценивали проверкой геля под ультрафиолетовым светом. РНК сшивали ультрафиолетом.

3.4.7 Гибридизация

Мембрану прегибридизовали с использованием буфера ULTRAhyb™ Ultrasensitive Hybridization Buffer при 68°C и гибридизовали в течение ночи с зондом, разбавленным в этом же буфере до объемной активности $2,19 \cdot 10^6$ срм/мл. После промывки мембрану подвергали автордиографии с усиливающим экраном в течение 5 дней, а экран сканировали с помощью Typhoon FLA 9500 (GE Healthcare, США).

3.5 Детекция экспрессии гена *SERPINA1* с противоположной цепи ДНК в локусе экзона 5

Проводили в два раунда на кДНК DU145 и HepG2 (пункт 3.2). Первый раунд предполагал однонаправленную ПЦР для дифференциации цепей ДНК с прямым 5'-AGAGGCCATACCCATGTCTATC-3' (амплификация “смыслового” продукта) либо обратным 5'-TTGGGTGGGATTCACCACTTTTCC-3' (амплификация “антисмыслового” продукта) ген-специфическими праймерами. В качестве полимеразы использовали высокоточную ДНК-полимеразу Tersus (PK121, Евроген, Россия). Условия однонаправленной ПЦР: предденатурация 94 °С 3 мин; 15 циклов: 94 °С -30 с, 62 °С - 30 с, 72 °С - 2 мин; 25 циклов: 94 °С -30 с, 60 °С - 30 с, 72 °С - 2 мин; финальная достройка 60 °С - 30 с, 72 °С - 5 мин. Второй раунд предполагал ПЦР-РВ с парой праймеров на участок пятого экзона, включающий контроль - референс с соответствующей исходной кДНК в соответствующем разбавлении. Во втором раунде использовали пару праймеров: прямой EX5F 5'-GCCATACCCATGTCTATC-3', обратный 5'-TTTGGGTGGGATTCACCA-3', - в условиях ПЦР-РВ идентичных таковым, указанным в пункте 3.2 с использованием готовой

смеси для ПЦР-PB 5X qPCRmix-HS SYBR (Евроген, Россия). Все эксперименты выполнены в трех параллельных повторах. Синтез дезоксирибоолигонуклеотидов (олигонуклеотидов) выполнен в компании “Синтол” (Москва, Россия).

3.6 3'-RACE (Rapid Amplification of cDNA Ends)

1 мкг очищенной суммарной РНК из клеточных линий, обработанной ДНКазой I (Thermo Fisher Scientific, США), использовали для реакции обратной транскрипции с ревертазой SuperScript III Reverse Transcriptase (Invitrogen, США) используя олигонуклеотид с адапторной областью 5'-GTGCCCTTCGTGCGC(T)₁₈(N)₃-3' (5065) *либо* 5'-GCGAGCACAGAATТААТАСГАСТСАСТАТАГГ(T)₁₂VN-3' (3dTA, протокол № 21 [330]), объем реакционной смеси – 20 мкл. После обратной транскрипции фермент инактивировали при 70 °С в течение 15 мин, реакционную смесь обработали РНКазой H (EN0201, Thermo Scientific, США). Полученную кДНК очистили от не прореагировавшего олигонуклеотида с помощью набора GenElute PCR Clean up kit (NA1020, Sigma, США). Для ПЦР использовали систему Expand High Fidelity PCR System (Roche, Швейцария). Пары праймеров для адаптора 5065: прямой EX4F 5'-АТСАСТААГГТСТТСАГСА-3' *либо* EX5F 5'-GCCАТАССАТГТСТАТС-3' в паре с обратным на адапторную область 5032 5'-GTGCCCTTCGTGCGCT-3' (рисунок 4.6, А). Пары праймеров для адаптора 3dTA из протокола № 21 [330]: прямой F1 5'-GGTCAAGTTCAACAACCCTTTG-3', F2 5'-ТАССААГТСТССССТСТТСАТГ-3', F3 5'-GAAAAGTGGTGAATCCCACCCA-3' *либо* F4 5'-AGAGGCCАТАССАТГТСТАТС-3' в паре с обратным 3AP1 5'-GCGAGCACAGAATТААТАСГАСТ-3' (рисунок 4.6, Б). Условия ПЦР: преденатурация 94 °С 2,5 мин; 10 циклов: 94 °С 15 с, 62 °С 30 с, 72 °С 2 мин; 25 циклов: 94 °С 15 с, 60 °С 30 с, 72 °С 2 мин, инкремент 5 с; 72 °С 7 мин. Продукты амплификации анализировали электрофоретически в 1,5% агарозном геле. Синтез дезоксирибоолигонуклеотидов (олигонуклеотидов) и секвенирование фрагментов ДНК выполнены в компании “Синтол” (Москва, Россия).

3.7 5'-RACE

5'-RACE был выполнен согласно протоколу №20 [330], используя тот же 5'-адаптер и адаптерные праймеры 5AP1 и 5AP2. Вкратце, суммарную РНК обработали щелочной фосфатазой (#M0290L, NEB, США), после чего ферментом TAP (табачной кислотной пирофосфатазой, T81050, Epicentre Biotechnologies, США). Затем 5'-адаптор (5'-GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA-3') лигировали с 5'-концами РНК с помощью T4 РНК-лигазы (M0204S, NEB, США). После окончательной экстракции фенол-хлороформом, РНК очистили с помощью набора GenElute PCR Clean up kit (NA1020, Sigma, США). Синтез кДНК осуществляли из 2,5 мкг суммарной РНК с использованием 3dTA адаптерного праймера (пункт 3.6) и SuperScript III RT (Invitrogen, США). Праймеры, использованные для вложенной ПЦР: 5AP1 5'-GCTGATGGCGATGAATGAACACTG-3', 5AP2 5'-

GAACACTGCGTTTTGCTGGCTTTGATG-3'; специфические для гена праймеры: R1 5'-TTGGGTGGGATTCACCACTTTTCC-3', R2 5'-CCATGAAGAGGGGAGACTTGGTAT-3', R3 5'-GACAAAGGGTTTGTGAACTTGACCTC-3', R4 5'-GGATAGACATGGGTATGGCCTCTA-3'.

Длинные кДНК были амплифицированы с помощью DreamTaq DNA Polymerase (EP0702, Thermo Scientific, США), короткие кДНК - с помощью полимеразы ДНК высокой точности Tersus (PK121, Евроген, Россия). Условия ПЦР для длинных продуктов был такой же, как в пункте 3.6; для коротких продуктов время элонгации в каждом цикле ПЦР составило 30 с. Продукты амплификации были вырезаны из агарозного геля, очищены и секвенированы. Синтез адапторного рибоолигонуклеотида и праймеров выполнены в компании "Синтол" (Москва, Россия); секвенирование фрагментов ДНК - в компании "Евроген" (Москва, Россия).

3.8 Извлечение и обработка пиков FANTOM5 CAGE и данных ENCODE ChIP-seq

Координаты пиков CAGE и значения экспрессии пиков в единицах TPM (Tags Per Million) для образцов фазы I FANTOM5 были извлечены с помощью инструмента извлечения таблиц (<https://fantom.gsc.riken.jp/5/tet/>). Координаты пиков были преобразованы в сборку hg38 с использованием утилиты Liftover на веб-платформе Galaxy (<https://usegalaxy.org/>) и визуализированы в геномном браузере UCSC (<http://genome.ucsc.edu/>). Для расчета попарных корреляций (матрицы корреляций) между значениями экспрессии CAGE пиков в логарифме по основанию 2 в образцах фазы I FANTOM5, мы использовали пакет R Hmisc, за которым последовала визуализация с помощью инструмента Morpheus (<https://software.broadinstitute.org/morpheus>) (рисунок 4.10, А). Кроме того, был использован пакет R PerformanceAnalytics для расчета матрицы корреляции значений экспрессии в логарифме по основанию 2, специфических для тех пиков CAGE, которые перекрываются с концами продуктов 5'-RACE, выявленных в нашем исследовании (рисунок 4.10, Б). При расчете корреляций значения экспрессии <1 TPM были исключены. Файлы ENCODE HepG2 ChIP-seq bigWig, использованные в нашем исследовании: ENCFF801BXZ (POLR2AphosphoS5); ENCFF940PKN (POLR2AphosphoS2); ENCFF323HAZ (POLR2A); ENCFF432CLC (H3K4me1); ENCFF992REP (H3K4me3); ENCFF401PIF (H3K27ac); ENCFF097LEQ (AGO2).

3.9 Анализ потенциала к трансляции альтернативных OPC

индивидуальных длинных и коротких транскриптов гена *SERPINA1*

3.9.1 Конструирование плазмид

ПЦР-продукт, не содержащий стартового кодона ATG eGFP, амплифицировали из вектора pEGFP-N1 с использованием прямого (сайты рестрикции подчеркнуты) 5'-AATAAGGATCCGTGAGCAAGGGCGAGGAG-3' и обратного 5'-

GAGTCGCGGCCGCTTTACTTGTACAGCTCGTCCA-3' праймеров и клонирован по сайтам рестрикции *Vam*HI (#FD0054, Thermo Scientific, США) и *Not*I (#FD0596, Thermo Scientific, США) вместо немутантного EGFP в вектор pEGFP-N1. Длинный фрагмент *SERPINA1* HLT для клонирования в мутантный вектор (названный pEGFP-N1(ATG-)) амплифицировали на кДНК HepG2 с использованием прямого праймера 5'-AATAAGCTAGCCTCCTCAGCTTCAGGCACCA-3' в сочетании с обратным праймером из экзона 5 5'-AATAAAGATCTTTTTTGGGTGGGATTCACCAC-3'. Длинные фрагменты DLT1 и DLT2 амплифицировали на кДНК DU145 с использованием прямого праймера 5'-AATAAGCTAGCACAATGACTCCTTTCGGTAAG-3' в паре с тем же обратным праймером, что и для HLT. Амплификация была выполнена с использованием DreamTaq DNA Polymerase (EP0702, Thermo Scientific, США). ПЦР-продукты для длинных вставок были клонированы в pEGFP-N1(ATG-) по сайтам *Bmt*I (#E457, SibEnzyme, Россия) и *Bgl*II (#E027, SibEnzyme, Россия).

Короткие фрагменты *SERPINA1* для клонирования в мутантный вектор амплифицировали на кДНК HepG2 с использованием прямых праймеров: 5'-AATAAGCTAGCAGGAAGAGGACTTCCACGTG-3' - для ST1; 5'-AATAAGCTAGCGAAAATGAAGACAGAAGGTCTG-3' - для ST2; 5'-AATAAGCTAGCACCCAAACTGTCCATTAAGT-3' - для ST3; 5'-AATAAGCTAGCAGCTGCTGGGGCCATGTT-3' - для ST4; - в паре с обратным праймером из экзона 5 5'-AATAAGGATCCTTTTTGGGTGGGATTCACCAC-3'. Амплификация была выполнена с использованием ДНК-полимеразы высокой точности Tersus (PK121, Евроген, Россия). ПЦР-продукты для коротких вставок клонировали в pEGFP-N1(ATG-) по сайтам *Bmt*I (#E457, SibEnzyme, Россия) и *Vam*HI (#ER0051, Thermo Scientific, США). Обработанные рестриктазами ПЦР-продукты и вектор очищали из геля, вектор обрабатывали щелочной фосфатазой TAP (SibEnzyme, Россия), реакцию лигирования проводили с использованием ДНК-лигазы T4 (Thermo Scientific, США). На следующий день после трансформации в компетентные клетки *E.coli* TOP10 (с добавлением в качестве селективирующего антибиотика канамицина ("Панэко") 25 мкг/мл) проводили ПЦР-скрининг колоний: для длинных вставок использовали прямой праймер к вектору CMVF 5'-CGCAAATGGGCGGTAGGCGTG-3' в паре с обратным из экзона 2 *SERPINA1* 5'-GGTGATCCTGATCATGGTG-3'; для коротких вставок - прямой CMVF в паре с обратным из экзона 5 *SERPINA1* 5'-AATAAGGATCCTTTTTGGGTGGGATTCACCAC-3'. Из колоний, несущих плазмиду со вставкой, затем наращивали ночную культуру (в 5 мл среды LB с канамицином 25 мкг/мл) и плазмиды очищали с использованием набора Plasmid Miniprep kit без эндотоксинов (#BC021, Евроген, Россия). Последовательности ДНК проверили секвенированием (Евроген, Россия), и непреднамеренные мутации не были обнаружены. Олигонуклеотиды синтезированы в компании "Евроген" (Россия).

3.9.2 Трансфекция

Клетки DU145 либо HepG2 снимали с использованием реагента TrypLE (Gibco, США) и высевали на культурные чашки диаметром 6 см количеством $1,25 \cdot 10^6$ клеток/чашку. На следующий день трансфецировали плазмиды с использованием реагента Lipofectamine 3000 (Invitrogen, США) согласно рекомендациям производителя. Через 24 часа после трансфекции обновили ростовую среду. Для подготовки секретома клетки промывали 5 раз по 2 минуты в среде без сыворотки, затем клетки оставляли в среде без сыворотки (с добавлением раствора пенициллина/стрептомицина) на дополнительные 24 часа. Для получения лизатов клеток и клетки оставляли в полной ростовой среде.

3.9.3 Подготовка секретомов

Кондиционированные среды (секретомы) центрифугировали при 500 g в течение 10 минут; надосадочные жидкости далее центрифугировали при 15000 g в течение 20 минут. Секретомы концентрировали примерно в 50 раз с помощью ультрафильтрации на центрифужном фильтре Amicon Ultra-4 Centrifugal filter unit с порогом в 3 кДа (UFC8003, Millipore, США), и концентраты хранились при температуре $-80\text{ }^{\circ}\text{C}$ до использования. Общую концентрацию белка измеряли с помощью набора для колориметрического определения белка на основе бицинхониновой кислоты (Pierce BCA Protein Assay Kit, #23227, Thermo Scientific, США).

3.9.4 Проточная цитофлуориметрия и сортинг клеток

Клетки анализировали с использованием прибора FACSAria SORP (BD Biosciences, США). Длины волн возбуждения/испускания для eGFP были следующими: Ex 488 нм, Em 505LP + 515/20BP. Данные анализировали и визуализировали с использованием программного обеспечения FACSDiva 6.2. Гейты для eGFP+ клеток выставляли на основе отрицательного контроля - клеток без добавления плазмиды. Сортировку живых клеток для дальнейшего анализа выполняли с использованием сопла размером 85 мкм и соответствующих настроек давления.

3.9.5 Подготовка лизатов клеток

Сортированные клетки помещали в лед, центрифугировали при 500 g в течение 8 минут, супернатанты тщательно удаляли, осадки клеток помещали в жидкий азот и хранили при $-80\text{ }^{\circ}\text{C}$ до использования. Осадки размораживали на льду и ресуспендировали в 1x буфере для нанесения образцов ($4\times$: 250 mM Tris, pH 8,5, 1 mM EGTA, 1,5 mM MgCl_2 , 8% SDS, 40% глицерин, 0,04% бромфенолового синего), с добавлением 100 mM DTT, коктейля ингибиторов протеаз без EDTA (11873580001, ROCHE, Франция) и 250 ед/мл бензоназы (E1014, Sigma, США).

3.9.6 Иммуноблоттинг

Процедура иммуноблоттинга практически идентична таковой, указанной в пункте 3.3, за исключением того, что мы использовали 11% разделяющий полиакриламидный гель,

а в качестве маркеров молекулярной массы белка использовали PageRuler Prestained Protein Ladder (SM0671, Fermentas, США) и Spectra Multicolor Broad Range Protein Ladder (26623, Thermo Scientific, США). Первичные антитела: анти-GFP-HRP в разведении 1:2000 (600-103-215, Rockland Immunochemicals, США); анти-AAT антитела в разведении 1:5000 (козы поликлональные антитела, полученные против полноразмерной молекулы AAT плазмы (LS-B1454, LifeSpan Biosciences, США). Анти-AAT антитела использовали в паре со вторичными антителами, конъюгированными с пероксидазой, в разведении 1:7000 (ab6741, Abcam, Великобритания). Коммерческий AAT (AATp, 178251, Calbiochem, Германия) использовался в качестве положительного контроля для антител anti-AAT. Мембрану после обработки анти-GFP-HRP антителами дополнительно инкубировали с антителами против альфа-тубулина в разведении 1:5000 (DM1A, ab7291, Abcam, Великобритания) в паре со вторичными антителами в разведении 1:3000 (1706516, Bio-Rad, США). Относительные интенсивности сигнала от eGFP, нормированные по альфа-тубулину из лизатов клеток и остаточному альфа-тубулину из секретомов, оценивали с помощью программы Image Lab 6.0 (Bio-Rad, США). Все образцы лизатов отсортированных eGFP⁺ клеток после трансфекции плазмидами с короткими вставками наносили из расчета $70 \cdot 10^3$ клеток в лунку, с длинными - $90 \cdot 10^3$, мутантным пустым вектором - эквивалентно указанным количествам, соответственно; немутантной векторной плазмидой pEGFP-N1 - $15 \cdot 10^3$, которые дополнительно обогащали не трансфицированными отсортированными клетками до $70 \cdot 10^3$ и $90 \cdot 10^3$, соответственно. Наносили по 25 мкг суммарных белков секретомов, приготовленных в том же буфере для нанесения без фермента бензоназы, за исключением pEGFP-N1 (6 мкг) - чтобы предотвратить быстрое насыщение сигнала от eGFP.

Экспрессию конструкций HLT, DLT1 и DLT2 в клеточных линиях DU145 и HepG2 сравнивали с помощью двухфакторного дисперсионного анализа (ANOVA) с использованием программы GraphPad Prism 8.0.1 (GraphPad Software Inc., США). Различия со значением $p \leq 0,05$ считались статистически значимыми. Данные представлены в виде среднего \pm стандартное отклонение (SD) из двух независимых экспериментов для каждой клеточной линии. Диаграммы были построены с использованием Microsoft Excel 16 (Microsoft, США).

3.9.7 Дегликозилирование белков ферментом PNFaseF

Осадки клеток после трансфекции конструкциями HLT, DLT1 и DLT2 размораживали на льду и ресуспендировали в 1x буфере для денатурации гликопротеинов, входящего в состав набора для дегликозилирования PNFaseF (P0704S, NEB, США), с добавлением 250 U/ml бензоназы (E1014, Sigma, США), 1 mM PMSF и коктейля ингибиторов протеаз (11873580001, ROCHE, Франция). 20 мкг белков секретома осаждали с помощью экстракции MeOH/CHCl₃; остаточные вещества выпаривали при 80°C в течение 2 минут, а осадки растворяли в 1x буфере для денатурации гликопротеинов. Затем мы следовали протоколу производителя. После детекции полос eGFP, мембраны инкубировали

в PBS с 0,05 % азиды натрия в течение ночи, затем окрашивали антителами против альфа-тубулина DM1A (ab7291, Abcam, Великобритания) в паре со вторичными антителами, конъюгированными пероксидазой хрена (#1706516, Bio-Rad, США), затем мембраны стрипировали по мягкому протоколу Abcam (mild stripping) и повторно окрашивали анти-ААТ антителами (LS-B1454, LifeSpan Biosciences, США), конъюгированными с пероксидазой хрена с помощью коммерческого набора HRP Conjugation Kit (ab102890, Abcam, Великобритания).

3.9.8 Сайт-направленный мутагенез

Указанные конструкции плазмид мутировали с помощью протокола “quick change”. И использованные праймеры для введения точечной мутации (подчеркнуты) были следующие: HLT ATG2off прямой 5'-СТАСAGCCTTTGCAAAGCTCTCCCTGGGGAC-3', обратный 5'- GTCCCCAGGGAGAGCTTTGCAAAGGCTGTAG-3'; DLT2 oATG2 прямой 5'- CAGCCTTTGCAATGGTCTCCCTGGGGACC-3', обратный 5' - GGTCCCCAGGGAGACCCATTGCAAAGGCTG-3'; DLT2 uORFoff прямой 5'- СACTGGGAGTCATCAAGTGCCTTGACTCGG-3', обратный 5'- CCGAGTCAAGGCACTTTGATGACTCCCAGTG-3'. 50 нг исходных плазмид амплифицировали с помощью Pfu ДНК-полимеразы (#EP0502, Thermo Scientific, США) в 50 мкл в условиях: предварительная денатурация при 95°C 1,5 мин; 14 циклов: 95°C - 30 с, 60°C - 1 мин, 72°C - 13 мин; 72°C 10 мин. Реакционные смеси замораживали, затем половину смеси обрабатывали 0,5 мкл DpnI (ER1705, ThermoScientific, США) в течение 1,5 ч при +37°C, за которым следовала инактивация фермента в течение 20 мин при 80°C. 6 мкл реакционных смесей использовали для трансформации, вся трансформированная бактериальная смесь использовалась для выращивания колоний с мутированными плазмидами. Плазмиды очищали с использованием набора для выделения плазмид без эндотоксинов Plasmid Miniprep (#BC021, Evrogen, Россия). Последовательности ДНК верифицировали секвенированием (Evrogen, Москва, Россия) и наблюдали только желаемые мутации.

3.10 Изучение внутриклеточной локализации химерных продуктов трансляции

3.10.1 Конфокальная микроскопия на живых клетках

Клетки высевали на стеклянные слайды с дном из 8 камер в плотности $37,8 \cdot 10^3$ клеток в 250 мкл ростовой среды. Трансфекция конструкциями HLT, DLT1 и DLT2 проводилась с использованием реагента Lipofectamine 3000 согласно рекомендациям производителя. Через 24 часа после трансфекции обновили ростовую среду, и через 48 часов после трансфекции клетки анализировали с помощью микроскопии *in vivo*. Клетки дополнительно окрашивали ДНК-красителем Hoechst 33342 и красным ЭПР-трекером (ER-Tracker Red, E34250, Invitrogen, США) за 20 минут до наблюдений. Образцы изучали с

использованием конфокального микроскопа Nikon C2+ на основе инвертированного корпуса Nikon Eclipse Ti-E. Клетки поддерживали в CO₂-инкубаторе (OKOLab, Италия) при 37 °C и 100 % влажности. Микроскопию проводили с использованием 60× NA 1.4 план-апохроматической масляно-иммерсионной линзы и лазеров с длинами волн 402, 488 и 562 нм для возбуждения флуоресценции Hoechst 33342, eGFP и ER-Tracker Red, соответственно. Размер пинхола был установлен 40 мкм, создавая оптическую толщину среза 0,46 мкм при длине волны возбуждения 488 нм. Изображения дальше компоновались с использованием NIS Elements Viewer (Nikon, Япония) и Photoshop (Adobe Inc., США).

3.10.2 Конфокальная микроскопия на препаратах фиксированных клеток

Клетки высаживали на ночь в плотности $90 \cdot 10^3$ клеток в 500 мкл ростовой среды на покровные стекла размером 9×9 мм в 24-луночном культурном планшете. Среду обновляли через 24 часа после трансфекции плазмидными конструкциями HLT, ST1-4. Через 24 часа после смены среды клетки фиксировали 3,5% растворе параформальдегида в PBS, затем пермеабелизовали в 0,25% TritonX-100 в PBS и блокировали с раствором BSA/глицина (1% BSA, 22,52 мг/мл глицина в 0,1% PBST). Для окрашивания использовали антитела анти-GFP-HRP в разведении 1:750 в паре со вторичными антителами DAGI-Alexa Fluor-488 (ab150129, Abcam) в разведении 1:500. Антитела разбавляли в 1% BSA в PBST и центрифугировали на максимальных оборотах в течение 10 минут в настольной микроцентрифуге MiniSpin (Eppendorf, Германия). Препараты окрашивали ДНК-красителем DAPI в PBS перед последней отмывкой. Образцы покрывали ProLong Gold antifade (P36930, Invitrogen). Микроскопию проводили как описано в пункте 3.10.1 с использованием лазеров с длинами волн 402 и 488 нм для возбуждения флуоресценции DAPI и eGFP, соответственно.

3.11 Детекция и изучение внутриклеточной локализации предполагаемых эндогенных продуктов трансляции

3.11.1 С длинных транскриптов *SERPINA1*

3.11.1.1 Детекция предполагаемой эндогенной внутриклеточной изоформы ААТ

Клетки высаживали в плотности $20 \cdot 10^3$ клеток в 500 мкл ростовой среды на покровные стекла размером 9×9 мм в 24-луночном культурном планшете и растили в течение 72 часов. Клетки фиксировали 3,5% раствором параформальдегида в PBS, затем пермеабелизовали в 0,25% Triton X-100 в PBS и блокировали с раствором BSA/глицина (1% BSA, 22,52 мг/мл глицина в 0,1% PBST). Образцы также предварительно блокировали нормальными козьими антителами (NGtIgG) в концентрации 50 мкг/мл (sc-2028, Santa CruzBiotechnology) в течение ночи, затем промывали и окрашивали конъюгатом анти-ААТ IgG-Cy3 (LS-B1454, LifeSpan Biosciences, Seattle, WA, USA) либо NGtIgG-Cy3 в

концентрации 3 мкг/мл с добавлением или без 30 мкг/мл немеченых NGtIgG. Конъюгацию с флуорофором проводили с использованием Су3-NHS-эфира (PA13101, Cytiva), а меченые антитела тщательно очищали ультрафильтрацией в PBS на центрифужном фильтре Amicon Ultra-0.5 Centrifugal filter unit (UFC5010, Millipore, США) с порогом 10 кДа. Мышинные моноклональные антитела против Sc-35 (Ab11826, Abcam) или анти-альфа-тубулиновые антитела (Ab7291, Abcam) в паре с DAMI-Alexa Fluor 488 (ab150105, Abcam) использовали для окрашивания ядерных спеклов и микротрубочек, соответственно. Микроскопию проводили как описано в пункте 3.10.1 с использованием лазеров с длинами волн 402, 488 и 562 нм для возбуждения флуоресценции DAPI, Alexa Fluor-488 и Су3, соответственно.

3.11.1.2 Проверка ассоциации предполагаемой изоформы ААТ с РНК

Для изучения возможной ассоциации предполагаемой внутриклеточной эндогенной изоформы ААТ также подготовили препараты после обработки клеток *in situ*. Клетки, выращенные на покровных стеклах, промыли в холодном транспортном буфере (ТВ) [331]: 20 mM Hepes, pH = 7,3, 110 mM ацетата калия, 5 mM ацетата натрия, 2 mM ацетата магния, 1 mM EGTA, 2 mM DTT и 1× ингибитор протеаз без EDTA (11873580001, ROCHE, Франция). Клетки пермеабелизовали 40 мкг/мл дигитонином (300410, Calbiochem, Германия) в ТВ в течение 6 мин на льду. После двукратной отмывки холодным ТВ клетки инкубировали в ТВ с или без 50 мкг/мл РНКазы А (#EN0531, Fermentas, США) в течение 30 мин при +37°C. После двукратной отмывки ТВ клетки были фиксированы 3,5% параформальдегидом в ТВ. Препараты дополнительно пермеабелизовали 0,1% Triton X-100 в PBS перед иммуноокрашиванием по вышеуказанной процедуре.

3.11.2 С коротких транскриптов SERPINA1

3.11.2.1 Получение кроличьих поликлональных антител против E5-эпитопа, кодируемого экзоном 5 гена SERPINA1

Синтетический пептид CGKVVNPTQK (E5) с чистотой 98,6% был приобретен в компании EZBiolab (США). Белый кролик Новой Зеландии был трижды иммунизирован конъюгатом E5-KLH согласно 70-дневному протоколу, предоставленному Thermo Scientific: первая иммунизация проводилась в полном адъюванте Фрейнда, за которой последовали три бустерные иммунизации в неполном адъюванте Фрейнда. Через две недели после последней иммунизации отобрали иммунную сыворотку, и с помощью ИФА определили титр (>1:364500) (рисунок 10П, А). Затем анти-E5 поликлональные антитела очистили с помощью аффинной хроматографии на смоле Sulfolink (Thermo Scientific, США) с ковалентно пришитым пептидом E5. Аффинно очищенные анти-E5 антитела проверили методом ИФА на иммобилизованных конъюгатах BSA-E5 и KLH-E5 и иммобилизованном коммерческом ААТр из плазмы, против отрицательных контролей - BSA и KLH (рисунок 10П, Б). Сыворотка того же кролика была отобрана и до иммунизаций и использована в пункте 3.11.2.3.

3.11.2.2 Проверка доступности эпитопа E5 в составе нативной молекулы ААТ методом ИФА

ААТр (178251, Calbiochem, Германия) биотинилировали с помощью биотин-NHS-эфира (H1759, Sigma) в молярном соотношении 1:1 и тщательно очищали ультрафильтрацией в буфере Tris. 96-луночный планшет для ИФА (655,061, Greiner Bio-One, США) покрывали анти-E5 антителами в концентрации 5 мкг/мл в 100 мМ карбонат/бикарбонатном буфере с pH=9.6 при +4 °С в течение ночи и заблокировали 1% BSA в TBS (20 мМ Tris, 150 мМ NaCl, pH 7,4). Все отмывки проводили трижды раствором TBST. Антиген добавляли в конечной концентрации 200 нг/мл в блокирующем буфере 1% BSA в TBST. В случае термической денатурации антиген ААТр-биотин кипятили, охлаждали и разбавляли до конечной концентрации в блокирующем буфере, затем центрифугировали в течение 10 мин при 18000 g. Далее добавляли конъюгат стрептавидин-HRP в разведении 1:10000 (Sigma, США). Сигнал проявили с помощью субстрата ТМВ (34021, Thermo Scientific, США), реакцию останавливали серной кислотой, и сигнал (A(450 нм)-A(650 нм)) считывали на приборе Infinite M200Pro (Tecan, Швейцария). BSA-биотин и NRbIgG в тех же условиях, что и специфические антитела, использовались в качестве отрицательных контролей. Данные представлены как среднее±стандартное отклонение по трем повторам.

3.11.2.3 Детекция эндогенных эпитопов E5 методом иммунофлуоресценции

Препараты фиксированных клеток, подготовленных по методике из пункта 3.11.1.1, окрашивали сывороткой кролика до иммунизации, анти-E5 антителами и NRbIgG (sc-2027, Santa Cruz Biotechnology, США), за которыми последовала инкубация со вторичными антителами DAPI-Alexa Fluor 555 (ab150062, Abcam, Великобритания). Для блокировки неспецифического взаимодействия кроличьего IgG фиксированные клетки предварительно инкубировали с 10% сывороткой кролика до иммунизации в течение 1,5 часов, далее препараты отмывали и окрашивали первичными анти-E5 антителами, конъюгированными с Cy3 флуорофором, либо NRbIgG-Cy3 в концентрации 5 мкг/мл. Конъюгация антител с Cy3 проводилась с использованием Cy3-NHS-эфира (PA13101, Cytiva), и меченые антитела тщательно очищались ультрафильтрацией в PBS с порогом 10 кДа. Мышиные моноклональные антитела против Sc-35 (Ab11826, Abcam) и B23 (MAB4500-C, Sigma) в паре с DAPI-Alexa Fluor 488 (ab150105, Abcam) использовались для окрашивания ядерных и ядрышковых мишеней. Микроскопию проводили как описано в пункте 3.10.1 с использованием лазеров с длинами волн 402, 488 и 562 нм для возбуждения флуоресценции DAPI, Alexa Fluor-488 и Alexa Fluor-555/Cy3, соответственно.

3.12 Эволюционный анализ

3.12.1 Сборка и выравнивание последовательностей белков

Нуклеотидные последовательности мРНК *SERPINA1* для 15 видов млекопитающих были извлечены из базы данных NCBI (<http://www.ncbi.nlm.nih.gov/>). Номера доступа последовательностей мРНК, использованных для анализа, можно найти в таблице 4.1 раздела 4.5.3. Нуклеотидные последовательности, соответствующие областям альтернативных ОРС, были переведены в аминокислотную последовательность с использованием MEGA X [332] по стандартной таблице кодонов. Нуклеотидные и белковые последовательности были выровнены с использованием Clustal W [333] в UGENE [334].

3.12.2 Тестирование на давление отбора

Давление отбора на *SERPINA1* анализировали путем применения метода максимального правдоподобия в пакете программ PAML, версия 4 [335]. Выбранные последовательности выравнивали с использованием RevTrans с настройками по умолчанию [336]. Филогенетическое дерево для анализа строилось с использованием метода максимального правдоподобия и модели Hasegawa-Kishino-Yano [337]. Исходное дерево для эвристического поиска получалось автоматически, применяя алгоритмы метода присоединения соседей (Neighbor-Join) и BioNJ к матрице попарных расстояний, оцененной с использованием подхода Maximum Composite Likelihood (MCL), а затем выбирая топологию с наиболее высоким значением логарифмической функции правдоподобия. Для моделирования различий в скоростях эволюции между сайтами использовалось дискретное Гамма распределение (5 категорий (+G, параметр = 0.8135)). Филогенетическое дерево строилось в MEGA X [332]. Приложение 'codeml' из PAML использовалось для обнаружения кодонов под положительным, отрицательным и нейтральным отбором. Модели, специфичные для сайтов (M0, M1, M2, M7 и M8), позволяют отношению ω изменяться между сайтами, но не между видами. Для обнаружения положительного отбора на конкретных сайтах контрольный файл был настроен как model = 0, NSsites = 0, 1, 2, 3, 7 и 8. Применялись модели M0 (одно отношение), M1 (нейтральная), M2 (отбор), M3 (дискретная), M7 (бета) и M8 (бета и ω). В контрольном файле использовались следующие параметры: runmode = 0, seqtype = 1, CodonFreq = 2, model = 0, icode = 0, fix_kappa = 0, kappa = 4, fix_omega = 0, omega = 2, fix_blength = 1. ncatG был установлен = 3 для модели M3 и = 10 для моделей M7 и M8.

3.12.3 Оценка максимального правдоподобия и тест отношения правдоподобия

Два пары сопряженных моделей M1–M2 и M7–M8 использовались для теста отношения правдоподобия (LRT). Оба теста LRT M1–M2 и M7–M8 имеют 2 степени свободы (d.f.). Распределение χ^2 для соответствующих степеней свободы было получено путем выполнения программы 'chi2' в пакете PAML.

3.13 Нокаут экзона 5 гена *SERPINA1* с помощью технологии CRISPR/Cas9

Гидовые РНК выбрали с помощью веб-инструмента CRISPOR [338], принимая в качестве мишени первые 320 п.н. экзона 5 гена *SERPINA1*. В качестве гидов использовались наиболее высокорейтинговые (с наибольшей предполагаемой специфичностью, но наименьшим количеством нецелевых мишеней): Гид-1 нацеливает на последовательность GTGCTGACCATCGACGAGAA в 5'-проксимальной области кодирующей последовательности экзона 5; Гид-2 - GGATGACATТАААГААGGG в 5'-проксимальной области 3'-НТО, таким образом удаляя всю область гена, кодирующую пептид. Гиды клонировали в вектор pHU6-gRNA (плаزمид # 53188, Addgene) с использованием синтетических олигонуклеотидов (Евроген, Москва, Россия) по сайтам рестрикции Bst_V2I (СибЭнзим, Новосибирск, Россия). Плазмиды очищали с помощью набора для очистки плазмид без эндотоксинов Plasmid Miniprep (#BC021, Евроген, Москва, Россия), а затем секвенировали (Евроген, Москва, Россия). Плазмиду pIRES2-EGFP-Cas9, любезно предоставленную Дашинимаевым Э.Б. (ведущим научным сотрудником Института биологии развития им. Кольцова РАН, Москва, Россия), очищали тем же набором.

300*10³ клеток HepG2 высаживали на культуральные чашки диаметром 3,5 см за 24 часа до трансфекции. Трансфекция проводилась с использованием реагента для трансфекции TurboFect (R0531, Thermo Scientific, США), следуя протоколу производителя. Через 48 часов после трансфекции клетки собирали, и клетки, содержащие eGFP, клонировали с помощью клеточного сортирования в 96-луночные культуральные планшеты, заполненные по 100 мкл полной ростовой среды. Примерно через 2 недели начали проводить ПЦР-скрининг клонов с помощью метода прямого ПЦР на клетках [339], не требующего выделения геномной ДНК, с использованием Taq ДНК-полимеразы (Евроген, Москва, Россия) и праймеров, фланкирующих удаленный участок: прямой 5'-ATCAGCCTTACAACGTGTCTCT-3', обратный 5'-AAGACAGCACTGTTACCTGGA-3'. Протокол ПЦР: 94°C - 3 мин; 40 циклов: 94°C - 30 сек, 57°C - 20 сек, 72°C - 30 сек; 72°C 5 мин. Продукты ПЦР анализировали с помощью электрофореза в 1,5% агарозном геле. Клоны с моноспецифическим продуктом 195 п.н. выбирали для дальнейшего исследования, наращивали и выделяли геномную ДНК с помощью набора для выделения ДНК из клеток, тканей и крови (Биолабмикс, Новосибирск, Россия). Из-за высокого содержания GC в целевой последовательности и меньшей длины продукта с мутацией фрагмент 195 п.н. амплифицировался моноспецифически даже у гетерозиготных клонов. Поэтому мутантные клоны дополнительно проверяли на отсутствие аллели дикого типа с использованием ПЦР с наборами праймеров из удаленной области в сочетании с фланкирующими праймерами (набор №: прямой, обратный): (1: 5'-ATCAGCCTTACAACGTGTCTCT-3', 5'-AAGACAGCACTGTTACCTGGA-3'); (2: 5'-CCCATGTCTATCCCCCG-3', обратный из набора 1); (3: 5'-AGAGGCCATACCCATGTCTATC-3', 5'-

TTGGGTGGGATTCACCACTTTTCC-3'); (4: 5'-GGTCAAGTTCAACAAACCSTTTG-3', обратный из набора 3); (5: прямой из набора 4, 5'-TGAGGAGCGAGAGGCAGT-3'); (6: 5'-TACCAAGTCTCCCCTTTCATG-3'; обратный из набора 5). Для наглядности мы приводим схему расположения праймеров на Рисунке 13П, А. Продукты анализировали с помощью электрофореза в 1,5% агарозном геле. Гомозиготные нокауты выбирали для дальнейшего анализа с помощью иммуноблоттинга и иммунофлуоресценции на E5-эпитоп (в соответствии с протоколами, описанными ранее в пунктах 3.3 и 3.11.2.3).

4. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Предпосылкой для темы данной диссертационной работы является выполнение проекта по поиску биомаркеров опухоли предстательной железы (ОПЖ). Проект предполагал анализ и сравнение протеомных карт сывороток пациентов с установленным диагнозом ОПЖ, а также пациентов с гиперплазией ПЖ и здоровых волонтеров. Сыворотки предварительно очищали от мажорного белка - альбумина - по стандартной методике аффинной хроматографии на колонке с пришитым лектином конканавалином А. Пул связавшихся белков использовали для получения протеомных карт методом двумерного электрофореза. При анализе протеомных карт обнаружили, что мажорные М4 и М6 гликоформы ААТ (рисунок 2.3, А), идентифицированного методом MALDI-TOF/TOF-масс-спектрометрии, при гиперплазии и ОПЖ более представлены, чем в сыворотке здоровых волонтеров [128]. Предположили, что при патологиях ПЖ повышается общий уровень ААТ и методом турбидиметрии показали, что концентрация ААТ в сыворотке крови больных ОПЖ ($2,13 \pm 0,28$ г/л), отобранной до хирургических вмешательств, достоверно выше таковой у здоровых волонтеров ($1,41 \pm 0,36$ г/л) и у пациентов с гиперплазией ПЖ ($1,63 \pm 0,21$ г/л), концентрация у которых занимала промежуточное значение [32]. На примере одного пациента с ОПЖ показали, что после гормональной терапии, до брахитерапии, уровень ААТ снизился до нормы [32]. Предположили, что вклад в повышение уровня ААТ в сыворотке вносят не только гепатоциты печени, но и сами опухолевые клетки. Опухолевая ткань содержит не только трансформированные клетки, но также ряд других типов клеток, таких как опухоль-ассоциированные иммунные клетки (лимфоциты, макрофаги, нейтрофилы, дендритные клетки), фибробласты, клетки прорастающих кровеносных сосудов. Этот факт, доступность, а также удобство работы, обусловили изучение экспрессии гена *SERPINA1* в клеточных линиях, происходящих из опухолей человека.

Исходя из имевшихся на момент и появлявшихся в процессе выполнения диссертационной работы данных об экспрессии гена *SERPINA1* и вдохновляясь обновляющимся списком функций полноразмерного ААТ и его С-концевых фрагментов, мы выбрали два основных и тесно связанных направления исследований. Ниже мы перечислим оба, не нарушая последовательности изложения данных о природе ААТ и его фрагментов в разделе “обзор литературы”, которая является более логически выстроенной, несмотря на то, что хронологически второе из излагаемых направлений определилось раньше.

Считается, что со всех альтернативно сплайсирующихся мРНК *SERPINA1*, содержащих главную ОРС, транслируется один и тот же секреторный белок. Мы задались вопросом: что, если это не так? Разнообразие свойств и ролей ААТ во многом обусловлено его способностью связываться с широким спектром молекулярных партнеров, а относительно невысокая константа связывания ААТ с некоторыми из них может компенсироваться высокой физиологической концентрацией серпина. Не исключено, что

на способность ААТ связывать те или иные “лиганды” влияет исходное конформационное состояние молекулы. Это предположение подтверждается тем наблюдением, что ААТ-Z, имея более гидрофобный профиль поверхности, более эффективно связывает жирные кислоты, чем ААТ-M вариант [249]. Мы предположили, что *in vivo*, для осуществления некоторых из рассмотренных в обзоре литературы функций, могут задействоваться не известные ранее конформационно дистинктные протеоформы/изоформы ААТ, длина полипептидной цепи которых соизмерима с таковой известного секреторного белка. Кроме того, как было упомянуто в обзоре литературы, есть спорадические данные, намекающие на большее разнообразие протеоформ ААТ, чем принято считать, и, что представляется весьма интригующим, на существование внутриклеточных изо/протеоформ (предполагаемая цитоплазматическая [39] и ядерная [40]). Поиску альтернативных протеоформ/изоформ ААТ и пути их образования посвящено одно из двух основных направлений диссертационной работы.

Второе направление связано с образованием С-концевых пептидов ААТ. Считается, что единственным механизмом их образования является протеолитическое расщепление предшественника - зрелой молекулы ААТ - во внеклеточном пространстве [55, 328]. Мы предположили, что помимо основного пути — протеолитического, связанного со специфическим расщеплением полноразмерного ААТ, возможно, существует альтернативный механизм синтеза пептидов, протекающий независимо от белка-предшественника. Эта же гипотеза фактически была изложена еще Церцеками: в своих исследованиях протуморогенных CRISPP пептидов они предположили, что участок гена *SERPINA1*, кодирующий пептиды (экзон 5), амплифицирован в геноме опухолевой клетки [340]. В качестве проверки своей гипотезы авторы провели дот-гибридизацию с использованием биотинилированных ДНК-зондов, комплементарных участкам, кодирующим N-концевую область ААТ и N-конец пептидов CRISPP (SIPP...), и обнаружили, что как на уровне геномной ДНК, так и на уровне мРНК, хемилюминесцентный сигнал от зонда на пептид-кодирующую область выше, чем на область, кодирующую N-конец ААТ, в опухолевых клетках по сравнению с нормальными (пары Malme3 нормальные фибробласты кожи - Malme3M клетки меланомы; WI-38 нормальные фибробласты легких - WI-38, трансформированные вирусом SV-40) [340]. На этом Церцеки закончили свои исследования и их гипотеза не имела дальнейшего развития. Мы предположили, что С-концевые пептиды образуются независимо от белка-предшественника путем трансляции с коротких транскриптов гена *SERPINA1*. Поиску спектра коротких транскриптов и синтезирующихся с них пептидных/белковых продуктов, а также возможных механизмов происхождения коротких транскриптов посвящено второе направление нашего исследования.

Итак, ретроспективная оценка проделанной работы способствовала осмыслению, что глобально она состоит из двух частей: изучение экспрессии длинных (содержащих

главную OPC) и предполагаемых коротких транскриптов (усеченных с 5'-конца и не содержащих главную OPC) гена *SERPINA1* в клеточных линиях, происходящих из опухолей человека. Забегая вперед, стоит отметить, что результаты по длинным и коротким транскриптам будут изложены параллельно вплоть до установления структуры их 5'-концов (начиная с пятого экзона и вплоть до 5'-концов транскриптов), поскольку они экспериментально связаны настолько, что едва делимы. Но прежде, чем приступить к поиску спектра транскриптов *SERPINA1*, ниже мы проанализируем перспективность заявленной цели, а также проведем скрининг клеточных линий для выбора клеточной модели исследования.

4.1 Анализ потенциальных альтернативных сайтов инициации трансляции в рамке с главной OPC *SERPINA1*

Поскольку поставленная нами цель в конечном итоге предполагает поиск альтернативных продуктов трансляции, для начала нам необходимо рассмотреть структуру гена *SERPINA1* на предмет наличия потенциальных альтернативных открытых рамок считывания, находящихся в рамке с *SERPINA1*, с учетом контекста Козак для их предполагаемых сайтов инициации трансляции. Поскольку определению структуры 5'-НТО будет посвящена одна из глав излагаемых результатов диссертации, для упрощения мы не будем рассматривать потенциальные старты трансляции в 5'-НТО гена, а сосредоточимся на его кодирующей области. Также, для простоты, мы рассмотрим только канонический стартовый кодон AUG, кодирующий метионин у эукариот, поскольку, будучи в хорошем контексте, он служит более сильным инициатором трансляции, чем неканонические старты [341] и поскольку альтернативная инициация трансляции в пределах кодирующей области гена (dTIS, downstream translation initiation site) в основном происходит именно с AUG кодона, в отличие от инициации в 5'-НТО, где первостепенную роль играет старт-кодон CUG [342].

Стартовый кодон главной OPC гена *SERPINA1* ATG-1 обладает субоптимальным контекстом Козак с остатком аденина в положении -3 (рисунок 4.1, А, В). Помимо главного ATG-1, девять дополнительных ATG кодонов распределены по всей кодирующей области гена *SERPINA1* (рисунок 4.1, А, В). Один из них - ATG-2 - расположен ниже ATG-1 во втором экзоне, четыре (ATG-3 - ATG-6) находятся в экзоне 3, и еще четыре расположены в экзоне 5. ATG2-6 ввиду контекста потенциально могли бы служить альтернативными сайтами инициации трансляции, в то время как не все, а три из четырех ATG кодона (ATG-7, 9 и 10), в экзоне 5 находятся в контексте Козак, подходящем для инициации трансляции (рисунок 4.1, В). С учетом общепринятой пороговой длины в 100 кодонов, потенциальные альтернативные OPC со стартами в пятом экзоне можно рассматривать как *изоформные* короткие OPC (икОPC) (изоформные по отношению к главной OPC) *SERPINA1* по классификации, предложенной в работе [343]. Отметим, что именно “слабый” ATG-8 подразумевался как инициатор синтеза и CRISPP пептидов Церцекон при раке [340], и

пептида С36 при воспалительных процессах [344]. Таким образом, “под подозрением” у нас остаются 8 из 9 потенциальных дополнительных стартовых кодонов ATG гена *SERPINA1*.

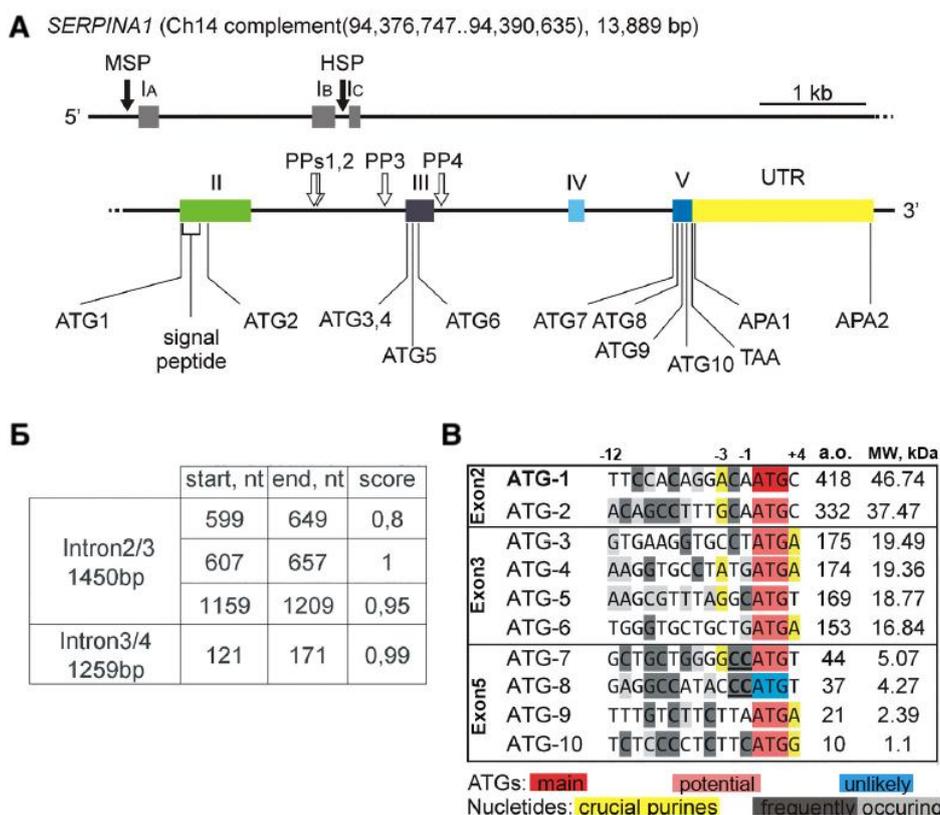


Рисунок 4.1. Оценка потенциала гена *SERPINA1* для альтернативной экспрессии. **А**) Схема гена *SERPINA1*. MSP - миелоид-специфический промотор, HSP - гепатоцит-специфический промотор, PPs - предполагаемые промоторы (предсказанные с использованием онлайн-источника https://www.fruitfly.org/seq_tools/promoter.html с пороговым значением = 0,8); APAs - альтернативные сайты полиаденилирования (см. раздел 4.3.1) [38, 122]. **Б**) Интронные позиции и предсказанные вероятности наличия предполагаемых промоторов. **В**) Анализ контекста Козак для главного и девяти предполагаемых альтернативных сайтов инициации трансляции. Главный стартовый кодон ATG-1 выделен жирным шрифтом. Нуклеотиды, часто встречающиеся на соответствующих позициях возле активных стартовых кодонов, согласно [345], закрашены темным (наиболее частые) и светло-серым (следующие по частоте). Ключевые пурины в -3 положении (предпочтительно А) и +4 (предпочтительно G) закрашены желтым. Помимо консенсусного остатка +4G, желтой заливкой выделены также остатки +4А, поскольку старт-кодоны с +4А и в отсутствие пуринов в положении -3 обладают потенциалом к инициации трансляции [346]). Благоприятный динуклеотидный мотив СС выделен подчеркнутым жирным шрифтом [347].

Если рассматривать предполагаемую альтернативную инициацию трансляции в пределах главной ОРС *SERPINA1*, приводящей к синтезу изоформы ААТ, максимально близкой по длине к известному секретируемому аналогу (то есть не являющейся малым белком или пептидом [348]), то на роль инициаторного кодона претендует следующий за первым, AUG-2. Поскольку AUG-1 находится в субоптимальном контексте Козак (С в положении +4), возможным механизмом трансляции предполагаемой изоформы может быть сканирование с уткой вдоль *длинного* транскрипта, содержащего главную ОРС [349, 350, 351]. Подобная протеоформа ААТ в работах не встречается.

Иначе обстоит дело с возможной инициацией трансляции со стартов в экзоне 5, которая приводила бы к синтезу С-концевых пептидов ААТ. Поскольку им предшествуют

главный и пять потенциальных старт-кодона, инициация трансляции икОПС за счет сканирования с утечкой вдоль полноразмерного (длинного) транскрипта *SERPINA1* не представляется возможной [352]. Это ограничение снимается, если транскрипт не содержит главную ОПС, то есть усечен с 5'-конца и представляет собой *короткий* транскрипт.

Действительно, в геномном браузере UCSC имеются несколько аннотированных коротких транскриптов на основе исследований, проведенных в 1980-х годах. X02920 включает EX3 (экзон 3) (частично), EX4 и EX5 [353]; V00496 - EX4 (частично) и EX5 [354]; M26123 содержит только EX5 [355]. Происхождение и функциональное значение этих транскриптов остаются неизвестными. Также, параллельно с нашими исследованиями, в кДНК-библиотеке лейкоцитов было подтверждено наличие коротких транскриптов *SERPINA1* (EX1A/1C-EX4-EX5 и EX1A/1C-EX5) стандартным ПЦР путем амплификации известных последовательностей из экзонов 1A/1C и экзона 5 гена *SERPINA1* [344]. Несмотря на то, что короткие транскрипты и содержат икОПС для синтеза С-концевых пептидов ААТ, на данный момент нет достоверных свидетельств их значимости в качестве источника этих пептидов.

Пептиды (таблица 2.2), точно соответствующие потенциальным продуктам трансляции предполагаемых икОПС *SERPINA1*, обнаружены *in vivo*. Так, SPAAT - пептид С44, состоящий из 44 а.о. и обнаруженный в плаценте [48, 49], соответствует икОПС-АТГ-7, а С21 и С10, найденные в моче беременных женщин с преэклампсией [356, 357] и беременных, в утробе которых развивается плод с синдромом Дауна [358], соответствуют икОПС-АТГ-9 и икОПС-АТГ-10, соответственно. Кроме того, предсказываются несколько внутригенных промоторов с высокими вероятностями в интронах ниже экзона 2 (рисунок 4.1, А, Б). Эти данные не противоречат, а скорее свидетельствуют в пользу альтернативной инициации трансляции, возможно, управляемой транскрипцией, которая может привести к образованию С-концевых пептидов ААТ.

4.2 Выбор модельной клеточной линии

4.2.1 *Определение относительного уровня экспрессии гена SERPINA1 в клеточных линиях, первичной культуре и тканях человека методом ПЦР в реальном времени*

Выбирая клеточную модель, мы руководствовались двумя параметрами: 1) клетки должны обладать существенной экспрессией гена *SERPINA1* и, поскольку С-концевые пептиды кодируются его пятым экзоном, 2) в этих клетках уровень экспрессии пептид-кодирующей области пятого экзона должна существенно превышать экспрессию вышестоящих экзонов.

Мы определили относительные уровни содержания двух участков мРНК *SERPINA1* в пулах кДНК, полученных от суммарных РНК ряда клеточных линий, иммортализованных и первичных культур и нормальных тканей человека, с помощью ПЦР в реальном времени

(ПЦР-РВ). Мы амплифицировали следующие два участка кДНК *SERPINA1*: стык экзонов 2 и 3 (EX2/3), что отражает экспрессию длинных транскриптов, кодирующих полноразмерный белок, и пептид-кодирующую область экзона 5 (EX5), что отражает суммарный уровень экспрессии длинных и предполагаемых коротких транскриптов *SERPINA1*. В приложении к диссертации приведен рисунок 1П со схемой расположения праймеров и оценкой их специфичности. Мы высчитали экспрессию предполагаемых коротких транскриптов в отдельности и построили диаграмму относительной экспрессии длинных и предполагаемых коротких транскриптов между исследуемыми образцами (рисунок 4.2). Результаты свидетельствуют в пользу существования коротких транскриптов в большинстве образцов (16 из 19). Стоит отметить, что по результатам ПЦР-РВ потенциальные короткие транскрипты в иммортализованной линии эпителия простаты RWPE-1 и в нормальных эндотелиоцитах пуповины HUVEC практически отсутствуют, в то время как в опухолевых клеточных линиях и в тканях можно ожидать их экспрессию. Ткани представлены набором клеток разных типов и их вклад в экспрессию того или иного типа транскриптов может отличаться, и данные представляют усредненный уровень экспрессии.

По полученным результатам в клеточной линии HepG2 максимальна не только экспрессия длинных транскриптов, как и ожидалось, но также и экспрессия предполагаемых коротких транскриптов. Из непеченочных клеток длинные транскрипты максимально экспрессируются в линии молочной железы MDA-MB-231, но по экспрессии предполагаемых коротких транскриптов эта линия уступает клеткам опухоли простаты DU145. Наибольшее соотношение экспрессии предполагаемых коротких транскриптов к длинным наблюдается в гормон-нечувствительных линиях опухоли простаты PC3 и в особенности DU145 (рисунок 4.2). Поскольку в клетках PC3 длинные транскрипты практически не экспрессируются (для EX2/3 среднее значение $St=35$), на этом этапе скрининга мы предварительно отобрали клеточную линию DU145 (для EX2/3 среднее значение $St=29,42$) наряду с печеночными клетками HepG2.

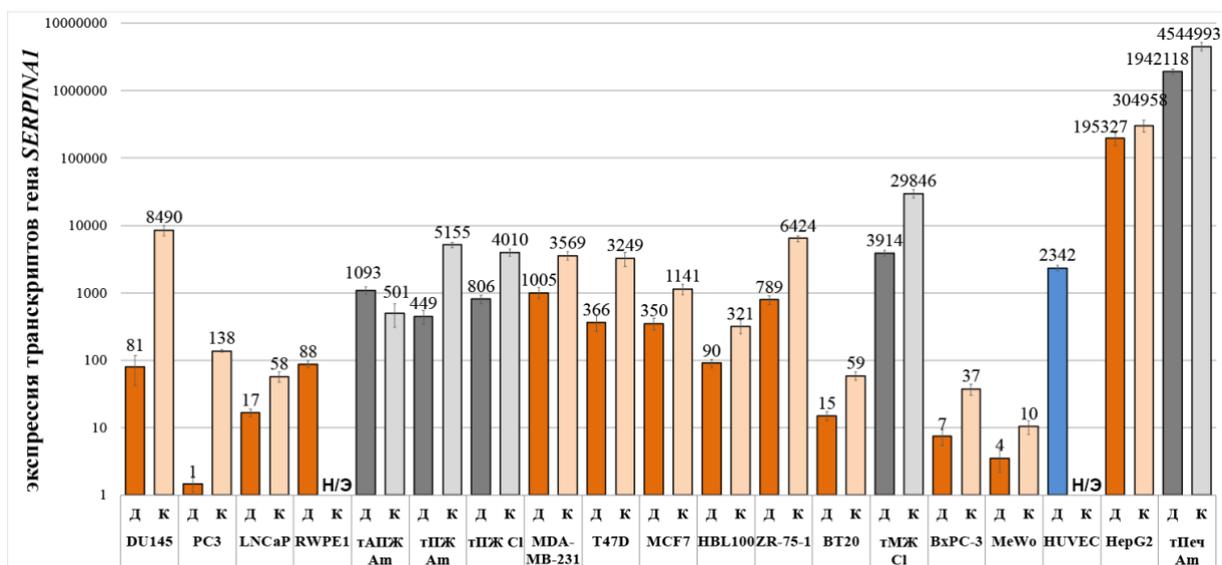


Рисунок 4.2. Относительный уровень экспрессии длинных (Д) и предполагаемых коротких (К) транскриптов гена *SERPINA1* в опухолевых клеточных линиях (предстательной железы DU145, PC3, LNCaP; молочной железы MDA-MB-231, T47D, MCF7, ZR-75-1, BT20; поджелудочной железы ВхРС-3; меланомы MeWo; гепатоцеллюлярной карциномы HepG2), иммортализованной клеточной линии эпителия нормальной предстательной железы RWPE-1 и спонтанно иммортализованной клеточной линии эпителия молочной железы HBL100 (оранжевые столбцы), первичных клетках HUVEC (синие столбцы), а также тканях человека (серые столбцы) производителей Clontech (Cl) и Ambion (Am): АПЖ- аденома предстательной железы, ПЖ - нормальная предстательная железа, МЖ - нормальная молочная железа, Печ- печень; т-ткань. В качестве гена домашнего хозяйства использовали GAPDH, экспрессия подсчитана относительно PC3 EX2/3. Данные представлены по трем повторам как среднее±стандартное отклонение.

4.2.2 Детекция длинных и коротких транскриптов *SERPINA1* с помощью нозерн-гибридизации

Результаты ПЦР-РВ свидетельствуют в пользу существования коротких транскриптов *SERPINA1*, содержащих участок экзона 5. Далее мы приступили к прямой детекции длинных и коротких транскриптов в пуле суммарных РНК клеточных линий и тканей с помощью нозерн-гибридизации. В качестве зонда для гибридизации использовали радиоактивно меченый кРНК зонд, комплементарный участку мРНК протяженностью 260 н.о., соответствующему 5 экзону *SERPINA1* (координаты в UCSC Genome Browser: chr14:94,378,432-94,378,692). В качестве положительных контролей использовали генно-инженерные транскрипты, содержащие участки 5 экзона *SERPINA1*: короткий (319 н.о.) - в кодирующей части полностью соответствующий зонду для гибридизации и длинный (1187 н.о.) - с координатами кодирующей части chr14:94,378,432-94,379,535 (UCSC Genome Browser). На рисунке 4.3, А представлен скан мембраны после нозерн-гибридизации на суммарной РНК из клеточных линий опухолевого происхождения. Наиболее высокий уровень экспрессии гена *SERPINA1* наблюдается в печеночной клеточной линии HepG2, что подтверждает результаты ПЦР-РВ (рисунок 4.2). Длинные мРНК *SERPINA1* детектируются в районе 3300 н.о. (содержат полноразмерный 3'-НТО длиной 1702 н.о.) и 1500 н.о. (содержат укороченный 3'-НТО за счет альтернативного полиаденилирования, см. раздел 4.3.1). Также, как и ожидалось из результатов ПЦР-РВ (рисунок 4.2), в области около 1000 н.о. и ниже детектируются короткие транскрипты, содержащие экзон 5 (рисунок 4.3, А). В других, не печеночных клеточных линиях короткие транскрипты практически не детектируются (рисунок 4.3, А). Среди них максимальная экспрессия длинных транскриптов наблюдается в гормон-нечувствительной клеточной линии опухоли простаты DU145, трижды негативной опухоли молочной железы MDA-MB-231 и HER2-отрицательной линии T47D (рисунок 4.3, А).

кишечник также обладает способностью к активному самообновлению [360]. Известно, что клетки проксимальных канальцев почек обладают значительной способностью к регенерации после острого повреждения почек: регенерация почки в значительной степени обеспечивается выжившими клетками канальцев, которые могут повторно вступать в клеточный цикл, пролиферировать и дифференцироваться в различные типы клеток, необходимые для восстановления канальцев [361]. Поскольку С-концевые пептиды ААТ повышают клеточную пролиферацию [313, 326, 327], в том числе в первичной культуре гепатоцитов крысы, не исключено, что пептиды, возможно происходящие путем инициации трансляции с коротких транскриптов, могут принимать участие в обновлении органов и тканей человека.

Детектируемые варианты коротких транскриптов могут различаться структурой 5'-концов, тогда как диффузность сигналов от коротких транскриптов может быть обусловлена разной длиной поли(А)-хвостов и/или высокой вариабельностью их 5'-концов. Обнаруженные длины коротких транскриптов позволяют заключить, что в них задействуется альтернативное полиаденилирование, приводящее к более короткой 3'-НТО (рисунок 4.1, А, см. раздел 4.3.1), что приводит к повышенной стабильности таких молекул РНК, в то время как полноразмерная 3'-НТО *SERPINA1* общей длиной в 1702 н.о. – потенциальная мишень для микроРНК [362].

Что касается тканей простаты и молочной железы, то экспрессия гена в них не столь ярко выражена, как в опухолевых клеточных линиях, происходящих из этих органов (сравните рисунок 4.3 А и Б), что противоречит данным ПЦР-РВ (рисунок 4.2, если сравнить экспрессию длинных транскриптов для клеточных линий и соответствующих тканей). Это может свидетельствовать о более сложной регуляции экспрессии изучаемого гена, поскольку, в отличие от нозерн-гибридизации, ПЦР-РВ не различает “продукты”, соответствующие кодирующей и антисмысловой цепям ДНК.

4.2.3 Экспрессия антисмысловой цепи ДНК в локусе пятого экзона гена *SERPINA1* с образованием предполагаемого антисмыслового транскрипта

Результаты относительной экспрессии длинных и коротких транскриптов в ПЦР-РВ и детекции транскриптов, содержащих экзон 5, в нозерн-гибридизации противоречат друг другу. Так, в клетках DU145 по результатам ПЦР-РВ в основном экспрессируются короткие транскрипты (рисунок 4.2) и их уровень ниже такового в HepG2 в среднем всего в 36 раз (при расчетном соотношении экспрессии длинных в 2423 раза), в то время как в нозерн-гибридизации в DU145 отчетливо детектируются только длинные транскрипты (рисунок 4.3, А). Принципиальная разница двух методов заключается в том, что в ПЦР-РВ амплифицируется продукт, который может соответствовать смысловой и/или антисмысловой цепи ДНК изучаемой области гена, в то время как в нозерне детекция цепь-специфична. Расхождение в результатах можно объяснить тем, что в локусе экзона 5 гена

SERPINA1 экспрессируется также обратная (антисмысловая) цепь ДНК, в результате синтезируется антисмысловой по отношению к мРНК *SERPINA1* транскрипт, содержащий частичную или полную последовательность экзона 5.

Для проверки этого предположения мы провели ПЦР на одноцепочечной кДНК DU145 и HepG2 в два раунда: 1) Для дифференциации транскриптов с антипараллельных цепей ДНК в первом раунде амплифицировали фрагменты в *однаправленной* ПЦР с использованием прямого (амплификация с ростом количества “смыслового” продукта в линейной прогрессии, праймер ложится на последовательность, кодирующую активный сайт молекулы ААТ М358-S359) *либо* обратного (амплификация с ростом количества “антисмыслового” продукта в линейной прогрессии, праймер ложится на последовательность, кодирующую С-конец ААТ) ген-специфического праймера; 2) ПЦР-РВ с парой праймеров на участок пятого экзона (кодирующий CRISPP пептиды), включающий контроль - референс с соответствующей исходной кДНК в соответствующем разбавлении. На рисунке 4.4 представлены результаты, свидетельствующие об экспрессии антипараллельной цепи ДНК во всех образцах: в клетках DU145 предполагаемый антисмысловый транскрипт экспрессируется в среднем в 3,5 раза слабее, чем смысловой, а в клетках HepG2 - в 3255 раз сильнее. Поскольку оценка относительной эффективности амплификации продуктов с ген-специфических праймеров в первом раунде *однаправленного* ПЦР не представляется возможной (в отличие от расчетной эффективности с коэффициентом 1,92 во втором раунде - ПЦР-РВ), эти относительные цифры следует воспринимать как качественный показатель, а не абсолютные величины. Тем не менее, они свидетельствуют в пользу экспрессии антисмысловой цепи ДНК в локусе экзона 5 *SERPINA1*.

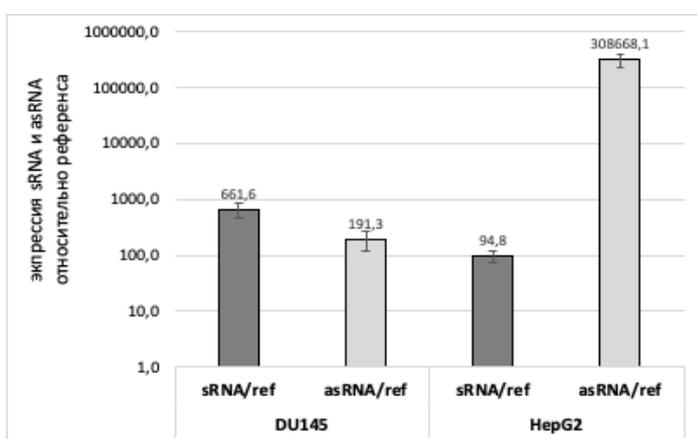


Рисунок 4.4. Экспрессия смысловой и антисмысловой цепей ДНК в локусе экзона 5 гена *SERPINA1* относительно соответствующего фона (референсного образца). asRNA - антисмысловый транскрипт, sRNA - смысловый транскрипт, Ref - референс, включенный во второй раунд ПЦР-РВ на исходных кДНК в соответствующем разбавлении. Данные представлены по трем повторам как среднее±стандартное отклонение.

При изучении транскриптома клеток человека обнаружено, что порядка 2900-6400 генов человека экспрессируются с противоположной цепи, синтезируя природные антисмысловые транскрипты (ПАТ, NAT - natural anti-sense transcripts), и их экспрессия может регулироваться альтернативными транскрипционными регуляторными элементами [363]. Антисмысловые транскрипты менее подвержены сплайсингу [363] и предполагается,

что большинство ПАТ регулируют экспрессию непосредственно гена, с противоположной цепи которого они экспрессируются, то есть являются цис-регуляторами экспрессии. Согласно механистической модели “транскрипционной коллизии”, сама транскрипция с противоположной геновой цепи ДНК, а не ПАТ как таковой, может являться фактором, препятствующим “смысловой” транскрипции, и эффект столкновения тем выраженной, чем больше (более 2000 п.н.) область перекрытия границ “смысловой” и “антисмысловой” транскрипции [364]. Однако, транскрипция в обоих направлениях может быть разрознена как во времени, так и в пространстве (происходить на разных аллелях).

ПАТ являются гетерогенной группой регуляторов экспрессии смысловых РНК [365]. ПАТ могут регулировать экспрессию на транскрипционном уровне, являясь эпигенетическими регуляторами: взаимодействуя с ДНК в соответствующем локусе и привлекая ДНК- или гистон-модифицирующие ферменты, ПАТ как ингибируют, так и активируют экспрессию смысловых РНК. Образуя РНК-РНК дуплексы со смысловым транскриптом в ядре, ПАТ направляют сплайсинг и изменяют баланс между сплайс-изоформами мРНК. На пост-транскрипционном уровне регуляции экспрессии гена, образуя РНК-РНК дуплексы в цитоплазме, ПАТ могут маскировать сайты связывания мРНК с микроРНК, изменять вторичную и/или третичную структуру мРНК, маскируя или демаскируя сайты связывания с РНК-связывающими белками, например, мРНК-стабилизирующим белком HuR, тем самым влияя на ее стабильность и трансляцию. В то же время, сам дуплекс ПАТ-мРНК может служить источником малых интерферирующих РНК, которые регулируют экспрессию обоих транскриптов [365]. Так, для мРНК *VASE1* (*beta*-secretase 1) было показано, что экспрессия его ПАТ повышается при болезни Альцгеймера, и ПАТ повышает стабильность мРНК путем изменения ее вторичной и/или третичной структуры и защиты комплементарных участков смысловой РНК в пределах ее ОРС от связывания с микроРНК-485-5p (но обычно микроРНК направлены против 3'-НТО), тем самым повышая выход белкового продукта [366].

Поскольку мы оценили возможность экспрессии с противоположной цепи только в участке пятого экзона, на данном этапе сложно судить о протяженности области перекрытия смыслового и антисмыслового транскриптов. Наши результаты показывают, что экспрессия антисмыслового транскрипта значительно сильнее смыслового в линии HepG2, а поскольку гепатоциты являются основным продуцентом ААТ, можно ожидать роль предполагаемого антисмыслового транскрипта *SERPINA1* в стабилизации смыслового и повышении выхода белка, возможно, путем, аналогичным вышеупомянутому, а именно за счет изменения вторичной/третичной структуры и/или маскирования сайтов связывания с микроРНК.

4.2.4 Антисмысловой транскрипт - возможный положительный регулятор синтеза ААТ. Содержание ААТ в секретах клеточных линий

Чтобы ответить на вопрос, как экспрессия антисмысловой цепи ДНК влияет на выход белкового продукта, мы провели корреляционный анализ экспрессии участков гена *SERPINA1* с уровнем белка в культуральных супернатантах (секретах) клеточных линий опухолевого происхождения.

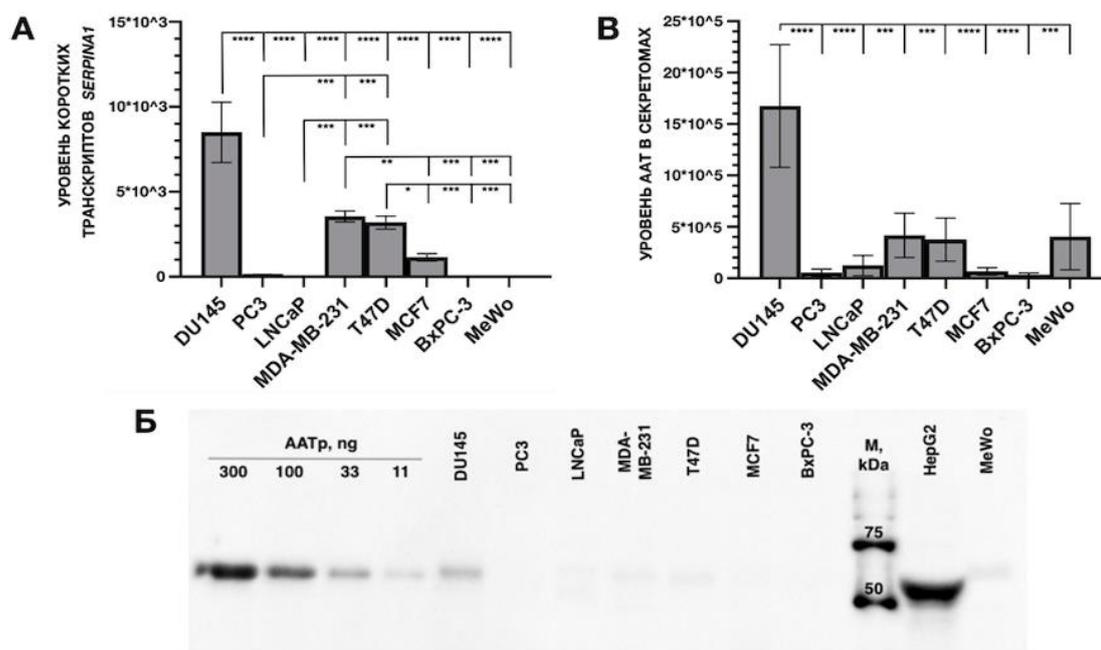


Рисунок 4.5. Экспрессия коротких транскриптов гена *SERPINA1* положительно коррелирует с уровнем секреторного ААТ в культивируемых клеточных линиях. А) Относительная экспрессия коротких транскриптов *SERPINA1* в выбранных клеточных линиях. Б) Детекция ААТ в секретах клеточных линий с помощью иммуноблоттинга с использованием антител против полноразмерного ААТ. Коммерческий препарат ААТ плазмы (ААТр) в указанных количествах и секретом HepG2 использовали в качестве положительных контролей. В) Денситометрический анализ относительного количества белка в секретах. Значимые отличия, выявленные с помощью множественного сравнения критерием Тьюки, на уровне РНК и белка: * - $p < 0.05$; ** - $p < 0.01$; *** - $p < 0.001$; **** - $p < 0.0001$. Данные представлены по трем повторам как среднее ± стандартное отклонение.

На рисунке 4.5 для удобства представления данных приведен профиль экспрессии коротких транскриптов *SERPINA1* из раздела 4.2.1 в тестируемых клеточных линиях, отобранных для анализа уровня ААТ в секретах. Чтобы избежать ложноположительной корреляции, данные для линии HepG2 исключили из статистического подсчета (эти данные не противоречат нашему выводу). Достоверность и значение коэффициента корреляции Пирсона ($R^2=0,8575$, $p=0,001$) показывает, что уровни коротких транскриптов и белка положительно коррелируют. Напротив, для экспрессии стыка EX2/3 и уровня белка ни положительной, ни отрицательной корреляции не выявлено ($R^2=0,0000005749$, $p=0,9986$). Кроме того, обнаружили, что надежность положительной корреляции общего уровня экспрессии пятого экзона (в составе и длинных и коротких, в том числе предполагаемых

антисмысловых, транскриптов) с уровнем секретируемого белка ниже, чем коротких транскриптов в отдельности, и составляет $R^2=0,7932$, $p=0,003$. Из полученных результатов можно сделать вывод о том, что антисмысловой транскрипт может служить положительным регулятором в конечном итоге синтеза ААТ.

К сожалению, транскрипция с противоположной цепи гена *SERPINA1* представляет собой совсем не исследованный феномен и в работах не встречается, но ее дальнейшее изучение и понимание механизмов как самой экспрессии с противоположной цепи ДНК, так и механизмов работы ПАТ *SERPINA1* могло бы иметь важное прикладное значение, а именно при разработке новых подходов к терапии дефицита ААТ. Забегая вперед, отметим, что и данные по гену *SERPINA1*, собранные нами из общедоступных баз данных FANTOM5 CAGE (Cap Analysis of Gene Expression), также свидетельствуют в пользу экспрессии с противоположной цепи гена, в том числе в локусе пятого экзона. По данным FANTOM5 CAGE старты транскрипции антисмысловых РНК обнаруживаются также во втором и третьем экзонах - их роль обсуждается в разделе 4.4. В пятом экзоне *SERPINA1* обнаруживаются сайты связывания с такими мРНК-стабилизирующими белками, как HuR, PTBP3 и IGF2BP2 (<http://rbpmap.technion.ac.il/>) [367], - не исключено, что его ПАТ обладает мРНК-защитной функцией.

На основании результатов, полученных в предыдущих разделах, мы выбрали две клеточные линии для дальнейшего изучения: линию гепатоцеллюлярной карциномы HepG2 как наиболее перспективный объект для выявления структуры и изучения потенциала к трансляции коротких смысловых транскриптов, содержащих экзон 5, и линию опухоли простаты DU145, в которой наблюдается существенная *de novo* экспрессия гена на уровне мРНК и максимальный выход белкового продукта среди линий не печеночного происхождения.

Две другие клеточные линии рака предстательной железы, PC3 и LNCaP, которые происходят из метастазов в кость и лимфатический узел, соответственно, демонстрируют незначительную экспрессию гена *SERPINA1*. Согласно данным, предоставленным Атласом Белков Человека (The Human Protein Atlas) [368], клетки Du145, действительно, проявляют максимальную экспрессию гена *SERPINA1* среди семи различных клеточных линий опухоли предстательной железы (<https://www.proteinatlas.org/ENSG00000197249-SERPINA1/cell+line>).

Несмотря на то, что здоровая ткань предстательной железы не выделяется по уровню экспрессии гена *SERPINA1* [122], есть данные об увеличении уровня ААТ в сыворотке крови пациентов с раком предстательной железы [369, 370, 280, 32], и опухолевые клетки могут вносить свой вклад. Поскольку прогрессирование рака предстательной железы тесно связано с повышением уровня ПСА, сериновой протеазы, можно предположить, что увеличение синтеза ААТ трансформированным эпителием предстательной железы служит для балансировки повышения уровня ПСА. ААТ, действительно, проявляет

ингибирующую специфичность по отношению к ПСА при высоком молярном избытке первого и отмечается положительная корреляция между содержанием ПСА и ААТ [370; 371]. Некоторые метастатические клетки рака предстательной железы, такие как DU145, хоть и утратили экспрессию ПСА [372], могли сохранить новоприобретенную экспрессию *SERPINA1* посредством механизмов эпигенетической памяти [373]. Это предполагает, что экспрессия *SERPINA1* может быть выгодна для клеток DU145. Поскольку клетки DU145 происходят из метастазов в мозг - явление, нехарактерное для рака простаты и насчитывающее всего 0,16% случаев метастазирования при этом заболевании [374], продукция ААТ могла бы быть для них преимущественной для преодоления гемато-энцефалического барьера. Экспрессия серпинов (нейросерпин/*SERPINI1* и ингибитор-2 активатора плазминогена/*SERPINB2*), ингибирующих активатор плазминогена, в клетках рака легкого и молочной железы способствует их метастазированию в мозг за счет ингибирования образования плазмина, который предотвращает метастазирование в мозг [375]. Наши данные свидетельствуют о том, что подобная промотирующая активность серпинов может распространяться и на другие виды рака. ААТ проявляет ингибирующую специфичность к плазмину [142]. Кроме того, известно, что С-концевой пептид ААТ, СААР48, нарушает структурную целостность эндотелия и приводит к потере функции эндотелиального барьера, что продемонстрировано снижением экспрессии VE-кадгерина, F-актина и разупорядочиванием примембранной локализации белка плотного контакта зонального окклюдина 1 (ZO-1) в клетках эндотелиоцитов в модели “печень на чипе” [322]. *SERPINA1* был найден среди восьми генов, уровень экспрессии которых снижается после обработки DU145 антителами против IL-6 [376]. Поскольку клетки DU145 экспрессируют IL-6 и компоненты его рецептора [377], IL-6 - основной активатор генов, кодирующих белки острой фазы в печени [116] - может действовать как аутокринный положительный регулятор в клетках DU145.

4.3 Идентификация структуры 3'- и 5'-концов длинных и коротких транскриптов

4.3.1 Идентификация структуры 3'-концов транскриптов методом 3'-RACE

Методом 3'-RACE с использованием прямых ген-специфических праймеров из пятого экзона и обратного праймера из 3'-адаптерной области получили три фрагмента, содержащих области альтернативного полиаденилирования (Alternative PolyAdenylation, APA) APA1, APA2 и APA3 (рисунок 4.6 А, Б). Во время выполнения этой части работы в UCSC геномном браузере в структуре аннотированных транскриптов был представлен 3'-НТО, соответствующий полиаденилированию по APA3 и состоящий из 1702 н.о. Сигнал к полиаденилированию (PAS) по этому сайту расположен в координатах 1678-1683 н.о. ниже стоп-кодона и представлен консенсусной последовательностью ААТААА [378].

Полиаденилирование по АРА2 осуществляется в положении 87 н.о. ниже стоп-кодона и задействует вариант консенсусного PAS - АТТААА, расположенного в координатах 56-61 н.о. ниже стоп-кодона. Несмотря на то, что АТТААА следующий по частоте реализации после консенсусного PAS (реже всего в 2,4 раза [378]), альтернативное полиаденилирование *SERPINA1* прицельно изучалось сравнительно недавно [38, 122] и в настоящее время (благодаря данным, полученным при секвенировании транскриптомов нормальных тканей человека в рамках проекта Genotype Tissue Expression Project (GTEx) [379]) в структуре транскриптов аннотируются оба АРА. Полиаденилирование по АРА1 происходит в положении 5 н.о. выше стоп-кодона и поэтому приводит к образованию мРНК без стоп-кодона (рисунок 4.6, В), а наиболее подходящий PAS (ААСААА, встречается в 11 раз реже консенсусного [378]) расположен в координатах 56-51 ниже стоп-кодона. Поскольку трансляция поли(А) хвоста в полилизин приводит к остановке рибосомы и в результате такие мРНК подвергаются деградации [380], можно утверждать, что *функциональным* проксимальным сайтом полиаденилирования является сайт АРА2.

С помощью ПЦР-РВ определили относительную экспрессию мРНК, содержащих участки, предшествующие сайтам АРА1-3 (рисунок 2П). Аналогично алгоритму подсчета экспрессии длинных и коротких транскриптов в отдельности, мы определили относительную экспрессию транскриптов, в которых реализуются проксимальный (АРА2) и дистальный (АРА3) сайты полиаденилирования (рисунок 4.6, Г). Уровень экспрессии участков, выбранных для АРА1 и АРА2, практически совпадает, следовательно, транскриптов, полиаденилированных по АРА1, крайне мало. По данным ПЦР-РВ в клетках DU145 и HepG2 преимущественно реализуется проксимальный сайт АРА2: экспрессия таких транскриптов выше, чем АРА3-транскриптов в среднем в 41 и 35 раз, соответственно. Длинный транскрипт, соответствующий АРА3, отчетливо детектируется в нозерн-гибридизации только в РНК HepG2 и тканей печени (рисунок 4,3, А, Б) (размер около 3300 н.о.), а в 3'-RACE - только в кДНК ткани печени (рисунок 4.6, А). Последнее может быть связано с тем, что в ткани печени доля длинных 3'-НТО выше таковой в клеточной линии HepG2, в которой амплифицируются только короткие продукты.

Укорочение 3'-НТО за счет выбора 5'-проксимального АРА ассоциировано с пролиферацией и клеточной трансформацией, наблюдается на ранних стадиях эмбриогенеза, в процессе репрограммирования в индуцированные плюрипотентные стволовые клетки, а также в тканях и органах с высокой пролиферативной активностью, в том числе опухолевых [381, 382, 383, 384]. И наоборот, транскрипты на поздних стадиях развития и в дифференцированных неделящихся клетках, как правило, содержат удлиненные 3'-НТО [385, 386]. Выбор проксимальных сайтов АРА приводит к потере сайтов связывания микроРНК в 3'-НТО, позволяя таким образом избежать репрессии трансляции. Наиболее вероятной представляется негативная регуляция, опосредованная miR-940, для которой в 3'-НТО гена *SERPINA1* идентифицировано 4 сайта связывания

[362]. Проксимальный сайт связывания miR-940 локализован в транскриптах, полиаденилированных по сайту APA2, остальные — в транскриптах с более протяженной 3'-НТО (APA3). Авторы полагают, что особенно “активен” дистальный сайт связывания miR-940, поэтому наибольшей репрессии трансляции посредством miR-940 подвержены транскрипты гена *SERPINA1* с более длинной 3'-НТО [362]. Кроме того, недавно обнаружено, что в клетках HepG2 при оверэкспрессии miR-320с (экспрессия которой повышена при легочных заболеваниях) ингибируется экспрессия изучаемого гена [387], и ее сайт связывания содержится только в транскриптах с длинным 3'-НТО.

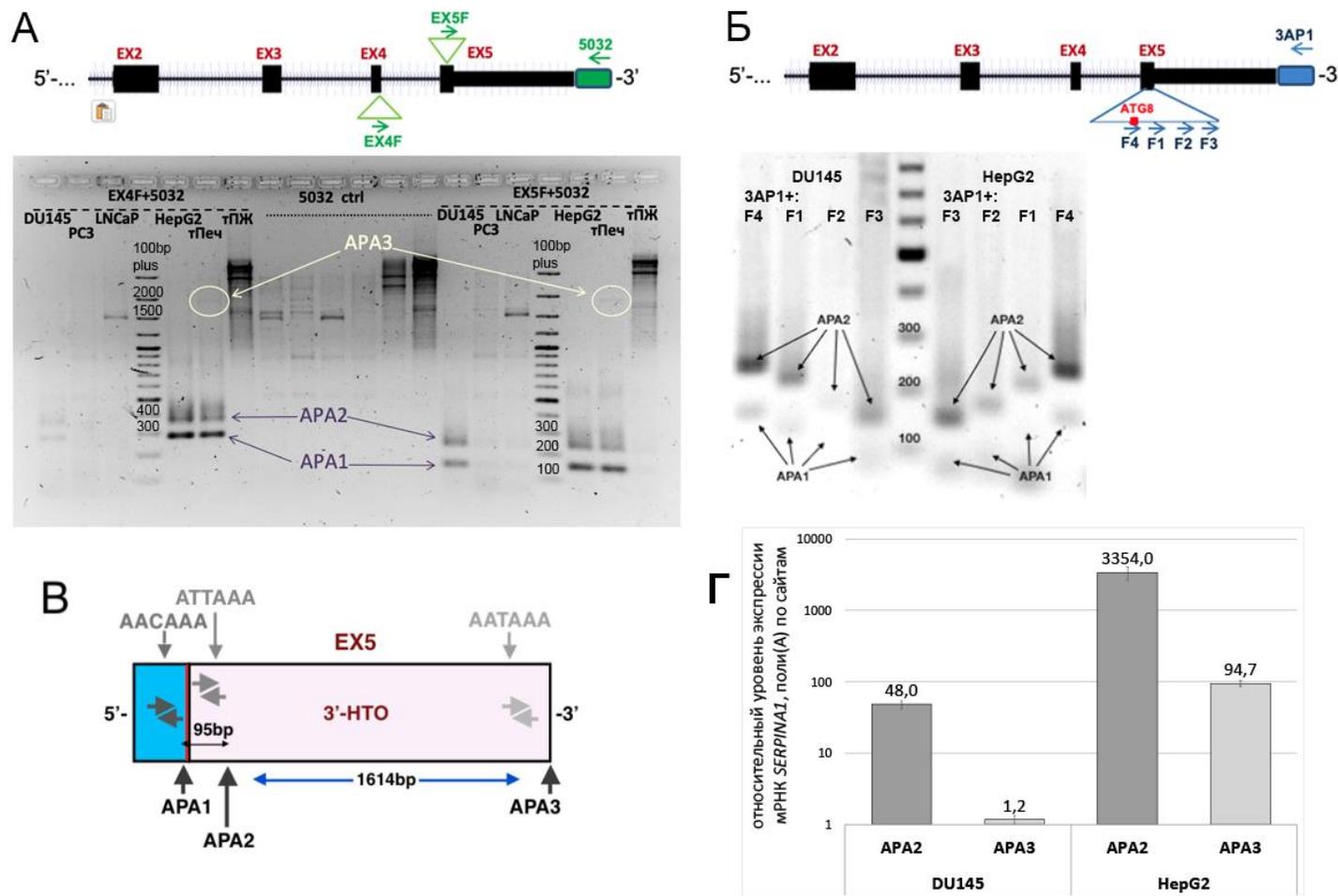


Рисунок 4.6. Анализ структуры 3'-концов транскриптов *SERPINA1* с помощью 3'-RACE. А) 3'-RACE на образцах из клеточных линий (DU145, PC-3, LNCaP и HepG2) и тканей печени (тПеч) и предстательной железы (тПЖ) от Ambion (США). Сверху - Схематическая иллюстрация кодирующих экзонов гена *SERPINA1*; зелеными стрелками отмечено расположение праймеров, использованных для амплификации кДНК; лигированный 3'-адаптер изображен зеленым прямоугольником. В реакции обратной транскрипции использовали адаптер 5065 и в ПЦР соответствующий обратный праймер 5032 (см. раздел 3.6). Снизу - Электрофореграмма полученных продуктов амплификации; пары праймеров соответствуют схеме сверху. Б) 3'-RACE на образцах из клеточных линий DU145 и HepG2; в реакции обратной транскрипции использовали адаптер 3dTA, в

ПЦР - набор прямых и обратного праймера на адаптер (3AP1) (см. раздел 3.6), схематически изображенных сверху. Прямой праймер F4 - 5'-удлинненный вариант праймера EX5F (А). Снизу - Электрофореграмма ПЦР продуктов; пары праймеров соответствуют схеме сверху. В) Схематическая иллюстрация относительного расположения обнаруженных альтернативных сайтов полиаденилирования APA1-3 (отмечены черными вертикальными стрелками, отмечены расстояния между APA) в структуре экзона 5. Прямоугольник с голубой заливкой соответствует кодирующей области экзона 5, красной вертикальной линией отмечен стоп-кодон; прямоугольник со светло-розовой заливкой - 3'-НТО. Г) Относительный уровень экспрессии транскриптов, полиаденилированных по APA2 и APA3, в отдельности. На схеме В) серыми горизонтальными стрелками отмечены пары праймеров, использованные в ПЦР-РВ.

Недавно обнаружили дополнительный механизм регуляции экспрессии *SERPINA1* за счет альтернативного полиаденилирования: дистальная часть 3'-НТО, содержащаяся только в АРА3-транскриптах, содержит сайт посадки РНК-связывающего белка QKI (Quaking homolog), препятствующего экспорту мРНК из ядра в цитоплазму для последующей трансляции [122].

Стоит отметить, что проксимальный сайт АРА2 является основным сайтом полиаденилирования в здоровых тканях/органах взрослого, экспрессирующих *SERPINA1* [122], а именно в тканях печени, крови, легких, тонком кишечнике, селезенке и почке, что может отражать потребность этих органов и тканей в ААТ. Среди перечисленных органов экспрессия транскриптов, полиаденилированных по дистальному сайту АРА3, максимальна в ткани печени [122], что, наиболее вероятно, связано с максимальной экспрессией гена как таковой. Экспрессия укороченного и длинного 3'-НТО *SERPINA1* в составе репортерной конструкции показала отличие по выходу белкового продукта в 50 раз [122]. Также, авторы показали, что с развитием хронической обструктивной болезни легких при наличии в генотипе S и Z-аллелей *SERPINA1* в ткани легких растет доля АРА3-мРНК [122]. В контексте дефицита ААТ выбор сайта АРА3, с одной стороны, усугубляет сам дефицит (как минимум на локальном уровне, - в легких), с другой стороны, потенциально может снижать цитотоксическую нагрузку ААТ-полимеров в гепатоцитах. Таким образом, альтернативное полиаденилирование является важным механизмом регуляции экспрессии гена *SERPINA1* на пост-транскрипционном уровне как в здоровых тканях, так и в трансформированных клетках.

4.3.2 Идентификация структуры 5'-концов транскриптов методом 5'-RACE

Методом 5'-RACE с использованием прямых праймеров на 5'-адаторную область и обратных из экзона 5 гена *SERPINA1* на кДНК из РНК клеточных линий DU145 и HepG2 получили фрагменты, соответствующие длинным и коротким транскриптам (Рисунок 4.7). 5'-RACE определил единственный длинный транскрипт в клетках HepG2 - HLT (HepG2 Long Transcript), являющийся классическим гепато-специфическим транскриптом с укороченным вариантом EX1C [41] и соответствующий 5'-концу аннотированного транскрипта NM_000295.5. По данным полнотранскриптомного анализа в клетках HepG2 экспрессируются все 11 вариантов мРНК *SERPINA1* [44], и по-видимому, в нашем эксперименте более эффективно амплифицировался продукт с наиболее короткой 5'-НТО, ввиду как меньшей его длины, так и большей его представленности в пуле кДНК HepG2.

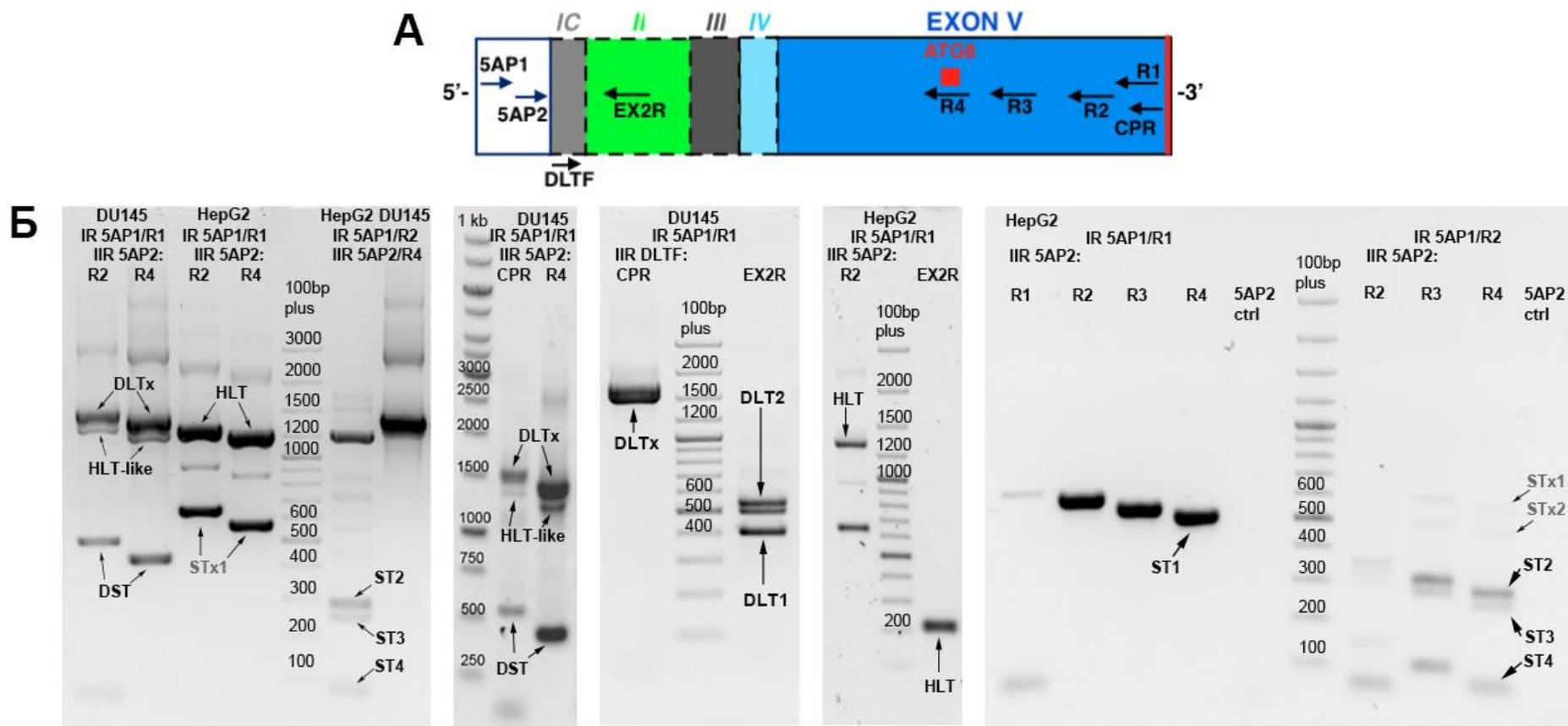


Рисунок 4.7. Анализ структуры 5'-концов транскриптов *SERPINA1* с помощью 5'-RACE. **A**) Схематическая иллюстрация экзонов гена *SERPINA1* (подписаны римскими цифрами) и лигированного 5'-адаптора (прямоугольник без заливки); на схеме указаны названия и относительное расположение используемых в анализе праймеров; красная вертикальная линия - стоп-кодон. **Б**) Сканы гелей с продуктами амплификации во втором раунде вложенной ПЦР. Подписаны ДНК маркеры и праймеры из схемы А), использованные в первом (IR) и во втором (IIR) раундах ПЦР. Наименования продуктов амплификации: D(L/S)T=DU145 Long/Short Transcript; DLTx - смесь продуктов DLT1 и DLT2; HLT=HepG2 Long Transcript; HLT-like - транскрипт, идентичный HLT; ST1-4=Short Transcript 1-4 - основные короткие продукты HepG2; STx - минорные короткие продукты HepG2, полученные в режиме амплификации кДНК коротких транскриптов.

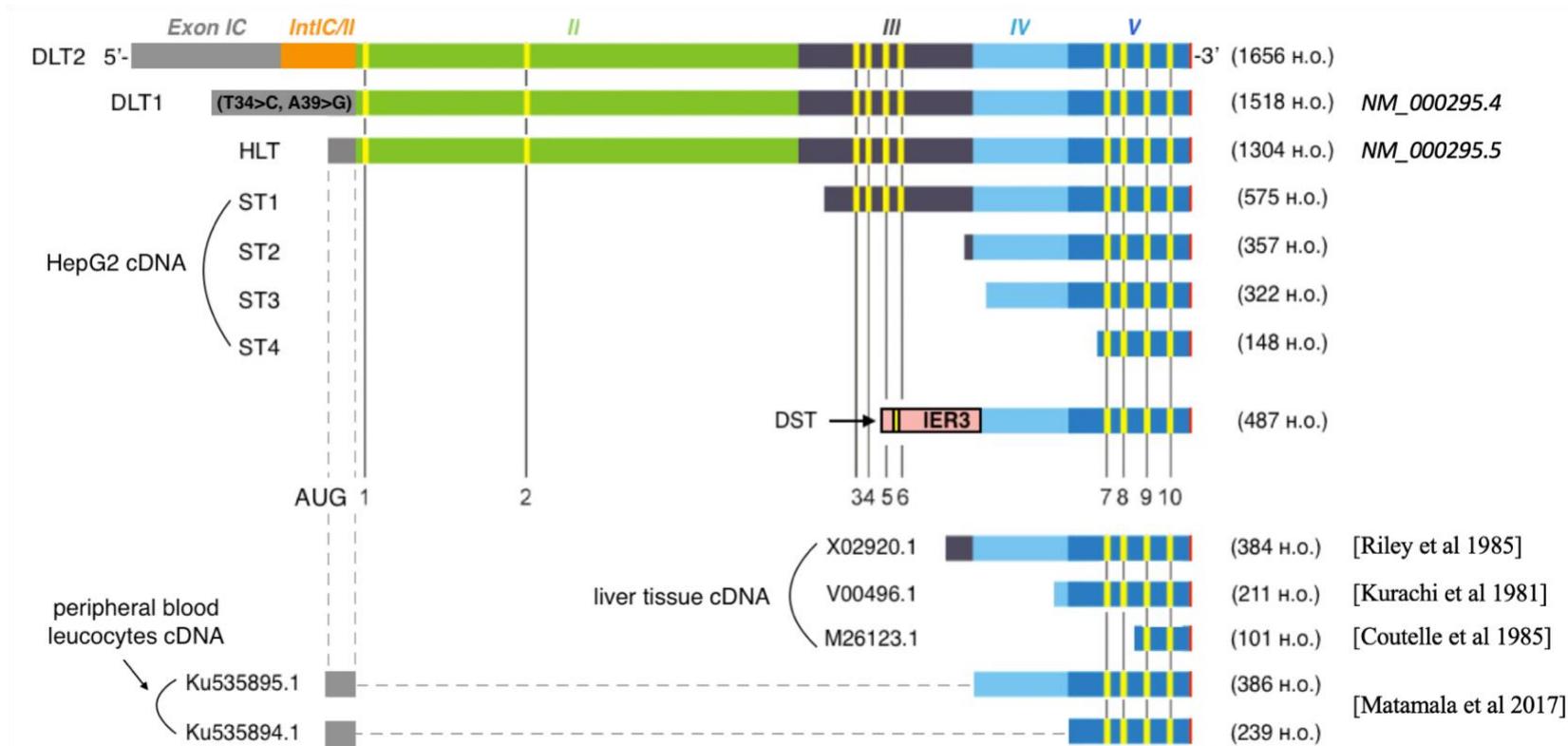


Рисунок 4.8. Схематические структуры 5'-концов транскриптов гена *SERPINA1* в клеточных линиях DU145 и HepG2. D(L/S)T=DU145 Long/Short Transcript; HLT=HepG2 Long Transcript; ST1-4=Short Transcript 1-4 из HepG2. Снизу приведены схематические структуры коротких транскриптов, встречающихся в других работах: X02920 [353]; V00496 [354]; M26123 [355]; Ku535895.1 и Ku535894.1 [344]. На схемах отмечены AUG кодоны желтыми вертикальными линиями и стоп-кодон красной вертикальной линией. Справа отмечены длины транскриптов до стоп-кодона включительно.

В отличие от печеночных клеток, основной длинный транскрипт DU145 (DLT, DU145 Long Transcript) содержит 5'-удлиненный вариант EX1C (с охватом всего интрона IV/IC) и соответствует 5'-НТО транскрипта NM_000295.4 (с заменами T34>C, A39>G). Стоит отметить, что NM_000295.4 аннотируется как вариант NM_000295.5 и не входит в “набор” из 11 транскриптов, указанных в работе [44]. Этот транскрипт был обозначен как DLT1 (рисунок 4.7, Б; рисунок 4.8). Примечательно, что недавно изоформа *SERPINA1* NM_000295.4 была обнаружена как одна из наиболее экспрессируемых вариантов мРНК в легких [121], - органе, который, как и печень, критически зависит от нормального функционирования ААТ [388]. Она составляет порядка 23% от общего количества транскриптов *SERPINA1*, экспрессируемых в легких [121].

Позже, при клонировании кДНК DLT1, амплификация с прямым праймером на самый 5'-конец транскрипта и обратным из пятого экзона выявила, что этот продукт представлен смесью продуктов (рисунок 4.7, третий скан слева). Взяв обратный праймер из второго экзона, мы показали, что в смеси как минимум три продукта, из которых выделили и секвенировали два - DLT1, обозначенный выше, и DLT2. Последний, по-видимому, является минорным, он не аннотирован и отличается частичным удержанием интрона, а именно интронных 138 н.о. смежных с 5'-концом первого кодирующего экзона 2. Мы его обозначили как DLT2 (рисунок 4.7, 4.8). Вероятность существования такого транскрипта подтверждается *in silico* предсказанием точного положения альтернативного акцептора сплайсинга (https://www.fruitfly.org/cgi-bin/seq_tools/splice.pl). Кроме того, мы провели сравнение этого интронного локуса с тегами экспрессируемых последовательностей (EST, expressed sequence tag) и обнаружили его экспрессию в фетальной печени (GenBank ID BX506843.1) и нормальной печеночной ткани (GenBank ID AV658834.1, [389]). Границы EX1C в DLT2 также соответствуют NM_000295.4 без замен, обнаруженных в DLT1. Таким образом, DLT1 и DLT2 могут происходить от разных аллелей *SERPINA1*. Клетки DU145 также производят транскрипты, идентичные HLT, но в меньших количествах (рисунок 4.7). Схематическая диаграмма этих транскриптов представлена на рисунке 4.8.

Продукт 5'-RACE, соответствующий короткому транскрипту DU145 (DST, DU145 Short Transcript), является химерой: с 5'-конца соответствует последовательности 5'-НТО и неполной кодирующей области мРНК гена *IER3* (ген раннего ответа 3, immediate early response 3), а с 3'-конца - участок неполного EX4 и EX5 *SERPINA1* (рисунок 4.7 и 4.8). Интересно, что между двумя составляющими нет сдвига рамки считывания и трансляция такого короткого транскрипта привела бы к образованию химерного белка, N-конец которого представлен последовательностью N-конца IEX1 - белкового продукта гена *IER3*, а С-конец - С-концом ААТ. Белок IEX1 состоит из 156 а.о., из которых в химерном продукте, размер которого случайно совпадает с размером IEX1, остаются N-концевые 47 а.о., через кодируемый за счет стыка остаток пролина соединенные со 108 С-

концевыми а.о. ААТ. Известно, что IEX1 поддерживает активность F1F0-АТФазы, направляя ее ингибитор IF1 на деградацию, снижает продукцию АФК, и тем самым регулирует тип метаболизма клетки, поскольку в IEX1-дефицитных клетках происходит переключение с окислительного фосфорилирования на гликолиз, в результате чего повышается продукция лактата (эффект Варбурга) [390]. IEX1 обладает проканцерогенными свойствами при опухоли молочной и поджелудочной желез и мочевого пузыря [391, 392, 393]. Длинная изоформа IEX1, которая содержит дополнительные 37 а.о. вдобавок к 156 а.о., защищает клетки от Fas- и TNF α - индуцированного апоптоза [394]. Химерный транскрипт может являться продуктом транссплайсинга транскриптов обоих генов. Поскольку он не детектируется в нозерн-гибридизации в клетках DU145 в явном виде, его базовая экспрессия (в отсутствие стимулов) ожидается на низком уровне. Химерный продукт амплифицировался только на матрицах в 5'-RACE и не выявлялся при попытке амплифицировать его из других препаратов кДНК DU145, поэтому мы исключили его из дальнейшего анализа.

В отличие от DU145, в клеточной линии HepG2 наблюдается несколько коротких продуктов амплификации (рисунок 4.7, ST, Short Transcript). Мы исключили из анализа минорные продукты, полученные в режиме амплификации кДНК коротких транскриптов (рисунок 4.7, STx, скан справа). Секвенирование основных коротких ПЦР-продуктов после второго раунда вложенной ПЦР выявило несколько вариантов 5'-концов коротких транскриптов HepG2: короткий транскрипт ST1 состоит из неполного EX3 (начиная с 37го н.о. экзона)-EX4-EX5; ST2 - EX3 (с 255го н.о.)-EX4-EX5; ST3 - EX4 (с 19го н.о.)-EX5; ST4 - EX5 (с 45го н.о.). Поскольку 5'-RACE проводили с предварительной элиминацией частично деградированных мРНК, учитывали только кэпированные транскрипты и использовали матрицы кДНК, полученные обратной транскрипцией с обратным праймером на поли(А)-хвост, то короткие транскрипты кэпированы и полиаденилированы.

Схематические структуры коротких транскриптов HepG2 представлены на рисунке 4.8. Помимо коротких транскриптов, идентифицированных нами, снизу представлены короткие транскрипты, встречающиеся в других работах ранее: V00496.1 - неполный EX4/EX5 [354]; X02920.1 - неполный EX3/EX4/EX5 [353]; M26123.1 - неполный EX5 [355], - найденных в библиотеках кДНК из печени человека. Эти транскрипты так же, как обнаруженные нами транскрипты HepG2, начинаются с внутренних участков кодирующих экзонов и, по-видимому, полиаденилированы по сайту APA2, поскольку в полной последовательности содержат соответствующий PAS (V00496.1, X02920.1) либо короткий участок за стоп-кодом (M26123.1), что позволяет нам предположить, что обнаруженные нами короткие транскрипты HepG2 имеют аналогичный паттерн полиаденилирования (рисунок 4.3, А). В лейкоцитах периферической крови стандартным ПЦР путем амплификации известных последовательностей из экзона 1C/1A и экзона 5 гена *SERPINA1* обнаружены транскрипты: KU535895.1 - неполный EX1C/EX4/EX5, KU535894.1 -

неполный EX1C/EX5, KU755450.1 - неполный EX1A/неполный EX5, KU755451.1 - неполный EX1A/EX4/EX5 [344], - на рисунке 4.8 представлены только варианты с EX1C. Поскольку в состав лейкоцитарных транскриптов входят некодирующие экзоны 1A/1C, их механизм образования - альтернативный сплайсинг. В случае с печеночными короткими транскриптами, начинающимися с внутренних участков кодирующих экзонов, механизм образования не так очевиден (транскрипция с альтернативных внутренних промоторов? - рисунок 4.1, А). Ниже мы попробуем его выявить, основываясь на общедоступных данных.

4.4 Досье *SERPINA1*, извлеченное из данных CAGE и Chip-Seq, выявляет возможный механизм образования коротких транскриптов

Поскольку мы определили структуры 5'-концов четырех вариантов коротких транскриптов HepG2 ST1-4, необычно начинающихся внутри кодирующих экзонов, мы решили оценить возможное происхождение этих транскриптов. Мы сопоставили их 5'-концы с известными пиками CAGE (Cap Analysis of Gene Expression), отмеченными в первой и второй фазах проекта FANTOM5 [395]. Три из четырех ST (ST1, 3 и 4), а также HLT подходят к известным пикам CAGE, однонаправленным к гену *SERPINA1*, по их 5'-концам (рисунок 4.9, А), в то время как 5'-конец ST2 отмечается на расстоянии 9 п.н. от ближайшего пика p20@SERPINA1, который мы далее рассматриваем как связанный с этим пиком. В клетках HepG2 ассоциированные с ST пики находятся в диапазоне экспрессии 3-6 TPM (Tags Per Million), что значительно ниже основного пика p1@SERPINA1 (16365 TPM), соответствующего области начала транскрипции с гепатоцит-специфического промотора и включающего 5'-конец HLT. Эти данные согласуются с нашими экспериментальными данными по нозерн гибридизации на суммарной РНК HepG2 (рисунок 4.3, А).

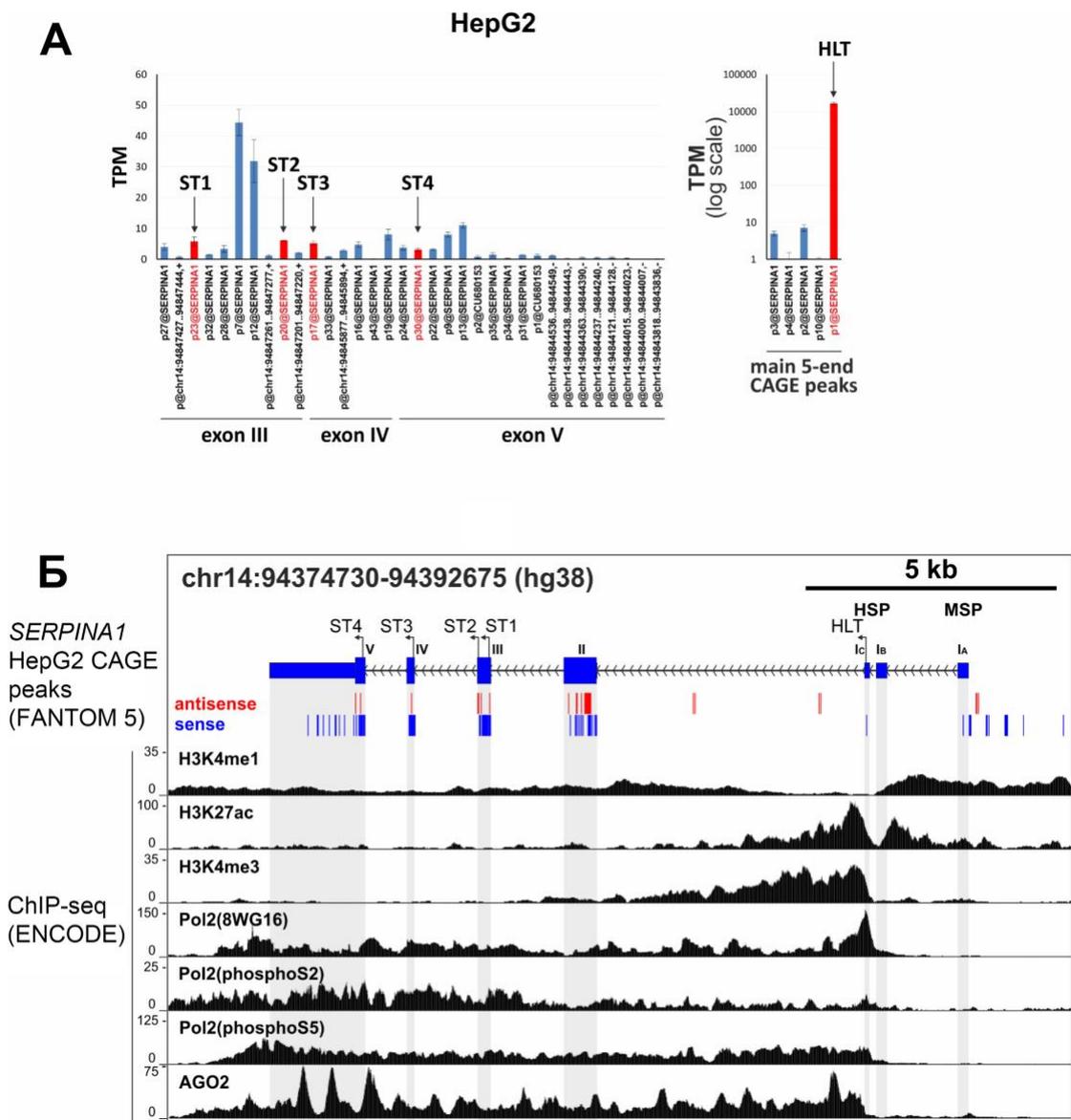


Рисунок 4.9. Пики CAGE FANTOM5 в локусе *SERPINA1*. А) Уровни экспрессии пиков CAGE, расположенных в экзонах 3, 4 и 5 гена *SERPINA1*. Пики, перекрывающиеся или прилегающие (ST2) к картированным в нашем исследовании 5'-концевым продуктам 5'-RACE клеток HepG2, отмечены красным. TPM - теги на миллион. Данные представлены в виде среднего \pm стандартное отклонение от трех повторов. Б) Распределение сигналов ChIP-seq меток хроматина энхансеров/промоторов, связанного с хроматином AGO2 и пиками CAGE в гене *SERPINA1* в клетках HepG2. MSP - миелоид-специфический промотор, HSP - гепатоцит-специфический промотор.

Возможным источником усеченных транскриптов *SERPINA1* могут быть активные цис-регуляторные элементы внутри гена, управляющие инициацией транскрипции. Наряду с промоторами, энхансеры также могут привлекать РНК-полимеразу II и начинать транскрипцию энхансерных РНК [396]. Внутригенные цис-регуляторные элементы, как правило, расположены в интронах, но экзонные энхансеры также изредка распознаются и могут потенциально служить источником 5'-концов экзонных транскриптов [397, 398, 399]. Мы проанализировали сигналы ENCODE ChIP-seq внутри локуса гена *SERPINA1* в клетках HepG2, полученные для модификаций гистонов H3K27ac (метки хроматина энхансеров/промоторов), H3K4me1 (энхансеры) и H3K4me3 (промоторы) и для РНК-полимеразы II. Результаты на рисунке 4.9, Б показывают отсутствие внутригенных

областей с выраженным локальным обогащением сигналами ChIP-seq от этих меток, следовательно - отсутствие явных цис-регуляторных элементов.

Мы проанализировали позиции пиков FANTOM5 CAGE в гене *SERPINA1* и прилегающих межгенных областей и обнаружили высокое содержание пиков, разбросанных по всем кодирующим экзонам (рисунок 4.9, Б), что указывает на высокую вариабельность коротких РНК *SERPINA1*. Обогащение экзонов, но не интронов, пиками CAGE, соориентированными с геном, поддерживает гипотезу посттранскрипционного происхождения коротких транскриптов *SERPINA1*, которые могут возникать путем энзиматического расщепления сплайсированного длинного транскрипта-предшественника с последующим вторичным кэпированием вновь образованных 5'-концов [400, 401, 402]. Мы предположили, что механизм посттранскрипционного процессинга длинного транскрипта-предшественника может основываться на расщеплении дуплекса смысловой и антисмысловой РНК - хорошо изученном клеточном процессе [403]. Действительно, мы также наблюдали обогащение экзонов пиками CAGE, ориентированными в противоположную сторону гена (соответствующие 5'-концам антисмысловых транскриптов, рисунок 4.9, Б), возможно, отражающими, что расщепление РНК-РНК дуплекса эндонуклеазами, такими как AGO2 [404], может способствовать формированию коротких транскриптов. Интересно, что AGO2 взаимодействует с основным транскрипционным аппаратом в хроматине [405] и обнаруживается методом ChIP-seq в теле гена *SERPINA1* в клетках HepG2 (рисунок 4.9, Б). 5'-концы антисмысловых РНК обнаруживаются и в экзоне 5 (рисунок 4.9, Б), что согласуется с нашими предыдущими данными (см. раздел 4.2.3).

Гипотеза посттранскрипционного происхождения коротких транскриптов *SERPINA1* предполагает, что их обилие напрямую зависит от представленности вариантов полноразмерного транскрипта *SERPINA1* и, следовательно, от активности основных пиков CAGE, экспрессированных с любого из известных промоторов гена *SERPINA1*. Действительно, было показано, что существует высокая корреляция между "активностями" экзонных пиков и 5'-концевых пиков генов [406]. Мы проверили профили экспрессии пиков CAGE *SERPINA1* на сходство между 889 образцами первой фазы проекта FANTOM5 (рисунок 4.10, А). Матрица сходства внутригенной, в основном экзонной, экспрессии пиков (включая соответствующие 5'-концам коротких транскриптов, найденных в нашем исследовании) показывает их хорошую корреляцию между собой и с основным гепатоцит-специфическим пиком p1@SERPINA1, образующим заметный кластер на матрице (рисунок 4.10, А, Б).

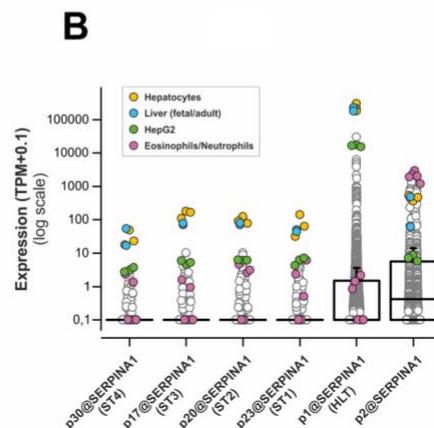
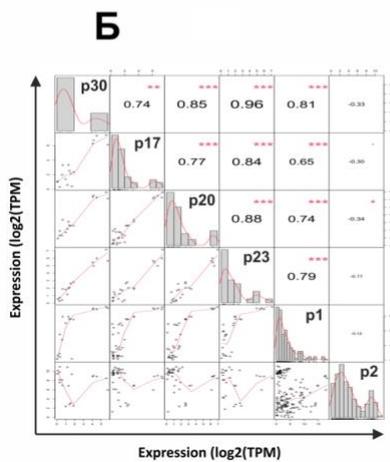
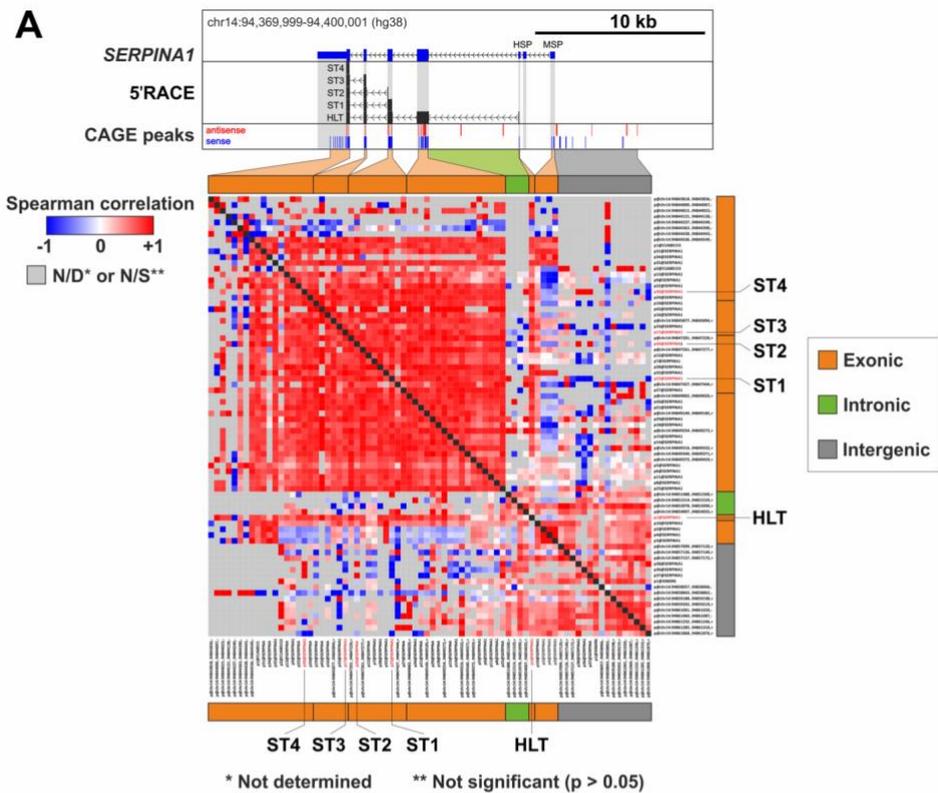


Рисунок 4.10. Экспрессия внутригенных пиков CAGE *SERPINA1* коррелирует с активностью гепатоспецифического промотора *SERPINA1*. **А**) Матрица корреляции, полученная в результате попарного сравнения значений экспрессии пиков CAGE среди 889 образцов (FANTOM5 фаза I). Матрица корреляции была визуализирована с помощью программного обеспечения morpheus (<https://software.broadinstitute.org/morpheus/>). **Б**) Попарные сравнения профилей экспрессии пиков CAGE, перекрывающихся с идентифицированными 5'-концами продуктов 5'-RACE. **В**) Точечная и ящичковая диаграммы, показывающие распределение уровней экспрессии для пиков CAGE, перекрывающихся с идентифицированными 5'-концами продуктов 5'-RACE среди 889 образцов (FANTOM5 фаза I). MSP - миелоид-специфический промотор, HSP - гепатоцит-специфический промотор.

Эти внутригенные пики оказались высокоспецифичными, и их экспрессия различима в клетках с чрезвычайно активным гепатоцит-специфическим промотором - эмбриональной и взрослой ткани печени, а также первичных гепатоцитах (рисунок 4.10, В). В то же время, "активности" тех же пиков не имеют значительной положительной

корреляции с более слабым вторым пиком p2@SERPINA1 (рисунок 4.10, Б), соответствующим экспрессии с альтернативного тканеспецифического промотора и имеющим наибольшую экспрессию в клетках врожденной иммунной системы, таких как эозинофилы и нейтрофилы (рисунок 4.10, В). В целом, приведенные данные позволяют предположить, что основным источником усеченных внутригенных транскриптов может быть посттранскрипционное расщепление длинных транскриптов *SERPINA1*, синтез которых направлен гепатоцит-специфическим промотором. Данное предположение, несомненно, требует дальнейшего изучения.

4.5 Изучение трансляции индивидуальных изоформ длинной мРНК *SERPINA1* в живых клетках *in vitro*: поиск продуктов трансляции и идентификация альтернативных ОРС

4.5.1 С длинных транскриптов *SERPINA1* синтезируется предполагаемая несекреторная изоформа ААТ

Повышение разнообразия белкового состава необходимо для удовлетворения потребностей клетки в условиях ограниченного размера генома при изменении условий среды. Инициация трансляции на альтернативных стартовых кодонах внутри рамки считывания представляет собой уникальный и энергосберегающий механизм, который позволяет получать различные изоформы белка, обладающие собственными свойствами, из одной и той же мРНК [351]. Этот механизм распространен у эукариот [407, 342] и продолжает обнаруживаться у транскриптов генов, которым ранее не присваивался [408, 409]. Не исключено, что среди этих генов может быть и изучаемый нами. Мы предположили, что длинные транскрипты *SERPINA1* могут транслироваться с альтернативного стартового кодона, в конечном итоге приводя либо к маскированию сигнального пептида (инициация трансляции выше и в рамке с главным стартовым кодоном, приводящей к N-концевому довеску), либо к его исключению (инициация трансляции ниже и в рамке с главным старт-кодоном). В результате будет синтезирована внутриклеточная изоформа ААТ.

Для проверки гипотезы о существовании альтернативных продуктов трансляции мы клонировали кДНК *SERPINA1*, соответствующие длинным транскриптам DLT1, DLT2 и HLT (рисунок 4.8), в вектор pEGFP-N1 (рисунок 4.11, А). С плазмид транскрибировалась химерная РНК, состоящая из идентифицированных транскриптов *SERPINA1* с 5'-конца (охватывая всю область до стоп-кодона), за которыми следовала последовательность, кодирующая eGFP, в рамке с *SERPINA1*. Для того, чтобы исключить эффективную трансляцию *EGFP* в отдельности, но выявить возможные альтернативные продукты *SERPINA1*, мы внесли мутацию в вектор pEGFP-N1, удалив старт-кодон ATG без сдвига рамки считывания, чтобы получить мутантную плазмиду pEGFP-N1(ATG-) (рисунок 4.11, А). Поскольку клетки DU145 производят все три длинные мРНК *SERPINA1*, мы в

основном проводили наше исследование на этих клетках. После трансфекции секретомы и сортированные флуоресцентные клетки анализировали на предмет eGFP+ продуктов с помощью иммуноблоттинга (рисунок 4.11, Б; полные сканы приведены на рисунке 3П). Для каждой конструкции в клеточных лизатах обнаружили несколько полос, соответствующих слитым белкам ААТ-eGFP в районе 69-88 кДа, в то время как одна полоса с более низкой электрофоретической подвижностью (99 кДа), соответствующая полностью гликозилированному ААТ (полностью созревшему в аппарате Гольджи), обнаруживалась в секретах после трансфекции HLT и DLT1. Отсутствие такой же полосы в секрете DLT2 может быть обусловлено более низкой эффективностью трансфекции этой плазмиды (которую лучше скорректировать как *пропорцию трансфицированных клеток, способных к трансляции DLT2*, ввиду клеточной гетерогенности, поскольку условия трансфекции были идентичными). Эти результаты предполагают, что помимо секреторного ААТ, могут также производиться и внутриклеточные протеоформы.

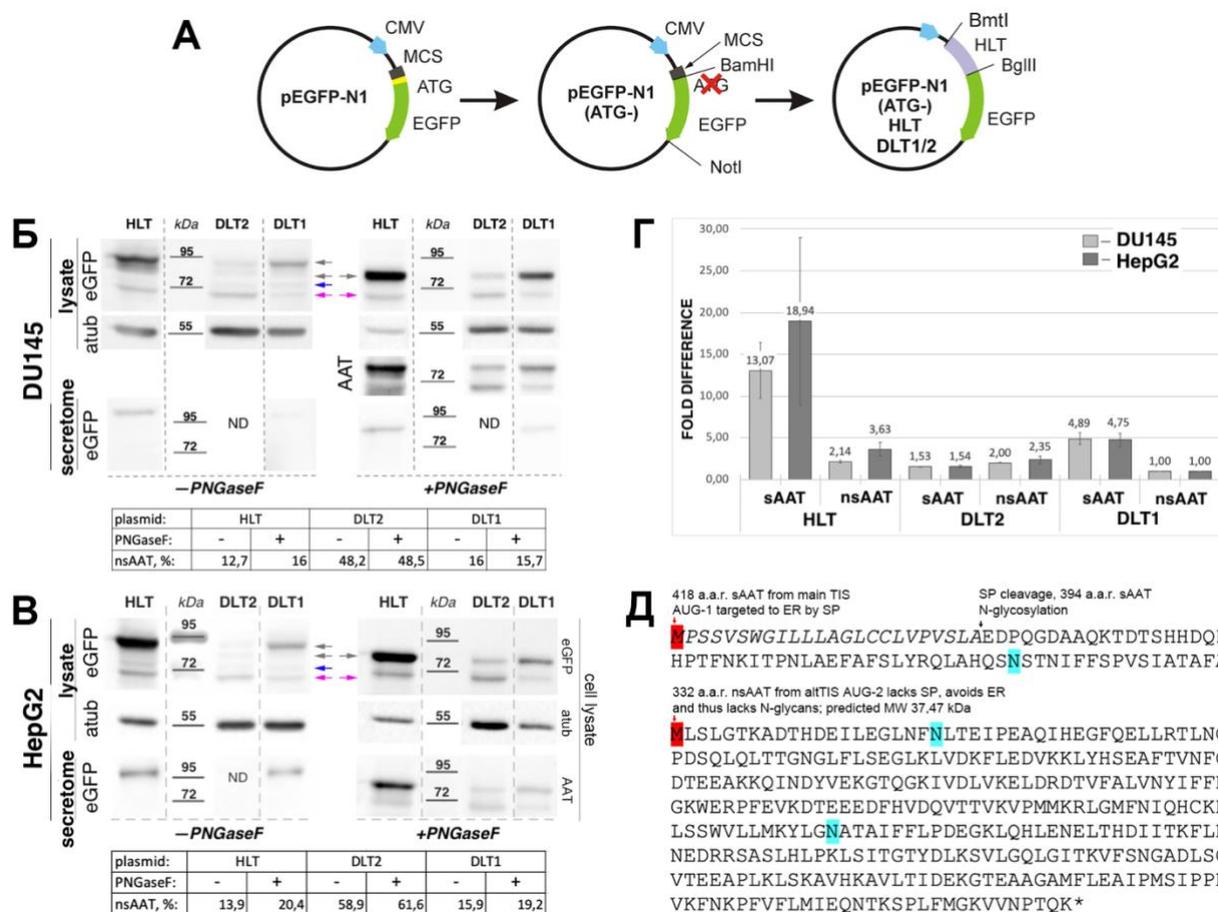


Рисунок 4.11. С длинных транскриптов *SERPINA1* синтезируется предполагаемая несекреторная негликозилированная изоформа ААТ (nsAAT). **А**) Схема клонирования кДНК в вектор pEGFP-N1(ATG-). **Б**) Иммуноблоттинг лизатов и секретомов трансфицированных клеток DU145 с и без обработки PNGaseF указывает на трансляцию негликозилированной изоформы nsAAT с альтернативного сайта инициации трансляции (altTIS) (Д). Серые стрелки - секреторный ААТ (sAAT), малиновая стрелка - предполагаемая изоформа nsAAT; синяя стрелка - слабо воспроизводимая полоса 75 кДа. Мембрану с образцами

+PNGaseF стрипировали и перепроверили антителами против полноразмерного ААТ для демонстрации продуктов в рамке *SERPINA1*. Доля интенсивности сигналов от nsAAТ-eGFP от общего сигнала eGFP до и после обработки PNGaseF показаны под изображениями сканов в %. **В**) Иммуноблоттинг после трансфекции HepG2. **Г**) Денситометрический анализ экспрессии sAAТ и nsAAТ с конструкций в клетках DU145 и HepG2; так как интенсивность альфа-тубулина зависит от конструкции, сигналы были нормированы к общей интенсивности Ponceau S в соответствующих дорожках, затем к полосе с минимальной экспрессией (nsAAТ с DLT1 для каждой клеточной линии); анализировались изображения до и после обработки PNGaseF; данные представлены как среднее \pm стандартное отклонение. Двухфакторный дисперсионный анализ определил различие в содержании белкового продукта между экспрессируемыми транскриптами ($p = 0,0003$), но не типами клеток ($p = 0,44$). **Д**) Полипептидная последовательность ААТ (UniProtKB ID P01009). Красная заливка указывает на остатки метионина, соответствующие секреторным и несекреторным изоформам, инициированным от основного TIS и предполагаемого altTIS, соответственно; к каждой изоформе приведена краткая пометка, обозначенная красной стрелкой; черная стрелка показывает место расщепления сигнального пептида (SP); ER - эндоплазматический ретикулум; GA - аппарат Гольджи; MW - молекулярная масса; голубая заливка указывает на остатки аспарагина, подвергающиеся гликозилированию в sAAТ; курсивом - сигнальный пептид секреции (SP).

Являются ли внутриклеточные протеоформы продуктами протеолитического расщепления зрелого секреторного ААТ, который затем эффективно эндоцитируется, или результатом трансляции с альтернативных ОРС, и, следовательно, их можно отнести к рангу "изоформы" [351], и сколько "изоформ по первичной структуре" синтезируется? Чтобы ответить на эти вопросы, мы дегликозилировали секреты и экстракты сортированных клеток с помощью фермента PNGaseF, который отщепляет почти все типы N-гликанов [410], поскольку секреторный ААТ подвергается N-гликозилированию в положениях 46, 83 и 247 в ЭПР [124, 125, 126]. После обработки PNGaseF, полосы eGFP+ продуктов размером 81-88 кДа из клеточных лизатов объединились в одну полосу 81 кДа (рисунок 4.11, Б), что указывает на различия их полисахаридного состава, а не первичной структуры, и продукты представляют собой *еще не* гликозилированный ААТ, промежуточные и зрелые гликоформы секреторного белка. Как и ожидалось, зрелый 99 кДа ААТ из секрета "ускорился" до 81 кДа после обработки ферментом, аналогично тому, что наблюдалось в клеточных лизатах. Полоса 75 кДа (рисунок 4.11, Б, синяя стрелка) плохо воспроизводилась после обработки PNGaseF, поэтому мы не рассматриваем ее как целевой несекреторный продукт. В отличие от нее, полоса в 69 кДа, соответствующая главному внутриклеточному продукту (рисунок 4.11, Б, малиновая стрелка), наблюдалась консистентно и, по-видимому, оставалась неизменной под воздействием фермента, что подтверждается незначительным изменением доли его сигнала от общего сигнала от eGFP в образцах с добавлением фермента или без него (рисунок 4.11, Б, таблица под сканом). Отсутствие полосы одинаковой интенсивности чуть ниже 69 кДа после дегликозилирования дополнительно подтверждает идентичность полосы 69 кДа до и после обработки (рисунок 3П).

Чтобы определить, является ли синтез внутриклеточной протеоформы ААТ специфическим для линии клеток эпителия простаты или потенциально указывает на более широкое биологическое явление (на данном этапе независимо от его ассоциации с

канцерогенезом), мы провели те же эксперименты по трансфекции плазмидных конструкций в линию HepG2, которая демонстрирует профиль секретируемых белков, очень схожий с таковым у первичных гепатоцитов человека [411] - основным источником ААТ в организме. HepG2 не только производят тот же 69 кДа слитой белок, но также демонстрирует аналогичное “поведение” трансфицированных НLT- и DLT1/2-содержащих конструкций, близко похожее на результаты, наблюдаемые в клетках DU145 (рисунок 4.11, В, Г). На основе двухфакторного дисперсионного анализа (two-way ANOVA) мы пришли к выводу, что содержание белков или соотношение секреторной к несекреторной протеоформе ААТ определяется самими транскриптами ($p=0.0003$ и $p=0.0026$, соответственно) и не зависит от типа клеток ($p=0.44$ и $p=0.54$ соответственно).

Поскольку 69 кДа слитой белок не наблюдается в секретомах, вероятно не содержит N-гликанов и отличается от "голого" eGFP на 38,7 кДа, он вряд ли является продуктом протеолитического деградации зрелого ААТ по нескольким причинам. Во-первых, разница молекулярной массы между основным секреторным ААТ-eGFP, наблюдаемым в клеточном лизате, и обработанным PNGaseF, составляет примерно 7 кДа, что в среднем дает 2,33 кДа на одну единицу N-гликана. Во-вторых, масса ААТ-компонента в слитом белке равная 38.7 кДа предполагает, что в его первичную структуру входит по крайней мере один, если не два, остатка аспарагина, которые обычно подвергаются гликозилированию (рисунок 4.11, Д). Следовательно, если бы 69 кДа ААТ-eGFP был продуктом деградации, при обработке PNGaseF был бы обнаружен сдвиг электрофоретической подвижности, поскольку разница в 2,33/4,66 кДа лежит в пределах разрешения 11% полиакриламидного геля для этого конкретного диапазона молекулярной массы.

Затем мы вернулись к первоначальной гипотезе исследования и рассмотрели возможность того, что источник 69 кДа слитого белка - трансляция с внутреннего альтернативного сайта инициации трансляции *SERPINA1*. Этот процесс приводил бы к потере сигнального пептида (рисунок 4.11, Д), который направляет новосинтезирующуюся цепь ААТ в просвет ЭПР для гликозилирования и секреции [412]. Следовательно, эта протеоформа ААТ — теперь считаемая предполагаемой *изоформой* [351] — оказалась бы негликозилированной и обошла бы секреторный путь. *SERPINA1* AUG-2 является наиболее вероятным кандидатом для альтернативной инициации трансляции, так как он канонический (потенциально более сильный, чем предполагаемые неканонические стартовые кодоны, расположенные вблизи и в рамке считывания с AUG-2), встроен в контекст Козак, подходящий для инициации трансляции (рисунок 4.1, В), и, что особенно любопытно, дает продукт из 332 а.о. с расчетной молекулярной массой полипептидной цепи 37,47 кДа (рисунок 4.11, Д).

4.5.2 Сайт-направленный мутагенез плазмидных конструкций ранжирует внутриклеточную протеоформу в несекреторную изоформу ААТ

Результаты иммуноблоттинга свидетельствуют в пользу того, что предполагаемая несекреторная изоформа ААТ может быть синтезирована с альтернативного стартового кодона AUG-2. Поскольку кодон ААG практически не может служить для инициации трансляции [413], мы мутировали предполагаемый альтернативный стартовый кодон ATG-2>AAG в конструкции HLT (что в случае DLT2 привело бы к крайне низкой “эффективности трансфекции”). Кроме того, мы оптимизировали контекст Козак предполагаемого альтернативного стартового кодона GCAATGC>GCAATGG в DLT2 (потому что эта конструкция будет более “чувствительна” к любым изменениям соотношения секреторного к внутриклеточному ААТ), чтобы повысить эффективность трансляции и сравнить полученные количества несекреторного ААТ.

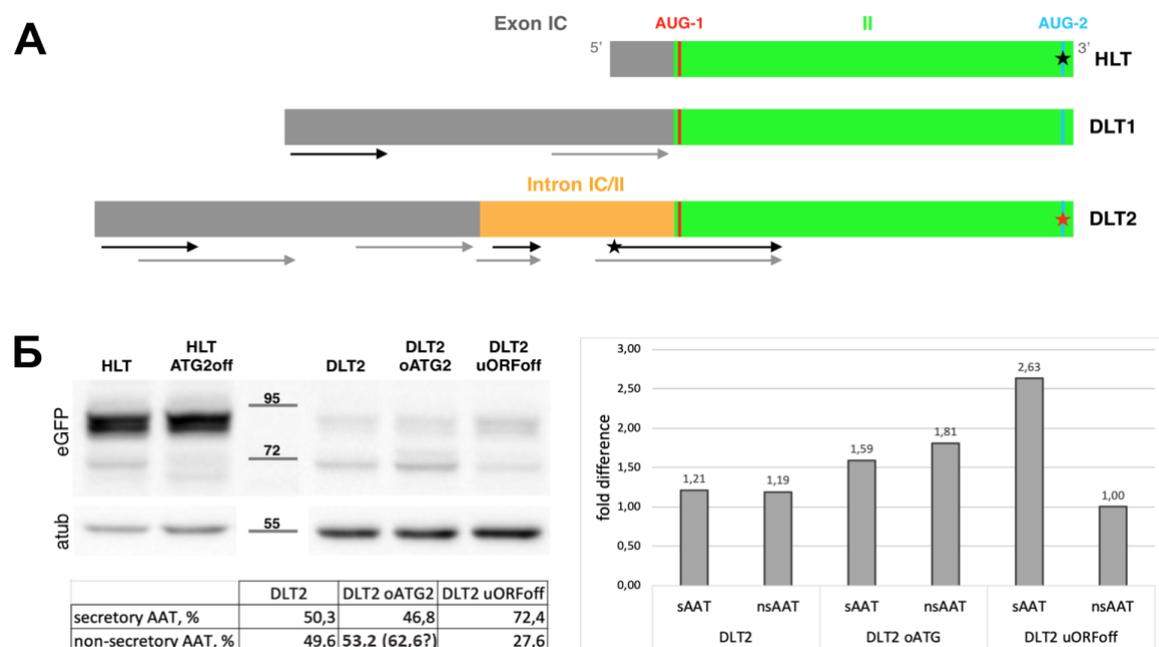


Рисунок 4.12. Идентификация альтОПС несекреторного ААТ. **А)** Схематическая иллюстрация расположения коротких ОПС в 5'-НТО длинных транскриптов *SERPINA1* (5'-ОПС). Черные стрелки - 5'-ОПС со старт-кодоном AUG, серые стрелки - 5'-ОПС с неканоническим старт-кодоном CUG. Показаны главный AUG-1 (красный) и предполагаемый альтернативный AUG-2 (голубой) в рамке с AUG-1. Звезды указывают на позиции, выбранные для мутации. Мутации, нарушающие (AUG>AAG) трансляцию или оптимизирующие контекст Козак, показаны черными или красными звездами, соответственно. На схеме сохранен относительный масштаб элементов. **Б)** Сайт-направленный мутагенез показывает, что внутриклеточная изоформа ААТ транслируется с *SERPINA1* AUG-2. **Слева** - иммуноблоттинг лизатов eGFP⁺ сортированных клеток после трансфекции плазмидами HLT, DLT2 и соответствующими мутантными плазмидами: ATG2off - мутация ATG-2>AAG, uORFoff - мутация предполагаемого перекрывающего основную ОПС 5'-ОПС ATG > AAG, oATG2 - ATG-2 с оптимизированным контекстом Козак (C>G в положении +4). Полные сканы на рисунке 4П; **Справа** - денситометрический анализ экспрессии секреторного (s) и несекреторного (ns) ААТ. Так как интенсивность альфа-тубулина зависит от экспрессируемой конструкции, сигналы были нормированы к общей интенсивности Ponceau S соответствующих дорожек, затем к полосе с минимальной экспрессии (nsAAT в DLT2 uORFoff).

Результаты иммуноблоттинга подтвердили, что AUG-2 служит стартовым кодоном для трансляции несекреторного ААТ, что подтверждается исчезновением соответствующей полосы (рисунок 4.12, Б). Кроме того, появившаяся новая полоса, мигрирующая ниже, предположительно соответствует продукту "включенной" трансляции с одного из нескольких близлежащих неканонических стартовых кодонов *SERPINA1* CUG, расположенных вблизи к AUG-2, которые находятся в подходящем контексте Козак (ΔUCCUGGAGGGCCUGΔ, где предполагаемые стартовые кодоны выделены курсивом, важные пуриновые остатки в позициях -3 и +4 подчеркнуты) (рисунок 4.12, Б). Как ожидалось, оптимизация контекста Козак у ATG-2 привела к увеличению доли внутриклеточного ААТ на несколько процентов (рисунок 4.12, Б (таблица под сканом), В). Это наблюдение дополнительно подтверждает роль AUG-2 в качестве альтернативного сайта инициации трансляции для синтеза несекреторного ААТ. Менее воспроизводимая следовая полоса 75 кДа может являться внутриклеточным ААТ-eGFP, подвергшимся посттрансляционной модификации, так как эта полоса становится более отчетливой после оптимизации контекста Козак ATG-2 (рисунок 4.12, Б). При этом сценарии, увеличение доли составило бы около 13%. На основании этих результатов мы можем сделать вывод о том, что помимо классической секреторной, несекреторная изоформа ААТ синтезируется путем инициации трансляции на альтернативном сайте инициации *SERPINA1*, а именно на AUG-2. Учитывая, что AUG-2 генерирует изоформу ААТ длиной 332 а.о. с молекулярной массой полипептидной цепи 37,47 кДа, что практически эквивалентно компоненту ААТ во внутриклеточном слитом белке, можно заключить, что эта новая изоформа ААТ не гликозилирована.

Итак, независимо от структуры 5'-НТО с длинных мРНК *SERPINA1* синтезируются две изоформы ААТ: секреторная и внутриклеточная, - по-видимому, за счет сканирования с утечкой, поскольку основной старт AUG-1 находится в субоптимальном контексте Козак (рисунок 4.1, В). С транскриптов с более длинной 5'-НТО (DLT2>DLT1>HLT) синтезируется меньшее количество секреторного белка ААТ-eGFP (рисунок 4.11, В). Это может быть объяснено регуляцией трансляции главной OPC 5'-OPC, предсказанных с помощью программы ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) с минимальной длиной OPC в 30 н.о. (рисунок 4.12, А). Из-за усеченного варианта EX1C (длиной 43 н.о.) в HLT отсутствуют 5'-OPC, поэтому он обладает максимальной эффективностью трансляции. DLT1 включает два потенциальных 5'-OPC (третий, присутствующий в NM_000295.4, отсутствует в DLT1 ввиду замены T34>C). В отличие от других вариантов мРНК, удержание части интрона в DLT2 приводит к появлению двух дополнительных пар котерминальных 5'-OPC (рисунок 4.12, А). Вторая пара перекрывает главную OPC (рисунок 4.12, А), что предполагает ее ощутимый вклад в общее ингибирование трансляции последней. Чтобы проверить ее влияние, мы нарушили предполагаемую инициацию трансляции с канонического варианта этой 5'-OPC (ATG>AAG) в DLT2. В

результате доля секреторного ААТ увеличилась на 22 % по сравнению с не мутантным DLT2 (Рисунок 4.12, Б, В). 5'-ОПС, представленные в DLT1/2, не являются триггерами трансляции с альтОПС, как, к примеру, в случае с *CEBPB* [414], а, скорее, регулируют частоту инициации трансляции с главного стартового кодона и практически не влияют на инициацию с AUG-2, поскольку мы наблюдали трансляцию с последнего в структуре HLT, 5'-НТО которого не содержит 5'-ОПС (рисунок 4.11, В). Мы сделали вывод о том, что клеточный аппарат регулирует баланс секреторной и внутриклеточной изоформ ААТ путем транскрипционного контроля на уровне альтернативного сплайсинга *SERPINA1* пре-мРНК. Поскольку *SERPINA1* транскрибируется тканеспецифически, преимущественно в печени, крови, легких, тонком кишечнике, селезенке и почках [43, 122, 121], можно ожидать, что разные ткани будут иметь разное соотношение обеих изоформ ААТ, предопределенное экспрессируемым 5'-НТО.

С помощью анализа на основе люциферазы в качестве репортера продемонстрировали, что некоторые 5'-ОПС ингибируют трансляцию с нижестоящей главной ОПС *SERPINA1* [44]. Недавнее исследование выявило влияние 5'-НТО транскрипта NM_000295.4 на эффективность трансляции главной ОПС, обнаружив, что он создает такую глобальную структуру, в которой оптимизируется трансляция главной ОПС путем повышения доступности сайта инициации последней [121]. Авторы предсказали, что именно этот транскрипт - один из наиболее эффективных с точки зрения синтеза ААТ [121]. В NM_000295.4 предсказываются три 5'-ОПС: первая 5'-ОПС начинается с канонического стартового кодона на непосредственном 5'-конце транскрипта (ACAUGA), который функционирует *in vitro*, но, по-видимому, не активен *in vivo* [44], и две 5'-ОПС с неканоническими стартами CUG в контекстах Козак (AAGCUGT и CCTCUGG), потенциально подходящих для инициации трансляции. Несмотря на то, что косвенно было показано, что обе неканонические 5'-ОПС, по-видимому, не транслируются [121], для дальнейшего подтверждения этих результатов было бы интересно провести мутационный анализ соответствующих стартовых кодонов на наших конструкциях. Наши данные указывают на то, что конструкция DLT1, имеющая почти идентичный 5'-НТО, как в NM_000295.4, действительно, является относительно сильным источником секреторного ААТ. Эффективность трансляции секреторного продукта с него всего в 2,67 и 4 раза ниже по сравнению с HLT в клетках DU145 и HepG2, соответственно. Для наиболее сложной конструкции DLT2 наблюдаемое снижение по сравнению с HLT составило 8,54 и 12,27 раз, соответственно.

4.5.3 Альтернативный сайт инициации трансляции *SERPINA1* и размер альтернативной ОПС консервативны среди млекопитающих

Мы проанализировали мРНК генов-ортологов *SERPINA1* человека из некоторых видов млекопитающих на предмет консервативности альтернативного стартового кодона AUG-2 *SERPINA1*, соответствующего контекста Козак, а также размера альтОПС. Для

анализа мы выбрали два из шести мышинных паралогов, которые являются ортологами человеческого гена - *SERPINA1a/b* [415]. Результаты, представленные в таблице 4.1, с одной стороны, отражают консервативность структуры полноразмерного ААТ за счет консервативности рассматриваемой кодируемой области, а с другой, свидетельствуют в пользу того, что внутриклеточная изоформа ААТ может иметь функциональную значимость.

Species	mRNA accession number	alt AUG to the main AUG start position	Kozak's context	altORF, codons
<i>Homo sapiens</i>	NM_000295.5	259	GCAA <u>AUGC</u>	332
<i>Pan troglodytes</i>	XM_009428360.3	259	GCAA <u>AUGC</u>	332
<i>Gorilla gorilla</i>	JX680603.1	259	GCAA <u>AUGC</u>	332
<i>Pongo abelii</i>	NM_001131225.1	259	GCAA <u>AUGC</u>	333
<i>Hylobates moloch</i>	XM_032135018.1	388	GCAA <u>AUGC</u>	332
<i>Macaca mulatta</i>	NM_001266017.2	259	GCAA <u>AUGC</u>	332
<i>Papio anubis</i>	XM_009212185.4	259	GCAA <u>AUGC</u>	332
<i>Oryctolagus cuniculus</i>	NM_001171081.1	244	GCCA <u>AUGC</u>	332
<i>Rattus norvegicus</i>	NM_022519.2	241	GCCA <u>AUGC</u>	331
<i>Mus musculus</i>	NM_001252569.1	310	GCTA <u>AUGC</u>	333
	NM_009244.4	241	GCTA <u>AUGC</u>	333
<i>Canis lupus familiaris</i>	NM_001080109.2	262	GCGA <u>AUGC</u>	332
<i>Sus scrofa</i>	NM_214395.2	268	GCCA <u>AUGC</u>	332
<i>Bos taurus</i>	NM_173882.2	253	GCGA <u>AUGC</u>	332
<i>Ovis aries</i>	NM_001009799.2	253	GCGA <u>AUGC</u>	332

Таблица 4.1. Анализ консервативности альтернативного сайта инициации трансляции длинных транскриптов *SERPINA1* среди видов млекопитающих. Позиции в третьем столбце рассчитаны относительно главного старт-кодона: указано положение остатка аденина в альтернативном AUG относительно главного. Последний столбец включает количество кодонов в альтОРС, исключая стоп-кодон, равный количеству кодируемых а.о.

4.5.4 Изучение внутриклеточной локализации химерных продуктов трансляции с длинных транскриптов и предполагаемой эндогенной внутриклеточной изоформы ААТ

4.5.4.1 Слитые белки изоформ ААТ-eGFP локализуются в разных клеточных компартментах

Если трансляция начинается с альтернативного стартового кодона, изменение N-конца белка может привести к его доставке в другой компартмент [351]. С помощью программы WoLF PSORT (<https://wolffpsort.hgc.jp>) для внутриклеточной изоформы ААТ предсказали его ядерно-цитоплазматическую локализацию (cyto: 24.5, cyto_nucl: 18.5, nucl: 5.5, extr: 1, cysk: 1), в то время как для секреторного ААТ - внеклеточную (extr: 22, E.R.: 9, plas: 1).

Мы провели микроскопию на живых клетках, используя сигналы от eGFP и ЭПР-трекера. HLT и DLT1, с которых в основном синтезируется секреторный ААТ, показали

цитоплазматическую локализацию слитого белка с ожидаемой солокализацией с ЭПР-трекером, в то время как спеклоподобные сигналы могут представлять транспортные везикулы, производные от ЭПР (рисунок 4.13). DLT2 со значительной долей внутриклеточной изоформы ААТ из-за низкого уровня секреторного аналога утратил обычный паттерн распределения ААТ в ЭПР, слитые белки распределены более диффузно и мигрировали между цитоплазмой и ядром (рисунок 4.13), в соответствии со спрогнозированной локализацией внутриклеточного ААТ. Такие же паттерны распределения классического и внутриклеточного ААТ-eGFP наблюдали и в клеточной линии НерG2 (рисунок 5П).

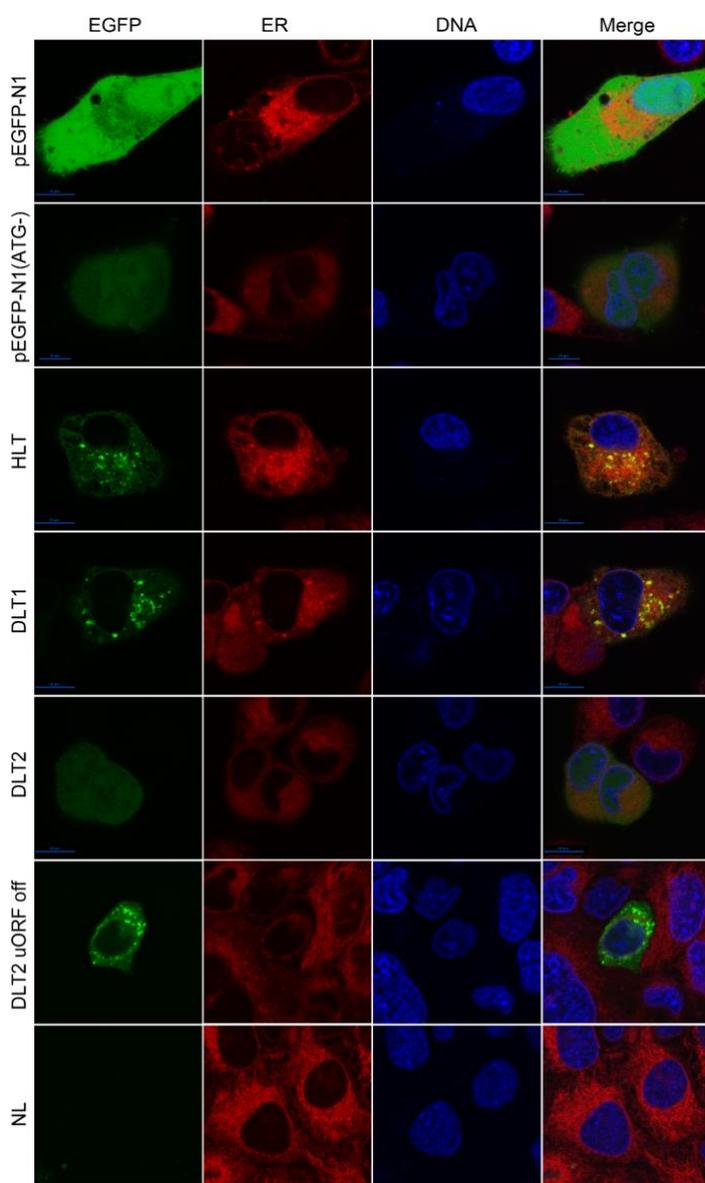


Рисунок 4.13. Изучение внутриклеточной локализации слитых белков ААТ-eGFP в клетках DU145 с помощью конфокальной микроскопии на живых клетках. Клетки трансфицировали конструкциями, экспрессирующими HLT, DLT1 и DLT2. DLT2 uORF off - мутация ATG>AAG 5'-ОПС, перекрывающейся с главной ОПС *SERPINA1* в конструкции DLT2, NL - отрицательный контроль без добавления плазмидной ДНК.

4.5.4.2 Предполагаемая эндогенная внутриклеточная изоформа ААТ накапливается в ядерных спеклах

Для выявления предполагаемого эндогенного внутриклеточного ААТ, мы окрасили фиксированные клетки с помощью Су3-меченых козьих поликлональных антител, выращенных против ААТ из плазмы человека. Козьи IgG сами по себе "липнут" к человеческим клеткам, особенно к их ядрам (рисунок 6П). Поэтому мы предварительно блокировали фиксированные клетки немечеными нормальными (неиммунными) козьими антителами и окрашивали их специфическими IgG-Су3 в смеси с 10-кратным молярным избытком немеченых нормальных козьих IgG, что создало конкурентные условия и повысило специфичность окрашивания. В этих условиях мы смогли уловить разницу между анти-ААТ IgG-Су3 и нормальными козьими IgG-Су3, служащими отрицательным контролем. Предполагаемая эндогенная изоформа ААТ проявляет характерную диффузную ядерно-цитоплазматическую локализацию (рисунок 4.14). Совместная окраска с анти-Sc-35 мышинными антителами показала, что предполагаемая усеченная изоформа ААТ накапливается в ядерных спеклах - специализированном субъядерном компартменте, обогащенном малыми рибонуклеопротеиновыми частицами и различными факторами сплайсинга (рисунок 4.14).

Для изучения возможной ассоциации предполагаемой изоформы с молекулами РНК мы обработали *in situ* пермеабелизованные клетки РНКазой А. Мы наблюдали заметное исчезновение сигнала после обработки ферментом по сравнению с необработанными клетками (рисунок 4.14), что указывает на то, что предполагаемый эндогенный ААТ может быть компонентом РНП-комплекса. В отличие от ААТ, распределение Sc35 оказалось практически нечувствительным к ферменту [416].

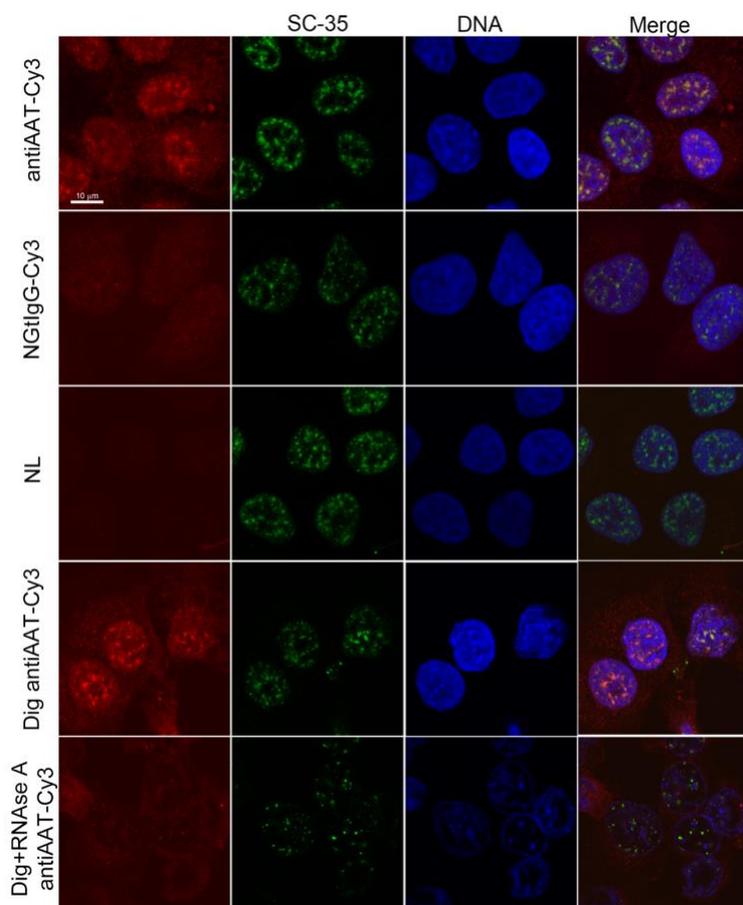


Рисунок 4.14. Выявление предполагаемой эндогенной изоформы ААТ в клетках DU145 методом иммунофлуоресцентной конфокальной микроскопии. Фиксированные клетки предварительно блокировали нормальными козьими IgG (NGtIgG), затем окрашивали с помощью анти-ААТ IgG-Cy3 и анти-Sc35 в смеси с немеченым NGtIgG с большим молярным избытком. В качестве отрицательного контроля использовали NGtIgG-Cy3 в эквивалентной концентрации. Нижние панели - клетки обрабатывали или не обрабатывали RNase A после *in situ* пермеабилзации дигитонином в транспортном буфере [331], фиксировали, блокировали и окрашивали анти-ААТ IgG-Cy3 и анти-Sc35 (в паре со вторичными IgG-Alexa 488) в конкурентных условиях.

Мы предприняли попытку детектировать эндогенный внутриклеточный ААТ с помощью иммуноблоттинга в клеточных экстрактах, но обнаружить ее нам не удалось с помощью антител, полученных против секреторного гликозилированного ААТ из человеческой плазмы, в то время как мы консистентно детектировали ААТ в секретах. Это можно объяснить крайне низким количеством изоформы ААТ и не оптимальным распознаванием используемыми антителами. Используя методы иммунофлуоресценции, мы обнаружили предполагаемую эндогенную изоформу ААТ в ядерных спеклах с помощью тех же антител, вероятно, благодаря сохранению некоторых конформационных эпитопов нативного ААТ. Полученные данные, вместе с эволюционной консервативностью альтОРС *SERPINA1* среди млекопитающих, предполагают, что изоформа ААТ может иметь биологическую значимость, возможно, проявляя функции, отличные от классической изоформы. Для идентификации эндогенной изоформы ААТ и раскрытия ее биологической роли необходимо получить антитела против усеченной и негликозилированной молекулы ААТ для увеличения шансов на ее специфическое распознавание. Это помогло бы "поймать" и охарактеризовать предсказанную *in vitro* эндогенную изоформу как из нормальных, так и из опухолевых тканей. Эти знания могли бы предоставить ценную информацию о возможном ее участии в развитии заболеваний или ее вкладе в поддержание нормальных клеточных функций.

На данном этапе сложно предсказать, какую конкретную функцию будет выполнять несекреторная изоформа ААТ. Несколько преждевременный, но интригующий вопрос - сохранит ли несекреторная ААТ основную антипротеолитическую активность серпина? Чтобы ответить на этот вопрос, нам следует более подробно рассмотреть пространственную структуру молекулы ААТ. Эта активность зависит от метастабильного состояния молекулы ААТ, которое проявляется через экспонирование RCL (а.о. 345–360) белка (рисунок 2.4 раздела 2.2.2 “обзора литературы”), который, после разрезания по сайту M358-S359, вставляется в β -лист А в качестве дополнительной, β -цепи 4 (s4A) [80, 417, 418]. N-концевые 62 а.о. (рассчитанные относительно зрелой молекулы 394 а.о.), отсутствующие во внутриклеточном ААТ, укладываются в структурные мотивы hA (α -спираль А)-s6B (β -цепь 6 β -листа В) - и 60 % hB в нативном секреторном ААТ [129]. β -лист В и β -лист С образуют гидрофобное ядро β -бочки В/С, поддерживающее кинетическую ловушку метастабильной структуры [419]. Остаток V55 hB, отсутствующий в усеченной изоформе ААТ, известен стабилизацией метастабильной конформации серпина [419]. Результаты экспериментов масс-спектрометрии обмена водорода на дейтерий (H/D-EX) указывают на то, что hA играет ключевую роль в выполнении ингибирующей функции ААТ: во время формирования комплекса ААТ-фермент hA, окружающий β -лист В, существенно разворачивается, и это структурное изменение вызывает открытие щели между s5A и s3A β -листа А, в который затем вставляется расщепленная RCL [129].

Некоторые подсказки можно найти из исследований по экспрессии рекомбинантного ААТ в бактериях и эндогенного бактериального серпина. Удаление первых 15 а.о. из ААТ при экспрессии в *E. coli* приводит к синтезу неактивного белка [420]. Бактериальный серпин тенгпин содержит дополнительный N-концевой довесок из 56 а.о. [421]. Удаление по меньшей мере 40 N-концевых а.о. приводит к синтезу серпина в неингибирующей латентной конформации [421]. Мутационный анализ демонстрирует, что каждый из N-концевых остатков N40, L41 и M42 критически необходим для стабилизации гидрофобного участка серпинового домена (L159, I162 и I170) и функционирует для кинетического захвата тенгпина в его нативном метастабильном состоянии, в то время как экспонирование этого гидрофобного участка на поверхности домена серпина, по-видимому, достаточно для осуществления конформационной перестройки [421]. Тот же гидрофобный участок (L131, V134 и A142) в нативном ААТ, вероятно, стабилизируется его N-концом и экспонируется после образования комплекса [129].

Вышеуказанные данные предполагают, что внутриклеточная изоформа ААТ будет находиться в латентно-подобной конформации, в которой мотивы, скрытые в нативной форме, экспонируются и могут обеспечивать потенциалы для неофункционализации молекулы. Стоит отметить, что латентный ААТ устойчив к полимеризации и остается

мономерным вплоть до 70 °C [132]. В отличие от нативного, латентный ААТ, приготовленный из ААТ плазмы, имеет более катодную электрофоретическую подвижность при электрофоретическом разделении в неденатурирующих условиях и изоэлектрическом фокусировании, что указывает на его дистинктные физико-химические свойства [132]. Учитывая, что внутриклеточная изоформа ААТ не содержит отрицательно заряженные углеводные остатки [127], ее ядерная локализация может быть результатом отличия от секреторного ААТ как по пространственной структуре, так и по физико-химическим свойствам.

Стоит внимания тот факт, что пептид, имитирующий N-конец тенгпина, может служить кофактором фолдинга для возвращения денатурированного усеченного серпина к его нативной конформации [421]; латентный ингибитор активатора пламиногена (РАI-1, *SERPINE1* человека) может быть “активирован” отрицательно заряженными фосфолипидами фосфатидилсеринем или фосфатидилинозитолом [422]; активная форма РАI-1 также может быть стабилизирована настолько малыми кофакторами, как ионы галогенидов [423]. Было бы интересно изучить, может ли усеченный ААТ также поддерживаться в ингибирующей конформации какими-либо кофакторами, присутствующими в ядре. Это бы расширило рассматриваемое поле действия усеченной изоформы ААТ к масштабам регуляции экспрессии всего генома, к примеру, на этапе клиппинга гистоновых хвостов - известного механизма эпигенетического контроля. Известно, что в клетках HeLa (в тканях опухоли шейки матки общий уровень клиппинга хвоста гистона H3 существенно ниже, чем в нормальных тканях) и эмбриональных стволовых клетках человека клиппинг хвоста гистона H3 преимущественно осуществляется не установленной сериновой протеазой [424, 425], в то время как в ворсинках кишечника он ассоциирован с дифференцировкой клеток и осуществляется катепсином L (цистеиновой протеазой) и трипсином, локализующимся в том числе в ядрах этих клеток [426]. Также не исключено, что альтернативный продукт трансляции с длинных транскриптов идентичен предполагаемой ядерной изоформе, накапливающейся в ядрах при опухоли яичников (к сожалению, к используемым в этой работе антителам на сайте производителя эпитоп не указывается) - см. раздел 2.5 “обзора литературы” [40].

В обзоре литературы мы приводим данные о том, что ААТ из плазмы связывает лиганды ядерных рецепторов - транскрипционных факторов и репрессоров, таких как РРАR (линолевая и олеиновая ненасыщенные жирные кислоты), RAR (полностью транс-ретиноевую кислоту), REV-ERB α и REV-ERB β (гем), и не исключено, также и RXR (9-цис-ретиноевая кислота, 9-цис-13,14-дигидроретиноевая кислота). Известно, что препарат Prolastin, используемый в заместительной терапии дефицита ААТ, содержит также ненасыщенные жирные кислоты в комплексе с ААТ, в отличие от препаратов других производителей [249]. Поскольку, в отличие от других, в нем обнаруживается ААТ в неингибирующей, латентной, конформации, можно ожидать, что именно такие молекулы,

обладая более гидрофобной поверхностью, и взаимодействуют с жирными кислотами. Также известно, что более гидрофобный ААТ-Z более эффективно связывает гидрофобные лиганды PPAR рецепторов, чем ААТ-М вариант [249], и, вероятно, эти свойства проявляются так же за счет мономеров ААТ-Z в латентно-подобной конформации (рисунок 2.7, В (2) раздела 2.3). Поскольку мы ожидаем, что внутриклеточная изоформа ААТ принимает латентно-подобную конформацию, можно предположить, что секреторный и внутриклеточный ААТ являются регуляторами ядерных рецепторов: первый через рецептор-опосредованный эндоцитоз доставляет лиганды внутрь клетки, а внутриклеточный - в ядро для коактивации ядерных рецепторов.

4.6 Изучение трансляции индивидуальных коротких транскриптов

SERPINA1 в живых клетках *in vitro*: поиск продуктов трансляции

4.6.1 Короткие транскрипты *SERPINA1* транслируются в живых клетках

Короткие транскрипты содержат предполагаемые альтернативные ОРС в рамке *SERPINA1*, начинающиеся в экзоне 3 и экзоне 5 (рисунок 4.8) со стартовыми кодонами (кроме AUG-8) в контексте Козак, подходящим для инициации трансляции (рисунок 4.1, В). Чтобы ответить на вопрос об их функциональности, мы клонировали кДНК, соответствующие коротким транскриптам ST1-4, в мутантный вектор pEGFP-N1(ATG-), как в разделе 4.5.1 (рисунок 4.15, А).

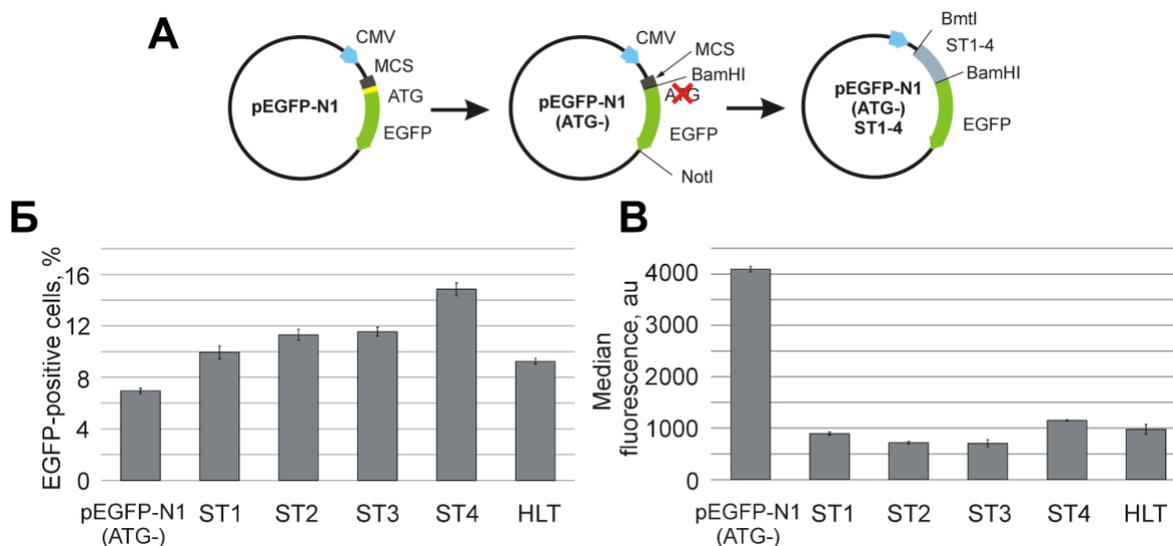


Рисунок 4.15. Трансляция обнаруженных гепато-специфических коротких транскриптов *SERPINA1 in vitro*. А) Схема клонирования кДНК, соответствующей коротким транскриптам. Б) Проточная цитофлуориметрия клеток HepG2 после трансфекции плазмидными конструкциями. Эффективность трансфекции и В) медианы флуоресценции eGFP-содержащих продуктов трансляции. Данные представлены как среднее±стандартное отклонение по двум независимым экспериментам. Диаграммы рассеяния представлены на Рисунке 7П.

После трансфекции мы проанализировали клетки методом проточной цитофлуориметрии (рисунок 7П) и сравнили эффективность трансфекции и медианы флуоресценции от разных конструкций (рисунок 4.15, Б, В). eGFP+ клетки, выявленные после трансфекции “пустым” вектором pEGFP-N1(ATG-), отражают количество клеток, способных транслировать транскрипты с неканонических стартовых кодонов, присутствующих в 5'-проксимальной области, кодирующей eGFP. Все плазмиды, содержащие вставки *SERPINA1*, давали больше eGFP+ клеток, чем мутантная плазида без вставки (рисунок 4.15, Б), что, возможно, отражает синтез пептидов, гомологичных С-концу ААТ. Чтобы прояснить последнее предположение, мы отсортировали живые eGFP+ клетки и сравнили продукты, содержащие eGFP, от конструкций и плазмид без вставки методом иммуноблоттинг на клеточных экстрактах и секретах (рисунок 4.16, А; рисунок 8П). Серпиновая компонента в слитом белке подтверждена детекцией сигнала от секрета НЛТ антителами против ААТ из плазмы, где, наряду с эндогенным секретируемым ААТ, мы наблюдали слитой продукт массой 100 кДа, - как и в случае использования анти-eGFP антител в этом образце, но отсутствующий в секрете от pEGFP-N1 (рисунок 8П). Так как для клонирования кДНК во всех конструкциях мы использовали обратный праймер на одну и ту же кодирующую область и подтвердили положение вставок секвенированием, мы продемонстрировали *SERPINA1* природу N-концов химерных продуктов.

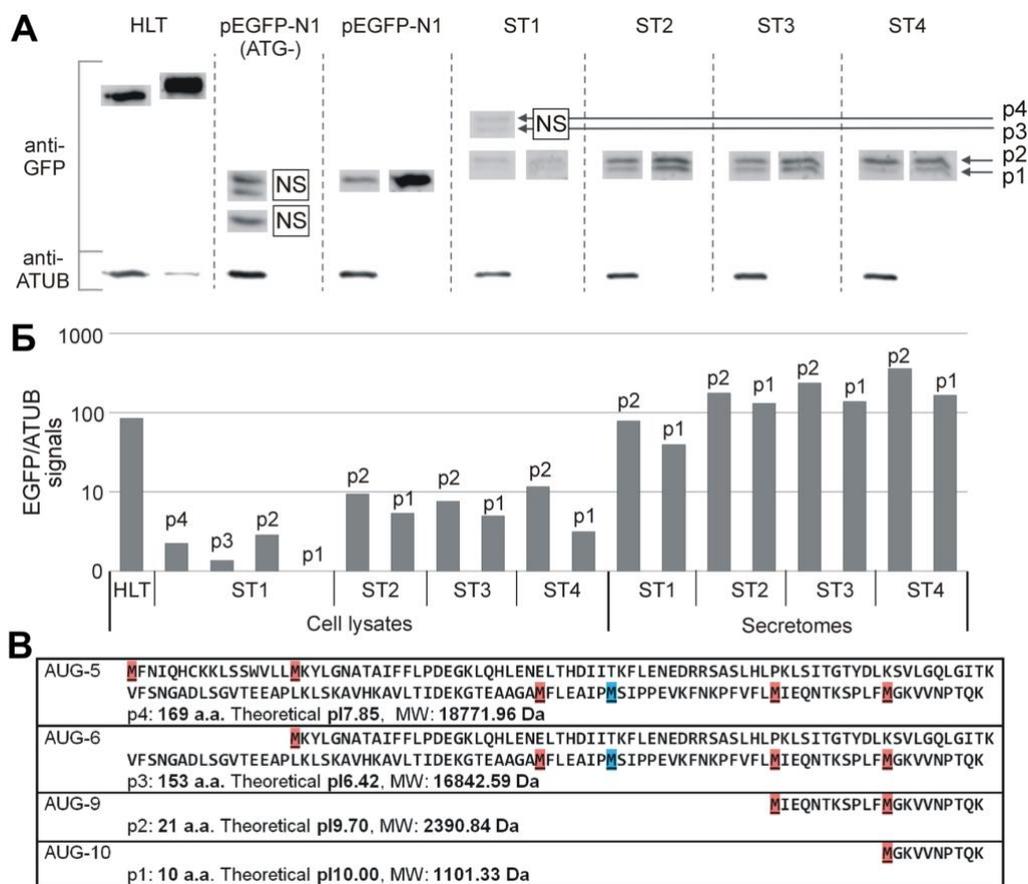


Рисунок 4.16. Выявление предполагаемых активных альтернативных OPC *SERPINA1*. **A**) Иммуноблоттинг лизатов отсортированных eGFP+ клеток и секретомов HepG2, трансфицированных плазмидами со вставками кДНК *SERPINA1*: HLT - длинный транскрипт, ST1-4 - короткие транскрипты; pEGFP-N1 (ATG-) - мутантная плаزمида без вставки (контроль). Для каждой конструкции **слева** представлены лизаты клеток, **справа** - секретомы. NS - не секретируется. Исходный скан представлен на Рисунке 8П. Сигналы проявлены антителами против eGFP, конъюгированными с пероксидазой хрена. **Б**) Относительная интенсивность сигнала eGFP, нормированного на альфа-тубулин (atub), отражает относительную эффективность соответствующих предполагаемых альтернативных OPC. Полученные отношения нормированы на минимальный сигнал (ST1-p1 продукт) и представлены в логарифмической шкале для лизатов клеток и секретомов. Отношение для секретомов HLT отсутствует из-за быстрого насыщения соответствующего сигнала eGFP. **В**) Последовательности предполагаемых продуктов трансляции, расчетные изоэлектрические точки (pI) и молекулярные массы продуктов p1-4. Расчеты проведены с использованием онлайн-программы (https://web.expasy.org/compute_pi/).

Как и ожидалось, полноразмерный ААТ-eGFP гликозилируется и секретируется. Внутриклеточный продукт HLT на 60,3 кДа выше, чем eGFP в отдельности, в то время как секретируемый продукт примерно на 9 кДа выше, чем внутриклеточный, что указывает на полное созревание остатков сахаров в аппарате Гольджи. Трансляция наблюдается также с мутантной плазмиды, которая не содержит главный OPC eGFP, что приводит к образованию трех основных полипептидных продуктов. Самый длинный является продуктом трансляции с неканонического стартового кодона, проксимального к 5'-концу кодирующей области eGFP, потому что он практически идентичен не мутантному eGFP по электрофоретической подвижности. Поскольку для флуоресценции eGFP критически важны остатки аминокислот 7–229 [427], более короткие продукты (с оценочной

минимальной разницей в 20 а.о. с самым длинным), по-видимому, вносят несущественный вклад в общую интенсивность флуоресценции. В то же время медианная интенсивность флуоресценции значительно выше в случае pEGFP-N1(ATG-) по сравнению с теми, которые содержат вставки *SERPINA1* (рисунок 4.15, В), что можно объяснить большим квантовым выходом “голой” молекулы eGFP.

В отличие от pEGFP-N1 и pEGFP-N1(ATG-), плазмиды ST1-4 дают продукты с большим молекулярным весом, указывая на трансляцию со стартов, расположенных в последовательности вставок. Детектированы четыре основных продукта ААТ, синтезируемых с коротких транскриптов *SERPINA1*: p1, p2, p3 и p4 (рисунок 4.16, А), которые имеют большую электрофоретическую подвижность, чем eGFP в отдельности, на 0,7, 2,2, 15,6 и 18,7 кДа, соответственно. Два обстоятельства позволили нам вывести рабочие альтернативы *SERPINA1*. Во-первых, вставка *SERPINA1* приводит к ингибированию трансляции мутантного eGFP с неканонических 5'-проксимальных стартов, наблюдаемых в случае мутантной плазмиды без вставки (рисунок 4.16, А; рисунок 8П). Во-вторых, экзон 4 содержит пять CUG, наиболее сильных из неканонических старт-кодонов [428], в подходящем контексте Козак и в рамке *SERPINA1*. Но поскольку нет продукта, соответствующего трансляции с экзона 4, ни один из них, по-видимому, не активен. Таким образом, можно предположить, что короткие транскрипты транслируются с наиболее сильных, канонических, старт-кодонов AUG [428]. Разница в электрофоретической подвижности по сравнению с eGFP указывает на то, что в ST1 есть две активные OPC со стартами в экзоне 3 (ATG-5 и 6, рисунок 4.1, А, В; рисунок 4.8; рисунок 4.16, В) и две изоформные (относительно главной) короткие OPC (икОПС) со стартами в экзоне 5 (ATG-9 и 10, рисунок 4.1, А, В; рисунок 4.16, В). В ST2-4 активны те же две икОПС. Оценочные длины продуктов p1-p4 составляют 10, 21, 153 и 169 а.о., соответственно (рисунок 4.16, В). В то же время в HLT активна преимущественно главная OPC, а также альтернативная OPC (см. раздел 4.5.2), с каноническими стартами в экзоне 2, AUG-1 и AUG-2, соответственно (рисунок 4.1, А, В; рисунок 4.8; рисунок 8П).

Мы также проверили слитые белковые продукты на способность к секреции. Мы обнаружили, что только пептидные продукты, происходящие из EX5, попадают во внеклеточное пространство. Более длинные (p3/4) усеченные протеоформы ААТ не секретируются, возможно, поскольку полипептиды, кодируемые EX3/EX4, влияют на биохимические свойства пептидов, происходящих из EX5, компенсируя их общий положительный заряд (pI: p1 = 10, p2 = 9,7, p3 = 6,42, p4 = 7,85), так, что они ингибируют способность к секреции. Действительно, положительный заряд на N-конце полипептидной цепи положительно влияет на эффективность секреции через путь, не зависящий от SRP [429]. Немутантный eGFP также секретируется, что наблюдалось ранее [430]. Поскольку eGFP, происходящий от pEGFP-N1(ATG-), не обнаруживается в культуральном супернатанте, мы предполагаем, что секреция слитых p1/p2 с мутантным eGFP отражает

естественное поведение С-концевых пептидов, которые являются драйверами секреции химерных продуктов.

Поскольку в образцах содержалось равное количество отсортированных eGFP+ клеток, мы оценили относительную эффективность трансляции альтернативных OPC, сравнивая интенсивности сигнала eGFP от синтезированных полипептидов (рисунок 4.16, Б). OPC в ST2-4, которые не содержат старт-кодона из EX3, проявляют более высокую общую эффективность трансляции, чем ST1, а продукты p1 и p2, по-видимому, являются основными продуктами, синтезированными с коротких транскриптов, вероятно, благодаря оптимальной структуре соответствующих 5'-НТО. Во всех коротких транскриптах OPC-p2 активнее, чем OPC-p1, в ST1 OPC-p4 активнее, чем OPC-p3, вероятно, обладая либо более выгодным контекстом Козак, либо более благоприятной локальной пространственной структурой РНК для распознавания рибосомой соответствующего сайта инициации трансляции (рисунок 4.8; рисунок 4.16, Б).

Наиболее изученным пептидом ААТ *in vitro* является С36 (см. таблицу 2.2 раздела 2.6 “обзора литературы”), который был предложен в качестве прямого продукта трансляции коротких транскриптов гена *SERPINA1* [344], а также подразумевался таковым в результате амплификации пептид-кодирующей области EX5 в геноме опухолевой клетки [340]. Наши результаты показывают, что короткие гепато-специфические транскрипты *SERPINA1* могут быть источником С-концевых пептидов ААТ С10 (p1) и С21 (p2), кодируемых EX5, но не С36. Это не удивительно, поскольку соответствующий потенциальный стартовый кодон находится в крайне слабом контексте Козак (ATG-8, рисунок 4.1, В). Таким образом, наши данные свидетельствуют о том, что С36 (SIPP...), скорее всего, образуется исключительно посредством протеолиза белка-предшественника, несмотря на правдоподобность того, что первый остаток метионина (MSIPP...) мог бы быть расщеплен метионинаминопептидазой, если бы он был непосредственно транслирован со своего собственного транскрипта [431, 432]. Тем не менее, чтобы сделать окончательный вывод о том, что более длинная рамка с ATG-7 (Рисунок 4.1, В) не активна, необходимо испытать другие тканеспецифические короткие транскрипты *SERPINA1* [344] или другой тип клеток.

Длинную мРНК *SERPINA1* можно рассматривать как транскрипт с «молчащими 3'-цистропами», принимая во внимание то, что их OPC котерминальны (имеют один и тот же стоп-кодон) и могут быть экспрессированы только при приближении к 5'-концу за счет продукции более коротких мРНК [352]. Действительно, согласно нашим данным, трансляция пептидов, гомологичных ААТ, с альтернативных OPC возможна только в случае коротких транскриптов. Это правило также было продемонстрировано после внесения мутации в pEGFP-N1, где альтернативные нижележащие рамки eGFP «включались» при удалении основного ATG. Поскольку мы устранили все некэпированные транскрипты в 5'-RACE, то есть идентифицированные короткие

транскрипты экзпированы, можно предположить, что они транслируются кэп-зависимым способом [433]. С каждого варианта усеченного транскрипта транслируются минимум две ОРС, вероятно, посредством сканирования с утечкой из-за неоптимального контекста Козак или, возможно, с отдельных по-разному свернутых молекул мРНК.

4.6.2 Сочетание 3'-проксимальных альтернативных ОРС *SERPINA1* представлено у приматов. икОРС *SERPINA1* - объект текущей эволюции млекопитающих

Далее мы поставили целью проанализировать наличие 3'-проксимальных альтОРС, в том числе икОРС, *SERPINA1* в генах-ортологах среди видов млекопитающих. Для анализа мы выбрали тот же набор видов, как в предыдущем разделе 4.5.3. Выравнивание показывает, что чем больше альтОРС по размеру, тем больше она представлена среди ортологов, и комбинация как минимум трех 5'-проксимальных (относительно рассматриваемых альтОРС) или всех четырех альтОРС достигается только у приматов (рисунок 4.17, А).

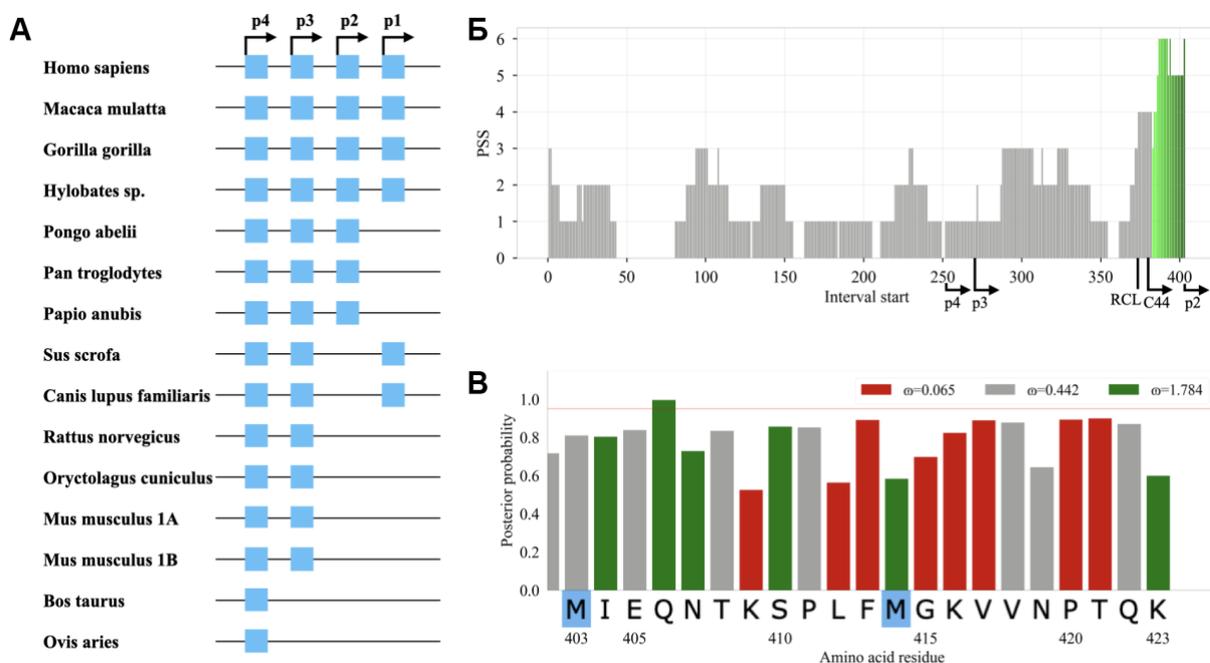


Рисунок 4.17. Консервативность альтОРС *SERPINA1* и давление отбора кодонов среди набора видов млекопитающих. **А**) Схематическая иллюстрация предполагаемых альтОРС в составе транскрипта ST1 и их представленность среди других видов. Все кодоны AUG (голубые прямоугольники) встроены в контекст Козак, подходящий для инициации трансляции. Список идентификационных номеров анализируемых мРНК представлен в таблице 4.1. **Б**) Гистограмма, представляющая количество положительно отобранных сайтов (PSS) в 21-мерных интервалах, смещенных на 1 а.о. к С-концу, идентифицированных моделью M3 с апостериорной вероятностью $P > 0.5$. Отмечен интервал, содержащий участок петли реактивного центра ААТ (RCL), а также начальная позиция пептида С44. Интервалы, начинающие перекрываться с p2, показаны в градиентном зеленом. **В**) Классы и их апостериорные вероятности для последних 21 а.о., соответствующих p2. Красная линия соответствует апостериорной вероятности = 0,95. Аминокислоты относятся к последовательности *Homo sapiens*. 5 пробелов на позициях 44, 46, 47,

309, 351 из-за вставок в других последовательностях привели к нумерации последней позиции 423 (а не 418, как у ААТ человека). Метионины выделены голубой заливкой.

Рисунок 4.17, А показывает, что, по сравнению с другими видами млекопитающих, приматы проявляют консервативность икОПС-р2, который наиболее активен среди альтОПС *SERPINA1* (рисунок 4.16, Б). Мы предположили, что икОПС-р2 был эволюционно отобран для проявления специфических новых функций соответствующего белкового продукта. Напомним, что в структуре полноразмерного секреторного ААТ этот участок полипептидной цепи спрятан внутри молекулы и представляет структурные мотивы s4B(частично)-s5B (рисунок 2.4, Б раздела 2.2.2 “обзора литературы”). Чтобы выяснить, был ли положительный отбор более смещен в сторону икОПС-р2, мы протестировали каждый кодон ортологичных *SERPINA1* генов млекопитающих на давление отбора с помощью сайт-специфических моделей (где сайты = кодон) PAML [335]. Мы провели тест отношения правдоподобия (LRT, likelihood ratio test) для подтверждения наличия позитивно отобранных сайтов (PSS) с помощью попарного сравнения на наборе моделей, отличающихся распределением ω -значений (где ω - нормализованное отношение несинонимичных нуклеотидных замен к синонимичным). LRT показал, что в случае *SERPINA1* модели, позволяющие сайтам быть под положительным отбором (M2 или M8), значительно лучше соответствуют данным, чем нейтральные модели (M1 или M7) ($p < 0.01$; таблица 1П приложения).

Затем мы поставили целью определить, какие конкретные кодоны *SERPINA1* положительно отобраны с использованием модели M3. Эта модель предполагает, что около 8,3% сайтов (35 кодонов) в главной ОПС *SERPINA1* подвергаются положительному отбору (таблица 2П). Апостериорные вероятности для каждого сайта *SERPINA1*, рассчитанные с помощью M3, показаны на рисунке 9П. С помощью модели M3 мы сравнили количество PSS в 21-мерных интервалах (поскольку пептид р2 состоит из 21 а.о.) с использованием подхода со скользящим окном, начиная с первого а.о. ААТ (первого кодона) и сдвигаясь к С-хвосту (3'-концу) на 1 аминокислоту (1 кодон). икОПС-р2 находится в топ-8 интервалов, ранжированных по содержанию PSS (рисунок 4.17, Б). Возможно, эти сайты важны для приобретения специфических функций р2. Получение специфических функций р2 (С21) приматов против соответствующих С-концевых фрагментов ААТ других видов млекопитающих может быть предсказано с помощью биоинформатической программы PredictProtein (таблица 3П, [434]). Интересно, что предполагаемый икОПС-AUG-7, кодирующий пептид С44 (MFLEA..., обнаруженный *in vivo* [48]), включает все топ-8 интервалов, обладая максимальным обогащением PSS. В нашем исследовании соответствующего продукта трансляции не обнаружено.

Учитывая, что икОПС-р2 встречается только у приматов (следовательно, может рассматриваться как ограниченная таксоном), и имеет относительно высокий процент PSS, мы предполагаем, что пептид может быть полезен для видов. Стоит отметить, что

синтетический пептид С20 не образует амилоидоподобные фибриллы, в отличие от С36 [207], то есть р2 существует в растворимой нетоксичной форме, что важно для сохранения ОРС со временем [435]. Функциональная значимость пептидов, кодируемых EX5, может быть предположена на основе данных об известных мутантных аллелях *SERPINA1*, связанных с дефицитом ААТ. Аллели (включая те, которые вызывают “нулл” состояние - полное отсутствие ААТ в сыворотке), которые бы влияли на структуру икОРС таким образом, что нарушали трансляцию пептидов или вызывали глобальные изменения в первичной структуре пептидов, если таковые имеются, найдены только у гетерозигот [436, 437]. Это свидетельствует в пользу того, что трансляция икОРС *SERPINA1* может быть важна для человека.

4.6.3 Изучение внутриклеточной локализации химерных продуктов трансляции с коротких транскриптов и предполагаемых эндогенных С-концевых пептидов ААТ

4.6.3.1 С-концевые пептиды ААТ-eGFP накапливаются в ядре

Для выявления клеточной локализации слитых белков мы проанализировали трансфицированные клетки после окрашивания их антителами против eGFP (рисунок 4.18). Немутантный eGFP проявляет интенсивное окрашивание ядер (рисунок 4.18, а), в то время как мутантный eGFP равномерно распределен между цитоплазмой и ядром, не проявляя различимых контуров ядра (рисунок 4.18, б).

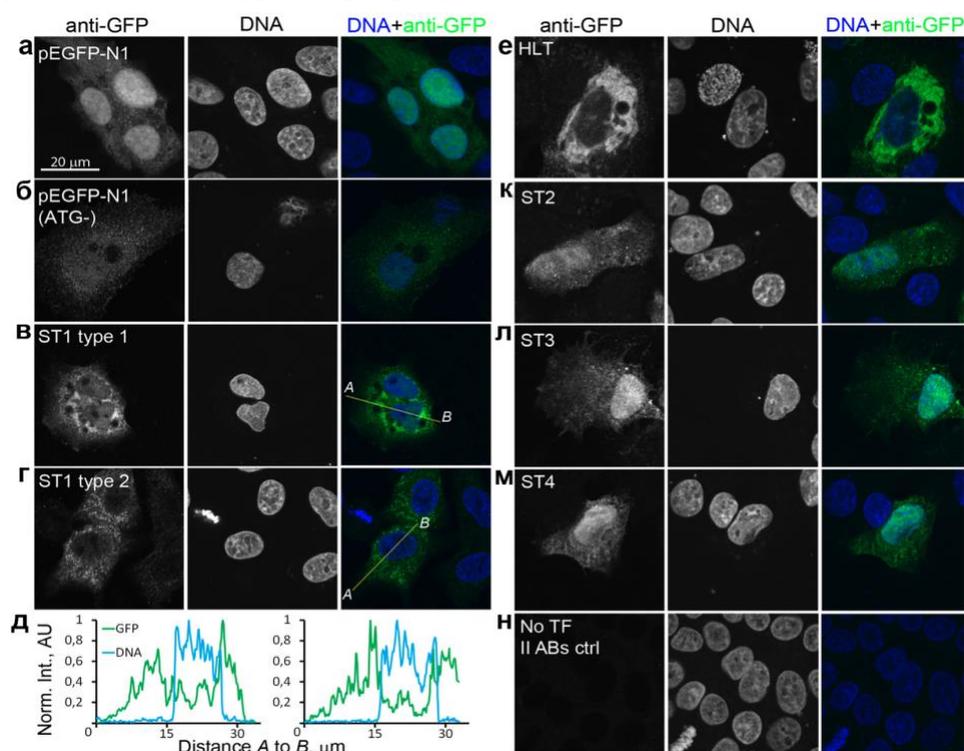


Рисунок 4.18. Иммунофлуоресцентный анализ полипептидных продуктов, транслированных после трансфекции клеток НерG2 плазмидами: немутантный pEGFP-N1 (а), мутантный pEGFP-N1 (ATG-) (б), ST1 (в, г) и соответствующие графики нормализованной интенсивности флуоресценции вдоль профиля линии (д, слева - тип 1, справа - тип 2), HLT (е), ST2–4 (к-м, соответственно). н - в

качестве отрицательного контроля использовались только вторичные антитела DAPI-Alexa Fluor-488, No TF - клетки без плазмиды.

Наиболее короткие продукты, которые идентичны от конструкций ST2–4 (p1 и p2), проявляют сильное ядерное обогащение в отличие от тех, что получены от родительской мутантной плазмиды (рисунок 4.18, к-м). Продукты ST1 дают два типа внутриклеточного распределения: с относительно более сильным ядерным окрашиванием (рисунок 4.18, в, д (слева)) и с почти отсутствующим сигналом в ядре (рисунок 4.18, г, д (справа)). Мы предполагаем, что мы видим только продукты p1 и p2 от ST1 в ядре, в то время как более крупные продукты не могут транслоцироваться в этот компартмент. Одно из возможных объяснений заключается в том, что полипептид, кодируемый EX3/EX4, влияет на общую молекулярную массу слитого белка, возможно, делая эти продукты неспособными преодолеть барьер ядерных пор. Однако это предположение, по-видимому, неверно, так как даже тетрамерный GFP может попасть в ядро [438]. Другое объяснение основывается на влиянии полипептидов, кодируемых EX3/EX4, на биохимические свойства пептидов, кодируемых EX5, делая их неспособными к транслокации в ядро. Как и ожидалось, продукт NLT локализуется в цитоплазме и предназначен для секреции (рисунок 4.16, А; рисунок 4.18, е).

4.6.3.2 Поиск эндогенных С-концевых пептидов ААТ, синтезируемых опухолевыми клетками: получение антител против С-концевого E5-эпитопа ААТ и детекция предполагаемых С-концевых пептидов ААТ в ядрышках

Транслируются ли короткие транскрипты *SERPINA1 in vivo*? Чтобы ответить на этот вопрос, мы получили кроличьи поликлональные антитела против последних 9 а.о. ААТ, которые бы распознали альтернативные усеченные протеоформы ААТ. Специфичность антител против E5-эпитопа (эпитоп, кодируемый EX5) оценивалась методом ИФА на конъюгате BSA-пептид E5 и против чистого BSA (рисунок 10П). Аффинно-очищенные антитела E5 анализировались на их способность отличать С-концевые пептиды ААТ от полноразмерного белка. Мы провели ИФА по "вылавливанию" ААТр (ААТ из плазмы человека) из раствора. Наши результаты показывают, что эпитоп E5 спрятан внутри белка и не доступен для распознавания без термически индуцированной денатурации антигена (рисунок 4.19, А), - таким образом, антитела против E5-эпитопа способны отличать усеченные протеоформы ААТ от целого нативного белка.

Мы обнаружили, что кроличьи IgG также имеют тенденцию связываться с клетками человека, особенно с их ядрами, неспецифически (рисунок 11П), как и в случае козьих IgG (рисунок 6П). Чтобы заблокировать взаимодействие кроличьих IgG, не имеющее отношение к их взаимодействию с E5-эпитопом, мы предварительно инкубировали образцы с неиммунной сывороткой того же кролика и окрашивали клетки флуоресцентно

мечеными первичными антителами. В отличие от отрицательного контроля, антитела против E5-эпитопа равномерно окрашивают цитоплазму и ядро и накапливаются в ядрышках (рисунок 4.19, Б-Г; рисунок 12П). Эпитопы E5 солокализуются с белком ядрышка B23 (также известным как nucleophosmin, NPM1, NO38 или numatrin) (рисунок 4.19, Г) и предположительно локализуются в гранулярном компоненте ядрышка [439].

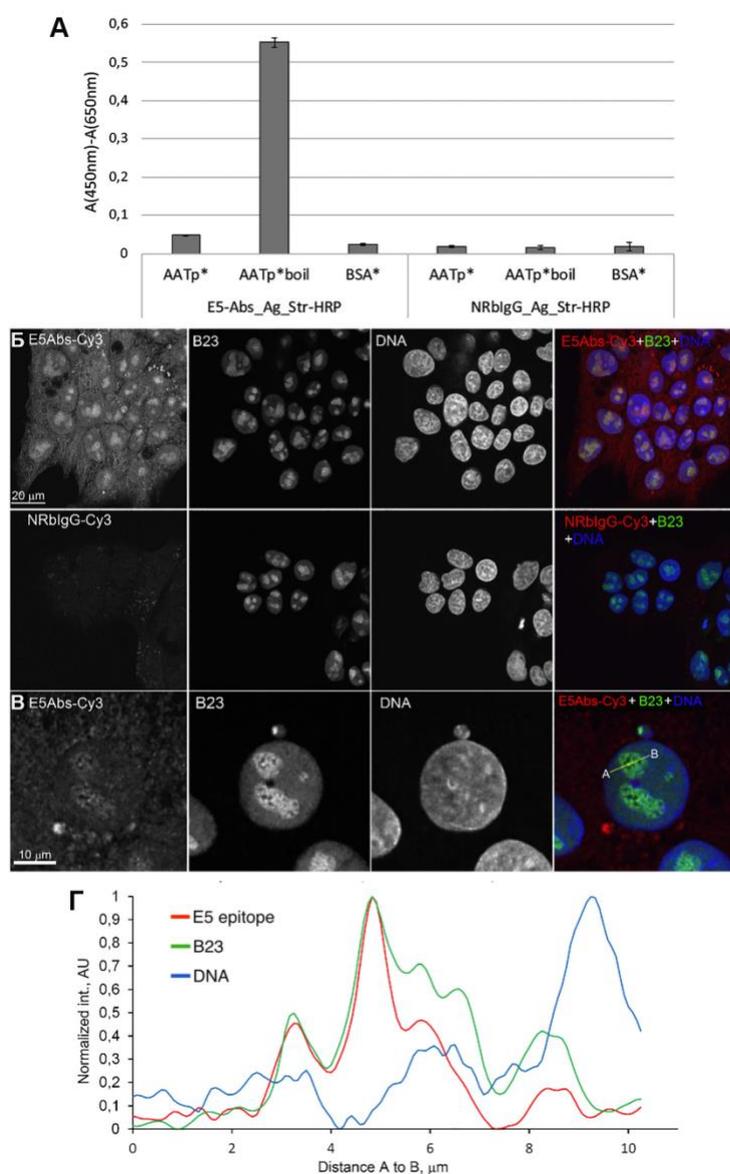


Рисунок 4.19. Выявление предполагаемых эндогенных С-концевых пептидов ААТ. **А)** ИФА, демонстрирующий, что антитела против E5-кодируемого эпитопа (E5-Abs) не узнают таковой в структуре нативного биотинилированного ААТ из плазмы (ААТр*, 1:1), но дают существенный сигнал после термической денатурации ААТр*. Отрицательные контроли: BSA* и нормальные кроличьи IgG (NRbIgG) на том же наборе антигенов. Ag=антиген, Str-HRP=конъюгат стрептавидина с пероксидазой хрена. Данные представлены как среднее±стандартное отклонение по трем повторам. **Б)** Прямой иммунофлуоресцентный анализ с использованием анти-E5 антител, конъюгированных с Cy3 (E5Abs-Cy3), в качестве отрицательного контроля - NRbIgG-Cy3. Фиксированные клетки предблокировали немеченой преиммунной кроличьей сывороткой. Анти-B23 антитела в паре со вторичными DAMI-Alexa Fluor-488 использовали для меченя ядрышек. **В)** Изображение ядра с большим увеличением. **Г)** График нормализованных интенсивностей флуоресценции вдоль профиля линии (от А до В, объединенное изображение из рисунка В), показывающий солокализацию эпитопов E5 и B23.

Основываясь на наших наблюдениях за внутриклеточным распределением слитых пептидов ААТ с eGFP, мы предполагаем, что именно пептиды p1 (C10) и/или p2 (C21) проникают в ядра и обогащают ядрышки. Как и в случае с внутриклеточной изоформой ААТ, внутриклеточное распределение слитых пептидов-eGFP отличается от распределения эндогенных эпитопов E5: первые почти равномерно распределяются внутри ядра (рисунок 4.18, к-м), в то время как последние накапливаются в ядрышках (рисунок 4.19, Б-Г), что может отражать влияние eGFP компонента. Стоит отметить, что предполагаемые эндогенные С-концевые пептиды ААТ, транскрируемые с коротких транскриптов, и предполагаемая внутриклеточная изоформа ААТ, транскрируемая с

длинных транскриптов, схожи в их проникновении в ядро, но внутри ядра локализуются в разных субъядерных компартментах (рисунок 12П).

Усеченные протеоформы белка могут локализоваться в другие клеточные компартменты [440], иметь антагонистические свойства по сравнению с длинным аналогом [441], а также выполнять функции и вовсе не типичные для полноразмерной протеоформы [442]. Поскольку продукты от конструкций ST1–4 имеют внутриклеточное распределение, отличное от распределения HLT, сопровождаемое накоплением предполагаемых эндогенных пептидов в клеточных ядрышках, можно предположить, что генерирование усеченных транскриптов *SERPINA1* посредством посттранскрипционного процессинга (см. раздел 4.4) диверсифицирует функциональный репертуар протеома гена *SERPINA1*. Эта диверсификация происходит благодаря прямому извлечению функционального потенциала С-конца ААТ, скрытого в нативном полноразмерном белке [134, 443].

Усеченные продукты р3 (С153) и р4 (С169), по-видимому, утрачивают способность к секреции. Мы можем предположить их возможную роль, экстраполируя известную функцию схожего аналога. Пептид SPAАТ (С44) (см. таблицу 2.2 раздела 2.6 “обзора литературы”) служит обратимым анти-протеолитическим агентом во внеклеточном матриксе ткани плаценты [49, 50], несмотря на то, что в его структуре сохранены всего лишь 65% от критически важной для ингибирования протеаз структуры петли реакционного центра (RCL, E32-G51 а.о. с С-конца ААТ [444]). В структуре р3 и р4 представлена вся последовательность RCL, что необходимо для *необратимого* ингибирования. Принимая во внимание цитоплазматическую локализацию р3 и р4, можно предположить, что эти протеоформы вносят вклад в анти-протеолитическую емкость клетки.

Сложнее предсказать роль р1 и р2, идущих в секрецию (что расширяет их поле действия), а также накапливающихся в ядре. Пептиды не только не имеют RCL, в них также отсутствуют мотивы с установленной функцией. Исключение составляет только *in vitro* исследование синтетического аналога р1 пептида С10 (см. таблицу 2.2 “обзора литературы”), повышающего уровень коллагена I типа в культуральном супернатанте [445].

Секреция коллагена селективно контролируется ядрышковым белком фибриларином, что предполагает регуляцию процессинга коллагена неизвестными механизмами, затрагивающими белки ядрышкового компартмента [446]. HSP47, кодируемый серпиновым геном *SERPINH1*, является проколлаген-специфичным шапероном, предотвращающим латеральную агрегацию своего “субстрата” [210] и повышающим его эффективность секреции [447]. Как мы ранее приводили в обзоре литературы, секреторный ААТ проявляет шапероноподобную активность по отношению к набору внутриклеточных мишеней в бесклеточной системе *in vitro* и механизм, к

сожалению, не известен [8]. Недавно обнаружили, что гранулярный компонент ядрышек (где мы детектировали E5-эпитопы) является резервуаром контроля качества развернутых под действием стресса ядерных белков [448], где, помимо HSP70 [448], B23 потенциально может служить шапероном [449]. Помимо свойства последнего связывать NLS-подобные положительно заряженные пептиды [450], B23 предотвращает агрегацию развернутых белков, и его “клиенты *in vitro*” частично перекрываются с таковыми, изученными для ААТ [449, 8]. Синтетические, даже короткие, пептиды производные от известных молекул шаперонов сохраняют шапероноподобную активность своих крупных аналогов [451, 452]. Известны также природные митохондриальные пептиды, проявляющие те же свойства [453]. Подытоживая, можно предположить, что и С-концевые пептиды ААТ, транслируемые с коротких транскриптов гена *SERPINA1*, могут представлять собой часть клеточного механизма контроля качества белковых структур. В свете этого предположения было бы интересно выявить минимальный мотив ААТ, сохраняющий шапероноподобные свойства.

Эффективность синтеза усеченных протеоформ ААТ существенно ниже, чем классического секреторного ААТ (рисунок 4.16, Б). Учитывая, что наша система предполагает оверэкспрессию транскриптов с CMV промотора, ожидается, что эндогенные продукты трансляции от коротких транскриптов будут присутствовать в гораздо меньших количествах. Малый размер основных продуктов (p1 и p2) и их низкая представленность существенно затрудняют их обнаружение [348], что, возможно, объясняет, почему этот феномен (экспрессия *SERPINA1* с иКОРС) не изучался ранее.

Мы многократно пробовали детектировать С-концевые пептиды ААТ в клеточных лизатах и секретах с помощью иммуноблоттинга после трис-трицинового электрофореза, но безуспешно. Однако отсутствие видимых сигналов не исключает существования продуктов трансляции коротких транскриптов. Возможным объяснением может быть как их низкая представленность, так и невозможность разделения коротких пептидов - основных продуктов трансляции - в полиакриламидном геле. Синтетический пептид С36 (SIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK, где красным выделены кислые а.о., синим - основные, зеленым - гидрофобные) consistently разделялся в геле, в то время как пептид С10, использованный нами для получения анти-E5 антител (CGKVVNPTQK, содержит дополнительный остаток цистеина для конъюгации с белком-носителем), нанесенный с лихвой, по-видимому, не связывается с додецилсульфатом натрия и не проникает в гель вовсе (рисунок 13П). В перспективе мы планируем провести иммунопреципитацию E5-эпитоп-содержащих протеоформ, совмещенную с их идентификацией в элюате методами масс-спектрометрии.

Мы также пробовали протестировать детектируемые предполагаемые С-концевые пептиды ААТ на их происхождение из экзона 5 гена *SERPINA1* с помощью геномного редактирования CRISPR/Cas9. Мы отобрали две сублинии НерG2 с гомозиготным

нокаутом практически всей кодирующей области экзона 5 (рисунок 14П). На более ранних пассажах в гомозиготных нокаутных клональных популяциях HepG2, 10E3 и 10E10, практически отсутствовал сигнал от E5-эпитопа в ядрышках (рисунок 15П), однако на более поздних пассажах сигнал не только восстановился, но стал заметно ярче, чем в родительских клетках (рисунок 16П). Полученные результаты не рассматриваются нами как достоверные и их можно объяснить срабатыванием компенсаторных механизмов, обеспечивающих генетическую целостность [454, 455].

Недавние исследования продемонстрировали, что механизмы генетической компенсации путем транскрипционной адаптации (адаптации путем включения транскрипции компенсаторных генов, являющихся паралогами либо обладающих схожей последовательностью с мутированным геном) срабатывают в случае, если не нарушена транскрипция мутированного гена, и обеспечиваются гомологией фрагментов деградированной аберрантной мРНК [456]. Поскольку в нашем случае сигнал от E5-эпитопа в мутантных клетках восстановился, в то время как компенсация обеспечивается не затронутой гомологичной последовательностью мутантной мРНК, логично предположить, что функции предполагаемых транскрибированных протеоформ ААТ могли быть замещены продуктами генов-паралогов *SERPINA1* (рисунок 14П, Г). Для того, чтобы верифицировать происхождение внутриклеточного E5-эпитопа, необходимо провести аналогичный эксперимент по нокауту *всей* области гена *SERPINA1*.

Заключение

Изученные нами особенности экспрессии гена *SERPINA1* выводят на новый уровень понимания его функционирования и свидетельствуют о том, что, по крайней мере в опухолевых клетках человека, задействуются “резервы” гена в виде альтернативных сайтов инициации трансляции. 5'-проксимальный альтернативный стартовый кодон AUG-2 реализуется в составе длинных транскриптов за счет сканирования с утечкой, в то время как 3'-проксимальные старты (предположительно AUG-5, 6, 9 и 10) - в отсутствие ограничений главной ОРС, то есть в составе коротких транскриптов, содержащих экзон 5 и образующихся, по-видимому, путем посттранскрипционного процессинга сплайсированного мРНК-предшественника. Данные свидетельствуют в пользу повышения биологического разнообразия продуктов гена *SERPINA1* за счет реализации альтернативных сайтов инициации трансляции, поскольку химерные продукты, синтезированные с главного и альтернативных стартовых кодонов, и соответствующие им предполагаемые эндогенные продукты трансляции локализуются в разных клеточных компартментах. Наши данные не только проливают свет на универсальность гена *SERPINA1*, но и дополняют представления об альтернативной инициации трансляции как регуляторном механизме экспрессии генов в целом. Дальнейшие исследования этого явления могут раскрыть новые функциональные роли укороченных изоформ ААТ и их участие в различных биологических процессах и/или развитии патологических состояний.

Выводы

1. Экспрессия длинных и предполагаемых коротких транскриптов, содержащих и не содержащих главную ОРС, соответственно, и содержание ААТ в секретах максимально в линии гепатоцеллюлярной карциномы HepG2, из линий не печеночного происхождения - в линии опухоли эпителия простаты DU145.

2. В препаратах суммарной РНК клеточных линий опухоли печени (HepG2), предстательной (DU145) и молочной железы (MDA-MB-231 и T47D) и тканей печени и почки детектируются длинные транскрипты. В HepG2 идентифицирован один, а в DU145 три варианта длинных мРНК, отличающихся структурой 5'-НТО. Длинные транскрипты преимущественно содержат укороченный вариант 3'-НТО.

3. С длинных транскриптов *SERPINA1* помимо секреторной транслируется внутриклеточная изоформа ААТ с внутреннего старт-кодона AUG-2 независимо от структуры 5'-НТО. 5'-НТО регулирует частоту инициации трансляции с главного AUG-1 и определяет соотношение обеих изоформ.

4. Химерный продукт трансляции с внутриклеточной изоформой ААТ мигрирует между ядром и цитоплазмой, в отличие от такового с секреторной, локализуя в ЭПР. Предполагаемая эндогенная внутриклеточная изоформа ААТ накапливается в ядерных спеклах.

5. В суммарной РНК HepG2 и тканей печени детектируются короткие транскрипты *SERPINA1*, содержащие экзон 5, генерируемые предположительно путем посттранскрипционного процессинга мРНК-предшественника. В HepG2 идентифицирована структура 5'-концов четырех вариантов коротких транскриптов, содержащих укороченную 3'-НТО.

6. Короткие транскрипты *SERPINA1* служат источником четырех альтернативных протеоформ ААТ, в том числе С-концевых пептидов, с предполагаемых стартов AUG-5, 6, 9 и 10. Химерные продукты, соответствующие пептидам ААТ, накапливаются в ядрах, предполагаемые эндогенные пептиды - в ядрышках.

ПРИЛОЖЕНИЕ

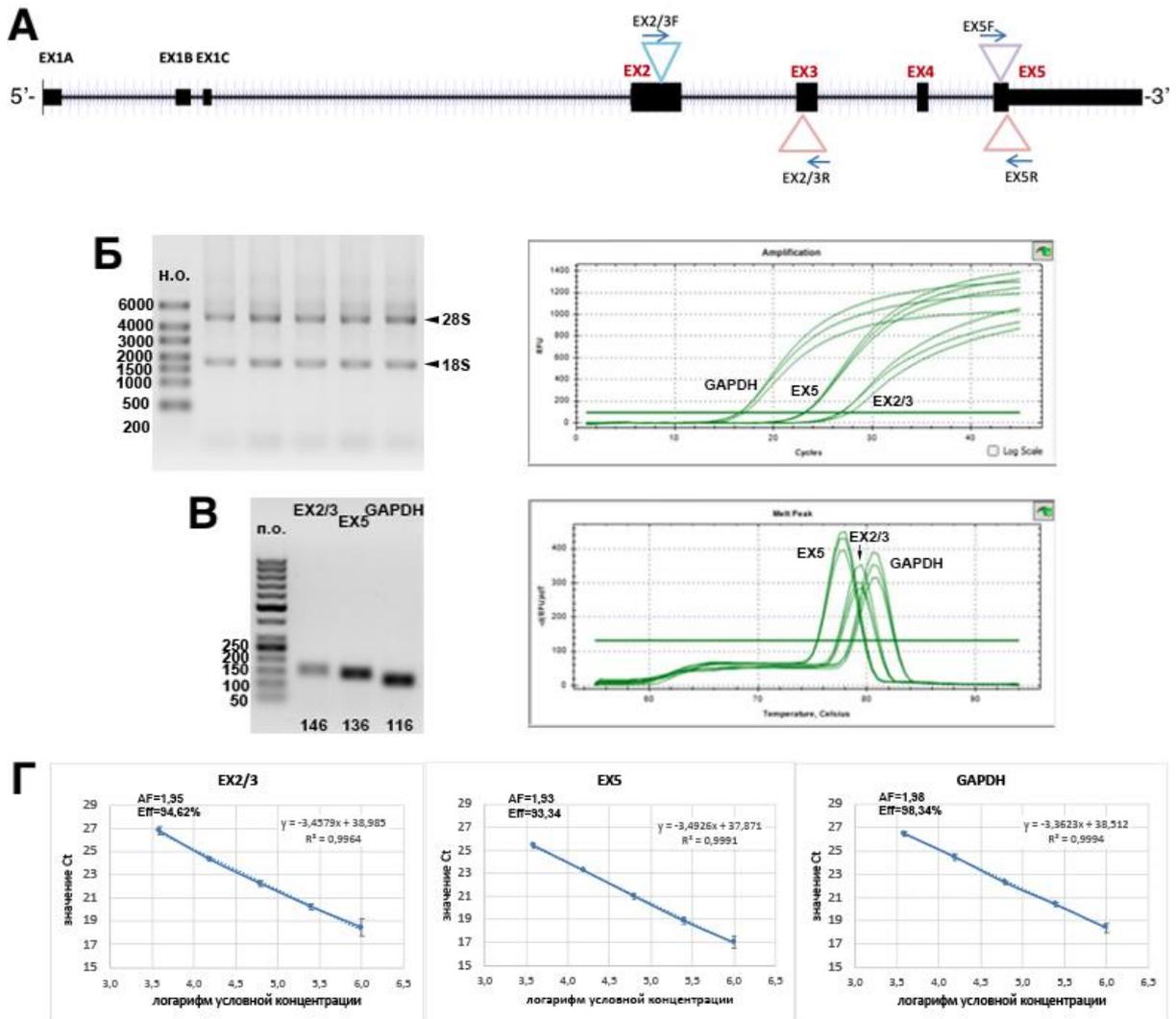


Рисунок 1П. ПЦР-РВ на разные кодирующие участки гена *SERPINA1*. А) Схематическое изображение положения праймеров. Б) слева - типичный скан агарозного геля после электрофоретического разделения препаратов суммарных РНК клеточных линий. справа - типичные кривые амплификации участков EX2/3, EX5 и GAPDH. В) слева - Электрофоретическое разделение продуктов амплификации, снизу подписаны ожидаемые длины ампликонов (п.н.); справа - кривые плавления продуктов амплификации. Г) Определение эффективности амплификации методом разведений. AF=фактор амплификации; Eff=эффективность амплификации.

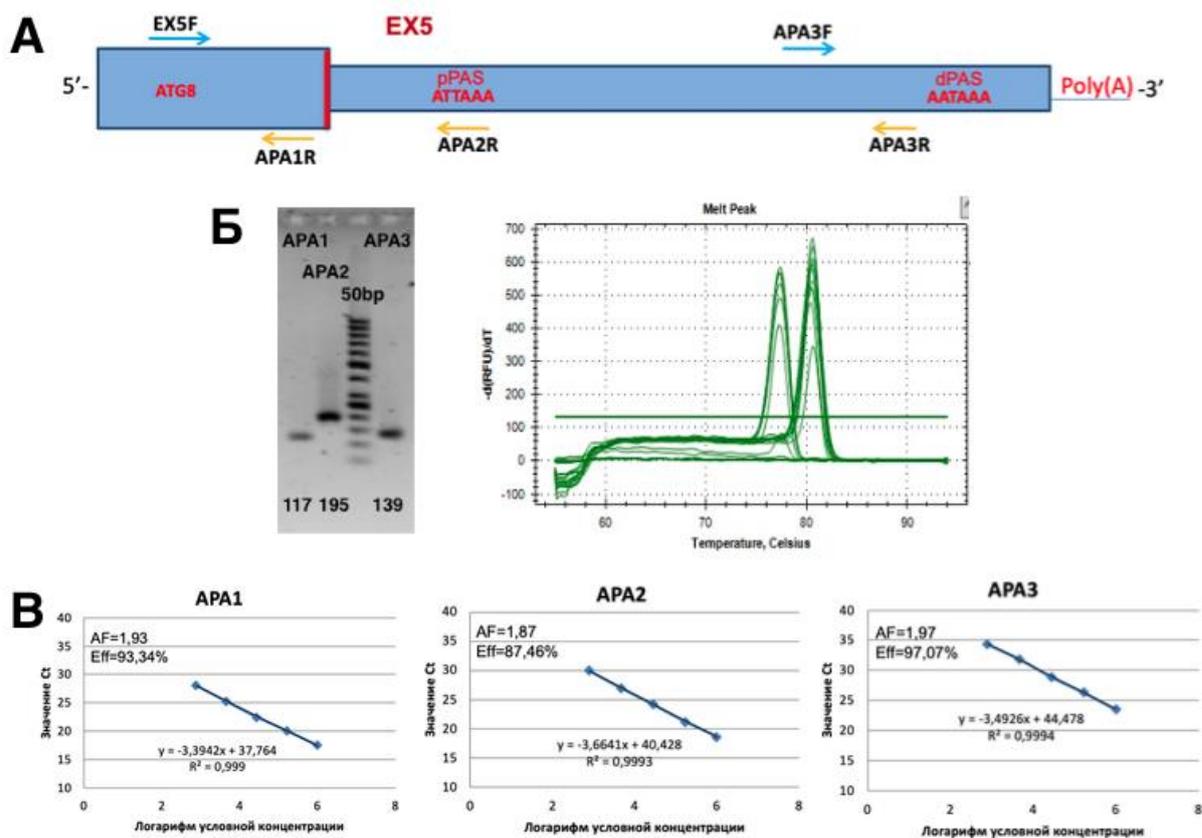


Рисунок 2П. ПЦР-РВ на разные сайты полиаденилирования (APA1-3), идентифицированные методом 3'-RACE. А) Схематическое изображение положения праймеров относительно экзона 5 гена *SERPINA1*. pPAS и dPAS - проксимальный и дистальный сигнал полиаденилирования. Б) слева - электрофоретическое разделение продуктов амплификации. Снизу отмечены ожидаемые длины ампликонов (п.н.). справа - кривые плавления продуктов амплификации. В) Определение эффективности амплификации методом разведений. AF=фактор амплификации; Eff=эффективность амплификации.

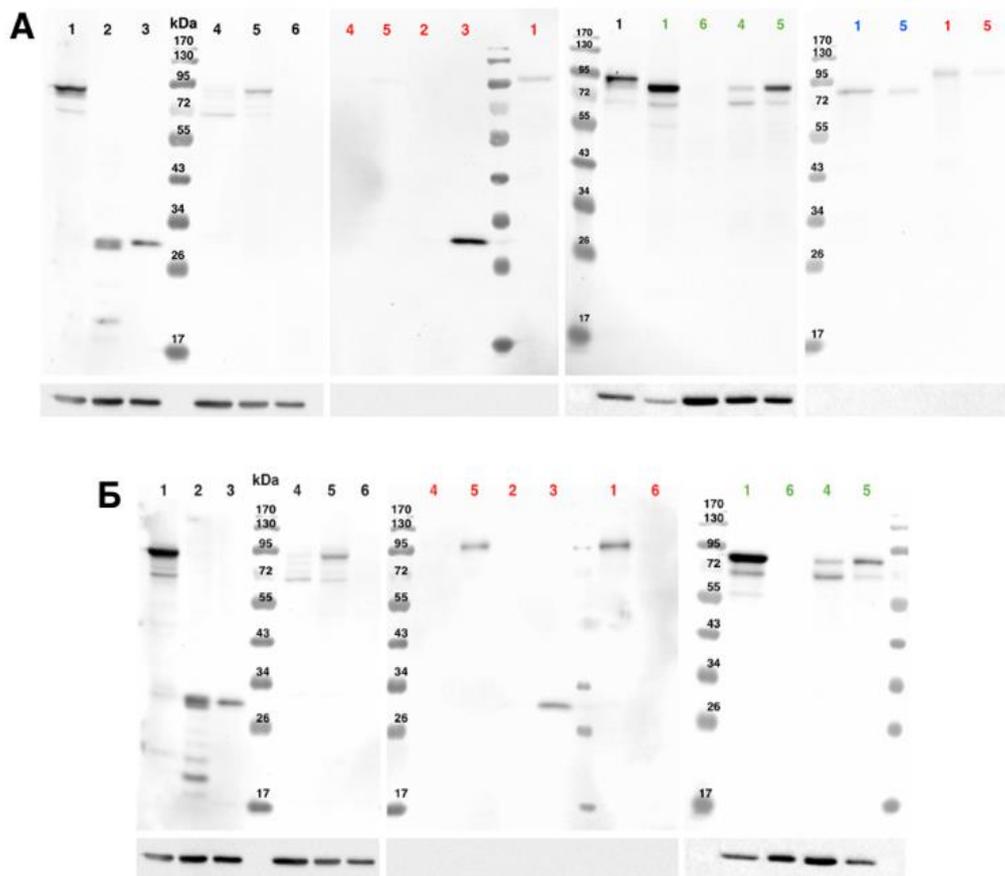


Рисунок 3П. Необрезанные изображения иммуноблоттинга из рисунка 11. А) лизаты eGFP⁺ клеток DU145 до (черные числа) и после (зеленые числа) обработки PNGaseF. Клетки были трансфицированы конструкциями: 1 - HLT; 2 - pEGFP-N1(ATG-); 3 - pEGFP-N1; 4 - DLT2; 5 - DLT1; 6 - отрицательный контроль без плазмиды. Красные и синие числа обозначают культурные супернатанты (секретомы) до и после обработки PNGaseF, соответственно. Сигналы проявлены с помощью поликлональных антител против GFP, конъюгированных с пероксидазой хрена. Нижняя панель - сигналы, проявленные антителами против альфа-тубулина. Б) тот же эксперимент по трансфекции на клеточной линии HepG2. Обозначения такие же, как в А).

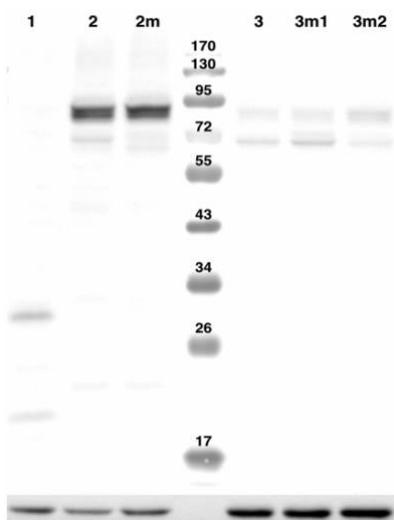


Рисунок 4П. Необрезанное изображение иммуноблоттинга из рисунка 12. Нанесены лизаты от равных количеств eGFP⁺ клеток DU145, полученные после трансфекции плазмидами: 1 - pEGFP-N1(ATG-), 2 - HLT, 2m - HLT с мутацией ATG-2>AAG, 3 - DLT2, 3m1 - DLT2 с оптимизированным контекстом Козак в ATG-2 (C>G в положении +4), 3m2 - DLT2 с мутацией uORF ATG>AAG, перекрывающей главную ORF *SERPINA1*. Сигналы проявлены антителами против GFP, конъюгированными с пероксидазой хрена. Нижняя панель - сигналы от альфа-тубулина.

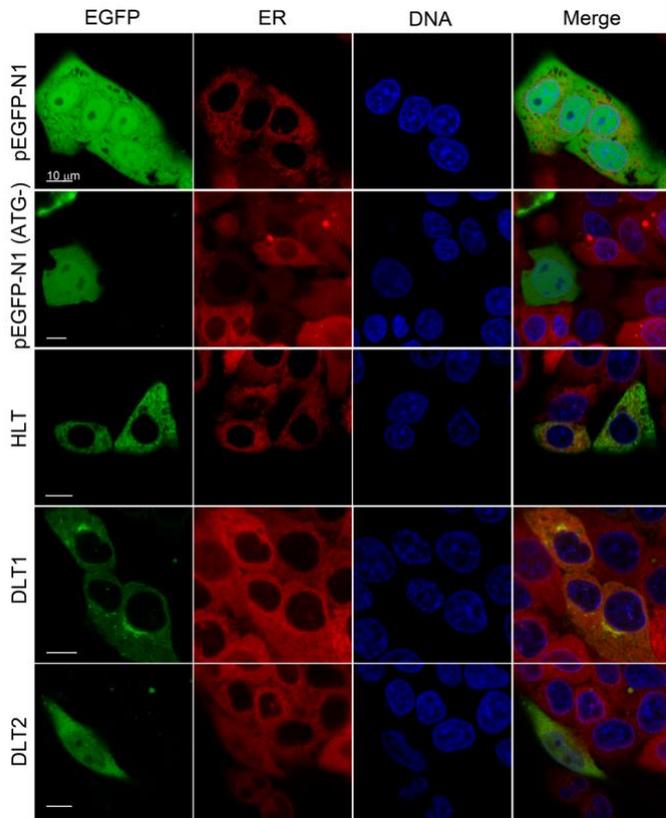


Рисунок 5П. Изучение внутриклеточной локализации слитых белков ААТ-eGFP в клетках HepG2 с помощью конфокальной микроскопии на живых клетках. Клетки трансфицировали конструкциями, экспрессирующими HLT, DLT1 и DLT2.

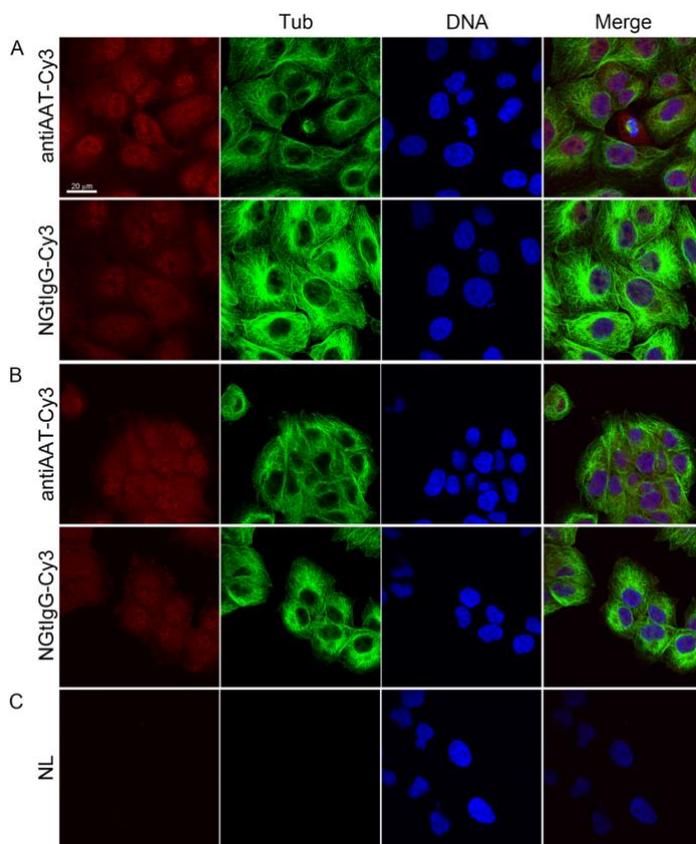


Рисунок 6П. Иммунофлуоресцентное окрашивание фиксированных клеток DU145 анти-ААТ антителами, мечеными Cy3 (antiAAT-Cy3). В качестве отрицательного контроля использовали нормальные козы антитела, меченые Cy3 (NGtIgG-Cy3). Образцы окрашены в неконкурентных условиях. А) Фиксированные клетки блокировали BSA-глицином в PBST; Б) Фиксированные клетки сначала блокировали как в А), затем предблокировали немечеными NGtIgG/BSA/PBST; В) фиксированные клетки без первичных антител, окрашены вторичными антителами DAMI-Alexa Fluor 488 (Donkey Anti-Mouse IgG), используемыми для первичных мышинных антител против альфа-тубулина.

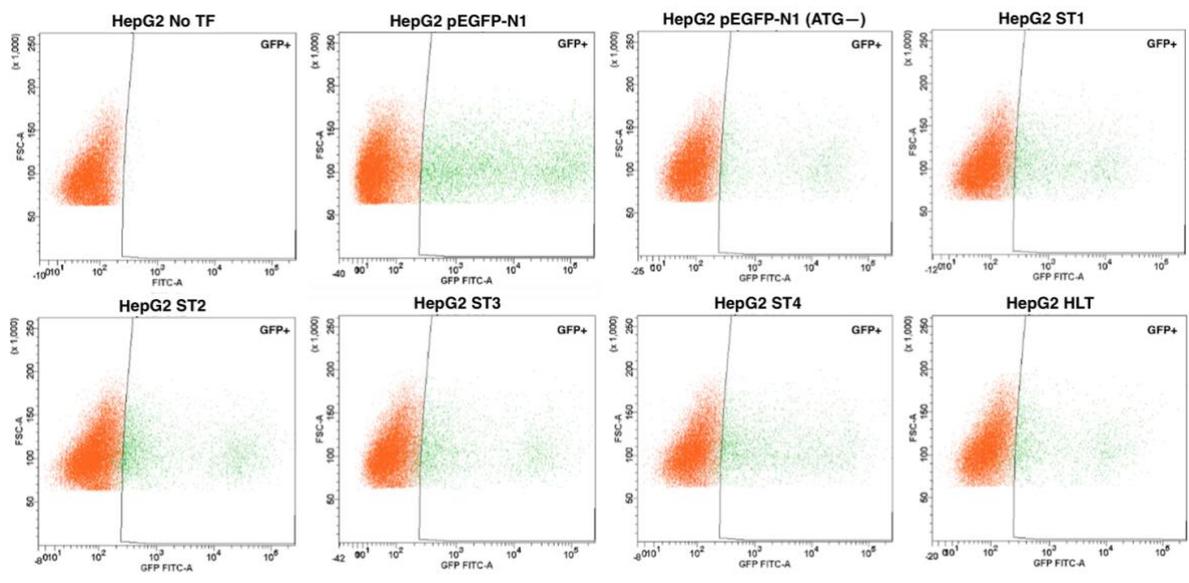


Рисунок 7П. Диаграммы рассеяния проточной цитофлуориметрии живых клеток HepG2, трансфицированных плазмидами pEGFP-N1, pEGFP-N1(ATG-), ST - плазмиды с клонированной кДНК, соответствующей коротким транскриптам ST1-4, HLT - кДНК, соответствующая длинному транскрипту HepG2. Гейт для отбора клеток был установлен на основе отрицательного контроля - клеток без трансфекции (No TF).

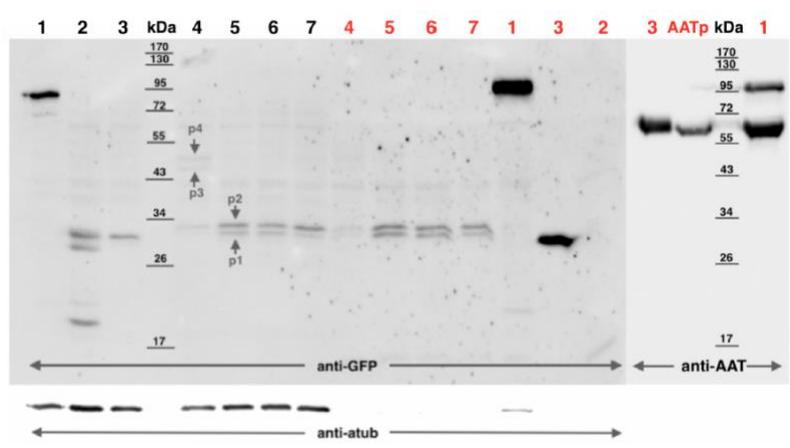


Рисунок 8П. Необрезанный скан иммуноблоттинга из рисунка 16. Клеточные лизаты (черные цифры) и секретомы (красный цифры) HepG2 после трансфекции: 1 - HLT, 2 - pEGFP-N1(ATG-), 3 - pEGFP-N1, 4-7 - ST1-4, соответственно. Четыре продукта, синтезируемые с коротких транскриптов, отмечены стрелками, начиная с самого короткого (p1) до наиболее длинного (p4). Сигналы

проявлены антителами, обозначенными для соответствующих дорожек снизу. Препарат ААТ из плазмы человека использован в качестве положительного контроля для антител против ААТ.

Model	2Δl	Df	P
M2-M1	11	2	4.087e-03
M7-M8	31,4	2	1.855e-07

Таблица 1П. Статистика отношения правдоподобия (2Δl) для теста положительного отбора.

Model	Likelihood	Average dN/dS	Parameters estimates		Positively selected sites
			Frequency	dN/dS	
M0, one-ratio	-7238,8	0,32516	p = 1	$\omega = 0.32516$	
M1, neutral	-7111,8	0,37843	p0: 0.73049 p1: 0.26951	ω 0: 0.14911 ω 1: 1.00000	
M2, selection	-7106,3	0,44322	p0: 0.72641 p1: 0.24668 p2: 0.02691	ω 0: 0.15801 ω 1: 1.00000 ω 2: 3.03810	101D, 108N, 205R, 240V, 249K, 308R, 329S, 333Q, 392E, 406Q
M3, discrete	-7099,5		p0: 0.44057 p1: 0.46233 p2: 0.09710	ω 0: 0.06485 ω 1: 0.44160 ω 2: 1.78380	2P, 7W, 21V, 39H, 114I, 128R, 150G, 155D, 183Q, 231E, 272L, 292N, 307D, 343G, 354A, 382L, 389I, 394K, 404I, 407N, 410S, 414M, 423K *43D,101D, 313S, 329S, 392E **108N, 205R, 240V, 249K, 308R, 333Q, 406Q
M7, beta	-7115,0		p = 0.47661 q = 0.88704		
M8, beta and ω	-7099,3		p0 = 0.92940 p1 = 0.07060 p = 0.72451 q = 1.79585	$\omega = 2.02592$	39H, 43D, 101D, 114I, 183Q, 240V, 249K, 313S, 329S, 392E, 394K, 410S *108N, 205R, 308R, 333Q, 406Q

Таблица 2П. Анализ эволюционного отбора с помощью моделей PAML. Положительно отобранные сайты обнаруживаются с апостериорной вероятностью $P > 0,5$ если не указано другого; * - апостериорная вероятность $0,99 > P > 0,95$; ** - апостериорная вероятность $P > 0,99$.

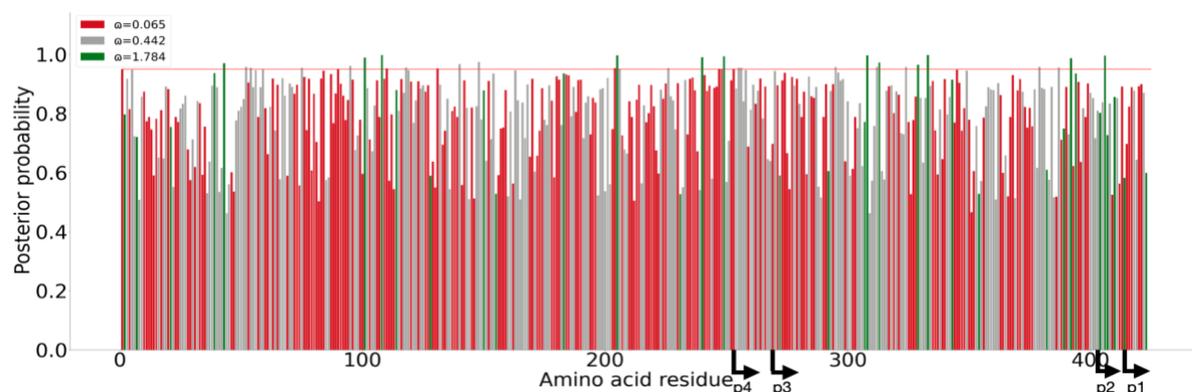


Рисунок 9П. Гистограмма, представляющая апостериорную вероятность для класса сайтов с наивысшими апостериорными вероятностями. Модель M3 предлагает три класса сайтов в пропорциях $p_0 = 0,441$, $p_1 = 0,462$ и $p_2 = 0,097$ с отношениями $\omega_0 = 0,065$, $\omega_1 = 0,442$ и $\omega_2 = 1,784$. Аминокислоты относятся к последовательности *Homo sapiens*. В позициях 44, 46, 47, 309, 351 имеются 5 пробелов из-за вставок в других последовательностях.

Mammalian species	Biological Process Ontology			Cellular Component ontology			Molecular Function Ontology		
	GO ID	GO Term	Reliability, %	GO ID	GO Term	Reliability, %	GO ID	GO Term	Reliability, %
Homo sapiens, Macaca mulatta, Gorilla gorilla, Hylobates moloch MIEQNTKSPLFMGK VVNPTQK	GO:1905117	regulation of ribonucleoside-diphosphate reductase activity	27	GO:0005634	nucleus	27	GO:1990846	ribonucleoside-diphosphate reductase inhibitor activity	27
	GO:0000077	DNA damage checkpoint	27	GO:0005737	cytoplasm	27			
	GO:0019222	regulation of metabolic process	27						
	GO:0007050	cell cycle arrest	27						
	GO:0043086	negative regulation of catalytic activity	27						
	GO:0006281	DNA repair	27						
Pongo abelii, Pan troglodytes MIEQNTKSPLFVGK VVNPTQK	GO:0006355	regulation of transcription, DNA-templated	27	GO:0005634	nucleus	27	GO:0003723	RNA binding	27
				GO:0005737	cytoplasm	27			
Papio anubis MIEQNTKSPLFIGK VVNPTQK	GO:0006355	regulation of transcription, DNA-templated	28	GO:0005737	cytoplasm	27	GO:0003723	RNA binding	28
				GO:0005634	nucleus	27			
Sus scrofa IYDTKTKAVLFMG KVMNPTQK	GO:0006412	translation	26	GO:0005840	ribosome	26	GO:0003735	structural constituent of ribosome	26

Canis lupus familiaris IVDRDTKSPLFMGK VVNPTQK	GO:0006412	translation	26	GO:0005840	ribosome	26	GO:0003735	structural constituent of ribosome	26
Rattus norvegicus IVESETQSPLFVGK IDPTR	GO:0034220	ion transmembrane transport	26	GO:0045202	synapse	26	GO:0005230	extracellular ligand-gated ion channel activity	26
				GO:0016021	integral component of membrane	26			
				GO:0005886	plasma membrane	26			
				GO:0045211	postsynaptic membrane	26			
Oryctolagus cuniculus IYSHEVKSPFVGK VVDPTQH	GO:0050896	response to stimulus	27	GO:0005615	extracellular space	27	GO:0001664	G protein-coupled receptor binding	27
	GO:0098657	import into cell	27	GO:0005576	extracellular region	27	GO:0051219	phosphoprotein binding	27
	GO:0009968	negative regulation of signal transduction	27				GO:0002046	opsin binding	27
	GO:0001932	regulation of protein phosphorylation	27				GO:0019904	protein domain specific binding	27
	GO:0007165	signal transduction	27						
	GO:0007601	visual perception	27						
	GO:0002031	G protein-coupled receptor internalization	27						
Mus musculus SERPINA1a IFEHTQSPIFLGKV	GO:0034220	ion transmembrane transport	27	GO:0005886	plasma membrane	27	GO:0005230	extracellular ligand-gated ion channel activity	27

VDPTHK				GO:0016021	integral component of membrane	27			
				GO:0045202	synapse	27			
				GO:0045211	postsynaptic membrane	27			
Mus musculus SERPINA1b IFEHTQSPIFVGKV VDPTHK	GO:0034220	ion transmembrane transport	27	GO:0016021	integral component of membrane	27	GO:0005230	extracellular ligand-gated ion channel activity	27
				GO:0005886	plasma membrane	27			
				GO:0045211	postsynaptic membrane	27			
				GO:0045202	synapse	27			
Bos taurus, Ovis aries LYDRNTKSPLFVGK VVNPTQA	GO:0034220	ion transmembrane transport	27	GO:0016021	integral component of membrane	27	GO:0005230	extracellular ligand-gated ion channel activity	27
				GO:0005886	plasma membrane	27			
				GO:0045211	postsynaptic membrane	27			
				GO:0045202	synapse	27			

Таблица 3П. Предсказание свойств и внутриклеточной локализации С21 пептидов ААТ млекопитающих с помощью онлайн программы PredictProtein (<https://predictprotein.org>) [434].

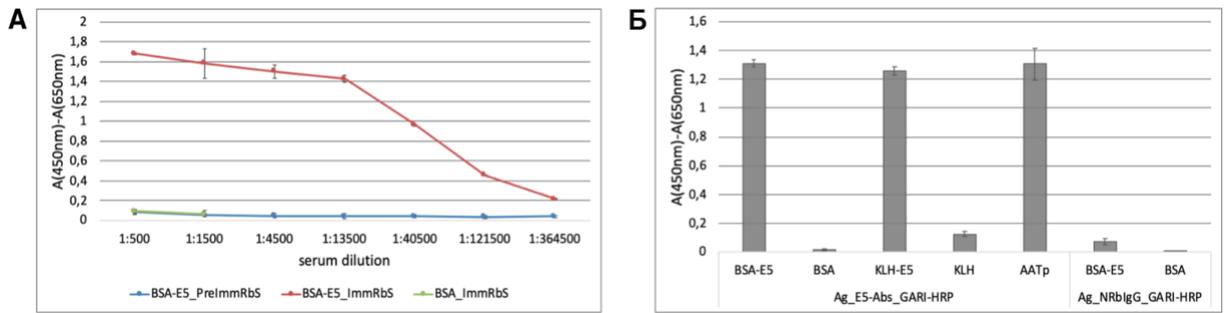


Рисунок 10П. Проверка полученных кроличьих поликлональных антител против E5-эпитопа ААТ. **А)** Непрямой ИФА на связывание IgG из преиммунной и иммунной кроличьей сыворотки (PreImmRbS и ImmRbS, соответственно) с иммобилизованным антигеном BSA-пептид E5 (200 нг конъюгата в лунку). Титр иммунной сыворотки оценен как минимум 1:364500. Отрицательный контроль - связывание антител из иммунной сыворотки при минимальном ее разведении с чистым BSA. **Б)** Непрямой ИФА на связывание аффинно-очищенных анти-E5 антител (2 мкг/мл) с E5-эпитопом в составе иммобилизованных конъюгатов BSA-E5 и KLH-E5 (по 200 нг/лунку) и ААТ из плазмы человека (ААТр) (1мкг/лунку). Отрицательный контроль - нормальные кроличьи антитела (NRbIgG). Ag=антиген. Данные представлены как среднее±стандартное отклонение по трем повторам.

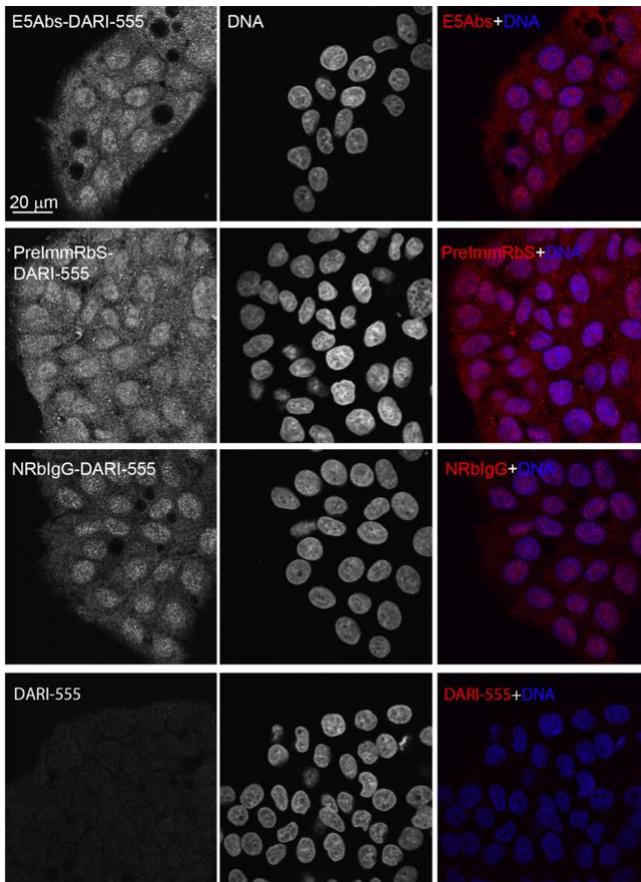


Рисунок 11П. Тестирование анти-E5 антител в непрямом иммунофлуоресцентном анализе. Отрицательные контроли: неиммунная сыворотка кролика (PreImmRbS) и нормальные кроличьи антитела (NRbIgG) показали аналогичный паттерн распределения, как и специфические антитела. Отрицательный контроль с использованием вторичных антител (DARI-Alexa Fluor-555) предполагает, что кроличьи IgG неспецифически “липнут” к клеткам HepG2.

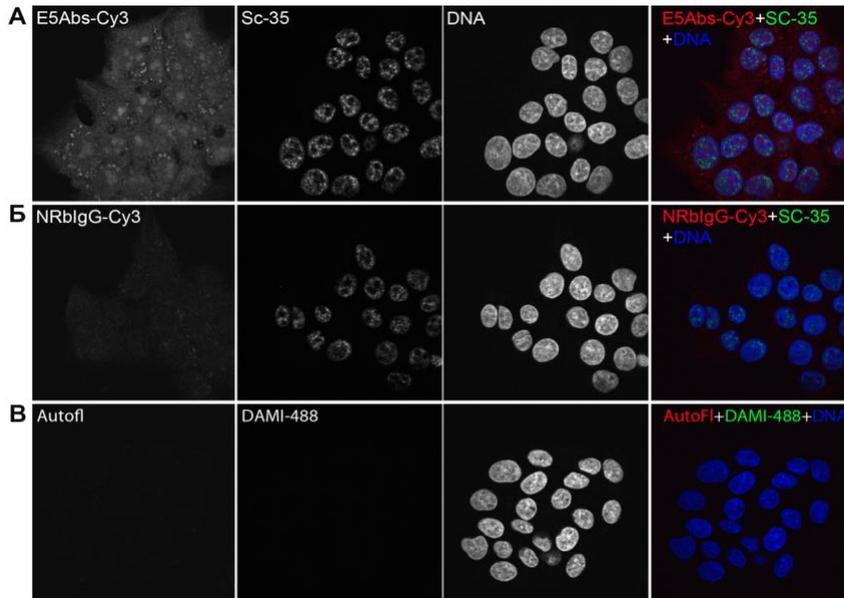


Рисунок 12П. Прямой иммунофлуоресцентный анализ эндогенных E5-эпитопов ААТ с использованием кроличьих антител против E5-эпитопа (E5-Abs-Cy3) (А), нормальных кроличьих антител (NRbIgG-Cy3) в качестве отрицательного контроля (Б). В) Контроль без первичных антител. Фиксированные клетки предварительно инкубировали с немеченой неиммунной сывороткой кролика для блокирования неспецифического связывания. Антитела против

маркера ядерных спеклов Sc-35 в паре со вторичными антителами (DAMI-Alexa Fluor-488) использовали для демонстрации локализации E5-эпитопов вне ядерных спеклов, в отличие от таковой у предполагаемой внутриклеточной изоформы ААТ, транскрируемой с длинных транскриптов *SERPINA1* (см. рисунок 4.14 раздела 4.5.4.2).

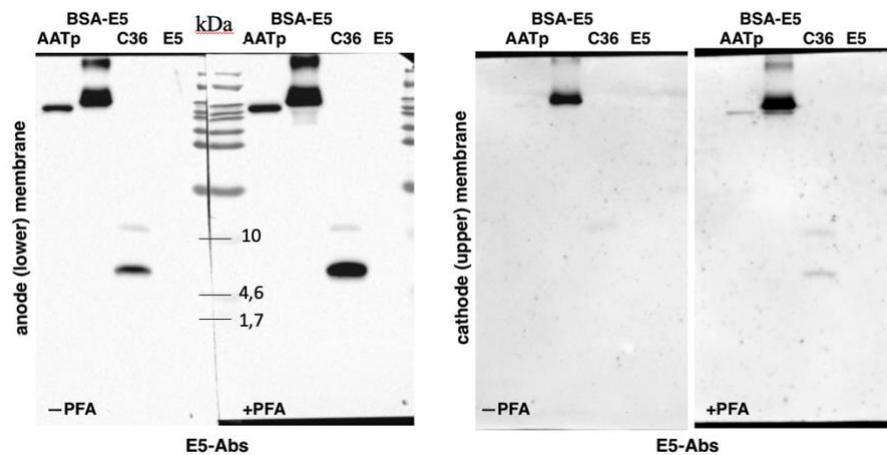


Рисунок 13П. Детекция E5-эпитопа с помощью анти-E5 антител (E5-Abs) в структуре ААТ из плазмы (ААТр), конъюгата BSA с синтетическим пептидом E5 (BSA-E5), синтетических пептидов C36 и E5 (=C10). Для разделения образцов использовали электрофорез в полиакриламидном геле (16%) в трис-трициновой системе в присутствии SDS. После переноса белков на мембраны, часть мембран после трехкратной отмычки в PBS были префиксированы 0,5 % раствором параформальдегида (PFA) в PBS в течение 30 минут для минимизации вымывания пептидов. Пептид E5 не детектируется ни на нижней (анодной), ни на верхней (катодной) нитроцеллюлозной мембране, в отличие от пептида C36.

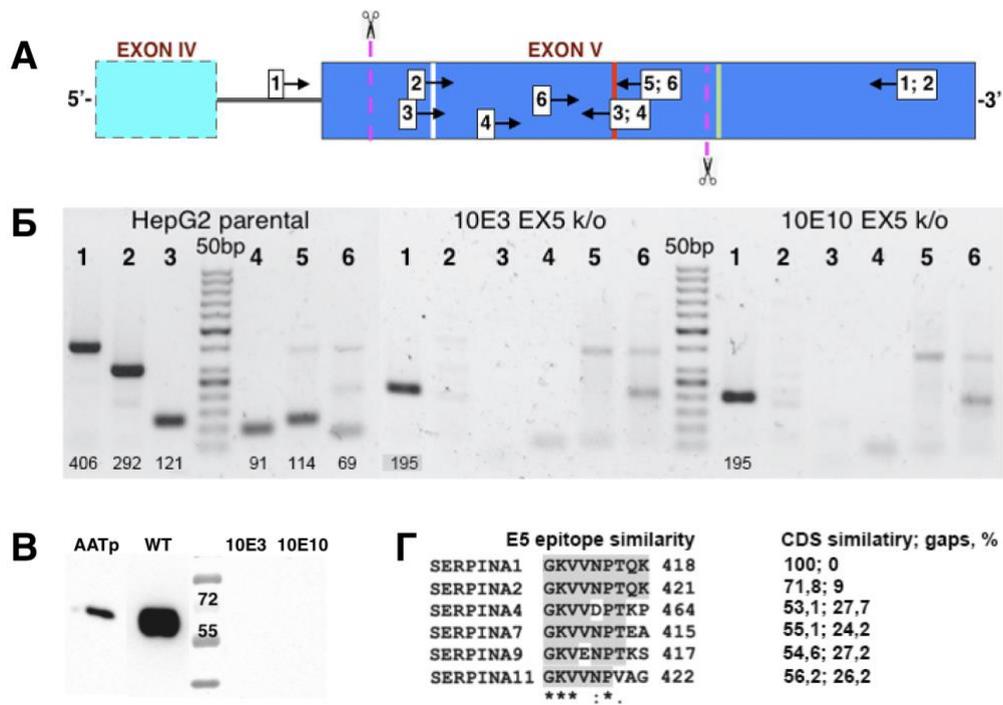


Рисунок 14П. Верификация сублиний НерG2 с нокаутом экзона 5 (EX5) *SERPINA1*. А) Схематическое изображение мест отжига используемых праймеров. Черными стрелками обозначены прямые и обратные праймеры, цифрами, соответствующими дорожкам на скане геля в Б), обозначены пары праймеров, используемые в реакции ПЦР; вертикальные малиновые штриховые линии - сайты редактирования; вертикальная белая линия - ATG-8; вертикальная красная линия - стоп-кодон; вертикальная светло-зеленая линия - проксимальный сигнал полиаденилирования. Б) Скан геля с продуктами амплификации на геномной ДНК. EX5 k/o - нокауты по пятому экзону гена *SERPINA1*. Цифрами снизу указаны ожидаемые длины ампликонов в п.н.; где не указано, продуктов амплификации не ожидалось. В) Иммуноблоттинг секретомов родительских клеток НерG2 (WT) и сублиний с нокаутом по EX5. Сигналы проявлены с использованием антител против полноразмерного ААТ. ААТр - ААТ, выделенный из плазмы крови. Г) Пять наиболее близких паралогов *SERPINA1*, включая псевдоген *SERPINA2* [457]. Слева - выравнивание С-концевых последовательностей продуктов генов-паралогов по E5 эпипептиду ААТ; справа - гомология (%) и пробелы (%) в последовательности кодирующих областей (CDS) генов-паралогов в сравнении с *SERPINA1*.

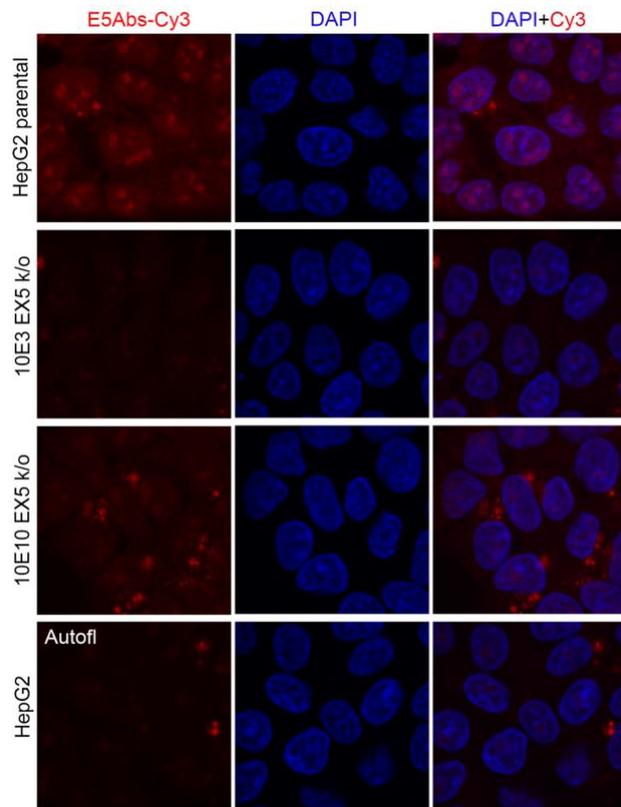


Рисунок 15П. Верификация предполагаемых С-концевых пептидов ААТ на происхождение из EX5 гена *SERPINA1* с помощью иммунофлуоресценции на препаратах фиксированных родительских клеток HepG2 и клонов с нокаутом EX5 *SERPINA1*. Клетки мечены конъюгатом первичных антител-Сy3 против E5-эпитопа после предварительной блокировки неиммунной сывороткой кролика. Клоны с нокаутом фиксированы на раннем пассаже.

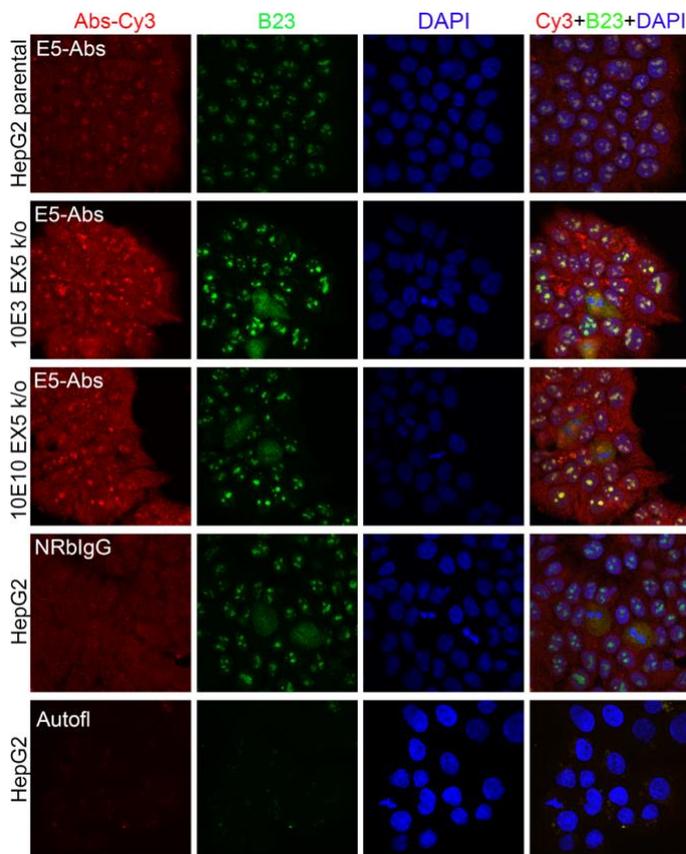


Рисунок 16П. Верификация предполагаемых С-концевых пептидов ААТ на происхождение из EX5 гена *SERPINA1* с помощью иммунофлуоресценции на препаратах фиксированных родительских клеток HepG2 и клонов с нокаутом EX5 *SERPINA1*. Клетки мечены конъюгатами первичных антител-Сy3 против E5-эпитопа (либо конъюгатом нормальных кроличьих антител NRbIgG в качестве отрицательного контроля) после предварительной блокировки неиммунной сывороткой кролика. Клоны с нокаутом фиксированы на более позднем пассаже, чем на Рисунке 13П, с разницей в один месяц.

Список литературы

1. Jonigk D, Al-Omari M, Maegel L, et al. Anti-inflammatory and immunomodulatory properties of α 1-antitrypsin without inhibition of elastase. *Proc Natl Acad Sci*. 2013;110(37):15007-15012. doi:10.1073/pnas.1309648110
2. Ehlers MR. Immune-modulating effects of alpha-1 antitrypsin. *Biol Chem*. 2014;395(10):1187-1193. doi:10.1515/hsz-2014-0161
3. Petrache I, Fijalkowska I, Medler TR, et al. α -1 Antitrypsin Inhibits Caspase-3 Activity, Preventing Lung Endothelial Cell Apoptosis. *Am J Pathol*. 2006;169(4):1155-1166. doi:10.2353/ajpath.2006.060058
4. Schwarz N, Tumpara S, Wrenger S, et al. Alpha1-antitrypsin protects lung cancer cells from staurosporine-induced apoptosis: the role of bacterial lipopolysaccharide. *Sci Rep*. 2020;10(1):9563. doi:10.1038/s41598-020-66825-w
5. Shapiro L, Pott GB, Ralston AH. Alpha-1-antitrypsin inhibits human immunodeficiency virus type 1. *FASEB J*. 2001;15(1):115-122. doi:10.1096/fj.00-0311com
6. Zhou X, Liu Z, Zhang J, Adelsberger JW, Yang J, Burton GF. Alpha-1-antitrypsin interacts with gp41 to block HIV-1 entry into CD4+ T lymphocytes. *BMC Microbiol*. 2016;16(1):172. doi:10.1186/s12866-016-0751-2
7. Wettstein L, Weil T, Conzelmann C, et al. Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection. *Nat Commun*. 2021;12(1):1726. doi:10.1038/s41467-021-21972-0
8. Zsila F. Inhibition of heat- and chemical-induced aggregation of various proteins reveals chaperone-like activity of the acute-phase component and serine protease inhibitor human α 1-antitrypsin. *Biochem Biophys Res Commun*. 2010;393(2):242-247. doi:10.1016/j.bbrc.2010.01.110
9. Niemietz C, Fleischhauer L, Sandfort V, Guttman S, Zibert A, Schmidt HH-J. Hepatocyte-like cells reveal novel role of SerpinA1 in transthyretin amyloidosis. *J Cell Sci*. Published online January 1, 2018. doi:10.1242/jcs.219824
10. Ferrarotti I, Thun GA, Zorzetto M, et al. Serum levels and genotype distribution of α 1 -antitrypsin in the general population. *Thorax*. 2012;67(8):669-674. doi:10.1136/thoraxjnl-2011-201321
11. de Serres F, Blanco I. Role of alpha-1 antitrypsin in human health and disease. *J Intern Med*. 2014;276(4):311-335. doi:10.1111/joim.12239
12. Yamashita T, Hashimoto S, Kaneko S, et al. Comprehensive Gene Expression Profile of a Normal Human Liver. *Biochem Biophys Res Commun*. 2000;269(1):110-116. doi:10.1006/bbrc.2000.2272
13. Perlmutter DH, Cole FS, Kilbridge P, Rossing TH, Colten HR. Expression of the alpha 1-proteinase inhibitor gene in human monocytes and macrophages. *Proc Natl Acad Sci*. 1985;82(3):795-799. doi:10.1073/pnas.82.3.795
14. du Bois RM, Bernaudin JF, Paakko P, et al. Human neutrophils express the alpha 1-antitrypsin gene and produce alpha 1-antitrypsin. *Blood*. 1991;77(12):2724-2730.
15. Hu C, Perlmutter DH. Cell-specific involvement of HNF-1 β in α 1 -antitrypsin gene expression in human respiratory epithelial cells. *Am J Physiol Cell Mol Physiol*. 2002;282(4):L757-L765. doi:10.1152/ajplung.00271.2001
16. Geboes K, Ray MB, Rutgeerts P, Callea F, Desmet VJ, Vantrappen G. Morphological identification of alpha-I-antitrypsin in the human small intestine. *Histopathology*. 1982;6(1):55-60. doi:10.1111/j.1365-2559.1982.tb02701.x
17. Bošković G, Twining SS. Local control of α 1-proteinase inhibitor levels: regulation of α 1-proteinase inhibitor in the human cornea by growth factors and cytokines. *Biochim Biophys Acta - Mol Cell Res*. 1998;1403(1):37-46. doi:10.1016/S0167-4889(98)00018-4

18. Laurell C-B, Eriksson S. The Electrophoretic α_1 -Globulin Pattern of Serum in α_1 - Antitrypsin Deficiency. *Scand J Clin Lab Invest.* 1963;15(2):132-140. doi:10.1080/00365516309051324
19. Lomas DA, Hurst JR, Gooptu B. Update on alpha-1 antitrypsin deficiency: New therapies. *J Hepatol.* 2016;65(2):413-424. doi:10.1016/j.jhep.2016.03.010
20. Sandström CS, Ohlsson B, Melander O, Westin U, Mahadeva R, Janciauskiene S. An association between type 2 diabetes and α_1 -antitrypsin deficiency. *Diabet Med.* Published online September 2008. doi:10.1111/j.1464-5491.2008.02584.x
21. Ghoreishi A, Mahmoodi M, Khoshdel A. Comparison of alpha 1- antitrypsin activity and phenotype in type 1 diabetic patients to healthy individuals. *J Fam Med Prim Care.* 2022;11(4):1377. doi:10.4103/jfmprc.jfmprc_905_21
22. Eriksson S, Carlson J, Velez R. Risk of Cirrhosis and Primary Liver Cancer in Alpha 1 - Antitrypsin Deficiency. *N Engl J Med.* 1986;314(12):736-739. doi:10.1056/NEJM198603203141202
23. Rudnick DA, Perlmutter DH. Alpha-1-antitrypsin deficiency: A new paradigm for hepatocellular carcinoma in genetic liver disease. *Hepatology.* 2005;42(3):514-521. doi:10.1002/hep.20815
24. Hiller A-M, Ekström M, Piitulainen E, Lindberg A, Rönmark E, Tanash H. Cancer risk in severe alpha-1-antitrypsin deficiency. *Eur Respir J.* 2022;60(4):2103200. doi:10.1183/13993003.03200-2021
25. Yang P. Alpha1-Antitrypsin Deficiency Carriers, Tobacco Smoke, Chronic Obstructive Pulmonary Disease, and Lung Cancer Risk. *Arch Intern Med.* 2008;168(10):1097. doi:10.1001/archinte.168.10.1097
26. Benkmann H-G, Hanssen P, Ovenbeck R, Goedde W. Distribution of Alpha-1-Antitrypsin and Haptoglobin Phenotypes in Bladder Cancer Patients. *Hum Hered.* 1987;37(5):290-293. doi:10.1159/000153720
27. Callea F, Stuyck JM, Massi G, et al. Alpha-1-Antitrypsin (AAT) Deposits in Gall Bladder Adenocarcinoma and Liver in Partial AAT Deficiency (Pi SZ Phenotype). *Am J Clin Pathol.* 1982;78(6):878-883. doi:10.1093/ajcp/78.6.878
28. Yang P, Cunningham JM, Halling KC, et al. Higher Risk of Mismatch Repair-Deficient Colorectal Cancer in α_1 -Antitrypsin Deficiency Carriers and Cigarette Smokers. *Mol Genet Metab.* 2000;71(4):639-645. doi:10.1006/mgme.2000.3089
29. Lindor NM, Yang P, Evans I, et al. Alpha-1-antitrypsin deficiency and smoking as risk factors for mismatch repair deficient colorectal cancer: A study from the colon cancer family registry. *Mol Genet Metab.* 2010;99(2):157-159. doi:10.1016/j.ymgme.2009.09.010
30. Callea F, Massi G, De Wolf-Peeters C, Lievens C, Desmet VJ. Alpha-1-antitrypsin phenotypes in malignant lymphoma. *J Clin Pathol.* 1982;35(11):1213-1215. doi:10.1136/jcp.35.11.1213
31. Daddi G, Mancini PA, Parola D, Contini A. Alfa-antitrypsin increase in lung cancer. *Boll Ist Sieroter Milan.* 1976;55(6):510-512.
32. Dokrunova AA, Sokolova OS. [Serum A1-antitrypsin levels in prostate pathology]. *Urologiia.* 2012;(5):77-80.
33. Hsu P-I, Chen C-H, Hsiao M, et al. Diagnosis of Gastric Malignancy Using Gastric Juice α_1 -antitrypsin. *Cancer Epidemiol Biomarkers Prev.* 2010;19(2):405-411. doi:10.1158/1055-9965.EPI-09-0609
34. Pérez-Holanda S, Blanco I, Menéndez M, Rodrigo L. Serum concentration of alpha-1 antitrypsin is significantly higher in colorectal cancer patients than in healthy controls. *BMC Cancer.* 2014;14(1):355. doi:10.1186/1471-2407-14-355
35. Bujanda L, Sarasqueta C, Cosme A, et al. Evaluation of Alpha 1-Antitrypsin and the Levels of mRNA Expression of Matrix Metalloproteinase 7, Urokinase Type Plasminogen Activator Receptor and COX-2 for the Diagnosis of Colorectal Cancer. Moschetta A, ed. *PLoS One.* 2013;8(1):e51810. doi:10.1371/journal.pone.0051810

36. Kataoka H, Nabeshima K, Komada N, Koono M. New human colorectal carcinoma cell lines that secrete proteinase inhibitors in vitro. *Virchows Arch B Cell Pathol Incl Mol Pathol*. 1989;57(1):157-165. doi:10.1007/BF02899077
37. Kataoka H, Seguchi K, Inoue T, Koono M. Properties of α 1 -antitrypsin secreted by human adenocarcinoma cell lines. *FEBS Lett*. 1993;328(3):291-295. doi:10.1016/0014-5793(93)80946-R
38. Maslakova AA, Telkov M V., Orlovsky I V., Sokolova OS. Comparative analysis of *SERPINA1* gene expression in tumor cell lines. *Moscow Univ Biol Sci Bull*. 2015;70(3):127-131. doi:10.3103/S0096392515030086
39. Shapira MG, Khalfin B, Lewis EC, Parola AH, Nathan I. Regulation of Autophagy by α 1-Antitrypsin: “A Foe of a Foe Is a Friend.” *Mol Med*. 2014;20(1):417-426. doi:10.2119/molmed.2014.00054
40. Normandin K, Péant B, Le Page C, et al. Protease inhibitor SERPINA1 expression in epithelial ovarian cancer. *Clin Exp Metastasis*. 2010;27(1):55-69. doi:10.1007/s10585-009-9303-6
41. Perlino E, Cortese R, Ciliberto G. The human alpha 1-antitrypsin gene is transcribed from two different promoters in macrophages and hepatocytes. *EMBO J*. 1987;6(9):2767-2771. doi:10.1002/j.1460-2075.1987.tb02571.x
42. Marsden, Mark D; Fournier REK. Organization and expression of the human serpin gene cluster at 14q32.1. *Front Biosci*. 2005;10(1-3):1768. doi:10.2741/1660
43. Lackey L, McArthur E, Laederach A. Increased Transcript Complexity in Genes Associated with Chronic Obstructive Pulmonary Disease. Eickelberg O, ed. *PLoS One*. 2015;10(10):e0140885. doi:10.1371/journal.pone.0140885
44. Corley M, Solem A, Phillips G, et al. An RNA structure-mediated, posttranscriptional model of human α -1-antitrypsin expression. *Proc Natl Acad Sci*. 2017;114(47). doi:10.1073/pnas.1706539114
45. Cercek L, Cercek B. Cancer-associated SCM-recognition, immunodefense suppression, and serine protease protection peptide. Part I. Isolation, amino acid sequence, homology, and origin. *Cancer Detect Prev*. 1992;16(5-6):305-319.
46. Cercek L, Cercek B. Cancer-associated SCM-recognition, immunodefense suppression, and serine protease protection peptide. Part II. Immunodefense suppressive effects of the CRISPPs peptide. *Cancer Detect Prev*. 1993;17(3):433-445.
47. Cercek L, Cercek B. Cancer-associated SCM-recognition, immunodefense suppression, and serine protease protection peptide. Part III. CRISPP peptide protection of serine proteases against inhibition. *Cancer Detect Prev*. 1993;17(3):447-454.
48. Niemann MA, Narkates AJ, Miller EJ. Isolation and Serine Protease Inhibitory Activity of the 44-Residue, C-Terminal Fragment of α 1 -Antitrypsin from Human Placenta. *Matrix*. 1992;12(3):233-241. doi:10.1016/S0934-8832(11)80066-1
49. Niemann MA, Baggott JE, Miller EJ. Binding of SPAAT, the 44-residue C-terminal peptide of alpha 1-antitrypsin, to proteins of the extracellular matrix. *J Cell Biochem*. 1997;66(3):346-357.
50. Niemann MA, Baggott JE, Miller EJ. Inhibition of human serine proteases by SPAAT, the C-terminal 44-residue peptide from α 1-antitrypsin. *Biochim Biophys Acta - Protein Struct Mol Enzymol*. 1997;1340(1):123-130. doi:10.1016/S0167-4838(97)00034-4
51. Münch J, Ständker L, Adermann K, et al. Discovery and Optimization of a Natural HIV-1 Entry Inhibitor Targeting the gp41 Fusion Peptide. *Cell*. 2007;129(2):263-275. doi:10.1016/j.cell.2007.02.042
52. Dichtl W, Moraga F, Ares MP., et al. The Carboxyl-Terminal Fragment of α 1-Antitrypsin Is Present in Atherosclerotic Plaques and Regulates Inflammatory Transcription Factors in Primary Human Monocytes. *Mol Cell Biol Res Commun*. 2000;4(1):50-61. doi:10.1006/mcbr.2000.0256
53. Chang W-C, Hsu P-I, Chen Y-Y, Hsiao M, Lu P-J, Chen C-H. Observation of peptide differences between cancer and control in gastric juice. *PROTEOMICS – Clin Appl*. 2008;2(1):55-62. doi:10.1002/prca.200780066

54. Zhou J, Trock B, Tsangaris TN, et al. A unique proteolytic fragment of alpha1-antitrypsin is elevated in ductal fluid of breast cancer patient. *Breast Cancer Res Treat.* 2010;123(1):73-86. doi:10.1007/s10549-009-0625-5
55. Blaurock N, Schmerler D, Hünninger K, et al. C-Terminal Alpha-1 Antitrypsin Peptide: A New Sepsis Biomarker with Immunomodulatory Function. *Mediators Inflamm.* 2016;2016:1-13. doi:10.1155/2016/6129437
56. Yost CC, Schwertz H, Cody MJ, et al. Neonatal NET-inhibitory factor and related peptides inhibit neutrophil extracellular trap formation. *J Clin Invest.* 2016;126(10):3783-3798. doi:10.1172/JCI83873
57. Rhee M, Davis P. Mechanism of Uptake of C105Y, a Novel Cell-penetrating Peptide. *J Biol Chem.* 2006;281(2):1233-1240. doi:10.1074/jbc.M509813200
58. Congote LF, Temmel N. The C-terminal 26-residue peptide of serpin A1 stimulates proliferation of breast and liver cancer cells: role of protein kinase C and CD47. *FEBS Lett.* 2004;576(3):343-347. doi:10.1016/j.febslet.2004.09.035
59. Fermi C, Pernossi L. Ueber die Enzyme. *Zeitschrift für Hyg und Infekt.* 1894;18(1):83-127. doi:10.1007/BF02216836
60. Hedin SG. An Antitryptic Effect of Charcoal and a Comparison between the Action of Charcoal and that of the Tryptic Antibody in the Serum. *Biochem J.* 1906;1(10):484-495. doi:10.1042/bj0010484
61. Hedin SG. Trypsin and Antitrypsin. *Biochem J.* 1906;1(10):474-483. doi:10.1042/bj0010474
62. Hussey RG, Northrop JH. A study of the equilibrium between the so called "antitrypsin" of the blood and trypsin. *J Gen Physiol.* 1923;5(3):335-351. doi:10.1085/jgp.5.3.335
63. Schultze HE., Göllner I, Heide K, Schönenberger M, Schwick G. Zur Kenntnis der α -Globuline des menschlichen Normalserums. *Zeitschrift für Naturforsch B.* 1955;10(8):463-473. doi:10.1515/znb-1955-0810
64. Schultze HE, Heide K, Haupt H. α 1-Antitrypsin aus Humanserum. *Klin Wochenschr.* 1962;40(8):427-429. doi:10.1007/BF01732842
65. Potempa J, Korzus E, Travis J. The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J Biol Chem.* 1994;269(23):15957-15960.
66. Carrell R, Travis J. α 1-Antitrypsin and the serpins: variation and countervariation. *Trends Biochem Sci.* 1985;10(1):20-24. doi:10.1016/0968-0004(85)90011-8
67. Huber R, Carrell RW. Implications of the three-dimensional structure of .alpha.1-antitrypsin for structure and function of serpins. *Biochemistry.* 1989;28(23):8951-8966. doi:10.1021/bi00449a001
68. Hunt LT, Dayhoff MO. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha1-proteinase inhibitor. *Biochem Biophys Res Commun.* 1980;95(2):864-871. doi:10.1016/0006-291X(80)90867-0
69. Irving JA. Phylogeny of the Serpin Superfamily: Implications of Patterns of Amino Acid Conservation for Structure and Function. *Genome Res.* 2000;10(12):1845-1864. doi:10.1101/gr.GR-1478R
70. Irving JA, Steenbakkens PJM, Lesk AM, Op den Camp HJM, Pike RN, Whisstock JC. Serpins in Prokaryotes. *Mol Biol Evol.* 2002;19(11):1881-1890. doi:10.1093/oxfordjournals.molbev.a004012
71. Law RHP, Zhang Q, McGowan S, et al. An overview of the serpin superfamily. *Genome Biol.* 2006;7(5):216. doi:10.1186/gb-2006-7-5-216
72. Silverman GA, Bird PI, Carrell RW, et al. The Serpins Are an Expanding Superfamily of Structurally Similar but Functionally Diverse Proteins. *J Biol Chem.* 2001;276(36):33293-33296. doi:10.1074/jbc.R100016200
73. Maas C, de Maat S. Therapeutic SERPINS: Improving on Nature. *Front Cardiovasc Med.* 2021;8. doi:10.3389/fcvm.2021.648349

74. Riahi Y, Siman-Tov R, Ankri S. Molecular cloning, expression and characterization of a serine proteinase inhibitor gene from *Entamoeba histolytica*. *Mol Biochem Parasitol*. 2004;133(2):153-162. doi:10.1016/j.molbiopara.2003.10.003
75. Spence MA, Mortimer MD, Buckle AM, Minh BQ, Jackson CJ. A Comprehensive Phylogenetic Analysis of the Serpin Superfamily. Echave J, ed. *Mol Biol Evol*. 2021;38(7):2915-2929. doi:10.1093/molbev/msab081
76. Goulas T, Ksiazek M, Garcia-Ferrer I, et al. A structure-derived snap-trap mechanism of a multispecific serpin from the dysbiotic human oral microbiome. *J Biol Chem*. 2017;292(26):10883-10898. doi:10.1074/jbc.M117.786533
77. Bao J, Pan G, Poncz M, Wei J, Ran M, Zhou Z. Serpin functions in host-pathogen interactions. *PeerJ*. 2018;6:e4557. doi:10.7717/peerj.4557
78. Janciauskiene S. Conformational properties of serine proteinase inhibitors (serpins) confer multiple pathophysiological roles. *Biochim Biophys Acta - Mol Basis Dis*. 2001;1535(3):221-235. doi:10.1016/S0925-4439(01)00025-4
79. Ohmuraya M, Yamamura K. The Roles of Serine Protease Inhibitor Kazal Type 1 (SPINK1) in Pancreatic Diseases. *Exp Anim*. 2011;60(5):433-444. doi:10.1538/expanim.60.433
80. Huntington JA, Read RJ, Carrell RW. Structure of a serpin–protease complex shows inhibition by deformation. *Nature*. 2000;407(6806):923-926. doi:10.1038/35038119
81. Khan MS, Singh P, Azhar A, et al. Serpin Inhibition Mechanism: A Delicate Balance between Native Metastable State and Polymerization. *J Amino Acids*. 2011;2011:1-10. doi:10.4061/2011/606797
82. Huntington JA. Serpin structure, function and dysfunction. *J Thromb Haemost*. 2011;9:26-34. doi:10.1111/j.1538-7836.2011.04360.x
83. Marijanovic EM, Fodor J, Riley BT, et al. Reactive centre loop dynamics and serpin specificity. *Sci Rep*. 2019;9(1):3870. doi:10.1038/s41598-019-40432-w
84. Gettins PGW, Olson ST. Exosite Determinants of Serpin Specificity. *J Biol Chem*. 2009;284(31):20441-20445. doi:10.1074/jbc.R800064200
85. Annand RR, Dahlen JR, Sprecher CA, et al. Caspase-1 (interleukin-1beta-converting enzyme) is inhibited by the human serpin analogue proteinase inhibitor 9. *Biochem J*. 1999;342 Pt 3(Pt 3):655-665. doi:10477277
86. Schick C, Pemberton PA, Shi G-P, et al. Cross-Class Inhibition of the Cysteine Proteinases Cathepsins K, L, and S by the Serpin Squamous Cell Carcinoma Antigen 1: A Kinetic Analysis. *Biochemistry*. 1998;37(15):5258-5266. doi:10.1021/bi972521d
87. Irving JA, Pike RN, Dai W, et al. Evidence That Serpin Architecture Intrinsically Supports Papain-like Cysteine Protease Inhibition: Engineering $\alpha 1$ -Antitrypsin To Inhibit Cathepsin Proteases. *Biochemistry*. 2002;41(15):4998-5004. doi:10.1021/bi0159985
88. Nakai A, Satoh M, Hirayoshi K, Nagata K. Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum. *J Cell Biol*. 1992;117(4):903-914. doi:10.1083/jcb.117.4.903
89. Hammond GL, Smith CL, Goping IS, et al. Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proc Natl Acad Sci*. 1987;84(15):5153-5157. doi:10.1073/pnas.84.15.5153
90. Flink IL, Bailey TJ, Gustafson TA, Markham BE, Morkin E. Complete amino acid sequence of human thyroxine-binding globulin deduced from cloned DNA: close homology to the serine antiproteases. *Proc Natl Acad Sci*. 1986;83(20):7708-7712. doi:10.1073/pnas.83.20.7708
91. Zhou A, Wei Z, Stanley PLD, Read RJ, Stein PE, Carrell RW. The S-to-R Transition of Corticosteroid-Binding Globulin and the Mechanism of Hormone Release. *J Mol Biol*. 2008;380(1):244-251. doi:10.1016/j.jmb.2008.05.012
92. Irving JA, Ekeowa UI, Belorgey D, et al. The Serpinopathie: Studying serpin polymerization in vivo. In: *Methods of enzymology. Serpin structure and evolution*. ; 2011:421-466. doi:10.1016/B978-0-12-385950-1.00018-3

93. Sharp HL, Bridges RA, Krivit W, Freier EF. Cirrhosis associated with alpha-1-antitrypsin deficiency: a previously unrecognized inherited disorder. *J Lab Clin Med.* 1969;73(6):934-939. doi:4182334
94. Lieberman J. Heterozygous and Homozygous Alpha-1-Antitrypsin Deficiency in Patients with Pulmonary Emphysema. *N Engl J Med.* 1969;281(6):279-284. doi:10.1056/NEJM196908072810601
95. De Maat S, Hofman ZLM, Maas C. Hereditary angioedema: the plasma contact system out of control. *J Thromb Haemost.* 2018;16(9):1674-1685. doi:10.1111/jth.14209
96. Marciniak E, Farley CH, DeSimone PA. Familial thrombosis due to antithrombin 3 deficiency. *Blood.* 1974;43(2):219-231. doi:Familial thrombosis due to antithrombin 3 deficiency
97. Roussel BD, Lomas DA, Crowther DC. Progressive myoclonus epilepsy associated with neuroserpin inclusion bodies (neuroserpinosis). *Epileptic Disord.* 2016;18(S2):103-110. doi:10.1684/epd.2016.0847
98. Long GL, Chandra T, Woo SLC, Davie EW, Kurachi K. Complete sequence of the cDNA for human .alpha.1-antitrypsin and the gene for the S variant. *Biochemistry.* 1984;23(21):4828-4837. doi:10.1021/bi00316a003
99. Hafeez W, Ciliberto G, Perlmutter DH. Constitutive and modulated expression of the human alpha 1 antitrypsin gene. Different transcriptional initiation sites used in three different cell types. *J Clin Invest.* 1992;89(4):1214-1222. doi:10.1172/JCI115705
100. Li Y, Zhou L, Twining SS, Sugar J, Yue BYJT. Involvement of Sp1 Elements in the Promoter Activity of the α 1-Proteinase Inhibitor Gene. *J Biol Chem.* 1998;273(16):9959-9965. doi:10.1074/jbc.273.16.9959
101. Sun Z, Yang P. Role of imbalance between neutrophil elastase and α 1-antitrypsin in cancer development and progression. *Lancet Oncol.* 2004;5(3):182-190. doi:10.1016/S1470-2045(04)01414-7
102. Ciliberto G, Dente L, Cortese R. Cell-specific expression of a transfected human α 1-antitrypsin gene. *Cell.* 1985;41(2):531-540. doi:10.1016/S0092-8674(85)80026-X
103. Monaci P, Nicosia A, Cortese R. Two different liver-specific factors stimulate in vitro transcription from the human alpha 1-antitrypsin promoter. *EMBO J.* 1988;7(7):2075-2087. doi:10.1002/j.1460-2075.1988.tb03047.x
104. De Simone V, Ciliberto G, Hardon E, et al. Cis- and trans-acting elements responsible for the cell-specific expression of the human alpha 1-antitrypsin gene. *EMBO J.* 1987;6(9):2759-2766. doi:10.1002/j.1460-2075.1987.tb02570.x
105. De Simone V, Cortese R. Transcription factors and liver-specific genes. *Biochim Biophys Acta - Gene Struct Expr.* 1992;1132(2):119-126. doi:10.1016/0167-4781(92)90001-G
106. Courtois G, Morgan JG, Campbell LA, Fourel G, Crabtree GR. Interaction of a Liver-Specific Nuclear Factor with the Fibrinogen and α 1 -Antitrypsin Promoters. *Science (80-).* 1987;238(4827):688-692. doi:10.1126/science.3499668
107. Courtois G, Baumhueter S, Crabtree GR. Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. *Proc Natl Acad Sci.* 1988;85(21):7937-7941. doi:10.1073/pnas.85.21.7937
108. Hardon EM, Frain M, Paonessa G, Cortese R. Two distinct factors interact with the promoter regions of several liver-specific genes. *EMBO J.* 1988;7(6):1711-1719. doi:10.1002/j.1460-2075.1988.tb03000.x
109. Li Y, Shen R-F, Tsai SY, Woo SLC. Multiple Hepatic trans -Acting Factors Are Required for In Vitro Transcription of the Human Alpha-1-Antitrypsin Gene. *Mol Cell Biol.* 1988;8(10):4362-4369. doi:10.1128/mcb.8.10.4362-4369.1988
110. Hu C, Perlmutter DH. Regulation of α 1 -antitrypsin gene expression in human intestinal epithelial cell line Caco-2 by HNF-1 α and HNF-4. *Am J Physiol Liver Physiol.* 1999;276(5):G1181-G1194. doi:10.1152/ajpgi.1999.276.5.G1181

111. Marsden MD, Fournier REK. Chromosomal Elements Regulate Gene Activity and Chromatin Structure of the Human Serpin Gene Cluster at 14q32.1. *Mol Cell Biol.* 2003;23(10):3516-3526. doi:10.1128/MCB.23.10.3516-3526.2003
112. Knoell DL, Ralston DR, Coulter KR, Wewers MD. Alpha 1-antitrypsin and Protease Complexation Is Induced by Lipopolysaccharide, Interleukin-1 β , and Tumor Necrosis Factor- α in Monocytes. *Am J Respir Crit Care Med.* 1998;157(1):246-255. doi:10.1164/ajrccm.157.1.9702033
113. Boutten A, Venembre P, Seta N, et al. Oncostatin M Is a Potent Stimulator of α 1 -Antitrypsin Secretion in Lung Epithelial Cells: Modulation by Transforming Growth Factor- β and Interferon- γ . *Am J Respir Cell Mol Biol.* 1998;18(4):511-520. doi:10.1165/ajrcmb.18.4.2772
114. Morgan K, Marsters P, Morley S, et al. Oncostatin M induced α 1-antitrypsin (AAT) gene expression in Hep G2 cells is mediated by a 3' enhancer. *Biochem J.* 2002;365(2):555-560. doi:10.1042/bj20011312
115. Morgan K, Scobie G, Marsters P, Kalsheker NA. Mutation in an α 1-antitrypsin enhancer results in an interleukin-6 deficient acute-phase response due to loss of cooperativity between transcription factors. *Biochim Biophys Acta - Mol Basis Dis.* 1997;1362(1):67-76. doi:10.1016/S0925-4439(97)00064-1
116. Castell J V., Gómez-Lechón MJ, David M, et al. Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. *FEBS Lett.* 1989;242(2):237-239. doi:10.1016/0014-5793(89)80476-4
117. Lisowska-Myjak B, Sygitowicz G, Wolf B, Pachecka J. Serum alpha-1-antitrypsin concentration during normal and diabetic pregnancy. *Eur J Obstet Gynecol Reprod Biol.* 2001;99(1):53-56. doi:10.1016/S0301-2115(01)00356-6
118. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet.* 2008;40(12):1413-1415. doi:10.1038/ng.259
119. Tung K-F, Pan C-Y, Chen C-H, Lin W. Top-ranked expressed gene transcripts of human protein-coding genes investigated with GTEx dataset. *Sci Rep.* 2020;10(1):16245. doi:10.1038/s41598-020-73081-5
120. Jiang W, Chen L. Alternative splicing: Human disease and quantitative analysis from high-throughput sequencing. *Comput Struct Biotechnol J.* 2021;19:183-195. doi:10.1016/j.csbj.2020.12.009
121. Grayeski PJ, Weidmann CA, Kumar J, et al. Global 5'-UTR RNA structure regulates translation of a *SERPINA1* mRNA. *Nucleic Acids Res.* 2022;50(17):9689-9704. doi:10.1093/nar/gkac739
122. Lackey L, Coria A, Ghosh AJ, et al. Alternative poly-adenylation modulates α 1-antitrypsin expression in chronic obstructive pulmonary disease. Zavolan M, ed. *PLOS Genet.* 2021;17(11):e1009912. doi:10.1371/journal.pgen.1009912
123. Sifers RN. Intracellular Processing of 1-Antitrypsin. *Proc Am Thorac Soc.* 2010;7(6):376-380. doi:10.1513/pats.201001-011AW
124. Carrell RW, Jeppsson J-O, Vaughan L, Brennan SO, Owen MC, Boswell DR. Human α 1 -antitrypsin: carbohydrate attachment and sequence homology. *FEBS Lett.* 1981;135(2):301-303. doi:10.1016/0014-5793(81)80805-8
125. Samandari T, Brown JL. A study of the effects of altering the sites for N -glycosylation in α -1-proteinase inhibitor variants M and S. *Protein Sci.* 1993;2(9):1400-1410. doi:10.1002/pro.5560020905
126. McCarthy C, Saldova R, Wormald MR, Rudd PM, McElvaney NG, Reeves EP. The Role and Importance of Glycosylation of Acute Phase Proteins with Focus on Alpha-1 Antitrypsin in Acute and Chronic Inflammatory Conditions. *J Proteome Res.* 2014;13(7):3131-3143. doi:10.1021/pr500146y

127. Kolarich D, Weber A, Turecek PL, Schwarz H-P, Altmann F. Comprehensive glyco-proteomic analysis of human α 1-antitrypsin and its charge isoforms. *Proteomics*. 2006;6(11):3369-3380. doi:10.1002/pmic.200500751
128. Докрунова А.А., Уласова Н.Ю., Крамина Т.Е., Альбицкий И.А., Хасигов П.З. СОС. Полиморфизм альфа1-антитрипсина при раке и гиперплазии предстательной железы. *Молекулярная медицина*. 2010;6:43-49.
129. Baek J-H, Yang WS, Lee C, Yu M-H. Functional Unfolding of α 1-Antitrypsin Probed by Hydrogen-Deuterium Exchange Coupled with Mass Spectrometry. *Mol Cell Proteomics*. 2009;8(5):1072-1081. doi:10.1074/mcp.M800365-MCP200
130. Brantly M, Nukiwa T, Crystal RG. Molecular basis of alpha-1-antitrypsin deficiency. *Am J Med*. 1988;84(6):13-31. doi:10.1016/0002-9343(88)90154-4
131. Levin EG, Santell L. Conversion of the active to latent plasminogen activator inhibitor from human endothelial cells. *Blood*. 1987;70(4):1090-1098.
132. Lomas DA, Elliott PR, Chang W-SW, Wardell MR, Carrell RW. Preparation and Characterization of Latent α 1-Antitrypsin. *J Biol Chem*. 1995;270(10):5282-5288. doi:10.1074/jbc.270.10.5282
133. Im H, Ahn H-Y, Yu M-H. Bypassing the kinetic trap of serpin protein folding by loop extension. *Protein Sci*. 2000;9(8):1497-1502. doi:10.1110/ps.9.8.1497
134. Dolmer K, Gettins PGW. How the Serpin α 1-Proteinase Inhibitor Folds. *J Biol Chem*. 2012;287(15):12425-12432. doi:10.1074/jbc.M111.315465
135. Simonovic I, Patston PA. The native metastable fold of C1-inhibitor is stabilized by disulfide bonds. *Biochim Biophys Acta - Protein Struct Mol Enzymol*. 2000;1481(1):97-102. doi:10.1016/S0167-4838(00)00115-1
136. Beatty K, Bieth J, Travis J. Kinetics of association of serine proteinases with native and oxidized alpha-1-proteinase inhibitor and alpha-1-antichymotrypsin. *J Biol Chem*. 1980;255(9):3931-3934.
137. Rao N V, Wehner NG, Marshall BC, Gray WR, Gray BH, Hoidal JR. Characterization of proteinase-3 (PR-3), a neutrophil serine proteinase. Structural and functional properties. *J Biol Chem*. 1991;266(15):9540-9548.
138. He S-H, Chen P, Chen H-Q. Modulation of enzymatic activity of human mast cell tryptase and chymase by protease inhibitors. *Acta Pharmacol Sin*. 2003;24(9):923-929.
139. Schapira M, Scott CF, Colman RW. Contribution of Plasma Protease Inhibitors to the Inactivation of Kallikrein in Plasma. *J Clin Invest*. 1982;69(2):462-468. doi:10.1172/JCI110470
140. Cohen AB. The interaction of α 1-antitrypsin with chymotrypsin, trypsin and elastase. *Biochim Biophys Acta - Enzymol*. 1975;391(1):193-200. doi:10.1016/0005-2744(75)90166-7
141. Poe M, Blake JT, Boulton DA, et al. Human cytotoxic lymphocyte granzyme B. Its purification from granules and the characterization of substrate and inhibitor specificity. *J Biol Chem*. 1991;266(1):98-103.
142. Churg A, Wang X, Wang RD, Meixner SC, Pryzdial ELG, Wright JL. Alpha1-antitrypsin suppresses TNF-alpha and MMP-12 production by cigarette smoke-stimulated macrophages. *Am J Respir Cell Mol Biol*. 2007;37(2):144-151. doi:10.1165/rcmb.2006-0345OC
143. Clemmensen I, Christensen F. Inhibition of urokinase by complex formation with human alpha1-antitrypsin. *Biochim Biophys Acta*. 1976;429(2):591-599. doi:10.1016/0005-2744(76)90307-7
144. Ellis V, Scully M, MacGregor I, Kakkar V. Inhibition of human factor Xa by various plasma protease inhibitors. *Biochim Biophys Acta*. 1982;701(1):24-31.
145. Janciauskiene S, Nita I, Subramaniam D, Li Q, Lancaster JR, Matalon S. α 1 -Antitrypsin Inhibits the Activity of the Matriptase Catalytic Domain In Vitro. *Am J Respir Cell Mol Biol*. 2008;39(6):631-637. doi:10.1165/rcmb.2008-0015RC
146. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun*. 1967;27(2):157-162. doi:10.1016/S0006-291X(67)80055-X

147. Owen MC, Brennan SO, Lewis JH, Carrell RW. Mutation of Antitrypsin to Antithrombin. *N Engl J Med*. 1983;309(12):694-698. doi:10.1056/NEJM198309223091203
148. Scott CF, Carrell RW, Glaser CB, Kueppers F, Lewis JH, Colman RW. Alpha-1-antitrypsin-Pittsburgh. A potent inhibitor of human plasma factor XIa, kallikrein, and factor XIIIf. *J Clin Invest*. 1986;77(2):631-634. doi:10.1172/JCI112346
149. TRAVIS J, MATHESON NR, GEORGE PM, CARRELL RW. Kinetic Studies on the Interaction of α 1 -Proteinase Inhibitor (Pittsburgh) with Trypsin-Like Serine Proteinases. *Biol Chem Hoppe Seyler*. 1986;367(2):853-860. doi:10.1515/bchm3.1986.367.2.853
150. Radisky ES, Lee JM, Lu C-JK, Koshland DE. Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proc Natl Acad Sci*. 2006;103(18):6835-6840. doi:10.1073/pnas.0601910103
151. Dementiev A, Dobó J, Gettins PGW. Active Site Distortion Is Sufficient for Proteinase Inhibition by Serpins. *J Biol Chem*. 2006;281(6):3452-3457. doi:10.1074/jbc.M510564200
152. Hopkins PCR, Stone SR. The Contribution of the Conserved Hinge Region Residues of .alpha.1-Antitrypsin to Its Reaction with Elastase. *Biochemistry*. 1995;34(48):15872-15879. doi:10.1021/bi00048a033
153. Chung H-S, Kim J-S, Lee SM, Park SJ. Role of the P2 residue of human alpha 1-antitrypsin in determining target protease specificity. Permyakov EA, ed. *PLoS One*. 2017;12(9):e0185074. doi:10.1371/journal.pone.0185074
154. Joslin G, Fallon RJ, Bullock J, Adams SP, Perlmutter DH. The SEC receptor recognizes a pentapeptide neodomain of alpha 1-antitrypsin-protease complexes. *J Biol Chem*. 1991;266(17):11282-11288.
155. Joslin G, Wittwer A, Adams S, Tollefsen DM, August A, Perlmutter DH. Cross-competition for binding of alpha 1-antitrypsin (alpha 1 AT)-elastase complexes to the serpin-enzyme complex receptor by other serpin-enzyme complexes and by proteolytically modified alpha 1 AT. *J Biol Chem*. 1993;268(3):1886-1893.
156. Kounnas MZ, Church FC, Argraves WS, Strickland DK. Cellular Internalization and Degradation of Antithrombin III-Thrombin, Heparin Cofactor II-Thrombin, and α 1-Antitrypsin-Trypsin Complexes Is Mediated by the Low Density Lipoprotein Receptor-related Protein. *J Biol Chem*. 1996;271(11):6523-6529. doi:10.1074/jbc.271.11.6523
157. Strickland DK, Muratoglu SC, Antalis TM. Serpin-Enzyme Receptors. In: ; 2011:17-31. doi:10.1016/B978-0-12-386471-0.00002-X
158. Casolaro MA, Fells G, Wewers M, et al. Augmentation of lung antineutrophil elastase capacity with recombinant human alpha-1-antitrypsin. *J Appl Physiol*. 1987;63(5):2015-2023. doi:10.1152/jappl.1987.63.5.2015
159. Sarkar A, Wintrode PL. Effects of glycosylation on the stability and flexibility of a metastable protein: The human serpin α 1-antitrypsin. *Int J Mass Spectrom*. 2011;302(1-3):69-75. doi:10.1016/j.ijms.2010.08.003
160. O'Brien ME, Murray G, Gogoi D, et al. A Review of Alpha-1 Antitrypsin Binding Partners for Immune Regulation and Potential Therapeutic Application. *Int J Mol Sci*. 2022;23(5):2441. doi:10.3390/ijms23052441
161. Hubbard RC, McElvaney NG, Sellers SE, Healy JT, Czerski DB, Crystal RG. Recombinant DNA-produced alpha 1-antitrypsin administered by aerosol augments lower respiratory tract antineutrophil elastase defenses in individuals with alpha 1-antitrypsin deficiency. *J Clin Invest*. 1989;84(4):1349-1354. doi:10.1172/JCI114305
162. Lockett AD, Brown MB, Santos-Falcon N, et al. Active Trafficking of Alpha 1 Antitrypsin across the Lung Endothelium. Ahmad S, ed. *PLoS One*. 2014;9(4):e93979. doi:10.1371/journal.pone.0093979
163. Lockett AD, Petrusca DN, Justice MJ, et al. Scavenger receptor class B, type I-mediated uptake of A1AT by pulmonary endothelial cells. *Am J Physiol Lung Cell Mol Physiol*. 2015;309(4):L425-34. doi:10.1152/ajplung.00376.2014

164. Zhou X, Liu Z, Shapiro L, Yang J, Burton GF. Low-density lipoprotein receptor-related protein 1 mediates α 1-antitrypsin internalization in CD4⁺ T lymphocytes. *J Leukoc Biol*. 2015;98(6):1027-1035. doi:10.1189/jlb.2A0515-209R
165. Subramaniam D, Zhou H, Liang M, Welte T, Mahadeva R, Janciauskiene S. Cholesterol rich lipid raft microdomains are gateway for acute phase protein, SERPINA1. *Int J Biochem Cell Biol*. 2010;42(9):1562-1570. doi:10.1016/j.biocel.2010.06.009
166. Johnson D, Travis J. The oxidative inactivation of human alpha-1-proteinase inhibitor. Further evidence for methionine at the reactive center. *J Biol Chem*. 1979;254(10):4022-4026.
167. Taggart C. Oxidation of either methionine super351 or methionine super358 in alpha 1-antitrypsin causes loss of anti-neutrophil elastase activity. *J Biol Chem*. Published online June 23, 2000. doi:10.1074/jbc.M004850200
168. Janciauskiene S, Eriksson S. The interaction of hydrophobic bile acids with the α 1 -proteinase inhibitor. *FEBS Lett*. 1994;343(2):141-145. doi:10.1016/0014-5793(94)80306-4
169. Vissers MC, George PM, Bathurst IC, Brennan SO, Winterbourn CC. Cleavage and inactivation of alpha 1-antitrypsin by metalloproteinases released from neutrophils. *J Clin Invest*. 1988;82(2):706-711. doi:10.1172/JCI113651
170. Winyard PG, Zhang Z, Chidwick K, Blake DR, Carrell RW, Murphy G. Proteolytic inactivation of human α 1 antitrypsin by human stromelysin. *FEBS Lett*. 1991;279(1):91-94. doi:10.1016/0014-5793(91)80258-5
171. Belorgey D, Bieth JG. Effect of Polynucleotides on the Inhibition of Neutrophil Elastase by Mucus Proteinase Inhibitor and α 1 -Proteinase Inhibitor. *Biochemistry*. 1998;37(46):16416-16422. doi:10.1021/bi981536o
172. Durantou J, Boudier C, Belorgey D, Mellet P, Bieth JG. DNA Strongly Impairs the Inhibition of Cathepsin G by α 1-Antichymotrypsin and α 1-Proteinase Inhibitor. *J Biol Chem*. 2000;275(6):3787-3792. doi:10.1074/jbc.275.6.3787
173. Ermolieff J, Boudier C, Laine A, Meyer B, Bieth JG. Heparin protects cathepsin G against inhibition by protein proteinase inhibitors. *J Biol Chem*. 1994;269(47):29502-29508.
174. Gupta VK, Gowda LR. Alpha-1-proteinase inhibitor is a heparin binding serpin: Molecular interactions with the Lys rich cluster of helix-F domain. *Biochimie*. 2008;90(5):749-761. doi:10.1016/j.biochi.2008.01.004
175. Stoller JK, Aboussouan LS. A Review of α 1 -Antitrypsin Deficiency. *Am J Respir Crit Care Med*. 2012;185(3):246-259. doi:10.1164/rccm.201108-1428CI
176. Foil KE. Variants of *SERPINA1* and the increasing complexity of testing for alpha-1 antitrypsin deficiency. *Ther Adv Chronic Dis*. 2021;12_suppl:204062232110159. doi:10.1177/20406223211015954
177. Blanco I, Bueno P, Diego I, et al. Alpha-1 antitrypsin Pi*Z gene frequency and Pi*ZZ genotype numbers worldwide: an update. *Int J Chron Obstruct Pulmon Dis*. 2017;Volume 12:561-569. doi:10.2147/COPD.S125389
178. Lomas DA, LI-Evans D, Finch JT, Carrell RW. The mechanism of Z α 1-antitrypsin accumulation in the liver. *Nature*. 1992;357(6379):605-607. doi:10.1038/357605a0
179. Elliott PR, Stein PE, Bilton D, Carrell RW, Lomas DA. Structural explanation for the deficiency of S α 1-antitrypsin. *Nat Struct Biol*. 1996;3(11):910-911. doi:10.1038/nsb1196-910
180. Lomas DA, Evans DL, Stone SR, Chang WSW, Carrell RW. Effect of the Z mutation on the physical and inhibitory properties of .alpha.1-antitrypsin. *Biochemistry*. 1993;32(2):500-508. doi:10.1021/bi00053a014
181. Kroeger H, Miranda E, MacLeod I, et al. Endoplasmic Reticulum-associated Degradation (ERAD) and Autophagy Cooperate to Degrade Polymerogenic Mutant Serpins. *J Biol Chem*. 2009;284(34):22793-22802. doi:10.1074/jbc.M109.027102
182. Lomas DA. New Therapeutic Targets for Alpha-1 Antitrypsin Deficiency. *Chronic Obstr Pulm Dis J COPD Found*. 2018;5(4):233-243. doi:10.15326/jcopdf.5.4.2017.0165
183. Nyon MP, Gooptu B. Therapeutic targeting of misfolding and conformational change in α 1 -antitrypsin deficiency. *Future Med Chem*. 2014;6(9):1047-1065. doi:10.4155/fmc.14.58

184. Miranda E, Pérez J, Ekeowa UI, et al. A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with $\alpha 1$ -antitrypsin deficiency. *Hepatology*. 2010;52(3):1078-1088. doi:10.1002/hep.23760
185. Yamasaki M, Sendall TJ, Pearce MC, Whisstock JC, Huntington JA. Molecular basis of $\alpha 1$ -antitrypsin deficiency revealed by the structure of a domain-swapped trimer. *EMBO Rep*. 2011;12(10):1011-1017. doi:10.1038/embor.2011.171
186. Behrens MA, Sendall TJ, Pedersen JS, Kjeldgaard M, Huntington JA, Jensen JK. The Shapes of Z- $\alpha 1$ -Antitrypsin Polymers in Solution Support the C-Terminal Domain-Swap Mechanism of Polymerization. *Biophys J*. 2014;107(8):1905-1912. doi:10.1016/j.bpj.2014.08.030
187. Faull S V., Elliston ELK, Goptu B, et al. The structural basis for Z $\alpha 1$ -antitrypsin polymerization in the liver. *Sci Adv*. 2020;6(43). doi:10.1126/sciadv.abc1370
188. Hidvegi T, Schmidt BZ, Hale P, Perlmutter DH. Accumulation of Mutant $\alpha 1$ -Antitrypsin Z in the Endoplasmic Reticulum Activates Caspases-4 and -12, NF κ B, and BAP31 but Not the Unfolded Protein Response. *J Biol Chem*. 2005;280(47):39002-39015. doi:10.1074/jbc.M508652200
189. Lawless MW, Greene CM, Mulgrew A, Taggart CC, O'Neill SJ, McElvaney NG. Activation of Endoplasmic Reticulum-Specific Stress Responses Associated with the Conformational Disease Z $\alpha 1$ -Antitrypsin Deficiency. *J Immunol*. 2004;172(9):5722-5726. doi:10.4049/jimmunol.172.9.5722
190. Ordóñez A, Snapp EL, Tan L, Miranda E, Marciniak SJ, Lomas DA. Endoplasmic reticulum polymers impair luminal protein mobility and sensitize to cellular stress in alpha 1 -antitrypsin deficiency. *Hepatology*. 2013;57(5):2049-2060. doi:10.1002/hep.26173
191. Chambers JE, Zubkov N, Kubánková M, et al. Z- $\alpha 1$ -antitrypsin polymers impose molecular filtration in the endoplasmic reticulum after undergoing phase transition to a solid state. *Sci Adv*. 2022;8(14). doi:10.1126/sciadv.abm2094
192. Littleton ET, Bevis L, Hansen LJ, et al. Alpha 1-antitrypsin deficiency, complement activation, and chronic liver disease. *J Clin Pathol*. 1991;44(10):855-858. doi:10.1136/jcp.44.10.855
193. Chapman KR, Burdon JGW, Piitulainen E, et al. Intravenous augmentation treatment and lung density in severe $\alpha 1$ antitrypsin deficiency (RAPID): a randomised, double-blind, placebo-controlled trial. *Lancet*. 2015;386(9991):360-368. doi:10.1016/S0140-6736(15)60860-1
194. Гембицкая Т.Е., Черменский А.Г. ПИНА. Генетически обусловленный дефицит альфа-1-антитрипсина и заболевания легких: диагностика, современные схемы лечения, перспективы организации помощи больным. *Практическая пульмонология*. 2018;4:67-73.
195. Toldo S, Seropian IM, Mezzaroma E, et al. Alpha-1 antitrypsin inhibits caspase-1 and protects from acute myocardial ischemia-reperfusion injury. *J Mol Cell Cardiol*. 2011;51(2):244-251. doi:10.1016/j.yjmcc.2011.05.003
196. Lockett AD, Van Demark M, Gu Y, et al. Effect of Cigarette Smoke Exposure and Structural Modifications on the α -1 Antitrypsin Interaction with Caspases. *Mol Med*. 2012;18(3):445-454. doi:10.2119/molmed.2011.00207
197. Jedicke N, Struever N, Aggrawal N, et al. Alpha-1-antitrypsin inhibits acute liver failure in mice. *Hepatology*. 2014;59(6):2299-2308. doi:10.1002/hep.27024
198. Al-Omari M, Korenbaum E, Ballmaier M, et al. Acute-Phase Protein $\alpha 1$ -Antitrypsin Inhibits Neutrophil Calpain I and Induces Random Migration. *Mol Med*. 2011;17(9-10):865-874. doi:10.2119/molmed.2011.00089
199. Lokuta MA, Nuzzi PA, Huttenlocher A. Calpain regulates neutrophil chemotaxis. *Proc Natl Acad Sci*. 2003;100(7):4006-4011. doi:10.1073/pnas.0636533100
200. Bergin DA, Reeves EP, Meleady P, et al. α -1 Antitrypsin regulates human neutrophil chemotaxis induced by soluble immune complexes and IL-8. *J Clin Invest*. 2010;120(12):4236-4250. doi:10.1172/JCI41196

201. Lockett AD, Kimani S, Ddungu G, et al. α 1 -Antitrypsin Modulates Lung Endothelial Cell Inflammatory Responses to TNF- α . *Am J Respir Cell Mol Biol*. 2013;49(1):143-150. doi:10.1165/rcmb.2012-0515OC
202. Ferry G, Lonchamp M, Pennel L, de Nanteuil G, Canet E, Tucker GC. Activation of MMP-9 by neutrophil elastase in an in vivo model of acute lung injury. *FEBS Lett*. 1997;402(2-3):111-115. doi:10.1016/S0014-5793(96)01508-6
203. Hansson SF, Andréasson U, Wall M, et al. Reduced Levels of Amyloid- β -Binding Proteins in Cerebrospinal Fluid from Alzheimer's Disease Patients. *J Alzheimer's Dis*. 2009;16(2):389-397. doi:10.3233/JAD-2009-0966
204. Zabel C, Chamrad DC, Priller J, et al. Alterations in the Mouse and Human Proteome Caused by Huntington's Disease. *Mol Cell Proteomics*. 2002;1(5):366-375. doi:10.1074/mcp.M200016-MCP200
205. Bohrmann B, Tjernberg L, Kuner P, et al. Endogenous Proteins Controlling Amyloid β -Peptide Polymerization. *J Biol Chem*. 1999;274(23):15990-15995. doi:10.1074/jbc.274.23.15990
206. Torricelli C, Capurro E, Santucci A, et al. Multiple plasma proteins control atrial natriuretic peptide (ANP) aggregation. *J Mol Endocrinol*. 2004;33(2):335-341. doi:10.1677/jme.1.01530
207. Janciauskiene S, Carlemalm E, Eriksson S. In vitro Fibril Formation from α 1 -Antitrypsin-Derived C-Terminal Peptides. *Biol Chem Hoppe Seyler*. 1995;376(7):415-424. doi:10.1515/bchm3.1995.376.7.415
208. Chowdhury P, Wang W, Lavender S, et al. Fluorescence Correlation Spectroscopic Study of Serpin Depolymerization by Computationally Designed Peptides. *J Mol Biol*. 2007;369(2):462-473. doi:10.1016/j.jmb.2007.03.042
209. Davids JW, Elthaher TSH, Nakai A, Nagata K, Miller AD. Modeling the Three-Dimensional Structure of Serpin/Molecular Chaperone HSP47. *Bioorg Chem*. 1995;23(4):427-438. doi:10.1006/bioo.1995.1029
210. Widmer C, Gebauer JM, Brunstein E, et al. Molecular basis for the action of the collagen-specific chaperone Hsp47/SERPINH1 and its structure-specific client recognition. *Proc Natl Acad Sci*. 2012;109(33):13243-13247. doi:10.1073/pnas.1208072109
211. Li Z, Alam S, Wang J, Sandstrom CS, Janciauskiene S, Mahadeva R. Oxidized α 1 -antitrypsin stimulates the release of monocyte chemoattractant protein-1 from lung epithelial cells: potential role in emphysema. *Am J Physiol Cell Mol Physiol*. 2009;297(2):L388-L400. doi:10.1152/ajplung.90373.2008
212. Carp H, Miller F, Hoidal JR, Janoff A. Potential mechanism of emphysema: alpha 1-proteinase inhibitor recovered from lungs of cigarette smokers contains oxidized methionine and has decreased elastase inhibitory capacity. *Proc Natl Acad Sci*. 1982;79(6):2041-2045. doi:10.1073/pnas.79.6.2041
213. Levine RL, Mosoni L, Berlett BS, Stadtman ER. Methionine residues as endogenous antioxidants in proteins. *Proc Natl Acad Sci*. 1996;93(26):15036-15040. doi:10.1073/pnas.93.26.15036
214. Feng Y-L, Yin Y-X, Ding J, et al. Alpha-1-antitrypsin suppresses oxidative stress in preeclampsia by inhibiting the p38MAPK signaling pathway: An in vivo and in vitro study. Gallyas F, ed. *PLoS One*. 2017;12(3):e0173711. doi:10.1371/journal.pone.0173711
215. Maslakova A.A., Rubtsov M.A., Efimova V.S., Sokolova O.S., Smirnova O.V. Orlovsky I.V. Alpha1-antitrypsin-derived C-terminal peptide is a potent oxidative stress inhibitor. In: *ASCB/EMBO Meeting. Mol Biol Cell Suppl*. ; 2017;28(26), 1259.
216. Hawkins P, McEnery T, Gabillard-Lefort C, et al. In vitro and in vivo modulation of NADPH oxidase activity and reactive oxygen species production in human neutrophils by α 1 -antitrypsin. *ERJ Open Res*. 2021;7(4):00234-02021. doi:10.1183/23120541.00234-2021
217. Magallón M, Castillo-Corullón S, Bañuls L, et al. Hypoxia Enhances Oxidative Stress in Neutrophils from ZZ Alpha-1 Antitrypsin Deficiency Patients. *Antioxidants*. 2023;12(4):872. doi:10.3390/antiox12040872

218. Malerba M. Neutrophilic inflammation and IL-8 levels in induced sputum of alpha-1-antitrypsin PiMZ subjects. *Thorax*. 2006;61(2):129-133. doi:10.1136/thx.2005.043471
219. Stone H, Pye A, Stockley RA. Disease associations in alpha-1-antitrypsin deficiency. *Respir Med*. 2014;108(2):338-343. doi:10.1016/j.rmed.2013.10.006
220. Pott GB, Chan ED, Dinarello CA, Shapiro L. α -1-Antitrypsin is an endogenous inhibitor of proinflammatory cytokine production in whole blood. *J Leukoc Biol*. 2009;85(5):886-895. doi:10.1189/jlb.0208145
221. Janciauskiene S, Larsson S, Larsson P, Virtala R, Jansson L, Stevens T. Inhibition of lipopolysaccharide-mediated human monocyte activation, in vitro, by α 1-antitrypsin. *Biochem Biophys Res Commun*. 2004;321(3):592-600. doi:10.1016/j.bbrc.2004.06.123
222. Gottlieb PA, Alkanani AK, Michels AW, et al. α 1-Antitrypsin Therapy Downregulates Toll-Like Receptor-Induced IL-1 β Responses in Monocytes and Myeloid Dendritic Cells and May Improve Islet Function in Recently Diagnosed Patients With Type 1 Diabetes. *J Clin Endocrinol Metab*. 2014;99(8):E1418-E1426. doi:10.1210/jc.2013-3864
223. Moraga F, Janciauskiene S. Activation of Primary Human Monocytes by the Oxidized Form of α 1-Antitrypsin. *J Biol Chem*. 2000;275(11):7693-7700. doi:10.1074/jbc.275.11.7693
224. Zhou X, Shapiro L, Fellingham G, Willardson BM, Burton GF. HIV Replication in CD4+ T Lymphocytes in the Presence and Absence of Follicular Dendritic Cells: Inhibition of Replication Mediated by α -1-Antitrypsin through Altered I κ B α Ubiquitination. *J Immunol*. 2011;186(5):3148-3155. doi:10.4049/jimmunol.1001358
225. Marcondes AM, Karopongse E, Lesnikova M, et al. α -1-Antitrypsin (AAT)-modified donor cells suppress GVHD but enhance the GVL effect: a role for mitochondrial bioenergetics. *Blood*. 2014;124(18):2881-2891. doi:10.1182/blood-2014-04-570440
226. Ozeri E, Rider P, Rigbi S, et al. Differential signaling patterns of stimulated bone marrow-derived dendritic cells under α 1-antitrypsin-enriched conditions. *Cell Immunol*. 2021;361:104281. doi:10.1016/j.cellimm.2020.104281
227. Lewis EC, Mizrahi M, Toledano M, et al. α 1-Antitrypsin monotherapy induces immune tolerance during islet allograft transplantation in mice. *Proc Natl Acad Sci*. 2008;105(42):16236-16241. doi:10.1073/pnas.0807627105
228. Schuster R, Motola-Kalay N, Baranovski BM, et al. Distinct anti-inflammatory properties of alpha1-antitrypsin and corticosteroids reveal unique underlying mechanisms of action. *Cell Immunol*. 2020;356:104177. doi:10.1016/j.cellimm.2020.104177
229. Guttman O, Freixo-Lima GS, Kaner Z, Lior Y, Rider P, Lewis EC. Context-Specific and Immune Cell-Dependent Antitumor Activities of α 1-Antitrypsin. *Front Immunol*. 2016;7. doi:10.3389/fimmu.2016.00559
230. Guttman O, Yossef R, Freixo-Lima G, Rider P, Porgador A, Lewis EC. α 1-Antitrypsin modifies general natural killer cell interactions with dendritic cells and specific interactions with islet β -cells in favour of protection from autoimmune diabetes. *Immunology*. 2015;144(3):530-539. doi:10.1111/imm.12403
231. Koulmanda M, Bhasin M, Hoffman L, et al. Curative and β cell regenerative effects of α 1-antitrypsin treatment in autoimmune diabetic NOD mice. *Proc Natl Acad Sci*. 2008;105(42):16242-16247. doi:10.1073/pnas.0808031105
232. Mizrahi M, Cal P, Rosenthal M, et al. Human Alpha-1-Antitrypsin Modifies B Lymphocyte Responses During Allograft Transplantation. *Immunology*. Published online July 2013:n/a-n/a. doi:10.1111/imm.12149
233. Lewis EC. Expanding the Clinical Indications for α 1-Antitrypsin Therapy. *Mol Med*. 2012;18(6):957-970. doi:10.2119/molmed.2011.00196
234. Guttman O, Baranovski BM, Schuster R, et al. Acute-phase protein α 1-anti-trypsin: diverting injurious innate and adaptive immune responses from non-authentic threats. *Clin Exp Immunol*. 2015;179(2):161-172. doi:10.1111/cei.12476

235. Madar T, Shahaf G, Sheiner E, et al. Low levels of circulating alpha-1 antitrypsin are associated with spontaneous abortions. *J Matern Neonatal Med.* 2013;26(18):1782-1787. doi:10.3109/14767058.2013.801955
236. Baron J, Sheiner E, Abecassis A, et al. α 1-Antitrypsin insufficiency is a possible contributor to preterm premature rupture of membranes. *J Matern Neonatal Med.* 2012;25(7):934-937. doi:10.3109/14767058.2011.600369
237. Miyamoto Y, Akaike T, Maeda H. S-Nitrosylated human α 1-protease inhibitor. *Biochim Biophys Acta - Protein Struct Mol Enzymol.* 2000;1477(1-2):90-97. doi:10.1016/S0167-4838(99)00264-2
238. Kaner Z, Engelman R, Schuster R, et al. S-Nitrosylation of α 1-Antitrypsin Triggers Macrophages Toward Inflammatory Phenotype and Enhances Intra-Cellular Bacteria Elimination. *Front Immunol.* 2019;10. doi:10.3389/fimmu.2019.00590
239. Lior Y, Zaretsky M, Ochayon DE, et al. Point Mutation of a Non-Elastase-Binding Site in Human α 1-Antitrypsin Alters Its Anti-Inflammatory Properties. *Front Immunol.* 2018;9. doi:10.3389/fimmu.2018.00759
240. Lior Y, Shtriker E, Kahremany S, Lewis EC, Gruzman A. Development of anti-inflammatory peptidomimetics based on the structure of human alpha1-antitrypsin. *Eur J Med Chem.* 2022;228:113969. doi:10.1016/j.ejmech.2021.113969
241. Kemper C, Chan AC, Green JM, Brett KA, Murphy KM, Atkinson JP. Activation of human CD4+ cells with CD3 and CD46 induces a T-regulatory cell 1 phenotype. *Nature.* 2003;421(6921):388-392. doi:10.1038/nature01315
242. O'Dwyer CA, O'Brien ME, Wormald MR, et al. The BLT1 Inhibitory Function of α -1 Antitrypsin Augmentation Therapy Disrupts Leukotriene B4 Neutrophil Signaling. *J Immunol.* 2015;195(8):3628-3641. doi:10.4049/jimmunol.1500038
243. Moreno J-A, Ortega-Gomez A, Rubio-Navarro A, et al. High-Density Lipoproteins Potentiate α 1 -Antitrypsin Therapy in Elastase-Induced Pulmonary Emphysema. *Am J Respir Cell Mol Biol.* 2014;51(4):536-549. doi:10.1165/rcmb.2013-0103OC
244. Ochayon DE, Mizrahi M, Shahaf G, Baranovski BM, Lewis EC. Human α 1-Antitrypsin Binds to Heat-Shock Protein gp96 and Protects from Endogenous gp96-Mediated Injury In vivo. *Front Immunol.* 2013;4. doi:10.3389/fimmu.2013.00320
245. Finotti P, Pagetta A. A heat shock protein70 fusion protein with α 1-antitrypsin in plasma of Type 1 diabetic subjects. *Biochem Biophys Res Commun.* 2004;315(2):297-305. doi:10.1016/j.bbrc.2004.01.058
246. Ohayon DE, Schuster R, Mizrahi MI, et al. Human α 1-antitrypsin attenuates injury-induced Inflammation through interacting with high mobility group box-1 (HMGB1). *J Immunol.* 2017;198(1_Supplement):82.31-82.31. doi:10.4049/jimmunol.198.Supp.82.31
247. Karnaukhova E, Krupnikova SS, Rajabi M, Alayash AI. Heme binding to human alpha-1 proteinase inhibitor. *Biochim Biophys Acta - Gen Subj.* 2012;1820(12):2020-2029. doi:10.1016/j.bbagen.2012.09.012
248. Madyaningrana K, Vijayan V, Nikolin C, et al. Alpha1-antitrypsin counteracts heme-induced endothelial cell inflammatory activation, autophagy dysfunction and death. *Redox Biol.* 2021;46:102060. doi:10.1016/j.redox.2021.102060
249. Frenzel E, Wrenger S, Brügger B, et al. α 1-Antitrypsin Combines with Plasma Fatty Acids and Induces Angiopoietin-like Protein 4 Expression. *J Immunol.* 2015;195(8):3605-3616. doi:10.4049/jimmunol.1500740
250. Varga T, Czimmerer Z, Nagy L. PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. *Biochim Biophys Acta - Mol Basis Dis.* 2011;1812(8):1007-1022. doi:10.1016/j.bbadis.2011.02.014
251. Karnaukhova E. Interactions of α 1-proteinase inhibitor with small ligands of therapeutic potential: binding with retinoic acid. *Amino Acids.* 2010;38(4):1011-1020. doi:10.1007/s00726-009-0309-9

252. Liu PT, Krutzik SR, Kim J, Modlin RL. Cutting Edge: All- trans Retinoic Acid Down-Regulates TLR2 Expression and Function. *J Immunol.* 2005;174(5):2467-2470. doi:10.4049/jimmunol.174.5.2467
253. Xu Q, Jia H, Ma L, et al. All-trans retinoic acid inhibits lipopolysaccharide-induced inflammatory responses in bovine adipocytes via TGF β 1/Smad3 signaling pathway. *BMC Vet Res.* 2019;15(1):48. doi:10.1186/s12917-019-1791-2
254. Bakdash G, Vogelpoel LT, van Capel TM, Kapsenberg ML, de Jong EC. Retinoic acid primes human dendritic cells to induce gut-homing, IL-10-producing regulatory T cells. *Mucosal Immunol.* 2015;8(2):265-278. doi:10.1038/mi.2014.64
255. Kliewer SA, Umesono K, Noonan DJ, Heyman RA, Evans RM. Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. *Nature.* 1992;358(6389):771-774. doi:10.1038/358771a0
256. Rühl R, Krzyżosiak A, Niewiadomska-Cimicka A, et al. 9-cis-13,14-Dihydroretinoic Acid Is an Endogenous Retinoid Acting as RXR Ligand in Mice. McCaffery P, ed. *PLOS Genet.* 2015;11(6):e1005213. doi:10.1371/journal.pgen.1005213
257. Raghuram S, Stayrook KR, Huang P, et al. Identification of heme as the ligand for the orphan nuclear receptors REV-ERB α and REV-ERB β . *Nat Struct Mol Biol.* 2007;14(12):1207-1213. doi:10.1038/nsmb1344
258. Mashiba S, Wada Y, Takeya M, et al. In Vivo Complex Formation of Oxidized α 1 - Antitrypsin and LDL. *Arterioscler Thromb Vasc Biol.* 2001;21(11):1801-1808. doi:10.1161/hq1101.098232
259. Diffenderfer MR, Schaefer EJ. The composition and metabolism of large and small LDL. *Curr Opin Lipidol.* 2014;25(3):221-226. doi:10.1097/MOL.0000000000000067
260. May P. The low-density lipoprotein receptor-related protein 1 in inflammation. *Curr Opin Lipidol.* 2013;24(2):134-137. doi:10.1097/MOL.0b013e32835e809c
261. Laudati E, Gilder AS, Lam MS, et al. The activities of LDL Receptor-related Protein-1 (LRP1) compartmentalize into distinct plasma membrane microdomains. *Mol Cell Neurosci.* 2016;76:42-51. doi:10.1016/j.mcn.2016.08.006
262. Schubert K, Collins LE, Green P, Nagase H, Troeberg L. LRP1 Controls TNF Release via the TIMP-3/ADAM17 Axis in Endotoxin-Activated Macrophages. *J Immunol.* 2019;202(5):1501-1509. doi:10.4049/jimmunol.1800834
263. Jardí M, Inglés-Esteve J, Burgal M, et al. Distinct patterns of urokinase receptor (uPAR) expression by leukemic cells and peripheral blood cells. *Thromb Haemost.* 1996;76(6):1009-1019.
264. Mantuano E, Brifault C, Lam MS, Azmoon P, Gilder AS, Gonias SL. LDL receptor-related protein-1 regulates NF κ B and microRNA-155 in macrophages to control the inflammatory response. *Proc Natl Acad Sci.* 2016;113(5):1369-1374. doi:10.1073/pnas.1515480113
265. Pocivavsek A, Mikhailenko I, Strickland DK, Rebeck GW. Microglial low-density lipoprotein receptor-related protein 1 modulates c-Jun N-terminal kinase activation. *J Neuroimmunol.* 2009;214(1-2):25-32. doi:10.1016/j.jneuroim.2009.06.010
266. Inubushi T, Kawazoe A, Miyauchi M, et al. Molecular Mechanisms of the Inhibitory Effects of Bovine Lactoferrin on Lipopolysaccharide-mediated Osteoclastogenesis. *J Biol Chem.* 2012;287(28):23527-23536. doi:10.1074/jbc.M111.324673
267. Mishra A, Yao X, Saxena A, et al. Low-density lipoprotein receptor-related protein 1 attenuates house dust mite-induced eosinophilic airway inflammation by suppressing dendritic cell-mediated adaptive immune responses. *J Allergy Clin Immunol.* 2018;142(4):1066-1079.e6. doi:10.1016/j.jaci.2017.10.044
268. Toldo S, Austin D, Mauro AG, et al. Low-Density Lipoprotein Receptor-Related Protein-1 Is a Therapeutic Target in Acute Myocardial Infarction. *JACC Basic to Transl Sci.* 2017;2(5):561-574. doi:10.1016/j.jacbts.2017.05.007
269. Griffiths SW, Cooney CL. Relationship between Protein Structure and Methionine Oxidation in Recombinant Human α 1-Antitrypsin. *Biochemistry.* 2002;41(20):6245-6252. doi:10.1021/bi025599p

270. Griffiths SW, King J, Cooney CL. The Reactivity and Oxidation Pathway of Cysteine 232 in Recombinant Human α 1-Antitrypsin. *J Biol Chem*. 2002;277(28):25486-25492. doi:10.1074/jbc.M203089200
271. Saklatvala J, Wood GC, White DD. Isolation and characterization of human plasma α 1-proteinase inhibitor and a conformational study of its interaction with proteinases. *Biochem J*. 1976;157(2):339-351. doi:10.1042/bj1570339
272. Gorovoy M, Gaultier A, Campana WM, Firestein GS, Gonias SL. Inflammatory mediators promote production of shed LRP1/CD91, which regulates cell signaling and cytokine expression by macrophages. *J Leukoc Biol*. 2010;88(4):769-778. doi:10.1189/jlb.0410220
273. Bai X, Bai A, Tomasicchio M, et al. α 1-Antitrypsin Binds to the Glucocorticoid Receptor with Anti-Inflammatory and Antimycobacterial Significance in Macrophages. *J Immunol*. 2022;209(9):1746-1759. doi:10.4049/jimmunol.2200227
274. Rowe RG, Weiss SJ. Navigating ECM Barriers at the Invasive Front: The Cancer Cell–Stroma Interface. *Annu Rev Cell Dev Biol*. 2009;25(1):567-595. doi:10.1146/annurev.cellbio.24.110707.175315
275. Friedl P, Wolf K. Tube Travel: The Role of Proteases in Individual and Collective Cancer Cell Invasion. *Cancer Res*. 2008;68(18):7247-7249. doi:10.1158/0008-5472.CAN-08-0784
276. Wolf K, Wu YI, Liu Y, et al. Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. *Nat Cell Biol*. 2007;9(8):893-904. doi:10.1038/ncb1616
277. Chio LF, Oon CJ. Changes in serum alpha1 antitrypsin, alpha1 acid glycoprotein and beta2 glycoprotein i in patients with malignant hepatocellular carcinoma. *Cancer*. 1979;43(2):596-604. doi:10.1002/1097-0142(197902)43:2<596::AID-CNCR2820430229>3.0.CO;2-R
278. Trichopoulos D, Tzonou A, Kalapothaki V, Sparos L, Kremastinou T, Skoutari M. Alpha1-antitrypsin and survival in pancreatic cancer. *Int J Cancer*. 1990;45(4):685-686. doi:10.1002/ijc.2910450419
279. Szepehcinski A, Poplawska-Wisniewska B, Debek E, et al. The evaluation of biological and clinical significance of alpha-1 antitrypsin in non-small cell lung cancer. In: *3.1 Molecular Pathology and Functional Genomics*. European Respiratory Society; 2015:PA4883. doi:10.1183/13993003.congress-2015.PA4883
280. El-Akawi ZJ, Abu-Awad AM, Sharara AM, Khader Y. The importance of alpha-1 antitrypsin (alpha1-AT) and neopterin serum levels in the evaluation of non-small cell lung and prostate cancer patients. *Neuro Endocrinol Lett*. 2010;31(1):113-116.
281. Niksić-Ivancić M, Oresković M. The role of alpha-1-antitrypsin in patients with laryngeal carcinoma. *Laryngol Rhinol Otol (Stuttg)*. 1983;62(6):246-248. doi:6412008
282. Konnova LA, Konnov BA, Mel'nikov LA, Zargarova OP. Antiprotease activity of the blood serum of patients with the disseminated form of breast cancer before and after combined treatment. *Vopr Onkol*. 1983;29(10):7-9.
283. Thompson DK, Haddow JE, Smith DE, Ritchie RF. Elevated serum acute phase protein levels as predictors of disseminated breast cancer. *Cancer*. 1983;51(11):2100-2104. doi:10.1002/1097-0142(19830601)51:11<2100::AID-CNCR2820511124>3.0.CO;2-J
284. Bhattacharyya J, Chaudhuri L. α 1-Antitrypsin: A possible tool for diagnosis of cervical cancer. *Biochem Med Metab Biol*. 1990;43(3):243-245. doi:10.1016/0885-4505(90)90030-5
285. Urquidi V, Goodison S, Ross S, Chang M, Dai Y, Rosser CJ. Diagnostic Potential of Urinary α 1-Antitrypsin and Apolipoprotein E in the Detection of Bladder Cancer. *J Urol*. 2012;188(6):2377-2383. doi:10.1016/j.juro.2012.07.094
286. Tahara E, Ito H, Taniyama K, Yokozaki H, Hata J. Alpha1-antitrypsin, alpha1-antichymotrypsin, and alpha2-macroglobulin in human gastric carcinomas: A retrospective immunohistochemical study. *Hum Pathol*. 1984;15(10):957-964. doi:10.1016/S0046-8177(84)80125-2

287. Higashiyama M, Doi O, Kodama K, Yokouchi H, Tateishi R. An evaluation of the prognostic significance of alpha-1-antitrypsin expression in adenocarcinomas of the lung: an immunohistochemical analysis. *Br J Cancer*. 1992;65(2):300-302. doi:10.1038/bjc.1992.60
288. Fukushima M, Fukuda Y, Kawamoto M, Yamanaka N. Elastosis in lung carcinoma: Immunohistochemical, ultrastructural and clinical studies. *Pathol Int*. 2000;50(8):626-635. doi:10.1046/j.1440-1827.2000.01103.x
289. Li Y, Miao L, Yu M, et al. α 1-antitrypsin promotes lung adenocarcinoma metastasis through upregulating fibronectin expression. *Int J Oncol*. 2017;50(6):1955-1964. doi:10.3892/ijo.2017.3962
290. El-Akawi. Alpha-1 Antitrypsin Blood Levels as Indicator for the Efficacy of Cancer Treatment. *World J Oncol*. Published online 2013. doi:10.4021/wjon663e
291. Huang H, Campbell SC, Nelius T, et al. α 1-antitrypsin inhibits angiogenesis and tumor growth. *Int J Cancer*. 2004;112(6):1042-1048. doi:10.1002/ijc.20494
292. Finlay TH, Tamir S, Kadner SS, Cruz MR, Yavelow J, Levitz M. alpha 1-Antitrypsin- and anchorage-independent growth of MCF-7 breast cancer cells. *Endocrinology*. 1993;133(3):996-1002. doi:10.1210/endo.133.3.8365378
293. Yavelow J, Tuccillo A, Kadner SS, Katz J, Finlay TH. α 1-Antitrypsin Blocks the Release of Transforming Growth Factor- α from MCF-7 Human Breast Cancer Cells1. *J Clin Endocrinol Metab*. 1997;82(3):745-752. doi:10.1210/jcem.82.3.3818
294. Zelvyte I, Stevens T, Westin U, Janciauskiene S. alpha1-antitrypsin and its C-terminal fragment attenuate effects of degranulated neutrophil-conditioned medium on lung cancer HCC cells, in vitro. *Cancer Cell Int*. 2004;4(1):7. doi:10.1186/1475-2867-4-7
295. Chang Y-H, Lee S-H, Liao I-C, Huang S-H, Cheng H-C, Liao P-C. Secretomic Analysis Identifies Alpha-1 Antitrypsin (A1AT) as a Required Protein in Cancer Cell Migration, Invasion, and Pericellular Fibronectin Assembly for Facilitating Lung Colonization of Lung Adenocarcinoma Cells. *Mol Cell Proteomics*. 2012;11(11):1320-1339. doi:10.1074/mcp.M112.017384
296. Kwon CH, Park HJ, Choi JH, et al. Snail and serpinA1 promote tumor progression and predict prognosis in colorectal cancer. *Oncotarget*. 2015;6(24):20312-20326. doi:10.18632/oncotarget.3964
297. Chang S-H, Cho K-C, Yu K-N, et al. Alpha 1-antitrypsin activates lung cancer cell survival by acting on cap-dependent protein translation, vesicle-mediated transport, and metastasis. *Oncotarget*. Published online July 19, 2016. doi:10.18632/oncotarget.10695
298. Wu D, Liu T, Deng S, et al. Alpha-1 Antitrypsin Induces Epithelial-to-Mesenchymal Transition, Endothelial-to-Mesenchymal Transition, and Drug Resistance in Lung Cancer Cells. *Onco Targets Ther*. 2020;Volume 13:3751-3763. doi:10.2147/OTT.S242579
299. Graziani V, Rodriguez-Hernandez I, Maiques O, Sanz-Moreno V. The amoeboid state as part of the epithelial-to-mesenchymal transition programme. *Trends Cell Biol*. 2022;32(3):228-242. doi:10.1016/j.tcb.2021.10.004
300. Rostenberg I, Guízar-Vázquez J, Peñaloza R. Altered carbohydrate content of alpha1-antitrypsin in patients with cancer. *J Natl Cancer Inst*. 1978;61(4):961-965.
301. Abd Hamid UM, Royle L, Saldova R, et al. A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression. *Glycobiology*. 2008;18(12):1105-1118. doi:10.1093/glycob/cwn095
302. Comunale MA, Rodemich-Betesh L, Hafner J, et al. Linkage Specific Fucosylation of Alpha-1-Antitrypsin in Liver Cirrhosis and Cancer Patients: Implications for a Biomarker of Hepatocellular Carcinoma. Ryu W-S, ed. *PLoS One*. 2010;5(8):e12419. doi:10.1371/journal.pone.0012419
303. Ochiai A, Yasui W, Tahara E. Growth-promoting effect of gastrin on human gastric carcinoma cell line TMK-1. *Jpn J Cancer Res*. 1985;76(11):1064-1071.

304. Perlmutter DH, Daniels JD, Auerbach HS, De Schryver-Kecsckemeti K, Winter HS, Alpers DH. The alpha 1-antitrypsin gene is expressed in a human intestinal epithelial cell line. *J Biol Chem*. 1989;264(16):9485-9490.
305. Tamir S, Kadner SS, Katz J, Finlay TH. Regulation of Antitrypsin and Antichymotrypsin Synthesis by MCF-7 Breast Cancer Cell Sublines*. *Endocrinology*. 1990;127(3):1319-1328. doi:10.1210/endo-127-3-1319
306. DeVito NC, Plebanek MP, Theivanthiran B, Hanks BA. Role of Tumor-Mediated Dendritic Cell Tolerization in Immune Evasion. *Front Immunol*. 2019;10. doi:10.3389/fimmu.2019.02876
307. Lejeune P-J, Mallet B, Farnarier C, Kaplanski S. Changes in serum level and affinity for concanavalin A of human α 1-proteinase inhibitor in severe burn patients: relationship to natural killer cell activity. *Biochim Biophys Acta - Gen Subj*. 1989;990(2):122-127. doi:10.1016/S0304-4165(89)80022-4
308. Wozniak A, Mila-Kierzenkowska C, Schachtschabel DO, et al. Activity of cathepsin D and alpha(1)-antitrypsin in the blood serum of patients with mammary carcinoma. *Exp Oncol*. 2005;27(3):233-237.
309. Woźniak B, Mila-Kierzenkowska C, Woźniak A, et al. The effect of combined therapy on activity of cathepsin D and alpha-1-antitrypsin in the blood serum of women with cervical cancer. *Eur J Gynaecol Oncol*. 2008;29(6):617-619.
310. Zietek Z, Iwan-Zietek I, Kotschy M, Tyloch F. [Activity of alpha 1-antitrypsin in blood of patients with prostatic tumor]. *Pol Tyg Lek*. 1996;51(6-9):89-90.
311. Mohammad S, Rastgar F, Sahebghadam-Lotfi A. Structural Features, Biological Functions of the Alpha-1 Antitrypsin and Contribution to Esophageal Cancer. In: *Squamous Cell Carcinoma*. InTech; 2012. doi:10.5772/30476
312. Jaberie H, Naghibalhossaini F. Recombinant production of native human α -1-antitrypsin protein in the liver HepG2 cells. *Biotechnol Lett*. 2016;38(10):1683-1690. doi:10.1007/s10529-016-2150-z
313. Zelvyte I, Sjögren H-O, Janciauskiene S. Effects of native and cleaved forms of α 1-antitrypsin on ME 1477 tumor cell functional activity. *Cancer Detect Prev*. 2002;26(4):256-265. doi:10.1016/S0361-090X(02)00090-9
314. Ossanna PJ, Test ST, Matheson NR, Regiani S, Weiss SJ. Oxidative regulation of neutrophil elastase-alpha-1-proteinase inhibitor interactions. *J Clin Invest*. 1986;77(6):1939-1951. doi:10.1172/JCI112523
315. Li W, Savinov AY, Rozanov D V., et al. Matrix Metalloproteinase-26 Is Associated with Estrogen-Dependent Malignancies and Targets α 1-Antitrypsin Serpin. *Cancer Res*. 2004;64(23):8657-8665. doi:10.1158/0008-5472.CAN-04-3019
316. Joslin G, Griffin GL, August AM, et al. The serpin-enzyme complex (SEC) receptor mediates the neutrophil chemotactic effect of alpha-1 antitrypsin-elastase complexes and amyloid-beta peptide. *J Clin Invest*. 1992;90(3):1150-1154. doi:10.1172/JCI115934
317. Janciauskiene S, Zelvyte I, Jansson L, Stevens T. Divergent effects of α 1-antitrypsin on neutrophil activation, in vitro. *Biochem Biophys Res Commun*. 2004;315(2):288-296. doi:10.1016/j.bbrc.2004.01.055
318. Moraga F, Lindgren S, Janciauskiene S. Effects of Noninhibitory α -1-Antitrypsin on Primary Human Monocyte Activation in Vitro. *Arch Biochem Biophys*. 2001;386(2):221-226. doi:10.1006/abbi.2000.2211
319. Janciauskiene S, Moraga F, Lindgren S. C-terminal fragment of α 1-antitrypsin activates human monocytes to a pro-inflammatory state through interactions with the CD36 scavenger receptor and LDL receptor. *Atherosclerosis*. 2001;158(1):41-51. doi:10.1016/S0021-9150(00)00767-X
320. Subramaniam D, Glader P, Vonwachenfeldt K, Burneckiene J, Stevens T, Janciauskiene S. C-36 peptide, a degradation product of α 1-antitrypsin, modulates human monocyte activation through LPS signaling pathways. *Int J Biochem Cell Biol*. 2006;38(4):563-575. doi:10.1016/j.biocel.2005.09.021

321. Johansson J, Gröndal S, Sjövall J, Jörnvall H, Curstedt T. Identification of hydrophobic fragments of α 1 -antitrypsin and Cl protease inhibitor in human bile, plasma and spleen. *FEBS Lett.* 1992;299(2):146-148. doi:10.1016/0014-5793(92)80234-8
322. Blaurock-Möller N, Gröger M, Siwczak F, et al. CAAP48, a New Sepsis Biomarker, Induces Hepatic Dysfunction in an in vitro Liver-on-Chip Model. *Front Immunol.* 2019;10. doi:10.3389/fimmu.2019.00273
323. Bigalke A, Neu C, Esper Treml R, Coldewey SM, Kiehntopf M. Fragments of alpha-1-antitrypsin in patients with severe COVID-19 and bacterial pulmonary sepsis. *Clin Chem Lab Med.* 2022;60(8):e187-e189. doi:10.1515/cclm-2022-0361
324. Bigalke A, Sponholz C, Schnabel C, Bauer M, Kiehntopf M. Multiplex quantification of C-terminal alpha-1-antitrypsin peptides provides a novel approach for characterizing systemic inflammation. *Sci Rep.* 2022;12(1):3844. doi:10.1038/s41598-022-07752-w
325. Takenouchi T, Munekata E. β -Amyloid peptide, substance P, and SEC receptor ligand activate cytoplasmic Ca²⁺ in neutrophil-like HL-60 cells: Effect of chemotactic peptide antagonist BocMLF. *Peptides.* 1995;16(6):1019-1024. doi:10.1016/0196-9781(95)00084-W
326. Cercek L, Carr BI, Siaw M, Cercek B. Effect of the cancer recognition, immunodefense suppression, and serine protease protection peptide on DNA synthesis in rat hepatocytes and human lymphocytes. *Cancer Detect Prev.* 1995;19(2):206-209.
327. Zelvyte I, Lindgren S, Janciauskiene S. Multiple effects of α 1-antitrypsin on breast carcinoma MDA-MB 468 cell growth and invasiveness. *Eur J Cancer Prev.* 2003;12(2):117-124. doi:10.1097/00008469-200304000-00005
328. Kataoka H, Uchino H, Iwamura T, Seiki M, Nabeshima K, Koono M. Enhanced Tumor Growth and Invasiveness in Vivo by a Carboxyl-Terminal Fragment of α 1-Proteinase Inhibitor Generated by Matrix Metalloproteinases. *Am J Pathol.* 1999;154(2):457-468. doi:10.1016/S0002-9440(10)65292-3
329. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001;29(9):45e - 45. doi:10.1093/nar/29.9.e45
330. Rio DC, Ares M Jr, Hannon GJ NT. *RNA: A Laboratory Manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2011. 171-179 (protocol №20); 180-185 (protocol № 21)
331. Adam SA, Marr RS, Gerace L. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol.* 1990;111(3):807-816. doi:10.1083/jcb.111.3.807
332. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Battistuzzi FU, ed. *Mol Biol Evol.* 2018;35(6):1547-1549. doi:10.1093/molbev/msy096
333. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007;23(21):2947-2948. doi:10.1093/bioinformatics/btm404
334. Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 2012;28(8):1166-1167. doi:10.1093/bioinformatics/bts091
335. Yang Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol.* 2007;24(8):1586-1591. doi:10.1093/molbev/msm088
336. Wernersson R. RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res.* 2003;31(13):3537-3539. doi:10.1093/nar/gkg609
337. Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol.* 1985;22(2):160-174. doi:10.1007/BF02101694
338. Haeussler M, Schönig K, Eckert H, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 2016;17(1):148. doi:10.1186/s13059-016-1012-2
339. Lomov NA, Viushkov VS, Zamalutdinov A V., Sboeva MD, Rubtsov MA. Direct ENIT: An easy and reliable tool for gRNA efficacy verification by tracking induced chromosomal translocation. *MethodsX.* 2020;7:101104. doi:10.1016/j.mex.2020.101104

340. Cercek L, Siaw M, Cercek B, Cercek B. A DNA probe study on the origin of the cancer recognition, immunodefense suppression and serine protease protection peptide. *Cancer Detect Prev.* 1995;19(4):325-330.
341. Ivanov IP, Loughran G, Sachs MS, Atkins JF. Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci.* 2010;107(42):18056-18060. doi:10.1073/pnas.1009269107
342. Lee S, Liu B, Lee S, Huang S-X, Shen B, Qian S-B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci.* 2012;109(37). doi:10.1073/pnas.1207846109
343. Guerra-Almeida D, Tschoeke DA, Nunes-da-Fonseca R. Understanding small ORF diversity through a comprehensive transcription feature classification. *DNA Res.* 2021;28(5). doi:10.1093/dnares/dsab007
344. Matamala N, Aggarwal N, Iadarola P, et al. Identification of Novel Short C-Terminal Transcripts of Human *SERPINA1* Gene. Strnad P, ed. *PLoS One.* 2017;12(1):e0170533. doi:10.1371/journal.pone.0170533
345. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 1987;15(20):8125-8148. doi:10.1093/nar/15.20.8125
346. Noderer WL, Flockhart RJ, Bhaduri A, Diaz de Arce AJ, Zhang J, Khavari PA, Wang CL. Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol Syst Biol.* 2014;10(8):748. doi:10.15252/msb.20145136
347. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 1986;44(2):283-292. doi:10.1016/0092-8674(86)90762-2
348. Sergiev P V., Rubtsova MP. Little but Loud. The Diversity of Functions of Small Proteins and Peptides – Translational Products of Short Reading Frames. *Biochem.* 2021;86(9):1139-1150. doi:10.1134/S0006297921090091
349. Kozak M. Initiation of translation in prokaryotes and eukaryotes. *Gene.* 1999;234(2):187-208. doi:10.1016/S0378-1119(99)00210-3
350. Smith E. Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation in vivo. *Nucleic Acids Res.* 2005;33(4):1298-1308. doi:10.1093/nar/gki248
351. Kochetov A V. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *BioEssays.* 2008;30(7):683-691. doi:10.1002/bies.20771
352. Kozak M. Pushing the limits of the scanning mechanism for initiation of translation. *Gene.* 2002;299(1-2):1-34. doi:10.1016/s0378-1119(02)01056-9
353. Riley JH, Bathurst IC, Edbrooke MR, Carrell RW, Craig RK. $\alpha 1$ -Antitrypsin and serum albumin mRNA accumulation in normal, acute phase and ZZ human liver. *FEBS Lett.* 1985;189(2):361-366. doi:10.1016/0014-5793(85)81056-5
354. Kurachi K, Chandra T, Degen SJ, et al. Cloning and sequence of cDNA coding for alpha 1-antitrypsin. *Proc Natl Acad Sci.* 1981;78(11):6826-6830. doi:10.1073/pnas.78.11.6826
355. Coutelle C, Speer A, Rogers J, Kalsheker N, Humphries S, Williamson R. Construction and partial characterization of a human liver cDNA library. *Biomed Biochim Acta.* 1985;44(3):421-431.
356. Kononikhin AS, Starodubtseva NL, Bugrova AE, et al. An untargeted approach for the analysis of the urine peptidome of women with preeclampsia. *J Proteomics.* 2016;149:38-43. doi:10.1016/j.jprot.2016.04.024
357. Starodubtseva N, Nizyaeva N, Baev O, et al. SERPINA1 Peptides in Urine as A Potential Marker of Preeclampsia Severity. *Int J Mol Sci.* 2020;21(3):914. doi:10.3390/ijms21030914
358. Shan D, Wang H, Khatri P, et al. The Urinary Peptidome as a Noninvasive Biomarker Development Strategy for Prenatal Screening of Down's Syndrome. *Omi A J Integr Biol.* 2019;23(9):439-447. doi:10.1089/omi.2019.0098
359. Wei Y, Wang YG, Jia Y, et al. Liver homeostasis is maintained by midlobular zone 2 hepatocytes. *Science (80-).* 2021;371(6532). doi:10.1126/science.abb1625

360. Liu Y, Chen Y-G. Intestinal epithelial plasticity and regeneration via cell dedifferentiation. *Cell Regen.* 2020;9(1):14. doi:10.1186/s13619-020-00053-5
361. Berger K, Bangen J-M, Hammerich L, et al. Origin of regenerating tubular cells after acute kidney injury. *Proc Natl Acad Sci.* 2014;111(4):1533-1538. doi:10.1073/pnas.1316177111
362. Hassan T, Smith SGJ, Gaughan K, et al. Isolation and identification of cell-specific microRNAs targeting a messenger RNA using a biotinylated anti-sense oligonucleotide capture affinity technique. *Nucleic Acids Res.* 2013;41(6):e71-e71. doi:10.1093/nar/gks1466
363. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The Antisense Transcriptomes of Human Cells. *Science (80-)*. 2008;322(5909):1855-1857. doi:10.1126/science.1163853
364. Osato N, Suzuki Y, Ikeo K, Gojobori T. Transcriptional Interferences in cis Natural Antisense Transcripts of Humans and Mice. *Genetics.* 2007;176(2):1299-1306. doi:10.1534/genetics.106.069484
365. Faghihi MA, Wahlestedt C. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol.* 2009;10(9):637-643. doi:10.1038/nrm2738
366. Faghihi MA, Zhang M, Huang J, et al. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* 2010;11(5):R56. doi:10.1186/gb-2010-11-5-r56
367. Qin H, Ni H, Liu Y, et al. RNA-binding proteins in tumor progression. *J Hematol Oncol.* 2020;13(1):90. doi:10.1186/s13045-020-00927-w
368. Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. *Science (80-)*. 2017;357(6352). doi:10.1126/science.aan2507
369. Ward AM, Cooper EH, Houghton AL. Acute phase reactant proteins in prostatic cancer. *Br J Urol.* 1977;49(5):411-418. doi:10.1111/j.1464-410x.1977.tb04168.x
370. Kuvibidila S, Rayford W. Correlation between serum prostate-specific antigen and alpha-1-antitrypsin in men without and with prostate cancer. *J Lab Clin Med.* 2006;147(4):174-181. doi:10.1016/j.lab.2005.11.012
371. Zhang WM, Finne P, Leinonen J, Stenman UH. Characterization and determination of the complex between prostate-specific antigen and alpha 1-protease inhibitor in benign and malignant prostatic diseases. *Scand J Clin Lab Invest Suppl.* 2000;233:51-58.
372. Scaccianoce E, Festuccia C, Dondi D, et al. Characterization of Prostate Cancer DU145 Cells Expressing the Recombinant Androgen Receptor. *Oncol Res Featur Preclin Clin Cancer Ther.* 2003;14(2):101-112. doi:10.3727/000000003108748658
373. Thiagalingam S. Epigenetic memory in development and disease: Unraveling the mechanism. *Biochim Biophys Acta - Rev Cancer.* 2020;1873(2):188349. doi:10.1016/j.bbcan.2020.188349
374. Hatzoglou V, Patel G V., Morris MJ, et al. Brain Metastases from Prostate Cancer: An 11-Year Analysis in the MRI Era with Emphasis on Imaging Characteristics, Incidence, and Prognosis. *J Neuroimaging.* 2014;24(2):161-166. doi:10.1111/j.1552-6569.2012.00767.x
375. Valiente M, Obenaus AC, Jin X, et al. Serpins Promote Cancer Cell Survival and Vascular Co-Option in Brain Metastasis. *Cell.* 2014;156(5):1002-1016. doi:10.1016/j.cell.2014.01.040
376. Zhong H, Davis A, Ouzounova M, et al. A Novel IL6 Antibody Sensitizes Multiple Tumor Types to Chemotherapy Including Trastuzumab-Resistant Tumors. *Cancer Res.* 2016;76(2):480-490. doi:10.1158/0008-5472.CAN-15-0883
377. Lou W, Ni Z, Dyer K, Twardy DJ, Gao AC. Interleukin-6 induces prostate cancer cell growth accompanied by activation of Stat3 signaling pathway. *Prostate.* 2000;42(3):239-242. doi:10.1002/(SICI)1097-0045(20000215)42:3<239::AID-PROS10>3.0.CO;2-G
378. Chen M, Wei R, Wei G, et al. Systematic evaluation of the effect of polyadenylation signal variants on the expression of disease-associated genes. *Genome Res.* 2021;31(5):890-899. doi:10.1101/gr.270256.120
379. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013;45(6):580-585. doi:10.1038/ng.2653

380. Saito S, Hosoda N, Hoshino S. The Hbs1-Dom34 Protein Complex Functions in Non-stop mRNA Decay in Mammalian Cells. *J Biol Chem.* 2013;288(24):17832-17843. doi:10.1074/jbc.M112.448977
381. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating Cells Express mRNAs with Shortened 3' Untranslated Regions and Fewer MicroRNA Target Sites. *Science (80-).* 2008;320(5883):1643-1647. doi:10.1126/science.1155390
382. Mayr C, Bartel DP. Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell.* 2009;138(4):673-684. doi:10.1016/j.cell.2009.06.016
383. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci.* 2009;106(17):7028-7033. doi:10.1073/pnas.0900028106
384. Ji Z, Tian B. Reprogramming of 3' Untranslated Regions of mRNAs by Alternative Polyadenylation in Generation of Pluripotent Stem Cells from Different Cell Types. Valcarcel J, ed. *PLoS One.* 2009;4(12):e8419. doi:10.1371/journal.pone.0008419
385. Mangone M, Manoharan AP, Thierry-Mieg D, et al. The Landscape of *C. elegans* 3'UTRs. *Science (80-).* 2010;329(5990):432-435. doi:10.1126/science.1191244
386. Licatalosi DD, Mele A, Fak JJ, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature.* 2008;456(7221):464-469. doi:10.1038/nature07488
387. Matamala N, Lara B, Gómez-Mariano G, et al. miR-320c Regulates *SERPINA1* Expression and Is Induced in Patients With Pulmonary Disease. *Arch Bronconeumol.* 2021;57(7):457-463. doi:10.1016/j.arbres.2020.03.006
388. Greene CM, Marciniak SJ, Teckman J, et al. α 1-Antitrypsin deficiency. *Nat Rev Dis Prim.* 2016;2(1):16051. doi:10.1038/nrdp.2016.51
389. Xu X-R, Huang J, Xu Z-G, et al. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. *Proc Natl Acad Sci.* 2001;98(26):15089-15094. doi:10.1073/pnas.241522398
390. Shen L, Zhi L, Hu W, Wu MX. IEX-1 targets mitochondrial F1Fo-ATPase inhibitor for degradation. *Cell Death Differ.* 2009;16(4):603-612. doi:10.1038/cdd.2008.184
391. Yang C, Trent S, Ionescu-Tiba V, et al. Identification of Cyclin D1- and Estrogen-Regulated Genes Contributing to Breast Carcinogenesis and Progression. *Cancer Res.* 2006;66(24):11649-11658. doi:10.1158/0008-5472.CAN-06-1645
392. Garcia MN, Grasso D, Lopez-Millan MB, et al. IER3 supports KRASG12D-dependent pancreatic cancer development by sustaining ERK1/2 phosphorylation. *J Clin Invest.* 2014;124(11):4709-4722. doi:10.1172/JCI76037
393. Ye J, Zhang Y, Cai Z, et al. Increased expression of immediate early response gene 3 protein promotes aggressive progression and predicts poor prognosis in human bladder cancer. *BMC Urol.* 2018;18(1):82. doi:10.1186/s12894-018-0388-6
394. Wu MX, Ao Z, Prasad KVS, Wu R, Schlossman SF. IEX -1L, an Apoptosis Inhibitor Involved in NF- κ B-Mediated Cell Survival. *Science (80-).* 1998;281(5379):998-1001. doi:10.1126/science.281.5379.998
395. Forrest ARR, Kawaji H, Rehli M, et al. A promoter-level mammalian expression atlas. *Nature.* 2014;507(7493):462-470. doi:10.1038/nature13182
396. Arnold PR, Wells AD, Li XC. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Front Cell Dev Biol.* 2020;7. doi:10.3389/fcell.2019.00377
397. Chen HP, Lin A, Bloom JS, Khan AH, Park CC, Smith DJ. Screening reveals conserved and nonconserved transcriptional regulatory elements including an E3/E4 allele-dependent APOE coding region enhancer. *Genomics.* 2008;92(5):292-300. doi:10.1016/j.ygeno.2008.07.009
398. Barthel KKB, Liu X. A Transcriptional Enhancer from the Coding Region of ADAMTS5. Abraham E, ed. *PLoS One.* 2008;3(5):e2184. doi:10.1371/journal.pone.0002184

399. Dong X, Navratilova P, Fredman D, Drivenes Ø, Becker TS, Lenhard B. Exonic remnants of whole-genome duplication reveal cis-regulatory function of coding exons. *Nucleic Acids Res.* 2010;38(4):1071-1085. doi:10.1093/nar/gkp1124
400. Fejes-Toth K, Sotirova V, Sachidanandam R, et al. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature.* 2009;457(7232):1028-1032. doi:10.1038/nature07759
401. Otsuka Y, Kedersha NL, Schoenberg DR. Identification of a Cytoplasmic Complex That Adds a Cap onto 5'-Monophosphate RNA. *Mol Cell Biol.* 2009;29(8):2155-2167. doi:10.1128/MCB.01325-08
402. Mercer TR, Dinger ME, Bracken CP, et al. Regulated post-transcriptional RNA cleavage diversifies the eukaryotic transcriptome. *Genome Res.* 2010;20(12):1639-1650. doi:10.1101/gr.112128.110
403. Bennett CF, Swayze EE. RNA Targeting Therapeutics: Molecular Mechanisms of Antisense Oligonucleotides as a Therapeutic Platform. *Annu Rev Pharmacol Toxicol.* 2010;50(1):259-293. doi:10.1146/annurev.pharmtox.010909.105654
404. Karginov F V., Cheloufi S, Chong MMW, Stark A, Smith AD, Hannon GJ. Diverse Endonucleolytic Cleavage Sites in the Mammalian Transcriptome Depend upon MicroRNAs, Drosha, and Additional Nucleases. *Mol Cell.* 2010;38(6):781-788. doi:10.1016/j.molcel.2010.06.001
405. Cernilogar FM, Onorati MC, Kothe GO, et al. Chromatin-associated RNA interference components contribute to transcriptional regulation in *Drosophila*. *Nature.* 2011;480(7377):391-395. doi:10.1038/nature10492
406. Nepal C, Hadzhiev Y, Previti C, et al. Dynamic regulation of the transcription initiation landscape at single nucleotide resolution during vertebrate embryogenesis. *Genome Res.* 2013;23(11):1938-1950. doi:10.1101/gr.153692.112
407. Ingolia NT, Lareau LF, Weissman JS. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell.* 2011;147(4):789-802. doi:10.1016/j.cell.2011.10.002
408. Wan J, Qian S-B. TISdb: a database for alternative translation initiation in mammalian cells. *Nucleic Acids Res.* 2014;42(D1):D845-D850. doi:10.1093/nar/gkt1085
409. Sapkota D, Lake AM, Yang W, et al. Cell-Type-Specific Profiling of Alternative Translation Identifies Regulated Protein Isoform Variation in the Mouse Brain. *Cell Rep.* 2019;26(3):594-607.e7. doi:10.1016/j.celrep.2018.12.077
410. Freeze HH, Kranz C. Endoglycosidase and glycoamidase release of N-linked glycans. *Curr Protoc Mol Biol.* 2010;Chapter 17:Unit 17.13A. doi:10.1002/0471142727.mb1713as89
411. Franko A, Hartwig S, Kotzka J, et al. Identification of the Secreted Proteins Originated from Primary Human Hepatocytes and HepG2 Cells. *Nutrients.* 2019;11(8):1795. doi:10.3390/nu11081795
412. Lodish HF, Kong N. Glucose removal from N-linked oligosaccharides is required for efficient maturation of certain secretory glycoproteins from the rough endoplasmic reticulum to the Golgi complex. *J Cell Biol.* 1984;98(5):1720-1729. doi:10.1083/jcb.98.5.1720
413. Wei J, Zhang Y, Ivanov IP, Sachs MS. The Stringency of Start Codon Selection in the Filamentous Fungus *Neurospora crassa*. *J Biol Chem.* 2013;288(13):9549-9562. doi:10.1074/jbc.M112.447177
414. Wethmar K, Bégay V, Smink JJ, et al. C/EBP β Δ uORF mice—a genetic model for uORF-mediated translational control in mammals. *Genes Dev.* 2010;24(1):15-20. doi:10.1101/gad.557910
415. Barbour K, Wei F, Brannan C, Flotte T, Baumann H, Berger F. The Murine α -Proteinase Inhibitor Gene Family: Polymorphism, Chromosomal Location, and Structure. *Genomics.* 2002;80(5):515-522. doi:10.1016/S0888-7543(02)96864-3

416. Sytnikova YA, Kubarenko A V., Schäfer A, Weber ANR, Niehrs C. Gadd45a Is an RNA Binding Protein and Is Localized in Nuclear Speckles. Xu S, ed. *PLoS One*. 2011;6(1):e14500. doi:10.1371/journal.pone.0014500
417. Loebermann H, Tokuoka R, Deisenhofer J, Huber R. Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J Mol Biol*. 1984;177(3):531-557.
418. Elliott PR, Lomas DA, Carrell RW, Abrahams JP. Inhibitory conformation of the reactive loop of α 1-antitrypsin. *Nat Struct Mol Biol*. 1996;3(8):676-681. doi:10.1038/nsb0896-676
419. Im H, Woo M-S, Hwang KY, Yu M-H. Interactions Causing the Kinetic Trap in Serpin Protein Folding. *J Biol Chem*. 2002;277(48):46347-46354. doi:10.1074/jbc.M207682200
420. Johansen H, Sutiphong J, Sathe G, et al. High-level production of fully active human alpha 1-antitrypsin in Escherichia coli. *Mol Biol Med*. 1987;4(5):291-305.
421. Zhang Q, Buckle AM, Law RHP, et al. The N terminus of the serpin, tengpin, functions to trap the metastable native state. *EMBO Rep*. 2007;8(7):658-663. doi:10.1038/sj.embor.7400986
422. Lambers JW, Cammenga M, König BW, Mertens K, Pannekoek H, van Mourik JA. Activation of human endothelial cell-type plasminogen activator inhibitor (PAI-1) by negatively charged phospholipids. *J Biol Chem*. 1987;262(36):17492-17496.
423. Stout TJ, Graham H, Buckley DI, Matthews DJ. Structures of Active and Latent PAI-1: A Possible Stabilizing Role for Chloride Ions. *Biochemistry*. 2000;39(29):8460-8469. doi:10.1021/bi000290w
424. Vossaert L, Meert P, Scheerlinck E, et al. Identification of histone H3 clipping activity in human embryonic stem cells. *Stem Cell Res*. 2014;13(1):123-134. doi:10.1016/j.scr.2014.05.002
425. Sandoval-Basilio J, Serafín-Higuera N, Reyes-Hernandez OD, et al. Low Proteolytic Clipping of Histone H3 in Cervical Cancer. *J Cancer*. 2016;7(13):1856-1860. doi:10.7150/jca.15605
426. Ferrari KJ, Amato S, Noberini R, et al. Intestinal differentiation involves cleavage of histone H3 N-terminal tails by multiple proteases. *Nucleic Acids Res*. 2021;49(2):791-804. doi:10.1093/nar/gkaa1228
427. Li X, Zhang G, Ngo N, Zhao X, Kain SR, Huang C-C. Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *J Biol Chem*. 1997;272(45):28545-28549. doi:10.1074/jbc.272.45.28545
428. Kearse MG, Wilusz JE. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev*. 2017;31(17):1717-1731. doi:10.1101/gad.305250.117
429. Guo H, Sun J, Li X, et al. Positive charge in the n-region of the signal peptide contributes to efficient post-translational translocation of small secretory preproteins. *J Biol Chem*. 2018;293(6):1899-1907. doi:10.1074/jbc.RA117.000922
430. Tanudji M, Hevi S, Chuck SL. Improperly folded green fluorescent protein is secreted via a non-classical pathway. *J Cell Sci*. 2002;115(19):3849-3857. doi:10.1242/jcs.00047
431. Sherman F, Stewart JW, Tsunasawa S. Methionine or not methionine at the beginning of a protein. *BioEssays*. 1985;3(1):27-31. doi:10.1002/bies.950030108
432. Wingfield PT. N-Terminal Methionine Processing. *Curr Protoc Protein Sci*. 2017;88(1). doi:10.1002/cpps.29
433. Hinnebusch AG. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu Rev Biochem*. 2014;83(1):779-812. doi:10.1146/annurev-biochem-060713-035802
434. Bernhofer M, Dallago C, Karl T, et al. PredictProtein - Predicting Protein Structure and Function for 29 Years. *Nucleic Acids Res*. 2021;49(W1):W535-W540. doi:10.1093/nar/gkab354
435. Dowling D, Schmitz JF, Bornberg-Bauer E. Stochastic Gain and Loss of Novel Transcribed Open Reading Frames in the Human Lineage. Aoife M, ed. *Genome Biol Evol*. 2020;12(11):2183-2195. doi:10.1093/gbe/evaa194
436. Ferrarotti I, Carroll TP, Ottaviani S, et al. Identification and characterisation of eight novel *SERPINA1* Null mutations. *Orphanet J Rare Dis*. 2014;9(1):172. doi:10.1186/s13023-014-0172-y

437. Seixas S, Marques PI. Known Mutations at the Cause of Alpha-1 Antitrypsin Deficiency an Updated Overview of *SERPINA1* Variation Spectrum. *Appl Clin Genet*. 2021;Volume 14:173-194. doi:10.2147/TACG.S257511
438. Wang R, Brattain MG. The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa. *FEBS Lett*. 2007;581(17):3164-3170. doi:10.1016/j.febslet.2007.05.082
439. Spector DL, Ochs RL, Busch H. Silver staining, immunofluorescence, and immunoelectron microscopic localization of nucleolar phosphoproteins B23 and C23. *Chromosoma*. 1984;90(2):139-148. doi:10.1007/BF00292451
440. Luo D, Mari B, Stoll I, Anglard P. Alternative Splicing and Promoter Usage Generates an Intracellular Stromelysin 3 Isoform Directly Translated as an Active Matrix Metalloproteinase. *J Biol Chem*. 2002;277(28):25527-25536. doi:10.1074/jbc.M202494200
441. Perry ME, Mendrysa SM, Saucedo LJ, Tannous P, Holubar M. p76MDM2 Inhibits the Ability of p90MDM2 to Destabilize p53. *J Biol Chem*. 2000;275(8):5733-5738. doi:10.1074/jbc.275.8.5733
442. Wakasugi K, Slike BM, Hood J, et al. A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc Natl Acad Sci*. 2002;99(1):173-177. doi:10.1073/pnas.012602099
443. Perlmutter DH, Glover GI, Rivetna M, Schasteen CS, Fallon RJ. Identification of a serpin-enzyme complex receptor on human hepatoma cells and human monocytes. *Proc Natl Acad Sci*. 1990;87(10):3753-3757. doi:10.1073/pnas.87.10.3753
444. Hyun Kyu Song, Kee Nyung Lee, Ki-Sun Kwon, Myeong-Hee Yu, Se Won Suh. Crystal structure of an uncleaved $\alpha 1$ -antitrypsin reveals the conformation of its inhibitory reactive loop. *FEBS Lett*. 1995;377(2):150-154. doi:10.1016/0014-5793(95)01331-8
445. Pascarella S, Tiberi C, Sabatino G, et al. Serpin A1 C-Terminal Peptides as Collagen Turnover Modulators. *ChemMedChem*. 2016;11(16):1850-1855. doi:10.1002/cmdc.201500472
446. Lefèvre F, Garnotel R, Georges N, Gillery P. Modulation of Collagen Metabolism by the Nucleolar Protein Fibrillarlin. *Exp Cell Res*. 2001;271(1):84-93. doi:10.1006/excr.2001.5359
447. Ishida Y, Kubota H, Yamamoto A, Kitamura A, Bächinger HP, Nagata K. Type I Collagen in Hsp47-null Cells Is Aggregated in Endoplasmic Reticulum and Deficient in N-Propeptide Processing and Fibrillogenesis. *Mol Biol Cell*. 2006;17(5):2346-2355. doi:10.1091/mbc.e05-11-1065
448. Frottin F, Schueder F, Tiwary S, et al. The nucleolus functions as a phase-separated protein quality control compartment. *Science*. 2019;365(6451):342-347. doi:10.1126/science.aaw9157
449. Szebeni A, Olson MOJ. Nucleolar protein B23 has molecular chaperone activities. *Protein Sci*. 2008;8(4):905-912. doi:10.1110/ps.8.4.905
450. Szebeni A, Herrera JE, Olson MOJ. Interaction of nucleolar protein B23 with peptides related to nuclear localization signals and its modulation by phosphorylation. *Biochemistry*. 1995;34(25):8037-8042. doi:10.1021/bi00025a009
451. Bhattacharyya J, Padmanabha Udupa EG, Wang J, Sharma KK. Mini- α B-Crystallin: A Functional Element of α B-Crystallin with Chaperone-like Activity. *Biochemistry*. 2006;45(9):3069-3076. doi:10.1021/bi0518141
452. Kho J, Pham PC, Kwon S, et al. De Novo Design, Synthesis, and Mechanistic Evaluation of Short Peptides That Mimic Heat Shock Protein 27 Activity. *ACS Med Chem Lett*. 2021;12(5):713-719. doi:10.1021/acsmchemlett.0c00609
453. Okada AK, Teranishi K, Lobo F, et al. The Mitochondrial-Derived Peptides, HumaninS14G and Small Humanin-like Peptide 2, Exhibit Chaperone-like Activity. *Sci Rep*. 2017;7(1):7802. doi:10.1038/s41598-017-08372-5
454. El-Brolosy MA, Stainier DYR. Genetic compensation: A phenomenon in search of mechanisms. Moens C, ed. *PLOS Genet*. 2017;13(7):e1006780. doi:10.1371/journal.pgen.1006780
455. Salanga CM, Salanga MC. Genotype to Phenotype: CRISPR Gene Editing Reveals Genetic Compensation as a Mechanism for Phenotypic Disjunction of Morphants and Mutants. *Int J Mol Sci*. 2021;22(7):3472. doi:10.3390/ijms22073472

456. El-Brolosy MA, Kontarakis Z, Rossi A, et al. Genetic compensation triggered by mutant mRNA degradation. *Nature*. 2019;568(7751):193-197. doi:10.1038/s41586-019-1064-z
457. Marques PI, Ferreira Z, Martins M, et al. SERPINA2 Is a Novel Gene with a Divergent Function from *SERPINA1*. Pastore A, ed. *PLoS One*. 2013;8(6):e66889. doi:10.1371/journal.pone.0066889

Благодарности

Прежде всего, я выражаю искреннюю благодарность МГУ имени М.В. Ломоносова и Биологическому факультету. Именно в стенах университета я вдохновлялась идеями, и здесь же мне посчастливилось встретить людей, которые бескорыстно принимали участие в развитии выбранного мной направления. Благодарю научного руководителя Рубцова Михаила Александровича за неоценимую помощь и поддержку, ценные советы, предоставленные мне свободу действий и возможность выполнить существенную часть работы над диссертацией на кафедре молекулярной биологии Биологического факультета МГУ.

Я признательна Смирновой Ольге Вячеславовне за доверие и доброе ко мне отношение, всестороннюю помощь, ценные советы и предоставленные мне свободу и возможность трудиться в стенах родного факультета. Благодарю Соколову Ольгу Сергеевну за поддержку моей работы в начале пути, - благодаря предоставленным ею свободе выбора и возможностям мне посчастливилось выбрать объект изучения.

Особую благодарность выражаю Орловскому Игорю Вячеславовичу за веру в меня и безграничное терпение, содействие и высокий профессионализм, ценные советы и своевременные напутствия. Отдельно благодарю Голышева Сергея Александровича за доверие, проявленный интерес к теме и критическое мышление, высокий профессионализм, многолетнюю работу по микроскопии, ценные советы, подготовку иллюстраций и конструктивные дискуссии. Благодарю Поташникову Дарью Марковну за высокопрофессиональную работу по проточной цитофлуориметрии, Катруху Ивана Алексеевича за проявленный интерес к теме, помощь и содействие в получении антител против E5-эпитопа, Дидыча Дмитрия Александровича за интерес к теме и выявление предполагаемого механизма происхождения обнаруженных в работе коротких транскриптов, Мойсенович Анастасию Михайловну за оперативную помощь в конфокальной микроскопии. Отдельно благодарю за содействие при выполнении работы и за рабочую атмосферу моих коллег из лаборатории молекулярной биологии МГУ, в особенности Вьюшкова Владимира Сергеевича за профессионализм и содействие в дизайне и сборке плазмидных конструкций и Замалутдинова Алексея Вячеславовича за эволюционный анализ. Благодарю Левицкого Сергея Алексеевича, Дмитриева Сергея Евгеньевича, Ломова Николая Андреевича, Карлову Марию Георгиевну, Новоселецкого Валерия Николаевича и Творогову Анну Владимировну за то, что делились со мной ценным опытом, мою коллегу Сиротину Наталью Сергеевну за поддержку. Особую благодарность выражаю Богомазовой Александре Никитичне, Ельчанинову Андрею Владимировичу и Лябину Дмитрию Николаевичу за согласие выступить в роли официальных оппонентов моей диссертационной работы. И, конечно, выражаю особую благодарность моей семье и друзьям за неоценимую поддержку на протяжении всего периода исследований.